CONTENTS

BERKELEY TECHNOLOGY LAW JOURNAL

A CONTINUATION OF HIGH TECHNOLOGY LAW JOURNAL
EFFECTIVE VOLUME 11

VOLUME 12 NUMBER 2 1997

ARTICLES

Misuse or Fair Use: That Is the Software Copyright Question
James A.D. White 251

Nonobviousness and the Biotechnology Industry: A Proposal for a Doctrine of Economic Nonobviousness
Karen I. Boyd 311

Protecting the Private Inventor Under the Peacetime Provisions of the Invention Secrecy Act
Sabing H. Lee 345

Cyberspace Self-Government: Town Hall Democracy or Rediscovered Royalism?
Henry H. Perritt, Jr. 413

COMMENT

Communications Tower Sitings: The Telecommunications Act of 1996 and the Battle for Community Control
Susan Lorde Martin 483

ARTICLE

MISUSE OR FAIR USE: THAT IS THE SOFTWARE COPYRIGHT QUESTION

JAMES A.D. WHITE †

TABLE OF CONTENTS

I. INTRODUCTION 252
II. THE UTILITARIAN ROLE OF INTELLECTUAL PROPERTY IN OUR ECONOMIC SYSTEM 253
A. The Utilitarian Public Policy Rationale For Intellectual Property Protection In The United States 254
B. Balancing The Rights Accorded Copyright And Patent Holders With Those Retained By The Public 256
C. Balancing The Rights Granted By Federal Intellectual Property Law With Those Accorded By State Law 264
III. THE ROLE OF INTELLECTUAL PROPERTY MISUSE IN ENFORCING UTILITARIAN PUBLIC POLICY 265
A. The History Of The Patent Misuse Doctrine 266
B. The Recent Emergence Of The Copyright Misuse Doctrine 268
IV. THE PROPER TEST TO ENHANCE UTILITARIAN PUBLIC POLICY: ANTITRUST "RULE OF REASON" OR MISUSE "SCOPE OF THE GRANT"? 272
A. The Misuse Analysis Provides Advantages Over An Antitrust Analysis In Balancing The Rights Granted 273
B. The Unique Status Of Software As A Copyrighted Work 280
C. A "Scope Of The Grant" Copyright Misuse Doctrine Should Be Applied To Computer Software 287
V. THE PAST USE OF OTHER JUDICIAL DOCTRINES TO LIMIT SOFTWARE COPYRIGHT IN MISUSE CONTEXTS 288

---

† J.D. 1997, Boalt Hall School of Law, UC Berkeley. The author also holds a Masters degree in Computer Science from Stanford University, and has over 10 years of experience as a software engineer. The author is an attorney at the Seattle law firm of Seed & Berry LLP, where he practices computer-related intellectual property law with an emphasis on software patent prosecution.

A. Recent Growth Of The Software Industry Creates A Need For Copyright Limitations 288
B. Employing Fair Use To Prevent Misuse 289
C. Employing Other Copyright Doctrines To Prevent Misuse 293
VI. APPLYING SOFTWARE COPYRIGHT MISUSE TO TRADITIONAL MISUSE CONTEXTS 298
A. Software Copyright Misuse And Tying Arrangements 298
B. Software Copyright Misuse And Restrictions On Competition 299
VII. APPLYING SOFTWARE COPYRIGHT MISUSE TO CURRENT SOFTWARE PRACTICES 305
A. The Licensing Rather Than Sale Of Software 305
B. License Restrictions On Reverse Engineering 307
VIII. CONCLUSION 309

I. INTRODUCTION

The emergence of the copyright misuse doctrine is one of the most significant trends in copyright law in recent years. A basic premise of our intellectual property system is that various rights granted to creators as an incentive to produce intellectual property must be carefully balanced with the rights retained by the public. The judicial doctrine of intellectual property misuse limits efforts of intellectual property creators to inappropriately extend their rights and alter this careful balance. While the patent misuse doctrine has for decades prevented patent holders from unduly extending the rights granted to them beyond the scope of the patent, an appellate court did not extend an analogous limitation to a copyright owner until 1990. In the ensuing period, a debate has emerged whether the copyright misuse doctrine should exist, and if so, whether it should exist outside the scope of an antitrust violation.

This paper argues that the copyright misuse defense should exist, and that at least in the context of computer software it should be separate from antitrust analysis. Computer software is distinct from other forms of copyrighted works in various ways, providing more market power to the copyright owner and impacting federal patent law in ways that require restraints on the actions of software copyright owners. In addition, application of traditional antitrust analysis does not extend well to the software industry because actions that evade antitrust law are nevertheless misuse of copyright.

After examining how courts have previously applied doctrines such as "fair use" to restrict over-reaching by software copyright owners, the paper concludes that a software copyright misuse doctrine would provide a more appropriate vehicle for defining the scope of rights

accorded to software copyright owners. Finally, the paper briefly employs a misuse analysis to consider various situations that occur in the software industry.

Part II of this paper begins by examining the role that intellectual property plays in our economic system, briefly tracing its history and examining the balance of rights between creators and the general public. Part III provides a brief history of the development of the intellectual property misuse doctrine. Part IV examines the advantages and disadvantages of an antitrust approach to copyright misuse, and after exploring the unique aspects of computer software relative to other copyrighted works, concludes that a software copyright misuse doctrine should exist independent of antitrust law. Part V considers the relatively recent arrival of the software copyright misuse doctrine, and concludes that, in the absence of such a doctrine in the past, courts have used other doctrines to limit software copyrights. Parts VI and VII briefly apply a copyright misuse analysis to various practices in the software industry.

II. THE UTILITARIAN ROLE OF INTELLECTUAL PROPERTY IN OUR ECONOMIC SYSTEM

Commentaries on the copyright misuse doctrine often refer exclusively to the patent misuse doctrine and antitrust laws when reaching competing conclusions as to the correct application of copyright misuse. Many proponents of a strong copyright misuse doctrine argue that patent misuse has a long history, and that the similarity of copyright law to patent law thus legitimizes a copyright misuse doctrine in analogous situations. Alternately, proponents of restricting copyright misuse to a purely antitrust analysis focus only on how a copyright misuse defense would affect the application of antitrust principles.

When examining the proper scope of copyright protection for software, it is instructive to explore the historical application of the copyright[1] and patent[2] laws in a broader sense, and to review other situations which require limitations on the rights of patent and copyright owners. This exploration will illustrate that the patent and copyright laws have indeed evolved together and that they are based on a common public policy of benefiting society through the encouragement of creation, discovery and dissemination of novel ideas and creative expression. It will also illustrate that this utilitarian goal is achieved through the grant of property rights in a limited monopoly to inventors and creators.

---

1. *See* 17 U.S.C. §§ 101-1101 (1997).
2. *See* 35 U.S.C. §§ 1-376 (1997).

Although society wants to provide incentives to create, underlying public policy necessitates the careful balancing of the rights granted to creators with the rights retained by the general public. The rights which are granted to creators are determined uniformly through statutory grants and limitations, but subjective judicial doctrines allow courts to modify these rights in individual situations. Enforcement of an appropriate balance prevents reordering of these rights, whether the reordering occurs through state law or by private action. While the misuse doctrine can be used to prevent such reordering, it is only one of many limitations placed on the rights of intellectual property holders and it should be applied in a manner consistent with these other limitations.

A. The Utilitarian Public Policy Rationale For Intellectual Property Protection In The United States

The current federal copyright and patent laws share underlying public policy rationales, and are the result of a similar evolution over time. As with much of the law in the United States, the impetus for federal patent and copyright laws originated with English law, which strove to benefit society by encouraging the creation of new inventions and new works of authorship.[3] In continuing this English tradition, the copyright and patent laws in the United States are designed to benefit society as a whole by providing incentives to creators.[4] These federal laws stem from a constitutional grant of power to Congress: "*To promote the Progress of Science and useful Arts*, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries."[5] The copyright/patent clause is notable in that it explicitly

---

3. During the sixteenth century, the English Crown was notorious for granting "letters patent" to individuals, which conveyed exclusive rights to produce, import, or sell items within the kingdom (extending even to common items such as salt and vinegar). In response to abuses, Parliament passed the Statute of Monopolies in the early seventeenth century, which permitted patents to be granted only to the creator of a new invention and only for fourteen years. In a similar manner, the Crown in the sixteenth century granted to the Stationers' Company an exclusive right to publish and print all published works. Parliament revoked that right by passing the Statute of Anne in the early 1700s, which gave authors the exclusive right to publish their works for twenty-eight years. .*See* 1 ERNEST B. LIPSCOMB, LIPSCOMB'S WALKER ON PATENTS §§ 1:1-1:2, 1:5-1:6 (3d ed. 1984) [hereinafter WALKER ON PATENTS]; WILLIAM F. PATRY, LATMAN'S THE COPYRIGHT LAW, 2-5 (1986).

4. The term "creators" as used in this paper is intended to include both authors and inventors, and "creative works" is intended to include both original works of authorship and novel inventions.

5. U.S. CONST. art. I, § 8, cl. 8 (emphasis added).

mentions the purpose behind the grant of power, an unusual occurrence in the Constitution.[6] By placing the copyright and patent grant of power together in this manner, and by explicitly incorporating the purpose of the grant in the Constitution, the framers indicated the importance of these doctrines and continued the English tradition of encouraging the creation of intellectual property.[7]

American courts have affirmed the utilitarian nature of intellectual property rights[8] by repeatedly emphasizing that the federal intellectual property rights granted to creators serve as an incentive to produce. "The

---

6. [The unusual inclusion of the reason for the grant of the power in the copyright/patent clause] doubtlessly was due to the fact that those who formulated the Constitution were familiar with the long struggle over monopolies so prominent in English history, where exclusive rights to engage even in ordinary business activities were granted so frequently by the Crown for the financial benefits accruing to the Crown only. It was desired that in this country any Government grant of a monopoly for even a limited time should be limited to those things which serve in the promotion of science and the useful arts.

In Re Yuan, 188 F.2d 377, 380 (C.C.P.A. 1951).

7. It is significant, we think, that the framers of our Constitution continued the English development of intellectual property law and considered in tandem those property rights protectable by copyrights and those protectable by patents. In giving Congress the power to create copyright and patent laws, the framers combined the two concepts in one clause, stating a unitary purpose—to promote progress .... [A] comment in The Federalist papers indicates the public policy behind the grant of copyright and patent powers is essentially the same.

Lasercomb America, Inc. v. Reynolds, 911 F.2d 970, 975 (4th Cir. 1990).

8. The intellectual property doctrine that arose in England emphasized the utilitarian nature of the intellectual property rights granted to inventors and authors. By granting a limited monopoly to these creators, they were given an incentive to produce inventions and works because the monopoly presumably allowed the creators to reap benefits from a useful creation that would compensate them for their effort and risk taken in producing the useful creation. Thus, benefits received by creators serve as an incentive to produce rather than as a reward for creations.

Conversely, the intellectual property doctrines that arose in continental Europe are often referred to as based on a moral rights theory. This theory grants rights to creators not as an incentive to produce, but as a recognition of the ownership in the creation that results from the labor and ingenuity in producing the creation, and stems in part from Locke's theory of property. *See generally* Jane C. Ginsburg, *A Tale Of Two Copyrights: Literary Property In Revolutionary France And America*, 64 TUL. L. REV. 991 (1990).

economic philosophy behind the clause empowering Congress to grant patents and copyrights is the conviction that encouragement of individual effort by personal gain is the best way to advance public welfare through the talents of authors and inventors in 'Science and useful Arts.'"[9] Thus, both patent law and copyright law derive their underlying authority from the common public policy of benefiting society by granting incentives for the production of novel ideas and creative expression.

B. Balancing The Rights Accorded Copyright And Patent Holders With Those Retained By The Public

In order to achieve the public policy objectives underlying copyright and patent law, it is necessary to limit the rights that are granted to the owners of copyrights and patents.[10] If the ultimate utilitarian goal is to increase the number of inventions and works of authorship that are available to the public, then legal monopolies granted to creators can have a negative short-term effect on this goal. To the extent that creators can restrict access or use of their creation, whether absolutely or through exorbitant monopoly rents, the public fares less well than if the same creation had been available without legal restrictions. However, the premise of utilitarian theory is that some incentive is necessary to induce the creation in the first place—that is, the public is better off with a new creation at a monopoly price than with no creation at all. Furthermore, the limited duration of the rights granted guarantees that at some point the creation will enter the public domain and become freely available to the public.

Since the legal incentives granted to creators deprive the public of a benefit in the short term, the utilitarian goal of the intellectual property system can only be fully achieved if the incentives granted to creators are the minimum necessary to spur the creation and dissemination of inventions and works of authorship. In an effort to achieve this optimal level of innovation at the lowest public costs, the rights granted to owners

---

9. *See* Mazer v. Stein, 347 U.S. 201, 219 (1953); *see also* Feist Publications, Inc. v. Rural Tel. Serv. Co., 499 U.S. 340, 349-50 (1991) ("The principal objective of copyright is not to reward the labor of authors, but 'to promote the Progress of Science and useful Arts'"); United States v. Paramount Pictures, Inc., 334 U.S. 131, 158 (1948); United States v. Loew's, Inc., 371 U.S. 38, 44-51 (1962); United States v. Dubilier Condenser Corp., 289 U.S. 178, 186 (1933); WALKER ON PATENTS, *supra* note 3, § 1.6.

10. It would be more accurate to state that rights are granted to a particular invention or work of authorship, and that these rights generally subsist in the current owner of the invention or work. For the sake of brevity, I will refer in this paper to the rights granted to creators.

of copyrights and patents are carefully balanced with the rights retained by the public. In a world with perfect information and zero transaction costs, the optimal solution could be attained by varying the rights granted to each creator on an individual basis. Thus, each creator could be compensated at the minimum level necessary to compensate for each creation produced, regardless of the underlying value of the creation.

Transaction costs render it impractical to grant rights to each creator on an individual basis. Therefore, the legal system instead imposes systemic constraints on the rights of creators. These constraints take the form of both statutory limitations and judicial doctrines. While some systemic constraints are objective limitations that apply to all creators equally, others are subjective or equitable in nature. To the extent that subjective and equitable constraints are imposed on an individual basis, these efforts can be viewed as an attempt to tune more finely the balance of rights granted to a particular creator. Conversely, courts can enforce subjective doctrines broadly and use them to alter the balance of rights granted to creators either for a broad class of transactions, or for a type of protected subject matter.

1. *STATUTORY DEFINITIONS OF COPYRIGHT AND PATENT RIGHTS*

The statutory grants of power under chapters 17 and 35 of the United States Code define federal intellectual property rights of creators. In the absence of such explicit statutory grants, a creator would not enjoy any federal intellectual property protection for their creations. Under the protections granted for both copyrights and patents, the federal statutes include broad provisions that grant power to the creator, and narrower provisions that restrict the granted power in a variety of specific ways.

A patent conveys the right to exclude others[11] from making, using, selling, offering for sale, and importing the patented invention.[12] This

11. The patent statute conveys rights to owners as a negative grant—while the owner can generally exclude others from engaging in prohibited activities, an owner is not statutorily entitled to practice his or her own invention. Thus, external limitations, such as regulations from an administrative body such as the Food and Drug Administration or legal restrictions from an inventor with a blocking patent, can prevent an inventor from practicing his or her own invention.

right can be enforced through a combination of injunctions,[13] damages,[14] and attorney's fees.[15] While these rights can convey significant economic power for a useful invention, the statutes limit the rights in various ways. As a threshold matter, a patent is only granted to inventions that are "worthy" of these legal protections. That is, to be patentable, an invention must have the appropriate statutory subject matter,[16] be novel,[17] be useful[18] and be non-obvious.[19] Further, the patent rights granted are of a limited duration.[20] In addition, the rights of patent owners are statutorily limited in other ways.[21] Through these statutory grants, patent law provides relatively broad rights that last for a relatively short time.

---

12. (a) Except as otherwise provided in this title, whoever without authority makes, uses, offers to sell, or sells any patented invention, within the United States or imports into the United States any patented invention during the term of the patent therefor, infringes the patent.

(b) Whoever actively induces infringement of a patent shall be liable as an infringer.

(c) Whoever offers to sell or sells within the United States or imports into the United States a component of a patented machine, manufacture, combination or composition, or a material or apparatus for use in practicing a patented process, constituting a material part of the invention, knowing the same to be especially made or especially adapted for use in an infringement of such patent, and not a staple article or commodity of commerce suitable for substantial noninfringing use, shall be liable as a contributory infringer.

35 U.S.C. § 271 (1997).

13. *See* 35 U.S.C. § 283 (1997).
14. *See* 35 U.S.C. § 284 (1997).
15. *See* 35 U.S.C. § 285 (1997).
16. *See* 35 U.S.C. § 101 (1997).
17. *See* 35 U.S.C. § 102 (1997).
18. *See* 35 U.S.C. § 101 (1997).
19. *See* 35 U.S.C. § 103 (1997).
20. *See* 35 U.S.C. §§ 154, 173 (1997).
21. For example, a patent owner cannot prevent another from using the owner's invention if it is in a vessel that is temporarily in the country. *See* 35 U.S.C. § 272 (1997). This is also true if the use of the invention is "solely for uses reasonably related to the development and submission of information under a Federal law which regulates the manufacture, use, or sale of drugs or veterinary biological products," or if the patent owner ties the use of the invention to a separate product and the patent owner has market power in the relevant market. *See* 35 U.S.C. § 271(d) & (e)(1) (1997).

A copyright, on the other hand, conveys the affirmative exclusive right to reproduce, distribute to the public, perform publicly, display publicly, and prepare derivative works of the copyrighted work.[22] Authors of works of visual art have additional rights of attribution and integrity,[23] and performers of live musical performances have additional rights to control the use of their performances.[24] Copyright protection is available for original works of authorship that are fixed in any tangible medium of expression[25] and lasts for a limited duration, but does not

---

22. Subject to sections 107 through 120, the owner of copyright under this title has the exclusive rights to do and to authorize any of the following:
(1) to reproduce the copyrighted work in copies or phonorecords;
(2) to prepare derivative works based upon the copyrighted work;
(3) to distribute copies or phonorecords of the copyrighted work ...;
(4) ... to perform the copyrighted work publicly;
(5) ... to display the copyrighted work publicly; and
(6) in the case of sound recordings, to perform the copyrighted work publicly by means of a digital audio transmission.

17 U.S.C. § 106 (1997).

23. *See* 17 U.S.C. § 106A (1997).

24. *See* 17 U.S.C. § 1101 (1997).

25. (a) Copyright protection subsists, in accordance with this title, in original works of authorship fixed in any tangible medium of expression, now known or later developed, from which they can be perceived, reproduced, or otherwise communicated, either directly or with the aid of a machine or device ....
(b) In no case does copyright protection for an original work of authorship extend to any idea, procedure, process, system, method of operation, concept, principle, or discovery, regardless of the form in which it is described, explained, illustrated, or embodied in such work.

17 U.S.C. § 102 (1997).

prevent independent creation of the same work.[26] As with patent law, the copyright statute provides various statutory limitations on copyright owners,[27] including special provisions for certain classes of users such as libraries[28] and instructors,[29] and special provisions for certain classes of copyrighted works such as computer programs.[30] In addition, section 109 provides limitations on a copyright owner's ability to control an item embodying a copyright after the item has been sold (the "first-sale" doctrine),[31] and section 107 codifies the equitable "fair use" limitation

---

26. (a) In General.—Copyright in a work created on or after January 1, 1978, subsists from its creation and, except as provided by the following subsections, endures for a term consisting of the life of the author and fifty years after the author's death ....

(c) Anonymous Works, Pseudonymous Works, and Works Made For Hire.—In the case of an anonymous work, a pseudonymous work, or a work made for hire, the copyright endures for a term of seventy-five years from the year of its first publication, or a term of one hundred years from the year of its creation, whichever expires first.

17 U.S.C. § 302 (1997).

27. *See* 17 U.S.C. § 108, §§ 110-120 (1997).

28. *See* 17 U.S.C. § 108 (1997).

29. *See* 17 U.S.C. § 110 (1997).

30. Notwithstanding the provisions of section 106, it is not an infringement for the owner of a copy of a computer program to make or authorize the making of another copy or adaptation of that computer program provided:

(1) that such a new copy or adaptation is created as an essential step in the utilization of the computer program in conjunction with a machine and that it is used in no other manner, or

(2) that such new copy or adaptation is for archival purposes only and that all archival copies are destroyed in the event that continued possession of the computer program should cease to be rightful.

Any exact copies prepared in accordance with the provisions of this section may be leased, sold, or otherwise transferred, along with the copy from which such copies were prepared, only as part of the lease, sale, or other transfer of all rights in the program. Adaptations so prepared may be transferred only with the authorization of the copyright owner.

17 U.S.C. § 117 (1997).

31. The owner of a lawfully made copy can sell or dispose of the copy (i.e., distribute it) or publicly display it without authorization of the owner of the copyright. *See* 17 U.S.C. § 109 (1997).

which permits certain types of uses of copyrighted works.[32] These statutory copyright grants provide relatively narrow rights that last for a relatively long time.

2. *JUDICIAL MODIFICATIONS OF COPYRIGHT AND PATENT RIGHTS*

While the copyright and patent statutes provide the primary definition of the rights granted to creators, the courts have adopted judicial doctrines which alter the scope of these rights. Over time, some of these doctrines were codified and became part of the statutory framework. For example, the "fair use" doctrine of copyright law originally arose judicially to restrict the rights of copyright owners, but is now statutorily codified.[33] Although it is possible for judicial doctrines to provide an objective expansion or limitation of a creator's rights, they generally are more subjective and equitable in nature. As such, courts can enforce these doctrines on an individual basis and can alter the rights granted to an individual creator to reflect the amount of protection that is needed or deserved. Alternatively, courts can enforce such doctrines broadly and use them in such a manner so as to alter the balance of rights

---

32. Notwithstanding the provisions of sections 106 and 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include—

(1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;

(2) the nature of the copyrighted work;

(3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and

(4) the effect of the use upon the potential market for or value of the copyrighted work.

The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.

17 U.S.C. § 107 (1997).

33. *See* 17 U.S.C. § 102 (1997).

granted to creators for a class of transactions or for a type of protected subject matter.[34]

Two notable judicial doctrines used in patent law are the requirement of non-obviousness for patentability[35] and the doctrine of equivalents for determining the breadth of claim scope.[36] While courts describe these doctrines as objective standards,[37] both doctrines involve subjective tests that provide results which are notoriously difficult to predict. For example, consider an invention that involves the combination of two techniques, A and B, and prior art references that describe either A or B but never their combination. An objective standard might hold that the combination of the two techniques is not obvious unless there is an explicit suggestion in the prior art to combine these techniques. Instead, the judicial determination is whether the combination of these techniques was obvious at the time of the invention to one of ordinary skill in the art. This determination is inherently subjective.

In the area of copyright law, the judicial doctrine of "fair use" is an extremely broad, subjective limitation on the rights of copyright holders. The fair use doctrine permits certain uses of copyrighted material (e.g. making copies of copyrighted material for news reporting or criticism), despite a literal violation of prohibited statutory activities. Although now codified in the copyright statute,[38] the determination of fair use is

---

34. It is not unusual for foreign copyright systems to vary the duration of the right granted based on the class of the work. *See* ROBERT A. GORMAN & JANE C. GINSBURG, COPYRIGHT FOR THE NINETIES 316 (4th ed. 1993). Even in the U.S., different terms are granted for utility and design patents, as well as for works of authorship depending on the situation under which the work was created. *See supra* notes 20 and 26.

35. The nonobviousness requirement, now codified in the patent statutes, creates a limitation on the rights of inventors by raising the standard necessary for an invention to achieve patentability. *See supra* note 19; Graham v. John Deere Co., 383 U.S. 1, 3-4 (1966) ("We have concluded that the 1952 Act was intended to codify judicial precedents.").

36. The doctrine of equivalents enhances the rights of patent holders by extending the scope of patented claims to cover products that have only insubstantial differences from the claimed invention. *See* Hilton Davis Chemical Co. v. Warner-Jenkinson Company, Inc., 62 F.3d 1512, 1516-17 (Fed. Cir. 1995), *rev'd on other grounds and remanded*, 1997 U.S. LEXIS 1476 (Mar 3, 1997).

37. *See id.* at 1519 (finding the doctrine of equivalents is an objective standard); Texas Instruments Inc. v. United States International Trade Commission, 988 F.2d 1165, 1178 (Fed. Cir. 1993) (holding the nonobvious requirement is an objective standard).

38. *See supra* note 32.

inherently subjective and thus dependent on judicial implementation.[39] While fair use has traditionally been viewed as a limitation on the rights of the copyright owner (i.e., a balancing of the rights that favors the public),[40] some scholars have conceived fair use in certain cases to merely reflect market failure; if the transaction costs to obtain permission to use a right are higher than the value of the right, it would not warrant granting that right.[41] Regardless of the underlying conception, such a broad-based doctrine gives significant power to the courts to modify the rights granted by copyright law. While the doctrine is intended to be a fact-intensive determination,[42] and thus variable in each case, sweeping pronunciations by the courts can be regarded as determinative of a broad class of similar situations.[43] In this manner, courts can use a powerful subjective doctrine such as fair use to alter the balance of rights granted to creators in a broad class of situations.

---

39. The factors enumerated in the section are not meant to be exclusive: "[S]ince the doctrine [of fair use] is an equitable rule of reason, no generally applicable definition is possible, and each case raising the question must be decided on its own facts." H.R. REP. NO. 94-1476, at 65 (1976), *reprinted in* 1976 U.S.C.C.A.N. 5659, 5679.

40. "In essence, the traditional concept of fair use excused reasonable unauthorized appropriations from a first work, when the use to which the second author put the appropriated material in some way advanced the public benefit, without substantially impairing the present or potential economic value of the first work." GORMAN & GINSBURG, *supra* note 34, at 548. Judge Birch wrote:

> Although the traditional approach is to view "fair use" as an affirmative defense, this writer, speaking only for himself, is of the opinion that it is better viewed as a right granted by the Copyright Act of 1976 .... As a statutory doctrine, however, fair use is not an infringement. Thus, since the passage of the 1976 Act, fair use should no longer be considered an infringement to be excused; instead, it is logical to view fair use as a right.

Bateman v. Mnemonics, Inc., 79 F.3d 1532, 1542 n.22 (11th Cir. 1996).

41. *See* Robert P. Merges, *Of Property Rules, Coase, And Intellectual Property*, 94 COLUM. L. REV. 2655 (1994); Wendy J. Gordon, *Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and Its Predecessors*, 82 COLUM. L. REV. 1600 (1982).

42. Harper & Row, Publishers, Inc. v. Nation Enterprises, 471 U.S. 539, 560-69 (1985).

43. For example, a court could effectively pronounce that a book review quoting 100 words or less from a full length novel is per se fair use. *See* Part V.B. *infra* for additional discussion.

C. Balancing The Rights Granted By Federal Intellectual Property Law With Those Accorded By State Law

The federal intellectual property laws are based upon a utilitarian public policy rationale that necessitates balancing the rights granted to creators with those retained by the public. The statutory framework of the copyright and patent laws, and their corresponding judicial doctrines, have evolved to adjust and protect this balance. However, the federal laws do not provide the only means to grant rights to creators or the public—both state laws and private contractual agreements can alter this balance of rights. Since our legal system grants significant deference to both state and private action, federal law will preempt such action only if necessary for an important federal purpose. Courts have recently faced issues of state and private actions that significantly alter the balance of rights granted under federal law, and have been forced to decide whether to preempt these actions. In resolving the preemption issue, courts have relied on the constitutional authority of federal intellectual property laws and the Supremacy Clause[44] of the Constitution, which dictate that state action, including state-enforced private action, may not encroach on that authority.

In the 1964 companion cases of *Sears, Roebuck & Co. v. Stiffel Co.*[45] and *Compco Corp. v. Day-Brite Lighting, Inc.*,[46] the Supreme Court gave a broad reading to the constitutional preemption of state laws relating to intellectual property. In both cases, state unfair competition laws prohibited the copying of products that were neither patented nor copyrighted. In *Sears*, the court stated that "[w]hen state law touches upon the area of these federal statutes, it is 'familiar doctrine' that the federal policy 'may not be set at naught, or its benefits denied' by the state law. [citation omitted.] This is true, of course, even if the state law is enacted in the exercise of otherwise undoubted state power."[47] The court recently reaffirmed this policy in 1989 by striking down a Florida law that prevented the copying by direct molding of unpatented boat hulls in *Bonito Boats, Inc. v. Thunder Craft Boats, Inc.*[48] The "state regulation of intellectual property must yield to the extent that it clashes

---

44. The United States "Constitution, and the Laws of the United States which shall be made in Pursuance thereof ... shall be the supreme Law of the Land; ... any Thing in the Constitution or Laws of any State to the Contrary notwithstanding." U.S. CONST. art. VI, cl. 2.

45. 376 U.S. 225 (1964).

46. 376 U.S. 234 (1964).

47. *Sears*, 376 U.S. at 229.

48. 489 U.S. 141 (1989).

with the balance struck by Congress in our patent laws."[49] Thus, there are clearly constitutional limits on the laws which states may enact and enforce. At a minimum, state laws will be preempted if they provide protection to uncopyrighted or unpatented works in such a manner as to give the author or creator of those works a property right equivalent to that granted by patent or copyright law.

Although some state laws alter the balance of federal intellectual property rights sufficiently to be preempted, the Supreme Court found that state trade secret law does not. In upholding an Ohio law of trade secrets, the Court held that a state law is void under the Supremacy Clause only if it stands as an obstacle to the accomplishment and execution of the full purposes and objectives of Congress.[50] However, the Court noted that limits existed on how far trade secret law could be enforced. "If a State, through a system of protection, were to cause a substantial risk that holders of patentable inventions would not seek patents, but rather would rely on the state protection, we would be compelled to hold that such a system could not constitutionally continue to exist."[51]

Thus, it is difficult to determine whether a particular state or private action should be allowed to alter the balance of federal intellectual property rights. The Supreme Court has stated that while state regulation of intellectual property, which clashes with the balance struck by Congress, must be preempted, such regulation will clash in this manner only if it stands as an obstacle to the accomplishment and execution of the full purposes and objectives of Congress.[52] It is clear, however, that consideration of the purposes and objectives of Congress will be the touchstone to maintain the careful balance of federal intellectual property rights in a preemption analysis.

III. THE ROLE OF INTELLECTUAL PROPERTY MISUSE IN ENFORCING UTILITARIAN PUBLIC POLICY

As previously noted, the utilitarian goal of intellectual property is accomplished by granting property rights in a limited monopoly to creators, and the underlying public policy necessitates the careful balancing of the rights and powers granted to the creator and the rights and powers retained by the general public. The judicial doctrine of intellectual property misuse was created to address situations in which

---

49. *Id.* at 152.

50. *See* Kewanee Oil Co. v. Bicron Corp., 416 U.S. 470 (1974).

51. *Id.* at 489.

52. *Id.* at 480-81.

the owner of an intellectual property right used his or her legal monopoly to create such an asymmetry in the balance of rights that the courts refused to enforce the normal intellectual property rights.[53]

A. The History Of The Patent Misuse Doctrine

The doctrine of intellectual property misuse first arose in the early 1900s in conjunction with the use of patents.[54] In the 1917 case of *Motion Picture Patents v. Universal Film Mfg. Co.*,[55] the patentee licensed its patented movie projector on the condition that the film used in the machine must be purchased from the patentee (a type of tying arrangement).[56] The Court found that:

> [s]uch a restriction is invalid because such a film is obviously not any part of the invention of the patent in suit; because it is an attempt, without statutory warrant, to continue the patent monopoly in this particular character of film after it has expired, and because to enforce it would be to create a monopoly in the manufacture and use of moving picture films, wholly outside of the patent in suit and of the patent law as we have interpreted it.[57]

In short, the Court denied relief to the patentee because the licensing restrictions attempted to extend the scope of the film projector patent into the unpatented area of film.

The Court extended the patent misuse doctrine in the seminal 1942 case *Morton Salt Co. v. G. S. Suppiger Co.*,[58] and in doing so addressed the relationship of the misuse doctrine to antitrust law. In *Morton Salt*, the plaintiff held a patent on a machine that placed salt tablets into canned goods, and required licensees of its machine to purchase the unpatented

---

53. [C]opyright owners are given power by virtue of a congressionally-sanctioned monopoly; users are safeguarded under the same statute by at least some congressional solicitude for their interests (*e.g.*, the fair use doctrine, limited terms, § 117). The proprietor's use of her property power to force a user to forego the rights that Congress intended him to have is the evil against which the misuse defense is aimed.

David Nimmer, *Brains And Other Paraphernalia Of The Digital Age*, 10 HARV. J.L. & TECH. 1, 24 n.103 (1996).

54. For a detailed look at the history of the patent misuse doctrine, see generally James B. Kobak, Jr., *A Sensible Doctrine Of Misuse For Intellectual Property Cases*, 2 ALB. L.J. SCI. & TECH. 1 (1992).

55. 243 U.S. 502 (1917).

56. *Id.* at 506.

57. *Id.* at 518.

58. 314 U.S. 488 (1942).

salt tablets for the machine exclusively from the plaintiff.[59] The defendant produced its own salt tablet-depositing machine that was modeled after the plaintiff's machine, and the plaintiff sued for direct infringement of its patent.[60] Among other defenses, the defendant argued that the plaintiff misused its patent through its licensing agreements. The Court, without ruling on validity or infringement, found that the plaintiff's tying arrangement constituted patent misuse. In so doing, the Court prevented the plaintiff from enforcing its patent,[61] despite the lack of any connection between the defendant's infringing machine and the plaintiff's tying actions.[62] In addition, the court found that an antitrust violation was not necessary for a finding of misuse.[63]

---

59. *Id.* at 491.

60. *Id.* at 490-91.

61. Equity may rightly withhold its assistance from such a use of the patent by declining to entertain a suit for infringement, and should do so at least until it is made to appear that the improper practice has been abandoned and that the consequences of the misuse of the patent have been dissipated.

Id. at 493.

62. It is the adverse effect upon the public interest of a successful infringement suit in conjunction with the patentee's course of conduct which disqualifies him to maintain the suit, regardless of whether the particular defendant has suffered from the misuse of the patent. Similarly equity will deny relief for infringement of a trademark where the plaintiff is misrepresenting to the public the nature of his product either by the trademark itself or by his label.

Id. at 494.

63. The Court of Appeals for the Seventh Circuit [found for the plaintiff] ... because it thought that respondent's use of the patent was not shown to violate § 3 of the Clayton Act, 15 U.S.C. § 14, 15 U.S.C.A. § 14, as it did not appear that the use of its patent substantially lessened competition or tended to create a monopoly in salt tablets .... [However, we reverse.] The Clayton Act authorizes those injured by violations tending to monopoly to maintain suit for treble damages and for an injunction in appropriate cases. 15 U.S.C. §§ 1, 2, 14, 15, 26, 15 U.S.C.A. §§ 1, 2, 14, 15, 26. But the present suit is for infringement of a patent. The question we must decide is not necessarily whether respondent has violated the Clayton Act, but whether a court of equity will lend its aid to protect the patent monopoly when respondent is using it as the effective means of restraining competition with its sale of an unpatented article .... *It is unnecessary to decide whether respondent has violated the Clayton Act, for*

The courts and Congress continued to develop and refine the doctrine of patent misuse after *Morton Salt*. While patent misuse has been analyzed in a number of factual settings,[64] two paradigmatic cases emerged where misuse was typically found—tying arrangements where the patentee requires the purchase of unpatented goods or services along with a patented product or process, and non-compete clauses that prevent a patent licensee from producing or selling competing goods.[65] The analysis was changed for tying arrangements in 1988 when Congress passed the Patent Misuse Reform Act[66] which modified § 271(d) of the Patent Act, and held that patent tying arrangements are no longer per se misuse.[67]

Thus, the patent misuse defense continues to be an important judicial doctrine, with some recent statutory modifications for certain types of misuse cases. However, the misuse doctrine has not been limited solely to the patent arena. It has also been applied in trademark law and has recently emerged in the area of copyright.

B. The Recent Emergence Of The Copyright Misuse Doctrine

Although the copyright misuse doctrine was articulated by the courts only recently, the doctrine has been implicitly recognized since the time of *Morton Salt*. In the 1948 case of *United States v. Paramount Pictures, Inc.*,[68] the Supreme Court affirmed the lower court's finding that

---

> *we conclude that in any event the maintenance of the present suit to restrain petitioner's manufacture or sale of the alleged infringing machines is contrary to public policy and that the district court rightly dismissed the complaint for want of equity.*

Id. at 490-94 (emphasis added).

64. *See* 5 DONALD S. CHISUM, ON PATENTS § 19.04[3] (1996) [hereinafter CHISUM] (noting twelve types of situations in which patent misuse has been considered, and speculates that the acts which will be viewed as constituting misuse will shift over time. The twelve situations are tying arrangements, covenants not to compete, package licensing, post-expiration royalties and restraints, royalties based on total sales (unrelated to amount of sales including patented item), refusals to license including excessive or discriminatory royalties, price fixing, territorial limitations including resale restraints, field-of-use and customer limitations, grant-back clauses, restraints on the patentee including covenants not to license, and suppression including compulsory licensing.).

65. *See id.*

66. Pub. L. No. 100-73, 102 Stat. 4674 (1988).

67. *See* 35 U.S.C. § 271(d) (1997); Lasercomb America, Inc. v. Reynolds, 911 F.2d 970, 976 n.15 (4th Cir. 1990) ("The primary effect of the Patent Misuse Reform Act is to eliminate the presumption that use of a patent license to create a tie-in is per se misuse.").

68. 334 U.S. 131 (1948).

Paramount's block-booking of movies was an antitrust violation by relying on the reasoning of the patent misuse cases.[69] The Court was more explicit in a later block-booking case, *United States v. Loew's Inc.*,[70] in analogizing the patent misuse cases to the area of copyright.[71] Furthermore, statements by other courts indicated that copyright misuse might have been recognized under different sets of facts.[72]

Despite this long history, the copyright misuse doctrine was not applied by an appellate court until 1990.[73] In that year, the Fourth

---

69. "[E]nlargement of the monopoly of the copyright was condemned below in reliance on the principle which forbids the owner of a patent to condition its use on the purchase or use of patented or unpatented materials. [Citations omitted]." *Id.* at 157.

70. 371 U.S. 38 (1962).

71. The requisite economic power is presumed when the tying product is patented or copyrighted. [Citations omitted]. This principle grew out of a long line of patent cases which had eventuated in the doctrine that a patentee who utilized tying arrangements would be denied all relief against infringements of his patent. Motion Picture Patents Co. v. Universal Film Mfg. Co., 243 U.S. 502; [citations omitted], Morton Salt Co. v. G. S. Suppiger Co., 314 U.S. 488; [citation omitted]. As the District Court [in *Paramount Pictures*] said, the result is to add to the monopoly of the copyright in violation of the principle of the patent cases involving tying clauses .... Accommodation between the statutorily dispensed monopoly in the combination of contents in the patented or copyrighted product and the statutory principles of free competition demands that extension of the patent or copyright monopoly by the use of tying agreements be strictly confined.

Id. at 45-50.

72. *See* Mitchell Bros. Film Group v. Cinema Adult Theater, 604 F.2d 852, 865 n.27 (5th Cir. 1979), *cert. denied*, 445 U.S. 917 (1980) ("It is ... likely that the public monopoly extension rationale of *Morton Salt* ... is applicable to copyright."); F.E.L. Publications, Ltd. v. Catholic Bishop of Chicago, 214 U.S.P.Q. 409, 413 n.9 (7th Cir. 1982) ("[I]t is copyright misuse to exact a fee for the use of a musical work which is already in the public domain"); United Telephone Company v. Johnson Publishing Co., 855 F.2d 604 (8th Cir. 1988); Broadcast Music, Inc. v. Moor-Law, Inc., 691 F.2d 490 (3rd Cir. 1982); Supermarket of Homes, Inc. v. San Fernando Valley Board of Realtors, 786 F.2d 1400 (9th Cir. 1986); BellSouth Advertising & Publishing Corp. v. Donnelley Information Publishing Co., Inc., 719 F. Supp. 1551 (S.D. Fla. 1988), *rev'd on other grounds*, 999 F.2d 1436 (11th Cir. 1993).

73. One district court case in 1949, M. Witmark & Sons v. Jensen, 80 F. Supp. 843 (D. Minn. 1948), *appeal dismissed*, 177 F.2d 515 (8th Cir. 1949), had previously upheld a copyright misuse defense and also found an antitrust violation in a situation involving the

Circuit decided *Lasercomb America, Inc. v. Reynolds*,[74] in which the plaintiff, Lasercomb, had developed and licensed software for the computer-aided design and manufacture of steel rule dies used to score paper for folding into cartons.[75] The defendant, Reynolds, had purchased four licenses for the software. However, rather than purchasing additional licenses, Reynolds chose to circumvent the copy protection devices included with the software and made three additional unauthorized copies which Reynolds used on its own systems.[76] Reynolds then created an almost exact copy of Lasercomb's software, and marketed this program in competition with Lasercomb's product.[77] Because the defendant purposefully copied the software and made misrepresentations to Lasercomb, there was no question of copyright infringement.[78] However, in addition to limiting copying of licensed software, Lasercomb also included a provision in its standard form contract that prevented a licensee from independently developing a competing software program for 100 years.[79] Although the defendants were not themselves subject to this licensing provision, other licensees of Lasercomb had agreed to this provision,[80] and the defendants argued that Lasercomb had misused its copyright through the use of this non-competition provision.

After a detailed analysis, the Fourth Circuit accepted the argument of the defendants, and found that the plaintiff's license agreements amounted to copyright misuse because the license provision inhibited independent creation for a lengthy period of time (the 100 year term was

---

blanket-license practices for musical compositions of the American Society of Composers, Authors and Publishers (ASCAP).

74. 911 F.2d 970 (4th Cir. 1990).

75. *Id.* at 971.

76. *Id.*

77. *Id.*

78. *Id.*

79. Licensee agrees during the term of this Agreement and for one (1) year after the termination of this Agreement, that it will not write, develop, produce or sell or assist others in the writing, developing, producing or selling computer assisted die making software, directly or indirectly without Lasercomb's prior written consent. Any such activity undertaken without Lasercomb's written consent shall nullify any warranties or agreements of Lasercomb set forth herein. The "term of this Agreement" referred to in these clauses is ninety-nine years.

Id. at 973.

80. *Id.*

potentially longer than the life of the copyright itself).[81] In so ruling, the court held that Lasercomb's copyright was unenforceable until the plaintiff had purged the effects of its misuse.[82] In its analysis, the court acknowledged that the existence of the copyright misuse defense was unclear, and proceeded to conduct an extensive analysis of the history of the misuse doctrine and the public policy underlying the patent and copyright laws. The court finally concluded that the patent misuse doctrine first enunciated in *Morton Salt* should be incorporated into copyright law,[83] and found copyright misuse despite the fact that the actions of the plaintiff had no direct effect on the defendants[84] and despite the fact that no antitrust violation had been found.[85] Since the

---

81. 911 F.2d at 978.

82. *Id.*

83. Although the patent misuse defense has been generally recognized since *Morton Salt*, it has been much less certain whether an analogous copyright misuse defense exists .... We are of the view, however, that since copyright and patent law serve parallel public interests, a "misuse" defense should apply to infringement actions brought to vindicate either right. As discussed above, the similarity of the policies underlying patent and copyright is great and historically has been consistently recognized. Both patent law and copyright law seek to increase the store of human knowledge and arts by rewarding inventors and authors with the exclusive rights to their works for a limited time. At the same time, the granted monopoly power does not extend to property not covered by the patent or copyright .... Thus, we are persuaded that the rationale of *Morton Salt* in establishing the misuse defense applies to copyrights.

Id. at 976-77.

84. [A]gain analogizing to patent misuse, the defense of copyright misuse is available even if the defendants themselves have not been injured by the misuse. In *Morton Salt*, the defendant was not a party to the license requirement that only Morton-produced salt tablets be used with Morton's salt-depositing machine. Nevertheless, suit against defendant for infringement of Morton's patent was barred on public policy grounds.

Id. at 979.

85. So while it is true that the attempted use of a copyright to violate antitrust law probably would give rise to a misuse of copyright defense, the converse is not necessarily true—a misuse need not be a violation of antitrust law in order to comprise an equitable defense to an infringement action. The question is not whether the copyright is being used in a manner violative of antitrust law (such as whether the licensing agreement is "reasonable"), but whether the copyright is being used in a

Lasercomb decision, four other circuit courts have recognized the doctrine of copyright misuse.[86]

IV. THE PROPER TEST TO ENHANCE UTILITARIAN PUBLIC POLICY: ANTITRUST "RULE OF REASON" OR MISUSE "SCOPE OF THE GRANT"?

Although most courts have accepted that some version of the copyright misuse doctrine should exist, questions remain over its application. One view holds that misuse should be found only when an antitrust violation occurs. In general, the antitrust laws are intended to prohibit situations that unreasonably harm competition, and can apply to various uses of intellectual property rights. Another view argues that the misuse doctrine should be applied when the scope of the intellectual property grant has been exceeded. This view asserts that the intellectual property laws have balanced the rights granted to creators with those granted to the public, and seeks to prevent the extension of creators' rights beyond the scope granted by the intellectual property laws. As such, this "scope of the grant" view of misuse is independent of antitrust criteria such as market power.

The extension of copyright to computer software significantly alters this analysis. Software possesses features that provide it greater market power than other copyrighted works. Moreover, the grant of copyright protection to computer software can create additional disruptions in the normal balancing of rights granted to creators. Due to these factors, the software arena needs a copyright misuse doctrine unencumbered by antitrust analysis in order to preserve the public policy rationales of copyright law.

> manner violative of the public policy embodied in the grant of a copyright.

Id. at 978.

86. The 4th, 5th, and 9th Circuits have explicitly recognized the doctrine of copyright misuse, while the 1st and Federal Circuits have cited *Lasercomb* with approval but have not applied the doctrine. DSC Communications Corp. v. DGI Technologies, Inc., 81 F.3d 597 (5th Cir. 1996); Triad Systems Corp. v. Southeastern Exp. Co., 64 F.3d 1330 (9th Cir. 1995); Data General Corp. v. Grumman Systems Support Corp., 36 F.3d 1147 (1st Cir. 1994); Atari Games Corp. v. Nintendo of America Inc., 975 F.2d 832 (Fed. Cir. 1992).

In the 2nd, 3rd, 6th, 7th, 8th, 10th, 11th, and DC Circuits the existence of the doctrine is an open question (either the issue has not been raised since *Lasercomb* or the courts have not had to rule on the issue). The 3rd, 7th, 8th, and 11th Circuits, however, have previously indicated approval of the doctrine. *See supra* note 72.

A. The Misuse Analysis Provides Advantages Over An Antitrust Analysis In Balancing The Rights Granted

The decisions in both *Morton Salt* and *Lasercomb* endorse a strong version of intellectual property misuse which measures the actions of the intellectual property owner against the scope of the grant of intellectual property rights accorded to the owner. This scope of the grant view of misuse asserts that the intellectual property laws have established a ceiling on the level of rights granted to creators, and seeks to prevent any extension of creators' rights beyond this ceiling. Since this analysis is not related to restraints on competition, this version of misuse is independent from antitrust. In addition, those opposed to an antitrust-based approach to misuse have argued that the goals of the intellectual property laws and the antitrust laws are different, and that a misuse offense can occur even if antitrust laws are not violated.[87]

1. *AN ANTITRUST ANALYSIS FOR INTELLECTUAL PROPERTY*

While the explicit use of the copyright misuse doctrine is relatively recent,[88] questions about the application of antitrust principles to patent and copyright situations have a long history. As was the case in *Morton Salt*,[89] it is not unusual for an antitrust law claim to accompany a misuse claim. In general, the antitrust laws are embodied in the Clayton Act[90] and the Sherman Act,[91] and are intended to prohibit situations that may harm competition.[92] It is generally necessary to show that the accused

---

87. *See supra* notes 63 and 85; *see also* Timothy H. Fine, *Misuse and Antitrust Defenses to Copyright Infringement Actions*, 17 HASTINGS L.J. 315 (1965); Richard Stitt, *Copyright Self-help Protection as Copyright Misuse: Finally, the Other Shoe Drops*, 57 UMKC L. REV. 899 (1989); Julie E. Cohen, *Reverse Engineering And The Rise Of Electronic Vigilantism: Intellectual Property Implications Of "Lock-Out" Programs*; 68 S. CAL. L. REV. 1091 (1995); David Scher, Note, *The Viability of the Copyright Misuse Defense*, 20 FORDHAM URB. L.J. 89 (1992).

88. *Lasercomb*, 911 F.2d at 970.

89. *See supra* note 63.

90. *See* 15 U.S.C. §§ 12-27 (1997).

91. *See* 15 U.S.C. §§ 1-7 (1997).

92. "Every contract, combination in the form of trust or otherwise, or conspiracy, in restraint of trade or commerce among the several States, or with foreign nations, is hereby declared to be illegal." 15 U.S.C. § 1 (1997); "Every person who shall monopolize, or attempt to monopolize, or combine or conspire with any other person or persons, to monopolize any part of the trade or commerce among the several States, or with foreign nations, shall be deemed guilty of a felony ...." 15 U.S.C. § 2 (1997);

has market power in the relevant market, and courts subsequently use a "rule of reason" to analyze whether the contested practice has "unreasonably" restrained competition in light of the circumstances.[93]

Since the enforcement of intellectual property rights restrain the abilities of others to act, these rights can be considered to be limited, legally-imposed monopolies. However, despite these monopoly-like results, these aspects of intellectual property rights are part of the careful balance of rights, and should not implicate antitrust laws. Thus, while the enforcement of a patent or copyright will always involve some restraint of competition, it will not be an unreasonable restraint if applied within the normal scope of the intellectual property right.

2. *A COMPARISON BETWEEN AN ANTITRUST ANALYSIS AND A MISUSE ANALYSIS*

While court decisions like *Morton Salt* and *Lasercomb* endorse the view that misuse can occur even without a violation of the antitrust laws, this view is not universal. The circuit courts are split on this point.[94] Congress also examined this issue when it enacted the Patent Misuse Reform Act in 1988.[95] In that year, Congress considered limiting patent misuse to those actions which violated the antitrust laws in 1988, but declined to take such action.[96] Adherents to the Chicago School of Law

---

> It shall be unlawful for any person engaged in commerce ... to lease or make a sale or contract for sale of goods, ... or fix a price charged therefor, ... on the condition, agreement, or understanding that the lessee or purchaser thereof shall not use or deal in the goods, ... or other commodities of a competitor or competitors of the lessor or seller, where the effect ... may be to substantially lessen competition or tend to create a monopoly in any line of commerce.

15 U.S.C. § 14 (1997).

93. For a more detailed discussion of antitrust law doctrine, *see* William M. Landes & Richard A. Posner, *Market Power In Antitrust Cases*, 94 HARV. L. REV. 937 (1981).

94. While five circuits have indicated approval of the copyright misuse doctrine and it is an open issue in the other circuits, the question remains whether or not the copyright misuse defense is considered to be independent of an antitrust analysis. The 1st, 4th, 5th, 9th, and Federal Circuits appear to agree with the *Lasercomb* and *Morton Salt* proposition that misuse is independent of antitrust law, while the 7th Circuit appears to favor requiring an antitrust approach to misuse. *See supra* note 86 and Reed-Union Corp. v. Turtle Wax, Inc., 77 F.3d 909 (7th Cir. 1996).

95. *See supra* note 66.

96. The Act is a compromise between the Senate and the House over the necessity for and content of reform of the misuse doctrine. A Senate-passed bill would have provided that no patent owner would be deemed

and Economics,[97] including Judges Posner and Easterbrook of the Seventh Circuit, also argue strongly for an antitrust-based approach to misuse analysis[98] and they have been joined by other commentators.[99]

Proponents of the scope of the grant approach have countered several of the antitrust proponents' arguments. Two typical arguments raised in support of an antitrust-based approach to misuse stem from a

---

> guilty of misuse 'by reason of his or her licensing practices or actions or inactions relating to his or her patent, unless such practices or actions or inactions, in view of the circumstances in which such practices or actions or inactions are employed, violate the antitrust laws.'

5 CHISUM § 19.04(1)(f) at 19-294 (citing S. Rep. 100-492, 100th Cong., 2d Sess. (Aug. 25, 1988)).

97. For additional discussion on the Chicago School of Economics, see Toshiko Takenaka, *Extending The New Patent Misuse Limitation To Copyright:* Lasercomb America, Inc. v. Reynolds, 5 SOFTWARE L.J. 739, 746-48 (1992); Kobak, *supra* note 54, at 27-28.

98. *See* USM Corp. v. SPS Technologies, Inc., 694 F.2d 505, 511-12 (7th Cir. 1982), where Judge Posner argued:

> The [patent misuse] doctrine arose before there was any significant body of federal antitrust law, and reached maturity long before that law ... attained its present broad scope .... [T]here is increasing convergence of patent-misuse analysis with standard antitrust analysis .... One still finds plenty of statements in judicial opinions that less evidence of anticompetitive effect is required in a misuse case than in an antitrust case .... But apart from the conventional applications of the doctrine we have found no cases where standards different from those of antitrust law were actually applied to yield different results .... If misuse claims are not tested by conventional antitrust principles, by what principles shall they be tested? Our law is not rich in alternative concepts of monopolistic abuse; and it is rather late in the day to try to develop one without in the process subjecting the rights of patent holders to debilitating uncertainty.

See also infra note 168.

99. *See* Takenaka, *supra* note 97; Tony Paredes, *Copyright Misuse And Tying: Will Courts Stop Misusing Misuse?*, 9 HIGH TECH. L.J. 271 (1994); Roger Arar, Note, *Redefining Copyright Misuse*, 81 COLUM. L. REV. 1291, 1311 (1981); Byron A. Bilicki, *Standard Antitrust Analysis and the Doctrine of Patent Misuse: A Unification Under the Rule of Reason*, 46 U. PITT. L. REV. 209 (1984); J. Dianne Brinson, *Patent Misuse: Time for a Change*, 16 RUTGERS COMPUTER & TECH. L.J. 357 (1990); Scott A. Miskimon, *Divorcing Public Policy From Economic Reality*, 69 N.C. L. REV. 1672 (1991); Philip Abromats, Comment, *Copyright Misuse and Anticompetitive Software Licensing Restrictions: Lasercomb America, Inc. v. Reynolds*, 52 U. PITT. L. REV. 629 (1991).

belief that antitrust law and intellectual property public policy are both based on similar goals (that of promoting consumer welfare through a combination of free competition and innovation),[100] and that antitrust principles provide more certainty than those of misuse.[101] Several commentators, however, have pointed out that the intellectual property goal of stimulating the creation and distribution of creative works is different from the antitrust goal of encouraging marketplace competition,[102] and that there are instances in which an antitrust inquiry alone fails to prevent all abuses of the intellectual property grant that may harm the public.[103]

An antitrust analysis can also be particularly difficult to apply to the area of software. While an antitrust analysis is often considered to be more objective and predictable than a "scope of the grant" determination, commentators have argued that the antitrust analysis is no more certain than a misuse analysis in the area of software.[104] In particular, it can be extremely problematic or impossible to define the relevant market for software goods. Since antitrust law depends on the concept of market power, it can thus be difficult to apply to the software industry. For

---

100. *See* Paredes, *supra* note 99, at 275; Donald F. Turner, *Basic Principles in Formulating Antitrust and Misuse Constraints on the Exploitation of Intellectual Property Rights*, 53 ANTITRUST L.J. 485, 485 (1984).

101. *See* Paredes, *supra* note 99, at 291.

102. *See* Ramsey Hanna, Note, *Misusing Antitrust: The Search For Functional Copyright Misuse Standards*, 46 STAN. L. REV. 401, 419-21 (1994) (while antitrust and intellectual property laws both have the ultimate objective of enhancing public welfare, they reach this objective through different intermediate goals which lead to different conclusions as to what practices to allow); Scher, *supra* note 87, at 98.

103. *See* Hanna, *supra* note 102, at 435-44 ("[T]he Supreme Court's current interpretation of antitrust law would hold harmless copyright holders engaging in a broad range of practices aimed at stifling product-based competition," particularly certain types of predatory pricing and non-pricing techniques in the software industry); Note, *Clarifying The Copyright Misuse Defense: The Role Of Antitrust Standards And First Amendment Values*, 104 HARV. L. REV. 1289 (1991); *see also* discussion in Part IV.A.3. *infra*.

104. *See* Hanna, *supra* note 102, at 431-32 ("[A]ntitrust analysis is particularly inadequate at measuring economic power in innovative dynamic industries, due to the difficulty of defining markets in industries with differentiated but highly substitutable goods."); Stephen J. Davidson & Nicole A. Engisch, *A Survey Of The Law Of Copyright Misuse And Fraud On The Copyright Office: Legitimate Restraints On Copyright Owners Or Escape Routes For Copyright Infringers?*, 448 PLI/PAT 489 (1996) (defining the relevant market for an antitrust analysis of most copyrighted works may be impossible, and the law of antitrust doesn't produce certain results because it varies by circuit and over time).

instance, software goods are frequently bundled with other products (including other software, hardware, or with services),[105] making measurement of market share difficult. Producers of software often leverage products in other seemingly unrelated markets, and use the desire for compatibility to achieve market power in seemingly unrelated areas.[106] Because these interconnections between unrelated markets in the software arena are difficult to define, the traditional antitrust analysis can lead to unpredictable results.

Thus, it has been argued that an antitrust-based approach to analyzing misuse is appropriate because of the common goals of antitrust law and intellectual property law, and the certainty provided by an antitrust analysis. However, many commentators have refuted these arguments, particularly as they apply to the software industry. Therefore, the justification for an antitrust-based approach to misuse is at best uncertain.

3. *ANTITRUST LAW DOES NOT APPLY WELL TO NETWORK ECONOMIC SITUATIONS*

While certainty and common goal arguments have been mentioned by commentators as reasons for an antitrust-based analysis for misuse, the primary thrust for an antitrust approach arises from the belief that antitrust's economic analysis allows for a more efficient utilization of the intellectual property rights granted.[107] For example, this argument (based on the Chicago School of Economics) asserts that an intellectual property right conveys only a limited amount of power, and that normally tying arrangements cannot extract more than that amount.[108] Moreover, ties

---

105. *See* Paredes, *supra* note 99, at 298, 309-15 ("[Software creators] are notorious for tying hardware, maintenance, and servicing to their software"). *But see* Thomas M. Jorde and David J. Teece, *Rule Of Reason Analysis Of Horizontal Arrangements: Agreements Designed To Advance Innovation And Commercialize Technology*, 61 ANTITRUST L.J. 579 (1993) (arguing that horizontal arrangements can be pro-competitive and a legitimate business strategy).

106. *See* discussion in Part IV.A.3. *infra*.

107. *See* Paredes, *supra* note 99, at 309-15; Kobak, *supra* note 54, at 25-32.

108. [T]he fixed sum [argument] is ... that a firm with market power may be able to gain its profit all from its own market, all from another, or from any combination thereof, but the total amount of restriction that the monopolist will profitably be able to impose is fixed regardless of the practice that is used.

Louis Kaplow, *Extension Of Monopoly Power Through Leverage*, 85 COLUM. L. REV. 515, 517-18 (1985).

can serve as a form of metering by allowing the seller to profit based on the amount of use (by extracting most or all of the profit from the tied item if it is consumed during use), and this metering can allow price discrimination which simultaneously maximizes access to the public and profits to the intellectual property owner.[109]

However, a relatively recent "network theory" of economics attacks these traditional antitrust economic assumptions. Adherents to network economic theory argue that traditional antitrust assumptions do not apply to certain industries (such as the software and Internet areas) and that traditional antitrust analysis is therefore inapplicable to these areas.[110] These adherents assert that in certain industries, systematic tendencies exist for inefficient technologies to become established and to resist replacement by superior alternatives.[111] These tendencies are particularly powerful if there is a strong need for compatibility between different purchasers or users of the products supplied. While it has long been clear that this phenomenon existed in situations where users were physically connected (e.g., a telephone system has little use if different people have different types of phones which cannot inter-communicate), commentators have recently realized that the same phenomenon can apply in the software industry where users need compatibility (e.g., users need to share files with other users and software programs need to interact with other software programs).[112] As users increasingly become interconnected through the Internet and intranets,[113] the problem in the

---

109. *See* Kobak, *supra* note 54, at 28-30.

110. *See generally* Mark A. Lemley, *Antitrust and the Internet Standardization Problem*, 28 CONN. L. REV. 1041 (1996); S.J. Liebowitz & Stephen E. Margolis, *Should Technology Choice Be A Concern Of Antitrust Policy?*, 9 HARV. J.L. & TECH. 283 (1996); Thomas A. Piraino, Jr., *The Antitrust Analysis Of Network Joint Ventures*, 47 Hastings L.J. 5 (1995); Kenneth W. Dam, *Some Economic Considerations In The Intellectual Property Protection Of Software*, 24 J. LEGAL STUD. 321 (1995).

111. *Id.*

112. Commentators at the Berkeley School of Law and Economics have pioneered much of this work. *See* Michael A. Cusumano, *Strategic Maneuvering And Mass Market Dynamics: The Triumph Of VHS Over Beta* (Consortium On Competitiveness And Cooperation, Haas School of Business, U.C. Berkeley Working Paper No. 90-95, 1990); Joseph Farrell, *Standardization and Intellectual Property*, 30 JURIMETRICS J. 35, 37 (1989); Peter S. Menell, *Tailoring Legal Protection for Computer Software*, 39 Stan. L. Rev. 1329 (1987); PAMELA SAMUELSON ET AL., *A Manifesto Concerning the Legal Protection of Computer Programs*, 94 COLUM. L. REV. 2308 (1994).

113. *See* discussion in Part V.A. *infra*.

software industry is exacerbated. These types of network situations do not fit well within the framework of antitrust law.

4. *THE ANTITRUST ANALYSIS DOES NOT ADDRESS THE PROPER SCOPE OF RIGHTS FOR CREATORS*

If this network theory of economics is correct, an antitrust analysis based on traditional economic theory will not identify many misuse situations in the software industry. But even if the analysis of the Chicago School is correct, allowing the full utilization of a granted right is only beneficial and efficient for society if that granted right conveys the appropriate degree of power. If the intellectual property owner is trying to enforce rights that go beyond those which they should rightfully possess based on public policy considerations, then allowing the full utilization of the owner's desired scope of rights will be counter-productive to the societal goals and public policy underlying the original granting of the right. Therefore, if enforcing a judicial doctrine such as intellectual property misuse is perceived to maintain the appropriate balance of rights between those granted to the creator and those retained by the public, then arguing for full utilization of granted rights is premature because the scope of the rights granted to a creator must be defined before the rights can be enforced.

However, using a judicial doctrine to define the appropriate scope of rights for a "scope of the grant" analysis has its own problems. In particular, such a scheme would create debilitating uncertainty if intellectual property owners do not know the extent of the rights granted to them until after judicial intervention in an enforcement action. In this situation, any benefits achieved by a careful balancing of rights would likely be overshadowed by the transaction costs associated with the system. Such a system would thus be unworkable if each individual creator needed to rely on a judicial pronouncement for a definition of his rights.

One solution, however, is to balance the rights for a class of works rather than for each individual work.[114] A small number of judicial decisions could then define the scope of the rights granted to creators of works in each class, and the resulting uncertainty would be no more than exists in any common law system of jurisprudence, including that of antitrust law. Of course, such a system would be practical only if the determination of class membership was simple, and would be reasonable only if the works in the class had shared features which altered the

---

114. Different grants for different classes of works are not unusual. *See supra* text accompanying note 34.

normal balance of rights and thus justified a judicial restriction on the standard statutory grant.

B. The Unique Status Of Software As A Copyrighted Work

As discussed above, some commentators argue that a judicial doctrine of intellectual property misuse is appropriate only if an antitrust-based analysis is used, primarily because an intellectual property owner should be able to engage in a full utilization of the intellectual property rights granted. Even if the underlying theory of economics behind this argument is valid as applied to the software industry, the argument is preempted if the misuse doctrine is needed to redefine the appropriate extent of rights that are granted in the first place. But if a judicial determination of rights engenders excessive uncertainty on the part of intellectual property owners and others, then the system will be unworkable. This can only be avoided if the scope of the rights granted can be generalized to a class of works, which then provides the necessary predictability for other works in that class. For the reasons outlined below, computer software is such a class of copyrighted works.

1. *THE UNEASY BALANCE OF SOFTWARE AND COPYRIGHT LAW*

An examination of the history of copyright protection for computer software demonstrates a tumultuous road. Because software does not fit well within the traditional confines of copyright law, many difficult issues have arisen, and many still remain to be resolved.

While some would argue that computer programming first began in the nineteenth century with Charles Babbage and his Analytical Engine, modern computer software traces its origins to the creation in the 1940s of the first electronic computer. In response to copyright issues raised by the increasing prevalence of computer software in the 1960s and 1970s, Congress established in 1974 a National Commission on New Technological Uses of Copyrighted Works (CONTU). Based on the 1978 CONTU final report, Congress amended the Copyright Act in 1980[115] to make explicit that copyright protection extended to computer software.

In the seventeen years since the 1980 amendments to the Copyright Act, the following major issues are among those that have arisen related to software copyrightability: whether software object code (as distinct

---

115. 17 U.S.C. § 101 (1997) (added definition of "computer program"); 17 U.S.C. § 117 (1997) (added section entitled "Limitations on exclusive rights: computer programs"); *see also supra* note 30.

from software source code) is copyrightable;[116] whether software copyright protection violates the *Baker v. Selden*[117] prohibition on protection for a system described in a copyrighted work;[118] whether software copyright protection violates 17 U.S.C. §102(b)'s prohibition on protection for ideas, procedures, systems or processes;[119] whether the structure, sequence and organization of a program are copyrightable (as opposed to the textual computer code);[120] whether computer menu command hierarchies are copyrightable;[121] whether the "look and feel" of computer screen displays are copyrightable;[122] whether temporary copies of software in Random Access Memory (RAM) are infringing;[123] whether reverse engineering of software is infringing;[124] and whether attempts to deny software purchasers the statutory rights granted to "owners" (such as the right to resale or to make backup copies) can be enforced through mass-market shrink-wrap licenses.[125]

For many of these issues the law is not yet clear, and new issues continue to arise in abundance, particularly with relation to software on the Internet. Although these examples do not suggest that copyright protection for software is inherently wrong, they do illustrate that fitting software into previously existing copyright doctrines has been akin to fitting a square peg in a round hole.[126] At a minimum, these problems illustrate that copyrighted software works share features that distinguish them from other classes of copyrighted works.

---

116. Apple Computer, Inc. v. Franklin Computer Corp., 714 F.2d 1240 (3rd Cir. 1983).

117. 101 U.S. 99 (1879).

118. *Apple*, 714 F.2d at 1250-51.

119. *Id.*

120. *See* Whelan Associates v. Jaslow Dental Laboratory, Inc., 797 F.2d 1222 (3rd Cir. 1986), *cert denied*, 479 U.S. 1031 (1987); Computer Associates International, Inc. v. Altai, Inc., 982 F.2d 693 (2nd Cir. 1992).

121. Lotus Development Corp. v. Borland International, Inc., 49 F.3d 807 (1st Cir. 1995), *aff'd by an equally divided Court*, 116 S. Ct. 804 (1996).

122. Apple Computer, Inc. v. Microsoft Corp., 35 F.3d 1435 (9th Cir. 1994), *cert denied*, 115 S. Ct. 1176 (1995).

123. MAI Systems Corp. v. Peak Computer, Inc., 991 F.2d 511 (9th Cir. 1993).

124. *See* Sega Enterprises v. Accolade, Inc., 977 F.2d 1510 (9th Cir. 1992); Atari Games Corp. v. Nintendo of America, Inc., 975 F.2d 832 (Fed. Cir. 1992); *see also infra* Part V.B.

125. ProCD v. Zeidenberg, 86 F.3d 1447 (7th Cir. 1996).

126. *Computer Assoc.*, 982 F.2d at 712 (2nd Cir. 1992).

2. *SOFTWARE IS A FUNCTIONAL WORK AND COPYRIGHT PROVIDES IT MARKET POWER & PATENT-LIKE PROTECTION*

Computer software embodies a group of works that is distinct from other copyrighted works.[127] One such distinction arises because software is a functional work. For other forms of copyrighted works such as books or paintings, perceiving the expression is the goal of the work. This is true even for a work such as a movie or a sound recording in which the item embodying the work (a film, a phonorecord, a compact disk, etc.) requires a machine to allow the user to perceive the work.

On the other hand, software's inherent value lies in its ability to perform a function or task when its instructions are executed on a machine.[128] The user almost always lacks interest in the particular expression that the software engineer used to accomplish the result (e.g., while another software engineer might comment on the aesthetic beauty of how Sally used a tail-recursive subroutine rather than a looping construct or how Bob structures his object classes to promote reusability of code, most users merely care whether the program accomplishes its task in a minimum amount of time without crashing the computer). While a computer screen is often needed to perceive the code for a computer program, only the very rare user acquires software out of a desire to see the expression that is present in the lines of software code. Thus, while users of other copyrighted works are interested in perceiving the expression, a software user is interested in the results that are accomplished when this expression is executed and performs a function.

---

127. Marshall Leaffer, *Engineering Competitive Policy and Copyright Misuse*, 19 U. DAYTON L. REV. 1087, 1090 (1994) (arguing that computer software differs significantly from other types of literary works); Paredes, *supra* note 99, at 309-15 (1994) (noting that research and development costs are higher for software than for other types of copyrighted works).

128. *See* Samuelson, *supra* note 112, at 2315-19; Hanna, *supra* note 102, at 409-10 ("[U]nlike other types of copyrightable subject matter, software serves a functional purpose" and superior functionality gives its producer a competitive edge); Leaffer, *supra* note 127; Note, *Clarifying The Copyright Misuse Defense: The Role Of Antitrust Standards And First Amendment Values*, 104 HARV. L. REV. 1289, 1299 (1991) ("Computer software, however, differs from other works of authorship in ways that make it more like a patented invention. What users value most in a computer program is rarely the expression contained in its coded instructions; rather, it is the utility of the program in accomplishing some purpose."); *see also* qad. inc. v. ALN Associates, Inc., 770 F. Supp. 1261, 1265 (N.D. Ill. 1991).

The extension of copyright protection to a functional work is a factor that grants greater power to a software copyright owner than to other copyright owners, and necessitates an adjustment of the scope of the rights granted to these copyright owners so that an appropriate balance between the owner and the public is maintained. Copyright protection for software can be used to withhold access to the ideas underlying the expression,[129] and can also prevent others from building compatible products thereby leveraging existing technology.[130]

In a network industry such as software, market conditions often dictate the need for a single standard, thus effectively destroying competition.[131] Even a small difference in initial market power can tip the balance of power between competing products in a new market segment. And once a product has larger market share than competitors, the network externalities that stem from the desire for compatibility with the installed user base act to perpetuate and increase the market share (i.e., software developers will increasingly support the standard to the exclusion of others, independent service organizations will arise to train and repair problems with that software, hardware vendors will preload the software on the computers that they sell, etc.). If an existing dominant software vendor can tie a new entry in the market to its existing dominant products or standards, new entrants in the marketplace have little opportunity to compete.[132] To the extent that software copyright is

---

129. *See infra* Part IV.B.3.

130. Computer software may be characterized as a "cumulative technology," an attribute it shares with other rapidly developing areas of technology. In "cumulative technology" industries, future advances build incrementally on previous recent developments .... If protection is given too broad a scope, it may retard the advancement of innovation in an industry (like the software industry) in which innovation proceeds through sequential development. Many economists would agree that getting technology to the market sooner will raise consumer welfare and encourage further development of new generations of products, particularly in industries with cumulative technological characteristics.

Leaffer, *supra* note 127, at 1095.

131. *See* Samuelson, *supra* note 112.

132. Many observers point to Microsoft as an example of this network economic phenomenon in the software industry. Microsoft received a dominant position in the personal computer operating system market with its MS-DOS® product (this initial dominant position arose from the endorsement by the then-dominant hardware manufacturer IBM, who did not at the time recognize the increasing importance of the software industry). Microsoft then used that near-monopoly in MS-DOS to leverage a

used to enhance and maintain this type of "monopoly", it can confer significant market power[133] and make software copyright protection more akin to that provided by patent law (which protects functional ideas).[134]

This software copyright market power creates a difficulty for those who argue that the protection afforded by copyright will not generally afford sufficient market power to justify a copyright misuse doctrine.[135]

---

near-monopoly in the Windows™ operating system market, which in turn provided a near-monopoly in the productivity applications market (such as Office, Word, Excel). Now, Microsoft is attempting to leverage its near-monopoly into dominance in the network and Internet arenas. A commentator sums up the industry's mix of admiration, jealousy, and disdain of Microsoft well:

> Microsoft is inevitable. It will dip into endlessly deep pockets to buy anything, or anyone, to feed its infinite appetites. It will out-work, under-price, and out-last any competition. It will own any market it chooses, and punish customers who buy from the competition. Whether you are a competitor or a customer, you will be assimilated. Resistance is futile.

Dan Gillmor, *Microsoft Empire Lacks Jedi Adept At Striking Back*, SAN JOSE MERCURY NEWS, Mar 24, 1997 at E1.

133. *See* Hanna, *supra* note 102, at 409-10, 427-32 ("Rivals may not be able to effectively compete by introducing functionally equivalent products if they lack access to necessary productive resources or expertise, or if they fear offending a dominant software producer's copyrights." In particular, if copyright law prevents a rival from reverse engineering software, significant market power can result.)

134. *Id.* at 415-16 ("Computer applications and operating systems, while considered 'literary works' for purposes of copyright, resemble patented inventions in that consumers value such computer programs for their functional utility, not their artistic expression."); P. Samuelson, *CONTU Revisited: The Case Against Copyright Protection for Computer Programs in Machine-Readable Form*, 1984 DUKE L.J. 663 (1984); Paul Goldstein, *Infringement of Copyright in Computer Programs*, 47 U. PITT. L. REV. 1119, 1128 (1986) (arguing that misuse should apply in the software context where the protection is for many practical purposes the same as patent law); Peter S. Menell, *An Analysis of the Scope of Copyright Protection for Application Programs*, 41 STAN. L. REV. 1045, 1102 (1989) (noting that courts should not allow owners of copyrights in user interfaces from misusing them).

135. The Court has historically treated tie-ins involving intellectual property more harshly than ties involving other goods or services.... [T]he Court presumed that patents and copyrights provided the seller with sufficient economic power over the tying product to foreclose competition in the tied market.... With regard to copyrights, this presumption of anticompetitive effects was not justified.

Paredes, *supra* note 99, at 302-03; "Since the requirements for a patent are more demanding than for copyrights, patents confer more market power than copyrights." *Id.* at 305;

To the extent that copyrighted software holds such market power, this argument falls short.

3. *COPYRIGHT LAW CAN BE USED TO EXTEND PATENT-LIKE PROTECTION TO SOFTWARE IDEAS AND TRADE SECRETS*

In addition to the functionality aspects of software which distinguish it from other copyrighted works, the distribution of software in object code format allows the copyright owner to prevent access to important parts of the work. These excluded parts include not only the original creative expression of the software engineer (which is embodied in the source code), but also the ideas that underlie the expression and to any trade secrets that are contained in the software.[136] In fact, computer software may be the only product that simultaneously receives patent, copyright, trademark and trade secret protection for the same aspects of the product. Even if a user is allowed and is able to reverse engineer the object code, the resulting reverse engineered end-product will contain less information than the original creative work for which the copyright owner obtained protection.[137] Courts and commentators alike have criticized

---

"[B]ased on the reduced scope of exclusionary power, at least in theory, of many copyrights, an inquiry into market power would seem even more germane for copyrights than for patents." *Kobak, supra* note 54, at 34 .

136. As many commentators have previously described, computer software is normally developed in a high-level computer language such as C++, Lisp, or Java. These languages are readable to a trained developer, and the resulting source code allows significant expression in such things as choices of names for variables and functions, or the specific manner in which a result is achieved. Generally, this source code is then compiled into object code, which is composed of 1s and 0s, and which can be executed by a computer but which does not convey useful information to a human. If only the object code is distributed by the software copyright owner, which is the normal practice, then a user will never be able to see the original source code that produced the object code. For more detail, see cases cited *supra* note 124.

137. It is sometimes possible to engage in a form of reverse engineering called decompilation whereby the goal is to produce the object code from the original source code. This is typically time-consuming and very difficult, and it is not possible to retrieve all of the original expression. Since reverse engineering of software will typically involve creating at least interim copies of the object code or code that is derived from the object code, this reverse engineering is arguably an act of copyright infringement. *See* Leaffer, *supra* note 127, at 1090-91 ("To develop either competitive or compatible products, interim copying into an intelligible medium may therefore be necessary to study a program's sequence and logic .... Prohibition of reverse engineering through decompilation would erect a serious obstacle to developers who legitimately desire to create compatible

the use of copyright law to prevent reverse engineering and thereby prevent access to expression and protect both underlying ideas[138] and trade secrets.[139]

---

software, which many would argue is essential to innovation in the computer industry."); *see also infra* Part V.B.

138. "[T]he unique attributes of computer software [allows] ... an author [to acquire] ... patent-like protection by putting an idea, process, or method of operation in an unintelligible format and asserting copyright infringement against those who try to understand that idea." *Leaffer, supra* note 127, at 1094.

Courts have also recently made broad statements, albeit at least partially in dicta.

> As discussed above, the fact that computer programs are distributed for public use in object code form often precludes public access to the ideas and functional concepts contained in those programs, and thus confers on the copyright owner a de facto monopoly over those ideas and functional concepts. That result defeats the fundamental purpose of the Copyright Act—to encourage the production of original works by protecting the expressive elements of those works while leaving the ideas, facts, and functional concepts in the public domain for others to build on.

Sega, 977 F.2d at 1527;

> An author cannot acquire patent-like protection by putting an idea, process, or method of operation in an unintelligible format and asserting copyright infringement against those who try to understand that idea, process, or method of operation. [Citations omitted.] The Copyright Act permits an individual in rightful possession of a copy of a work to undertake necessary efforts to understand the work's ideas, processes, and methods of operation.

Atari, 975 F.2d 842; *see also* David A. Rice, *Public Goods, Private Contract and Public Policy: Federal Preemption of Software License Prohibitions Against Reverse Engineering*, 53 U. PITT. L. REV. 543, 611 (1992) (noting that reverse engineering clauses in software licenses should be constitutionally preempted by copyright law).

139. In *Vault Corporation v. Quaid Software Ltd.*, 655 F. Supp. 750 (E.D. La. 1987), *aff'd Vault Corporation v. Quaid Software Ltd.*, 847 F.2d 255 (5th Cir. 1988), a case criticized by some commentators, the court preempted a software license, which disallowed reverse engineering of the software, in part because the contractual provision would prevent access to unprotectable trade secrets and thus was violative of patent law.

> Vault's cause of action based on the Louisiana trade secrets act is not preempted by the federal Copyright act, on its face. *See* Kewanee Oil Co. v. Bicron Corp., 416 U.S. 470, 94 S. Ct. 1879, 40 L. Ed. 2d 315 (1974). State trade secret protection is much less effective than copyright and patent protection, and so is not preempted by federal law. One of the major weaknesses is the inability of trade secret laws to protect the

C. A "Scope Of The Grant" Copyright Misuse Doctrine Should Be Applied To Computer Software

As discussed above, the scope of the rights granted for intellectual property works is determined by a combination of statutory and judicial doctrines, and is designed to provide a balance between the rights granted to creators and those retained by the public. If a class of copyrighted works shares features that provide an excess of power to the copyright owner, a judicial determination is needed to realign the extent of rights that are granted. As discussed earlier, computer software is such a class of copyrighted works and owners of computer software copyrights thus possess an excess of rights relative to what is appropriate based on public policy. The misuse doctrine is the appropriate vehicle to realign the rights granted because it focuses on the scope of the rights granted and upholds the public policies underlying the intellectual property laws.

Adherents of an antitrust-based approach to copyright misuse rely on arguments that do not fit well within the domain of software copyright misuse. The copyright goal of promoting creative works is different from the antitrust goal of promoting competition, and copyright and antitrust laws achieve their goals through different mechanisms. In addition, defining the relevant markets for an antitrust-based approach to copyrighted software is extremely difficult, thus making the application of antitrust law uncertain. More importantly, the argument that copyrights do not convey sufficient market power to necessitate a misuse doctrine does not apply to software. Moreover, the argument to allow full utilization of the rights granted is preempted if the misuse doctrine is needed to define the appropriate level of rights that should be granted.

Because software is a class of copyrighted works that can provide greater power to the owner than is true of other classes of copyrighted works, the copyright misuse doctrine is a necessary vehicle to correct the imbalance of power between owners of copyrighted software works and the general public. As the *Lasercomb* court determined, the copyright

---

> owner against discovery of the trade secret by independent invention, accidental disclosure, or reverse engineering.

Id. at 475-76, 94 S. Ct. at 1883. "The Louisiana Software License Enforcement Act ... has gone beyond trade secrets law by outlawing reverse engineering." *Vault*, 655 F. Supp. at 763; *see also* Foresight Resources Corp. v. Pfortmiller, 719 F. Supp. 1006 (D. Kan. 1989) (finding that trade secret protection achieved through license restriction on decompiling program code was preempted, because it granted greater rights than Copyright Act in precluding exercise of section 117 rights); Rice, *supra* note 138, at 611-13 (arguing that reverse engineering clauses in software licenses should be preempted by patent law because they unduly protect trade secrets).

misuse doctrine should be applied without the need for an antitrust-based approach, and courts should instead rely on the scope of the copyright grant as determined by public policy.

V. THE PAST USE OF OTHER JUDICIAL DOCTRINES TO LIMIT SOFTWARE COPYRIGHT IN MISUSE CONTEXTS

This paper has argued that software is a unique form of copyrighted work that confers significant market power to the copyright owner. But if the copyright misuse doctrine is truly needed to mitigate the power of software copyright, the question arises as to why the first appellate decision enforcing misuse did not occur until 1990,[140] with only a gradual acceptance of the misuse doctrine in the time since. The answer to the question lies both in the recent emergence of the software industry itself and in the past use of other judicial doctrines by courts to address misuse-like situations.

A. Recent Growth Of The Software Industry Creates A Need For Copyright Limitations

The software copyright misuse doctrine's recent emergence can partially be attributed to the relatively recent ascent of the computer software industry itself. While high school students can probably not conceive of a world without computers (including their Sega Saturn™ and Nintendo-64™ game machines), it is important to realize that the first personal computer was not introduced until the late 1970s,[141] less than 20 years ago. Despite a current $150 billion a year packaged software global market that continues to grow at a rate near 15% a year,[142] this market was only $30-35 billion a year in 1990[143] and merely $18 billion a year in 1985.[144] More strikingly, the World Wide Web on the Internet did not even come into existence until the early 1990s. Yet Internet market revenue has grown to over $5 billion a year in 1995, and is projected to

---

140. *See supra* note 74.

141. While the introduction of the IBM PC in the early 1980s is sometimes considered to be the beginning of the personal computer industry, the Apple II computer, as well as other computers with more limited uses such as the Altair and the Radio Shack TRS-80, were available in the late 1970s.

142. *Server Market Is The Sweet Spot*, ASIA COMPUTER WEEKLY, July 29, 1996.

143. Mark S. Basham, *Computer Software Stocks: Core Segment Holdings In A Growth Stock Portfolio*, EMERGING & SPECIAL SITUATIONS 4, March 15, 1991.

144. Jim Mitchell, *Software Battle Entering Round 2: Copyright Trouble Looms Again*, THE DALLAS MORNING NEWS, May 25, 1985 at G1.

reach over $40 billion a year before the turn of the century[145] and $200 billion a year by 2006.[146] Thus, it is not surprising that a 1990 case, based on actions that occurred beginning in 1983,[147] was the first case in which the software copyright misuse doctrine was applied.

B. Employing Fair Use To Prevent Misuse

While relatively young, the software industry is highly competitive. This competitiveness dictates that at least some producers of software will use every tool at their disposal (including intellectual property rights), even if their use is questionable use or a misuse of an intellectual property right. Due to the paucity of precedent for applying the software copyright misuse doctrine, and perhaps limited by the pleadings of the parties, courts have in the past used other judicial doctrines to address instances of copyright misuse. In particular, the subjective and equitable doctrine of copyright fair use[148] has proven to be a powerful tool.

The most visible examples of the courts' use of fair use in the area of software arise with respect to reverse engineering[149] of software[150] in the

---

145. Margaret Ryan, *Midyear Forecast/ Annual Report*, ELECTRONIC ENGINEERING TIMES, July 1, 1996 at 25.

146. Louis Connor, *Can You Trust Web Transactions?; Digital Signatures And Other Security Measures Can't Prevent Crime-But They Can Deter It*, COMMUNICATIONS WEEK, Jan 15, 1996 at 41.

147. *Lasercomb*, 911 F.2d 970 (4th Cir. 1990).

148. *See* discussion in Part II.B.2. *supra.*

149. *See supra* note 137.

150. While reverse engineering cases are the most visible software fair use cases, other courts have applied fair use principles in other contexts. *See* Mitel, Inc. v. Iqtel, Inc., 896 F. Supp. 1050 (D. Colo. 1995) (finding that even if telephone call controllers' software program command codes were copyrightable, competitor's use of same command codes was "fair use" despite its commercial nature because manufacturer controlled between 75% and 90% of call controller market and competitor needed to copy the codes in order to be competitive); Triad Systems Corp. v. Southeastern Express Co., 31 U.S.P.Q.2d 1239 (N.D. Cal. 1994) (finding that if service organization can show that it needs to make backup copies of plaintiff's copyrighted software while reformatting customers' computer hard drives, it is likely to prevail on copyright infringement claim as a fair use, but summary judgment is not appropriate when a factual dispute over the necessity exists. While the court later found at 36 U.S.P.Q.2d 1028 that it was not fair use to load the plaintiff's diagnostic software into RAM while servicing the computer, the court did not appear to find a copyright violation with regard to the backup copies). *See also* Nimmer, *supra* note 53, at 23 (arguing that it should be fair use for an independent service organization to make

Sega[151] and *Atari*[152] cases. Since the legal analysis of the Federal Circuit in *Atari* is similar to that of the Ninth Circuit in *Sega*, only *Sega* will be discussed here. In *Sega*, the plaintiff, Sega, was a video game manufacturer which produced both game consoles and game cartridges that ran in the consoles, and the defendant, Accolade, was a game manufacturer who independently produced game cartridges that were designed to be compatible with the Sega consoles.[153] While Sega licensed some vendors to produce compatible games, Accolade was not a licensee. Instead, Accolade reverse engineered Sega game cartridges to discover the necessary software interfaces to make their games compatible with the Sega consoles. After discovering the interfaces, Accolade included the required 4-letter initialization code in their own game cartridges so that they could interact with the Sega console.[154] Sega attempted to use copyright law to prevent competition from Accolade in two ways—it claimed that Accolade's copying of software that occurred incidental to the reverse engineering process was infringement, and that the use of the 4-letter initialization code in Accolade's game cartridges was infringement.[155] The court rejected the proposition that these actions were copyright infringements, and instead found that Accolade's actions were protected by fair use.[156]

While the case was thus decided under the doctrine of fair use, there is a strong argument that the application of fair use was inappropriate, and that the misuse doctrine would have been more appropriate.[157] The

---

temporary copies of diagnostic software in RAM while performing repair or maintenance of customers' computers).

151. *See supra* note 124.

152. *Id.*

153. *Sega*, 977 F.2d at 1514.

154. *Id.* at 1514-16.

155. *Id.* at 1517.

156. "We conclude that where disassembly is the only way to gain access to the ideas and functional elements embodied in a copyrighted computer program and where there is a legitimate reason for seeking such access, disassembly is a fair use of the copyrighted work, as a matter of law." *Id.* at 1527-28. In addition, the court stated that the 4-letter code was undeserving of protection due to its functionality and its shortness. *Id.* at 1524-25.

157. It would appear that the court was limited by the pleadings of the parties—while copyright misuse was argued at the district court level, it did not appear to be an issue on appeal. 785 F. Supp. 1392 (N.D. Cal. 1992). In the *Atari* case, the Federal Circuit did consider the copyright misuse defense, but the willfully deceptive practices of the defendant in that case provided the plaintiff with a stronger equitable case. In the recent *Bateman v. Mnemonics* case involving reverse engineering of software, the 11th Circuit raised the possibility of misuse sua sponte, but concluded that infringement had not occurred due to

first indication that the fair use doctrine was the incorrect vehicle to use is that the traditional fair use analysis did not fit well with the defendant's commercial actions,[158] and so the court needed instead to rely heavily on public policy to justify its conclusion.[159] The second indication that the fair use analysis was inappropriate was the fact that the court found it necessary to make a broad pronouncement as a matter of law that would

---

the necessity to allow software compatibility without making an express finding on the legal doctrine used. 79 F.3d 1532, 1547 (11th Cir. 1996) and *infra* note 164.

158. Section 107 of the Copyright Act lists factors to be considered in determining whether a particular use is a fair one. Those factors include: (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyrighted work; (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and (4) the effect of the use upon the potential market for or value of the copyrighted work. *See* 17 U.S.C. § 107 (1997). The statutory factors are not exclusive.

The *Sega* court stretched the limits of reason in concluding that the first, third, and fourth factors all supported Accolade. The court argued that while Accolade copied the Sega code for the commercial purpose of creating a competing product, it was "a legitimate, essentially non-exploitative purpose, and that the commercial aspect of its use can best be described as of minimal significance." *Sega*, 977 F.2d at 1522-23. In addition, despite the statement by the court that if "the copying resulted in the latter effect [diminishing potential sales, interfering with marketability, or usurping the market], all other considerations might be irrelevant," the court concluded that there was no evidence of harm to Sega since consumers might buy Sega's games as well as Accolade's (even if they are of the same genre, like a football game). *Id.* at 1523.

As the court itself admitted, "[w]e are not unaware of the fact that to those used to considering copyright issues in more traditional contexts, our result may seem incongruous at first blush. To oversimplify, the record establishes that Accolade, a commercial competitor of Sega, engaged in wholesale copying of Sega's copyrighted code as a preliminary step in the development of a competing product." *Id.* at 1527.

159. [W]e are free to consider the public benefit resulting from a particular use notwithstanding the fact that the alleged infringer may gain commercially .... Public benefit need not be direct or tangible, but may arise because the challenged use serves a public interest. In the case before us, Accolade's identification of the functional requirements for Genesis compatibility has led to an increase in the number of independently designed video game programs offered for use with the Genesis console. It is precisely this growth in creative expression, based on the dissemination of other creative works and the unprotected ideas contained in those works, that the Copyright Act was intended to promote.

Id. at 1523.

affect software copyright owners as a class,[160] a result that was seemingly inconsistent with the fact-intensive nature of the traditional fair use analysis.[161] But the most telling indication that the fair use doctrine should not have been used was that the court went out of its way to denounce the actions of the plaintiff,[162] an unnecessary departure from the focus on the defendant in the normal fair use analysis. Since the court found it necessary to restrict the rights available to software copyright owners as a class and thus adjust the balance of rights, the fair use doctrine was inappropriate. With its focus on public policy and limiting

---

160. In light of the public policies underlying the Act, we conclude that, when the person seeking the understanding has a legitimate reason for doing so and when no other means of access to the unprotected elements exists, such disassembly is as a matter of law a fair use of the copyrighted work.

Id. at 1514. After *Sega* and *Atari*, it is now a commonly-held belief that reverse engineering of software is not copyright infringement if the reverse engineering is necessary to discover underlying software interfaces in the code that are required to build competing or compatible products. Thus, a right to prevent what would otherwise be copyright infringement has been withdrawn from software copyright owners as a class. *See* Leaffer, *supra* note 127, at 1087 ("*Sega Enterprises, Ltd. v. Accolade, Inc.* ... held that reverse engineering for purposes of developing non-infringing competing or compatible software is a fair use of copyrighted software").

161. Early in its opinion, the *Sega* court stated that "consideration of the unique nature of computer object code thus is more appropriate as part of the case-by-case, equitable 'fair use' analysis authorized by section 107 of the Act." *Sega*, 977 F.2d at 1520.

162. "[A]n attempt to monopolize the market by making it impossible for others to compete runs counter to the statutory purpose of promoting creative expression and cannot constitute a strong equitable basis for resisting the invocation of the fair use doctrine." *Id.* at 1523-24.

> Here, both parties agree that there is a misuse of a trademark, both agree that there is unlawful mislabeling, and both agree that confusion may result. The issue, here, is—which party is primarily responsible? Which is the wrongdoer—the violator? ... The facts are relatively straightforward and we have little difficulty answering the question. ... [W]e hold that Sega is primarily responsible for any resultant confusion. ... Sega seeks once again to take advantage of its trademark to exclude its competitors from the market. The use of a mark for such purpose is inconsistent with the Lanham Act.

Id. at 1528-30. "Sega argues that ... section 117 of the Act ... constitutes a legislative determination that any copying of a computer program other than that authorized by section 117 cannot be considered a fair use of that program under section 107. That argument verges on the frivolous." *Id.* at 1520-21.

over-reaching by copyright owners, the misuse doctrine instead should have been applied.

Thus, the *Sega* decision illustrates that when the court focuses on the actions of the plaintiff rather than those of the defendant, appeals to underlying public policy, treats software as a special class of copyrighted works,[163] and makes broad pronouncements of law, the court is misapplying the fair use doctrine. Since the court is effectively performing a misuse analysis that limits the rights of an entire class of copyrighted works, direct application of the misuse doctrine would provide more consistent and predicable results.

C. Employing Other Copyright Doctrines To Prevent Misuse

While the judicial doctrine of copyright fair use provides a powerful tool for courts to use in preventing misuse of software copyrights, the courts have other tools at their disposal. In *Bateman v. Mnemonics, Inc.*, a recent case related to reverse engineering of software, the Eleventh Circuit articulated this well:

> Whether the [copyright] protection is unavailable because these factors render the expression unoriginal, nonexpressive per 17 U.S.C. § 102(b), or whether these factors compel a finding of fair use, copyright estoppel, or misuse, the result is to deny copyright protection to portions of the computer program. Thus, we today join these other circuits in finding that external considerations such as compatibility may negate a finding of infringement.[164]

An example of another judicial doctrine that has been used to limit the rights of software owners who attempt to leverage their copyrights beyond allowable limits is the copyright preemption doctrine.[165]

---

163. [N]ot all copyrighted works are entitled to the same level of protection. ... Computer programs pose unique problems for the application of the "idea/expression distinction" that determines the extent of copyright protection .... Because of the hybrid nature of computer programs, there is no settled standard for identifying what is protected expression and what is unprotected idea in a case involving the alleged infringement of a copyright in computer software ... computer programs are also unique among copyrighted works in the form in which they are distributed for public use... . Because Sega's video game programs contain unprotected aspects that cannot be examined without copying, we afford them a lower degree of protection than more traditional literary works.

Id. at 1524-26.

164. 79 F.3d 1532, 1547 (11th Cir. 1996).

165. [A]ll legal or equitable rights that are equivalent to any of the exclusive rights within the general scope of copyright as specified by section 106 in

However, copyright misuse principles often provide a superior framework to restrain plaintiffs who abuse their copyright privileges. In *ProCD, Inc. v. Zeidenberg*,[166] a software manufacturer attempted to contractually restrict the uses of uncopyrightable telephone directory information[167] which could be downloaded from its product. While the Seventh Circuit court found that the shrink-wrap contract contained inside the product box was not preempted (in a ruling that was led by Judge Easterbrook and which relied heavily on a Chicago School economic analysis),[168] the district court judge had conversely found preemption due to the invalid extension of copyright power.[169] If the plaintiff truly was trying to extend

---

> works of authorship that are fixed in a tangible medium of expression and come within the subject matter of copyright as specified by sections 102 and 103, ... are governed exclusively by this title. ... [N]o person is entitled to any such right or equivalent right in any such work under the common law or statutes of any State.

17 U.S.C. § 301(a) (1997).

166. 86 F.3d 1447 (7th Cir. 1996).

167. While the telephone listings were not copyrightable, they were within the "subject matter" of copyright, and thus subject to § 301 preemption.

168. The court spent several pages reviewing economic theory, explaining why the restrictions in the shrink-wrap contract allowed price discrimination ("[t]o make price discrimination work, however, the seller must be able to control arbitrage") which was beneficial to consumers. *ProCD*, 86 F.3d at 1449-52. The court's statements indicating that "[c]ompetition among vendors, not judicial revision of a package's contents, is how consumers are protected in a market economy" and that "[t]erms and conditions offered by contract reflect private ordering, essential to the efficient functioning of markets" demonstrate the 7th Circuit's laissez-faire approach to intellectual property. *Id.* at 1453, 1455. One assumes that even the Seventh Circuit would still uphold traditional restrictions on contractual extensions of intellectual property rights, despite being "consensual" private orderings, such as the disallowance of required royalty payments that extend beyond the term of the intellectual property right. *See* Brulotte v. Thys Co., 379 U.S. 29 (1964) (holding that a patent owner cannot require royalty payments beyond the end of the patent term); F.E.L. Publications, Ltd. v. Catholic Bishop of Chicago, 214 U.S.P.Q. 409, 419 n.9 (7th Cir. 1982) ("[I]t is copyright misuse to exact a fee for the use of a musical work which is already in the public domain." (citing Mercoid Corp. v. Mid-Continent Investment Co., 320 U.S. 661, 60 U.S.P.Q. 21 (1994))).

169. Contracts that are consistent with the copyright law's goals of self-protection should be upheld. Rightful owners should be able to define the limits of permissible copying or modification of their works. ... It is only when a contract erects a barrier on access to information that under copyright law should be accessible that § 301 operates to protect copyright law from individually crafted evasions of that law. ...

the power of its copyright beyond the bounds allowed by public policy, the misuse doctrine provides a better framework with which to analyze these activities.

While the Seventh Circuit did not believe that the actions of the plaintiff in the *ProCD* case were an impermissible extension of copyright law, other courts have found differently under analogous situations. In the *Vault* case,[170] the Fifth Circuit found several provisions of the Louisiana Software License Enforcement Act (SLEA)[171] to be subject to preemption by the Copyright Act because the SLEA authorized contractual duplication or extension of rights granted under copyright law.[172] The SLEA allowed producers of software to impose various

---

> Plaintiff's license agreement is an attempt to avoid the confines of copyright law and of Feist, 499 U.S. 340, 111 S. Ct. 1282, 113 L. Ed. 2d 358. Its prohibition on the distribution of public information cannot be squared with the purposes of copyright law or with plaintiff's own compilation of data .... Plaintiff cannot use a standard form contract to make an end run around copyright law. Its contract claim is preempted by § 301.

908 F. Supp. 640, 658-59 (W.D. Wis. 1996); *see also* Rice, *supra* note 138, at 658-59 (explaining that the far-reaching public policy Section 301 implements clearly requires preemption of contract-based protection of expression where the effect is to secure rights in that expression which are greater than, equal to, or supplemental of those which section 106 secures); Robert P. Merges, *Intellectual Property and the Costs of Commercial Exchange: A Review Essay*, 93 MICH. L. REV. 1570, 1611 (1995) (stating that reverse engineering clauses in standard form software contracts, such as shrink-wrap licenses, should be preempted by patent law because they unduly protect trade secrets and act as a form of private legislation); Christopher Celentino et al., Vault Corp. v. Quaid Software Ltd.: *Invalidating Shrink-Wrap Licenses?*, 2 J.L. & TECH. 151, 162 (1987) (noting that *Vault* opinion has merit if it means that contracts providing rights greater than or equivalent to those under copyright law are preempted).

170. *See supra* note 139.

171. LA. REV. STAT. ANN. § 51:1961 (West 1987).

172. § 1964(2) of the SLEA prohibits copying "for any purpose." This section is in direct violation of the Copyright Act which permits the making of archival copies from copies of computer software and permits copying which is "an essential step in the utilization of the computer program." 17 U.S.C. § 117. *The SLEA has granted greater protection than the Copyright Act.*

> In addition, *§ 1964(3) prohibits "translating, reverse engineering, decompiling, disassembling, and/or creating derivative works based on the computer software."* The Copyright Act grants the owner of the copyright the right to prepare derivative works based upon the copyrighted work.

contractual terms upon software purchasers, provided that the terms were included with the software in a license agreement that comported with SLEA. The district court identified four areas in which the SLEA could be preempted by the copyright law, both constitutionally and under section 301.[173] It found that SLEA provided rights which exceeded those of the copyright law by allowing for prohibition of copying for any purpose, creating a bar against copying that was unlimited in duration, and allowing protection for any computer program, regardless of its originality.[174] In addition, the district court found that SLEA provided equivalent rights to those of the copyright law by allowing the prohibition of the creation of derivative works.[175] The Fifth Circuit, while affirming the reasons set forth by the district court, added yet another reason for constitutional preemption of SLEA—that the prohibition against reverse engineering does not allow the owner of a computer program to exercise the § 117 right to adapt the program as an essential step in its utilization,

---

> 17 U.S.C. § 106(2). The right to prepare derivative works is an exclusive right under the federal Copyright Act and *the SLEA cannot provide "an equivalent right."* 17 U.S.C. § 301(a). ...
>
> ....
>
> ... Congress has taken action to afford copyright protection to computer software. In this situation *the Sears-Compco preemption doctrine, as well as § 301 of the Copyright Act, are both applicable.*
>
> *The Louisiana Software License Enforcement Act creates a perpetual bar against copying any computer program licensed pursuant to its provisions.* The federal Copyright Act, on the other hand, grants protection against unauthorized copying for the life of the author plus fifty years. 17 U.S.C. § 302(a). The Louisiana act also places no restrictions on the programs which may be protected under its provisions. However, under Section 102 of the federal Copyright Act, only "original works of authorship" can be protected. *The Louisiana Software Act allows any computer program, original or not, to be protected from copying.*
>
> Since the Louisiana Software Act has "touch[ed] upon the area" of the federal patent and copyright law, the provisions of the PROLOK licensing agreement are unenforceable to the extent they are contrary to the policies of the federal Copyright Act. Fantastic Fakes, Inc. v. Pickwick International, Inc., 661 F.2d at 481-83.

Vault, 655 F. Supp. at 762-63 (emphasis added), *aff'd*, Vault Corp. v. Quaid Software Ltd., 847 F.2d 255 (5th Cir. 1988).

173. *Vault*, 655 F. Supp. at 762-63.

174. *Id.*

175. *Id.*

and thus gives rights to the owner of the copyright that exceeds those granted under the copyright law.[176] Thus, the circuit courts are split on what constitutes an impermissible extension of copyright under a preemption analysis. Since the misuse analysis focuses more directly on public policy, it is more appropriate than preemption to address such extensions of copyright law.

In addition to the use of the preemption doctrine, some courts have followed the *Bateman* suggestion to find works "nonexpressive per 17 U.S.C. § 102(b)" when allowing copyright protection would unduly extend the reach of copyright law.[177] For example, the West Publishing Company has long attempted to prevent competitors from launching similar software products by claiming copyright protection on the "star pagination" used in its compilations of uncopyrightable judicial opinions. In two recent cases, the courts have found the West star pagination to be undeserving of copyright protection,[178] again focusing on the inappropriate actions of the alleged copyright owner.

Thus, in only a short period of time, owners of software copyrights have used various techniques to extend their rights. Courts have addressed these actions using a variety of copyright doctrine analyses, including fair use, copyright preemption and section 102(b) "non-expressiveness." The erection of such disparate barriers to prevent abuses of software copyrights demonstrates the existence of a significant problem, and furthermore, the use of a single uniform approach would provide greater consistency and predictability. Since copyright misuse focuses on the actions of copyright owners and relies on underlying public policy considerations, a uniform approach to future examples of misuse under the penumbra of the software copyright misuse doctrine is a solution to the problem.

---

176. *Vault*, 847 F.2d at 270.

177. Section 102(b) states that copyright protection does not extend to any idea, procedure, process, system, or method of operation. *See supra* note 25.

178. Matthew Bender & Co. v. West Pub. Co., No. 94 Civ. 0589, 1997 WL 266972 (S.D.N.Y. May 19, 1997) (holding that West may not attempt to use copyright in its pagination to prevent its Hyperlaw competitor from entering the market); United States v. Thomson Corp., 949 F. Supp. 907, 928 (D.D.C. 1996) (upholding the requirement that West license its electronic pagination service because it is unlikely that the pagination is copyrightable after *Feist*, and even if it is copyrightable, the enforcement of the right "might significantly impede access to the text of court decisions, discourage the publication of secondary law products and hamper competition in the print, CD-ROM, and online research services markets").

VI. APPLYING SOFTWARE COPYRIGHT MISUSE TO TRADITIONAL MISUSE CONTEXTS

Having argued that software copyright misuse is a necessary and useful doctrine, the question next arises as to the proper application of the doctrine. Although the proper scope of the doctrine will ultimately be decided through a series of cases that develop underlying principles, it is appropriate to consider possible areas of application. One starting point arises from the established case law in the patent misuse area. As described previously, the two paradigmatic patent misuse cases involve tying arrangements (where the patentee requires the purchase of unpatented goods or services along with a patented product or process), and non-compete clauses (that prevent a patent licensee from producing or selling competing goods).[179]

A. Software Copyright Misuse And Tying Arrangements

Due to the complicated nature of software tying arrangements, a general resolution of the issues involved is beyond the scope of this paper. Although it was fairly easy to find an illegal tying arrangement between unpatented salt tablets and the patented salt-depositing machinery in *Morton Salt* (because it is easy to treat the salt tablets and the salt-depositing machinery as distinct products), it is more difficult to determine whether a computer system sold with the hardware and operating system bundled together is such a tying arrangement. In many past business situations, one company produced both the computer hardware and the only operating system that functioned on the hardware, thus requiring both parts to achieve a minimally useful machine. In the software industry today, it is common practice to bundle productivity applications together into "office" suites. However, this is bundling only because we conceptually conceive of the different parts of the suite as independent applications—as we increasingly move toward a computing world where data is king, the notion of distinct applications may disappear.[180]

---

179. *See supra* Part III.A.

180. In a data-centric model of computing, being increasingly embraced as one aspect of object-oriented programming, the data is the focal point and various operations are performed on the data to manipulate it as needed. For example, if office suites had originally been created around this notion, it is likely that a single document type would have been created by the suite. If a user desired to add words to the document, then word-processing functions would be used, and if a user desired to perform calculations or add graphics, then other corresponding functions would be used. Even if these different types of

These "tying" situations may be acceptable (or not tying at all), because it is easy to treat these tied products as a single product. At a minimum, a strong functional relationship between the tied products provides a necessity-based justification for the tying. However, in other types of software tying situations, the software copyright owner attempts to forbid incidental copying of software to prevent competition in goods or services that are unrelated to the software. These situations will be addressed below in relation to restrictions on competition.

B. Software Copyright Misuse And Restrictions On Competition

1. *USE OF NON-COMPETITION AGREEMENTS IN SOFTWARE LICENSES*

Not surprisingly, the first software copyright misuse case arose due to an attempted restraint on competition. As described above, the plaintiff in the *Lasercomb* case had included a provision in its standard form contract that not only limited copying of the licensed software, but also prevented licensees from independently developing a competing software program for 100 years.[181] The Fourth Circuit held that this restriction constituted copyright misuse. Other courts have found similarly,[182] and little argument has been advanced that an alternative policy would be desirable.

2. *ENJOINING INCIDENTAL RAM COPIES TO PREVENT DEVELOPMENT OF PRODUCTS*

Since the 1993 decision in *MAI* (which held that temporary copies of software made in RAM while the computer is turned on and running are sufficiently fixed to be infringing),[183] companies have tried to use copyright law to control various third party activities involving temporary copies, including the development of competing products and competing services. Since copyright law is designed to protect only the creative expression embodied in the source code, the functional effects of the object code should not be included within that protection unless they

---

functions were in fact embodied in different application programs (which wouldn't be required), this would be transparent to the user.

181. *See supra* Part III.B.

182. *See* Tamburo v. Calvin, No. 94 C 5206, 1995 WL 121539 (N.D. Ill. Mar. 17, 1995) (finding that a software license agreement which required licensees to forgo developing "any product, whether or not it is computer-based, that competes with [the TBS software]" constituted copyright misuse).

183. *See supra* note 123 and *infra* Part VI.B.3.

are integrally tied to the creative expression. Arguably, enjoying the functional results of an executing RAM copy of software is sufficiently related to the copy itself that the creation of the RAM copy is copyright infringement. However, when the creation of the RAM copy is incidental to the activity being controlled, a strong argument exists that it would be misuse to use the RAM copy to restrict the activity. For most computers, merely pressing the power switch to turn it on causes the operating system software to be loaded into RAM, thus creating an infringing copy.[184] Furthermore, in order to use a computer program it must be loaded into RAM,[185] whether the stored copy of the program normally resides on the computer's hard drive, on a floppy disk, on a CD-ROM, or on another machine on a computer network. Even if these RAM copies are infringing under a *MAI*-type rationale, the question remains whether the activities of the copyright owner constitute copyright misuse.

In the 1996 case of *DSC Communications Corp. v. DGI Technologies, Inc.*,[186] the Fifth Circuit found one such example likely to constitute

---

184. Arguably, even authorized users of most software products are actually infringers, because their shrink-wrap licenses rarely allow RAM copies to be made. However, companies have little incentive to charge legitimate customers with copyright infringement, and commentators have pointed out that a strong argument can be made for an implied license in these situations. *See* Nimmer, *supra* note 53, at 20-21 (noting that "any sane observer of the copyright scene would say that [loading a legitimate copy of software into computer RAM] has been implicitly licensed and is therefore noninfringing" despite the literal act of reproduction); *ProCD*, *supra* note 125, at 648-50 (finding that 17 U.S.C. § 117 allows copies of software created as an essential step in the utilization of the computer program in conjunction with a machine, and this includes both copies made in RAM and copies on the computer hard disk when made by a legitimate user).

185. There is a form of client-server computing in which the program is loaded into the RAM of a host computer and executed there, while the results of the execution display on the user's computer as if the user's computer were executing the program (this idea is used by some versions of the network computer or NetPC concept that has recently been popularized in the press, although other versions download an application from a host computer but run it locally). Even in this situation, a RAM copy is created somewhere, and the mere act of displaying results on the user's computer may create an infringing copy in a frame buffer or video RAM on the user's computer.

While outside the scope of this paper, other issues arise with respect to Web browsers. Any time that a web page is accessed, a copy of the page is made in the RAM on the local computer. Increasingly, copies of images, sound recordings, and even executable applets are included in this process. It is only a matter of time before someone attempts to treat a browser-based copy as an infringing copy under copyright law to restrict some type of activity.

186. 81 F.3d 597 (5th Cir. 1996).

copyright misuse. The plaintiff, DSC, manufactured telephone switching systems which were a specialized form of computer, including an operating system, at least one central computer processor and removable microprocessor cards.[187] When the cards are inserted into the system, they must "boot up", and in doing so they download a copy of the operating system into their RAM.[188] The defendant, DGI, was developing compatible microprocessor cards to run in DSC systems. In order to test its product, DGI needed to download the operating system into the card in the same manner that would occur during normal operation of the system.[189] In the context of a preliminary injunction requested by DSC to prevent DGI from making these copies, the Fifth Circuit found that DSC was unlikely to succeed on the merits because it was likely that their activities constituted copyright misuse.[190]

The principles advocated in this paper support the Fifth Circuit analysis. DSC attempted to extend the rights granted to one work to prevent the creation of other independent works. Moreover, the removable add-in cards were unrelated to the creative aspects of the operating system software. A contrary holding would thwart copyright law's underlying goal of fostering creation and innovation. The development of the add-in card by DGI required substantial creativity and innovation, and its use of the operating system for compatibility purposes was incidental. Granting a monopoly over products such as the add-in cards was not the type of incentive envisioned by the framers of

---

187. *Id.* at 598-99.

188. *Id.* at 599.

189. *Id.*

190. DSC seems to be attempting to use its copyright to obtain a patent-like monopoly over unpatented microprocessor cards. ... In order to ensure that its card is compatible, a competitor such as DGI must test the card on a DSC phone switch. ... If DSC is allowed to prevent such copying, then it can prevent anyone from developing a competing microprocessor card, even though it has not patented the card. The defense of copyright misuse "forbids the use of the copyright to secure an exclusive right or limited monopoly not granted by the Copyright Office," including a limited monopoly over microprocessor cards. ... Of course, we do not hold that DGI will successfully avail itself of the copyright misuse defense. After a trial on the merits, the district court may well decide that DSC did not commit copyright misuse, or that DGI cannot avail itself of the defense because it has "unclean hands."

Id. at 601.

the copyright laws for rewarding the creativity expressed in creating a software literary work.

3. *USE OF INCIDENTAL RAM COPIES TO PREVENT DEVELOPMENT OF SERVICES*

Although the Fifth Circuit held that extending copyright in functional RAM copies to prevent the creation of *products* unrelated to the copyright was misuse, several courts have found otherwise when *services* unrelated to the copyright were obstructed. However, this distinction between products and services is not warranted,[191] particularly as our economy becomes increasingly information-based and service-based rather than product-based.

The *MAI* case illustrates a blatant example of extending copyright in temporary RAM copies to unrelated services. In *MAI*, the Ninth Circuit found that the defendant had committed copyright infringement by merely turning on a MAI-produced computer at the direction of a MAI customer who had purchased the computer. [192] The plaintiff, MAI, manufactured and sold computers that included operating system software and other diagnostic software developed by MAI.[193] The defendant, Peak, provided servicing of computers, including routine maintenance and emergency repairs, that included turning on the computer and viewing the system error log (which is part of the operating system).[194] Turning on the computer was found to be sufficient to constitute copyright infringement, since a copy of the MAI operating

---

191. *See* Nimmer, *supra* note 53, at n.108 (arguing that the plaintiff's attempt in *MAI* to prevent defendant from committing acts of software reproduction that were essential for the proper competitive purpose of performing hardware diagnostics is analogous to the plaintiff's attempt in DSC to prevent the defendant from committing acts of software reproduction that were essential for the proper competitive purpose of developing a competing microprocessor card); Eastman Kodak Co. v. Image Technical Services, Inc., 504 U.S. 451 (1992) (finding that a genuine issue of material fact existed on whether micrographic equipment manufacturer unlawfully monopolized and attempted to monopolize sale of service and parts for machines, when manufacturer's policies limited availability to independent service organizations (ISOs) of replacement parts for its equipment, and made it more difficult for ISOs to compete with the manufacturer in servicing such equipment).

192. 991 F.2d 511 (9th Cir. 1993).

193. *Id.* at 513.

194. *Id.* at 518.

system was made in RAM as part of the boot process of the computer, even if the diagnostic software was not used by Peak.[195]

A consideration of slight variations in the fact pattern illustrates the absurdity of this holding. Under the scenario posited by the court, if the Peak repairman had arrived at the customer's site and the computer was fortuitously on, no copyright infringement would have occurred.[196] Alternately, infringement would have been avoided if the repairman had stood over the shoulder of an authorized user while the user pressed the on button. A legal analysis that determines whether a computer service can be provided or whether copyright infringement has occurred based on who presses the computer "on" button seems ridiculous. In a paraphrase of the *DSC* court, *MAI* is attempting to use its copyright to obtain a patent-like monopoly over unpatented maintenance services, and the defense of copyright misuse forbids the use of the copyright to secure an exclusive right or limited monopoly not granted by the Copyright Office.

The absurdity of the *MAI* result exists even if the copyrighted software was more closely related to the service of performing maintenance on a computer. If the defendant had made a RAM copy of diagnostic software rather than operating system software, then at least there would be some logical relationship between the copy being made and the service being performed. However, while this modification to the *MAI* scenario would at least provide a tenuous relationship between the activities of producing the software and providing the service, the result is no less absurd. The on-site repairman will now have to stand over the

---

195. Peak concedes that in maintaining its customer's computers, it uses MAI operating software "to the extent that the repair and maintenance process necessarily involves turning on the computer to make sure it is functional and thereby running the operating system." It is also uncontroverted that when the computer is turned on the operating system is loaded into the computer's RAM. As part of diagnosing a computer problem at the customer site, the Peak technician runs the computer's operating system software, allowing the technician to view the systems error log, which is part of the operating system, thereby enabling the technician to diagnose the problem. [FN4]

FN4. MAI also alleges that Peak runs its diagnostic software in servicing MAI computers. Since Peak's running of the operating software constitutes copyright violation, it is not necessary for us to directly reach the issue of whether Peak also runs MAI's diagnostic software.

Id.

196. Once the temporary RAM copy was created, the use of the copy would not seem to implicate any of the exclusive rights accorded the copyright owner.

shoulder of an authorized user until the user starts up the diagnostic software (rather than the operating system software), thus creating an authorized RAM copy which the repairman can use.

The problem with the *MAI* ruling is that it discriminates against different classes of users. A customer should be authorized to use its software in a normal fashion and a service provider should be entitled to act as an agent for the customer and take the same actions.[197] Instead in *MAI*, an inane distinction occurs based on the status of the repairman as an employee (if the company is big enough to afford a full-time system administrator) or an independent contractor. If a software producer has provided diagnostic software to a customer for the customer's use, later attempts to enjoin incidental copying in order to create a monopoly in servicing the hardware is an undeserved extension of the copyright.

Despite the obvious attempt on the part of MAI to extend its exclusive rights to an unrelated service, an important distinction can be made between computer servicing cases[198] and the *DSC* case. The competition being restrained in the computer servicing cases involves the provision of a service rather than the production of a good. Arguably, however, the distinction between goods and services should be irrelevant

---

197. One commentator argues that not only should such a software copy be considered a necessary copy under section 117 (one created when employing someone to run diagnostic software that is too complicated to be run by oneself), but that "a party who offers both hardware and software should be construed to have conveyed an implicit license to engage in diagnostics ... in order to make it function appropriately" and "[r]equiring customers to forego by contract the right that Congress accorded them by statute ... constitutes misuse." *Id.* at 23-25.

198. Several other courts have considered similar situations, and those who have considered a copyright misuse defense have rejected it. *See* Triad Sys. Corp. v. Southeastern Express Co., 64 F.3d 1330 (9th Cir. 1995); Service & Training, Inc. v. Data Gen. Corp., 963 F.2d 680 (4th Cir. 1992); Advanced Computer Services of Mich., Inc. v. MAI Sys. Corp., 845 F. Supp. 356 (E.D. Va. 1994). *But see* Data General Corp. v. Grumman Sys. Support Corp., 36 F.3d 1147, 1169-70 (1st Cir. 1994) ("Although DG correctly notes that the misuse in *Lasercomb* (conditioning a copyright license on a non-competition agreement) is not identical to the misuse alleged in this case (tying access to ADEX to the purchase of DG service and refusing to license ADEX to [third-party maintenance companies]), the reasoning of *Lasercomb* does not turn on the particular type of anti-competitive behavior alleged"). However, since Grumman had not argued a non-antitrust copyright misuse claim, the court declined to find one). In addition, at least one commentator has criticized these rulings. *See* Nimmer, *supra* note 53, at n.98 ("Yet none [of the courts that have considered copyright misuse in connection with a diagnostic computer service case] articulates a satisfactory formulation of what constitutes misuse in order to find the conduct outside its scope").

in today's society. As our society becomes increasingly specialized and the portion of our economy devoted to providing services grows, it is difficult to argue that an innovative service is less valuable from a public policy perspective than an innovative good. For example, the innovative service from Federal Express to provide overnight delivery of packages at economical prices has revolutionized how business is done. In addition, there is no principled reason why the creation of services cannot be covered under the constitutional authorization to promote "writings."[199]

Thus, attempts to use copyright protection to stifle competition should be met with judicial disapproval, regardless of whether the competition comes in the form of goods or services. If the underlying copy of the software from which the RAM copy was created is itself a legitimate copy, then the incidental copy created in RAM should not be treated as infringement for the purposes of restricting competition. Attempts to alter this result through license agreements with customers should be treated as an undue restraint that constitutes misuse of the rights granted by copyright law.

VII. APPLYING SOFTWARE COPYRIGHT MISUSE TO CURRENT SOFTWARE PRACTICES

While considering the paradigmatic patent misuse cases is an appropriate starting point in defining the necessary scope of software copyright misuse, it is also necessary to consider those aspects of software which make it unique as a copyrighted work and to impose some limits where the unique nature of software grants excessive power to software copyright owners. The intent of this section is not to provide an exhaustive list or an in-depth analysis, but rather to suggest areas that deserve special judicial attention.

A. The Licensing Rather Than Sale Of Software

The widespread practice of "licensing" software to users rather than selling it is troubling in the context of mass-market software products.[200] By definition, such "licensing" occurs because software

---

199. If "writings" can include three-dimensional sculptures, sound recordings, mask works, object code that is unreadable to humans, and buildings, it is difficult to see why it cannot include innovative computer maintenance services. Computer maintenance services are a particularly apropos example because of the creativity needed to diagnose and repair a complex system such as a computer.

200. While there is nothing inherently unique about software that allows it to be licensed to the exclusion of other off-the-shelf consumer products, this licensing phenomenon has not yet widely occurred with other products. Thus, while it is possible

copyright owners believe that it grants them greater rights than they receive under copyright law. If this were not true, companies would not invest the time and energy to develop and propagate the licenses. However, copyright misuse restrictions on this practice would collide head-on with one of the basic premises of our capitalistic economy—that contract law provides the ability to achieve the most efficient ordering of assets and rights through private agreements in the free marketplace.[201]

Although it is unclear whether mass-market[202] form licenses are generally enforceable,[203] software companies are undoubtedly attempting

---

that you could agree when you "license" a book in a bookstore that you could read it only once, not allow anyone else to read it and not transfer it to anyone else, there is a strong argument that such a "license" would not be allowed under copyright's "first-sale" doctrine (particularly if the terms of the "license" were not available to you when you made your purchase).

201. Of course, a distinction can be drawn between mass-market form licenses and those negotiated software contracts between two informed parties with relatively equal bargaining power, but these types of negotiated contracts are becoming increasingly rare in the software arena. Moreover, it is possible that some rights are too important to allow a contractual waiver. For example, a standard clause in a form contract that acknowledged and accepted the unconscionability of the contract is unlikely to be enforced.

202. Mass-market form licenses have been referred to in the past as "shrink-wrap" licenses, but this name is quickly losing its meaning in an on-line world where software is distributed electronically (some commentators have used the term "click-on" licenses more recently). The electronic means of contracting does provide some advantages, however. Not only can assent to contract terms be demonstrated more readily (even if they are unconscionable), the lower transaction costs in both locating customers and in providing a variety of products should presumably allow software manufacturers to price-discriminate on the basis of the rights received. While this author is not aware of any such products, it would be possible to pay one price for software that includes the right to resell the software to another after the purchaser is done with it, another price for software that includes the right to reverse engineer it, and still another price for software that allows use only by you on a single computer. To the extent that a licensee has some modicum of choice in the rights received, the inherent unconscionability of mass-market licenses in a world where as a practical matter there are often not substitutes available for the product is lessened.

203. Until the recent decision in the *ProCD* case by the Seventh Circuit, *supra* notes 125 and 168, most legal commentators believed that shrink-wrap licenses were unenforceable, particularly if the license could not be read before the sale was completed. *See supra* notes 169 and 172. Now, at least in the Seventh Circuit, "[c]ompetition among vendors" is the only protection for "consumers ... in a market economy." One of the underlying principles of the proposed Article 2B to the Uniform Commercial Code is that mass-market licenses should be broadly enforceable. U.C.C. § 2B-308 (Proposed Draft 1996) ("(a) Except as otherwise provided ..., a party adopts the terms of a mass market license if, before or

to alter the careful balance of rights granted between copyright owners and the public. In particular, by characterizing purchasers of off-the-shelf (or off-the-net) software as licensees rather than owners, the copyright owners are explicitly attempting to alter the rights granted to "owners" of software products under the "first-sale" provisions of section 109[204] and under the computer-program specific rights granted to users in section 117.[205] Allowing this practice would thwart the careful balance of rights struck by Congress.

B. License Restrictions On Reverse Engineering

While the licensing of software can generally lead to abuses of power and the alteration of the basic rights granted under the copyright law, mass-market contractual attempts to prevent reverse engineering are even more bothersome. As discussed earlier, courts have generally allowed reverse engineering in the absence of contractual restrictions.[206] Therefore, the question becomes whether this balance of rights can be altered contractually.

Although the courts have not yet decided this issue, they have discussed the related question of the patent-like protection that copyright law can provide to ideas and trade secrets when software is distributed in object code format.[207] This question involves competing principles. On one hand, a software copyright owner should not be able to receive the benefits of copyright law without having to disclose the underlying ideas and expressions in the source code. On the other hand, trade secrets have commercial value and businesses have a legitimate desire to protect them.

In *Kewanee*, the Supreme Court held that trade secret law was not preempted by patent law[208] and emphasized the important role that reverse engineering played for trade secrets embodied in products.[209] The court in *Bonito Boats* recently returned to this theme:

---

within a reasonable time after beginning to use the information pursuant to an agreement, the party signs or otherwise by its behavior manifests assent to a mass market license.").

204. *See supra* note 31.

205. *See supra* note 30.

206. *See supra* Part V.B.

207. *See supra* Part IV.B.3.

208. *See supra* Part II.C.

209. While trade secret law does not forbid the discovery of the trade secret by fair and honest means, e.g., independent creation or reverse engineering, patent law operates 'against the world,' forbidding any use of the invention for whatever purpose for a significant length of time .... Where patent law acts as a barrier, trade secret law functions relatively

> [I]n essence, the Florida law prohibits the entire public from engaging in a form of reverse engineering of a product in the public domain. This is clearly one of the rights vested in the federal patent holder, but has never been a part of state protection under the law of unfair competition or trade secrets .... The protections of state trade secret law are most effective at the developmental stage, before a product has been marketed and the threat of reverse engineering becomes real. During this period, patentability will often be an uncertain prospect, and to a certain extent, the protection offered by trade secret law may 'dovetail' with the incentives created by the federal patent monopoly.[210]

Despite the doubts of the Court in *Kewanee* that owners of patentable inventions will rely on trade secret law instead of patent law, that is exactly what happened in the early days of software patents when it was unclear if a patentee had to disclose actual source code in a patent application in order to enable the invention and describe its best mode. In this uncertain situation, many software owners recognized that a combination of trade secret law and copyright law provided stronger protection than patent law. The combination of trade secret law and copyright law allowed the software owners to use their software without fear of reverse engineering. Presumably, the widespread use of copyright law to enhance the power of trade secret law could serve to eviscerate the use of software patents, and if so, should be preempted by the patent laws.

Like the protection provided by copyright law, contracts may be used to protect ideas and trade secrets embodied in software. Restrictions in mass-market licenses can routinely prevent reverse engineering by the general public, and thus can greatly enhance the power of trade secret law. If so, the balance between trade secrets and patents will have shifted sufficiently that these clauses will be an "obstacle to the accomplishment and execution of the full purposes and objectives of Congress"[211] and should be struck down as misuse.[212]

> as a sieve. The possibility that an inventor who believes his invention meets the standards of patentability will sit back, rely on trade secret law, and after one year of use forfeit any right to patent protection, § 35 U.S.C. § 102(b), is remote indeed. ... We conclude that the extension of trade secret protection to clearly patentable inventions does not conflict with the patent policy of disclosure.

Kewanee, 416 U.S. at 490-92.

210. *Bonito Boats*, 489 U.S. at 160-61.

211. *Kewanee*, *supra* note 50, at 479.

212. *See also* Robert P. Merges, *Intellectual Property and the Costs of Commercial Exchange: A Review Essay*, 93 MICH. L. REV. 1570 (1995).

As a final consideration, in an era of increasing international harmonization of intellectual property law, it is important to note that the European Union has rejected the use of contract law to abrogate various legitimate rights of software users.[213] Software copyright owners cannot contractually restrict the right to make back-up copies,[214] the right to attempt to determine the ideas underlying the program,[215] or the right to reverse engineer software in order to create competing or compatible software.[216] It is instructive that the Europeans, who normally grant greater rights to authors than does the United States, do not allow contractual restrictions on reverse engineering when it is necessary for interoperability.

VIII. CONCLUSION

The United States' federal patent and copyright laws are based on a common utilitarian public policy of benefiting society through the encouragement of the creation, discovery and dissemination of novel ideas and creative expression. The grant of property rights in a limited monopoly to creators achieves these goals, and the underlying public policy necessitates the careful balancing of the rights granted to the creator with those retained by the general public. Enforcement of the intellectual property misuse doctrine prevents reordering of these rights, whether through state law or by individual owners of intellectual property.

While the misuse doctrine can be used in a general manner as described above, it can also be used to redefine the rights granted to creators when a class of protected works provides excessive power to the owners of such works. The misuse doctrine is the appropriate vehicle to

---

213. Council Directive 91/250/EEC, 1991 O.J. (L122) 1-2.

214. "The making of a back-up copy by a person having a right to use the computer program may not be prevented by contract insofar as it is necessary for that use." *Id.* at art. 5, cl. 2.

215. *Id.* at art. 5, cl. 3.

216. Reverse engineering is allowed if it is indispensable to achieve interoperability with some program (even if not the program reverse engineered), provided that a) the acts are performed by someone having a legitimate copy of the software, b) the information necessary to achieve interoperability has not previously been readily available, and c) the reverse engineering is confined to the parts of the program which are necessary to achieve interoperability.

Id. at art. 6, cl. 1 and art. 9, cl. 1.

be used in such a situation because it focuses on upholding the public policies underlying the intellectual property laws. However, a judicial redefinition of rights is justified only if it can avoid creating undue uncertainty on the part of intellectual property right owners as to the rights that they possess.

Computer software is a class of works that is unique among copyrighted works. Because copyright law provides excessive power to software copyright owners, the copyright misuse doctrine is needed to retain the balance of rights necessary to accomplishing the public policy underlying the intellectual property laws. The excessive power received by software copyright owners arises as a combination of the functional aspects of computer software, the ability to protect underlying ideas and expression when distributing software in object code format, the ability to protect trade secrets in a matter akin to patent law when distributing software in object code format, and unique aspects of the software industry which provide network externalities to some producers of software. For related reasons, a software copyright misuse doctrine separate from antitrust law is needed.

Due to the recent emergence of the copyright misuse doctrine, courts have applied other doctrines to address situations in which software copyright owners have attempted to unduly extend their deserved monopolies. Whether through fair use, intellectual property preemption, non-copyrightability or findings of misuse, the abuses by software copyright owners should be equitably limited. While the courts can generally arrive at the fair result through other doctrines, application of these doctrines creates suspect legal reasoning that does not always extend well to novel problems that the courts must face. Thus, a uniform software copyright misuse doctrine will provide a more appropriate vehicle for defining the scope of rights accorded to software copyright owners.

In developing a law of software copyright misuse, courts should be especially wary of explicit attempts to alter the statutorily-defined balance of rights granted, and should disallow attempts to restrain competition. In particular, the use of non-competition agreements and copyright restrictions on incidental RAM copies to prevent the development of unrelated goods or services should be unenforceable due to their inherent misuse. Similarly, the use of contracts to significantly limit the rights retained by purchasers of mass-market software warrants judicial limitations. While software is an incredibly useful product that deserves some protection, the rights granted must be limited so that the public retains its fair allocation.

ARTICLE

NONOBVIOUSNESS AND THE BIOTECHNOLOGY INDUSTRY: A PROPOSAL FOR A DOCTRINE OF ECONOMIC NONOBVIOUSNESS

KAREN I. BOYD †

TABLE OF CONTENTS

I. INTRODUCTION 312
II. WHY SHOULD BIOTECHNOLOGY RECEIVE SPECIAL TREATMENT? 313
A. Cultural Aspects of the Biotechnology Industry 313
B. Economic Aspects of the Biotechnology Industry 316
C. Biotechnology: An Industry That Should Be Developed 318
III. PATENT POLICY SHOULD BE CHANGED TO ENCOURAGE THE BIOTECHNOLOGY INDUSTRY 319
A. Possible Sources for a Biotechnology Policy: Current Regulation of the Industry 319
B. An Argument for the Use of Patent Law 321
IV. THE NONOBVIOUSNESS REQUIREMENT SHOULD BE USED TO CHANGE PATENT POLICY 323
A. Basic Patent Law 323
B. The Administrative Scheme 324
C. The Nonobviousness Requirement Should Be Used to Effect a Biotechnology Industrial Policy 326
V. THE NONOBVIOUSNESS REQUIREMENT 328
A. The General Requirement of Nonobviousness 328
B. Overview of Molecular Biology and Biotechnology 329

---

† J.D. 1996, Boalt Hall School of Law, UC Berkeley. The author also holds a Masters degree in molecular, cellular and developmental biology from UCLA, and a Bachelors degree in environmental biology from California State University Northridge. The author served as a law clerk to the Honorable Paul R. Michel, U.S. Court of Appeals for the Federal Circuit, from 1996-97 and is currently an associate at Fish & Richardson, P.C.

C. The Current Nonobviousness Doctrine Applied to Biotechnology Inventions 334
VI. A PROPOSED MODIFICATION OF THE NONOBVIOUSNESS REQUIREMENT WHICH WILL GIVE AN INCENTIVE TO THE BIOTECHNOLOGY INDUSTRY 337
A. The Proposal 337
B. The Proposal Responds to the Policy Needs of the Biotechnology Industry, and Its Implementation Is Feasible 339
VI. SUMMARY 343

I. INTRODUCTION

In the 23 years since the groundbreaking scientific discovery that made biotechnology possible, the biotechnology industry has grown to a 4 billion dollar per year industry that employs almost 100,000 people in 700 firms.[1] Conservative predictions estimate the industry will have 30 billion dollars in sales by the year 2000. American biotechnology is clearly a global leader in both biomedical and agricultural inventions, and the United States enjoys a significant trade surplus in biotechnological products.

In the words of Dr. Bernadine Healy, director of the National Institutes of Health, "[t]o maintain our world leadership in biotechnology, the United States must see the support of progress in biomedical research as one of its highest Federal priorities."[2] This support, however, need not come from monetary subsidies or regulatory relief (which seem unlikely in the current political climate). Rather, the U.S. Patent and Trademark Office, in conjunction with the Court of Appeals for the Federal Circuit, stands in a unique position to give needed support to the industry through a change in patent policy. Virtually cost-free to the public fisc, making patents slightly easier to get will satisfy the policy needs of the biotechnology industry and will be logically defensible.

Part II argues for the need to encourage the biotechnology industry. It considers arguments that focus on specific cultural and economic needs of the biotechnology industry, as well as why our society should give this particular industry special treatment. Part III suggests that patent law is

---

1. LYNN G. ZUCKER, ET AL., INTELLECTUAL CAPITAL AND THE BIRTH OF THE U.S. BIOTECHNOLOGY ENTERPRISES 1 (Nat'l Bureau of Econ. Res. Working Paper No. 4653, 1994).

2. Biotechnology Development and Patent Law: Hearing Before the Subcomm. on Intellectual Property and Judicial Admin. of the Comm. on the Judiciary, 102d Cong. (1992) [hereinafter Hearing] (prepared statement of Bernadine Healy, director, NIH, Dept. of Health and Human Services).

the best vehicle to implement a biotechnology industrial policy. Part IV provides an overview of the patent system, discussing both the law and the administrative scheme. Part IV concludes that modifying the nonobviousness requirement, in particular, would benefit the biotechnology industry. Part V discusses the current nonobviousness doctrine. After a discussion of the general doctrine, it gives a brief and non-technical explanation of molecular biology and biotechnology, thereby laying the groundwork for the ensuing discussion of the nonobviousness requirement as it applies to biotechnological inventions.

Part VI proposes a modification of the nonobviousness requirement for biotechnological inventions. It suggests the nonobviousness inquiry should take into consideration not only *technical* nonobviousness, but also *economic* nonobviousness. Finally, the nonobviousness standard proposed herein explains certain outcomes in recent biotechnology case law better than the doctrine on which those cases purport to stand.

II. WHY SHOULD BIOTECHNOLOGY RECEIVE SPECIAL TREATMENT?

In order to support the conclusion reached in this article—that biotechnological inventions should be given patents more easily in order to foster the industry's growth—two underlying assumptions must be stated. First, the industry *needs* special treatment to become fully developed; that is, something is wrong with allowing the market to determine how much investment biotechnology receives. Second, society *should* encourage biotechnology for equitable and economic reasons. The first assumption is supported by an analysis of the culture and economy of the industry, the second by an analysis of the potential benefits industry brings to society.

A. Cultural Aspects of the Biotechnology Industry

The biotechnology industry is characterized by several interrelated properties. Perhaps the most obvious of these is that the industry is primarily made up of small, single-product start-up companies. This paper argues that this important characteristic is related to, and mandated by, two other properties of the industry: its highly educated workforce and the close relationship between basic and applied science in the field.

This interrelationship is illustrated by the evolution of the first biotechnology company: Genentech. The basic scientific discovery which makes the industry possible is a method of making functional chimeric DNA (recombinant DNA) discovered by Stanley Cohen at Stanford University and Herbert Boyer at the University of California, San Francisco in 1973, for which a patent was filed in 1974 and granted in

1980.[3] Genentech was founded by Robert Swanson, a venture capitalist, and Hebert Boyer and successfully produced genetically engineered insulin by 1977.[4] What is important to note is that the potential commercial value of the academic research was immediately apparent and continues to be so. For example, BRCA1, the gene responsible for most breast cancers, was discovered through collaboration between scientists at Vanderbilt University School of Medicine and the University of Washington;[5] the gene responsible for early onset Alzheimer's Disease was discovered at the University of Toronto;[6] the gene for neurofibromatosis (Elephant Man Disease) was discovered at the University of Michigan;[7] the gene for Duchenne Muscular Dystrophy was discovered at Baylor College of Medicine;[8] the gene for Huntington's Disease was discovered at Johns Hopkins University.

Because of the close association between academic laboratories and industrial laboratories, biotechnology companies developed a culture that borrows several features of the university setting. The academic culture in which scientists train produces men and women who are accustomed to working in highly focused, small, independent laboratories. Although the university itself is a large institution, the laboratories are highly autonomous, self-funded functional units. Most laboratory research funding is from outside the university (through government grants and private endowments).

As a result of the influence of academic research on the biotechnology industry, the "research ethos" is encouraged, with publication and sharing of results encouraged.[9] Additionally, companies are relatively small and focused, not unlike some academic research laboratories.

An interdisciplinary team at the University of California, Los Angeles has done empirical research that bears out at least part of this

---

3. *See* Stanley N. Cohen et al., *Process for Producing Biologically Functional Molecular Chimeras*, U.S. Patent No. 4,237,224 (1980).

4. David V. Goeddel et al., *Direct Expression in Escherichia Coli of a DNA Sequence Coding for Human Growth Hormone*, 281 NATURE 544, October 18, 1979.

5. *See* Natalie Angier, *Surprising Role Found for Breast Cancer Gene*, N.Y. TIMES, March 5, 1996, at C1.

6. *See Scientists Discover Gene That Causes Alzheimer's in the Young*, N.Y. TIMES, June 29, 1995, at A17.

7. *See* Natalie Angier, *Scientists Discover the Gene in a Nervous System Disease*, N.Y. TIMES, July 14, 1990, at A1.

8. *See* Harold M. Schmeck Jr., *Battling the Legacy of Illness*, N.Y. TIMES GOOD HEALTH MAGAZINE, Apr. 29, 1990, at 36.

9. *See id.*

intuitive description. Lynne Zucker, Marilynn Brewer, and Michael Darby asked whether the relationship among scientists, universities, venture capital, and other economic factors influenced the founding of biotechnology companies.[10] They examined what they termed "intellectual capital" and found that this intellectual capital is critical to the birth of a biotechnology company. A person has intellectual capital if "she embodies a specialized body of knowledge which enables the individual to earn supranormal returns on the cost of obtaining that knowledge."[11] The authors say that biotechnology is intellectual capital intensive because radically new laboratory techniques need to be learned first hand, thus making the initial discoverers of the techniques and their apprentices extremely valuable.[12] Because intellectual capital dissipates as a technique becomes more widely understood and applied, it is "by its nature a transient property of disequilibria."[13] Despite this transience, intellectual capital is critical to the birth of biotechnology companies because these companies come into existence before new technological knowledge is diffused.[14] Using the metric of publications of gene sequences, the authors found that intellectual capital is geographically concentrated.[15] Understandably, these centers of concentration are frequently universities.[16] The researchers found that the concentration of intellectual capital, in time and place, is an accurate predictor of the birth of biotechnology companies.[17]

By following the academic model, biotechnology loses some of the economic advantages of large industries, such as economies of scale. However, it retains a feature critical to successful innovation in high-tech industries: an atmosphere of creativity and intellectual freedom. The need for this atmosphere argues for an economic policy for biotechnology that will compensate for the economic losses that accompany this academic atmosphere.

Because biotechnology requires a very highly skilled work force, there is a need for the industry to continue to lure academic scientists to industry. To do this, the industry must create an environment similar to that in a university, and as a result, must be provided with viable economic incentives to maintain a high level of innovation.

---

10. Zucker, et al., *supra* note 1, abstract.
11. *Id.* at 5.
12. *Id.*
13. *Id.* at 6.
14. *See id.*
15. *Id.* at 7.
16. *See id.* at 11-12.
17. *Id.* at 11.

B. Economic Aspects of the Biotechnology Industry

In addition to having cultural attributes (and the concomitant economic problems) that need to be protected in order to remain innovative, the biotechnology industry has purely economic considerations that suggest the need for special policies. For instance, venture capital is an important source of funding for biotechnology start-ups. One commentator has gone so far as to say that "[b]iotechnology has emerged as an industry largely because of one economic institution: venture capital."[18] Others have been more cautious, merely noting that "the existence of the venture capital industry in America has had a significant effect on the development of the biotech industry." [19]

Regardless of the exact quanta of importance that venture capital has had in the development of the biotechnology industry, as realities such as technical difficulties and limited markets have become apparent, venture capitalists have become much less willing to invest. The increased competition in the venture capital market has necessitated the development of a metric for differentiating between biotechnology firms. Because venture capitalists are typically not experts with respect to the technology, intellectual property rights have become the proxy for technological value. It is therefore important that the patent system award patents to deserving inventions so that the companies developing those inventions can attract venture capital, thereby supporting the biotechnology industry.

Related to this argument is the suggestion made by Professor Robert Merges that very expensive and very risky inventions tend to be pursued at less than socially desirable levels.[20] Biotechnology inventions are frequently both risky and expensive. Professor Merges takes a statistically rigorous approach, through economic modeling, to show how this risk will result in less than optimal investment. Nonetheless, a less mathematical approach is also instructive.[21] Any given invention will have a certain likelihood of success and a certain expected market payoff. The market payoff must be discounted by the probability of success in order to decide whether it is economically wise to pursue an invention. However, these are not the only factors in the equation. Inventors and, perhaps more importantly, investors, are more or less risk averse. For instance, given two investments, one with a 100% chance of a $1 million payoff, and the other with a 50% chance of a $50,000 loss and

---

18. MARTIN KENNEY, BIOTECHNOLOGY: THE UNIVERSITY-INDUSTRIAL COMPLEX 133 (1986).

19. Zucker, et al., *supra* note 1, at 15.

20. Robert P. Merges, *Uncertainty and the Standard of Patentability*, 7 HIGH TECH. L.J. 1, 47 (1992).

21. This intuitive approach is based on Prof. Merges' discussion, *id.* at 43-55.

a 50% chance of a $2,100,000 payoff, it is not surprising that some investors will prefer the former.[22] They will invest in the former despite the fact that it has a lower absolute payoff ($1 million vs. (0.5)(-$50,000) + (0.5)($2,100,000), or $1,025,000). The preference will stem instead from the first investment's lower risk. This risk aversion can be overcome only by offering a greater return. In other words, investors will more heavily discount a high risk investment than a low risk one.

This very basic risk analysis is complicated when there are a number of different possible outcomes, each with a certain probability of occurring. For instance, two investments, one with a 50% chance of winning $20 and a 50% chance of losing $10, and the other with a 1% chance of winning $1000, a 1% chance of winning $100, a 1% chance of losing $500 and a 96% chance of breaking even, both have a predicted return of $5.[23] Intuitively, one can understand the greater appeal of the first investment, again because of its lower risk. This intuition can be articulated with a statistical measure called variance.[24] In this context, variance, as its name implies, describes how widely varying a set of possible outcomes is. The higher the variance, the higher the risk.[25] Biotechnology inventions, because they have many hurdles to clear before commercial success, can be described as high variance. There must be positive *in vitro* results, animal tests, and three or four stages of clinical trials before ever even entering the market. Also, biotechnology products must compete with traditional, often less expensive, alternatives. Finally, there is always the possibility of a high payoff when a biotechnology product is found to be useful in a multitude of contexts.[26]

Furthermore, risk aversion also results in heavier discounting of extremely expensive inventions. To understand this, one need only imagine oneself going into a gambling casino. Although the odds of winning a hand of blackjack are exactly the same whether you bet $5 or $500, the more risk averse gambler will intuitively discount the chance of winning the $500 hand and be less likely to wager that amount of money.[27]

---

22. EDWIN MANSFIELD ED., MICROECONOMICS 565 (5th ed. 1985).

23. *See* Merges, *supra* note 20, at 143.

24. Variance describes how closely a collection of values is clustered around the mean value. It is defined as the average squared deviation of the observations from the mean. *See* DAVID T. SUZUKI ET AL., AN INTRODUCTION TO GENETIC ANALYSIS 645 (4th ed. 1989).

25. Merges, *supra* note 20, at 43.

26. For example, Epo, a product of Amgen, was initially intended to be used for patients undergoing dialysis. It has since been shown to be useful in treating cancer.

27. The gambling analogy breaks down somewhat since it does not take into account the adrenaline rush that people who bet $500 a hand are seeking.

The implication of this analysis is clear. Because biotechnology inventions are high cost and risky, they will be pursued at less than optimal levels because their perceived payoff is discounted disproportionately to their actual probability of success. This should be compensated for by giving an added incentive (that is, increasing the total payoff) to pursue these expensive, risky inventions.

C. Biotechnology: An Industry That Should Be Developed

The economic needs of the biotechnology industry due to expense, risk, and cultural needs, however, do not answer the more fundamental question of whether biotechnology is an industry that *should* be encouraged. This argument is two pronged: technological and economic.

The field of biotechnology offers society extremely important technological advances. Medically, biotechnology promises treatment and prevention of previously incurable and inevitable disease. Biotechnology also makes possible early detection and genetic screening. Moreover, even in areas where traditional medicine has made great advances, biotechnology shows great promise. For instance, while the advent of antibiotics made a tremendous difference in society's health, now antibiotic resistant bacteria are threatening this medical mainstay. Biotechnology is attacking this problem in ways impossible with traditional techniques.[28] Agriculturally, biotechnology has the potential of increasing food production, limiting the need for fertilizers and pesticides that harm the environment, and increasing the quality of products that reach the market.

Economically, biotechnology also offers important advantages to the United States. In the words of one commentator: "the spectacular innovations in recombinant DNA technology introduced in the early 1970s, and subsequently refined beyond all expectations, have transformed molecular biology into one of the 'high-tech' fields that supposedly presage the future economic and professional base of Western society."[29] Biotechnology offers high wage jobs and, currently, American biotechnology enjoys a distinct advantage in the worldwide market because virtually all the major discoveries have been made here, and the American post-graduate educational system is regarded as producing some of the finest molecular biologists. In order to retain this economic advantage, the United States should give extra incentives to the industry in order to keep up with the policies in other parts of the world that support biotechnology.

---

28. *See* Lawrence M. Fisher, *Biotech Counterattack on Resistant Bacteria; Small Companies Leading in Research*, N.Y. TIMES, Apr. 26, 1996, at C1.

29. NATALIE ANGIER, NATURAL OBSESSIONS 28-29 (1988).

In conclusion, our system should be revised to give biotechnology companies an extra push. The necessary culture of the industry, combined with risk averse, investment strategies have led to less than optimal innovation. This innovation is socially desirable for both equitable and economic reasons.

III. PATENT POLICY SHOULD BE CHANGED TO ENCOURAGE THE BIOTECHNOLOGY INDUSTRY

The standard industrial policy mechanisms are direct subsidization and regulatory relief. Direct subsidization of the biotechnology industry, while perhaps appealing as a transparent, controllable method for supporting the industry, is not the best choice. First, it is indiscriminate. Direct subsidies are to industries, rather than to individual inventions that would otherwise not be pursued at the desired levels. Second, in today's fiscal climate, direct subsidization is politically unlikely. Third, direct subsidization of other industries has not worked well in the past. In fact, there is no reliable model in American politics for how to make subsidization work.

Regulatory relief is equally problematic. Part III of this article will review the agencies that currently regulate the biotechnology industry in the public health and safety arena. It will then discuss how a modification of the patent system, rather than regulatory relief in the classic sense of the term, would best meet the needs of the industry.

A. Possible Sources for a Biotechnology Policy: Current Regulation of the Industry

The biotechnology industry is currently regulated by at least eight administrative agencies with overlapping and sometimes conflicting authority.[30] The Biotechnology Research Subcommittee (BRS) of the interagency Committee of Health and Life Sciences, established in 1990,[31] coordinates the regulation of biotechnology by the Food and Drug Administration (FDA), National Institutes of Health (NIH), United States Department of Agriculture (USDA), Environmental Protection Agency (EPA), National Science Foundation (NSF), and Occupational

---

30. *See* Sandra H. Cuttler, *Commentaries: The Food and Drug Administration's Regulation of Genetically Engineered Human Drugs*, 1 J. PHARM. & L. 191, 204 (1992).

31. *See* Karen Goldman Herman, *Issues in the Regulation of Bioengineered Food*, 7 HIGH TECH. L.J. 107, 120 (1992). The BRS replaced the Biotechnology Science Coordinating Committee (BSCC), established in 1986, and has similar responsibilities to the BSCC. *Id.* This paper refers to both committees collectively as the BRS.

Safety and Health Administration (OSHA).[32] Any regulatory relief for the biotechnology industry would have to be coordinated and approved by the BRS. The U.S. Patent and Trademark Office (PTO) also regulates the biotechnology industry, but is not coordinated by the BRS.

The BRS is responsible for effecting the "Coordinated Framework for the Regulation of Biotechnology" (coordinated framework), which was a statement of federal biotechnology regulation policy promulgated by the Office of Science and Technology Policy in 1986.[33] Under the coordinated framework, existing agencies regulate biotechnological products.[34] Products are regulated according to type, and if a product has aspects that bring it under the regulatory umbrella of more than one agency, the BRS designates a lead agency.[35] For instance, a tomato plant genetically engineered to create its own pesticide could be regulated as a pesticide by the EPA, a plant by the USDA, or a food product by the FDA.[36]

The FDA's regulatory purpose is to ensure the safety and efficacy of drugs and the safety of foods.[37] As an agency, the FDA is a potential candidate for the implementation of regulatory relief because certain types of drugs receive special treatment under the Food, Drug and Cosmetic Act. For instance, under intense pressure by the AIDS lobby, the FDA has devised special, expedited examination procedures and special experimental use provisions for drugs for the treatment of AIDS. These procedures apply to a lesser extent to drugs for the treatment of other immediately life-threatening diseases. Also, so called "orphan drugs" that treat diseases affecting fewer than 20,000 people have special protections.

These types of regulatory relief, however, are unlikely to be extended to biotechnological drugs as a general class. The policies behind the expedited procedures for AIDS drugs do not apply to drugs that treat diseases from which people are not dying in the pre-approval interim. Furthermore, diseases that affect larger numbers of people do not need extra incentives to induce the development of treatments. Finally, given the FDA's administrative purpose of ensuring that drugs are safe and

---

32. *See* Cuttler, *supra* note 30, at 204. OSHA's regulatory power over biotechnology is only incident to its regulatory power over all industry, so it will not be discussed in this paper.

33. *See* Herman, *supra* note 31, at 119.

34. *See id.*

35. *See id.*

36. *See id.* at 121. This is a hypothetical example, so the BRS has not designated a lead agency.

37. *See* Food, Drug, and Cosmetic Act, 21 U.S.C. §§ 301-392 (1972).

effective, it is unlikely that biotechnological products will be given any special regulatory relief.[38]

Likewise, the other agencies regulating the biotechnology industry are not in a viable position to grant the industry regulatory relief. OSHA's goals are too tangential to the purposes of the industry; the USDA and EPA have regulatory purposes too important to compromise through regulatory relief (safety of the food supply and protection of the environment, respectively); the NIH and NSF are primarily research and research funding agencies that already support most of the academic molecular biology research underpinning the entire industry. For these, and other reasons, an extended discussion of which are beyond the scope of this paper, standard regulatory relief for the industry is not advisable. Instead, this paper advocates the use of patent policy to give an additional incentive to biotechnology research and investment.

B. An Argument for the Use of Patent Law

Patent law is unique among the administrative schemes regulating biotechnology in that it has a constitutional basis. The Constitution states that Congress shall have the power, "[t]o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries." [39] It is also unique because, as the constitutional language reflects, the regulatory system is designed to promote the progress of science, rather than protect public health and safety. This underlying purpose makes patent law an ideal forum for the institution of a policy to give incentives to the biotechnology industry. Rather than subsidizing the industry or giving regulatory relief, a "fine tuning" of the patent scheme for the industry is squarely within political grasp.

Patents are granted for two reasons: to give an incentive to innovators to invent, and to get innovators to disclose their technological findings.[40] The disclosure goal has been characterized as a contract

38. One exception to this may be user fees. The FDA has recently restructured their fee system so that the agency can depend on user fees to fund itself to a greater extent. This has resulted in much higher user fees. Biotechnology companies, generally small and with limited funding, may be able to lobby for reduced user fees for companies below a particular size. This type of relief would not, however, be limited to biotechnology companies.

39. U.S. CONST. art. I, § 8, cl. 8.

40. *See* ARTHUR R. MILLER & MICHAEL H. DAVIS, INTELLECTUAL PROPERTY IN A NUTSHELL 15 (1990).

between the inventor and society.[41] Society grants a 20-year monopoly[42] to the inventor in exchange for the inventor's disclosure of how to practice the invention.[43] This information is included in the patent application, which becomes publicly available as soon as the patent is granted. Thus, when the patent expires, others can use the information in the patent for society's benefit. Background information, the use of which does not infringe the patent, can be used to society's benefit even during the patent term. The incentive goal has as its premise that certain socially desirable inventions will not be pursued unless the inventor is given an extra economic incentive. The patent monopoly is this incentive. According to then Chairman of the House of Representatives Subcommittee on Intellectual Property, William J. Hughes, "[p]atent law is a powerful economic incentive that can actually determine the amount of capital investment and research activity that occurs within a particular area."[44]

Patent law is a particularly appropriate vehicle for a biotechnology policy because patents are extremely important to the industry. Like other pharmaceutical industries, the real risk in biotechnology is that what appears to be a promising product will not actually work well enough to be approved by the FDA and become commercially successful. This result occurs because pharmaceuticals deal with complex biological systems that are not completely understood.[45] Also like other pharmaceuticals, biotechnological products, once perfected and shown to be commercially viable, are easy to copy.[46] Thus, there is a real danger that without patent protection an inventor's return on her investment in a risky endeavor will be siphoned off by copyists. Because of this, the pharmaceutical industry, including biotechnology, is one of the few industries that literally could not survive without patent protection.[47]

---

41. *See* ROBERT P. MERGES, PATENT LAW AND POLICY 513 (1992).

42. Actually, patent rights are not quite a monopoly. A patent gives the patent-holder the right to prevent others from making, using, selling or importing the patented invention. It does not give the patent holder an affirmative right to practice the invention. Thus, although a drug may be patented, it must still be licensed by the FDA before it can be marketed. A patent also does not obligate the patentee to make, use, or sell the invention, nor does it obligate the patentee to license the invention. *See* MILLER & DAVIS, *supra* note 40, at 12-14.

43. *See* MERGES, supra note 41, at 513.

44. *Hearing, supra* note 2 (opening statement of William J. Hughes).

45. *See* C. TAYLOR & Z. SILBERSTON, THE ECONOMIC IMPACT OF THE PATENT SYSTEM: A STUDY OF THE BRITISH EXPERIENCE 252 (1973).

46. *See id.* at 244-45.

47. *See* Levin et al., *Appropriating the Returns from Industrial Research and Development,* 1987 BROOKINGS PAPERS ECON. ACTIVITY 783.

Patent policy changes are the proper strategy for encouraging the biotechnology industry for three main reasons. First, the goals of the patent system (to encourage innovation and disclosure) are not undermined by supplying additional incentive to the biotechnology industry. Second, the patent system is set up as a economic incentive, which means it can easily address the economic needs of the industry. Third, because patents are demonstrably critical to the industry as it now stands, a change in patent law will readily be recognized as good for the industry and thus elicit quick and positive responses.

IV. THE NONOBVIOUSNESS REQUIREMENT SHOULD BE USED TO CHANGE PATENT POLICY

The administrative scheme for granting patent rights requires that an invention be new, useful, and nonobvious patentable subject matter. Part IV introduces this scheme and argues that the nonobviousness requirement should be changed in order to increase biotechnology innovation.

A. Basic Patent Law

The basic requirement for patenting an invention that falls within patentable subject matter is that the invention be new, useful, and non-obvious. Patentable subject matter includes processes, machines, manufactures, compositions of matter, and improvements of any of these.[48] These categories have been judicially interpreted so that laws of nature, mental processes, intellectual concepts, ideas, natural phenomena, mathematical formulae, methods of calculation, fundamental truths, original causes, and motives are not patentable subject matter.[49] The novelty requirement provides that any invention known, used, published, or patented domestically, or patented or published abroad is unpatentable.[50] The utility criterion, set out in 35 U.S.C. § 101, requires only that an invention be "useful." This requirement has been interpreted to mean that an invention must actually *do* something that the inventor has shown, not that he or she merely suspects, and that the invention actually does what it purports to do.[51] The first utility requirement would render unpatentable, for instance, a machine with moving parts that merely spins around without serving any purpose (although to amuse or to entertain is considered a valid purpose for patentability inquiries).[52]

48. *See* 35 U.S.C. § 101 (1984).

49. *See In re* Bergy, 596 F.2d 952 (C.C.P.A. 1979).

50. *See* 35 U.S.C. § 102 (1984).

51. *See* MERGES, *supra* note 41, at 147.

52. *See id.*

The second utility requirement would render unpatentable, for instance, a perpetual motion machine (which, under commonly accepted laws of physics, cannot do what it purports to do).[53] The utility requirement is not typically a particularly onerous hurdle for a patentee to overcome.[54] The non-obviousness requirement states that "if the differences between the [invention] and the prior art[55] are such that the [invention] as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art" the invention is unpatentable.[56] The non-obviousness requirement is typically thought to be the greatest hurdle to patentability and has been called "the ultimate condition of patentability."[57]

B. The Administrative Scheme

The administrative scheme for patents follows a slightly different model from the typical regulatory agency. The Patent and Trademark Office grants patents. Agency employees called examiners examine patent applications and determine whether or not to grant patents based on the law briefly outlined above.[58] Through a procedure called prosecution, the examiner and patent applicant carry on a written and sometimes oral discourse, arguing about patentability and amending the patent application. Ultimately, the examiner either grants the patent or issues a final rejection.[59] However, there is little final, about a "final" rejection.

The applicant can appeal the examiner's decision to the Board of Patent Appeals and Interferences, which reviews the rejection.[60] If the Board affirms the rejection, the applicant can appeal the rejection to the Court of Appeals for the Federal Circuit.[61]

The Federal Circuit is an Article III court, which was established in 1982.[62] While its jurisdiction is not completely patent related, the main

---

53. *See id.* at 148.

54. *See id.* at 147.

55. "[Prior art] ... includes any relevant knowledge, acts, descriptions and patents which pertain to, but predate, [the] invention in question." BLACK'S LAW DICTIONARY 828 (Abridged 6th ed. 1991).

56. 35 U.S.C § 103 (1984).

57. *See* NONOBVIOUSNESS—THE ULTIMATE CONDITION OF PATENTABILITY (J. Witherspoon ed. 1980).

58. *See* MERGES, *supra* note 41, at 30-31.

59. *See id.* at 31.

60. *See id.* at 32.

61. *See id.*

62. *See id.* at 9.

reason for the creation of the court was to bring all patent questions to the same appellate court. Previously patent decisions were appealed to whatever circuit was geographically appropriate.[63] Because patent law is highly technical, both legally and scientifically, and because the Supreme Court rarely hears patent questions, prior to the formation of the Federal Circuit, there were many circuit splits in patent law and the doctrine was muddled.[64] Generally, the Federal Circuit is perceived to have accomplished its mission of bringing coherence and consistency to patent law.[65] The court is also perceived to be quite pro-patent, in that it generally upholds the patentability of the invention and the validity of patents.[66]

The Federal Circuit reviews the decision of the Board de novo for questions of law, while factual findings are reviewed for clear error.[67] If the Federal Circuit upholds the Board's decision to affirm the examiner's final rejection, the applicant does have recourse in the Supreme Court. Because the Supreme Court so rarely grants certiorari to patent issues, though, the Federal Circuit is generally regarded as the court of last resort for patent questions.

If the examiner grants the patent and the patent issues, there is still an opportunity for the validity of the patent to be challenged. The most common forum for this is in an infringement action.[68] Invalidity is an absolute defense to patent infringement.[69] While a patent carries a presumption of validity, this presumption is rebuttable under a clear and convincing evidence standard. Patent infringement actions can be brought in any Federal District Court (there is exclusive federal jurisdiction for patent law). Any appeals are to the Federal Circuit.[70]

---

63. *See id.*

64. *See id.*

65. *See id.*

66. *See id.*

67. *See In re* Vaeck, 947 F.2d 488, 493 (Fed. Cir. 1991); *In re* Woodruff, 919 F.2d 1575, 1577 (Fed. Cir. 1990).

68. Less common, but not rare, is a re-examination proceeding where a third party challenges the validity of an issued patent outside of an infringement action. This is an administrative proceeding that essentially seeks a declaratory judgment of invalidity. *See* MERGES, *supra* note 41, at 32.

69. *See id.*

70. *See id.* at 9.

C. The Nonobviousness Requirement Should Be Used to Effect a Biotechnology Industrial Policy

1. *NOVELTY AND PATENTABLE SUBJECT MATTER*

A quick analysis of the requirements for patentability shows that the nonobviousness requirement has the greatest opportunity for changes to effect a policy to benefit the biotechnology industry. Despite continuing controversy over the ethics of patenting living things,[71] it is clear that biotechnological inventions, including living subject matter, are patentable subject matter.[72] Save for the remote possibility of congressional action prohibiting patents on biotechnological inventions, this is not likely to change.[73] Likewise, because novelty is such a low threshold requirement, it too presents no real challenge to biotechnology inventions.

2. *UTILITY*

Utility is an area that has seen some interesting developments in biotechnology patents. The National Institutes of Health (NIH) attempted to patent a collection of gene sequences. The patent application was originally rejected for lack of utility. The subject matter of the application requires some explanation. Craig Venter, a scientist at NIH involved with the Human Genome Project, made a cDNA library of all the genes expressed in the human brain. In lay terms, he collected the information that codes for every cellular function in the brain. He then sequenced parts of each of these cDNAs. The utility rejection was based mostly on the fact that Venter did not know what any of the sequences did in the brain, and therefore the sequences lacked utility. Because of the nature of the experiment, though, he had parts of sequences for every gene used in brain function. Some, if not all, of the genes will eventually be found to be important for understanding the way the brain functions. Thus there is a definite argument that there is at least latent utility in the sequences.

---

71. Jeremy Rifkin of the Foundation on Economic Trends, along with others such as Dr. Michael W. Fox, have consistently opposed both the patenting of genetically engineered organisms, and the underlying genetic engineering itself. These oppositions are on ethical grounds, and focus on the profanation of all life when genes are manipulated. Rifkin's views have been almost universally dismissed by the scientific community. *See* DR. MICHAEL W. FOX, SUPERPIGS AND WONDERCORN: THE BRAVE NEW WORLD OF BIOTECHNOLOGY ... AND WHERE IT ALL MAY LEAD 22-26 (1992).

72. *See* Diamond v. Chakrabarty, 447 U.S. 303 (1980).

73. The scope of patentable subject matter has always been a one-way ratchet. Once a subject matter is declared patentable, courts have not revisited the issue.

This utility question has not been judicially resolved. NIH has since withdrawn the patent applications on policy grounds. This does not mean, however, that the issue is a dead letter. Dr. Venter has left NIH and is continuing his work, including attempts at patenting gene sequences of unknown function, at an independent lab. Other inventors are also attempting to attain similar patents. Because PTO proceedings are closed, the resolution of this very interesting question about utility will have to wait until the issue finishes winding its way through the administrative process.

Aside from the fascinating open question regarding the utility of gene sequences of unknown function, most biotechnology inventions routinely clear the utility hurdle. According to the recent Federal Circuit opinion in *In re Brana*,[74] and the Patent and Trademark Office's interpretive Utility Examination Guidelines,[75] utility for a biotechnology invention can be shown by proof of clinical utility in humans or animals, or by *in vitro* testing. *In vitro* testing is the first step in determining whether a given drug has any biological activity. The court in *Brana* focused on the fact that it is the FDA's responsibility to determine safety and efficacy in humans; usefulness for patent purposes can be shown with much less rigorous data.[76]

The policies behind the utility requirement, and the economic implications of where the requirement is set, are quite interesting. The basic premise of the requirement is that society should not suffer the "embarrassment" of a monopoly unless society gets something in return. Unless an invention has some measure of usefulness, society has received nothing for its grant of patent rights. Pulling against this force for increasing the quantum of utility required for patentability is the patent policy of giving an incentive for innovation. The earlier in the research process patent rights are granted, the more of that process is protected, and the more likely it is that people will be enticed to innovate. There is a lower limit, however, to this incentive-driven pull for lower utility requirements. At some point, competitors will begin to engage in inefficient "racing" behaviors, where the goal becomes patent rights rather than quality inventions. This lower limit has arguably been reached in the biotechnology context. There is no lesser way to demonstrate biological activity than to do *in vitro* testing. Utility is a low hurdle for biotechnological inventions to overcome, and probably should not be further reduced for biotechnology inventions.

---

74. 51 F.3d 1560 (Fed. Cir. 1995).

75. 60 Fed. Reg. 36263 (July 14, 1995).

76. 51 F.3d at 1568.

V. THE NONOBVIOUSNESS REQUIREMENT

A. The General Requirement of Nonobviousness

The basic nonobviousness inquiry was set out in *Graham v. John Deere Co.*[77] The Court held that § 103 obviousness is a factual inquiry under which "the scope and content of the prior art are to be determined; differences between the prior art and the claims at issue are to be ascertained; and the level of ordinary skill in the pertinent art resolved."[78] Nonobviousness is then to be determined given these considerations.[79] The test is focused on the unpredictability of the patentee's invention. The standard is whether a person of ordinary skill in the art, knowing all of the prior art, would have a reasonable expectation that the invention would work.[80] If that person would have expected the invention to work, the invention is obvious and non-patentable.[81]

An invention can be deemed obvious if an inventor merely pieced together aspects from separate prior art references—there need not be a single reference that makes the invention obvious. However, there must have been a "suggestion" in the prior art to combine the prior art elements in the way in which they are combined in the patentee's invention.[82] Furthermore, an invention will not be found obvious if it was merely obvious to try.[83] Again, the standard is a reasonable expectation of success. If an invention is both obvious to try and a person skilled in the art would reasonably expect it to succeed at some level, the invention can nonetheless be found nonobvious if the invention works much better than expected.[84] If the prior art suggests the combination of a number of variables, and parameters are given for each variable, a specific combination of variables can be found nonobvious.[85] For this type of invention to be nonobvious, there cannot be a reasonable expectation of success for all possible combinations nor can there be a suggestion in the prior art to combine the variables in the way that the invention combines

---

77. 383 U.S. 1 (1966).

78. *Id.* at 694.

79. *See id.*

80. *See* Loctite Corp. v. Ultraseal, 781 F.2d 861, 874 (Fed. Cir. 1985).

81. *See id.*

82. *See* Lindemann Maschinenfabrik GmbH v. American Hoist & Derrick Co., 730 F.2d 1452, 1462 (Fed. Cir. 1984).

83. *See, e.g., In re* Fine, 837 F.2d 1071, 1075 (Fed. Cir. 1988); Merck & Co. v. Biocraft Labs., 874 F.2d 804, 807 (Fed. Cir. 1989).

84. *See* Novo Industri A/S v. Travenol Labs., 677 F.2d 1202, 1208 (7th Cir. 1982).

85. *See In re* O'Farrell, 853 F.2d 894 (Fed. Cir. 1988).

them.[86] Instead, the test is again a reasonable expectation of success. As a whole, the "obvious to try" doctrine seems to be a particular application of the reasonable expectation of success standard.

This paper will next turn to the application of the nonobviousness requirement to biotechnological inventions. To give this discussion some context, a brief description of the technology, and the underlying molecular biology, is in order. This introduction is not meant to be comprehensive. Because of this simplicity, many of the exceptions to and nuances of general principles have been omitted from this discussion. Where it is scientifically and intellectually irresponsible to do this, the exceptions or nuances have been noted in footnotes.

B. Overview of Molecular Biology and Biotechnology

Biotechnology is used in three main ways. First, a biologically active molecule, such as an antibody or a peptide hormone, made naturally in small amounts in a given organism, can be produced in large quantities using bacteria. Second, an individual organism that is malfunctioning can be corrected by adding a correct copy of a malfunctioning gene (somatic gene therapy). Third, an entire line of organisms with a unique or corrected characteristic can be made by altering genes in reproductive cells (transgenic animals).

1. BASICS

There are three main classes of molecules important to biotechnology, each intimately related to the others: DNA,[87] RNA,[88] and proteins. DNA has been described as the "genetic blueprint"[89] and every cell in an organism contains a complete copy of all the DNA that defines the organism. This complete copy of DNA is called the genome.[90] Thus, each cell has a complete copy of the instructions necessary for the development, maintenance, and reproduction of the organism.

However, not all genes function in all cells. Those parts of the DNA that a nerve cell, for example, needs to function as a specialized cell, are "transcribed" into messenger RNA (mRNA) for subsequent translation by

---

86. *See* Merck & Co. v. Biocraft Labs., 874 F.2d 804 (Fed. Cir. 1989).

87. Deoxyribonucleic acid. *See* CHRISTOPHER WILLS, THE WISDOM OF THE GENES 328 (1989).

88. Ribonucleic acid. *See* LARRY GONICK AND MARK WHEELIS, THE CARTOON GUIDE TO GENETICS 132 (1983).

89. EDWARD YOXEN, THE GENE BUSINESS: WHO SHOULD CONTROL BIOTECHNOLOGY? 59 (1983).

90. *See* WILLS, *supra* note 87, at 83.

cellular machinery into a protein.[91] Thus, the mRNA present in a cell is a direct reflection of the genes that are functioning in that particular cell type. Because different genes function in different cell types, the spectrum of mRNA will vary from cell type to cell type. The proteins translated from mRNA are the workhorses of the cell. Although some protein, indeed the protein with which we are perhaps most familiar, is merely structural (for instance, the protein in hair or cartilage),[92] by far the most important role for proteins is enzymatic.[93] Enzymes are required for catalyzing every biochemical reaction that happens in every cell, in every organism.[94] For instance, the protein DNA polymerase helps the cell to copy its DNA so the cell can divide.[95]

2. *WHAT CAN BIOTECHNOLOGY DO?*

Lest this all seem hopelessly circular (DNA begot RNA; RNA begot protein; protein begot DNA ...), one must remember the myriad other tasks for cells that proteinaceous enzymes must catalyze: building the cellular machinery, respiring, metabolizing, breaking down of waste products, producing cellular secretions, and communicating with other cells.[96] Biotechnology is the technology that uses the understanding of molecular biology to manipulate living organisms in any of these, or countless other functions.

For instance, biotechnology can be used to treat diseases that are caused by malfunction in any one of these cellular tasks.[97] If DNA is damaged, the RNA does not carry the right information. The protein that is then made does not function properly. Malfunctioning proteins result in disease. Cancer is a breakdown in the regulation of cellular division.[98] Cells, instead of dividing at just the rate required to replace dead cells, begin to divide uncontrollably into cancer cells.[99] Lactose intolerance is due to a deficiency in lactase, which breaks down milk sugar. Some

---

91. *See id.* at 20.

92. *See* ANGIER, *supra* note 29, at 41

93. *See id.*

94. *See* GONICK AND WHEELIS, *supra* note 88, at 113.

95. *See* DAVID FREIFELDER, MOLECULAR BIOLOGY 224 (2d. ed. 1987).

96. *See* GONICK AND WHEELIS, *supra* note 88, at 113-14.

97. *See* ANGIER, *supra* note 29, at 71.

98. *See id.*

99. *See* SUZUKI ET AL., *supra* note 24, at 595. Cancer cells are "dedifferentiated" cells, which are generic. They contrast with differentiated cells, such as lung cells, liver cells, skin cells, muscle cells. Cells differentiate very early in embryonic development. Dedifferentiation is a hallmark of cancer. *Id.*

genetic diseases, like SCIDs or lactose intolerance, are wholly inherited;[100] others, like cancer, are often partly due to environmental factors damaging otherwise healthy DNA.[101]

3. *HOW DOES BIOTECHNOLOGY DO IT?*

Biotechnology attacks these problems in several ways. The most straightforward way is to use another organism, usually bacteria, to make lots of whatever proteins people need.[102] For instance, the first biotechnology company, Genentech, set out to treat diabetes.[103] It was known that diabetics lacked the enzyme insulin and were therefore unable to metabolize sugar properly.[104] The traditional treatment involved isolating insulin from bovine and porcine pancreases (available in large quantities from slaughterhouses), and injecting it into diabetics.[105] Although this worked, and still works, reasonably well for most diabetics, because bovine and porcine insulin are slightly different from human insulin, some diabetics have allergic reactions.[106] Clearly, the administration of human insulin would be superior.[107] Just as clearly, for ethical reasons, you couldn't isolate human insulin from human pancreases.

Genentech approached the problem via recombinant DNA technology. First, scientists isolated the gene (DNA) that codes for insulin. Then they took that gene and put it into a bacteria's genome. By using special inducers of transcription, they were able to induce the bacteria to transcribe the insulin gene and translate it into insulin. The insulin could then be isolated and administered to diabetics. Given the

---

100. Genetic diseases, like SCIDs or Tay-Sachs, can appear in children even when the parents are unaffected because every cell actually contains two complete copies of the genome, one inherited from each parent. Usually, one copy of a gene can make enough of a protein for the child to be healthy. It is only when a child in unlucky enough to inherit a bad gene from both parents that the disease will result. There are, however, some genetic diseases where only one bad copy is enough to result in disease. Huntington's disease is a particularly insidious example of this. It is insidious because the disease does not manifest itself until a victim is in their 40s or 50s. People can therefore pass this debilitating disease onto their children without realizing it.

101. *See* ANGIER, *supra* note 29, at 142-43.

102. *See* YOXEN, *supra* note 89, at 11 (1983).

103. *See id.* at 90.

104. *See id.*

105. *See id.* at 91.

106. *See id.*

107. *See id.*

state of the knowledge and technology at the time, each of these steps was an amazing technical breakthrough.

4. *CURRENT TECHNOLOGICAL CAPABILITIES OF BIOTECHNOLOGY*

Given the great strides of the technology in the last few years, Genentech's amazing feat of the early 80s would now be almost trivial. Indeed, once certain scientific principles had been discovered, and technical problems solved, application of those technical solutions to new scientific problems was made theoretically obvious (although a certain amount, sometimes a great amount, of trial and error is required for any given problem to be solved). To understand what biotechnology is currently capable of, a few more facts about the interrelationships of DNA, mRNA, and protein are needed.

DNA is a double stranded molecule which looks like a ladder with each rung split in half. Each half-rung is "complementary" to its matching half-rung. Only one vertical half of the ladder carries protein coding information. RNA is a single stranded molecule complementary to the coding strand of DNA. Because of the relationship between DNA and mRNA, if you know the sequence of the DNA of a gene, the sequence of the mRNA can be deduced exactly. Conversely, knowing the sequence of the RNA allows you to deduce the sequence of the DNA exactly.

The relationship between mRNA (and, since it is informationally equivalent, DNA) and protein is more complicated. The ladder half-rungs of RNA are four different types of chemical bases. Protein is made from sequences of 20 different kinds of amino acids.[108] Three consecutive DNA bases code for each amino acid in the sequence, providing 64 different possible codons for 20 amino acids. As a result of this "degeneracy" of the code (that is, that a given amino acid may have been coded for by one of several possible mRNA/DNA sequences of three bases), while one can deduce the protein amino acid sequence exactly from a given RNA or DNA sequence, for a given the protein sequence, there are many possible mRNA/DNA sequences.[109] For instance, bovine insulin has 51 amino acids.[110] Given an average of three possible base sequences for each amino acid, there are billions[111] of possible DNA sequences.

This does not mean, however, that if only the protein sequence is known, the DNA/RNA sequence cannot be found. As little as a

---

108. *See id.* at 135.

109. *See id.* at 146.

110. *See* SUZUKI ET AL., *supra* note 24, at 286.

111. Specifically, 3^{51}, or 2.2×10^{24}.

sequence of DNA encoding a sequence of 6 amino acids will only be found once in the entire human genome. Using straightforward technology, every possible DNA combination that will code for that 6 amino acid sequence can be chemically synthesized. This is about 700 different 18-base DNA sequences. You can then use this collection of small pieces of DNA sequence as a kind of very specific molecular magnet to retrieve the rest of the gene. Only the one sequence that matches the actual genetic sequence in the organism will retrieve the full length gene, but all 700 sequences can be tried together in one experiment.

Thus, if one were trying to produce recombinant insulin, one could sequence a piece of the naturally-occurring insulin, synthesize a collection of all the possible DNA sequences of that piece, and use the collection of DNA sequences to retrieve the rest of the gene. You could even start with bovine insulin and use that to retrieve the human insulin gene because bovine and human insulin are very close to identical. You could then insert the human insulin gene into a bacteria and allow the bacteria to divide and make buckets of human insulin. Diabetics could inject this insulin with no fears of allergic reaction. This type of protein-production biotechnology is also used in non-medical applications. For instance, bovine somatotropin is produced by genetically engineered bacteria to be injected into cows to increase their milk production.[112]

Biotechnology is not limited to this type of protein manufacture. It can also be used in so-called "gene therapy." There are two types of gene therapy. For example, a diabetic's problem is that his or her pancreas does not make insulin because the gene for insulin is partially missing or mutated. One type of gene therapy would attack the problem by changing the pancreas of the diabetic. A few pancreas cells could be removed and the correct insulin gene could be inserted into the cells. The cells could then be placed back into the patient's pancreas, where they could divide and produce insulin.

While not being pursued currently in the medical context, a second type of gene therapy, "germ-line gene therapy," is being pursued in agricultural and environmental contexts. For instance, certain tomatoes on the market have themselves been genetically engineered to retard spoilage.[113] Rather than genetically engineering each individual tomato, the gametes of a tomato were altered thereby, creating an entirely new line of genetically engineered plants. Research is being done to create plants that make their own pesticides: pesticides that cannot wash off and contaminate the environment, are resistant to plant viruses, and have increased nutritional value.[114]

---

112. *See* Herman, *supra* note 31, at 112.

113. The "Flavr-Savr" tomato, developed by Calgene in Davis, California.

114. *See* Herman, *supra* note 31, at 109-11.

C. The Current Nonobviousness Doctrine Applied to Biotechnology Inventions

In Part VI of this paper, a modification of the nonobviousness requirement for biotechnological inventions is proposed. As a backdrop to this proposal, a discussion of the current biotechnology nonobviousness doctrine is in order. This will give a basic doctrine upon which the modification will overlay, and which will serve as an example to demonstrate the superiority of the proposed modification over the current doctrine.

In re Deuel[115] is the most recent pronouncement of the Federal Circuit with regard to the nonobviousness requirement for biotechnology inventions. The patent applicants (referred to here collectively as "Deuel") appealed to the Federal Circuit from a decision of the Board of Patent and Trademark Appeals and Interferences upholding a final rejection based on the examiner's finding of obviousness.[116]

The invention in *Deuel* is much like the description of the finding of the insulin gene discussed above. Deuel had purified a protein called heparin-binding growth factor from bovine uterine tissue.[117] He was interested in this protein because it stimulates cell division and, thus, may be useful in facilitating the repair and replacement of damaged or diseased tissue.[118] He figured out the sequence of amino acids for a small piece of the protein and then synthesized a collection of DNA fragments that represented all the possible DNA sequences for the heparin-binding growth factor.[119] He then used this collection of DNA sequences to pull out the complete gene from a collection of all the genes transcribed in the bovine uterus.[120] He also used the collection of short DNA sequences to isolate the human heparin-binding growth factor gene from a collection of genes transcribed in human placentas (recall that human and bovine genes are generally quite similar).[121] He then sequenced both genes and deduced from the sequences what the complete protein amino acid sequences are.[122] This last step may seem odd since he could have simply found the entire protein sequence from the isolated protein that he started

---

115. 51 F.3d 1552 (Fed. Cir. 1995).

116. *Id.* at 1552.

117. *Id.* at 1555.

118. *Id.* at 1554. Note that this activity is exactly the *opposite* of what one would be looking for to treat cancer, where there is uncontrolled cell division. Heparin-binding growth factor may thus also be interesting to cancer researchers.

119. *Id.* at 1555.

120. *Id.*

121. *Id.*

122. *Id.*

with. However, determining exact amino acid sequences is technically difficult and time consuming. In contrast, sequencing DNA is relatively straightforward. Also, recall that once a DNA sequence is known, deducing the protein sequence is trivial. Since Deuel was going to find the gene and sequence it anyway, it made much more sense to wait to find out the entire protein sequence.

Deuel claimed in his patent application the human and bovine DNA sequences, as well as the deduced amino acid sequences. The examiner issued a final rejection based on a finding that the invention was obvious. The rejection was based on the fact that a partial amino acid sequence for heparin-binding growth factor had previously been published.[123] The examiner reasoned, and the Board agreed, that this published sequence, combined with the routine nature of finding the gene and complete protein sequence, rendered the invention obvious.[124]

The issue confronted by the Federal Circuit was "whether the combination of a prior art reference teaching a method of gene cloning, together with a reference disclosing a partial amino acid sequence of a protein, may render DNA ... molecules encoding the protein prima facie obvious under Section 103."[125] The focus of the inquiry was whether a holding of obviousness could only be supported by the actual physical structure of the DNA and protein being made obvious by the prior art or whether obvious methods of finding that structure are sufficient to support the holding of obviousness.[126]

The court held that the claimed DNA and protein sequences were nonobvious.[127] The reasoning was based primarily on analogizing molecular biological inventions to chemical inventions. In chemical cases, the teaching of a particular compound, combined with a suggestion to make a specific kind of change to that compound, renders the new compound obvious.[128] The court noted that because the prior art partial

---

123. 51 F.3d at 1555. There was actually some dispute as to whether the published partial amino acid sequence was for heparin-binding growth factor or not. The published sequence was of a protein found in bovine brain (not uterus, as Deuel found here) and exhibiting similar properties to heparin-binding growth factor. The Federal Circuit assumed the proteins were the same for purposes of their discussion.

124. *Id.*

125. *Id.*

126. *Id.*

127. *Id.* at 1560.

128. *See, e.g., In re* Jones, 919 F.2d 688, 692 (Fed. Cir. 1990) (en banc) ("[s]tructural similarity between claimed and prior art subject matter, ... where the prior art gives reason or motivation to make the claimed compositions creates a prima facie case of obviousness"), *cert. denied*, 500 U.S. 904 (1991); *In re* Grabiak, 769 F.2d 727, 731-32 (Fed. Cir. 1985) ("[T]here

amino acid sequence was not a DNA sequence, it could not render the DNA sequence structure obvious.[129] Furthermore, the court reasoned, because there were so many possible DNA sequences that could potentially code for the protein, a person of ordinary skill in the art could not have determined the DNA sequence without actually doing the experiment that Deuel did.[130] The court stated, "[w]hat cannot be contemplated or conceived cannot be obvious."[131]

The court rejected the argument that the genetic code relationship between DNA and proteins rendered the claimed compounds obvious.[132] Citing *In re Baird*,[133] the court stated that the disclosure of a "broad genus of [compounds] does not necessarily render obvious each [species] compound within its scope."[134] In other words, the disclosure of the protein sequence, which could be used to deduce the billions of *possible* DNA sequences, does not render the *particular* DNA sequence actually found in the organism obvious. Importantly, the court did not reach the question whether the prior art disclosure of a small protein with a very limited number of possible DNA sequences would make each sequence obvious.[135]

The court also rejected the argument that a person with ordinary skill in the art would, given the prior art, know how to make the claimed compounds. The court stated that, because it was compounds, and not the methods to make the compounds, that were being claimed, the fact that the methods were known was irrelevant.[136] This was based on the decision in *In re Bell*.[137]

---

must be adequate support in the prior art for the ester/thioester change in structure, in order to complete the PTO's prima facie case [of obviousness].").

129. *Deuel*, 51 F.3d at 1559.

130. *Id.*

131. *Id.*

132. *Id.*

133. 16 F.3d 380 (Fed. Cir. 1994). *Baird* dealt with a prior art reference that disclosed a general chemical formula with many variables in it. The applicant claimed a compound that was one of the more than 100 million different possible chemicals under the prior art reference. The court held that the applicant's compound was nonobvious because it worked well and the prior art reference made no suggestion that this particular compound would work well, or even at all.

134. *Id.*

135. *Id.* The court cited *In re Petering*, 301 F.2d 676 (C.C.P.A. 1962), where a prior art reference disclosing a genus with 20 species was held to render obvious each of the 20 species.

136. *Deuel*, 51 F.3d at 1559.

137. 991 F.2d 781, 785 (Fed. Cir. 1993).

In *Bell*, the court held that the prior art disclosure of the complete amino acid sequence of insulin-like growth factors I and II (IGF I and II), in combination with a prior art patent entitled "Method for Cloning Genes," did not make the cloning of the IGF I and II genes obvious.[138] The fact situation, however, was not entirely analogous to that in *Deuel*. The court discussed the vast numbers of possible DNA sequences and the lack of suggestion of the structure of the particular DNA sequence in the prior art. However, the court based its holding, at least in part, on the fact that the "Method for Cloning Genes" patent counseled the use of a short string of amino acids with unique DNA codes, thereby limiting the amount of degeneracy that needed to be addressed. In *Bell*, the actual amino acids used to construct the DNA probes did not have unique DNA sequences. The court stated that "we cannot say that [the patent] fairly suggests that its teachings should be combined with those of [the amino acid sequences], since it nowhere suggests how to apply its teachings to amino acid sequences without unique codons."[139] Because the cloning method relied on by the PTO in *Deuel* did not depend on the use of the unique DNA sequences for amino acids, this reasoning from *Bell* does not apply to *Deuel*.

VI. A PROPOSED MODIFICATION OF THE NONOBVIOUSNESS REQUIREMENT WHICH WILL GIVE AN INCENTIVE TO THE BIOTECHNOLOGY INDUSTRY

A. The Proposal

The current test for nonobviousness is a test for technical nonobviousness. This paper proposes an additional test for economic nonobviousness. Under this test, an invention need only be technically *or* economically nonobvious to be patentable. To understand this distinction, imagine a proposed invention that has a 1% chance of becoming *commercially* successful (recall that under the traditional test for nonobviousness, the invention need only be *technically* successful, not commercially successful). Imagine further that it will take $5 million to research and develop this invention to the point where commercial viability is apparent. If the expected payoff is $501 million, the invention is efficient to pursue because expected payoff, discounted by the probability of success is less than the investment risked. However, because of additional discounting of the expected payoff due to the natural, rational risk aversion of a rational economic actor, this socially desirable invention will not be pursued. Thus, although it might be

138. *Id.* at 783.

139. *Id.* at 784 (internal quotation marks omitted).

conceded that an invention is "obvious" in the sense that it is technically *possible*, an invention may be economically, or relatively, nonobvious because a rational economic actor would not pursue it.

Recognizing this type of economic nonobviousness does two things. First, it reflects the fact that inventions are not made in an economic vacuum. Technical nonobviousness gives inventors an incentive to go out on a technological limb; economic nonobviousness gives investors an incentive to go out on an economic limb. Given two possible inventions with the same expected return (projected return discounted by the chance of failure), investors are more likely to invest in the less risky invention. If the riskier invention is a socially desirable one, however, giving the extra incentive of greater ease in patent acquisition will shift investment toward that socially desirable invention. Second, assuming that expected payoff takes into account the normal chances of getting a patent on the invention, making the patent easier to get would increase the expected payoff. This increase results both because of the patent monopoly rights granted and because a patent can be an early property right that will discourage competition even before the invention is developed. Thus, the recognition of economic nonobviousness creates the needed additional incentive so that socially desirable inventions will be pursued.

Of course, this example assumes that the patent office can determine *ex ante* both the probability of commercial success and the potential commercial payoff. This assumption is patently false. What the example serves to illuminate is the existence of a class of inventions that would not be granted patents under the current doctrine of nonobviousness, but which should nonetheless be granted patents in a perfectly rational system. What is needed, then, is a proxy for determining which inventions are likely to fall into this class.

Because biotechnological inventions are expensive to pursue and have a high risk of ultimate commercial failure, the classification of an invention as biotechnological can serve as this proxy. While this will not address the issue of which inventions will ultimately prove to be commercially successful it will pick out those inventions which are likely to be under-pursued due to risk aversion.

In addition to specifying which inventions will receive special treatment, the type of special treatment needs to be defined. One alternative might be to determine nonobviousness based on actual, individual expense.[140] For instance, an invention might be classified as economically nonobvious if it actually cost more than 150% of the average cost for the industry.[141] This type of system is administratively intractable and fraught with opportunity for deception or manipulation.

---

140. *See* MERGES, *supra* note 41, at 418.

141. *See* Merges, *supra* note 20, at 1.

It would encourage inefficient spending to increase the possibility of patent protection and the "average" cost for the industry would spiral upward. Although this system may appear theoretically sound, practical considerations must prevail. Furthermore, this alternative does nothing to address the issue of an entire *industry* which is expensive and risky to pursue. A test based on actual numbers needs a baseline for comparison; the baseline in the biotechnology industry is too high to make this alternative workable.

Instead, I propose a definition of economic nonobviousness which is much closer in spirit to the current definition of technical nonobviousness. The test should be whether one skilled in the art, *considering the cost and likelihood of commercial success*, would have regarded the invention as an obvious one to pursue. This is a highly subjective test. Nonetheless, because it is modeled on the test for technical nonobviousness, judges, patent practitioners, inventors, and investors already have the requisite experience. Indeed, it would actually be easier to prove economic nonobviousness, because it lends itself to hard evidence much more easily than technical nonobviousness.

Precise definitions of economic nonobviousness aside, the mere existence of a label of economic nonobviousness will be helpful. Currently, it appears that when the PTO or Federal Circuit is faced with a technically obvious invention that the judges or examiners intuitively recognize should have patent protection (that is, they intuitively respond to the economic nonobviousness of the invention), they are forced to bend the technical nonobviousness doctrine to fit what they perceive to be the only just outcome. By giving the agency and the court a label for what they are doing, the practice can be legitimized and brought out so that it can be discussed without subterfuge.

B. The Proposal Responds to the Policy Needs of the Biotechnology Industry, and Its Implementation Is Feasible

The economic nonobviousness standard gives an extra incentive to biotechnological inventions. By acknowledging that there are types of inventions that deserve special patent treatment because they are particularly expensive or risky, it gives the needed incentive to overcome the risk aversion that is otherwise problematic in the industry. In doing so, it gives economic support to the industry as whole. The economic advantage of finding some inventions economically nonobvious will spill over to other inventions in the intellectual property portfolio.

Because this modification will support the industry as a whole, it is more likely that the biotechnology industry will be able to raise the funds it needs to maintain its academic laboratory model. Without this change, it is more likely that the current spate of mergers and acquisitions of

biotechnology companies will continue, thus undermining the otherwise preferable research model. As discussed above, the model of the financially and economically independent, focused, academic lab must be seen as vitally important to the continued successes in biotechnology. By granting patent rights to economically risky inventions, the economic nonobviousness doctrine allows small biotechnology companies to remain independent. Without this type of patent protection, it is more likely that this type of research would be forced into large conglomerates, where economic risk could be spread, but at the expense of the culture which has been responsible for so many successes.

The addition of the proposed doctrine of economic nonobviousness reflects the economic realities that are so important to the business of innovation. It also reflects the fact a person of ordinary skill in the art would take into account economic realities when deciding if a given invention had a reasonable expectation of success. The purpose of patents is to encourage innovation, and the recognition of the economics of innovation will facilitate this purpose. Ignoring economics is illogical and not mandated by the statute.

Also, the concept of economic nonobviousness is not inconsistent with the *current* doctrine of nonobviousness. For instance, section 103 states that an invention's nonobviousness "shall not be negatived by the manner in which the invention was made."[142] This sentence was included in order to abolish the "flash of genius" test previously used in judicial determinations of nonobviousness.[143] That is, "brute force" or trial and error inventions can be nonobvious. This language tacitly allows economic factors to be considered. Currently this language is used only to support inventions which *could have been* made through a flash of genius, but happened to have been made through hard work. Nonetheless, it is not inconsistent to say that trial and error invention that one of ordinary skill in the art would not have done (for technical *or* economic reasons), is not "negatived by the manner in which [it] was made."[144] The secondary considerations of commercial success, failure of others, and long-felt need also import some economic considerations to the traditional nonobviousness test.

The economic nonobviousness doctrine also compensates for some unfairness to biotechnology inherent in the traditional nonobviousness test. The "ordinary skill in the art" of biotechnology is very high because of the high degree of education required even to be a technician in the field. Because the nonobviousness of an invention is determined with

142. 35 U.S.C. § 103 (1995).

143. *See* Graham v. John Deere Co., 383 U.S. 1, 15 (1966) (citing the "flash of genius test" in Cuno Corp. v. Automatic Device Corp., 314 U.S. 84 (1941)).

144. 35 U.S.C. § 103 (1995).

reference to the technical skill of "one of ordinary skill in the art," this makes a legitimate finding of nonobviousness quite difficult. By incorporating the real-life economic concerns of "one of ordinary skill in the art," the doctrine of economic nonobviousness compensates for the difficulty in finding technical nonobviousness in a field of high ordinary skill.

Another major advantage of the proposed economic nonobviousness requirement is that the implementation is politically feasible. Such a policy theoretically could be implemented in one of three ways: through examiner guidelines issued by the PTO after notice and comment, through judicial interpretation, or through congressional action. Because the proposed modification would constitute a real change in the law, and because the PTO issued examiner guidelines are not technically rulemaking and thus do not have the force of law, it is probably inappropriate for this modification to be implemented by the PTO. However, either of the other two possible methods of implementation is not only proper, but possible.

The Federal Circuit is a patent policy making body. As discussed earlier, there is generally little oversight by the Supreme Court. Consequently, any doctrinal changes made by the Federal Circuit are unlikely to be overruled. Since the use of economic nonobviousness is justifiable logically and legally, it is entirely possible that the Federal Circuit could choose to implement it. Moreover, the Federal Circuit has shown a willingness to make doctrinal changes in the nonobviousness area. For instance, the Supreme Court, in *Graham v. John Deere Co.*,[145] mentioned that "[s]uch secondary considerations as commercial success, long felt but unsolved needs, failure of others" might have "relevancy" to the question of nonobviousness.[146] The Federal Circuit has drastically increased the importance of secondary considerations or objective indicia. In *Hybritech, Inc. v. Monoclonal Antibodies, Inc.*,[147] the court stated that secondary considerations "must be considered before a conclusion on obviousness is reached."[148]

Congressional action in this area is also possible. Congress' willingness to carve out special considerations for particular industries is evidenced by the recent amendments to the nonobviousness statute with regard to biotechnology inventions. Under the amendment, signed into law in November, 1995, an otherwise obvious biotechnological process

145. 383 U.S. at 17-18.

146. *Id.* at 17-18.

147. 802 F.2d 1367 (Fed. Cir. 1986), *cert. denied*, 480 U.S. 947 (1987).

148. *Id.* at 1379.

will be considered nonobvious if it uses or results in a product that is nonobvious.[149]

Between these two implementation strategies, judicial implementation is preferable. There are several arguments for this type of implementation. First, the Federal Circuit knows more about patent law and is therefore likely to make better decisions about it than Congress. Second, a judicially created doctrine can more easily change over time because it does not need to worry about political majorities. Consequently, the doctrine can evolve with the needs of the industry. Third, the Federal Circuit's current approach to nonobviousness is actually quite similar to the modification proposed here. Economic considerations are being made *sub rosa*; they should be made explicitly.

By any honest interpretation of the nonobviousness statute, the decisions in *Deuel*[150] and *Bell*[151] are wrong. Given the sequence, or partial sequence of a protein, along with known molecular cloning techniques, any person of ordinary skill in the art would reasonably expect to have success in cloning the gene that encodes the protein. Granted, there may be technical difficulties, but given infinite time and financial resources, which the traditional technical nonobviousness test appears to give, one would expect eventually to find the gene.

The court's focus on the fact that a particular structural product is claimed is misplaced. While comparing structures of simple organic compounds in chemical cases makes some sense scientifically, it makes no sense to draw this distinction in biotechnological cases, because DNA and proteins are informationally related. The court's focus on the different chemical structure of the two molecules is akin to saying an English translation of a French book is nonobvious.

The method/product distinction is also insupportable. It ignores the scientific reality that method and product are intimately related. It is disingenuous to say that a person of ordinary skill in the art could not take a known starting product, and apply a known method to create a new product. The distinction drawn by the court very much raises form over function. Furthermore, it encourages manipulative claiming in order to avoid a finding of obviousness. Because methods and products are functionally related, virtually any product can be claimed as a method and vice versa. The court's distinction gives a perverse incentive for inventions to be claimed in non-straightforward ways.

It is also possible that, after the economic nonobviousness doctrine has been developed judicially, it could be legislatively ratified. This is what happened with the current doctrine of nonobviousness. The Patent

149. 35 U.S.C. § 103(b) (1995).

150. 51 F.3d 1552 (Fed. Cir. 1995).

151. 991 F.2d 781 (Fed. Cir. 1993).

Act prior to 1952 required only that an invention be novel and useful. The requirement of nonobviousness was judicially created. The doctrine was tied to the statutory language by the argument that if obvious, there was no true "invention."[152] After over a century of judicial evolution, the requirement of nonobviousness was added in the Patent Act of 1952.

VI. SUMMARY

Biotechnology offers great promise to the health and welfare of the American people, as well as great promise to the American economy. It creates high wage jobs and enjoys a huge trade surplus. Perhaps more importantly, it offers tangible, non-economic advantages. Disease treatment and prevention, and more efficient production of more nutritious food with limited environmental impact are made possible by biotechnology.

However, as the realities of technical difficulties and economic limitations have become apparent, the promise of biotechnology has moved further from our reach. In response to the needs of the industry for a culture that encourages innovation and research and development funds for expensive and risky invention, the patent law should be changed to give certain biotechnology inventions an easier time of getting a patent.

By creating a judicial doctrine of economic nonobviousness for biotechnological inventions, the needs of the industry can be met without sacrificing health and safety regulation or burdening the public fisc. The proposed doctrine is logically sound, consistent with patent policy, and implementation is feasible.

152. Hotchkiss v. Greenwood, 52 U.S. 248 (1850).

ARTICLE

PROTECTING THE PRIVATE INVENTOR UNDER THE PEACETIME PROVISIONS OF THE INVENTION SECRECY ACT

SABING H. LEE †

TABLE OF CONTENTS

I. INTRODUCTION 346
II. THE LEGISLATIVE HISTORY OF INVENTION SECRECY 348
A. The First Invention Secrecy Act 348
B. Invention Secrecy During World War II 349
C. The Creation of Peacetime Invention Secrecy 351
D. Evaluation of Peacetime Invention Secrecy as Legislative Policy 353
III. THE MECHANICS OF INVENTION SECRECY 358
A. Procedure for Issuing a Secrecy Order 359
B. The Effect of a Secrecy Order 363
C. Inventors' Rights and Remedies Under the Act 366
D. Fairness Analysis 369
IV. CONSTITUTIONAL QUESTIONS 377
A. Constitutional Authority for Invention Secrecy 378
B. First Amendment Analysis 379
C. Fifth Amendment Analysis 399
V. CONCLUSION 409

† Associate, Knobbe, Martens, Olson & Bear, LLP, Newport Beach, CA; J.D., 1997, Boalt Hall School of Law, University of California at Berkeley; B.S. & M.S., 1994, University of California, Los Angeles. Special thanks to Professor Robert P. Merges and Stephan W. Gruen for their comments and suggestions.

I. INTRODUCTION

The Invention Secrecy Act of 1951[1] provides that whenever "the publication or disclosure of the invention by the granting of a patent ... would be detrimental to the national security ... the Commissioner [of Patents] shall order that the invention be kept secret ...."[2] Such an order prohibits the inventor from publishing or disclosing any material information relating to the invention.[3]

Since 1917, when the first invention secrecy act was created,[4] until the end of a twenty-nine year national emergency in 1979,[5] virtually all of the secrecy orders existing in this country were authorized under legislation designed to protect the United States against the threat of conflict during times of war or national emergency.[6] With the recent end to the Cold War, however, the threat to national security diminished significantly.[7] Thus, one may ask whether the invention secrecy doctrine, which was created specifically in response to the threat of war, can be applied to authorize secrecy orders during a time of peace.

The government's increasing propensity to impose peacetime secrecy orders underscores the need to reassess their validity under changed conditions. Since 1983, the number of government-imposed secrecy orders increased by about 40 percent, from 3,900 in 1983 to 6,033 in 1993.[8] This latter figure is comparable to the number of orders in effect at the height of World War II.[9] This increase may be partially attributed to

---

1. Invention Secrecy Act of 1951, ch. 4, 66 Stat. 3 (1952) (codified as amended at 35 U.S.C. §§ 181-188 (1994)).

2. 35 U.S.C. § 181 para. 3 (1994).

3. *See* 35 U.S.C. §§ 182, 186 (1994).

4. Act of Oct. 6, 1917, ch. 95, 40 Stat. 394 (1917).

5. President Truman declared a national emergency on December 16, 1950. Proclamation No. 2914, 3 C.F.R. 99 (1949-1953). This national emergency was terminated by the National Emergencies Act of 1976, Pub. L. No. 94-412, Title I, § 101, 90 Stat. 1255 (1976).

6. *See* discussion *infra* Parts II.A-C.

7. The United States still faces potential threats from countries such as France, China, Japan, Israel, Iraq, and Russia. *See* Ralph Vartabedian, *Most Promising U.S. Technology Still Kept Secret*, L.A. TIMES, July 13, 1993, at A1, A12. However, none of these threats has ever created a serious threat of actual conflict reaching American soil.

8. *See* Steven Aftergood, *Invention Secrecy Criteria Disclosed*, SECRECY & GOV'T BULL., Nov. 1994 (Federation of American Scientists) (visited Nov. 23, 1997) <http://www.awpi.com/IntelWeb/US/S-GB/041.html>.

9. During World War II, the number of secrecy orders in effect peaked at 8,293 on December 31, 1944. *See* H.R. REP. NO. 1540, 96th Cong., 2d Sess. 37 (1980). From 1951 to

the Cold War threat during the 1980s and partially attributed to the general increase in the number of patents being issued. However, the number of secrecy orders is still unusually high for a country that currently faces no major conflict.[10]

Invention secrecy orders negatively impact private interests. Although invention secrecy originally affected only military employees and military technologies, in recent years the Invention Secrecy Act has increasingly been applied to private inventors.[11] Many of these inventors develop dual-use technologies, i.e., inventions with both military and commercial purposes.[12] In 1991, over three-quarters of all new secrecy orders—506 out of 774—were issued to private inventors.[13] These orders covered technologies such as computer hardware, advanced ceramics, laser systems, semiconductor manufacturing technologies, automated process control systems, highly specialized software, video display technology, space photography, industrial plating, and advanced sensors.[14] Thus, one may ask whether peacetime secrecy orders place unnecessary burdens upon private inventors of primarily non-military inventions.

This article addresses general questions regarding the necessity and justifications for an invention secrecy doctrine during peacetime, and specifically, whether the Invention Secrecy Act as currently enacted adequately protects the rights of private inventors. This article is divided into three parts. The first part of the article traces the legislative development of the invention secrecy doctrine, focusing on its historical justifications and the interests that the legislature intended to protect. The second part explores the mechanical aspects of the Act, paying special attention to how a secrecy order is imposed and how inventors may be compensated for the issuance of a secrecy order, in order to

---

1958, the number of secrecy orders rose from 3,435 to 6,149, and remained between 4,100 and 5,100 for the next twenty years. *See id.* at 1-2.

10. *See Even After Cold War, Patents Remain Secret*, INSIDE R&D (Technical Insights, Inc.), June 3, 1992, *available in* 1992 WL 2799306.

11. *See id.*; *see also* Gary L. Hausken, *The Value of a Secret: Compensation for Imposition of Secrecy Orders Under the Invention Secrecy Act*, 119 MIL. L. REV. 201, 202 (1988) (citing Interview with John Raubitchek, Patents, Copyrights, and Trademarks Division, Office of the Judge Advocate General, Department of the Army (Feb. 26, 1987)).

12. *See Even After Cold War, Patents Remain Secret*, *supra* note 10.

13. *See* Edmund Andrews, *Cold War Secrecy Still Shrouds Inventions*, S.F. CHRON., May 23, 1992, at A23. This is in comparison to 43 of 250 secrecy orders in 1979. *See id.*

14. *See id.*; *see* Vartabedian, *supra* note 7, at A12 (the push toward commercial applications is signified by President Clinton's proposed twenty billion dollar effort to help convert the defense industry to commercial enterprises).

ascertain whether the Act sufficiently minimizes burdens on private inventors. The final part applies constitutional scrutiny to the peacetime provisions of the Invention Secrecy Act under the First and Fifth Amendments to determine whether the burdens placed on private inventors are constitutional. The article concludes that while invention secrecy remains an important peacetime policy, the government should limit its application to ensure that an inventor's patent rights are not unduly burdened.

II. THE LEGISLATIVE HISTORY OF INVENTION SECRECY

A. The First Invention Secrecy Act

The power of the government to keep certain inventions secret to protect national security is well-established in the United States. Since World War I, the government has been concerned that "those inventions which are of most use to the Government during a time of war are also those which would, if known, convey useful information to the enemy."[15] For this reason, in preparation for the United States' entry into World War I, Congress passed the Act of October 6, 1917.[16] The 1917 Act stated, in part:

> That whenever during a time when the United States is at war the publication of an invention by the granting of a patent might, in the opinion of the Commissioner of Patents, be detrimental to the public safety or defense or might assist the enemy or endanger the successful prosecution of the war he may order that the invention be kept secret and withhold the grant of a patent until the termination of the war ....[17]

At the heart of this provision is Congress' desire to protect the "public safety or defense" *during wartime*. The legislative history accompanying the Act also emphasizes this point.[18] When World War I ended in 1918, the necessity for invention secrecy, at least for the time being, also expired.

While the 1917 Act was in effect, the burdens of invention secrecy fell most directly on individuals seeking patents for inventions with potential national security implications. For inventions which the Commissioner of Patents deemed secrecy to be appropriate, the government issued a secrecy order temporarily preventing the inventor

---

15. S. REP. NO. 119, 65th Cong., 1st Sess. at 1 (1917).

16. Act of Oct. 6, 1917, ch. 95, 40 Stat. 394 (1917).

17. *Id.*

18. *See* S. REP. NO. 119, *supra* note 15, at 1.

from obtaining a patent until the end of the war.[19] The government enforced strict penalties for the violation of secrecy orders: by publishing an invention or filing for a patent in a foreign country without proper permission, inventors could forever lose their right to obtain a patent.[20]

Placing the burden of invention secrecy on inventors created two problems for Congress. First, the mere threat of a secrecy order might have discouraged an inventor from even filing for a patent, thereby depriving the government of potentially useful inventions. Second, inventors' rights required protection under the Act. The Senate Report accompanying the 1917 Act recognized these problems and emphasized the need "to stimulate invention, and provide adequate protection to owners of patents."[21] The solution was a system of compensation:

> When an applicant whose patent is withheld as herein provided and who faithfully obeys the order of the Commissioner of Patents above referred to shall tender his invention to the Government of the United States for its use, he shall, if and when he ultimately received [sic] a patent, have the right to sue for compensation in the Court of Claims ....[22]

By providing for a system of compensation, the Act attempted: (1) to stimulate invention, by eliminating the deterrent effects that the threat of a secrecy order might have on an inventor thinking about applying for a patent; and (2) to protect inventors' rights, by substituting monetary compensation for the reward that an inventor otherwise might have received from timely receipt of a patent. The inventor received no compensation, however, for those inventions that created a national security threat but for which the government had no use. By conditioning the right to compensation on the tendering of the invention and the use of the invention by the government, Congress meant only to stimulate inventions that the government might actually use for its own purposes.

The duration of a secrecy order was limited to the period of the war. Therefore, most secrecy orders lasted only about a year. Because the war ended so quickly, the Act of 1917 did not substantially affect inventors.

B. Invention Secrecy During World War II

After World War I, the invention secrecy doctrine lay dormant for more than two decades. In 1940, in preparation for the United States' entry into World War II, the statute was renewed by the Act of July 1,

---

19. *See* Act of Oct. 6, 1917, ch. 95, 40 Stat. 394 (1917).

20. *See id.*

21. S. REP. NO. 119, *supra* note 15, at 1.

22. Act of Oct. 6, 1917, ch. 95, 40 Stat. 394 (1917).

1940.[23] The 1940 Act basically reinstated the essential provisions of the 1917 Act. Congress ensured that the effect of invention secrecy would be temporary, giving the Act a duration of only two years.[24] The specific two-year period was necessary because at the time of enactment the United States had not yet entered the war, so Congress needed a way to ensure that the 1940 Act would have a definite end. On June 16, 1942, after the United States had officially entered the war, Congress amended the 1940 Act to provide that its provisions remain in effect "during the time when the United States is at war."[25] Thus, like the 1917 Act, the 1940 Act authorized the use of secrecy orders only during wartime.

The invention secrecy doctrine of World War II contained several notable changes from the World War I doctrine. In particular, the Act of August 21, 1941,[26] amending the 1940 Act, placed several additional limitations on inventors' rights. The 1941 Act reemphasized the restriction found in the 1917 Act, that to apply for a patent abroad, the applicant must obtain a license from the Commissioner of Patents.[27] Any person found in noncompliance with this licensing procedure was barred from receiving a United States patent.[28] Stricter penalties were created for violation of a secrecy order. In addition to abandonment of the invention, an inventor could be fined $10,000 or be imprisoned if convicted of willful publication or disclosure of an invention under secrecy order, or of willful filing in a foreign country of a patent application.[29]

During World War II, the government formalized the decision-making process guiding the issuance of secrecy orders. In August 1940, the Commissioner of Patents requested the Secretary of War and the Secretary of the Navy to create the Army and Navy Patent Advisory Board (ANPAB).[30] ANPAB assisted the Patent Office in determining when a patent application might relate to issues of national defense.[31] In 1948, ANPAB was renamed the Armed Services Patent Advisory Board

---

23. Act of July 1, 1940, ch. 501, 54 Stat. 710 (1940).
24. *See id.*
25. Act of June 16, 1942, ch. 415, 56 Stat. 370 (1942).
26. Act of Aug. 21, 1941, ch. 393, 55 Stat. 657 (1941).
27. *See id.*
28. *See id.*
29. *See id.*
30. *See* H.R. REP. NO. 1540, *supra* note 9, at 37.
31. *See id.*

(ASPAB).[32] ASPAB still exists and continues to provide the Patent Office with substantial guidance on when to issue a secrecy order.[33]

C. The Creation of Peacetime Invention Secrecy

The end to World War II in 1945 brought with it the general end to secrecy orders that had been imposed during the war. On November 30, 1945, the Commissioner of Patents issued a rescission order to release most of the patent applications from secrecy order status.[34] This removed 6,575 secrecy orders.[35] The Secretary of Commerce explained that because the applications have both military and commercial significance, the owners of these applications should to be able to exploit their inventions commercially.[36]

Despite the end to actual conflict, national security continued to be a significant concern. Cold War tensions were mounting, as indicated by President Truman's September 23, 1949 announcement that the Soviet Union had tested an atomic bomb.[37] In 1950, the start of the Korean War again brought war-related national security issues to the forefront. By 1951, the United States still had not declared an end to the war with Germany, nor had it signed the Japanese Peace Treaty.[38] As a result, certain patent applications continued to remain under secrecy orders past 1945. These secrecy orders were properly issued under the 1940 Act and its amendments because the national emergency declared prior to World War II was not terminated until 1952.[39] As of December 31, 1945, 799 applications were still kept secret by the government.[40] However, the government continued to issue new secrecy orders for applications

---

32. *See id.*

33. *See* 37 C.F.R. § 5.1; 35 U.S.C. § 181 para. 3; Aftergood, *supra* note 8.

34. *See* H.R. REP. NO. 1540, *supra* note 9, at 46.

35. *See id.* at 47.

36. *See id.* at 46.

37. *See id.* at 50.

38. *See* S. REP. NO. 1001, H.R. REP. NO. 1028, 82d Cong., 1st Sess. (1951), *reprinted in* 1952 U.S.C.C.A.N. 1322.

39. The formal end to the hostilities was declared by President Truman on December 31, 1946. Proclamation No. 2714, 3 C.F.R. 99 (1943-1948), *reprinted in* 61 Stat. 1048 (1946). The national emergency declared in preparation for World War II was terminated on April 28, 1952. Proclamation No. 2974, 3 C.F.R. 158 (1949-1953), *reprinted in* 66 Stat. ch. 31 (1952).

40. *See* H.R. REP. NO. 1540, *supra* note 9, at 47.

deemed to carry particular national security concerns.[41] By 1951, the number of secrecy orders grew to 2,395.[42]

The Department of Defense repeatedly requested that Congress grant renewed invention secrecy authority.[43] In response to these requests, Congress repealed the acts of 1917, 1940, 1941, and 1942, and passed the Invention Secrecy Act of 1951.[44] The 1951 Act contained several significant changes. A very important change was the creation of the semi-permanent secrecy order during peacetime:

> An invention shall not be ordered kept secret and the grant of a patent withheld for a period of more than one year. The Commissioner shall renew the order at the end thereof, or at the end of any renewal period, for additional periods of one year upon notification by the head of the department or the chief officer of that agency who caused the order to be issued that an affirmative determination has been made that the national interest continues so to require.[45]

Under the new doctrine, a year-long secrecy order, capable of indefinite renewal so long as the national interest requires, can be issued during peacetime.

The 1951 Act also effected several other important changes. First, the Act set out specific terms for invention secrecy during times of war or national emergency. During wartime, a secrecy order "shall remain in effect for the duration of hostilities and one year following cessation of hostilities."[46] During a declared state of national emergency, a secrecy order "shall remain in effect for the duration of the national emergency and six months thereafter."[47] The national emergency provision was especially important because in 1950, President Truman declared a national emergency that would last until March 1979.[48] Therefore, it was not until 1979 that the Invention Secrecy Act operated under its peacetime provisions. Second, the 1951 Act expanded inventors' rights by giving them the right to appeal a secrecy order to the Secretary of Commerce[49] and by reducing the foreign filing restrictions contained in the previous acts. The Act prohibited inventors from filing outside the

---

41. *See id.* at 47-48.

42. *See id.* at 47.

43. *See id.* at 1.

44. Invention Secrecy Act of 1951, ch. 4 §§10 & 11, 66 Stat. 3 (1952) (codified as amended at 35 U.S.C. §§ 181-188 (1994)).

45. 35 U.S.C. § 181 para. 4 (1994).

46. *Id.*

47. *Id.*

48. *See supra* note 5.

49. *See* 35 U.S.C. § 181 para. 3 (1994).

United States within six months of filing in the U. S. without a license from the Secretary of Commerce.[50]

Perhaps the most significant change with respect to inventors' rights was a revision of the compensation provisions. The new provisions stated in relevant part that:

> An applicant ... whose patent is withheld as herein provided, shall have the right ... to apply to the head of any department or agency who caused the order to be issued *for compensation for the damage caused by the order of secrecy and/or for the use of the invention by the Government* ....[51]

Under the revised statute, inventors could receive compensation for the damage caused by the order of secrecy, and did not have to first tender their inventions to the government as the previous acts required.[52] Through these new provisions, the Department of Defense gave expanded rights to inventors in order to extend the invention secrecy doctrine.

D. Evaluation of Peacetime Invention Secrecy as Legislative Policy

This part analyzes whether there is sufficient justification for peacetime invention secrecy. This part also examines the degree to which the Act conflicts with the principles of the patent system. Throughout the analysis, particular attention will be placed on invention secrecy hearings that took place in 1980 before the House Committee on Government Operations.[53] These hearings and the subsequent report entitled "The Government's Classification of Private Ideas,"[54] are significant because they came immediately after the national emergency was lifted in 1979, when peacetime invention secrecy took effect for the first time.[55] Several arguments made therein raise significant questions regarding the legitimacy of peacetime invention secrecy policy. This part concludes with a proposal for revising invention secrecy policy to better balance the competing goals of national security and patent right interests.

---

50. *See* 35 U.S.C. § 184 (1994).

51. 35 U.S.C. § 183 (1994) (emphasis added).

52. *See* S. REP. NO. 1001, H.R. REP. NO. 1028, *supra* note 38, *reprinted in* 1952 U.S.C.C.A.N. at 1323; *see also* H.R. REP. NO. 1540, *supra* note 9.

53. *The Government's Classification of Private Ideas: Hearings Before a Subcomm. of the House Comm. on Government Operations*, 96th Cong., 2d Sess. (1980) [hereinafter *Hearings*].

54. H.R. REP. NO. 1540, *supra* note 9.

55. In 1979, government agencies were required to determine affirmatively for each patent application subject to secrecy order that secrecy was still required. The review of these patent applications resulted in the renewal of 3,300 orders. *See Hearings*, *supra* note 53, at 19.

1. *THE NATIONAL SECURITY INTEREST DURING PEACETIME*

National security was Congress' primary reason for enacting invention secrecy legislation. "Inventions useful in war are made and developed during times of peace and it is important to prevent knowledge of such inventions being disclosed during times of peace as well as times of war."[56] The legitimacy of this rationale is weakened, however, by the fact that when the legislation was created, Congress was not actually legislating in anticipation of peace, but in fear of war.

a. Congress Has Failed to Give Modern Justification for Peacetime Secrecy Orders

The most significant problem with Congress' rationale for peacetime invention secrecy derives from the fact that the 1951 legislation was based on conditions different from those existing today. When the Invention Secrecy Act was created, the country was just beginning to recover from World War II, and the threat of war remained a significant concern.[57] The Department of Defense firmly believed in the need for peacetime invention secrecy due to continued threats to the national security, specifically, the Soviet Union's testing of an atomic bomb and the Korean War.[58] These circumstances influenced Congress when it decided to enact invention secrecy for peacetime.[59]

Although the United States continues to be threatened by terrorist attacks and possible outbreaks of war, no *concrete and continuous* threat analogous to that which weighed heavily on the minds of the 1950s' legislators exists today. These outdated motivations do not justify current issuance of secrecy orders. This is not to say that the rationale stated in the legislative history of the 1951 Act—that inventions useful in war are made during times of peace—no longer has any bearing on the justification of secrecy orders.[60] However, for Congress to continue to apply this rationale, it should make some minimal findings on modern

---

56. S. REP. NO. 1001, H.R. REP. NO. 1028, *supra* note 38, *reprinted in* 1952 U.S.C.C.A.N. at 1323-24.

57. *See id.* at 1322.

58. *See* H.R. REP. NO. 1540, *supra* note 53, at 50; *see also supra* text accompanying notes 37-39.

59. Congress relied extensively on the input of the Department of Defense in deciding to create the Invention Secrecy Act. *See id.* at 1, 50, 55; S. REP. NO. 1001, H.R. REP. NO. 1028, *supra* note 38, *reprinted in* 1952 U.S.C.C.A.N. at 1323-24.

60. For an analysis of the sufficiency of the old rationale with respect to modern secrecy orders, see discussion *infra* Part II.D.1.

circumstances that necessitate such a policy. Absent such findings, peacetime invention secrecy operates without legitimacy.

Furthermore, when Congress created the peacetime secrecy order, it was not actually legislating in anticipation of peace. The 1980 Report states: "Congress never set down a rationale for invention secrecy in peacetime. It avoided that issue in legislating the Invention Secrecy Act of 1951 by granting secrecy orders a lifetime six months beyond the duration of President Truman's December 1950 proclamation of national emergency."[61] Congress failed to justify peacetime invention secrecy because all secrecy orders that were to be issued under the new act were authorized under the national emergency provision. For this reason, the rationale offered by Congress in the legislative history of the Act only supports the need for a national emergency provision, and not a peacetime provision. When the national emergency ended twenty-seven years after the Act was created, it was all but forgotten that Congress had never truly justified peacetime orders. Congress today should "[m]ake the necessary findings and declaration of public policy that would justify the exercise of invention secrecy powers in peacetime."[62]

b. Peacetime Invention Secrecy May Still Be Necessary

The wartime rationale still might provide useful arguments supporting peacetime invention secrecy. Peacetime publication of certain inventions could be used by enemies against the United States during times of war. Some inventions may be so dangerous that their mere disclosure will create a threat to national security. For instance, the disclosure of the workings of a bomb may allow a terrorist access to information necessary for an attack on the United States. Even though the original rationale for peacetime secrecy orders may have been based on outdated national security concerns, the substance of the rationale provides a compelling argument for why peacetime invention secrecy continues to be necessary. The issuance of secrecy orders not only prevents foreign adversaries from obtaining important military technology, but it also allows the United States government to use the secret technology to its own advantage should war actually occur. Therefore, to protect the country against future national security risks, invention secrecy remains necessary even during times of peace.

---

61. *See* H.R. REP. NO. 1540, *supra* note 9, at 3.

62. *Id.* at 32.

2. *THE CONFLICTS BETWEEN INVENTION SECRECY AND THE PATENT SYSTEM*

The importance of protecting the rights of the inventor, like protecting national security, has been recognized since the beginning of invention secrecy.[63] The Constitution also recognizes the importance of inventors by empowering Congress to create patent laws.[64] However, inventors' rights have never received the same degree of protection as national security interests under the Invention Secrecy Act. Invention secrecy "conflicts with the principles of the patent system."[65] This necessitates an examination of the effect that secrecy orders place on inventors' rights and the overall goals of the patent system.

The 1980 Report identifies the primary dilemma faced by inventors: "Invention secrecy ... ensnares the inventors who work outside of the classified information community. It gives these nonmember inventors the choice of presenting their discoveries to the public without ownership protection, or of trying to obtain a patent and thereby risking Government confiscation of their ideas."[66] The private inventor easily could have published his invention in academic journals without fear of a secrecy order.[67] The inventor also could have pursued trade secret protection in lieu of seeking a patent. By applying for a patent instead, the inventor risks receiving a secrecy order which not only prevents him or her from obtaining a patent, but also prohibits all publication or disclosure relating to the invention.

Despite these problems, the national security interest at the heart of invention secrecy policy should not be completely compromised for the sake of inventors' rights. If the disclosure of an invention results in a breach of national security, then the issuance of secrecy orders benefits national security and such a benefit outweighs the costs imposed on inventors. Moreover, the imposition of secrecy orders on private inventors is not that severe. For example, "secrecy orders in many cases [create] an extension of the patent monopoly for ... conceivably ... substantial periods beyond the ... term of the issued patent."[68] Also, inventors receive some compensation.[69] Thus, as a matter of legislative policy, the burdens placed on inventors are justified.[70]

---

63. *See* discussion *supra* Part I.A.

64. *See* U.S. CONST. art. I, § 8, cl. 8.

65. *See* H.R. REP. NO. 1540, *supra* note 9, at 1.

66. *Id.* at 2.

67. *See id.* at 21.

68. *See* H.R. REP. NO. 1540, *supra* note 9.

69. *See* 35 U.S.C. § 183 (1994). When the peacetime provisions were introduced, many of the rules became more sympathetic to inventors by creating the right to appeal a secrecy

However, the problem remains that invention secrecy inherently conflicts with the interests of the patent system. Patents are not only designed to benefit inventors, but also are intended to benefit the public as well. "[A patent] is the reward stipulated for the advantages derived by the public for the exertions of the individual, and is intended as a stimulus to those exertions."[71] The suppression of inventions temporarily deprives the public of information that it would otherwise have obtained sooner. Inventions with purely military applications may not serve the public interest because the government is the only party with a need for these inventions.[72] However, because invention secrecy no longer affects only military applications,[73] a secrecy order may deprive the public of inventions with useful commercial applications. In addition, the mere threat of a secrecy order may discourage the inventor from filing for a patent, despite the promise of compensation, thereby frustrating the constitutional goal of promoting the useful arts.[74] The value to the public lies not only in the actual use of the invention but in the dynamic scientific knowledge stimulated by studying the invention. Consequently, the secrecy order impedes the ability of the patent system to stimulate new and useful inventions.

3. *RESOLVING THE CONFLICT*

The need to promote the patent interest directly conflicts with the need to protect national security during peacetime. Peacetime invention secrecy can be justified because, in general, the public is better served by legislation that protects the national security than legislation that protects the patent interest. However, because the national security interest is less compelling during peacetime, certain limitations should be placed on the legislation in order to provide the public with the most patent protection possible.

---

order and loosening the restrictions on foreign filings and receiving compensation. *See supra* Part I.C.

70. This is not to say that all legislation restricting inventors' rights in the name of national security can be justified. While the mere imposition of a secrecy order may be justifiable, the procedures used by the government must not be unduly burdensome. For a discussion of whether the mechanics of invention secrecy do in fact minimize the burdens on inventors, see *infra* Part III.

71. Grant v. Raymond, 31 U.S. (6 Pet.) 218, 242 (1832).

72. *See* H.R. REP. NO. 1540, *supra* note 9, at 7.

73. *See* discussion *supra* Part I.D.

74. *See* U.S. CONST. art. I, § 8, cl. 8.

The appropriate balance can be obtained by limiting peacetime invention secrecy only to inventions posing an imminent and likely threat to national security.[75]

> From its inception in 1917, invention secrecy was premised on the *fact or imminent* prospect of war. The Invention Secrecy Act of 1951 extended it in the expectation of a formal end to World War II hostilities, which would have unveiled existing secrecy orders, and the Korean conflict, which implied a need for new ones. Now, invention secrecy thrives on the presumption that war is not merely possible, but *likely*.[76]

Historically, Congress has required a credible threat to national security—such as the *actual* threat or presence of war—to invoke the protections of a secrecy order.[77] A modern application of the statute should recognize that threats to national security can exist even in the absence of a threat of war. Therefore, by requiring an imminent and likely breach of national security for the issuance of a secrecy order, Congress can ensure that secrecy orders will be imposed only when absolutely necessary.

The Invention Secrecy Act as currently enacted lacks any such limitations. The Act only requires that "the publication or disclosure of the invention by the granting of a patent ... [would] be detrimental to the national security."[78] Because the actual decision to issue a secrecy order lies in the discretion of the interested government agency, these agencies must be guided by precisely written legislation in order to correctly issue secrecy orders.[79] Congress should revise the statute to require a finding of an imminent and likely threat to security as a condition precedent.

III. THE MECHANICS OF INVENTION SECRECY

This part explores the decision-making process involved in imposing a secrecy order and the role of the government agency. This part also examines the remedies available to inventors, with an emphasis on how the government determines compensation. The part concludes with a discussion of the fairness of placing such burdens on inventors.

---

75. Support for a similar approach may be found in the First Amendment cases dealing with prior restraints. In such cases, infringement upon a First Amendment right by a prior restraint can only be justified when the disclosure creates direct, immediate, and irreparable harm. *See* discussion *infra* Part III.B.2.

76. *See* H.R. REP. NO. 1540, *supra* note 9, at 3 (emphasis added).

77. *See* discussion *supra* Part I.A.-C.

78. 35 U.S.C. § 181 para. 2 (1997). For a discussion of the standards used in imposing secrecy orders, see discussion *infra* Part II.A.

79. *See* discussion *infra* Part II.A.

A. Procedure for Issuing a Secrecy Order

The government's procedure for imposing a secrecy order depends in large part on whether or not the government holds a property interest in the invention:

> The phrase 'property interest' is intended to include the ownership of all rights in the invention or to a lesser interest therein such as, for example, cases where the foreign rights are retained by the inventor, or where the Government is entitled only to the interest of one or more joint inventors, and not to the interest of all the joint inventors.[80]

The 1980 Hearings defined a government property interest as including "inventions made by government employees either as part of their normal duties or on their own behalf, on which patent applications have been filed by the government," and "inventions made by government contractors during performance of their contractual duties."[81] Under Executive Order 9424,[82] all government-owned or government-controlled interests in patent applications are required to be registered in the Patent Office's Government Register.[83] Therefore, by referring to the register, it is easy to determine whether a government property interest exists. When the appropriate property interest is found, the issuance of a secrecy order is governed by the first paragraph of section 181 of the statute. When the government has no such property interest, secrecy orders are issued pursuant to paragraphs two and three of section 181.

1. WHEN THE GOVERNMENT HAS A PROPERTY INTEREST

The decision to issue the secrecy order falls within the discretion of the government agency having the interest in the invention. The statute's only limitation on this discretion is the requirement that the agency find that publication or disclosure of a patent "might" create a threat to the national security:

> Whenever publication or disclosure by the grant of a patent on an invention in which the Government has a property interest might, in the opinion of the head of the interested Government agency, be detrimental to the national security, the Commissioner upon being so notified shall order that the invention be kept secret and shall

---

80. S. REP. NO. 1001, H.R. REP. NO. 1028, *supra* note 38, *reprinted in* 1952 U.S.C.C.A.N. at 1321.

81. *Hearings, supra* note 53, at 450 (prepared statement of the Armed Services Patent Advisory Board (ASPAB), Department of Defense).

82. 3 C.F.R. 303 (1943-1948).

83. *See* 37 C.F.R. § 7.1(a) (1996).

> withhold the grant of a patent therefor under the conditions set forth hereinafter.[84]

Whether the information contained in an application is classified in part determines when disclosure might create a national security threat. For in-house government research, where the government agency files the patent application itself, a secrecy order may issue only if the application is properly classified under the provisions of Executive Order No. 12,065.[85] Applications classified under the above provisions are those which "reasonably could be expected to cause identifiable damage to the national security."[86] The government agency files the classified patent application in the Patent Office with security markings, thereby notifying the Patent Office to handle the application in accordance with appropriate security requirements.[87] The Patent Office then waits for the filing agency to request imposition of a secrecy order.[88]

Applications relating to contractor research also receive secrecy orders if they contain classified information. If the contract is classified, the contractor must give a copy of the application to the government agency after filing the patent application.[89] Then the agency obtains the secrecy order.[90] Contractor patent applications containing classified material must contain security markings as in the case of government-filed applications. For contractor applications that do not contain classified material, a secrecy order does not issue unless the Patent Office feels that another government agency may be interested in the subject matter. If this is the case, the application is subject to the screening process used for inventions in which the government has no property interest.[91]

---

84. 35 U.S.C. § 181 para. 1 (1994).

85. *See Hearings, supra* note 53, at 451 (prepared statement of the Armed Services Patent Advisory Board (ASPAB), Department of Defense); *id.* at 740 (Department of the Army, Office of the Judge Advocate General, Memorandum for the Record, Subject: Categories of secrecy order cases and related claims (Feb. 27. 1980)).

86. Exec. Order No. 12,065, 43 Fed. Reg. 28,949 (Section 1-104) (1978).

87. *See Hearings, supra* note 53, at 740 (Department of the Army, Office of the Judge Advocate General, Memorandum for the Record, Subject: Categories of secrecy order cases and related claims (Feb. 27. 1980)).

88. *See id.*

89. *See id.*

90. *See id.* at 741. In some cases the Patent Office may refer the application to the defense agencies before being notified to issue the secrecy order. In these cases the Patent Office will send an abbreviated disclosure to the agencies, calling the agency's attention to the existence of an application relating to a particular contract. *See id.*

91. *See* discussion *infra* Part III.B.2.a.

Once the agency decides to issue a secrecy order, it requests the Commissioner of Patents to impose the order.[92] For defense agencies, this request is made by the Armed Services Patent Advisory Board (ASPAB), which acts as a clearinghouse for military patent applications.[93] Once ASPAB makes the request on behalf of the interested agency, the Commissioner of Patents must issue the secrecy order.[94]

2. *WHEN THE GOVERNMENT DOES NOT HAVE A PROPERTY INTEREST*

With respect to those inventions in which the government does not have a property interest, the 1951 Act states:

> Whenever the publication or disclosure of an invention by the granting of a patent, in which the Government does not have a property interest, might, in the opinion of the Commissioner, be detrimental to the national security, he shall make the application for patent in which such invention is disclosed available for inspection to the Atomic Energy Commission, the Secretary of Defense, and the chief officer of any other department or agency of the Government designated by the President as a defense agency of the United States.
>
> ....
>
> If ... disclosure of the invention by the granting of a patent therefor would be detrimental to the national security, the Atomic Energy Commission, the Secretary of a Defense Department, or such other chief officer shall notify the Commissioner and the Commissioner shall order that the invention be kept secret and shall withhold the grant of a patent ... for such period as the national interest requires ....[95]

The Patent Office plays a much larger role in the determination of invention secrecy where the government has no property interest in the invention. The decision to impose a secrecy order is a two-stage process, with the first stage taking place in the Patent Office and the second stage taking place in a particular government defense agency.

---

92. *See* Manual of Patent Examining Procedure § 115 (6th ed. 1996), at 100-09 [hereinafter MPEP] ("[f]or those applications in which the Government has a property interest, responsibility for notifying the Commissioner of the need for a Secrecy Order resides with the agency having that interest").

93. *See Hearings, supra* note 53, at 451 (prepared statement of the Armed Services Patent Advisory Board (ASPAB), Department of Defense, before the Government Information and Individual Rights Subcommittee of the House Committee on Government Operations).

94. *See* 37 C.F.R. § 5.2(a) (1996).

95. 35 U.S.C. § 181 para. 2, 3 (1994).

a. Patent Office Procedure

All patent applications begin their prosecution in the Licensing and Review Branch of the Special Laws Administration Group,[96] also known as "the secret group" or "Group 220."[97] Here examiners categorize all applications as chemical, electrical, or mechanical inventions.[98] Then, examiners with appropriate security clearances and technical backgrounds inspect each patent application to determine if it contains "subject matter the disclosure of which might impact the national security ...."[99] Examiners screen applications not only for claimed subject matter but also for other matters, such as the incidental description of the use of the invention in the specification.[100] To aid in the determination, ASPAB, the Department of Energy, and the National Aeronautics and Space Administration (NASA) provide the Patent Office with guidelines.[101]

By far, the most important of these agency guidelines is ASPAB's Patent Security Category Review List.[102] The ASPAB List consists of twenty-one categories of inventions, each containing specific items which are identified as being of current security interest to different defense agencies. The agencies identified in the ASPAB list are the Army, the Navy, the Air Force, the Atomic Energy Commission, NASA, and the National Security Agency.[103] Items contained on the list include military devices as well as items with commercial applications.[104] If subject matter in a patent application corresponds to an item on the list, the Patent Office informs the agency, which then can view the application.[105] The Department of Energy uses a similar type of list to aid in the Patent

---

96. *See Hearings, supra* note 53, at 2 (statement of Rene D. Tegtmeyer, Assistant Commissioner for Patents, Patent and Trademark Office, Department of Commerce).

97. *See* H.R. REP. NO. 1540, *supra* note 9, at 17.

98. *See Hearings, supra* note 53, at 2-3 (statement of Rene D. Tegtmeyer, Assistant Commissioner for Patents, Patent and Trademark Office, Department of Commerce).

99. MPEP § 115, at 100-10 (6th ed. 1996).

100. *See id.*

101. *See id.*

102. ARMED SERVICES PATENT ADVISORY BOARD, PATENT SECURITY CATEGORY REVIEW LIST (1971) [hereinafter ASPAB LIST]. The list was declassified in 1994 at the Freedom of Information Act request of Michael Ravnitzky. *See* Aftergood, *supra* note 8. About 3% of all patent applications fall into areas identified by the list. *See id.*

103. *See* ASPAB LIST.

104. *See id.*

105. *See Hearings, supra* note 53, at 12 (statement of Rene D. Tegtmeyer, Assistant Commissioner for Patents, Patent and Trademark Office, Department of Commerce).

Office's examination. The Department of Energy Category Guide List relies on sections 151(c) and 152 of the Atomic Energy Act of 1954.[106]

b. Defense Agency Procedure

The Patent Office forwards applications containing subject matter deemed a possible national security threat to the interested defense agencies. The inspection of applications by the agencies must be performed "only by responsible representatives authorized by the agency to review applications."[107] An agency's decision to impose a secrecy order is governed by the statutory standard that "the publication or disclosure of the invention by the granting of a patent ... *would* be detrimental to the national security."[108] Because this is the only statutory restriction on the agency's discretion, the agency essentially is left to its own devices to make its decisions. Currently, military agencies impose secrecy orders on about five to ten percent of the applications they review.[109] After making its decision, an agency using the ASPAB List simply informs ASPAB that it wants a secrecy order, and ASPAB instructs the Commissioner of Patents and Trademarks to issue the order.[110] Other nonmilitary agencies take it upon themselves to notify the Patent Office.[111]

B. The Effect of a Secrecy Order

All secrecy order cases are examined for patentability like any other application.[112] When an application subject to a secrecy order is poised for allowance, the Patent Office issues a notice of allowability (Form D-10).[113] This ends the prosecution of the application until the secrecy order is rescinded.[114] Thus, even if the application is found to contain patentable subject matter, no patent will issue on the application until the government rescinds the secrecy order. Moreover, an interference will not be declared where one or more of the conflicting cases is classified or

106. Codified at 42 U.S.C. §§ 2181-82 (1994).

107. 37 C.F.R. § 5.1 (1996).

108. 35 U.S.C. § 181 para. 3 (1994) (emphasis added).

109. *See* Aftergood, *supra* note 8.

110. *See Hearings*, *supra* note 53, at 451 (prepared statement of the Armed Services Patent Advisory Board (ASPAB), Department of Defense).

111. *See* MPEP § 115, at 100-09.

112. *See id.* § 130, at 100-14.

113. *See id.*

114. *See id.*

under secrecy order.[115] In the case of a final rejection, an appeal before the Board of Patent Appeals and Interference will not be heard until the secrecy order is lifted, unless specifically ordered otherwise by the Commissioner.[116]

A secrecy order restricts disclosure of the invention or dissemination of information in the patent application.[117] A secrecy order prevents the Commissioner of Patents from granting a patent on the application until the order is lifted.[118] The Commissioner also orders "that the invention be kept secret"[119] by sending the inventor a notice of the order. The notice instructs the inventor that "the subject matter or any material information relevant to this application, including unpublished details of the invention, shall not be published or disclosed to any person not aware of the invention prior to the date of this order, including any employee of the principals."[120]

The particular type of secrecy order that issues determines the extent of the restraint on publication or disclosure. There are three possibilities. First, a "Secrecy Order and Permit for Foreign Filing in Certain Countries" applies to patent applications containing technical data whose export is controlled by guidelines established by the Department of Defense.[121] These applications can be filed in certain countries with which the United States has reciprocal security agreements.[122] The second type of secrecy order, known as a "Secrecy Order and Permit for Disclosing Classified Information," is appropriate where the patent application contains technical information that is properly classified or classifiable under a security guideline and where the owner has a security agreement with the Department of Defense.[123] These secrecy orders treat the information contained in the application like other classified material.[124] The third type of secrecy order is known

---

115. *See id.*

116. *See id.*

117. *See Hearings, supra* note 53, at 176 (letter from Sidney A. Diamond, Commissioner for Patents and Trademarks, U.S. Patent and Trademark Office, to Richard Preyer, Chairman, Government Information and Individual Rights Subcommittee of the Committee on Government Operations (May 3, 1980)).

118. *See* 35 U.S.C. § 181 para. 1, 3 (1994).

119. *Id.*

120. *Hearings, supra* note 53, at 128 (notice of secrecy order from Director, Special Laws Administration Group, Patent and Trademark Office).

121. *See* MPEP § 120, at 100-09.

122. *See id.* at 100-10.

123. *See id.* at 100-09.

124. *See id.* at 100-10.

simply as a "Secrecy Order" and is used where the technical information is properly classifiable and the patent application owner does not have a Department of Defense security agreement.[125] Secrecy orders generally apply whenever the other types of orders do not, and include orders issued by directors of agencies other than the Department of Defense.[126] These orders prohibit disclosure to anyone without express written consent from the Commissioner.[127]

All secrecy orders expose an inventor to the penalty provisions for violation of the order. Should the inventor publish or disclose an invention subject to a secrecy order, or file for a patent on that invention in a foreign country without the consent of the Commissioner, the invention shall be held to be abandoned.[128] An inventor who, without due authorization, willfully publishes or discloses the invention, or who willfully files a foreign patent application, "shall, upon conviction, be fined not more than $10,000 or imprisoned for not more than two years, or both."[129]

A peacetime secrecy order lasts for only one year.[130] The government agency sponsoring the order may renew the order for additional periods of up to one year upon notice to the Patent Office "that an affirmative determination has been made that the national interest continues so to require."[131] Thus, a secrecy order may continue indefinitely. A secrecy order ends either when it is not renewed, or when the Commissioner is notified by the sponsoring agency "that the publication or disclosure of the invention is no longer deemed detrimental to the national security."[132] With respect to secrecy orders authorized by ASPAB in particular, the patent application must be circulated among ASPAB members for their individual consideration of whether the secrecy order should be rescinded.[133] Once a secrecy order is rescinded, the Patent Office issues a notice of allowance to the applicant.[134]

---

125. *See id.* at 100-09.

126. *See id.* at 100-10.

127. *See id.* at 100-09. However, these secrecy orders often contain a "Permit A" which relaxes the disclosure restrictions. *See id.* at 100-10.

128. *See* 35 U.S.C. § 182 (1994).

129. 35 U.S.C. § 186 (1994).

130. *See* 35 U.S.C. § 181 para. 4 (1994).

131. *Id.*

132. *Id.*

133. *See Hearings, supra* note 53, at 458 (prepared statement of the Armed Services Patent Advisory Board (ASPAB), Department of Defense).

134. *See* 37 C.F.R. § 5.3(c) (1996).

C. Inventors' Rights and Remedies Under the Act

An inventor under a secrecy order has two options. First, the inventor may attempt to contest the order as either erroneous or overly broad. Second, the inventor may apply for compensation under section 183 of the Act. The procedures used in pursuing these two options are discussed in detail below. Normally, only the private inventor will be interested in pursuing these options. For inventors whose inventions are controlled by the government, challenges to secrecy orders are rare because the government itself either owns the invention or is authorized to prevent publication of the classified information.[135] Further, the Invention Secrecy Act's compensation provisions prohibit compensation to persons "who, while in the full-time employment or service of the United States, discovered, invented, or developed the invention on which the claim is based."[136] Accordingly, the options discussed below primarily benefit the private inventor with no government affiliation.

1. *CONTESTING THE ORDER OF SECRECY*

An inventor who believes that an erroneous secrecy order has been issued has several options to seek rescission of the order. The Manual of Patent Examining Procedure recommends that the applicant directly contact the agency sponsoring the secrecy order to discuss what would render the secrecy order unnecessary.[137] Alternatively, the applicant may follow a more formalized procedure and petition the Commissioner of Patents for a rescission.[138] Such a petition "must recite any and all facts that purport to render the order ineffectual or futile if this is the basis of the petition."[139] Finally, the most formal method for the applicant is to appeal to the Secretary of Commerce to rescind the secrecy order.[140] An appeal cannot be made until a petition for rescission has been made and denied by the Commissioner, and must be taken within sixty days of the denial.[141] The Secretary of Commerce, or officers designated by the Secretary, hear and decide the appeal.[142]

---

135. *See Hearings, supra* note 53, at 741 (Department of the Army, Office of the Judge Advocate General, Memorandum for the Record, Subject: Categories of secrecy order cases and related claims (Feb. 27, 1980)).

136. 35 U.S.C. § 183 (1994).

137. *See* MPEP § 120, at 100-11 (6th ed. 1996).

138. *See* 37 C.F.R. § 5.4 (1996).

139. *Id.*

140. *See* 35 U.S.C. § 181 para. 3 (1994).

141. *See* 37 C.F.R. § 5.8 (1996).

142. *See id.*

Besides a rescission, the applicant may seek a permit to disclose or modify the order. This procedure is necessary to make disclosure of the contents of the application to certain categories of individuals or to file patent applications in foreign countries.[143] The government agencies review these requests and are likely to grant them where the country in which the applicant desires a foreign filing is obligated by agreement to maintain secrecy.[144]

2. *COMPENSATION*

Compensation is the primary remedy for an inventor. The Invention Secrecy Act gives an inventor the right to "compensation for the damage caused by the order of secrecy and/or for the use of the invention by the Government, resulting from his disclosure."[145] For an inventor to qualify for compensation, the patent application must, except for the secrecy order, be in condition for allowance.[146]

There are two ways in which an inventor may seek compensation under the Act. First, the claimant may immediately apply to the agency causing the order for a settlement agreement.[147] Such claims must be filed after the date of first use of the invention by the government.[148] If a settlement agreement cannot be reached, the head of the agency may award the applicant a sum not exceeding 75% of the amount the head of the agency deems fair.[149] The claimant then has the right to bring suit against the United States in the United States Court of Federal Claims or in the district court in which the claimant resides for an amount which when added to the settlement award will result in just compensation.[150] Even if no settlement amount is given, the claimant still may bring suit.[151] The second method by which an inventor can obtain compensation is to wait for the secrecy order to expire and the patent to issue. Provided that the inventor did not apply for compensation under the first set of procedures described above, the inventor then can bring suit in the United States Court of Federal Claims.[152]

---

143. *See* 37 C.F.R. § 5.5 (1996).

144. *See Hearings, supra* note 53, at 3 (prepared statement of the Armed Services Patent Advisory Board (ASPAB), Department of Defense).

145. 35 U.S.C. § 183 (1994)

146. *See id.*

147. *See id.*

148. *See id.*

149. *See id.*

150. *See id.*

151. *See* Robinson v. United States, 236 F.2d 24, 26 (2d Cir. 1956).

152. *See* 35 U.S.C. § 183 (1994).

Under either of these methods, section 183 specifies that the inventor shall be entitled to "just compensation." Court determinations of what constitutes "just compensation" under section 183 are rare.[153] In *Constant v. United States (Constant I)*, the patentee Constant sought compensation in the Court of Claims for damages resulting from the imposition of a secrecy order on his invention for a method for encoding radar signals.[154] The court stated:

> The Government erroneously urges that the only rights section 183 protects are the rights accorded by 35 U.S.C. § 154 to the patent grant—"the right to exclude others from making, using or selling the invention throughout the United States" .... The statute does not say that its provisions are confined to that precise area, and its language as to compensation has much wider and more general phrasing. The core of the legislation is recovery of all "damage caused by the order of secrecy and/or for the use of the invention by the Government."[155]

The court recognized that even though the statute employs the term "just compensation," compensation should not be limited to an eminent domain theory.[156] *Constant I* therefore encourages liberal awards of compensation under section 183.

For claims of damages based solely on the order of secrecy, courts consider a wide range of factors. In *Constant I*, the patentee only sought compensation for damages caused by the issuance of the secrecy order and not for any use by the government. This secrecy order never actually affected the issuance of Constant's patent, because it was issued and removed before the patent application was even determined to be allowable.[157] Constant alleged that because of the secrecy order: (1) he was unable to obtain loans necessary for the development of his invention, (2) he lost prospective users and licensees for his invention because he could not demonstrate that his invention represented a superior technology in the field, and (3) he expended substantial

---

153. Compensation for government infringement of patents may also be relevant to this determination. An inventor may sue the government for infringement of a patent under the provisions of 28 U.S.C. § 1498 for reasonable and entire compensation. *See* 28 U.S.C. § 1498(a) (1994). Section 183 of the Invention Secrecy Act implicitly recognizes the similarities in the two types of compensation by stating: "In a suit under the provisions of this section the United States may avail itself of all defenses it may plead in an action under section 1498 of title 28." 35 U.S.C. § 183 (1994). For a discussion of damages for government infringement, see *infra* text accompanying notes 371-83.

154. Constant v. United States (Constant I) 617 F.2d 239, 239-44 (Ct. Cl. 1980).

155. *Id.* at 243 n.10.

156. *See id.* at 242.

157. *See id.*

attorneys' fees in attempts to have the secrecy order rescinded.[158] The court found that all of these damages could be compensable, provided that the loss is supported by "real concrete evidence of damage."[159] Because Constant failed to provide such evidence, however, no compensation was granted.[160]

For claims based on use by the government, courts have awarded compensation on a reasonable royalty basis.[161]

> In the determination of a reasonable royalty rate for the computation of a fair award of damages, such factors as the limited marketability of the product (thus requiring that the entire compensation be obtained from the Government) must be equated with assumption of risk in providing capital for the production of the invention and other similar variables (which factors would tend to depress the allowable royalty rate).[162]

Thus, like the situation where damages are claimed only for the orders of secrecy, compensation for use by the government is based on a wide variety of factors that are often considered in conventional patent infringement cases. In both situations, claimants have the potential to recover substantial compensation awards. However, if damages are too speculative in either situation, courts may be reluctant to grant any compensation whatsoever.[163]

D. Fairness Analysis

Having examined the mechanics of the Invention Secrecy Act, the analysis now turns to a discussion of whether the procedures used by the government are sufficiently fair to minimize the burdens placed on inventors.[164] The issue is addressed in three parts: (1) whether inventors are properly treated when issued a secrecy order, (2) whether the specific effects of a secrecy order are fair, and (3) whether inventors are given

---

158. *See id.* at 244.

159. *Id.*

160. *See* Constant v. United States (*Constant II*), 1 Cl. Ct. 600, 609 (1982), *aff'd*, 714 F.2d 162 (Fed. Cir. 1983).

161. *See* Farrand Optical Co. v. United States, 197 F. Supp. 756, 773 (S.D.N.Y. 1961), *modified*, 325 U.S. 328 (2d Cir. 1963).

162. *Id.* at 777 n.2.

163. *See Constant II*, 1 Cl. Ct. at 609. A further discussion of the compensation issue is contained *infra* Part II.D.3. That part examines whether the requirement of proof of actual damages is fair to inventors whose inventions are suppressed by secrecy order.

164. The discussion in this part assumes that the provisions of the Act are constitutional. For a discussion of the Act's potential constitutional problems, see discussion *infra* Part III.

sufficient options to counter and obtain remedies for the issuance of a secrecy order.

1. THE DECISION TO ISSUE A SECRECY ORDER

The fairness of imposing a secrecy order on an invention depends on whether the decision-making process results in an erroneous secrecy order. In other words, has the appropriate agency applied the proper discretion in determining whether national interests require a secrecy order? When the government has a property interest in the invention, the only statutory limit on an agency's decisionmaking is the requirement that publication or disclosure by grant of the patent "*might* ... be detrimental to the national security."[165] The use of the term "might" is significant because it reflects a relatively low standard for the agency to satisfy in order to find a danger to national security. The statute essentially places no limitations on the discretion to issue the order. While this lack of congressional guidance may seem unfair to the individual inventor, it is important to remember that in this situation, the government agency holds the property interest in the invention, not the individual inventor. The inventor has virtually no expectation of any individual patent rights.

The more problematic situation arises in the application of secrecy orders to inventions in which the government does not have property interests. The statute imposes a higher standard, such that these inventions can be subject to secrecy order only if the government agency determines that their publication or disclosure "*would* be detrimental to the national security."[166] Since peacetime secrecy orders were first issued, the number of secrecy orders, especially with respect to private inventors, has steadily increased while the threat of war has arguably declined.[167] If the number of secrecy orders issued to private inventors continues to increase during peacetime, it may be necessary to enact a more stringent statutory standard to ensure that orders are issued only when absolutely necessary. Rewording the statute to require that a likely and imminent threat to national security exists for the issuance of a secrecy order would suffice to increase protection for private inventors.[168]

---

165. 35 U.S.C. § 181 para. 1 (1994) (emphasis added).

166. 35 U.S.C. § 181 para. 3 (emphasis added).

167. *See supra* text accompanying notes 8-14; *see also Hearings, supra* note 53, at 453 (prepared statement of the Armed Services Patent Advisory Board (ASPAB), Department of Defense). In 1979, the Commissioner of Patents and Trademarks estimated that about 10-20% of secrecy orders were imposed on applications in which the government had no property interest. *See id.* at 455.

168. *See* discussion *supra* part II.A.

Congress should also prompt the agencies to establish more rigorous internal procedures to justify the issuance of all secrecy orders.

2. *THE EFFECTS OF A SECRECY ORDER*

Because invention secrecy reflects justifiable legislative policy,[169] its consequent preclusion of publication or disclosure of an invention does not in itself create an excessive burden on inventors. However, whether provisions regarding the duration of secrecy orders and penalties for their violation fail to minimize the burdens on inventors requires further analysis.

a. Secrecy Orders of Indefinite Duration

With respect to duration, the main problem is that a peacetime secrecy order, though one year in length, can be renewed indefinitely. A secrecy order of unspecified length creates tremendous burdens by depriving the inventor of opportunities to exploit the invention. In addition, an inventor whose invention is subjected to a secrecy order of indefinite duration receives very little notice as to when the invention might finally be available for commercialization.

In response, the government asserts that certain inventions create such a large threat to the national security that only a permanent secrecy order will suffice. The most common technology deemed to justify secrecy orders of extended duration is cryptology.[170] Several secrecy orders in effect during the 1980 Hearings covered cryptologic inventions made in the 1930s.[171] Moreover, the government argues that the compensation under the Act provides inventors with an adequate remedy.[172]

Imposing a higher standard on the determination of when national security requires invention secrecy would allow secrecy orders to issue when truly necessary and would limit the imposition of indefinite orders. This higher standard should be imposed on both the initial decision for a secrecy order and the subsequent review for renewal. The statute currently states that renewal should take place "upon notification by the head of the department or the chief officer of the agency who caused the order to be issued that an affirmative determination has been made that

---

169. *See* discussion *supra* part II.D.

170. For a general discussion of the problems of cryptology, see David Kahn, *Cryptology Goes Public*, FOREIGN AFF., Fall 1979, at 141; *see also* H.R. REP. NO. 1540, *supra* note 9, at 62-120.

171. *See* H.R. REP. NO. 1540, *supra* note 9, at 70.

172. For a discussion of whether the compensation provisions themselves are sufficient, *see* discussion *infra* Part III.D.3.

the national interest continues so to require."[173] The statute should be modified to emphasize that only inventions *continuing* to create a likely and imminent threat to the national security are eligible for renewal. By doing so, Congress can ensure that only those inventions with grave national security implications will remain under semi-permanent secrecy orders.

Second, Congress should consider the possibility of an additional type of secrecy order during peacetime. A *permanent* secrecy order may be appropriate to encompass those inventions that would always have national security implications, regardless of the presence of war. The benefit of a permanent secrecy order would be to inform inventors that they will never have any chance to exploit their invention, so they can try to sell it to the government or obtain compensation under the Act for the government's use. This proposal would also eliminate the administrative hassle of having to renew these secrecy orders every year. However, this type of secrecy order must only apply to those inventions having major national security implications.

b. Extension Period for Secrecy Orders

The GATT[174]-related changes to the patent law create another problem with the potentially infinite duration of a secrecy order. Before GATT, once a secrecy order was removed from an otherwise allowable patent application, a patent would issue with a seventeen-year term from date of issuance. Even though patent issuance would be delayed while the secrecy order was in effect, inventors still retained the full length of their patent term. Under the newly revised section 154 of Title 35, a patent's term now ends twenty years from the date of filing.[175] This creates the problem that a secrecy order effectively shortens the length of the patent term.

Congress resolved this problem by allowing an extension, whereby the term of the issued patent may be extended for the period of the delay, but in no circumstances for more than five years.[176] By limiting the extension term to only five years, inventors issued with secrecy orders lasting longer than five years effectively lose time to exploit their patents. The legislative history of this revision does not explain why Congress chose this five-year period.[177]

---

173. 35 U.S.C. § 181 para. 4 (1994).

174. General Agreement on Tariffs & Trade, Oct. 30, 1947, 61 Stat. A-11, T.I.A.S. 1700, 55 U.N.T.S. 194.

175. *See* 35 U.S.C. § 154(a)(2) (1994).

176. *See* 35 U.S.C. § 154(b)(1) (1994).

177. *See* S. REP. NO. 412, 103d Cong., 2d Sess. (1994).

The government may argue that although a secrecy order is in effect, the inventor still is able to plan for exploitation and further development of the invention during this time and therefore should not be unjustly enriched by a long extension period as well as the full term of the patent. This still does not address those inventors faced with secrecy orders that never end.[178] Nevertheless, because the legislative history offers no reasons for choosing a five-year period, Congress should either justify its use of only a five-year extension, or allow a further extension for secrecy orders lasting longer than five years.

c. The Penalty Provisions

A final point regarding the fairness of the effects of a secrecy order relates to the penalty provisions of the Act. The Act prohibits the violation of a secrecy order with provisions for abandonment, criminal fines, and possible imprisonment. An argument may be made that these provisions are too harsh and place an unjustifiable burden on inventors. Assuming these provisions are constitutional, they are appropriate and necessary for ensuring that secrecy orders will be obeyed, thereby fulfilling the legislative purpose of the Act.[179]

3. *THE FAIRNESS OF INVENTORS' REMEDIES*

The final consideration in the fairness analysis is whether an inventor possesses sufficient remedies once a secrecy order is imposed. Fairness in this context is measured by the ease with which an inventor can contest the imposition of an erroneous secrecy order or obtain compensation for the imposition of a legitimate order.

a. The Availability of Review for Secrecy Order Decisions

The Act provides several options for an inventor to contest a secrecy order. These include appealing the order to the agency authorizing the order, to the Patent Office, or to the Secretary of Commerce.[180] The ability to appeal to the Secretary of Commerce is important because it allows the inventor to have his claim heard by a party completely unrelated to the issuance of the order. Given this

178. At the time of the 1980 Hearings, the longest secrecy order still in effect was issued in 1942 on a 1940 application. *See* H.R. REP. NO. 1540, *supra* note 9, at 165.

179. For a discussion of the First Amendment problems associated with these provisions, *see* discussion *infra* Part III.

180. *See* discussion *supra* Part III.C.1.

abundance of options, an inventor seems to have sufficient avenues to contest unjustified secrecy orders.[181]

What an inventor lacks is the ability to obtain judicial review. The Administrative Procedure Act (APA) provides that "[a] person suffering legal wrong because of agency action, or adversely affected or aggrieved by agency action within the meaning of a relevant statute, is entitled to judicial review thereof."[182] Agency action generally must not be "arbitrary, capricious, an abuse of discretion, or otherwise not in accordance with law."[183] With respect to military agency decisions in particular, review "must be extremely deferential because of the confluence of the narrow scope of review under the APA and the military setting."[184] Judicial review is not available when the "agency action is committed to agency discretion by law."[185] In *Webster v. Doe*,[186] the Court held that a provision of the National Security Act of 1947, giving the Director of the Central Intelligence Agency the discretion to terminate any officer or employee when deemed necessary or advisable in the interests of the United States, precluded judicial review under the APA because the statute clearly showed Congress' intent to commit individual employee discharges to the Director's discretion.[187] Similarly, the fact that Congress has firmly placed the decision to impose secrecy orders in the hands of the relevant military agencies indicates that judicial review should not be available for these agency decisions. Accordingly, an inventor forced to comply with a secrecy order is left to pursue only those remedies clearly defined by the statute.[188]

b. The Availability of Compensation

The Invention Secrecy Act authorizes compensation for both "the damage caused by the order of secrecy and/or for the use of the invention by the Government, resulting from his disclosure," and provides several

---

181. A further discussion of the fairness of these procedures is contained in the procedural due process analysis, discussed *infra* Part III.C.1.

182. 5 U.S.C. § 702 (1994).

183. 5 U.S.C. § 706(2)(A) (1994).

184. Henry v. United States Dept. of Navy, 77 F.3d 271, 272 (8th Cir. 1996).

185. *See* 5 U.S.C. § 701(a)(2) (1994).

186. 486 U.S. 592 (1988).

187. *See id.* at 601.

188. Section 701(a)(2) of the APA does not automatically preclude judicial review of constitutional claims. The Court in *Webster* stated that Congress' intent to preclude judicial review of constitutional claims must be clear. *See id.* at 603. Absent such a clear intent, a petitioner may still bring claims challenging the constitutionality of agency action.

procedures by which the inventor obtains the statutory compensation.[189] In reality, awards for compensation have been exceedingly rare. The 1980 Report indicates that between 1945 and 1979, twenty-nine administrative claims for compensation were filed with the Department of Defense.[190] Of these twenty-nine, only nine claims led to the receipt of any amount of compensation.[191] For every thousand secrecy orders requested by ASPAB, only one claim for compensation is filed.[192] The 1980 Report points out that only those claimants "who are financially strong enough and persistent enough can collect eventually."[193]

Two factors explain the difficulty in obtaining compensation. First, the requirement of actual damages, as discussed in *Constant I*, poses a major barrier to inventors' compensation.[194] Constant maintained that this requirement would "have a 'chilling effect' upon inventorship in this country, and, further, will render § 183 meaningless by making it a cause of action without a remedy."[195] Because it is inherently difficult to determine the market value of an invention having commercial potential when disclosure is barred by secrecy order,[196] government agencies may be inclined to minimize the compensation awarded to an inventor for a secrecy order.

Second, ASPAB officials continue to construe the Act in such a way that awards for secrecy order damage claims are not appropriate "where the invention was suppressed but not used by the Government and the Government was the sole intended market for the invention."[197] This construction directly contradicts the language of the statute which authorizes compensation awards for the damage caused by the order, without requiring any government use.[198] Again, the main problem is that these agencies have too much discretion to refuse to grant compensation:

> Agencies can specify the form and content of an administrative claim, and may attempt to require by regulation nearly as much supporting evidence for the claim as they could expect to collect by

---

189. *See* 35 U.S.C. § 183 (1994); discussion *supra* Part III.C.2.

190. *See* H.R. REP. NO. 1540, *supra* note 9, at 6.

191. *See id.*

192. *See id.*

193. *Id.* at 8.

194. *See* Constant v. United States, 617 F.2d 239, 244 (Ct. Cl. 1980); Constant v. United States, 1 Cl. Ct. 600, 609 (1982), *aff'd*, 714 F.2d 162 (Fed. Cir. 1983).

195. *Supreme Court Asked to Review Intellectual Property Decisions*, 26 PATENT, TRADEMARK & COPYRIGHT J. 570 (1983).

196. *See* H.R. REP. NO. 1540, *supra* note 9, at 29.

197. *Id.* at 28.

198. *See* 35 U.S.C. § 183 (1994).

> rule of discovery in a court of law. Agencies have little or no incentive to settle a claim.[199]

Although a suit may be brought by an inventor in district court or the United States Court of Federal Claims,[200] it is unfair to make an inventor resort to this more expensive and complex avenue when the claim could be settled satisfactorily at a lower level.

To remedy these problems, agencies should be encouraged to be flexible in determining a compensation award under section 183. The 1980 Report points out that the Internal Revenue Service's engineering and valuation branch is able to appraise patents and other property for gift and inheritance tax purposes.[201] Perhaps a better approach to estimating just compensation when there is little proof of actual damages should come from the provisions of the Atomic Energy Act of 1954. Section 154 of the Atomic Energy Act states: "No patent shall hereafter be granted for any invention or discovery which is useful solely in the utilization of special nuclear material or atomic energy in an atomic weapon. Any patent granted for any such invention or discovery is revoked, and just compensation shall be made therefor."[202] This deprivation of a patent interest has a similar effect to the Invention Secrecy Act. Therefore, in practical terms, guidelines issued under the Atomic Energy Act for just compensation may be the most appropriate with respect to the present situation. To determine just compensation, the Atomic Energy Act establishes the following standards:

> (1) In determining a reasonable royalty fee ..., the Commission shall take into consideration
>
> (A) the advice of the Patent Compensation Board;
>
> (B) any defense, general or special, that might be pleaded by a defendant in an action for infringement;
>
> (C) the extent to which, if any, such patent was developed through federally financed research; and
>
> (D) the degree of utility, novelty, and importance of the invention or discovery, and may consider the cost to the owner of the patent of developing such invention or discovery or acquiring such patent.
>
> (2) In determining what constitutes just compensation as provided for in section 2181 of this title ... the Commission shall take into account the considerations set forth in paragraph (1) of this subsection and the actual use of such invention or discovery. Such

199. *See* H.R. REP. NO. 1540, *supra* note 9, at 7.

200. *See* 35 U.S.C. § 183 (1994).

201. *See* H.R. REP. NO. 1540, *supra* note 9, at 29 n.54.

202. 42 U.S.C. § 2181(a) (1994).

> compensation may be paid by the Commission in periodic payments or in a lump sum.[203]

Case law that reviews the granting of just compensation under these provisions emphasizes that so long as an inventor retains some rights in the invention, it is inappropriate to deny compensation completely.[204] Thus, any time the government does not hold a property interest in an invention, the inventor should recover compensation. Moreover, the Atomic Energy Act provisions do not require proof of actual damages for obtaining compensation. Rather, just compensation in some amount will always be due through consideration of the factors listed above.[205]

Incorporating these provisions of the Atomic Energy Act into the Invention Secrecy Act's compensation system creates a good solution to the problems discussed above. First, to determine a reasonable royalty for the imposition of a secrecy order, factors such as the degree of utility, novelty, and importance of the invention can be considered even without proof of actual damages. Congress should emphasize that compensation based on these factors can be granted even without use of the invention by the government. Second, to remedy the problem that government agencies receive too much discretion in making a determination of just compensation, the Invention Secrecy Act should follow the Atomic Energy Act's model and create a Patent Compensation Board.[206] The use of a separate committee to determine just compensation ensures a fair and unbiased compensation determination.

IV. CONSTITUTIONAL QUESTIONS

The 1980 Hearings raised First and Fifth Amendment questions relating to the Act.[207] A witness testifying for the Justice Department stated:

> [I]t is perfectly true that any flat prohibition on private speech raises an issue under the First Amendment, but we are dealing here with a prohibition, § 186, that has never been tested. There has never been a prosecution under § 186. We have no judicial opinion to guide us. In advance of litigation, undisciplined by facts, the expression of views on the First Amendment issues that might be presented by a prosecution under this statute would be difficult in any event and would be either self-serving or prejudicial from the

---

203. 42 U.S.C. § 2187(c) (1994).

204. *See* Hobbs v. United States, 376 F.2d 488, 493 (5th Cir. 1967) (holding that even though the government had obtained shop rights in an invention, the inventor still retained rights in the invention which entitled him to some amount of just compensation).

205. *See id.*

206. *See* 42 U.S.C. § 2187(a) (1994).

207. *See* H.R. REP. NO. 1540, *supra* note 9, at 27.

> standpoint of the Department's duty to enforce the statute. As regards the Fifth Amendment issues, I note simply that the statute provides both an administrative and a judicial remedy for damages caused by the secrecy procedure.[208]

This statement reflects an overall reluctance on the part of the government to confront the constitutional problems created by the Invention Secrecy Act. In fact, the constitutionality of the statute has never been tested.[209] This part thus analyzes whether the provisions of the Invention Secrecy Act are constitutional under the First and Fifth Amendments.

A. Constitutional Authority for Invention Secrecy

Before exploring the potential constitutional violations under the First and Fifth Amendments, it is useful to identify the Congress' constitutional authority for invention secrecy. There are three possibilities. First, the Constitution empowers Congress to "provide for the common Defense and general Welfare of the United States."[210] Because invention secrecy is primarily directed toward protecting the national security, invention secrecy legislation is clearly authorized under this power.

Second, Congress has the power "[t]o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries."[211] This provision, which authorizes the creation of a patent system, does not grant inventors an absolute patent right. Rather, it gives Congress extremely broad discretion to decide how the patent system should be formulated. Invention secrecy can therefore simply be viewed as a legitimate part of the patent system's operation.

Third, the power of eminent domain may authorize the government statute. The Court has long recognized that the federal government may constitutionally take private property,[212] provided that the taking is rationally related to a public purpose and just compensation is paid.[213]

---

208. *Hearings, supra* note 53, at 252 (testimony of H. Miles Foy, Senior Attorney-Adviser, Office of Legal Counsel).

209. *See* H.R. REP. NO. 1540, *supra* note 9, at 2; U.S. CONG., OFFICE OF TECHNOLOGY ASSESSMENT, *Science, Technology, and the First Amendment*, OTA-CIT-369 (Jan. 1988) 48 [hereinafter *OTA Report*].

210. U.S. CONST. art. I, § 8, cl. 1.

211. U.S. CONST. art. I, § 8, cl. 8.

212. *See* Kohl v. United States, 91 U.S. 367, 372 (1875).

213. *See* Hawaii Housing Auth. v. Midkiff, 467 U.S. 229, 241 (1984). Although the Constitution does not explicitly recognize the power of eminent domain, the Supreme Court has interpreted the Fifth Amendment Takings Clause to be an implicit recognition of the

Whether the imposition of a secrecy order meets these requirements is the subject of the Takings Clause analysis below.[214] If these requirements are met, Congress' invention secrecy legislation can be viewed as an appropriate exercise of the government's eminent domain power.

There are inherent limits to Congress' power. "[T]he Constitution requires that the powers of government 'must be so exercised as not, in attaining a permissible end, unduly to infringe' a constitutionally protected freedom."[215] Although Congress may have created invention secrecy under legitimate constitutional authority, problems may still arise under other constitutional provisions. The Court has stated: "When Congress' exercise of one of its enumerated powers clashes with those individual liberties protected by the Bill of Rights, it is our "delicate and difficult task" to determine whether the resulting restriction on freedom can be tolerated."[216] Because of the constitutional problems identified at the 1980 Hearings, the analysis that follows embarks on the "delicate and difficult task" of determining whether the liberties protected by the First and Fifth Amendments are unduly burdened by Congress' invention secrecy legislation.

B. First Amendment Analysis

Because it enables punishment for publication or disclosure of certain details of an invention, the Invention Secrecy Act[217] appears to directly violate the First Amendment's literal directive. The First Amendment states that "Congress shall make no law ... abridging the freedom of speech, or of the press ...."[218] These are not absolute

---

eminent domain power, which was specifically created as a check on the government's exercise of this power. *See Kohl*, 91 U.S. at 372.

214. *See* discussion *infra* Part III.C.2.

215. Aptheker v. Secretary of State, 378 U.S. 500, 509 (1964) (quoting Cantwell v. Connecticut, 310 U.S. 296, 304 (1940)). In *Aptheker*, the issue before the Court was the constitutionality of section 6 of the Subversive Activities Control Act of 1950, which made it a felony for members of a Communist organization to apply for, use or attempt to use a passport. *See id.* at 501-02. Despite the fact that the legislation was based on Congress' power to safeguard the national security, the Court held that that the evil Congress sought to control swept "too widely and too indiscriminately across the [right to travel] liberty guaranteed in the Fifth Amendment." *Id.* at 514.

216. United States v. Robel, 389 U.S. 258, 264 (1967) (holding a section of the Subversive Activities Control Act making it unlawful for a member of a Communist-action organization to engage in any employment in any defense facility unconstitutional because it proscribes rights of association protected by the First Amendment).

217. *See* 35 U.S.C. §§ 181, 182, 186 (1994).

218. U.S. CONST. amend. I.

freedoms. However, certain types of speech receive less protection than others. In certain circumstances the government may also restrict the freedom of speech or the press.[219] For these reasons, the constitutionality of invention secrecy depends on the specific degree of First Amendment protection given to expression restricted by the Act, as well as the nature and reasons for the government regulations placed on the expression.

The following discussion consists of three parts. The first part discusses the general issue of protection, analyzing how the First Amendment protects the type of expression proscribed by the Invention Secrecy Act. The second and third parts examine whether the government restrictions created by the Act can overcome the protection granted by the First Amendment, so as to render these restrictions constitutional. In particular, the second part discusses the constitutionality of the apparent system of prior restraints created by the imposition of secrecy on an invention. In the case that no prior restraints are created by the Act, the third part considers whether the Act's penalty provisions are constitutional as subsequent punishments.[220]

1. *FIRST AMENDMENT PROTECTION OF SCIENTIFIC EXPRESSION*

In order to determine the degree of First Amendment protection appropriate for expression proscribed by the Act, it is first necessary to ascertain the type of expression at issue. The purpose of a secrecy order is to prevent an inventor from publishing or disclosing any material information relating to certain patent applications or their subject matter.[221] Accordingly, the analysis that follows involves two inquiries: (1) how does the First Amendment protect scientific expression in general, and (2) considering the nature of the expression proscribed by the Act, does the First Amendment recognize any exceptions that would permit a finding that this particular expression is unprotected?

---

219. *See OTA Report, supra* note 209, at 38.

220. Remember that protection under the First Amendment is a technical term, determinative mainly of the level of scrutiny to be applied to restrictions imposed on the expression. Further, whether speech is classified as protected or not is only of consequence in a subsequent punishment analysis, and not in a prior restraint analysis. Despite the apparent irrelevance of the protection issue to prior restraints, the discussion of protection precedes the prior restraint analysis in order to emphasize the constitutional importance of scientific expression and to provide background in the form of several national security-related cases that are discussed in the context of First Amendment protection.

221. *See supra* text accompanying notes 117-20.

a. Protecting Scientific Expression in General

The expression of scientific information has historically never received explicit First Amendment protection. However, the Founding Fathers did recognize that science and technology were important areas deserving attention and support.[222] Many of the Founding Fathers, including James Madison, Benjamin Franklin, and Alexander Hamilton, were well-educated in science and technology, and understood the importance of scientific freedom.[223] Therefore, when the first ten amendments were added to the Constitution in 1791, scientific freedom was presumed to be included under the First Amendment's broad protections for speech and the press.[224]

A few constitutional scholars have argued that the original purpose of the First Amendment separation of church and state was to give scientific activity or communications the special status given to political comment.[225] This is a minority view, and it has often been argued that scientific expression is less important than political speech because it does not advance the same types of governmental reform ideas.[226] Although scientific information does not contribute to the marketplace of ideas in the same manner as political speech, it does have a significant impact on the free flow of ideas. The importance of scientific speech lies in its factual nature. Facts form the basis of discussion, and discussion of ideas is one of the most important of all First Amendment goals.[227] In addition, technical information is important to informed political debate on a wide range of issues. "Society must collectively participate in evaluating and debating each new technology. If sound policy is to result, technical information is essential in the debate."[228] Furthermore, the free dissemination of all types of scientific information reduces the appearance of government secrecy, and results in greater public trust of political decision-making regarding technology and related issues.[229] Greater public awareness and understanding that arises from the

---

222. *See OTA Report, supra* note 209, at 37.

223. *See id.*

224. *See id.*

225. *See* Steven Goldberg, *The Constitutional Status of American Science*, 1979 U. ILL. L.F. at 1-6.

226. *See OTA Report, supra* note 209, at 37.

227. *See generally* Abrams v. United States, 250 U.S. 616, 630 (Holmes, J., dissenting).

228. A. DEVOLPI ET AL., BORN SECRET: THE H-BOMB, THE *PROGRESSIVE* CASE, AND NATIONAL SECURITY 12 (1981).

229. *See id.* at 5.

dissemination of technology creates better-informed discussions regarding how the government should control technology.[230]

The free dissemination of scientific information furthers the constitutional goal of promoting the progress of the useful arts. For instance, scientists require a free flow of scientific information in order to conduct their research effectively:

> A fundamental tenet of scientific methodology is that basic scientific research results or new scientific theories should be published, widely disseminated, and thoroughly argued, and the results replicated. In part this is in order to share knowledge with other scientists for the benefit of people in general. More immediately, it provides a test and means of validation.[231]

Better scientific research leads to technological advancements that benefit the public. Patents are particularly effective in disseminating scientific information because they provide thorough descriptions of how inventions are made and how they can best be implemented.[232] "The biggest contribution the patent system makes to progress is to induce a steady flow of contributions and to secure their continuous disclosure."[233] Because scientific expression serves important constitutional goals, the fact that it arguably carries less political significance than other forms of expression does not mean that it deserves less protection by the First Amendment.

b. Exceptions to First Amendment Protection

The Court has recognized only a few specific categories of speech that deserve a lesser degree of First Amendment protection.[234] For example, the Court has held that advocacy of imminent illegal conduct, fighting words, and obscenity are unprotected by the First Amendment.[235] The Court has also recognized intermediate categories deserving of lesser protection, including commercial speech, near obscene and offensive

---

230. *See id.* at 12; *see also* Bernstein v. United States Department of State, 922 F. Supp. 1426, 1436 n.17 (N.D. Cal. 1996) (holding that source code is protected speech deserving of a high degree of constitutional protection); *but see* United States v. The Progressive, Inc., 467 F. Supp. 990, 994 (W.D. Wis. 1979) ("this Court can find no plausible reason why the public needs to know the technical details about hydrogen bomb construction to carry on an informed debate on this issue"), discussed *infra* part III.B.3.

231. *OTA Report*, *supra* note 209, at 38.

232. *See* 35 U.S.C. § 112 (1994).

233. *In re* Nelson and Shabica, 280 F.2d 172, 182 (C.C.P.A. 1960).

234. *See, e.g.*, Chaplinsky v. New Hampshire, 315 U.S. 568, 571-72 (1942).

235. *See generally* LAURENCE H. TRIBE, AMERICAN CONSTITUTIONAL LAW §§ 12-9 to 12-17, at 841-928 (2d ed. 1988).

speech, and defamation.[236] General scientific expression does not fall into any of these categories, and therefore, should remain protected under the First Amendment despite its arguably nonpolitical nature.[237]

The government may argue that because of the national security implications of inventions issued with secrecy orders, any expression relating to these inventions falls within the exception for the advocacy of illegal action. This argument is not based on the allegation that scientific expression constitutes advocacy, but rather on the fact that many of the cases discussing the advocacy exception rely on national security as the primary justification for leaving certain types of speech unprotected. In other words, the government would argue for a *national security* exception to First Amendment protection. An analysis of the relevant cases will demonstrate that there are important limitations on when national security interest can justify a deprivation of First Amendment protection.

i. The Clear and Present Danger Test

The government's most compelling arguments that certain types of scientific speech may be unprotected by the First Amendment derive from case law from the World War I period. Several cases at that time were concerned with the advocacy of action which would be detrimental to the war effort. For instance, in *Schenck v. United States*,[238] the Court considered the conviction of Schenck, who had mailed documents contesting the constitutionality of the draft and encouraged those drafted to assert their opposition.[239] The government charged Schenck with attempting to cause insubordination in the armed forces and to obstruct

---

236. *See id.* § 12-18, at 928.

237. One commentator has argued that the expression contained in a patent application should be classified as commercial speech, because a patent protects an inventor's rights while he is selling his invention for profit. *See* Lee Ann Gilbert, *Patent Secrecy Orders: The Unconstitutionality of Interference in Civilian Cryptography Under Present Procedures*, 22 SANTA CLARA L. REV. 358-61 (1982). However, this analysis was used only as a "worst case" scenario to demonstrate that even under a lower standard of protection, the Invention Secrecy Act would still be found unconstitutional. Besides, the connection between scientific expression and commercial speech is tenuous at best, for the reason that commercial speech has generally been interpreted to cover only activities akin to advertising. *See, e.g.*, Central Hudson Gas v. Public Service Comm'n, 447 U.S. 557 (1980).

238. Schenck v. United States, 249 U.S. 47 (1919).

239. *See id.* at 51. The document contained statements such as: "Do not submit to intimidation"; "If you do not assert and support your rights, you are helping to deny or disparage rights which it is the solemn duty of all citizens and residents of the United States to retain"; and "You must do your share to maintain, support and uphold the rights of the people of this country." *Id.*

the recruiting and enlistment service of the United States in violation of the Espionage Act of June 15, 1917.[240] The Court, per Justice Holmes, held that Schenck's speech was unprotected by the First Amendment.[241] To determine when speech should be protected, the Court should ascertain "whether the words used are used in such circumstances and are of such a nature as to create a *clear and present danger* that they will bring about the substantive evils that Congress has a right to prevent."[242] Because Schenck's speech constituted a clear and present danger, it was not protected by the First Amendment.[243]

Despite the fact that *Schenck* involved the advocacy of illegal action, the clear and present danger test, as formulated by Justice Holmes, did not require that there be an intent on the part of the speaker to bring about any particular sort of action. Rather, "[i]t is only the present danger of immediate evil *or* an intent to bring it about that warrants Congress in setting a limit to the expression of opinion where private rights are not concerned."[244] In other words, speech could be found to be unprotected merely because it creates an immediate danger to national security.[245] Thus, the government may argue that certain types of scientific information, especially information relating to military technology, may give an enemy such an advantage as to create an immediate danger to the United States' security.

However, this sort of threat only may be considered imminent during times of war. "When a nation is at war many things that might be said in time of peace are such a hindrance to its effort that their utterance will not be endured so long as men fight and that no Court could regard them as protected by any constitutional right."[246] For the government to succeed on its argument, it faces the difficult burden of proving during

---

240. *See id.* at 49.

241. *See id.* at 52.

242. *Id.*

243. For additional opinions by Justice Holmes regarding the clear and present danger test, see Frohwerk v. United States, 249 U.S. 204 (1919); *see generally*, Debs v. United States, 249 U.S. 211 (1919).

244. Abrams v. United States, 250 U.S. 616, 628 (1919) (Holmes, J., dissenting) (emphasis added).

245. This is not to say that the Court would never consider the intent of the speaker. *See* Debs v. United States, 249 U.S. 211 (1919) (jury was instructed that they could not find defendant guilty unless the words had as their natural tendency and reasonably probable effect to obstruct the recruiting service and unless the defendant had the specific intent to do so in mind). The intent of the speaker need not be examined if there was an immediate danger.

246. *Schenck*, 249 U.S. at 52.

peacetime that the danger created by the expression is so imminent that such expression should be proscribed to protect the country's national security.[247]

Since *Schenck*, the Court has reformulated its standard for determining when First Amendment protections should be curtailed for the sake of national security. In *Dennis v. United States*,[248] the Court purported to apply the clear and present danger test in determining whether to uphold the convictions of defendants for violation of the conspiracy provisions of the Smith Act. The standard used by the Court, however, focused more on the nature of the evil than did the standard articulated by Justice Holmes: "In each case [courts] must ask whether the gravity of the 'evil,' discounted by its improbability, justifies such invasion of free speech as is necessary to avoid the danger."[249] Under this interpretation, if the gravity of the evil is great enough, First Amendment protection can be withheld, even if the evil is not imminent. This interpretation of the clear and present danger test lends greater support to the government's argument for less protection for security-related scientific information. By removing the requirement of imminence and focusing on the gravity of the danger, the government can emphasize that the evil created by the potential publication of certain inventions justifies depriving the inventor of First Amendment protection.

ii. The Modern Advocacy Standard

Despite the force of the government's arguments based on both *Schenck* and *Dennis*, under modern Supreme Court holdings the expression at issue in a patent application probably still would be protected. The Court in *Brandenburg v. Ohio*[250] established a much more stringent test for when speech would be unprotected. In considering the constitutionality of the Ohio Criminal Syndicalism Act, which proscribed the advocacy of violence as a means of accomplishing industrial or political reform, the Court held that the State may proscribe such advocacy only where the advocacy "is directed to inciting or producing imminent lawless action and is likely to incite or produce such action."[251] Under this newly

247. For an argument that in certain situations scientific expression can satisfy an "imminence" test, see United States v. The Progressive, Inc., 467 F. Supp. 990, 996 (W.D. Wis. 1979) (holding that publication teaching how to make a hydrogen bomb does create a threat of direct, immediate, and irreparable injury), discussed *infra* text accompanying notes 286-88.

248. Dennis v. United States, 341 U.S. 494 (1951).

249. *Id.* at 510 (citing Dennis v. United States, 183 F.2d 201, 212 (1950) (L. Hand, C.J.)).

250. 395 U.S. 444 (1969).

251. *Id.* at 447.

articulated standard, the Court struck down the Ohio statute because it failed to distinguish mere advocacy, which is protected, from incitement to imminent lawless action, which is not.[252]

Using the *Brandenburg* standard, the scientific information contained in patent applications will be unprotected only if it: (1) is directed to producing imminent lawless action, and (2) is likely to produce such action. Unlike previous conceptions by the Court, the test looks not only to the effect of the speech, but to the *intent* of the speaker as well. It is highly unlikely that information contained in a patent application carries an intent to produce any sort of illegal action. Further, a peacetime violation of an order is unlikely to promote imminent lawless action. For these reasons, scientific information contained in patent applications will rarely satisfy the first prong of the *Brandenburg* standard, and therefore, this speech should remain protected.

The government may attempt to argue that the *Brandenburg* test only applies to advocacy, and does not apply to scientific expression which does not purport to advocate anything. The government would then argue for a return to the more favorable *Dennis* test which does not require any intent on the part of the speaker. This interpretation fails to consider case law subsequent to *Brandenburg* which indicates that its test should not only apply to advocacy, but to expression in general.[253] Moreover, *Brandenburg* clearly evidences the Court's intent to dismiss the "gravity of the evil" test used in *Dennis* in favor of a test considering the *likelihood* of *imminent*, lawless action. Accordingly, the *Dennis* test no longer should be applied to determine whether speech is unprotected.[254]

c. Summary of First Amendment Protection

Scientific expression was meant to be protected by the First Amendment. Scientific expression of the type proscribed by the Invention Secrecy Act is not unprotected under the modern standard of *Brandenburg* because the expression is not directed to imminent lawless action. It is important to remember that a finding of protected or unprotected speech is merely the first step in the First Amendment analysis. The national security interest may still be considered even after a finding that speech is protected. For prior restraints, this interest is

---

252. *See id.* at 448-49.

253. *See* Texas v. Johnson, 491 U.S. 397, 409 (1989) (defining the *Brandenburg* test in terms of "expression," not advocacy).

254. *Dennis* can still be applied in other contexts, such as in considering whether the presumption against prior restraints can be overcome. *See* Nebraska Press Ass'n v. Stuart, 427 U.S. 539 (1976), discussed *infra* note 298.

considered regardless of whether speech is protected.[255] For subsequent punishments, this interest is considered in the context of determining whether strict scrutiny can be satisfied.[256]

2. *THE CONSTITUTIONALITY OF THE APPARENT SYSTEM OF PRIOR RESTRAINTS CREATED BY INVENTION SECRECY*

The most significant type of government restriction authorized by the Invention Secrecy Act is the apparent system of prior restraints created by the imposition of a secrecy order. The term prior restraint is used "to describe administrative and judicial orders forbidding certain communications when issued in advance of the time that such communications are to occur."[257] Regardless of whether or not the speech is protected, such restrictions have been strongly disfavored by the Supreme Court: "In determining the extent of the constitutional protection, it has been generally, if not universally, considered that it is the chief purpose of the guaranty to prevent previous restraints upon publication."[258] Because of this strong policy against prior restraints, the government bears a heavy burden of showing justification for the imposition of a system of prior restraints.[259] Courts consider two issues: (1) whether invention secrecy establishes a system of prior restraints, and (2) if so, whether the presumption of invalidity against such a system can be overcome.[260]

a. Does the Invention Secrecy Act Create a Prior Restraint?

The initial consideration is whether the imposition of a secrecy order creates a system of "prior restraints" as that term has been interpreted by the Court. Generally, the two types of prior restraints recognized by the

---

255. *See* discussion *infra* Part III.B.2.b.

256. *See* discussion *infra* Part III.B.3.

257. Alexander v. United States, 509 U.S. 544, 550 (1993) (citation omitted).

258. Near v. Minnesota, 283 U.S. 697, 713 (1931) (holding that a statute providing for the abatement, as a public nuisance, of a malicious, scandalous and defamatory periodical was an unconstitutional restraint on publication).

259. *See* New York Times Co. v. United States, 403 U.S. 713, 714 (1971) (per curiam) (citations omitted).

260. Note that this analysis will only apply to private inventors. The Court has repeatedly held that where an individual is in a position of trust with the government, an injunction against dissemination by that person may be appropriate even without proof of a substantial danger. *See Hearings, supra* note 53, at 246 (testimony of H. Miles Foy, Senior Attorney-Adviser, Office of Legal Counsel).

Court are licensing schemes[261] and government orders or injunctions.[262] The Invention Secrecy Act creates both.

First, the Act requires that all patent applications be screened by the Secret Group of the Patent Office.[263] This censorship process allows the government to decide which patents will be disclosed to the public before any publication takes place. Thus, the Invention Secrecy Act operates like a licensing scheme by determining in advance which patents are worthy of publication.

Second, a secrecy order *is* a government order restraining publication. The order prevents the Commissioner of Patents from publishing the patent, and it restrains the individual inventor from publishing or disclosing any material information contained in the application. The statute explicitly authorizes the Commissioner of Patents to order that an "invention be kept secret."[264] The following is an example of such an order taken from the 1980 Hearings:

> You are hereby notified that your application as above identified has been found to contain subject matter, the unauthorized disclosure of which might be detrimental to the national security, and *you are ordered in nowise to publish or disclose the invention or any material information with respect thereto,* including hitherto unpublished details of the subject matter of said application, in any way to any person not cognizant of the invention prior to the date of the order, including any employee of the principals, but to keep the same secret except by written consent first obtained of the Commissioner of Patents and Trademarks, under the penalties of 35 U.S.C. (1952) 182, 186.[265]

Inventors receive an actual order from the government to refrain from publication or disclosure. The Code of Federal Regulations recognizes that inventors are the ultimate recipients of secrecy orders, stating that "[t]he secrecy order is directed to the applicant, his successors, any and all assignees, and their legal representatives."[266] Accordingly, it is implicit in the operation of invention secrecy that a secrecy order creates a prior restraint on inventors' individual expression.

To dispute these findings, the government may argue that because a secrecy order is not directed toward any pre-planned publication or disclosure by the inventor, it cannot create a prior restraint. Before the government issues a secrecy order it may not yet have asserted a specific

---

261. *See, e.g.*, Freedman v. Maryland, 380 U.S. 51 (1965).

262. *See, e.g.*, Near v. Minnesota, 283 U.S. 697 (1931).

263. *See* discussion *supra* Part III.A.

264. 35 U.S.C. § 181 para. 1, 3 (1994).

265. *Hearings, supra* note 53, at 126 (notice of secrecy order from Director, Special Laws Administration Group, Patent and Trademark Office) (emphasis added).

266. 37 C.F.R. § 5.2 (1996).

intent to publish or disclose the contents of the inventor's application.[267] Therefore, a secrecy order does not prohibit the inventor from taking any particular action of which the government is already aware. Rather, a secrecy order is more like a subsequent punishment in that its penalties come into play *only if* an inventor decides to publish or disclose the inventor's invention. If the inventor already has decided not to publish or disclose, the inventor can never be restrained by the secrecy order.

A secrecy order is also like a subsequent punishment in that the punishment for its violation comes solely from the statute. Specifically, violations of a secrecy order are punishable only by the provisions of section 182 and 186 of the Act. Secrecy orders are therefore unlike the typical prior restraints of court injunctions or temporary restraining orders which are punishable by contempt proceedings.

The line between prior restraints and subsequent punishments is not always easy to draw. To resolve the issue in the present case, it is necessary to return to the policy underlying the doctrine against prior restraints: "The special vice of a prior restraint is that communication will be suppressed, either directly or by inducing excessive caution in the speaker, before an adequate determination [has been made] that it is unprotected by the First Amendment."[268] Regardless of whether an inventor actually plans to publish or disclose any information, a secrecy order chills the ability to express this information. A prior restraint can be distinguished from a subsequent punishment by its particular effect of suppressing expression. For instance, the Court in *Near v. Minnesota*, in finding an injunction authorized by a nuisance statute to be an unconstitutional prior restraint, emphasized that "[t]he object of the statute is not punishment, in the ordinary sense, but suppression ...."[269] Similarly, although the Invention Secrecy Act *punishes* inventors for publishing or disclosing certain information, the primary purpose of the statute is to *suppress* information. Therefore, secrecy orders create prior restraints on expression.

---

267. Inventors do, however, have the expectation that their issued patents will be printed by the Patent Office. This publication cannot be considered individual expression, nor can inventors assert a mandatory right to the publication, because the Commissioner alone possesses the discretion to decide whether patents should be published or not. *See* 35 U.S.C. § 11 (1994) ("[t]he Commissioner *may* print, or cause to be printed ... [p]atents, including specifications and drawings, together with copies of the same") (emphasis added).

268. Pittsburgh Press Co. v. Pittsburgh Comm'n on Human Relations, 413 U.S. 376, 390 (1972).

269. Near v. Minnesota, 283 U.S. 697, 711 (1931).

b. Can the Interest in Protecting National Security Overcome the Heavy Burden Against Prior Restraints?

Once a prior restraint has been identified, the government must meet a heavy burden to justify it.[270] This burden may be overcome only under exceptional circumstances, such as when the nation is at war, when publication of obscene materials is threatened, or when one incites acts of violence or the forceful overthrow of the government.[271] The following analysis examines when a national security interest—such as the one at the heart of the Invention Secrecy Act—may be used to overcome the presumption of invalidity of a prior restraint.

i. The New York Times Test

In *New York Times Co. v. United States*,[272] the United States sought to enjoin the New York Times and the Washington Post from publishing the contents of a classified study entitled the "History of U.S. Decision-Making Process on Vietnam Policy."[273] The Court, by a 6-3 vote, held *per curiam* that the government failed to meet its burden to justify the restraint.[274] Although the Court did not reach a majority rule, the concurring opinions of five of the Justices reflect the Court's overall reluctance to use a national security interest to justify a prior restraint.[275]

Two of the Justices in *New York Times* supported the view that prior restraints are always unconstitutional. Justice Black stated:

> The guarding of military and diplomatic secrets at the expense of informed representative government provides no real security for our Republic. The Framers of the First Amendment, fully aware of both the need to defend a new nation and the abuses of the English and Colonial Governments, sought to give this new society strength and security by providing that freedom of speech, press, religion, and assembly should not be abridged.[276]

Justice Douglas' opinion was similar in tone to Justice Black's opinion: "Secrecy in government is fundamentally anti-democratic, perpetuating bureaucratic errors. Open debate and discussion of public issues are vital

---

270. *See* New York Times Co. v. United States, 403 U.S. 713, 714 (1974) (per curiam) (citations omitted).

271. *See* Near v. Minnesota, 283 U.S. at 716 (citations omitted).

272. 403 U.S. 713 (1974).

273. *Id.* at 714.

274. *See id.*

275. Justice Marshall also concurred in the decision, but focused on a separation of powers rationale rather a First Amendment analysis in reaching his decision. *See id.* at 740-41 (Marshall, J., concurring).

276. *Id.* at 719 (Black, J., concurring).

to our national health. On public questions there should be 'uninhibited, robust, and wide-open' debate."[277]

In contrast to Justice Black's and Justice Douglas' absolute prohibition of prior restraints, Justices Brennan, Stewart and White took a more moderate approach to the justification of prior restraints. These Justices recognized that prior restraints are highly disfavored, but in limited circumstances national security may overcome the presumption of invalidity. Justice Brennan stated that "publication must inevitably, directly, and immediately cause the occurrence of an event kindred to imperiling the safety of a transport already at sea ...."[278] Similarly, Justices Stewart and White refused to allow the prior restraint because the publication would not "surely result in direct, immediate, and irreparable damage to our Nation or its people."[279] Thus, because the article at issue did not create a sufficient danger to national security, the Court found the prior restraint invalid.

Given the manner in which the votes were distributed in the *New York Times* case, it appears that the standard used should at least be that supported by Justices Brennan, Stewart and White.[280] Under this standard, during peacetime most inventions subject to secrecy order do not create "direct, immediate, and irreparable damage" to national security. Although it is conceivable that a few inventions may cause such damage,[281] even the most critical military inventions would rarely create *direct, immediate and irreparable* damage. Therefore, secrecy orders as applied to these inventions should be unconstitutional.

ii. The Effect of the *Progressive* Decision

Although no Supreme Court case discusses the constitutionality of a restraint on scientific information, one lower court decision has addressed many issues similar to those raised under the Invention Secrecy Act. In *United States v. Progressive, Inc.*, Progressive planned to publish an article entitled "The H-Bomb Secret: How We Got It, Why We're Telling It."[282] This article described the operation of a hydrogen bomb.[283] The government's fear was that publication of the article could give a foreign country information that would accelerate its creation of an atomic bomb.

---

277. *Id.* at 724 (Douglas, J., concurring).

278. *Id.* at 726-27 (Brennan, J., concurring).

279. *Id.* at 730 (Stewart, J., joined by White, J., concurring) (emphasis added).

280. This standard is appropriate because Justices Black and Douglas supported an even harsher rule against prior restraints.

281. *See supra* text accompanying notes 286-88.

282. 467 F. Supp. 990, 998 (W.D. Wis. 1979).

283. *See id.*

It sought a preliminary injunction against publication, communication, or disclosure of any restricted data contained in the article.[284] In granting the injunction, the court stated:

> In view of the showing of harm made by the United States, a preliminary injunction would be warranted even in the absence of statutory authorization because of the existence of the likelihood of direct, immediate and irreparable injury to our nation and its people.[285]

Accordingly, by applying the standard approved by Justices Brennan, Stewart, and White in *New York Times*, the court in *Progressive* found that the government met its heavy burden of showing justification for the prior restraint.

The result in *Progressive* is instructive because the case involved scientific information similar in nature to the information contained in a patent application. The information contained in The Progressive's article described how to make a weapon, much like a patent application teaches how to make an invention. Therefore, the court's use of the *New York Times* test of direct, immediate, and irreparable damage justified the previous application of the test to inventions on which secrecy orders are imposed. Note, however, the court in *Progressive* did find that the national security threat was great enough to warrant the prior restraint. These contrary results can be reconciled by considering the gravity of the threat at issue in *Progressive*: "What is involved here is information dealing with the most destructive weapon in the history of mankind, information of sufficient destructive potential to nullify the right to free speech and to endanger the right to life itself."[286] The unique threat created by publishing technical information on how to build a hydrogen bomb satisfied the "direct, immediate, and irreparable damage" test. "A mistake in ruling against the United States could pave the way for thermonuclear annihilation for us all."[287] Because the scientific information contained in many patent applications does not create as great of a threat, the *New York Times* test will not be satisfied and secrecy orders as applied to them would be unconstitutional.[288]

---

284. *See id.* at 999.

285. *Id.* at 1000 (citing New York Times Co. v. United States, 403 U.S. 713, 730 (1974) (Stewart, J., concurring)).

286. *Id.* at 995.

287. *Id.* at 996.

288. In some respects, even though *Progressive* purports to apply the *New York Times* test, in fact it appears to have applied a test considering the gravity of the harm, such as the one used in *Dennis v. United States*. *See supra* text accompanying notes 248-49. There is support from the Supreme Court for using such a method. In *Nebraska Press Association v. Stuart*, 427 U.S. 539, 562 (1976), Chief Justice Burger applied the test used in *United States*

iii. Potential Limitations on Applying the *New York Times* Test

The *New York Times* standard supports two limitations on the publication or disclosure of inventions. First, several of the justices in *New York Times* emphasized the importance of political speech[289] and the right of the press to disseminate news.[290] Invention secrecy implicates few of these factors. As noted above, scientific expression, despite its political implications, does not generally contribute to the marketplace of ideas in the same manner as political speech.[291] Moreover, the proscription on publication of an invention does not affect the right of the organized press to disseminate news. Second, a few of the justices in *New York Times* indicated that their decisions might have been different if the prior restraints were authorized under statute. For instance, Justice White stated:

> But I nevertheless agree that the United States has not satisfied the very heavy burden that it must meet to warrant an injunction against publication in these cases, at least in the absence of express and appropriately limited congressional authorization for prior restraints in circumstances such as these.[292]

Justice Stewart took a similar view.[293] Because a secrecy order is made pursuant to the Invention Secrecy Act, there may be more reason to allow the prior restraint.

Progressive seems to support these arguments, at least in part, by stating:

---

v. Dennis, where the Court shall determine whether "the gravity of the 'evil,' discounted by its improbability, justifies such invasion of free speech as is necessary to avoid the danger." 183 F.2d 201, 212 (2d. Cir. 1950), *aff'd*, 341 U.S. 494 (1951). *Nebraska Press*, however, dealt specifically with the problems of containing trial publicity. *See* 427 U.S. at 542. It is therefore unclear whether the Supreme Court would condone an application of the *Dennis* test to issues of national security.

289. *See New York Times*, 403 U.S. at 720 (Black, J., concurring) (emphasizing the importance in maintaining "the opportunity for free political discussion" (quoting De Jonge v. Oregon, 299 U.S. 353, 365 (1937) (Hughes, C.J.)); *id.* at 724 (Douglas, J., concurring) (stressing the importance of promoting debate over Vietnam policy).

290. *See id.* at 717 (Black, J., concurring) ("Both the history and language of the First Amendment support the view that the press must be left free to publish news, whatever the source, without censorship, injunctions, or prior restraints"); *id.* at 730-31 (White, J., concurring) ("I concur in today's judgments, but only because of the concededly extraordinary protection against prior restraints enjoyed by the press under our constitutional system").

291. *See supra* text accompanying notes 225-30.

292. *New York Times*, 403 U.S. at 731 (White, J., concurring).

293. *See id.* at 730 (Stewart, J., concurring).

> They [defendants] believe publication will provide the people with needed information to make informed decisions on an urgent issue of public concern. However, this Court can find no plausible reason why the public needs to know the technical details about hydrogen bomb construction to carry on an informed debate on this issue.[294]

This statement implies that the court was more willing to justify the prior restraint because the information contained in the article had less First Amendment value. Furthermore, the court did seem to place great value on the fact that a particular statute was involved in authorizing the restraint:

> A final and most vital difference between [this case and *New York Times*] ... is the fact that a specific statute is involved here. Section 2274 of the Atomic Energy Act prohibits anyone from communicating, transmitting or disclosing any restricted data to any person "with reason to believe such data will be utilized to injure the United States or to secure an advantage to any foreign nation."[295]

The fact that Congress, through the Atomic Energy Act, authorized the imposition of the prior restraint provided legitimacy to the decision that the expression should be restrained. Because the Invention Secrecy Act provides statutory authority for invention secrecy, the same sort of argument can be made to justify a secrecy order as a prior restraint.

In spite of these arguments, the fact remains that the court in *Progressive* actually applied the *New York Times* test of direct, immediate, and irreparable damage. The consideration of additional factors was only necessary in *Progressive* because the court attempted to justify prohibiting disclosure on a subject matter so inherently dangerous to national security. Therefore, the *New York Times* test should still be the primary standard for determining whether the prior restraint created by a secrecy order is constitutional. The consideration of the First Amendment value of scientific expression and the statutory authorization of the prior restraint may prove helpful in determining the outcome when the result reached under the *New York Times* test may not be so clear. For instance, for those borderline inventions in which it is unclear whether the gravity of the harm may cause direct, immediate, and irreparable damage, these factors may tip the scales in favor of allowing the prior restraint. However, when the publication or disclosure of an invention clearly does not create direct, immediate, and irreparable damage, consideration of these additional factors is unnecessary. Accordingly, because many inventions would easily fail the *New York Times* test, secrecy orders as applied to these inventions would be unconstitutional.

---

294. *Progressive*, 467 F. Supp. at 994.

295. *Id.*

As a final point, the government may argue that military agencies typically receive a large amount of discretion in issues of national security, so the Court should defer to the agency's judgment regarding the necessity of a particular secrecy order. "[U]nless Congress specifically has provided otherwise, courts traditionally have been reluctant to intrude upon the authority of the Executive in military and national security affairs."[296] Because of this discretion, if a government agency makes a determination that an invention creates direct, immediate, and irreparable damage, the Court would be unlikely to question this decision. However, the standard currently imposed by the Invention Secrecy Act is not as high as the *New York Times* test. When an agency issues a secrecy order, it has only determined that publication or disclosure of the invention would cause a threat to national security.[297] The agency has not explicitly found direct, immediate, or irreparable damage. Without such findings, the Court need not defer to the agency's judgment, and should still find a secrecy order unconstitutional.

3. *THE CONSTITUTIONALITY OF SUBSEQUENT PUNISHMENTS CREATED BY THE INVENTION SECRECY ACT*

In addition to the possible prior restraint created by the imposition of a secrecy order, the Invention Secrecy Act also imposes two types of subsequent punishments on inventors. If an inventor violates a secrecy order by publishing or disclosing information relating to his invention, section 182 calls for abandonment of the patent,[298] and section 186 imposes criminal fines and/or imprisonment if this publication or disclosure is willful.[299] A discussion of these subsequent punishments may seem unnecessary given the fact that the prior restraint analysis will often be dispositive of the issue of whether a secrecy order is constitutional. If a secrecy order is an unconstitutional prior restraint, it is unnecessary to analyze whether the penalty provisions enforcing the order are unconstitutional as well. Likewise, if a secrecy order does create a prior restraint, but is constitutional, then the subsequent punishment provisions should also be constitutional because the First

296. Department of Navy v. Egan, 484 U.S. 518, 530 (1988) (citations omitted).

297. *See* 35 U.S.C. § 181 (1994).

298. 35 U.S.C. § 182 (1994).

299. 35 U.S.C. § 186 (1994).

Amendment treats prior restraints much more harshly than subsequent punishments.[300]

Subsequent punishment analysis remains necessary only when no prior restraints have been found on the expression under consideration. The penalty provisions contained in sections 182 and 186 then become the primary government restrictions at issue in the constitutional analysis.[301] Unlike the prior restraint analysis, the constitutionality of the subsequent punishment provisions depends in large part on whether the expression proscribed by the Act is protected.[302] When speech is protected, the Court has repeatedly held that content-based restrictions on expression are subject to strict scrutiny.[303] The penalty provisions at issue here fit the definition of content-based restrictions because they restrict expression on the basis of subject matter.[304] Under the strict scrutiny standard, the government regulations must be: (1) necessary to serve a compelling state interest, and (2) narrowly drawn to achieve that interest.[305]

The Court has established that national security is a compelling interest. "The Government has a compelling interest in protecting ... the secrecy of information important to our national security ...."[306] The

---

300. *See* Alexander v. United States, 509 U.S. 544, 554 (1993). The reason for the harsher treatment given to prior restraints derives from the particular danger that prior restraints pose to expression:

A system of prior restraint is in many ways more inhibiting than a system of subsequent punishment: It is likely to bring under government scrutiny a far wider range of expression; it shuts off communication before it takes place; suppression by a stroke of the pen is more likely to be applied than suppression through a criminal process; the procedures do not require attention to the safeguards of the criminal process; the system allows less opportunity for public appraisal and criticism; the dynamics of the system drive toward excesses, as the history of all censorship shows.

Nebraska Press Ass'n v. Stuart, 427 U.S. 539, 589-90 (1976) (Brennan, J., concurring) (citing T. EMERSON, THE SYSTEM OF FREEDOM OF EXPRESSION 506 (1970)).

301. Even if the Invention Secrecy Act does create a prior restraint, the analysis of subsequent punishments still makes for a useful academic discussion.

302. *See* discussion *supra* Part III.B.

303. *See* Denver Area Educational Telecommunications Consortium, Inc. v. Federal Communications Commission, 116 S.Ct. 2374, 2413 (1996) (Kennedy, J., concurring and dissenting). Even if scientific expression is completely unprotected, the government may still not impose content-based restrictions on the expression. *See* R.A.V. v. City of St. Paul, 505 U.S. 377, 383-84 (1992).

304. *See* TRIBE, *supra* note 235, § 12-3, at 803.

305. *See* Widmar v. Vincent, 454 U.S. 263, 269-70 (1981).

306. Snepp v. United States, 444 U.S. 507, 509 n.3 (1980).

constitutional analysis turns on whether the penalty provisions imposed by the Invention Secrecy Act are narrowly tailored to serving the national security interest. In *United States v. Robel*,[307] the Court discussed how such a determination should be made: "We have ruled only that the Constitution requires that the conflict between congressional power and individual rights be accommodated by legislation drawn more narrowly to avoid the conflict."[308] The provisions of the Invention Secrecy Act should attempt to minimize the imposition placed on First Amendment values. Inventors would argue, however, that these provisions do not meet this goal, since under the ambiguous and uncertain standards established for the imposition of secrecy orders,[309] certain inventions may have secrecy orders imposed on them even though they do not create a substantial threat to the national security. Inventors would also argue that the standards used by the Invention Secrecy Act create problems of overbreadth for these very same reasons.[310] Specifically, the subsequent punishment provisions have the potential to proscribe speech that does not pose a significant national security threat; therefore, these provisions are not narrowly tailored to protecting the national security interest.

The problem with this argument is that the statute defers secrecy order decisions to the judgment of the interested government agency. Since the penalty provisions are dependent on the issuance of specific orders, the constitutionality of invention secrecy should not be determined on the basis of the statute in its entirety. Instead, the Court must employ a case-by-case inquiry into each individual order to determine whether strict scrutiny is satisfied. Therefore, the individual inventor cannot claim overbreadth or a lack of narrowly tailored legislation on the basis of secrecy orders affecting other inventors.

In most individual cases strict scrutiny should be easy enough to satisfy. Again, national security has clearly been identified as a

---

307. 389 U.S. 258 (1967). *Robel* involved section 5(a)(1)(D) of the Subversive Activities Control Act, which made it unlawful for any member of a Communist-action organization "to engage in any employment in any defense facility." *Id.* at 259-60 (citing 64 Stat. 992, 50 U.S.C. § 784(a)(1)(D)). The purpose of this section was "to reduce the threat of sabotage and espionage in the Nation's defense plants." *Id.* at 264. In finding the section unconstitutional, the Court emphasized that section 5(a)(1)(D) "cut deeply into the right of association," and put the appellee "to the choice of surrendering his organizational affiliation, regardless of whether his membership threatened the security of a defense facility." *Id.* at 264-65.

308. *Id.* at 267-68 n.20.

309. *See* discussion *supra* Part III.

310. *See* Broadrick v. Oklahoma, 413 U.S. 601 (1973) (requiring overbreadth to be substantial).

compelling interest.[311] In addition, recall that the courts will typically defer to executive authority with respect to issues of national security.[312] Thus, even for secrecy orders with marginal national security implications, the Court would be unlikely to question a military determination that the national security interest is compelling.[313] The individual secrecy order is also narrowly tailored because its prohibition is limited to information *material* to the subject matter of the invention.[314] For these reasons, strict scrutiny can be satisfied and the subsequent punishment provisions would be constitutional.[315]

---

311. *See* Snepp v. United States, 444 U.S. 507, 509 n.3 (1980).

312. *See* Department of Navy v. Egan, 484 U.S. 518, 530 (1988) (citations omitted).

313. Note the difference between deferring to an agency determination that the national security interest is compelling, discussed herein, with deferring to an agency determination that publication or disclosure creates direct, immediate, and irreparable damage, discussed *supra* text accompanying notes 296-97. In the present situation, it is appropriate to defer to the agency's judgment because the mere act of imposing a secrecy order reflects the agency's belief that there is a compelling interest in protecting the national security. In the situation before, it was inappropriate to defer to the agency's judgment because the agency had never made a determination that there was direct, immediate, and irreparable damage. Also note that the deference given to national security interests does have its limits. In *United States v. Robel*, the Court rejected an argument that Congress' war making power gave it broad discretion to enact a statute limiting First Amendment rights: "[T]he phrase 'war power' cannot be invoked as a talismanic incantation to support any exercise of congressional power which can be brought within its ambit. '[E]ven the war power does not remove constitutional limitations safeguarding essential liberties.'" *Robel*, 389 U.S. at 263-64 (quoting Home Bldg. & Loan Ass'n v. Blaisdell, 290 U.S. 398, 426 (1934)). Although *Robel* involved congressional legislation and not executive decision-making, the significance of this statement should not be ignored. Government action relying on the national security interest should not be exercised indiscriminately. When the government agency uses its discretion to impose a secrecy order on an invention, it must take care not to overly burden First Amendment liberties.

314. *See supra* text accompanying note 120.

315. Because strict scrutiny was found to be satisfied, it is unnecessary to go into great detail of the other arguments the government could have made. Suffice it to say that the government could also have relied on the nonpolitical nature of the expression to argue that a lesser standard of review should apply. The basis for such an argument would be that the penalty provisions are not in fact content-based regulations. Although the government action proscribes the expression on the basis of the particular content of the patent applications, the government is not actually concerned with the message of the expression. Rather, the government is only concerned with the effect that the speech would have on the national security. In cases involving content-neutral regulations on speech, a lesser standard of review applies. *See generally* TRIBE, *supra* note 235, § 12-23, at 977-86.

C. Fifth Amendment Analysis

The Fifth Amendment states: "No person shall ... be deprived of life, liberty, or property, without due process of law; nor shall private property be taken for public use, without just compensation."[316] These two provisions, known as the Due Process Clause and the Takings Clause, respectively, impose certain requirements on the government before it can deprive or take property away from a private individual. The analysis that follows considers whether the procedures promulgated by the Invention Secrecy Act satisfy the requirements imposed by these two clauses.

1. *PROCEDURAL DUE PROCESS*

The protections provided by the Due Process Clause of the Fifth Amendment have been described as protecting against two types of government actions:

> So-called "substantive due process" prevents the government from engaging in conduct that "shocks the conscience," or interferes with rights "implicit in the concept of ordered liberty." When government action depriving a person of life, liberty, or property survives substantive due process scrutiny, it must still be implemented in a fair manner. This requirement has traditionally been referred to as "procedural" due process.[317]

With respect to the Invention Secrecy Act, the imposition of a secrecy order survives substantive due process scrutiny because the government action involved is not of the type that "shocks the conscience" or interferes with rights "implicit in the concept of ordered liberty."[318] Therefore, the focus of the due process analysis is on the procedural aspects of the imposition of a secrecy order.

a. The Due Process Property Interest

The individual affected by the action must first establish a protected interest in life, liberty, or property. Case law has long established that patents are considered a property interest for due

---

316. U.S. CONST. amend. V.

317. United States v. Salerno, 481 U.S. 739, 746 (1987) (citations omitted).

318. A review of the substantive due process case law is beyond the scope of this article. Suffice it to say that the Court has taken an extremely limited view of what types of rights are deemed fundamental in order to receive substantive due process protection. *See* Rochin v. California, 342 U.S. 165, 172 (1952); Palko v. Connecticut, 302 U.S. 319, 325-26 (1937).

process purposes.[319] Invention secrecy, however, creates a unique situation because no patent has actually been issued. Instead, an inventor maintains an interest only in the pending patent application, or in the invention itself.

The Court has broadly defined what constitutes property under the Due Process Clause. Property is not limited to the "actual ownership of real estate, chattels, or money,"[320] but extends to "interests that a person has already acquired in specific benefits."[321] The scope of this protection has been described by the Court as follows:

> To have a property interest in a benefit, a person clearly must have more than an abstract need or desire for it. He must have more than a unilateral expectation of it. He must, instead, have a legitimate claim of entitlement to it .... [Property interests] are created and their dimensions are defined by existing rules or understandings that stem from an independent source such as state law—rules or understandings that secure certain benefits and that support claims of entitlement to those benefits.[322]

Applying these statements to the present situation, it appears that an inventor might not have a legitimate property interest in a pending patent application. Since a property interest can only exist for "already acquired" benefits, an inventor who only possesses a patent application has not yet received any of the benefits that derive from holding a patent.[323] The provisions of the Invention Secrecy Act negate any notion that the rules or understandings established by the patent laws entitle an inventor to the benefits of a patent, because the Act clearly states that no patent shall issue on an application with an imposed secrecy order.[324] An inventor applying for a patent has therefore been put on notice that

---

319. *See* Patlex Corp. v. Mossinghoff, 758 F.2d 594, 599 (Fed. Cir. 1985); Consolidated Fruit Jar Co. v. Wright, 94 U.S. 92, 96 (1876) ("[A] patent for an invention is as much property as a patent for land."); 35 U.S.C. § 261 (1994) ("[S]ubject to the provisions of this title, patents shall have the attributes of personal property").

320. Board of Regents of State Colleges v. Roth, 408 U.S. 564, 571-72 (1972) (holding that a non-tenured university professor who was not rehired had no property interest in employment at the university because the statutory terms of the professor's employment set out a specified end date).

321. *Id.* at 576.

322. *Id.* at 577.

323. The primary benefit of holding a patent is the ability to *exclude* others from making, using, offering to sell or selling the subject matter of the patent. *See* 35 U.S.C. § 271 (1994). This *exclusive* property interest is only created once the patent is issued; an inventor cannot sue another for infringement based on an otherwise allowable patent application.

324. *See* 35 U.S.C. § 181 para. 1, 3 (1994).

should his or her application receive a secrecy order, he or she shall not be able to obtain a patent. Accordingly, the inventor cannot legitimately claim that he or she is entitled to benefits from the application.

However, another approach is to consider the property interest as existing in the invention itself. There is no doubt that an inventor has a property interest in his own invention that can be protected by the Fifth Amendment. The statute itself, by using the phrase "an invention in which the Government has a property interest,"[325] recognizes that an invention can constitute property. However, a secrecy order does not deprive the inventor of the actual invention. Rather, it deprives the inventor of the use of the invention by requiring him to keep the invention secret.[326] Still, this deprivation of use effectively takes away "already acquired" benefits that the inventor has obtained in the invention, because prior to applying for the patent, the inventor had the right to publish, disclose and freely use the invention. After the secrecy order is issued, the inventor can no longer partake in these activities. Therefore, a secrecy order deprives the inventor of a property interest protected by the Due Process Clause.[327]

b. What Process Is Deserved?

The next question in the analysis is whether the provisions of the Invention Secrecy Act establish proper procedures for the deprivation of that interest. The Court has stated:

> This Court consistently has held that some form of hearing is required before an individual is finally deprived of a property interest. The "right to be heard before being condemned to suffer grievous loss of any kind, even though it may not involve the stigma and hardships of a criminal conviction, is a principle basic to our society." The fundamental requirement of due process is the opportunity to be heard "at a meaningful time and in a meaningful manner."[328]

The language of the Court suggests that some sort of pretermination hearing is required to satisfy procedural due process. The Invention Secrecy Act, however, provides for no significant procedures before the issuance of the secrecy order. Rather, the Act contains only procedural

---

325. 35 U.S.C. § 181 para. 1 (1994).

326. *See* 35 U.S.C. § 181 para. 1, 3 (1994).

327. This view is supported by Court decisions in which an invention has been deemed to be private property for the purpose of eminent domain. *See infra* text accompanying note 351; *see also* Ruckelhaus v. Monsanto, 467 U.S. 986, 1001 (1984) (applying the same definition of property in a Takings Clause analysis as has been used in the procedural due process context).

328. Mathews v. Eldridge, 424 U.S. 319, 333 (1976) (citations omitted).

safeguards that occur subsequent to the issuance of the secrecy order, such as a right to appeal the secrecy order,[329] a right to petition for modification[330] or rescission,[331] and a right to obtain compensation.[332]

The procedural due process requirement has not been interpreted to strictly require a hearing before final termination of the interest. Instead, "due process is flexible and calls for such procedural protections as the particular situation demands."[333] The Court's preferred approach is to apply a balancing test considering three factors:

> First, the private interest that will be affected by the official action; second, the risk of an erroneous deprivation of such interest through the procedures used, and the probable value, if any, of additional or substitute procedural safeguards; and finally, the Government's interest, including the function involved and the fiscal and administrative burdens that the additional or substitute procedural requirement would entail.[334]

The following section applies this test to the imposition of a secrecy order to determine whether pretermination procedures are required, or whether the existing post-termination procedures are sufficient.

The first factor that should be considered is the private interest. A secrecy order temporarily prevents an inventor from receiving a patent on an invention that is otherwise patentable.[335] The private interest affected is the inventor's interest in obtaining this patent, or other benefits that can be derived from the use of the invention. More specifically, the inventor's main interest is obtaining a patent sooner rather than later. The Court has held that the need for hearings *prior* to termination of benefits is important only for those people on the very margin of subsistence, emphasizing the importance of financial need.[336] An inventor waiting for a patent does not have a presumption of an important financial need. Further, in a social security disability benefits case, the Court held that the "sole interest is in the uninterrupted receipt of this source of income pending final administrative decision on [the] claim."[337] The inventor who has not yet received a patent has also not yet obtained any benefits from the potential patent, and therefore does not face the problem of

---

329. *See* 35 U.S.C. § 181 (1994).

330. *See* 37 C.F.R. § 5.5 (1996).

331. *See* 37 C.F.R. § 5.4 (1996).

332. *See* 35 U.S.C. § 183 (1994).

333. Morrissey v. Brewer, 408 U.S. 471, 481 (1972).

334. *Eldridge*, 424 U.S. at 335 (citations omitted).

335. *See* 35 U.S.C. § 181 (1994).

336. *See* Goldberg v. Kelly, 397 U.S. 254, 264 (1970).

337. *Eldridge*, 424 U.S. at 340.

having financial benefits interrupted. Consequently, the inventor's private interest in obtaining a pretermination hearing is minimal.

The second factor to be considered is "the fairness and reliability of the existing pretermination procedures, and the probable value, if any, of additional procedural safeguards."[338] This includes assessing any possible bias of the group making the termination decision.[339] Prior to the imposition of the secrecy order, there are very few safeguards used by either the government agency requesting the order or by the Commissioner of Patents. The Act gives the interested government agency discretion to decide whether a secrecy order is appropriate. Although this may show bias, additional procedures probably would not be helpful. The government agency is in the best position to determine if the secrecy order is necessary. Moreover, the Court has given government agencies broad discretion in determining the scope of procedural due process: "In assessing what process is due ..., substantial weight must be given to the good-faith judgments of the individuals charged by Congress with the administration of ... programs that the procedures they have provided assure fair consideration of the entitlement claims of individuals."[340] Given these factors, the addition of a pretermination hearing to contest a secrecy order would be unlikely to decrease the probability of an erroneous deprivation of property interest.

The third factor to be considered in the balancing test is the government interest. This includes consideration of "the administrative burden and other societal costs that would be associated with requiring, as a matter of constitutional right, an evidentiary hearing ...."[341] The Court has recognized in the context of disability benefit termination that the administrative and societal costs may not be insubstantial.[342] In addition, the Court has expressed reluctance to increase these costs when significant discretion should be given to the administrative agencies responsible for carrying out the termination of benefits.[343] Because of the considerable number of secrecy orders imposed every year, requiring an evidentiary hearing for each one prior to its imposition would place significant costs on the government. The government's interest in protecting the national security is also important because it justifies the need to impose secrecy orders quickly before inventors release any

338. *Id.* at 343.

339. *See id.* at 344.

340. *Id.* at 349.

341. *Id.* at 347.

342. *See id.*

343. *See id.* at 348.

potentially dangerous information. Therefore, the third factor strikes in favor of retaining the existing procedures.

Based on the three factors discussed above, the costs of creating a pretermination hearing are too great to justify. To satisfy procedural due process, post-termination procedures must merely provide inventors with notice of the secrecy order and a meaningful opportunity to be heard.[344] Inventors receive constructive notice of the secrecy order through the statute itself,[345] and actual notice from the Patent Office once it imposes a secrecy order.[346] Inventors receive an opportunity to be heard by an unbiased party through an appeal to the Secretary of Commerce.[347] Accordingly, the Invention Secrecy Act does comport with the flexible requirements of procedural due process.

2. *THE TAKINGS CLAUSE*

The government's power of eminent domain gives it the right to take private property for public use, provided that it pays just compensation.[348] The particular questions involved in the Takings Clause analysis are: (1) do inventors have a property interest protected by the Takings Clause, (2) if so, do the regulations created by the Invention Secrecy Act effect a taking of that property interest, (3) if there is a taking, is it a taking for a public use, (4) if there is a taking for public use, does the statute adequately provide for just compensation?[349]

a. Do Inventors Have a Protected Property Interest?

The first question in the Takings Clause analysis is very similar to the initial question posed in the procedural due process analysis. The only distinction made in the language of the Fifth Amendment between the Due Process Clause and the Takings Clause is that in the takings context, "*private* property" must be taken.[350] Although certain inventions may be government-controlled and therefore not private, the focus of this

344. *See* Eisen v. Carlisle & Jacquelin, 417 U.S. 156, 174 (1974).

345. *See* Texaco Inc. v. Short, 454 U.S. 516, 532 (1982) (stating that to satisfy due process, "a legislature need do nothing more than enact and publish the law, and afford the citizenry a reasonable opportunity to familiarize itself with its terms and to comply"). Before the ASPAB List was declassified, some inventors could have argued that they received no notice that their inventions might cover technology contained in the list. Now, however, because the list has been declassified, inventors can no longer make this argument.

346. *See supra* text accompanying note 120.

347. *See* 37 C.F.R. § 5.8 (1996).

348. *See* U.S. CONST. amend. V.

349. *See* Ruckelshaus v. Monsanto Co., 467 U.S. 986, 1000-01 (1984).

350. *See* U.S. CONST. amend. V (emphasis added).

analysis is on those secrecy orders imposed on private inventors. The Court has recognized that inventions can generally be the subject of eminent domain:

> Many inventions relate to subjects which can only be properly used by the government, such as explosive shells, rams, and submarine batteries to be attached to armed vessels. If it could use such inventions without compensation, the inventors could get no return at all for their discoveries and experiments. It has been the general practice, when inventions have been made which are desirable for government use, either for the government to purchase them from the inventors, and use them as secrets of the proper department; or, if a patent is granted, to pay the patentee a fair compensation for their use.[351]

Moreover, the Supreme Court has used the same definition of property with respect to the Takings Clause as it has used in the procedural due process cases.[352] For many of the same reasons described in the procedural due process analysis, inventors possess a protected property interest under the Fifth Amendment's Takings Clause.[353]

b. Do Government Regulations Effect a Taking of Property?

Court decisions in the context of eminent domain and secrecy orders indicate that a Fifth Amendment taking is possible. In *Constant v. United States (Constant II)*,[354] the Court of Claims held that there can be no Fifth Amendment taking from the mere issuance of a secrecy order.[355] Most important to the case was the fact that the inventor did not allege any specific use of the invention by the government.[356] This lack of factual support for the takings claim was crucial to the court's decision. The court implied that had the takings claim been properly established by alleging government use, the court would be disposed to act upon it.[357] However, where there is only an allegation of damage from the mere imposition of a secrecy order, without an allegation of government use, the inventor's sole recourse is the compensation procedure set out in the Invention Secrecy Act. "An inventor whose patent has been withheld under a secrecy order has a right created by statute, sections 181-88, to obtain compensation from the United States, and the statute sets forth

---

351. James v. Campbell, 104 U.S. 356, 358 (1881).

352. *See Monsanto*, 467 U.S. at 1001 (quoting Board of Regents v. Roth, 408 U.S. 564, 577 (1972)).

353. *See* discussion *supra* Part III.C.1.a.

354. *Constant v. United States (Constant II)*, 16 Cl. Ct. 629 (1989)

355. *See id.* at 632.

356. *See id.*

357. *See id.* at 633.

procedures whereby this right may be realized."[358] Thus, *Constant II* supports the contention that if an inventor can allege actual use of an invention as described in a patent application, the government action moves from a mere regulation to a compensable taking.

Moreover, very rarely will an inventor pursue the takings claim without first attempting to obtain compensation under section 183. In fact, the court in *Constant II* ordered that the inventor take steps to request a modification of the secrecy order before it would even consider the takings claim.[359] This order comports with Supreme Court authority: "Equitable relief is not available to enjoin an alleged taking of private property for a public use, duly authorized by law, when a suit for compensation can be brought against the sovereign subsequent to the taking."[360] In addition, because compensation under section 183 is broader than under an eminent domain theory,[361] inventors will prefer pursuing the section 183 remedy first.

There are indications that because of the remedy available under section 183, a court may refuse entirely to hear a taking claim. Almost every case dealing with secrecy orders addresses compensation under the provisions of section 183 rather than the Fifth Amendment.[362] In *Hornback v. United States*, Hornback, who was issued with a secrecy order for a patent application on a missile guidance system, filed a complaint in district court alleging entitlement to damages under 35 U.S.C. § 183 and under the Fifth Amendment.[363] With respect to the Fifth Amendment argument, the District Court held that "the Fifth Amendment takings

---

358. *Id.* at 632.

359. *See id.* at 632-33.

360. *Monsanto*, 467 U.S. at 1016 (holding that Monsanto's takings claim against the Environmental Protection Agency was not ripe for resolution because Monsanto did not yet pursue compensation authorized under the Tucker Act).

361. *See supra* text accompanying note 156.

362. *See* Farrand Optical Co. v. United States, 325 F.2d 328, 335 (2d Cir. 1963) (refusing to allow just compensation interest because such interest is only allowed when the claimant is entitled to just compensation under the Fifth Amendment); Halpern v. United States, 258 F.2d 36, 39-40 (2d Cir. 1958); Constant v. United States, 617 F.2d 239, 241-42 (Ct. Cl. 1980) ("use of the words 'just compensation' in a federal statute does not necessarily mean that a constitutional taking is involved"); *see also* Radioptics, Inc. v. United States, 621 F.2d 1113, 1126-27 (Ct. Cl. 1980) (holding that a mere prohibition on disclosure or publication of classified information contained in a research proposal, without a prohibition on use of the information, was not a significant enough imposition to cause a taking of the information).

363. Hornback v. United States, 16 F.3d 422, 1993 WL 528066 (Fed. Cir. 1994) (unpublished disposition).

argument is inappropriate, because 35 U.S.C. § 183 provides the exclusive remedy to inventor-owners for damages claimed as the result of a secrecy order imposed by the government."[364] Although the Federal Circuit affirmed the District Court's decision, this opinion was unpublished and determined to be unquotable as precedent. Thus, it is difficult to ascertain the significance of the *Hornback* decision. On one hand, the decision itself supports the contention that there can be no Fifth Amendment taking from the imposition of a secrecy order with or without government use. On the other hand, the fact that the Federal Circuit de-published the opinion (and the United States Supreme Court denied certiorari) may imply that the higher courts may not be so opposed to such a Fifth Amendment taking claim.

Under general takings law, the government regulation at issue under the Invention Secrecy Act does seem to possess the typical characteristics necessary for finding a taking. In particular, the fact that a secrecy order deprives an inventor of virtually all economic use of the invention weighs very heavily toward finding a compensable taking.[365] Therefore, despite the uncertainties surrounding some of the case law as to whether a taking claim is available for the imposition of a secrecy order, the very nature of the regulation, especially when accompanied by government use, strongly supports that such a claim should be recognized. Court decisions have required that before an inventor may even pursue the takings claim, he must exhaust his options under section 183 of the Act. However, owing to the inadequacies of such compensation,[366] if the section 183 remedy fails the inventor may still want to pursue compensation under an eminent domain theory. Accordingly, even though courts may be reluctant to hear a claim under the Takings Clause, inventors should continue to argue that such compensation is available.[367]

c. Is There a Taking for Public Use?

For a constitutional government taking to occur, the taking must be for a "public use." The public use requirement is a very broad one, allowing a taking so long as it is rationally related to a conceivable public

---

364. *Id.*

365. *See* Lucas v. South Carolina Coastal Council, 505 U.S. 1003, 1015-16 (1992) (implying that the Court will forego the typical case-specific inquiry and will find a per se taking where a property owner has been deprived of all economically viable use of his land).

366. *See* discussion *supra* Part II.D.3.

367. For further arguments that a patent application cannot be the subject to a Fifth Amendment taking, see generally Hausken, *supra* note 11, at 245.

purpose.[368] "The scope of the 'public use' requirement of the Takings Clause is 'coterminous with the scope of a sovereign's police powers.'"[369] The protection of national security falls within the scope of these police powers. "The Fifth Amendment implicitly sanctions the taking of private property for public use by requiring only that just compensation be paid therefore .... [P]ublic use includes not only what is necessary for national security but also what is needed for maintaining public health and safety."[370] Because a secrecy order serves to protect the national security, the public use requirement is satisfied.

d. Has Just Compensation Been Paid?

Having taken the private property, the government must pay just compensation. The best approach for determining the proper amount is modeled on cases in which the government has been ordered to pay just compensation for patent infringement.[371] In the Federal Court of Claims, Title 28 U.S.C. section 1498 authorizes a cause of action for government infringement on an inventor's patent.[372] Section 1498 "is essentially an Act to authorize the eminent domain taking of a patent license, and to provide just compensation for the patentee."[373] In these cases, recovery should be "reasonable and entire compensation" for the use and manufacture of the invention.[374]

The Court of Claims has applied the "reasonable and entire compensation" guideline to find that a reasonable royalty is the only appropriate method for awarding just compensation.[375] The court rejected the use of remedies under Title 35 because it "would grant plaintiff a recovery in excess of the just compensation required by the fifth amendment, and in excess of the reasonable and entire compensation contemplated by Congress with the passage of § 1498."[376]

---

368. *See* Hawaii Housing Auth. v. Midkiff, 467 U.S. 229, 241 (1984).

369. Ruckelhaus v. Monsanto, 467 U.S. 986, 1014 (1984) (quoting *Midkiff*, 467 U.S. at 240).

370. *See* H.R. REP. NO. 1540, *supra* note 9, at 25 (quoting PETER D. ROSENBERG, PATENT LAW FUNDAMENTALS 177 (1977)).

371. Just compensation cannot be measured under cases interpreting that term for purposes of section 183. *Constant I* recognized that compensation under section 183 should not be limited to damages under an eminent domain theory. *See* Constant I, 617 F.2d 239, 240, 242 (Ct. Cl. 1980). Therefore, compensation under a takings claim is more limited.

372. 28 U.S.C. § 1498(a) (1994).

373. Leesona v. United States, 599 F.2d 958, 966 (Ct. Cl. 1979).

374. *See* 28 U.S.C § 1498(a) (1994).

375. *See Leesona*, 599 F.2d at 973.

376. *Id.* at 969.

Awards for increased damages, attorney's fees, savings to the government, and lost profits are therefore inappropriate.[377] However, consideration of these types of remedies may be appropriate in applying what the court calls the "comparative royalty technique."[378] This technique "compute[s] the award by estimating a reasonable royalty on a proper compensation base, and then test[ing] this award by an examination of other available measures—savings to the government, lost profits, etc."[379] The initial reasonable royalty is based on a hypothetical meeting between a willing buyer and a willing seller.[380] The other factors are then taken into account to "renegotiate" the reasonableness of the amount.[381] Additionally, a patentee is entitled to delay-damages, i.e., damages for the government's delay in payment of a reasonable royalty.[382] Section 1498(a) of Title 28 also provides that "reasonable and entire compensation shall include the owner's reasonable costs, including reasonable fees for expert witnesses and attorneys, in pursuing the action ...."[383]

Therefore, if an inventor cannot obtain adequate compensation under section 183 of the Invention Secrecy Act, the inventor should pursue compensation in an amount calculated by the aforementioned procedures on the theory that a secrecy order does effect a Fifth Amendment taking for which just compensation must be paid.

V. CONCLUSION

While invention secrecy may have been justified during wartime, the application of invention secrecy during peacetime raises questions regarding the necessity of the doctrine and its effects on inventors' rights. As a matter of legislative policy, the peacetime provisions of the Invention Secrecy Act suffer from a lack of support. The peacetime provisions rely more upon their wartime foundations than upon any articulated justification to protect the national security during peacetime. This article calls for Congress to justify why the peacetime provisions of the Invention Secrecy Act continue to be necessary. This does not mean that the peacetime provisions are completely *unnecessary*. However, invention secrecy should not be concerned with mere speculation

---

377. *See id.* at 966-71.

378. *Id.* at 973.

379. *Id.*

380. *See* Tektronix, Inc. v. United States, 552 F.2d 343, 349 (Ct. Cl. 1977) (citing Georgia-Pacific Corp. v. U.S. Plywood-Champion Papers, Inc., 446 F.2d 295 (2d Cir. 1971)).

381. *See Leesona*, 599 F.2d at 973.

382. *See id.* at 979.

383. 28 U.S.C. § 1498(a) (1994).

regarding a possible threat to national security. Instead, invention secrecy should only be applied when a threat to national security is both imminent and likely. Adopting such a policy will guarantee the protection of national security, while also ensuring the continued promotion of the goals of the patent system and the protection of inventors' rights.

This article makes two proposals regarding the duration of secrecy orders. First, for those inventions that the government predicts would always cause harm to the national security, a permanent secrecy order is appropriate. Permanent secrecy orders benefit inventors because they remove any doubt in the inventors' minds that they might someday be able to exploit their inventions. It also allows inventors to pursue compensation immediately for the imposition of the orders. Second, because the duration of a patent is measured twenty years from the date of *filing*, Congress should consider using a extension period longer than five years to ensure that inventors subjected to secrecy orders of extended duration can exploit their patents for a full term.

The system of compensation established by the Invention Secrecy Act is flawed. Inventors very rarely receive compensation for the imposition of a secrecy order because of the difficulties of proving actual damages and the excessive discretion given to the government agencies. Compensation procedures should be revised to make compensation easier to obtain. Proof of actual damages should not be required. An appropriate solution should consider several factors—such as the degree of utility, novelty, and importance of the invention—which can be used to calculate compensation without proof of actual damages and a separate board to adjudicate disputes.

Constitutional problems present the most serious questions. With respect to the First Amendment, a secrecy order creates a prior restraint on expression. Many secrecy orders issued during peacetime would be found unconstitutional because such orders do not possess the requisite direct, immediate, and irreparable danger to national security. However, a few inventions may satisfy this test if they create extreme danger; the invention of a bomb is an appropriate example. In the unlikely case that courts view secrecy orders only as a subsequent punishment on expression, secrecy orders can satisfy strict scrutiny—and are therefore constitutional—because of a compelling interest in protecting the national security, and the deference given to the executive branch in making national security-related decisions.

The Fifth Amendment problems raised under the Invention Secrecy Act are less severe than the First Amendment problems. Procedural due process can be satisfied through the Act's post-termination procedures providing inventors with notice and an opportunity to be heard. However, the Takings Clause presents the larger problem. When the

government issues a secrecy order and subsequently uses the invention for its own purposes, a taking has occurred for which just compensation must be paid. However, some cases have indicated that because the Invention Secrecy Act already provides compensation under section 183, the court will not even hear the takings claim. This remains an unresolved issue. The courts or the legislature must provide an answer as to whether a valid claim under the Takings Clause can be recognized. Until this issue is resolved, inventors should not be precluded from obtaining compensation under the Takings Clause.

There are signs that government agencies may have finally begun to limit their use of secrecy orders: "In agreement with the Clinton administration policy ... the Pentagon is preparing to limit its use of the ISA and allow greater flexibility."[384] A suggestion has even been made to place the authority for issuing secrecy orders in the Commerce Department instead of the Pentagon.[385] Nevertheless, secrecy orders continue to place significant burdens on private inventors during peacetime. Moreover, secrecy orders may be unconstitutional. For these reasons, the government today should examine the recommendations made throughout this article in an effort to reanalyze the necessity of peacetime invention secrecy and reformulate its operation.

---

384. *Pentagon to Overhaul the Invention Secrecy Act*, INTELLIGENCE NEWSLETTER, Apr. 1, 1993 (Indigo Publications).

385. *See generally Patent Office Tries to Silence Inventors* (National Public Radio broadcast, Sept. 14, 1992).

ARTICLE

CYBERSPACE SELF-GOVERNMENT: TOWN HALL DEMOCRACY OR REDISCOVERED ROYALISM? †

HENRY H. PERRITT, JR. ††

TABLE OF CONTENTS

I. INTRODUCTION 414
II. EXAMPLES OF CYBERSPACE DISPUTES 417
III. IS SELF-GOVERNANCE DESIRABLE? 419
A. Self-governance may be more efficient 420
B. Networks need different rules and procedures 420
C. Open networks escape enforcement of conventional rules 422
D. Self-governance promotes voluntary compliance 424
E. Conclusion 425
IV. IS SELF-GOVERNMENT LEGALLY FEASIBLE? 425
A. Basic legal frameworks 426
B. Contract: the framework for autonomy 433
C. The limits of contract 435
V. THREE EXAMPLES OF CYBERSPACE SELF-GOVERNMENT 437
A. A.c.e.n.a.: an example of self-government 438

† This article is the first of a trilogy of articles considering the relationship between the Internet and regulation. This article considers self-regulation. A second article will consider regulation of the Internet through traditional legal institutions. The third will consider use of the Internet to facilitate governance through traditional legal institutions and new international institutions following traditional institutional models.

†† Dean and Professor of Law, Chicago-Kent College of Law, Illinois Institute of Technology; member of the bar, Virginia, Pennsylvania, District of Columbia, Maryland, United States Supreme Court; J.D., 1975, Georgetown University Law Center; S.M., 1970, MIT; S.B., 1966, MIT. The author appreciates stimulating contributions from several members of the Villanova Law School faculty, particularly John Hyson, Joseph Dellapenna, and Richard Turkington, and from a continuing dialogue with David R. Johnson about governance of Cyberspace.

B. Is A.c.e.n.a. fair? 440
C. Self-government institutions proposed by IAHC 442
D. Economic royalism: proprietary power 449
E. Conclusion 450
VI. COMPARISON WITH OTHER SELF-GOVERNING COMMUNITIES 451
A. Introduction 451
B. Involuntary membership models 452
C. Voluntary membership models 456
D Conclusion 463
VII. WHAT REMAINS TO BE DONE? 464
A. Completing the contractual web 464
B. Immunities and community boundaries 469
C. International treaty 476
VIII. CONCLUSION 476
IX. APPENDIX: CRITERIA FOR AUTONOMY 479

I. INTRODUCTION

Growing interest in the Global Information Infrastructure—the Clinton Administration's Information Superhighway—has given rise to suggestions that some or all of this "cyberspace" should be self-governing—autonomous with respect to the regular law.[1] Cyberspace, the set of electronic network communities, may be distinct enough to have its own law and legal institutions—a system of "cybergovernment." This self-governance may be more efficient for cyberspace. However, the rules and/or the adjudicatory techniques for applying the rules may need to be different from those of the surrounding community. In any event, compliance with the basic norms of the community may be higher when members of the cyberspace subcommunity participate in self-governance. Each of these criteria can be evaluated separately with respect to the three basic activities of governance: rulemaking (legislation), rule application (adjudication), and enforcement. More or less autonomy may be appropriate depending on whether one considers rulemaking, adjudication, or enforcement.

---

1. *See, e.g., White House Paper on Electronic Commerce* (released June 30, 1997) <http://www.iitf.nist.gov/eleccomm/exec_sum.htm>; *Bonn Declaration*, (visited Nov. 24, 1997) <http://www2.echo.lu/bonn/final.html>.

The emergence of new social communities in Internet newsgroups and on public electronic bulletin boards has already attracted comment.[2] Some markets are currently almost completely electronic in cyberspace,[3] and already govern themselves. In general, there are numerous communities that enjoy powers of self-government. Many instances of self-government are so commonplace as to escape notice. Virtually every citizen of a modern state is a member of multiple private organizations: bar associations, national fraternities, and non-profit organizations. All of these organizations exercise some powers of self-governance. Usually, there is little controversy over the application of special bodies of substantive law and the use of specialized institutions to resolve intra-organizational disputes pursuant to charters and bylaws of these organizations. It may seem strange that something can be law without being adopted by a legislature or a court, but it happens all the time and has for centuries.[4]

Self-government—legal autonomy—may also be appropriate for some new electronic communities, although it is extremely unlikely that self-governance will result just because some of the communications occur through new electronic channels. But, when all of the functions of a particular market or of other commercial communities are contained within electronic communications systems, the result is something like a community, whereby the participants may qualify for self-governance.[5] Nevertheless, while it is possible, and in this author's view desirable, for

---

2. *See* Reid Kanaley, *Transforming the Internet Into a World Wide Safety Net*, PHILADELPHIA INQUIRER, Jan. 17, 1995, at A1 (reporting on the Internet's role as a psychological safety net of support groups and crisis intervention techniques for thousands of people contemplating suicide and experiencing other distress); Peter H. Lewis, *Strangers, Not Their Computers, Build a Network in Time of Grief*, N.Y. TIMES, Mar. 8, 1994, at A1 (describing economic and personal support by members of computer forum for family of former member of forum killed in robbery).

3. Markets are economic electronic communities, and some satisfy important needs of their participants. Participants in some markets only have transitory attachments. Attendance at a single auction is an example. Participants in other markets have more than transitory attachment. Someone who regularly sells magazine articles to a group of competing publishers is an example. Certain information markets are almost completely electronic, typified by vendors such as WESTLAW, LEXIS, and Dialog. International financial markets relating to wholesale funds transfers and clearance of credit card transactions also are mostly electronic.

4. *See* discussion *infra* Part VI.

5. The idea of community presupposes shared interests and activities. "Community: 2. a group of people living together as a smaller social unit within a larger one, and having interests, work, etc. in common ...." WEBSTER'S NEW WORLD DICTIONARY 288 (2d ed. 1972).

net participants to make up their own rules and establish their own institutions of government, merely doing this does not necessarily assure them of an immunity or exemption from regular law.[6]

This article considers theoretical legal frameworks for autonomy of open networks, based upon models from other relatively autonomous communities. The article evaluates four possible justifications for electronic community self-governance, and considers sovereignty and contractual frameworks for self-government. The article reviews three attempts at self-governance: one in the alt.current-events.net-abuse newsgroup, another proposed by the International Ad Hoc Committee (IAHC), and a third represented by traditional proprietary services. The article evaluates the justification for such autonomy and inventories the major steps to be taken before credible exercises in self-government can

---

6. Terminology is a problem in talking about self-governance. The main problem arises with respect to what the traditional legal system should be called. This paper refers to it as the "traditional" community or legal system. The traditional system could also be thought of as the "surrounding" or "larger" community or legal system, but that suggests that an electronic community is entirely contained within one traditional legal system or community. While "traditional" does not communicate the precise relationship of potentially self-governing systems with other legal systems, it should be understood as referring to the legal system that ordinarily will govern some or all of the activities of the electronic community being discussed.

A glossary of technical terms may be useful at this point:

Open network refers to a computer network to which anyone may connect, as distinguished from a *closed network*, on which connections are limited to a predetermined group. *Proprietary networks* are instances of closed networks, on which connections are limited to those who have paid a fee. Proprietary networks sometimes use *proprietary protocols* for digital communications between the connected computers, further limiting the class that can connect. Open networks almost always use *open protocols* such as TCP/IP, which defines the Internet. *Open architecture* refers to the configuration of an open network.

A *network services provider* offers a means for connecting a computer to a network, as by providing a dial-up telephone number connected to a modem, which is, in turn, connected to the Internet. An *internet service provider* (ISP) is a type of network services provider. A network administration entity is a person or organization that undertakes to perform network support functions, such as assigning user names, domain names, IP addresses, and e-mail addresses to allow computers to connect to the network and to use its services.

A *network community* is a group of interdependent persons or entities that communicate with each other predominantly via a computer network. The means of communication include newsgroups, e-mail lists, Web pages, and markets and forums organized through the Web.

occur. Then the article briefly explores historical and contemporary models for such arrangements.

Based upon the theoretical frameworks, the justifications, and the models, the article concludes that considerable autonomy can be achieved through contractual arrangements featuring arbitration accompanied by choice of customary substantive law. However, there may be difficulties in defining community boundaries, in implementing effective enforcement mechanisms, and in avoiding antitrust problems in the electronic network context. The article identifies the major points of tangency between regular legal systems and new Internet systems of government. It evinces that an independent legal system for the Internet is most likely to exist if the countries of the world negotiate a kind of "hands-off" treaty, committing themselves to defer to private Internet governing institutions meeting certain criteria, and empowering existing multilateral institutions to play certain ministerial roles. The article concludes by observing that the cybergovernment inquiry is but a subset of a broader range of issues presented by the informal, conversational, and frequently transient nature of electronic transactions in a legal context that has traditionally stressed formality and paper records.

At least three recent law review articles have explicitly considered the possibility of self-governance for electronic communities.[7] This article has a broader scope than the preceding articles in that it considers self-governance and legal autonomy in cyberspace along with self-governance and legal autonomy for other types of private associations. This article also links the basic idea of cyberspace self-governance to recently-proposed mechanisms for private registration of Internet domain names.

II. EXAMPLES OF CYBERSPACE DISPUTES

There is nothing new about the possibility of disputes arising in digital electronic networks. Nor is there anything new about private

---

7. *See generally* William S. Byassee, *Jurisdiction of Cyberspace: Applying Real World Precedent to the Virtual Community*, 30 WAKE FOREST L. REV. 197 (1995) (emphasizing the importance of recognizing differences between on-line and physical interactions, and discussing that autonomous jurisdiction is a utopian solution); Henry H. Perritt, Jr., *President Clinton's National Information Infrastructure Initiative: Community Regained?*, 69 CHI.-KENT L. REV. 991 (1994) (Charles Green Lecture) [hereinafter Perritt, *Community Regained*] (exploring the role of new computer and communications technologies in undermining traditional communities and facilitating new ones); Henry H. Perritt, Jr., *Dispute Resolution in Electronic Network Communities*, 38 VILL. L. REV. 349 (1993) [hereinafter Perritt, *Dispute Resolution*].

governance of such networks, including the resolution of such disputes. Figure 1 shows three common disputes.

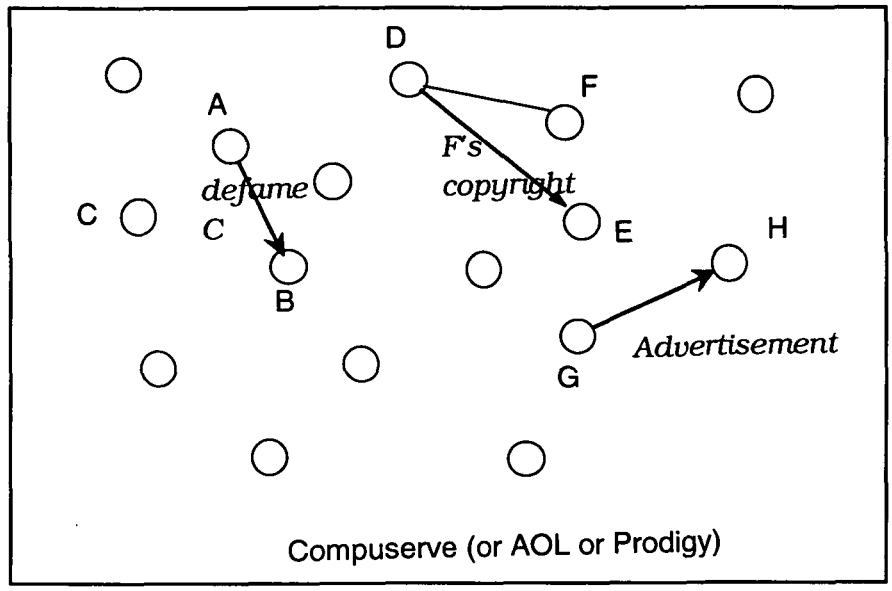

Figure 1.[8]

What is new is that a growing proportion of communications is taking place across the boundaries of proprietary network systems like CompuServe, America Online, and Prodigy, as shown in Figure 2.

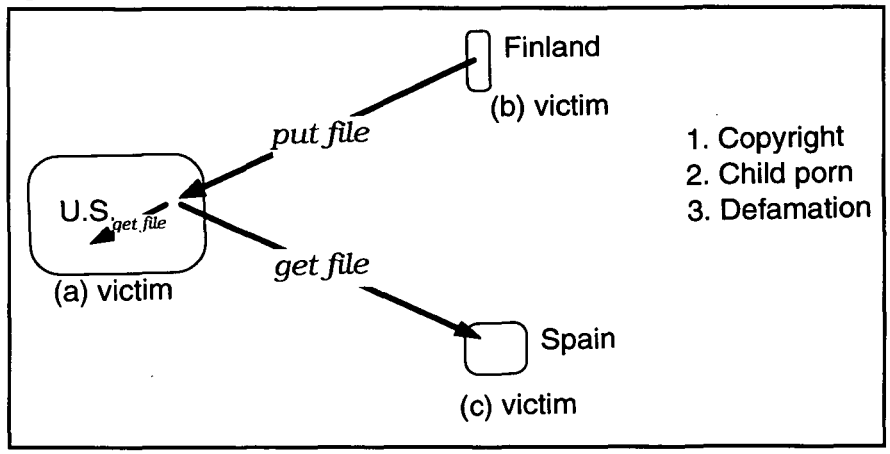

Figure 2.

8. One network user, "D," may take intellectual property belonging to another, "F," and send it to a third party, "E." One network user, "A," may defame another, "C," in a communication to "B." "G" may offend "H" by sending him an unsolicited advertisement.

A person in Finland can place a file on a computer in the United States, which can then be retrieved by someone in Spain, or in the United States. The file could infringe copyrights, contain child pornography, or be defamatory. The victim might be in the United States, in Finland, or in Spain. This means that any one network-administration entity has lost control over the activities that may give rise to controversy. In the Figure 1 scenarios, the network service provider could expel the wrongdoer. In the new open architectures, however, any one network service provider may not even know who the wrongdoer is, let alone have any control over the resource the wrongdoer wants, and the deprivation of which would represent an effective sanction.

On the one hand, the shift to open networks invites self-government, because the capacity of traditional, nationally based legal institutions to regulate the problems illustrated in Figure 1 is diminished by the transnational character of such networks. On the other hand, the capacity of the service providers—the logical organizers of regimes of self-government—is also lower in such networks because the service providers do not have the same control as they did over their own proprietary networks.

III. IS SELF-GOVERNANCE DESIRABLE?

Merely because it is conceivable that electronic communities might be self-governing, and because models for self-governance exist within recognized theoretical frameworks, does not mean that self-governance is desirable. Subcommunities within larger legal communities exercise powers of self-governance for one or more of four reasons: self-governance is more efficient; the rules and/or the adjudicatory techniques for applying the rules need to be different from those of the surrounding community; it is impracticable to apply the rules of the surrounding community; or compliance with basic norms of the community is higher when members of the subcommunity participate in self-governance. Each of these criteria can be evaluated separately with respect to the three basic activities of governance: rulemaking (legislation), rule application (adjudication), and enforcement. More or less autonomy may be appropriate depending on which of these factors one considers.[9] The following analysis of the criteria argues that self-governance is desirable for electronic communities.

---

9. The four justifications stated above are not mutually exclusive. For instance, efficiency concerns surface when one considers any of the other justifications. Moreover, the fourth justification (voluntary compliance) is a way of dealing with the third (unenforceability).

A. Self-governance may be more efficient

Self-governance may simply be a more efficient way of making and enforcing specialized rules and of enforcing rules of the larger legal system. The electronic community can enforce a norm that is supported by broad consensus in the larger society about what substantive rules ought to apply to conduct. When this occurs, it is easy for the larger legal system to defer to a self-governing community, because community institutions make exactly the same decisions that the traditional legal institutions would. It makes no real difference whether internal institutions or traditional institutions apply it; the law is the same either way.

Self-governance by electronic communities may be more efficient than governance directly through larger community mechanisms because of the inherent availability of more efficient communication technologies in electronic communities. Proposals for new rules can be published almost instantly to members of the electronic communities, and they or their representatives can debate the desirability of the proposed rules without having to assemble physically. Application of existing rules can also be more efficient using information technology because of easier detection, prompter notice, and electronic adjudication of rule violations. Indeed, an adjudicatory tribunal to hear arguments and evidence can be convened electronically. Further, the tribunal could deliberate electronically (when multiple decision-makers are involved), and make its decisions known electronically. Electronic communities may also have greater efficiency in imposing sanctions for rule violation because of the ease with which a violator can be denied access to electronic community resources.[10] A "judgment" can be executed simply by invalidating a user's password for a closed system and by removing her Internet address from routers in open systems.

B. Networks need different rules and procedures

Self-governance is desirable if different rules or adjudicatory outcomes for electronic communities are important, compared with the surrounding communities. This criterion is met when the matters addressed by self-governance are highly specialized. Specialization militates in favor of deference to the electronic community by the traditional institutions. Traditional institutions are unable, as a practical matter, to take the time to master the complexities of the specialized subject matter.

---

10. *But see infra* Part VII.A.3 (describing limited sanctions available in electronic communities).

The second criterion is also met when nobody in the traditional community really cares what the internal community does. For instance, while no one in the larger society or political system really cares about the rules for earning merit badges in the Boy Scouts, or the seniority rules in a collective bargaining agreement, members of the internal communities do. When no one cares, it is easier to defer to self-governance.

1. RULES:

Both the need for, and the indifference to, specialized rules exist with respect to cyberspace. The most obvious example relates to purely technical issues, such as enhancements to basic e-mail, Internet routing, Web protocols, and to netiquette rules of subject matter for newsgroups.

The case is hard to make, however, that members of electronic communities should be subject to different rules with respect to conduct that causes harm outside their own communities. It is implausible to assume that no one in the surrounding community will care when members of the electronic community cause harm beyond the electronic community boundaries. The larger community will certainly insist that its rules, intended to address harm to its members, be enforced within electronic communities as well as elsewhere. Clear examples are copyright infringement and use of domain names that conflict with trademarks.

A second possibility for independent rules involves contract formation. The members of an electronic community could agree that contracts for the sale of goods or services could be formed in a particular way.[11] For example, members could agree that exchange of electronic data interchange (EDI) transaction sets, or of tokens satisfying a predefined standard (for authentication using public key encryption)[12] forms a contract.

---

11. *See* RESTATEMENT (SECOND) OF CONTRACTS § 30 (1979) (permitting the offeror to specify how the offer may be accepted). Under this rule, offerors in an electronic network community could all specify the same manner of acceptance. The result would be the same as a contract formation rule for the community, such as discussed in the text.

12. "Public key encryption involves mathematical algorithms that factor large numbers. Through the use of appropriate algorithms, it is possible to obtain two numbers, called *keys*, one of which creates an encrypted message from plain text, and the other of which recovers the plain text from the encrypted version. One of these keys is held by a user of the technique and not disclosed to anyone else. This is called that user's *private key*. The other number, a key associated with the private key, is disclosed publicly. This is that user's *public key*. The public and private keys can be used together either to protect privacy in the content of a message, or to construct digital signatures ... or both."

A third possibility would be to have specialized rules for potentially offensive communications including obscene and pornographic communications. The rules could require that such communications be directed to particular parts of the electronic community, access to which is limited so as to admit only those over a certain age.[13]

Fourth, rules for payments could prescribe how offers, acceptances, and payment orders are to be authenticated, and how the risk of forgery and insolvency are to be borne. The result would be like bank clearinghouse rules.

2. *PROCEDURES:*

Electronic communities need specialized adjudicators to produce better results at a lower cost. Different adjudicatory results could arise from the use of specialized decision-makers in electronic communities. Specialized adjudicators could understand specialized rules for electronic communities better than adjudicators in larger communities who have less contact with the specialized rules. Specialized adjudicators could also understand particular factual contexts within which disputes arise over rule application. For example, a dispute might arise over loss-allocation under an electronic community payments rule. In such a case, an adjudicator who understands public key encryption would be in a better position to appreciate fault in the handling of public and private keys resulting in a forgery. Or, a specialized adjudicator could appreciate the failure of the manager of a reserved area to give notice of potentially offensive contents in the manner prescribed by the community rule. In such a specialized "zoning" case, knowledge of the workings of the boundaries of newsgroups and Web spaces would improve decision-making.

C. Open networks escape enforcement of conventional rules

Self-governance may be desirable because it is impractical to apply rules of larger communities.[14] One situation in which the impracticability

---

HENRY H. PERRITT, JR., LAW AND THE INFORMATION SUPERHIGHWAY 394 (1996) (emphasis in original) [hereinafter PERRITT, INFORMATION SUPERHIGHWAY].

The public and private keys can be used together either to protect privacy in the content of a message, or to construct digital signatures ... or both." *Id.* at 394.

13. *But see* Reno v. American Civil Liberties Union, 117 S.Ct. 2329, 2336-38 (1997) (stating that such verification mechanisms were "effectively unavailable to a substantial number of Internet content providers." (citations omitted)).

14. A similar situation led to the development of certain rules in admiralty. *See* Gordon W. Paulsen, *An Historical Overview Of The Development Of Uniformity In International Maritime Law*, 57 TUL. L. REV. 1065, 1066-67 (1983) (history of admiralty shows

criterion is satisfied is when the boundaries of electronic communities cross the geographic boundaries of traditional sovereigns. This occurs with rapidly increasing frequency as the Internet becomes the model for computer networks, handling a wide variety of commercial and personal communications and delivering commercially valuable information and services. In such network communities, harm occurring in one geographically defined jurisdiction frequently results from conduct occurring in a different geographically defined jurisdiction.

Internet transactions regularly cross national boundaries. [15] Such cross-border communications raise questions of the enforceability of export restrictions, the limitations on public access to public information, intellectual property protection, and the liability for injurious content.

The international nature of these transactions create problems that cannot be dealt with by traditional legal systems. Even if a jurisdiction in which the injury occurs asserts jurisdiction and chooses a plausible body of substantive law, it may lack the means of enforcing its decision, because the actor is somewhere beyond its reach. When conduct traditionally considered criminal is involved, the problem is more acute because of the absence of transitory crimes in traditional jurisprudence.[16] It is unusual for geographically defined legal systems to prosecute for crimes committed in other places, except by artificially redefining the place of commission to be the place of injury.

---

that the motivation for a separate legal system was the need of commerce for international uniformity).

15. This international characteristic is true not only of the Internet; it is also true of multinational businesses. However, the Internet poses greater problems for traditional law enforcement because it permits the effects of conduct occurring elsewhere to be felt within a traditional state without any conduct occurring in that state. Usually a multinational business has some physical presence in the state where its effects are felt.

16. *See* State v. Jones, 443 A.2d 967, 970 (Md. Ct. Spec. App. 1982) (courts of one state may not hear prosecution for crime committed against laws of another state); Bruce Church, Inc. v. United Farm Workers, 816 P.2d 919, 926 (Ariz. 1991) (distinguishing civil and criminal practices); State v. Miller, 755 P.2d 434, 436 (Ariz. 1988) (international law determines whether state may impose criminal penalties for conduct occurring elsewhere). *But see* Lauritzen v. Larsen, 345 U.S. 571, 585 (1953) (law of flag covers even criminal conduct under maritime law); United States v. Noriega, 746 F. Supp. 1506, 1512 (S.D. Fla. 1990) (state may criminalize conduct occurring elsewhere but having effects in prosecuting jurisdiction); Rios v. State, 733 P.2d 242, 244 (Wyo. 1987) (state may prosecute for child custody offense committed elsewhere by actors never within state when effects are felt on custodial parent in state); State v. Mazzadra, 258 A.2d 310, 314 (Conn. 1969) (holding that theft of automobile was a transitory crime for which defendants could be prosecuted in Connecticut although the theft occurred in New York).

Electronic network communities, on the other hand, may find it much easier to enforce rules. For electronic networks in which the attachment is primarily social, threat of exclusion from the network may be a powerful enough incentive to induce compliance with the rules. For electronic networks in which the attachment is primarily economic, the growing availability of low transaction cost methods of making payment potentially facilitate enforcement. A large producer may be required to post a bond, as for certification authorities used in digital authentication systems. Smaller participants, like consumers, are likely to have credit on the network, either through having paid for cybermoney or having arranged for secured electronic credit card transactions. A condition of network participation could be that the consumer must place some of this credit at risk in order to enable fines or civil penalties to be imposed through appropriate adjudicatory procedures.

Self-governing electronic communities can use these methods to deal with conduct occurring in their communities regardless of the place at which it occurs. Also, electronic communities can impose punishments and effectuate compensatory remedies regardless of the geographic place where the community member engaging in the conduct violating the rule is found.

D. Self-governance promotes voluntary compliance

An important advantage of democratic or other representative political systems from a utilitarian perspective, is that they are more likely than authoritarian systems to induce voluntary rule-compliance by citizens. This is so not only because of greater participation by those bound by the rules, but also because specialized rules are less likely to produce bizarre results than general rules drawn from traditional communities. Compliance with the rules imposed by surrounding legal systems may be low in certain electronic communities because the rules are not perceived as fitting the realities of the communities, or because enforcement of the rules by the regular legal institutions is impractical. In either or both of these situations, compliance may improve with standards of behavior that are acceptable to (while not identical to the rules of) the surrounding community, if a measure of self-governance is allowed to participants in the electronic communities. For instance, electronic community participants who wish to exchange messages containing potentially offensive content, might be willing to comply with rules requiring clear notice, the exclusion of minors, and other restrictions on access, while they would be unwilling to comply with prohibitions on exchanging such messages.

E. Conclusion

The likelihood of autonomy for electronic communities is greatest when specialized rules and adjudicators are needed, and traditional communities are indifferent to their content. In such situations, the inherent likelihood that a specialized legal system will be more efficient, that it will induce greater voluntary compliance, and that it will regulate behavior that otherwise would escape regulation, tilt the political balance in favor of autonomy.

IV. IS SELF-GOVERNMENT LEGALLY FEASIBLE?

All modern legal systems proceed from the foundational premise that only entities possessing sovereignty can make, apply, and enforce law. Despite the association of sovereignty with national governments, the reality is that governance is dispersed among a rich variety of public and private institutions. Most people in industrialized society work for employers who administer private systems of workplace governance. Most money moves in complex clearinghouse systems set up and administered by private banks. Most industrial production and commerce takes place in private contractual webs. Much social and religious life transpires in private associations. The increased importance of international human rights, trade, and environmental law has drawn upon the energy and expertise of thousands of non-governmental organizations (NGOs), such as Amnesty International and Greenpeace, to provide information and analysis to treaty based institutions.

In theory, however, these private governments derive their power from the traditional sovereigns and are always subject to the sovereign imposing new regulations and enforcing them. The relationship between private governments and traditional sovereigns is determined by traditional laws or regulations enacted by traditional sovereigns, by constitutions defining the power of traditional sovereigns, or by international treaty.

Self-governance can be realized in at least three forms: (1) immunity from the application of surrounding legal standards, (2) immunity from the enforcement power of traditional legal institutions, and (3) recognition of the prescriptive and adjudicatory acts of the autonomous community.

The feasibility of self-government depends on the traditional community's respect for, and deference to, community law. Traditional communities generally respect party autonomy exercised through contractual agreements. Thus, respect for community law can be earned by having a contractual framework for electronic communities. Although the enforcement of contractual autonomy may depend on traditional institutions, such dependence can be mitigated by internalizing

enforcement. The crucial elements of a self-governing community are completeness, the availability of coercive power to enforce community decisions, and a contractual framework expressing the norms, procedures, and institutional competencies.

A. Basic legal frameworks

Assessing the feasibility of autonomy for private electronic communities requires an understanding of the points of tangency between these communities and traditional sovereigns—the scenarios in which someone challenges the autonomy of the private community in the courts of traditional sovereigns. Ultimately, autonomy for private communities depends on comity being afforded them by traditional sovereigns,[17] which in turn is more likely if the private communities are "complete," in the sense that they offer the entire spectrum of rulemaking, adjudication, and enforcement.

Political autonomy originates in physical power. Nation-states are politically autonomous because they have the military power to keep themselves that way. The geographic scope of political units has historically depended on the reach of military technology, and the social cohesion necessary to use it. Sovereignty is formally associated with nation-states that have the practical ability to assert physical power to coerce compliance with their law within defined borders and with respect to a defined class of persons.[18] New nations, such as Bosnia-Herzegovina,

---

17. "Comity ... is the recognition which one nation allows within its territory to the legislative, executive or judicial acts of another nation, having due regard both to international duty and convenience, and to the rights of its own citizens or of other persons who are under the protection of its laws." Joel R. Paul, *Comity in International Law*, 32 HARV. INT'L L.J. 1, 8 (1991). Mr. Paul criticizes comity as an imprecise concept, meaning little more than choice of law to some analysts, a discretionary doctrine of public international law to others, and a basis for insisting on reciprocity for still others.

18. *See* RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW OF THE UNITED STATES § 201 (1987) ("Under international law, a state is an entity that has a defined territory and a permanent population, under the control of its own government, and that engages in, or has the capacity to engage in, formal relations with other such entities.").

Under international law, a state has:

(a) sovereignty over its territory and general authority over its nationals;

(b) status as a legal person, with capacity to own, acquire, and transfer property, to make contracts and enter into international agreements, to become a member of international organizations, and to pursue, and be subject to, legal remedies;

(c) capacity to join with other states to make international law, as customary law or by international agreement.

are created, and old nations such as the Soviet Union, disappear, but the birth of a sovereign state is a momentous occasion in diplomacy and international law.

Network communities quite clearly are not entitled to status as traditional sovereigns because they lack a defined territory, a permanent population, and mechanisms for exerting physical coercive power.[19] But new sovereigns can be created by delegation of power from traditional sovereigns. The European Union and the United States came into existence as sovereign entities through delegation of powers from nation-states through treaties and constitutions. Such delegation for network communities is, however, unlikely to occur within the foreseeable future.

Smaller, more or less autonomous communities, have long existed within power-maintained sovereign political units. Their establishment and continued existence has always depended on the sufferance of the sovereign. For example, fairs, cities, universities, and guilds existed under English law because of the grant of patents from the King. The patent defined the powers of the community it authorized.[20] This type of

---

Id. at § 206. "Sovereignty" is a term used in many senses and is much abused. As used here, it implies a state's lawful control over its territory generally to the exclusion of other states, authority to govern in that territory, and authority to apply law there. "The sovereignty of a state is reflected also in immunity for the state and its public property from certain exercises of authority by other states." *Id.* at § 206, cmt b.

19. "The Second Circuit has limited the definition of 'state' to entities that have a defined territory and a permanent population, that are under the control of their own government, and that engage in, or have the capacity to engage in, formal relations with other entities." Kadic v. Karadzic, 70 F.3d 232, 239 n.2 (2d Cir. 1995) (quoting Klinghoffer v. S.N.C. Achille Lauro, 937 F.2d 44, 47 (2d Cir.1991)).

20. "[T]he settlers had emigrated from an England that was localist in political organization: early seventeenth-century English towns, boroughs, counties and guilds still operated to a great extent as self-governing (although partially overlapping) entities." Jeremy Elkins, *University of Chicago Law School Roundtable Conference: Constitutions and "Survivor Stories" Declarations Of Rights*, 3 U. CHI. L. SCH. ROUNDTABLE 243, 255 (1996); *see also* Joan C. Williams, *The Invention Of The Municipal Corporation: A Case Study In Legal Change*, 34 AM. U. L. REV. 369, 374 (1985):

> Groups that the law identified as aggregate corporations seem unrelated to the modern eye: chartered boroughs, companies of merchants, including guilds, and universities. What did these groups share that caused them to be identified as corporations, while other groups, such as villages and towns, were not? The answer is that 'incorporated' entities were corporations because they shared a special relationship to feudal society: each of the major English 'corporations' developed from the late feudal practice of granting charters to groups that wanted to 'opt out' of feudal obligations. This division between groups that were corporations and

delegation exists in modern legal systems in the form of corporate and municipal charters. More generally, other types of private communities exercise a form of sovereignty under contracts mutually delegating attributes of sovereignty retained by the community members. The traditional sovereign allows this kind of private sovereignty by allowing freedom of contract. However, such community autonomy is dependent on the traditional sovereign for its very existence. It exists only if traditional sovereign institutions recognize the community's autonomy.

Let us examine what would happen if there were to be a clash between an autonomous electronic community and the traditional community in which it sits. Assume a deputy sheriff shows up at the door to seize a computer, to demand copies of certain files, or to arrest a natural person for certain conduct. Members of a network community are unlikely to prevail in physical resistance to the deputy sheriff. The law enforcement agent almost always can call upon superior force. The community enjoys autonomy only because it can claim privileges or immunities recognized by the traditional sovereign. For instance, if the deputy sheriff intrudes, the autonomous community prevails in a subsequent legal proceeding for trespass, conversion, or violation of civil rights.

But community autonomy is rarely a bilateral test between the traditional sovereign and the private community. It usually involves a three-way contest between private interests in which the traditional sovereign is the arbiter. The deputy sheriff in the hypothetical was sent by a court and acted pursuant to a writ or warrant. The warrant or writ was issued by a traditional sovereign's court on the request of a private plaintiff or public prosecutor acting on a private complaint. The prosecutor acted pursuant to authority granted by the traditional sovereign. The private plaintiff may be from within the community or from outside it. The community-autonomy question may have been tested long before the sheriff showed up at the door of the electronic community. In such cases, one party seeks to deny community autonomy. For example, an electronic association sued by a present member asserting a violation of its constitution, would defend on the grounds that the court in which the suit is filed must defer to internal

---

groups that were not was the second anachronistic aspect of English corporation law.

See also Joel Edan Friedlander, *Corporation And Kulturkampf: Time Culture As Illegal Fiction*, 29 CONN. L. REV. 31, 76 (1996) (explaining the conflict between the view that groups such as corporations enjoy status as actual person and the orthodox view that they are artificial persons with only such existence as is recognized by the traditional states).

tribunals on constitutional issues.[21] In other words, it asserts adjudicatory autonomy, while the private plaintiff asks the traditional court to deny autonomy of the private tribunals by deciding the case on the merits. A member of the electronic community accused of intellectual property infringement would defend on the grounds that the court in which the lawsuit is brought must defer to community rules which grant her a privilege with respect to the intellectual property, while the plaintiff claims that any contractual privilege is voided by traditional sovereign law on intellectual property. A criminal defendant would defend on the grounds that the traditional criminal statute should be interpreted with deference to electronic community rules, while the prosecutor on behalf of the victim would argue that no such deference is appropriate.[22] Less conventionally, the defendants may assert that they are immune from suit or from prosecution because of their community membership and the nature of the claim, while the plaintiffs appeal to traditional sovereign authority and deny the existence of immunity.[23]

---

21. *See, e.g.*, Blackshire v. NAACP, 673 N.E.2d 1059, 1061 (Ill. App. 1996) (reversing trial court for inappropriately interfering in internal affairs of private association; law of private associations requires judicial deference to authorized decisions of internal bodies); Georgopoulos v. Teamsters, 942 F. Supp. 883, 895 (S.D.N.Y. 1996) (holding that federal statute does not authorize judicial intervention into internal union affairs except when necessary to enforce minimum statutory standards).

22. *See, e.g.*, United States v. Morris, 928 F.2d 504, 509 (2d Cir. 1991) (defining criminal conduct in terms of what is authorized by private computer system). The suggested defense also might arise if a pornography prosecution were defended on the grounds that the electronic community defines the standards and that the material was not pornographic under those standards.

23. There are several types of immunity. The most basic type is sovereign immunity. Such immunity was rooted in the "perfect equality and absolute independence of sovereigns." Schooner Exchange v. McFaddon, 11 U.S. (7 Cranch) 116, 137, (1812). In recent decades, the former absolute view of sovereign immunity has evolved into a restricted view, which accommodates the reality that many sovereigns engage in commercial activities, as to which they should not necessarily be treated as states. *See generally* RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW § 451 intro. note (1986) (providing immunity to states from the jurisdiction of the courts of other states, except for "claim arising out of activities of the kind that may be carried on by private persons"). Considerations of judicial administration supplement international law in immunizing certain witnesses and chattels from service of process or execution. *See* RESTATEMENT (SECOND) OF CONFLICT OF LAWS § 83 (1969). Charities historically were immune from tort liability based on the rationales that their resources should not be diverted from charitable purposes, that the doctrine of respondeat superior was inapplicable, or that persons accepting benefits from charities waived tort claims. *See* RESTATEMENT (SECOND) OF TORTS

Such trilateral contests arise in several relevant contexts:

- conduct by community members that harm the legally recognized interests of nonmembers (e.g. defamation or intellectual property infringement);[24]
- intra-community conduct that offends non-waivable traditional community standards (e.g. racial, gender, or disability-motivated adverse decisions, or life- or personal-injury-threatening conduct within the community);[25]
- denials of membership under circumstances that would constitute a legal wrong under traditional law,
- expulsions from membership and post-expulsion efforts to collect fines or penalties from past members that would either offend traditional legal standards or would necessitate resort to traditional legal institutions for enforcement.[26]

In several of these examples, it is a member of the autonomous community who seeks to avoid self-governance. A community member may go "outside" because she thinks that traditional institutions, procedures, or substantive law will give her a better result on an access, authorship, or authentication issue.[27] For example:

- a present member of the association files a lawsuit in a traditional court asserting breach of contract based on an alleged violation of the association's constitution;

---

§ 895(e) & cmts. (1979) (reviewing history and justifications for immunity and repudiating it as a general rule).

24. *See, e.g.*, United States v. LaMacchia, 871 F. Supp. 535, 536-37 (D. Mass. 1994) (example of the use of anonymous file transfer protocol area to exchange software violating copyrights of non-participants); Cubby, Inc. v. CompuServe, Inc., 776 F. Supp. 135, 139-42 (S.D.N.Y. 1991) (use of computer service to defame non-participant).

25. *See, e.g.*, United States v. Alkhabaz, 104 F.3d 1492, 1493 (6th Cir. 1997) (use of electronic mail system to discuss abduction of classmate).

26. *See, e.g.*, Cyber Promotions, Inc. v. Apex Global Info. Serv., Inc., No. Civ.A 97-5931, 1997 WL 634384 (E.D. Pa. Sept. 30, 1997) (granting preliminary injunction against termination of service to mass mailer in violation of Internet access service contract); CompuServe, Inc. v. Cyber Promotions, Inc., 962 F. Supp. 1015, 1028 (S.D. Ohio 1997) (granting preliminary injunction enjoining mass mailer from sending unsolicited advertisements to subscribers); Cyber Promotions, Inc. v. America Online, Inc., 948 F. Supp. 456, 459-60, 464-65 (E.D. Pa. 1996) (denying preliminary injunction against use of tool allowing subscribers to block junk e-mail).

27. *See* Perritt, *Community Regained*, *supra* note 7, at 991 (explaining that controversies over access, authorship, and authentication are the major ones requiring legal attention as the national information infrastructure develops). Authentication includes electronic signatures and other protections against forgery and repudiation of legally significant messages.

- a member of the association files a lawsuit in a traditional court alleging that the association's conduct violates a statute newly-enacted by a traditional legislature which explicitly applies to association conduct.

Outsiders also have grievances against communities and their members. What happens when an outsider wants access but is denied? What happens when an outsider infringes intellectual property generated within the electronic community? What happens when an outsider masquerades as a member and gets involved in an authentication controversy? Some examples of this are:

- a nonmember brings a lawsuit against a member of the association for intellectual property infringement;
- a prosecutor from a traditional jurisdiction commences a criminal prosecution against an association member for intra-association conduct that prima facie violates a traditional criminal statute.

The need to define boundaries between traditional sovereigns and autonomous private communities is analogous to the need to resolve inter-sovereign conflicts in international law.[28] The question of community autonomy in cyberspace depends upon whether the court to which the claim is presented defers to community law, either by recognizing an immunity for a particular defendant (unlikely) or by recognizing community substantive or procedural law. Often the question of autonomy is a choice of law question. Should the non-community court apply its own law (or the law of another conventional sovereign), or should it apply community law?

It is useful to consider how such situations are dealt with in the international realm. Boundaries of autonomy in international law are defined by the jurisdiction to prescribe,[29] the jurisdiction to adjudicate,[30] and the jurisdiction to enforce.[31] These three types of jurisdiction are useful benchmarks for private communities as well.

---

28. *See id.* at 1009.

29. *See* RESTATEMENT (THIRD) OF FOREIGN RELATIONS LAW § 401(a) (1987) (listing categories of jurisdiction); *id.* at §§ 402-03 (listing bases of and limitations on jurisdiction to prescribe); *id.* at § 461 (immunity of foreign state from jurisdiction to prescribe).

30. *See id.* at § 401(b) (1987) (describing jurisdiction to prescribe); *id.* at § 421 (describing jurisdiction to adjudicate); *id.* at § 451 (describing the immunity of a foreign state from jurisdiction to adjudicate).

31. This tripartite classification of types of jurisdiction is an innovation of the third Restatement of Foreign Relations. *See id.* at § 401 rptr. nt. 2 (1987). The second Restatement subdivided jurisdiction into the jurisdiction to prescribe and the jurisdiction to enforce. *See* RESTATEMENT (SECOND) OF FOREIGN RELATIONS § 6 (1965); *see also* Laker

Respect for electronic community law depends upon the electronic community having a contractual framework sufficient in scope to bind those wishing to avoid the effect of community decisions. Communities stand a better chance of being recognized either as sovereigns or as contractual communities if they offer relatively complete legal systems of their own. Incomplete systems must rely on traditional legal systems to perform the missing functions. To the extent that such external dependence exists, the community is less autonomous. Completeness, and thus autonomy, depends upon the capacity to perform functions essential to any legal system. As Joseph Raz has observed, "[t]he three most general and important features of the law are that it is normative, institutionalized, and coercive."[32] Cyberlaw—the legal system of electronic communities—is eligible for recognition as a separate legal system to the extent that it contains these three features. Electronic communities must offer normative rules for conduct; they must institutionalize rulemaking and rule application; they must sanction rule violators.[33]

Professor Hart observed that legal rules fall into one of two classes: primary rules, which impose duties; and secondary rules, which define powers to make and apply primary rules.[34] Primary rules pertain to the normative dimension. Secondary rules institutionalize and channel coercive forces. Cyberlaw is a complete legal system to the extent that it has both types of rules. Any claim for self-regulation in cyberspace must be tested according to these criteria—the existence of rulemaking, adjudication, and coercive enforcement means.

Airways, Ltd. v. Pan American World Airways, Inc., 604 F. Supp. 280, 292 (D.D.C.), *aff'd*, 731 F.2d 909 (D.C. Cir. 1984) (relating comity to extent of jurisdiction to prescribe, adjudicate, and enforce); Joel P. Trachtman, *Conflict of Laws and Accuracy in the Allocation of Government Responsibility*, 26 VAND. J. TRANSNAT'L. L. 975, 1046 n.288 (1994) (describing the three components of jurisdiction recognized by the Restatement); Christopher C. Joyner & Wayne P. Rothbaum, *Libya and the Aerial Incident at Lockerbie: What Lessons for International Extradition Law?*, 14 MICH. J. INT'L L. 222, 234-38 (1993) (using three bases of jurisdiction to explore extraterritorial application of United States law); Bruce Zagaris & David R. Stepp, *Criminal and Quasi-Criminal Customs Enforcement Among the U.S., Canada and Mexico*, 2 IND. INT'L & COMP. L. REV. 337, 338-42 (1992) (discussing bases of jurisdiction in U.S. and Mexican Law).

32. Joseph Raz, The Concept of a Legal System 3 (2d ed. 1980).

33. Offering normative rules is an assertion of jurisdiction to prescribe. Formalizing mechanisms for rulemaking and rule application involve assertion of jurisdiction to prescribe and to adjudicate, respectively. Sanctioning rule violators asserts jurisdiction to enforce.

34. *See* H.L.A. HART, THE CONCEPT OF LAW 78-79 (1961). Hart's secondary rules define legislative (rulemaking) and adjudicatory institutions and powers.

B. Contract: the framework for autonomy

Private contract is the most appropriate source of autonomy for electronic communities. Indeed, treaties and constitutions, the traditional sources of sovereignty, can be understood as contracts among sovereign states and sovereign people respectively. Most of the examples of private legal communities reviewed in Part IV involve private contractual webs to define the community and to allocate legal power within it.

Much can be done through conventional contracts to set up communities to which sovereigns will defer. Bank clearing house systems, WESTLAW licensing agreements, and collective bargaining agreements are good examples of contractual arrangements that establish internal governance mechanisms for the parties to the contract. Assuming a valid contract can be formed among the members of an electronic community, as discussed in this section, such a contract can achieve the criteria identified by Raz: norms, institutionalization, and coercion. A community contract can set specialized standards for conduct within the community. By providing for arbitration, such a contract can arrange for application of these rules through specialized community institutions. Indeed, it can arrange for on-line, cyberspace-based common law courts through appropriate arbitration clauses. Such a contract can also provide directly for coercive enforcement, by specifying liquidated damages, by requiring posting of a bond against which penalties may be imposed, or by providing for expulsion from the community (with or without forfeiture of property left within the community, e.g., intellectual property). In other words, contracts can provide the framework for a complete legal system.

Most of the models of self-governance[35] (all except the military one) depend upon private contracts as the normative, institutionalizing, and (to a lesser extent) coercive source of law. Even constitutional and international arrangements use documents similar to contracts in some ways to express the delegated powers. Parties to purely private contracts can achieve some immunity from outside legal institutions by waiving application of traditional law and recourse to traditional legal institutions. Thus, contract principles are a natural starting point for the establishment of an independent electronic community.

But the contractual nature of some electronic communities may be problematic.[36] Some electronic communities are anonymous, have

---

35. *See infra* Part VI.

36. Autonomy based on contract requires the presence of the elements of an enforceable contract: capacity to contract, offer, acceptance, and consideration. *See* PERRITT, INFORMATION SUPERHIGHWAY, *supra* note 12, at 379 (explaining the formal prerequisites of contractual obligation).

rapidly shifting membership, and may exist for any particular member only for as long as it takes her to send a request to a World Wide Web or news server and receive an item of information as a response. In these communities, there is no negotiation and no ongoing social relationship. There may be a contract, but it may be so brief in duration that it may be an intellectual stretch to say that the consumer of these services joins a community and agrees to participate in self-government. In this type of electronic community, contractual models associated with standard form contracts unilaterally issued by one party are most relevant.[37] This is because one party in these anonymous electronic communities will almost certainly publish the electronic equivalent of a standard form contract to which the participants will become parties to some extent. Standard form contracts have become inevitable as managerial direction has replaced market forces for a vast range of commercial transactions.[38] Professor Rakoff describes the reality that consumers almost never read the terms of standard form contracts, and that it would be eccentric to insist on changes. The drafting organizations would almost certainly not agree to changes,[39] and neither the drafting organization nor the consumer really expects the lawyer-crafted terms of the standard form to be followed.[40] In these settings, it may be unclear whether someone involved with community resources really is a "member" of the community, subject to its normative rules, institutions, and enforcement mechanisms, and within any shield of immunity and deference.[41]

In these circumstances, Professor Rakoff and others have suggested new rules of contract enforcement. Professor Slawson suggested that the standard terms in a form contract be enforced only when they are consistent with the reasonable expectations of the parties.[42] Slawson would determine the reasonable expectations according to the nature of the transaction. Both Slawson and Rakoff reviewed leading cases

---

37. *See* W. David Slawson, *Standard Form Contracts and Democratic Control of Lawmaking Power*, 84 HARV. L. REV. 529, 530, 532 (1971) (estimating that standard form contracts account for 99 percent of all contracts made).

38. *See* Todd D. Rakoff, *Contracts of Adhesion: An Essay in Reconstruction*, 96 HARV. L. REV. 1173, 1223 (1983).

39. *See id.* at 1225.

40. *But see* ProCD v. Zeidenberg, 86 F.3d 1447, 1455 (7th Cir. 1996) (enforcing the terms of shrink-wrap license).

41. *See* Perritt, *Dispute Resolution*, *supra* note 7, at 352.

42. *See* Slawson, *supra* note 37.

explaining how their suggestions are not revolutionary departures from what courts actually do when confronted with standard form contracts.[43]

Making standard contracts unenforceable, however, engenders uncertainty. A better approach may be to construct a different contract regime, in which contract terms posted in some formal way and subject to review or challenge might be presumptively valid, but not otherwise. This approach borrows the concept from insurance regulation that the standard contract is generally subject to review by the insurance commissioner before it can be used with purchasers of insurance. It also borrows from ERISA,[44] which requires that employee benefits plans be published and filed with the Department of Labor. Whenever someone offers a contract defining a self-governing electronic community, that person can specify the way in which the offer is to be accepted and can also indicate what sort of an exchange is sought by the offeror (by the content of the offer and the circumstances under which it is made).[45] The offer can specify that it may be accepted by conduct (for example, hitting the enter key on one's computer), or by making a promise (such as giving a credit card number representing an implied promise to pay a stated subscription fee).

When the party to whom the offer is addressed (the offeree) engages in the specified conduct or makes the promise, she accepts the offer.[46] Typically, this conduct or promise also constitutes the offeree's half of the desired exchange, frequently called consideration.[47]

C. The limits of contract

Notwithstanding the power of contract, the contract theory has important limitations as a source of community autonomy. One limitation is political and another is legal. Politically, contractual communities are porous and may be impermanent, as compared to

---

43. *See* Rakoff, *supra* note 38.

44. *See* 29 U.S.C. §§ 1021-24 (1994) (requiring publication and filing of employee benefit plans).

45. *See* RESTATEMENT (SECOND) OF CONTRACTS § 30, cmt. a ("The offeror is the master of his offer .... The terms of the offer may limit acceptance to a particular mode."); *id.* at § 60 ("If an offer prescribes the place, time or manner of acceptance its terms in this respect must be complied with in order to create a contract.").

46. *See, e.g., id. at* § 32, cmt. a ("In case of doubt an offer is interpreted as inviting the offeree to accept either by promising to perform what the offer requests or by rendering performance, as the offeree chooses.").

47. *See* JOHN EDWARD MURRAY, JR., MURRAY ON CONTRACTS, 51 (1974) (consideration noted among six essential elements to formation of a contract).

sovereign communities. Private disputes tend to drift into public forums. For instance, the private character of collective bargaining has been significantly eroded by an expansion of legislatively defined individual employee rights, enforced by public institutions.[48] Injured persons seek relief in whatever forums seem most likely to produce the relief they desire. When an injured person is outside an electronic community, that person will probably press for relief from traditional institutions.

This limitation of contract can be mitigated to some extent by internalizing the enforcement function. Then enforcement of private community norms does not depend on the willingness of a traditional court to enforce a contract; the private community enforces it directly. Internalization of the enforcement function reduces the dependence of the self-governing community on traditional legal institutions to enforce its decisions, although it may increase the possibility of legal liability in traditional forums for the injury resulting from internalized enforcement.[49] The possibilities for internalized enforcement are greatly enlarged by the possibility of the privatization of Internet domain registration, under the International Ad Hoc Committee (IAHC) recommendations.[50] Someone who does not obey the rules or who flouts a decision can be denied an Internet domain name, effectively excluding him from the Internet (or at least from the part of the Internet within the scope of that domain registry).

The legal limitations on contracts concern the extent of comity; limits on the prescriptive jurisdiction of the private community. As Judge Posner once wrote, "If a consent decree provided that a violator could be punished by having his ears cut off, the judge could not sign it."[51] Despite the strong tendency for courts to enforce private arbitration agreements and arbitration awards, they are not enforceable when they contravene public policy.[52] Moreover, it is not clear that private arbitrators may be given authority to award punitive damages.[53] Further,

---

48. *See* Henry H. Perritt, Jr., Employee Dismissal Law and Practice 191-272 (3d ed. 1993).

49. Excluding someone from a subnetwork may constitute a breach of contract. Blocking someone's messages may be a tort. Collective enforcement may be a combination in restraint of trade in violation of the antitrust laws. *See infra* Part VII.B.1.

50. *See infra* Part V.C.

51. Donovan v. Robbins, 752 F.2d 1170, 1176 (7th Cir. 1985).

52. *See, e.g.*, United Paperworkers v. Misco, 484 U.S. 29, 43 (1987) (reversing refusal to enforce arbitration on public policy groups, but stating general principle).

53. See John Y. Gotanda, Awarding Punitive Damages in International Commercial Arbitrations in the Wake of Mastrobuono v. Shearson Lehman Hutton, Inc., 38 HARV.

contracts "in restraint of trade" are unenforceable,[54] and conduct undertaken pursuant to private contractual arrangements may produce tort liability.[55] The boundaries of self-government are determined by the scope of such liability, and by the limits of contract enforceability. Many of the limitations discussed above can be addressed by carefully designing a contractual system with all of the features necessary for a completely legal system, as per the Raz formulation. Hence, despite these limitations, contracting is the best way to achieve autonomy for electronic communities.

V. THREE EXAMPLES OF CYBERSPACE SELF-GOVERNMENT

Several forms of self-government already exist in cyberspace. Others have been proposed. The existing forms range from the (mostly) benevolent dictatorships exercised centrally by the proprietors of America Online (AOL) and Compuserve, to the much more loosely organized "netiquette" of Internet newsgroups. The newsgroup, alt.current-events.net-abuse (a.c.e.n.a.), is one of the most highly developed examples of the latter type of self-government. Proposals for a private, international mechanism for domain name registration by the International Ad Hoc Committee (IAHC) provide a comprehensive framework for self-government on a larger scope than has heretofore been experienced or proposed. Examination of both a.c.e.n.a. and IAHC's frameworks, and their comparison with the "royalist" frameworks operated by proprietary services, reveals some of the issues that must be confronted in any system of private self-government for cyberspace.

The legal frameworks for all three examples are contractual, explicitly in the case of America Online and the IAHC recommendations, and implicitly in the case of a.c.e.n.a. The contractual frameworks in all three cases are complete, in that they authorize rulemaking, adjudication, and enforcement. They contemplate coercive measures: termination of service by AOL, expulsion from the Internet by revoking domain names

---

INT'L L.J. 59, 61 (1997) (explaining that punitive damages in arbitration are allowed under United States law but not under many foreign sovereigns).

54. *See* 15 U.S.C. § 1 (1994).

55. *But see* Cubby, Inc. v. CompuServe, Inc., 776 F. Supp. 135, 141 (S.D.N.Y. 1991) (unsuccessful action against electronic information service for defamatory statements made under contractual arrangement); Zeran v. America Online, Inc., 958 F. Supp. 1124, 1128-35 (E.D. Va. 1997) (finding negligent defamation claim against information provider preempted; because statements made by party under contract with defendant); Religious Technology Center v. Netcom Online Communications Co., 907 F. Supp. 1361, 1383 (N.D. Cal. 1995) (denying summary judgment on contributory copyright infringement claim based on material posted under contract with defendant).

in the IAHC recommendations, and direct "killing" of messages in the case of a.c.e.n.a.

A. A.c.e.n.a.: an example of self-government[56]

A.c.e.n.a. is a democratic mechanism for enforcing certain netiquette rules against undesired postings popularly called "spam." The a.c.e.n.a newsgroup and several other conversation "forums" exist on a world wide conference and exchange system called USENET. Although often mentioned in conjunction with the Internet, USENET is a distinct system of cooperating Internet nodes; not all Internet nodes participate in USENET.[57] Each newsgroup addresses a particular subject. USENET developed a body of rules or conventions accepted by most traditional users, often referred to as "netiquette."[58] Commercial advertising violates netiquette, and frustrates the intent of USENET to channel communications into subject-specific groups.

Despite this netiquette convention, on April 12, 1994, the Phoenix-based law firm of Cantor & Seigel (C&S) sent a message (often called a "post") advertising its legal services to thousands of newsgroups. The response was virtually instantaneous, as thousands of users voiced their disgust in discussions on newsgroups such as news.admin.misc. and news.admin.policy. USENET subscribers were outraged by the

---

56. The research and initial drafting of this section was done by Sean P. Lugg, Villanova University School of Law, Class of 1996, December 19, 1997. Mr. Lugg is a law clerk to the author. For background information on a.c.e.n.a., see Scott Southwick, *The news.admin.net-abuse FAQ File* (visited Nov. 23, 1997) <http://www.bluemarble.net/~scotty/nana-history.html> and Scott Southwick & J.D. Falk, *The Net Abuse FAQ* (visited Nov. 23, 1997) <http://www.cybernothing.org/faqs/net-abuse-faq.html>.

57. USENET is a collection of several thousand discussion groups called "newsgroups." Participants in USENET feed newsgroup updates to each other, so that a human user can add a comment to a newsgroup by "posting it" on his own computer, and the USENET system then propagates that new posting and all others like it to other USENET computers so that within a day or so, the new postings are available on the newsgroup throughout the Internet. There is no entity that owns or controls USENET; it is a collection of cooperating computer administrators.

58. *See generally* Sally Hambridge, *Netiquette Guidelines* (visited Nov. 23, 1997) <http://www.cybernothing.org/cno/docs/rfc1855.html> (summarizing netiquette rules, including general rule against posting messages inconsistent with character of newsgroups or mailing lists).

commercialization of the system. C&S were "flamed"[59] by thousands who alleged, inter alia, violations of USENET conventions, and disregard for netiquette. Unaffected by these protests,[60] and realizing the vast, low cost advertising potential of the USENET,[61] C&S announced their intentions to form an advertising company, Cybersell.[62] Because flaming failed to end the practice, a search for other, more coercive, means was initiated.

A.c.e.n.a. was established on April 25, 1994 to channel concerns about such USENET abuses.[63] The most prominent assailant of identified "spam" is Cancelmoose(TM), an anonymous member of the newsgroup.[64] Operating from a European site, Cancelmoose effectively rids the network of bothersome postings by means of "cancelbots," [65] or cancel messages. A response is posted soon thereafter, notifying the newsgroup of the cancellation. Furthermore, a message is sent to the "spammer," the individual or group who posted the message, notifying him of the action,

---

59. "Flaming" is the practice of besieging an individual with electronic or paper mail to voice disagreement to a posted message.

60. Acting contrary to the beliefs expressed by a consensus of USENET users is a violation of USENET conventions. This proposition is inferable from the general rules of netiquette.

61. C&S were able to transmit their message to approximately 30 million users in less than 90 minutes, with modest cost to the firm.

62. C&S stated that their goal was to make commercial advertising pervasive on the Internet. To accomplish this goal, they planned to create the advertising company Cybersell. *See* Peter H. Lewis, *Arizona Lawyers Form Co. for Internet Advertising*, N.Y. TIMES, May 7, 1994 at A1.

63. The newsgroup alt.current-events.net-abuse was "chartered" on April 25, 1994, less than two weeks after the initial C&S post. A.c.e.n.a. was replaced in November 1996 by the news.admin.net-abuse.* groups. *See* Scott Southwick & J.D. Falk, *The Net Abuse FAQ* (visited Nov. 27, 1997) <http://www.cybernothing.org/faqs/net-abuse-faq.html>.

Although formed for discussion of net abuses generally, "spamming" is the only occurrence which has been deemed "net-abuse" by consensus. Although the definition of "spam" varies, the generally accepted description is "the same article, or essentially the same article, posted an unacceptably high number of times to one or more newsgroups." *Id.*

64. *See id.* Cancelmoose, who now has a home page on the Web—http://www.cm.org/—has left his original activities to others. *See* Scott Southwick & J.D. Falk, *The Net Abuse FAQ* (visited Nov. 27, 1997) <http://www.cybernothing.org/faqs/net-abuse-faq.html>.

65. *See id.*

the reasons for the action, and what steps to take in the future to avoid a similar occurrence.

The readership of a.c.e.n.a consists predominantly of news administrators who can set filters that control the flow of messages to and from the site. Cancelmoose's cancel messages contain a readily detectable signature which enables site administrators to screen the cancels if desired. There is widespread approval of the actions of Cancelmoose by those active in the newsgroup. Furthermore, those disapproving a particular cancel maintain the ability to disregard the cancel messages by reconfiguring the receiving site.[66]

B. Is A.c.e.n.a. fair?

A.c.e.n.a. is interesting because it exhibits attributes of rulemaking (deciding what the norms of acceptable use are), adjudication (deciding whether a particular message violates the norms), and enforcement (Cancelmoose's cancellation of messages determined to violate the norms).

The main questions with respect to rulemaking are not procedural; they concern representation. How does one know that an electronic group like a.c.e.n.a has legitimate rulemaking power? What is the likelihood that the views represented in that discussion group adequately reflect the views of those to be bound by the rules? One answer, of course, is that the a.c.e.n.a. newsgroup is accessible to anyone using the Internet, and only persons using the Internet will be bound by the rules. In other words, anyone who is bound had an opportunity to participate, and failure to participate is not a persuasive argument for not being bound.

The adjudicatory function is somewhat trickier. Needed flexibility could be lost if lawyers (and others) insist on imposing the details of a modern civil procedural system on the adjudicatory process in cyberspace. The usual question in evaluating adjudication is compliance with due process. To avoid the risk of violating due process, it is appropriate to consider the evolution of adjudication in the Anglo-American tradition. Such an evolutionary perspective reveals the flexibility of the due process concept. Initially, the adjudicatory decision-makers were persons with actual knowledge of the facts.[67] The earliest juries had virtually plenary power to decide the case, without the constraints of modern notions of the fact-law distinction, and they also

---

66. Cancelmoose's cancel messages contain identifiers that may be easily recognized and disregarded by proper configurations of the receiving computer system.

67. *See* F.W. Maitland, The Constitutional History of England 46 (1908).

were witnesses to the conduct giving rise to the dispute. So the basic idea was that the legal system gathered together a group of people from the community who had personal knowledge of what went on and then permitted them to decide whether the conduct should be punished. That is not too far removed from the situations in a.c.e.n.a when most of the participants of the discussion of a particular incident have seen the offending message for themselves.

It was not always feasible, however, to assemble a jury that already knew what went on. Thus, it was necessary to develop methods for the disputants to tell their stories to the adjudicatory decision-makers. There are, of course, a variety of ways of telling stories, some effective and some not, some faithful to the real facts and some not. The notion of story telling to the adjudicatory decision-makers gradually evolved into formal mechanisms for determining who is entitled to tell a story (usually a professional lawyer) and rules for deciding how the story could be told—rules of evidence. The core idea of the modern jury trial, however, is not to be found in the definitions of the legal profession or in the current versions of the rules of evidence or civil procedure. The core ideas are to be found in the concept of giving each side an opportunity to tell its story, so that the people with the greatest interest in developing the story fully from the two opposing perspectives can do so. With that as a guide, a.c.e.n.a can be evaluated more fully. Arguably, its openness permits both accusers and defenders of a message to tell their stories to the decision-makers—the net-administrator participants on a.c.e.n.a. Viewed thus, a.c.e.n.a. satisfies the test for fair adjudication.

The enforcement function is perhaps trickiest of all, because it is here that the risk of an unaccountable invasion of private rights is greatest. A private adjudicatory decision does little harm if there is no coercive enforcement. It is important that due process have occurred before the deprivation represented by enforcement. In this regard, it is useful to look to traditional approaches to private individuals' authority to arrest (seizure of the person) or to seize property pursuant to judicial decree. The Statutes of Winchester identified private individuals as significant actors in the criminal justice process.[68] The role of the private citizen extended beyond simply protection of his own possessions; individuals owed a duty to society to join in the attempts to apprehend

---

68. *See* M. CHERIF BASSIOUNI, CITIZENS ARREST: THE LAW OF ARREST, SEARCH, AND SEIZURE FOR PRIVATE CITIZENS AND PRIVATE POLICE 9 (1977). "The Statutes of Winchester, enacted in 1285, formalized much of England's practice in matters of criminal justice and rules of apprehension." Furthermore, "the role of private persons in criminal justice was significant." *Id.* at 9.

criminals.[69] Private citizen arrests, searches, and seizures have traditionally been upheld under statutory or common law principles of citizens' arrests.[70]

Cancelmoose acts pursuant to the consensus of the participants in a.c.e.n.a. The consensus formed in a.c.e.n.a. can be viewed as the equivalent of a combination of a jury verdict and a warrant or a judgment. Viewed thus, Cancelmoose is equivalent to a deputy sheriff executing an arrest warrant after a criminal conviction, or a private party actually under color of a judgment. Using such analogies, we can see that the enforcement mode of a.c.e.n.a. is legitimate and "fair."

The other side of the coin, however, is that Cancelmoose does not enjoy a status equivalent to that of a public officer such as a sheriff (a sheriff is not self-appointed). And the a.c.e.n.a. process "authorizing" Cancelmoose to act is much more fluid and informal than the highly formal process of receiving a jury verdict and entering judgment on it. These differences animate arguments that a.c.e.n.a. is not "fair."

C. Self-government institutions proposed by IAHC

In 1997, the International Ad Hoc Committee (IAHC) proposed a comprehensive plan for self-government in a limited subject-matter—Internet domain names. This plan, based on an international web of private contracts and backed up by arbitration, is the most comprehensive yet proposed for a private, international system of Internet governance.

The IAHC was formed at the initiative of the Internet Society (ISOC)[71] and at the request of the Internet Assigned Numbers Authority

---

69. *See id.* at 9 ("Not only was it the right of any person to apprehend offenders, there was also a positive duty to drop all work when the 'hue and cry' was raised, and to 'join immediately in the pursuit'; and a private person was required to take part in the community institution of the 'hue and cry.'" (quoting J. HALL, THEFT, LAW AND SOCIETY, 162 (2d ed. 1952)); *see also* John Simon, Note, Tennessee v. Garner: *The Fleeing Felon Rule*, 30 ST. LOUIS U. L.J. 1259, 1263 (1986) (describing historical practice of outlawry; once one was declared an outlaw, every citizen had a duty to apprehend, and if necessary, to kill the outlaw).

70. *See* BASSIOUNI, *supra* note 68, at 87-95 (providing an index of state citizen's arrest statutes).

71. See Donald M. Heath, *Written Testimony of Donald M. Heath to U.S. House of Representatives Committee on Science Subcommittee on Basic Research For: Hearing on Internet Domain Names* (visited Oct. 10, 1997) <http://www.house.gov/science/heath_9-30.html>; see generally Internet Society (last modified Oct. 8, 1997) <http://www.isoc.org/>.

(IANA).[72] In addition, the IAHC was supported by the Internet Architecture Board,[73] the International Telecommunications Union,[74] the International Trademark Association,[75] and the World Intellectual Property Organization (WIPO).[76] Beginning work in November 1996, the IAHC was to "define, investigate, and resolve issues resulting from international debate over a proposal to establish global registries and additional generic Top-Level Domains."[77] IAHC sought comments from a wide variety of people and organizations and issued a final report with associated draft legal documents in February 1997. This report recommended changes in top-level domains for the Internet and a complete reorganization of the mechanisms for administering Internet domain names. The constitutional document, the generic Top-Level Domain-Memorandum of Understanding (gTLD-MoU), was signed in Geneva on May 1, 1997 and deposited with the Secretary General of the International Telecommunications Union.[78] With the signing of the

---

72. *See generally Internet Assigned Numbers Authority* (visited Sept. 13, 1997) <http://www.isi.edu/iana/>.

73. "The [Internet Architecture Board (IAB)] is responsible for defining the overall architecture of the Internet .... The IAB also serves as the technology advisory group to the Internet Society, and oversees a number of critical activities in support of the Internet." *The Internet Engineering Task Force: Glossary* (visited Oct. 11, 1997) <http://www.ietf.cnri.reston.va.us/glossary.html#IAB>.

74. *See generally International Telecommunication Union* (last modified Sept. 30, 1997) <http://www.itu.int>. The ITU is a treaty based upon inter-governmental organization, concerned with international telecommunications regulation. *See id.*

75. "[The International Trademark Association (INTA)] is an association of trademark owners and advisors worldwide. INTA is dedicated to the support and advancement of trademarks and related intellectual property concepts as essential elements of effective national and international commerce." *INTA Online* (visited Oct. 11, 1997) <http://www.inta.org>.

76. *See generally The World Intellectual Property Organization (WIPO)* (last modified Oct. 7, 1997) <http://www.wipo.org/eng/index.htm>. The World Intellectual Property Organization is a treaty-based intergovernmental organization providing a framework for multi-national negotiation of intellectual property treaties. *See id.*

77. Donald M. Heath, Written Testimony of Donald M. Heath to U.S. House of Representatives Committee on Science Subcommittee on Basic Research, For: Hearing on Internet Domain Names (visited Oct. 10, 1997) <http://www.house.gov/science/heath_9-30 html>.

78. The International Telecommunications Union is an entity with some advantages to nongovernmental participants because it permits full scale participation by such entities in its deliberations. This is not true of most international multilateral organizations.

gTLD-MoU, the IAHC was dissolved and replaced by the Interim Policy Oversight Committee (IPOC). As of May 15, 1997, 110 entities had signed or indicated their intent to sign the gTLD-MoU, although there is much controversy over the inclusion of some entities on that list.[79]

A Domain Name System (DNS) is an essential component in the Internet's operation. It permits use of human-friendly addresses for nodes connected to the Internet such as "kentlaw.edu," "law.vill.edu," "cilp.org," "fcc.gov," and "ibm.com."[80] The DNS functions through domain name servers that translate the human-friendly names into IP addresses (such as 153.104.15.250) through a series of interconnected domain name tables maintained on DNS servers. Tens of thousands of DNS servers are linked in a kind of hierarchical distributed look-up service.[81]

The IAHC was formed because of a growing set of controversies over the DNS as it now exists. The popularity and commercialization of the Internet has meant that multiple entities sometimes want to use the same domain name. Sometimes this occurs because the same few letters can signify more than one well known company, product, or service, or because some persons have registered domain names for the primary purpose of selling them to enterprises with which they appear to be associated. Many of the controversies relate to trademarks and service marks, as when enterprise A uses a domain name that is the same as a trademark registered to enterprise B. At the same time, Internet users outside the United States increasingly are restless with U.S. dominance of the DNS, a result of the Internet's origins in the U.S. Department of Defense.

The IAHC recommendations reflected the IAHC mandate to ameliorate conflicts over top level domains. They proposed a non-governmental solution to provide for competition among registries, and

---

79. See generally Donald M. Heath, *Written Testimony of Donald M. Heath to U.S. House of Representatives Committee on Science Subcommittee on Basic Research, For: Hearing on Internet Domain Names* (visited Oct. 10, 1997) <http://www.house.gov/science/heath_9-30 html>; The Generic Top Level Domain Memorandum of Understanding (visited Oct. 10, 1997) <http://www.gtld-mou.org/>. In private conversations with the author in the Summer of 1997, some entities shown as subscribed to the IAHC recommendations questioned whether they knowingly consented to be signatories.

80. The characters after the period in the examples given are Top Level Domains (TLD) signifying respectively two educational institutions, a non-profit organization, a United States governmental body, and a commercial enterprise.

81. If one DNS server does not know a domain name for which it is asked to supply the IP address, it refers the request to another DNS server with broader knowledge of that part of the Internet domain.

to develop an open process.[82] The recommendations addressed the administration of domain name assignments and the behavior of the distributed look-up service that maps human-friendly names into IP addresses. In addition to recommending the definition of seven new top-level domains, the IAHC report declared that "the Internet top-level domain space is a public resource." The administration of this public resource presents public policy issues, and should be carried out in an open and public manner "in the interest and service of the public."[83]

Of particular significance for this article, the IAHC recommended a new governance structure based on several memoranda of understanding, which both public and private sector entities were invited to sign. The gTLD-MoU—the constitutional document—became effective when it was signed by the IANA and ISOC. "Stewardship of the gTLD space was assigned to the gTLD DNS Policy Oversight Committee ("POC") comprising members named by the ISOC, IANA, [Internet Architecture Board], [International Telecommunications Union], International Trademark Association, WIPO and [the Council of Registrars]."[84]

Other memoranda created several regulatory bodies to carry out domain name policy. The Council of Registrars (CORE) was established by a Memorandum of Understanding (CORE-MoU), signed by multiple competing globally-dispersed registrars. CORE operates as a Swiss non-profit association. A gTLD DNS Policy Advisory Body (PAB) was formed from public and private sector consultation and oversees POC and CORE activities.[85] "Changes to policy can be initiated by POC and enabled upon the agreement of ISOC and IANA, with the review of PAB and CORE." One could regard the legislative initiative function as residing with POC, subject to revision and possible veto by PAB and CORE.

Two international treaty-based organizations also play a role in implementing the IAHC recommendations. The International

---

82. See generally Donald M. Heath, *Written Testimony of Donald M. Heath to U.S. House of Representatives Committee on Science Subcommittee on Basic Research, For: Hearing on Internet Domain Names* (visited Oct. 10, 1997) <http://www.house.gov/science/heath_9-30 html>.

83. *Final Report of the International Ad Hoc Committe: Recommendations for Administration and Management of gTLDs* (visited Sept. 13, 1997) <http://www.gtld-mou.org/draft-iahc-recommend-00.html>.

84. *See generally* Donald M. Heath, *Written Testimony of Donald M. Heath to U.S. House of Representatives Committee on Science Subcommittee on Basic Research, For: Hearing on Internet Domain Names* (visited Oct. 10, 1997) <http://www.house.gov/science/heath_9-30 html>.

85. *See Generic Top-Level Domain (gTLD-MoU) Technical Meeting* (visited Oct. 11, 1997) <http://www.gtld-mou.org/press/pab-2.html>.

Telecommunication Union agreed to act as the depository for the gTLD-MoU and to publish the list of signatories.[86] WIPO supports a dispute resolution mechanism for challenges of any domain name applicant's right to hold and use a second level domain name under the rules of the WIPO (Geneva) Arbitration and Mediation Center. WIPO would administer a new system of Administrative Domain Name Challenge Panels (ACPs). "These panels do not substitute for national or regional sovereign courts; they have authority over the domain names only, not the parties. Unlike courts, however, the challenge panels would have the ability to exclude certain names, such as world-wide famous trademarks, from all of the CORE gTLDs."[87]

Article 7 of the CORE-MoU reinforces WIPO's function. Registration agreements and application forms for assignment of secondary level domain names must include clauses that bind the registrars to follow ACP decisions and that bind applicants to submit to WIPO mediation, decision by an ACP and arbitration.[88] The WIPO Center must notify CORE of any results and decisions of ACP, mediation or arbitration proceedings that require action.[89]

Appendix D of the gTLD-MoU provides substantive guidelines for administrative domain name challenge panels. Under the gTLD-MoU, ACPs and the associated mediation and arbitration mechanism only have jurisdiction over claims regarding use of a second level domain name that is identical or closely similar to an alphanumeric string that is deemed to be internationally known and for which demonstrable intellectual property rights exist.[90]

---

86. *But see* Bruno Giussani, *Cybertimes: International Council to Take Up Issue of Domain Names*, N.Y. TIMES, June 18, 1997, at A1 (reporting on opposition to ITU role by Internet service providers).

87. *Final Report of the International Ad Hoc Committee: Recommendations for Administration and Management of gTLDs* (visited Oct. 11, 1997) <http://www.gtld-mou.org/draft-iahc-recommend-00.html>.

88. *See Memorandum of Understanding for the Internet Council of Registrars* ("Core-MoU") (visited Oct. 11, 1997) <http://www.gtld-mou.org/docs/cor-mou.htm>.

89. *See id.*

90. "Once an alpha numeric string has been deemed, for purposes of this policy, to be internationally known, and existing intellectual property rights have been demonstrated, an exclusion could be decided by an ACP, subject to consideration of rights held by others." *[Revised] Substantive Guidelines Concerning Administrative Domain Name Challenge Panels* (visited Nov. 7, 1997) <http://www.gtld-mou.org/docs/racps.htm>. These guidelines are reserved, pending further public discussion on the details of the Substantive Guidelines.

ACP procedures would allow for two types of exclusion. First, the second level domain name which was challenged could be excluded (that is, from the particular gTLD in which it was registered without the authorization of the owner of the intellectual property right). Second, a broader exclusion from some or all of the CORE gTLDs could be applied for, in 'exceptional cases.' Such cases would include at least trademarks which are globally known.[91]

Procedurally, any person can file a challenge requesting either exclusion or transfer of the requested second level domain name to the challenger.[92] Appendix D provides criteria for ACPs to determine if a challenge has been established successfully.[93] The ACP determinations, however, are of limited effect. "A determination of an ACP shall carry no precedential weight in any later national or regional court proceeding."[94] Appeals are permitted, although Appendix D is unclear as to what body has jurisdiction over the appeal. Presumably it is the same or another ACP.[95] Clearly, a de novo hearing by national or regional courts is contemplated.[96]

Unfortunately, the dispute resolution machinery proposed by the IAHC is limited to disputes over domain name assignment, especially those disputes that raise trademark or unfair competition issues. Moreover, it is an optional procedure, with a resort to national courts remaining available. As explained above, agreement on even this limited arrangement has been elusive. No doubt, agreement would be even harder to obtain with respect to a broader dispute resolution procedure and more ambitious use of domain names as leverage to enforce a broader set of international norms.

Notwithstanding these practical difficulties, it is useful to consider the possibility of using Internet domain names as a means of enforcing international norms in general. The growing importance of domain names in the Internet may provide the basis for a broader enforcement mechanism based on the IAHC recommendations, and may ultimately obviate the need for reliance on traditional legal institutions.

---

91. *Id.*

92. *See id.*

93. *See id.*

94. *Id.*

95. *See id.*

96. "Any dispute which has been submitted to an ACP may be brought, at any time before, during or after the administrative challenge procedure, to a national or regional court, which would hear the dispute under its normal jurisdictional and substantive rules." *Id.*

Would such a system work? Domain names as the centerpiece of a new private governance mechanism can serve some of the traditional purposes of legal remedies[97] reasonably well, but not others. On the one hand, revoking a domain name is a poor way of compensating a victim. Even if a domain name is awarded to a complaining party, that provides no compensation for past infringement of the trademark. There is nothing in the proposed IAHC machinery, no matter how far it is extended, that would serve the compensation purpose well. On the other hand, revocation will exclude the target from the Internet, and that possibility may have economic consequences serious enough to represent a major deterrent. If an entity believes it will be put out of business if it violates rules, it will avoid violating those rules. Finally, revocation of domain names is a very effective means of preventing further misconduct by the target; without a domain name, the target cannot repeat any further misconduct through the Internet.

A complete system for using domain name revocation as a remedy for enforcing international adjudicatory decisions requires at least three elements: rules for prescriptive jurisdiction, rules for adjudicatory jurisdiction, and rules for assuring compliance with the final order requiring that a domain name be revoked. The rules for prescriptive and adjudicatory jurisdiction have already been worked out.[98] Such rules are necessary to determine which substantive norms and which adjudicatory decisions would be entitled to enforcement through the domain name system. When the substantive norms and the adjudicatory decisions emanate from international institutions, their jurisdiction would be determined according to the documentary sources of their power. Currently, those sources are treaties.

That leaves the need to assure that the "sheriff" in this new legal domain obeys the "writ of execution." The obligation to obey the writ of execution would be expressed much as the obligation to obey decisions of ACPs and IAHC arbitration is expressed in the existing Memorandum of Understanding: once a decision to revoke a domain name is reached by the designated body, any registrar in the system must revoke the domain name. A registrar who declines to fulfill that obligation would lose its status as registrar. The integrity of this system depends upon the continued willingness of everyone within the hierarchical chain of

---

97. Remedies in law are intended to achieve at least three ends. First, damages compensate the victim. Second, remedies are intended to deter misconduct by punishing actors; knowledge of the possibility of such penalties deters misconduct. Third, remedies such as injunctions and incarceration are intended to block further misconduct by the actor.

98. *See supra* Part IV.A.

registrars to live up to their contractual commitments. As the scope of rules and decisions to be enforced by this means increases, however, the degree of compliance by registrars who intend to comply may diminish.

D. Economic royalism: proprietary power

Proprietary forms of private governance prevail in many parts of cyberspace. For example, a service provider such as America Online or CompuServe enforces unilaterally adopted rules by withdrawing the service of users who violate the rules. Most proprietary providers publish relatively complete sets of rules for use of the service.[99] Violation of the rules constitutes a trespass[100] and may justify termination of the service under contract.[101] Some proprietary providers use software tools that enforce the rules.[102] Use of a proprietary service over the objections of the proprietor is a trespass and is enjoinable.[103] Federal courts have deflected arguments that proprietary providers must provide access under the First Amendment of the United States Constitution or under the antitrust laws.[104] Thus, governmental authority in these situations is based on the power over private property.

Despite performance of these governance functions, proprietary services are not subject to constitutional constraints applicable to

---

99. *See* CompuServe, Inc. v. Cyber Promotions, Inc., 962 F. Supp. 1015, 1024 (S.D. Ohio 1997) (referring to policy statement limiting uses of service).

100. *See id.* at 1024 (granting preliminary injunction on trespass theory: connecting to Internet is no more a relinquishment of power over service provider's private property than any invitation to business customer is a relinquishment of power over inviter's premises).

101. *See* Cyber Promotions, Inc. v. Apex Global Info. Servs., Inc., No. Civ.A 97-5931, 1997 WL 634384, at *3 (E.D. Pa. Sept. 30, 1997) (recognizing general rule, but enjoining termination of service before expiration of 30-day contractual notice).

102. *See* Cyber Promotions, Inc. v. America Online, Inc., 948 F. Supp. 456, 459-60 (E.D. Pa. 1996) (describing software permitting users to block unsolicited e-mail); *CompuServe*, 962 F. Supp. at 1017 (describing orders to cease and desist, followed by use of software blocking devices).

103. *See* CompuServe, Inc. v. Cyber Promotions, Inc., 962 F. Supp. 1015, 1017 (S.D. Ohio 1997) (granting injunction on trespass theory).

104. *See* Cyber Promotions, Inc. v. America Online, Inc., 948 F. Supp. 456, 457-58 (E.D. Pa. 1996) (reviewing result in earlier First Amendment ruling, and summarizing ineffectiveness of antitrust argument).

traditional governmental entities.[105] Nevertheless, to the extent it limits its power by contract, a proprietor must follow its own rules.[106]

If one has a purely contractual framework within which rules are made and enforced, as in the three cases cited in the notes to this section, the likelihood of state action is de minimis. The only remedy of someone disadvantaged by the private dispute resolver would be for breach of contract, as in *Apex*, or a related tort claim such as fraudulent misrepresentation or intentional interference with contractual relations. On the other hand, when the dispute resolution mechanism is sanctioned by statute, as in the Fair Credit Reporting Act, the situation looks more like *Flagg Brothers v. Brooks*,[107] where the self help repossession was sanctioned by Article 9 of the Uniform Commercial Code, as adopted by the New York legislature. But in *Flagg Brothers*, the Supreme Court held that private conduct within a framework established by statute insufficiently engages the power of the state to represent state action. State action occurs only when enforcement powers of the state are used by private entities. One of the strongest examples of private enforcement is the landlord's common law right of distress: the power and privilege of seizing personal property on leased premises as a remedy for tenant nonpayment of rent. Exercise of the distress remedy generally has not been viewed as constituting state action, unless officers of the state such as deputy sheriffs assist the landlord.[108] Thus, designers of private electronic governmental mechanisms have greater autonomy when their arrangements are purely contractual, and correspondingly less when the last step in the private process is resort to public judicial machinery.

E. Conclusion

Among the three patterns of Cyberspace self-governance that have begun to emerge, the a.c.e.n.a. approach is the most democratic, but tends toward anarchy because of low institutionalization and diffusion of coercive power. The proprietary approach avoids those vices but concentrates power in the hands of one party and provides few channels

---

105. *See id.* at 464.

106. *See Apex*, 1997 WL 634384, at *3 (granting injunction against termination of service, based on failure to observe contractual notice period).

107. 436 U.S. 149 (1978).

108. *See e.g.*, Smith v. Chipman, 348 P.2d 441, 442 (Or. 1960); *see also* Shane J. Osowski, *Alaska Distress Law in the Commercial Context: Ancient Relic or Functional Remedy?*, 10 ALASKA L. REV. 33, 45-48 (1993); Douglas Ivor Brandon et al., Special Project, *Self Help: Extra-Judicial Rights, Privileges and Remedies in Contemporary American Society*, 37 VAND. L. REV. 845, 937, 1040 (1984).

traditional governmental entities.[105] Nevertheless, to the extent it limits its power by contract, a proprietor must follow its own rules.[106]

If one has a purely contractual framework within which rules are made and enforced, as in the three cases cited in the notes to this section, the likelihood of state action is de minimis. The only remedy of someone disadvantaged by the private dispute resolver would be for breach of contract, as in *Apex*, or a related tort claim such as fraudulent misrepresentation or intentional interference with contractual relations. On the other hand, when the dispute resolution mechanism is sanctioned by statute, as in the Fair Credit Reporting Act, the situation looks more like *Flagg Brothers v. Brooks*,[107] where the self help repossession was sanctioned by Article 9 of the Uniform Commercial Code, as adopted by the New York legislature. But in *Flagg Brothers*, the Supreme Court held that private conduct within a framework established by statute insufficiently engages the power of the state to represent state action. State action occurs only when enforcement powers of the state are used by private entities. One of the strongest examples of private enforcement is the landlord's common law right of distress: the power and privilege of seizing personal property on leased premises as a remedy for tenant nonpayment of rent. Exercise of the distress remedy generally has not been viewed as constituting state action, unless officers of the state such as deputy sheriffs assist the landlord.[108] Thus, designers of private electronic governmental mechanisms have greater autonomy when their arrangements are purely contractual, and correspondingly less when the last step in the private process is resort to public judicial machinery.

E. Conclusion

Among the three patterns of Cyberspace self-governance that have begun to emerge, the a.c.e.n.a. approach is the most democratic, but tends toward anarchy because of low institutionalization and diffusion of coercive power. The proprietary approach avoids those vices but concentrates power in the hands of one party and provides few channels

---

105. *See id.* at 464.

106. *See Apex*, 1997 WL 634384, at *3 (granting injunction against termination of service, based on failure to observe contractual notice period).

107. 436 U.S. 149 (1978).

108. *See e.g.*, Smith v. Chipman, 348 P.2d 441, 442 (Or. 1960); *see also* Shane J. Osowski, *Alaska Distress Law in the Commercial Context: Ancient Relic or Functional Remedy?*, 10 ALASKA L. REV. 33, 45-48 (1993); Douglas Ivor Brandon et al., Special Project, *Self Help: Extra-Judicial Rights, Privileges and Remedies in Contemporary American Society*, 37 VAND. L. REV. 845, 937, 1040 (1984).

for involvement by service users. It represents a kind of economic royalism, that may invite resort to traditional institutions to limit the power of proprietors.[109] The best of the three is the contractual web proposed by the IAHC because it is complete, democratic, and provides an appropriate degree of institutionalization.

VI. COMPARISON WITH OTHER SELF-GOVERNING COMMUNITIES

A. Introduction

Many autonomous self-governing communities exist within or separate from national states. As noted above, most autonomous communities owe their existence to grants of power from the national sovereign. Sometimes the grant is explicit, as in royal patents for the City of London, the East India Company, or the American provinces of Pennsylvania and Maryland. Sometimes it is implicit, having evolved through the common law. For example, the common law has worked out a kind of prescriptive and adjudicatory autonomy[110] for religious orders and certain internal matters of corporate governance. The following sections examine several models of self-governing communities. Mechanisms of self-governance for the Internet are likely to draw on these models, including their mostly contractual frameworks and the bases for limiting their scope.

The interesting thing about the following models is that, unlike traditional sovereigns, their boundaries are rarely marked by geography. Rather, other techniques are used for defining community membership. Voluntary membership models rely on consent to represent acceptance of a contractual framework for self-governance. By contrast, involuntary membership models must rely on some other legal justification for binding members to community norms and decisions, even when the remainder of their structure is contractual.

In the following sections, each model is assessed against Raz's three criteria for legal systems: the existence of normative rules, the existence of institutions, and the existence of coercive mechanisms. These criteria map roughly into the prescriptive, adjudicatory, and enforcement modes

---

109. *See, e.g.*, Cyber Promotions, Inc. v. America Online, Inc., 948 F. Supp. 456 (E.D. Pa. 1996); Cyber Promotions, Inc. v. Apex Global Info. Servs., Inc., No. Civ.A. 97-5931, 1997 WL 634384 (E.D. Pa. Sept. 30, 1997).

110. Prescriptive authority is authority to make rules. Adjudicatory authority is authority to decide cases. *See supra* Part IV.A (considering prescriptive and adjudicative immunity in the international context).

of jurisdiction.[111] As the section on legal frameworks explained, complete legal systems—those possessing all there criteria—are more likely to achieve autonomy. In addition, the sections also illustrate the ways in which traditional sovereigns limit the boundaries of community power, while affording immunities to activities at the core of the communities.

B. Involuntary membership models

1. *COLLECTIVE BARGAINING MODEL*

Collective bargaining refers to the process of making and enforcing terms and conditions of employment and other workplace rules through institutions established by a contract (a collective bargaining agreement) negotiated between an employer or group of employers and one or more trade unions representing their employees. After a majority of the employees within a "bargaining unit" have selected a union representative, a collective bargaining agreement binds all the employees within the "bargaining unit" regardless of whether they would prefer to negotiate individual contracts of employment with different terms.[112] Thus, membership can be involuntary.

An employment relationship covered by collective bargaining is a strong example of a self-governing community under American law. Even though specialized agencies have been set up at the federal level to establish the boundaries of collective bargaining communities and may, in limited circumstances, designate the representatives of employees,[113]

---

111. The Raz elements relate to the existence and comprehensiveness of a legal system. *See* RAZ, *supra* note 32, at 1-2 (explaining that a complete theory of legal system seeks to solve four problems, including existence and membership). They are in a sense attributes of sovereignty. The jurisdictional models of prescription, adjudication, and enforcement are concerned with the scope of power of a sovereign, usually assuming that a sovereign exists. *See* RESTATEMENT (THIRD) OF FOREIGN RELATIONS § 401 (1987) (describing three types of jurisdictions as "limitations" on state power in international law).

112. *See* Machinists Lodge 19 v. Soo Line R. Co., 850 F.2d 368, 375 (8th Cir. 1988) (stating general rule and noting exceptions).

113. *See, e.g.*, 29 U.S.C. § 159 (1994) (granting the National Labor Relations Board jurisdiction to define the bargaining unit); *see also* National Labor Relations Board v. Gissel Packing Co., 395 U.S. 575, 610-16 (1969) (describing situations in which a court may issue a bargaining order establishing the union as bargaining representative); Gourmet Foods, Inc. v. Warehouse Employees of St. Paul, 270 N.L.R.B. 578 (1984) (holding that the National Labor Relations Board does not have authority to establish a union as a bargaining representative when the union never had majority support within the bargaining unit).

these agencies neither define the rules governing the employment relationship[114] nor resolve individual disputes over terms and conditions of employment.[115] The Supreme Court has characterized a community covered by collective bargaining as a specialized community unto itself.[116]

Collective bargaining communities have a considerable immunity from state tort law.[117] The immunity is not absolute, and state law may be applied when it involves deeply rooted state interests.[118] The federal antitrust laws also give way to collective bargaining community decisions, as long as they are made within the traditional scope of workplace governance.[119]

Collective bargaining has normative rules expressed in collective bargaining agreements. Most collective bargaining agreements are comprehensive in this regard, including rules on every major subject of workplace governance, although they typically have "management rights clauses" allowing the employer considerable discretion to make specified entrepreneurial decisions. Collective bargaining has its own set of institutions—periodic negotiation for making rules and grievance arbitration for resolving disputes over rule application and enforcement. Collective bargaining has coercive mechanisms. It channels the

---

114. *See* National Labor Relations Board v. Insurance Agents Int'l Union, 361 U.S. 477, 486-88 (1960) (slowdown concurrent with labor negotiations does not constitute a refusal to bargain collectively; Congress intended that parties to collective bargaining have wide range of discretion).

115. *See* H.K. Porter Co. v. National Labor Relations Board, 397 U.S. 99,106-09 (1970) (holding that the National Labor Relations Board does not have the authority to compel the acceptance of any contractual provision in a collective bargaining agreement).

116. *See, e.g.*, United Steelworkers of America v. Warrior & Gulf Navigation Co., 363 U.S. 574, 579 (1960).

117. *See, e.g.*, San Diego Building Trades v. Garmon, 359 U.S. 236, 239-48 (1959) (holding state law against secondary pressure preempted); Lodge 76, Int'l Assoc. of Machinists and Aerospace Workers v. Wisconsin Employment Relations Comm'n, 427 U.S. 132, 144 (1976) (areas not addressed by federal law nevertheless are shielded from state regulation because Congress meant for them to be unregulated by law).

118. *See, e.g.*, Lingle v. Norge Div. of Magic Chef, 486 U.S. 399, 406-10 (1988) (public policy tort claim for wrongful dismissal is not preempted by federal enforcement of collective bargaining agreement covering employee when state claim is completely distinct).

119. *Compare* Brown v. Pro Football, Inc., 116 S. Ct. 2116, 2120-23 (1996) (nonstatutory antitrust exemption extended to unilateral imposition of compensation after impasse in bargaining), *with* Connell Construction Co. v. Plumbers Local 100, 421 U.S. 616, 626-35 (1975) (antitrust exemption did not extend to "pre-hire" agreement negotiated before any represented employees were in bargaining unit).

employer's power to discipline employees and terminate employment, and it organizes and channels the union's ability to put economic pressure on employers by striking. The union's strike weapon is less directly related to rule violation, although some collective agreements have exceptions to no-strike clauses that are triggered when employers violate rules set forth in the collective agreement.

Collective bargaining has all three Raz factors. The collective agreement institutionalizes workplace governance, it articulates norms, and it provides for coercive enforcement through strikes and termination of employment. As a model for network self-governance, collective bargaining is interesting because of its completeness and because of the limited immunities from antitrust and tort law enjoyed by its participants.

2. *MILITARY MODELS*

Military law governs military communities.[120] It is unique among the examples considered in this article because it goes the farthest in establishing immunities from the civil law of the traditional national community. Under United States military law, for example, a member of the armed forces is not subject to criminal or civil process for conduct associated with the performance of his duty. Military and naval communities historically have enjoyed such substantial immunity from the application of civilian law.[121] The immunity extends to military forces of both de facto and de jure governments.[122]

Nevertheless, members of military and naval forces are not completely immunized from civilian law. They may be charged with,

---

120. Membership in military societies is involuntary both because of the common historical practice of conscription, in which initial membership is involuntary, and because a member of the military establishment even if she was a volunteer at the outset is not free to terminate her membership unilaterally during its term.

121. *See* Underhill v. Hernandez, 65 F. 577, 581-83 (2d Cir. 1895) (reversing damages judgment against Venezuelan officer for harm done to an American citizen during revolution and reviewing cases establishing proposition that military officer enjoys sovereign immunity in international law).

122. A de facto government "exists where a portion of the inhabitants of a country have separated themselves from the parent state, and established an independent government. The validity of its acts, both against the parent state, and its citizens or subjects, depends entirely upon its ultimate success" Williams v. Druffy, 96 U.S. 176, 186 (1877).

arrested for, and tried for crimes committed within their forces.[123] Under some circumstances, writs of habeas corpus may issue to military authorities to show why a member of the military or naval force is detained.[124] Although damage actions are not allowed,[125] injunctions may issue for violations of civil rights.[126] Most of these cases involve internal military and naval disputes where the arguments against civil court intrusion are substantial. A fortiorari, members of otherwise autonomous military communities should be subjected to the surrounding legal system when the civil courts seek to adjudicate conduct in which a military authority injures a civilian.[127]

Military societies have normative rules for the conduct of individual members in the form of regulations and standing orders. They institutionalize rulemaking and enforcement through the chain of command and through rules defining the scope of the powers in each level of the chain of command, and they have courts marshal for dealing with rule violations. Military societies have coercive mechanisms that are employed directly against rule violators, including loss of pay, incarceration, and expulsion (discharge) from the service. They thus satisfy all three Raz criteria.

---

123. *See, e.g.*, United States v. Bevans, 16 U.S. (3 Wheat) 336, 386 (1918) (reversing conviction of marine sentry for murder committed aboard a ship of war because state courts had jurisdiction).

124. *See, e.g.*, Parisi v. Davidson, 405 U.S. 34, 35 (1972) (holding that habeas corpus may issue to inquire into basis for keeping conscientious objector in the service).

125. *See, e.g.*, Miller v. Newbauer, 862 F.2d 771, 774 (9th Cir. 1988) (well settled that no damages action may be pursued); Walden v. Bartlett, 840 F.2d 771, 773-74 (10th Cir. 1988) (affirming denial of damages but reversing denial of injunction for military prisoner alleging due process violations in connection with disciplinary proceedings); Knutson v. Wisconsin Air Nat'l Guard, 995 F.2d 765, 769-70 (7th Cir. 1993) (noting that there is some level of judicial review). *But see* Tigue v. Swain, 585 F.2d 909, 914 (8th Cir. 1978) (denying absolute immunity for alleged libel and false imprisonment by military officer).

126. *See, e.g., Walden*, 840 F.2d at 774-75 (reversing denial of injunction against military officials for alleged due process violations in connection with military disciplinary proceedings); *Knutson*, 995 F.2d at 770-71 (canvassing cases and concluding that no per se rule exempts military decisions from injunctive relief).

127. *See* B. ZOBEL, THE BOSTON MASSACRE 241-94 (1970) (describing the trial of British soldiers in regular civilian courts for "Boston massacre" resulting in the acquittal of most of them).

C. Voluntary membership models

1. *RELIGIOUS COMMUNITIES*

Religious communities long have enjoyed autonomy from secular sovereigns. In the United States, religious community autonomy is guaranteed by the free exercise clause of the First Amendment.[128] However, deference and immunity are limited to certain levels of subject matter. The deference relates only to matters of religious doctrine or policy, and to rulemaking and adjudication over internal discipline and government which has been interpreted to include "matters of discipline, faith, internal organization or ecclesiastical rule, custom, or law."[129] Other matters may be addressed with more or less autonomy pursuant to the contract rules applicable to other private associations.

Despite large measures of autonomy, traditional sovereigns impose boundaries on religious communities. Religious organizations may be liable for fraud for statements made outside the religious context[130] and for intentional infliction of emotional distress when they act coercively far beyond the bounds of customary religious practices.[131]

Religious communities can be complete Raz systems because they institutionalize, they articulate norms, and they coerce compliance by the prospect of expulsion from membership and from religious grace.

2. *PRIVATE ASSOCIATION MODELS*

Private associations like fraternities, churches, athletic leagues, country clubs, the Boy Scouts of America, and trade associations are largely self-governing, both with respect to rulemaking and adjudication.[132] One of the justifications for limited self-governance by private associations is freedom of association—a type of privacy interest.

---

128. *See* U.S. CONST. amend I.

129. Primate and Bishop's Synod of the Russian Orthodox Church Outside Russia v. Russian Orthodox Church of the Holy Resurrection, Inc., 617 N.E.2d 1031, 1033 (Mass. 1983) (describing a "neutral principles of law" analysis).

130. *See, e.g.*, Christofferson v. Church of Scientology of Portland, 644 P.2d 577, 598 (Or. 1982) (remanding for new trial, rejecting outrageous conduct liability because of voluntary nature of plaintiff's membership, and articulating rule limiting fraud liability to statements not involving religious matters).

131. *See, e.g.*, Wollersheim v. Church of Scientology, 66 Cal. Rptr. 2d 1, 6-18 (Cal. 1989) (affirming in material part judgment on jury verdict against religious organization).

132. *See* Note, *State Power And Discrimination By Private Clubs: First Amendment Protection For Nonexpressive Associations*, 104 HARV. L. REV. 1835, 1847 (1991) (articulating basic propositions).

The courts get involved only to enforce compliance with association rules.[133]

Private associations vary in the degree to which they have the three Raz attributes. Most private associations have normative rules, although their scope may be relatively narrow, limited to matters directly concerning the association rather than a broader range of human conduct. They have mechanisms for making rules to be recognized as such and usually have institutional arrangements for applying and enforcing rules. Coercive mechanisms in private associations usually are limited to expulsion, but some religious associations also subject rule violators to spiritual penalties or social penalties like shunning. Some non-religious associations like country clubs may subject rule violators to forfeiture of membership fees, which resemble a kind of security bond in this respect.

3. *CLEARINGHOUSE FUNCTIONS*

Clearinghouses handle the successive negotiation of checks and other financial instruments from the payee's bank to the drawee's bank, resulting in the eventual debiting of the drawer's account and crediting of the payee's account. Clearinghouses exist for electronic funds transfers and credit card transactions, as well as personal checks. The communication is mostly electronic, with paper instruments following later, if at all.

Bank clearinghouse functions are performed pursuant to clearinghouse bylaws and rules, which are contracts among participating banks. Although Article 4 of the Uniform Commercial Code supplies default rules for clearinghouse functions, these statutory rules routinely are altered by the clearinghouse rules.

Bank clearinghouses thus are examples of self-governing electronic communities. This example of self-governance is likely to be extended as electronic payment systems become more popular with buyers and sellers

---

133. *See, e.g.*, Rowland v. Union Hills Country Club, 757 P.2d 105, 108-09 (Ariz. 1988) (reversing summary judgment for country club officers because of factual question whether club followed bylaws in expelling members); Straub v. American Bowling Congress, 353 N.W.2d 11, 13 (Neb. 1984) (rule of judicial deference to private associations and compliance with association requirements, counseled affirmance of summary judgment against member of bowling league who complained his achievements were not recognized). *But see* Wells v. Mobile County Bd. of Realtors, Inc., 387 So. 2d 140, 142-45 (Ala. 1980) (claim of expulsion of realtor from private association was justiciable and bylaws, rules, and regulations requiring arbitration were void as against public policy; reversing declaratory judgment for defendant association).

of goods and services in cyberspace.[134] Nevertheless, there is nothing about the bank clearinghouse experience that suggests the extension of self-governance to persons or entities not actually signatories to the clearinghouse contract. Nor is there any indication that these communities are immune from rules developed by the surrounding legal system to address concerns of members of that larger community.

Financial clearinghouses have normative rules on the limited subject matter of allocating the risk of dishonor and setting time limits for the settlement functions in their financial communities. They have institutional mechanisms for applying the rules, although it is hard to find examples of clearinghouse agreements that provide for actual adjudicatory mechanisms for rule violations. Instead, the sanction for rule violation is to bear the loss. Enforcement takes the form of expulsion or bearing the loss.

4. *CORPORATION MODELS*

The corporation is an example of a private association that enjoys powers of self-governance, which may be enforced by traditional law, subject to certain limitations. In this respect, corporations are like the other private associations discussed above.[135] However, the corporate structure does offer advantages over other forms of business enterprises.

Of primary interest is the limited liability that the corporate form provides to its members.[136] Limited liability is premised upon the basic principles of agency.[137] In many instances, the principle will shield its agents from liability. For example, in certain circumstances, only the corporation, and not its agents, are liable on contracts made in the

---

134. *See generally* Henry H. Perritt, Jr., *Legal and Technological Infrastructures for Electronic Payment Systems*, 22 RUTGERS COMPUTER & TECH. L.J. 1 (1996).

135. Corporations have normative rules with respect to the allocation and commitment of corporate resources. They usually have rules relating to conduct in the workplace or on behalf of the corporation. They have institutions for rulemaking—usually the board of directors and a variety of management committees. They also may have formal institutional mechanisms for rule application and enforcement, although this also may be handled less formally through the managerial chain of command, with each supervisor applying and enforcing rules as to her subordinates. Coercion is limited to exclusion from the community, demotion, or repudiation of an action or decision.

136. *See* Stockmar v. Warrec Co., 844 F. Supp. 103 (D. Conn. 1994) (holding corporate officers not personally liable under state wage payment statute based on legislative intent).

137. *See generally* Restatement (Second) Agency (1958).

corporation's name.[138] However, there are other complex situations where the agent may be held personally liable. For example, the corporation and the agent, or both, may be liable for torts and crimes depending on the status of each party and the context in which such tort or crime occurred.[139] This may also affect "foreign" corporations (corporations incorporated in one state but doing business in another state), and thus must comply with certain formalities as defined by the other state's domestic law.[140] Absent compliance with these requirements, corporate actors may be found individually liable not only for torts and crimes, but also for contractual obligations into which they enter on the corporation's behalf. The same basic concepts for the treatment of foreign corporations apply internationally.[141]

The nature of limited liability for members of corporate communities is expressed by the legal fiction that a corporation is person. The corporation is treated as a separate legal entity,[142] which results in a

---

138. *See id* at §§ 140-43; HAROLD G. REUSCHLEIN & WILLIAM A. GREGORY, HANDBOOK ON THE LAW OF AGENCY § 118 at 182 (1979) (disclosing principal protects agent from liability).

139. *See id.* § 124 at 193-94 (acting for principal does not exculpate agent from tort liability). *See* United States v. Wise, 370 U.S. 405, 416 (1962) (reversing dismissal of Sherman Act indictment against individual corporate officer); *Compare* Bourgeois v. Commonwealth, 227 S.E.2d 714, 718-19 (Va. 1976) (holding corporate president not criminally liable for grand larceny absent proof he actually participated), *with* United States v. Dotterweich, 320 U.S. 277, 285-86 (1943) (upholding conviction of corporate president for criminal violations of Federal Food Drug and Cosmetic Act despite lack of proof of personal knowledge or participation).

140. *See* Gradison v. Ohio Oil Co., 156 N.E.2d 80, 83 (Ind. 1959) (construing state statute as granting qualifying foreign corporations all of the powers of domestic corporations); Cincinnati, Indianapolis & W.R.R. Co. v. Barrett, 94 N.E.2d 294, 296-97 (Ill. 1950) (holding foreign corporation that acquired domestic railroad not exempt from payment of registration fee merely because domestic railroad was exempt).

141. *See generally* LUCIE A. CARSWELL & XAVIER DE SARRAU, LAW & BUSINESS IN THE EUROPEAN SINGLE MARKET § 4.02 at 4-7 (1993) (explaining that liability is joint and several under European community law unless certain formalities are satisfied); *id.* § 4.03 at 4-11 (describing five formal requirements for the incorporation of a company); *id.* § 4.09[3] at 4-65 (describing European economic integrated grouping ("EEIG") as a kind of corporate joint venture operating across boundaries within the European community, but liabilities are joint and several thus negating much of the purpose). Inter-partner contracts purporting to limit liability are ineffective as against third parties for EEIGs. *Id.* at 4-70.

142. Acceptance of the concept that a corporation is an entity separate from its shareholders or members long antedates the development of limited liability for shareholders, which occurred in the middle of the nineteenth century in England, when law developed new structures to allow capital aggregation to exploit new technologies

tiered structure whereby management powers and limited liability may co-exist in a single individual in a corporation.[143] This legal fiction permits—at least in many instances—a third-party victim to be made whole through a legal claim against the corporation as an entity. The entity theory has received virtually universal legal acceptance, and the fictitious person so created has been given many of the constitutional protections available to individuals.[144]

The entity approach, however, has been subject to some criticism, especially in the context of the multinational enterprise. While in a strictly legal sense, a multinational enterprise can most simply be characterized as "an aggregate of corporate entities, each having its own juridical identity and national origin, but each in some way interconnected by a system of centralized management normally exercising its control from the seat of primary ownership," a multinational enterprise "has no coherent existence as a legal entity." [145] There is, however, a body of law emerging which is applicable to multinational corporations.[146] One such area consists of the rules of international law dealing with expropriation.[147] Nevertheless, efforts at establishing international codes and guidelines are relatively weak in their influence thus far, primarily because they are not legally binding.[148]

---

and larger markets made possible by new technologies. *See* P. Blumberg, The Law of Corporate Groups: Procedural Law 1-2 (1983).

143. Theoretically, a corporation consists of three tiers: (1) the shareholders who are traditionally viewed as the ultimate owners of the enterprise, (2) the board of directors, who are the managers of the corporation's affairs, and (3) the officers, who traditionally act as an officer, a director, and a shareholder. HAMILTON, CASES AND MATERIALS ON CORPORATIONS 9 (1994).

144. *See, e.g.*, Virginia Pharmacy Bd. v. Virginia Citizens Consumer Council, Inc., 425 U.S. 748 (1976) (finding that a corporation has a First Amendment right to free speech); Helicopteros Nacionales de Columbia v. Hall, 466 U.S. 408 (1984) (holding a corporation is entitled to due process). Immunity in the contract and criminal areas is justified by two rationales: (1) the practicability of enforcement; and (2) the perception that corporate institutional liability is more likely to result in internalization of societal goals and the mobilization of corporate bureaucratic institutional mechanisms to enforce traditional legal standards.

145. CYNTHIA DAY WALLACE, LEGAL CONTROL OF THE MULTINATIONAL ENTERPRISE 2 (1982).

146. *See id.* at 3-4 (posing question of whether international regime of regulation of multinational corporations is desirable).

147. *See id.* at 249-294.

148. *See id.* at 300.

In the future, one area in which truly international corporations may emerge is in Europe. The European Union encourages member states to adopt domestic corporation laws that conform to standards set by a European Union directive. Thus, the company incorporating in one member state can do business in other member states without discrimination. It will be interesting to evaluate European Union mechanisms for incorporation to determine if a European Union corporation can achieve corporate status vis-à-vis member states without being separately incorporated in each member state.

5. *THE LAW MERCHANT*

The law merchant was a transnational private law based not on any single national law but on mercantile customs generally accepted by trading nations.[149] The law merchant (lex mercatoria)[150] originated in the pre-Christian era in the Mediterranean, spread through Europe in the Middle Ages, and was reinforced through Admiralty and Maritime law and through Roman and Canon law. Despite the influences of several bodies of law, business dealings "rested on mutual confidence and good faith to an extent unknown to civil life."[151] By the end of the eleventh century, the law merchant began to be formalized and incorporated into the codes of certain conventional governments.[152] Eventually, the law merchant was applied by the Admiral's Court in England and published in Italian, French, Latin, Dutch, and German as the "Consulato del Mare."[153]

More or less independently, a body of commercial law developed in so-called "fair courts." Annual fairs took place in various places on the continent of Europe, attracting traders from Africa, Russia, and the Middle East as well as Europe. Each fair had a dispute resolution body that heard commercial disputes among the participating merchants.

---

149. *See* Harold J. Berman & Colin Kaufman, *The Law of International Commercial Transactions (Lex Mercatoria)*, 19 HARV. INT'L L. J. 221, 224-29 (1978).

150. Lex mercatoria actually is a broader concept than the law merchant. Philip De Ly described lex mercatoria as consisting of "self regulatory rules of professional organizations, usages, customs, general conditions, usual contractual clauses and techniques, arbitration rules, arbitral case law, general principles of private law and general principles of conflict of laws." FILIP DE LY, INTERNATIONAL BUSINESS LAW AND LEX MERCATORIA 221 (1992)

151. W. BEWES, THE ROMANCE OF THE LAW MERCHANT 14 (1923).

152. *See* Mark Garavaglia, *In Search of the Proper Law in Transnational Commercial Disputes*, 12 N.Y.L. SCH. J. INT'L & COMP. L. 29, 34-35 (1991).

153. *See id.* at 35.

While the crown might appoint a judge to guide the proceedings, the juries consisted of merchants participating in that particular fair.[154]

By the late sixteenth and early seventeenth centuries, when national sovereigns began to encroach on the traditional law merchant, the law merchant governed a special class of people (merchants) in special places (fairs, markets, and seaports). It was distinct from local, feudal, royal, and ecclesiastical law. Its special characteristics were that (1) it was transnational; (2) its principal source was mercantile customs; (3) it was administered not by professional judges but by merchants themselves; (4) its procedure was speedy and informal; and (5) it stressed equity, in the medieval sense of fairness, as an overriding principle.[155]

Thereafter, certain factors led to the diminished importance of the law merchant as a separate legal system. These factors included the rise of nationalism, competition between different kinds of courts for legal business, the tendency of traders to settle down and conduct their affairs from a particular place rather than traveling from fair to fair, and the incorporation of certain substantive principles of the law merchant into municipal law.[156] Nevertheless, as Mark Garavaglia has explained, the law merchant survives in modern commercial law under the guidance of international arbitration, commercial arbitration, and the Uniform Commercial Code.[157]

Thus, until the seventeenth century, the law merchant was an independent legal system with its own normative rules, its own institutions, and its own coercive measures. After that time, it lost the latter two features but retained its own normative rules.[158] Nevertheless, Professor Philip De Ly has cautioned that modern international business

---

154. *See id.* at 36-39 (describing fair courts).

155. *Id.* at 33 n.10. *But see* DE LY, *supra note* 150, at 17-19 (expressing doubt on whether law merchant ever was completely separate from national legal systems).

156. *See* Garavaglia, *supra* note 152, at 38-39; *see* DE LY, *supra* note 148, at 17 (explaining that the substantive absorption of law merchant by common law dates to 1756 when Chief Justice Mansfield began to qualify trade custom as legal rules applicable to all citizens).

157. *See* Garavaglia, *supra* note 152, at 40-55, 79-102 (describing international arbitration, concepts of law merchant in American commercial law, the emphasis on trade usages and regular practices in Uniform Commercial Code, and American attitudes toward international commercial arbitration); Parsons & Whittemore Overseas Co. v. Societe Generale, 508 F.2d 969, 973-77 (2d Cir. 1974) (rejecting public policy challenge to international arbitration decision).

158. Mr. Garavaglia's work does not make it entirely clear whether the fair courts imposed their own sanctions or relied upon traditional legal institutions to enforce their judgments. *See* Garavaglia, *supra* note 152, at 36-38 (describing fair courts and state facilitation of fair court proceedings).

law is not really an autonomous legal system in the sense that it "exists outside national legal systems; rather within national systems, it has some features of its own," derived from international origins and leading to a uniformity of international business law.[159] Lex mercatoria does not have a monopoly on resolving transnational business disputes and may need national law to enforce decisions applying its rules. Yet, lex mercatoria is an independent body of law that can be applied by national courts under choice-of-law rules of contract provisions. Because of the tendency of national courts to apply their own law, international arbitration represents the best forum within which to apply and enhance the status of lex mercatoria as a complete legal system.[160] Of course, the content of lex mercatoria must be known in order for it to play an enhanced role. One way to achieve this is through the publication of arbitral awards.[161] One author described lex mercatoria as consisting of "self regulatory rules of professional organizations, usages, customs, general conditions, usual contractual clauses and techniques, arbitration rules, arbitral case law, general principles of private law and general principles of conflict of laws."[162] The distinction between reliance on these sources of law and proof of customs and usages is subtle.[163]

Lex mercatoria is significant not only as a model of community autonomy, but also as a legal doctrine that may legitimate recognition of electronic community "law," as explained in Part III of this article. This is so because it is the clearest example of satisfaction of the four criteria that justify self-governance for electronic communities.[164] Lex mercatoria, however is not a complete Razian system because it lacks its own institutions and coercive measures.

D. Conclusion

All of the models exhibit some degree of autonomy because traditional sovereigns defer to them. All of the models, except lex mercatoria, are relatively complete Razian systems because all institutionalize their internal law, develop their own norms, and employ coercive enforcement power through expulsion from membership. Collective bargaining and military models have stronger coercive

---

159. DE LY, *supra* note 150, at 209-10 (1992).

160. *See id.* at 16.

161. *See id.* at 225.

162. *Id.* at 221.

163. *See id.* (describing how usages must be proven, while customs as rules of law may not need to be proven).

164. *See supra* Part III.

enforcement tools like the strike and direct physical action against military members. Corporations are less complete Razian systems because they employ coercive measures only against some constituents.

The frameworks for self-governance in all but the military communities are contractual. Antitrust and tort immunities are recognized by traditional sovereigns for many of the communities. Autonomy has its limits, however, in every model. Certain means and purposes justify application of traditional law, overriding the autonomy otherwise enjoyed by the community.

Aspects of all of these examples provide exemplars for electronic communities. Their utility depends, however, on working out institutional details for electronic counterparts and on developing appropriate immunities to define the boundary between electronic communities and traditional legal systems.

VII. WHAT REMAINS TO BE DONE?

Self-governance for electronic network communities is desirable: it is legally feasible within a contractual framework; examples already exist in parts of cyberspace; and rich models exist in the form of other types of private associations and communities. Establishment of comprehensive systems of self-governance for the Internet requires fleshing out contractual webs, defining antitrust and tort immunities according to established doctrine and newly articulated criteria for autonomy, and eventual development of a treaty framework.

A. Completing the contractual web

1. *DEVELOPING NORMS FOR ELECTRONIC COMMUNITIES*

The section on the legal feasibility of self-governance observed that completeness enhances deference by traditional legal institutions. Complete legal systems include rulemaking, adjudication, and enforcement. Earlier sections explained how adjudication can be provided through arbitration, and how denial of access to net resources—through the domain registry—can provide enforcement. That leaves rulemaking as the most significant challenge for developers of a comprehensive contractual web for self-governance.

Self-governing communities must have institutions to serve as sources of law. Institutions in cyberspace for rulemaking would exist on top of and in parallel with geographic-based institutions like state and federal courts and international rulemaking and adjudicatory institutions.

One possibility is to establish an electronic structure for a continuing plebiscite, such as that represented by a.c.e.n.a.[165] Alternatively, and more formally, electronic communities could identify a dozen or fewer experts in the norms and values of conduct in cyberspace to be rule-makers.[166] The IAHC recommendation includes such a mechanism in its policy committee.[167] In other contexts, the lawmakers need not be members of the community.

These "wise men and women" might function like the American Law Institute, publishing a restatement of appropriate principles to guide conduct in cyberspace. They also might function as arbitrators.[168] They also might play ancillary roles like being called as expert witnesses by the regular courts presented with cyberspace disputes. There is no reason that a single panel of experts cannot serve multiple communities as the lawmakers for those communities. The concept is roughly like a state adopting a particular section of a restatement written by the American Law Institute. The state, acting through its legislature or courts, reaches outside its own institutions to incorporate a doctrine developed for use by multiple communities.

More formally, one can follow other aspects of the model suggested by the IAHC, and write a kind of constitution that builds representative rulemaking institutions in a hierarchy defined by the domain name registration system for the Internet. The limitations of this model relate to the fluidity of "citizenship"—the composition of constituencies—at lower levels of the hierarchy.

2. *DEFINING MEMBERSHIP AND BOUNDARIES*

One of the greatest difficulties in formulating a means for electronic community self-governance is the difficulty in defining the boundaries of that community. The most fruitful source of guidance for defining a self-governing community is contract law. Determining the class of parties to the contract defines the boundaries of the community. The issue often arises when an individual files a lawsuit to compel arbitration or a community member might file a motion to dismiss a lawsuit for failure to exhaust arbitration remedies because a community may choose

---

165. *See supra* Part V.A-B.

166. The same individuals chosen to be rule-makers also could be dispute resolvers.

167. *See supra* notes 72-96 and accompanying text.

168. Arbitration is usually thought of as an adjudicatory mechanism that applies preexisting rules. However, there is no bright-line between rulemaking and adjudication in the common law tradition. Arbitrators and common law tradition can make law by elaborating and extending basic principles.

arbitration as the first step in its self-governance scheme. A court hearing either type of suit must decide if the reluctant party has agreed to arbitration. Only parties to the arbitration agreement are bound to arbitrate.

One can draft an arbitration agreement that represents a multilateral contract among all the members of a community, but the agreement will not effect individuals outside the community. Most existing models for self-governance present situations in which there is little doubt who is a member of the community, and thus little doubt as to the boundaries of the community's powers of self-government. For instance, nation-states are geographically defined, and international law places great emphasis on geographic boundaries in determining the reach of sovereignty. Further, the involuntary models, collective bargaining and military societies, have formal rules for determining who is a member of the community: induction and swearing in the case of military institutions and definition of an appropriate bargaining unit or craft or class in the case of collective bargaining. Finally, in the voluntary models, including private associations and bank clearinghouses, the act of joining and submitting to the rules of the private community defines the membership.

Electronic communities do not ordinarily have geographic boundaries, and thus that technique for defining membership and the boundaries of governance is unavailable. Furthermore, it is unlikely that traditional legal systems will provide formal rules to define membership along the pattern of the collective bargaining model or the military model until electronic community self-governance has been a reality for a considerable period of time.[169] The private association and clearinghouse models, which focus on a voluntary act of joining, appear to provide the best starting point for analysis of electronic community formation.

There may be problems in adapting conventional tests for contract formation to some electronic communities. While the act of subscription to America Online, CompuServe, or Lexis Counsel Connect may be unambiguous, it is not clear what the relevant act is in "joining" an Internet newsgroup or a community whose activities are carried on through Web-based postings. Does one "join" and thereby become subject to the rules of that newsgroup simply by reading the back postings from a newsgroup on one occasion, or by clicking into a Web site? If so, for what period of time is one member subject to that community's rules? Perhaps, only when one is reading and posting. This

169. Compelling submission to private governance may be politically unpalatable until there is more empirical evidence of the desirability of such compulsion.

answer satisfies the specialization justification for autonomy.[170] However, the transitory nature of membership in this example defeats almost any conceivable sanctioning power that the community could have. The sanction would only be effective if the community's resources are so attractive that exclusion is effective.

Notwithstanding the previous discussion, one should not overplay the importance of precise community membership definition. One can generally define boundaries in relation to the practical nature of the community. Even traditional national states have variable membership. As new people become citizens, others renounce their citizenship, and aliens come and go, sometimes these individuals fall within the power of the state and sometimes not. Moreover, nation-states have mechanisms for bringing ex-members within their communities again forcibly, such as extradition and reciprocal enforcement. Similarly, in the collective bargaining context, the class of employees covered by a collective agreement changes constantly as new persons are hired, and existing employees retire or are terminated. While all communities must define some reasonably ascertainable boundaries, these boundaries are defined in relation to the practical nature of the community. It might be quite feasible to define membership in certain electronic communities as the traffic moving through the community facilities at any given time.

The risk of sweeping significant numbers of people under the jurisdiction of private legal institutions to which they have not consented in fact, and with which they may be unfamiliar, will exert pressure on traditional legal institutions—legislatures, courts, and agencies—to draw narrower boundaries. When plausible boundaries, albeit fuzzy ones, are definable, then the arbitration mechanism can interpret those boundaries in particular cases.

3. *LEGALIZING*[171] *COERCIVE ENFORCEMENT OF COMMUNITY RULES*

There are two mechanisms for autonomous enforcement of community decisions: (1) execution against some asset made available as a security, such as a bond posted by member of a networked community or intellectual property left within the community; and (2) expulsion or exclusion from the community. The first of these mechanisms is available only if the network community requires the posting of some security as a precondition for membership. It might be feasible to require providers of

---

170. *See supra* Part III.B.

171. As I use the term, "legalizing" signifies recognizing privileges or immunities for self-governing activities that otherwise would produce liability in traditional legal institutions.

services to become members and post security before they are entitled to use network resources, like routers and network access points. Network communities might also require consumers to provide authorization to charge a credit card in advance. The limitation of this approach to enforcement, however, is that public key encryption[172] will permit a large volume of very small commercial transactions on open networks. In this context, consumers are unlikely to give authorization for enforcement security. Internet domain names present the most interesting possibility for an appropriate property interest. As discussed above, the IAHC proposal for privatizing and internationalizing domain name assignment and registration[173] shows how such a property interest might be the focus of coercive enforcement.

The second approach is to exclude community members who break the rules. Exclusion or expulsion is the most effective means of coercive enforcement of community norms. It is found in all the models for self-governance discussed above, except for lex mercatoria.[174] When social or economic resources available only through membership are valuable and more or less unique, the threat of exclusion provides a powerful incentive for rule compliance. Although the availability of competitive alternatives reduces the likelihood of antitrust liability for exclusion,[175] these same alternatives reduce the effect of exclusion as an enforcement technique. If a rule violator can just as easily go to another node on the Internet or to another service provider and get the same thing, exclusion is not very coercive. Nevertheless, the members of the network community may not focus on whether exclusion inflicts significant injury on the violator. Rather, they may seek to keep him out of that community and thus eliminate the possibility of his causing further injury within the community.

Effective enforcement of electronic community norms is easier when the community has reasonable solidarity. Solidarity is characterized by several important preconditions to informal community governance. Most important among these are the likelihood of continuing relationships among the people making, enforcing, and violating the rules as well as the existence of multidimensional relationships in the

---

172. *See supra* note 12.

173. Revocation of a domain name is an effective means for expelling someone from the Internet. Accordingly the IAHC report provides for effective enforcement—at least if solidarity can be maintained. *See supra* notes 72-96 and accompanying text.

174. *See supra* Part VI.

175. *See infra* Part VII.B.

community.[176] While the first of these prerequisites may be met in electronic communities, the second usually is not. Participants in electronic network communities may have continuing relationships, but their relationship is unidimensional; it involves only a particular type of communication and none of the other important human activities. This unidimensionality greatly weakens the force of informal community sanctions, such as social disapprobation by other members of the community and ultimately expulsion from the community.[177] If a violator of network community norms gets expelled, he simply can connect to another network. At least, he can do this if the market structure is competitive.

B. Immunities and community boundaries

Vigorous self-governance in cyberspace will involve conduct that ordinarily could give rise to liability imposed by traditional sovereign institutions. In particular, antitrust liability may result from rulemaking involving competitors and tort liability may result from accusations and testimony in adjudicatory proceedings. Robust self-governance depends on recognition of appropriate antitrust and tort immunities, as are already enjoyed by other self-governing communities.[178] Their availability to electronic communities usefully may be conditioned on certain criteria sketched out in an October 8, 1997 meeting in Washington, DC, presided over by this author, described in the Appendix, *infra*.

---

176. Multidimensionality is not fully explanatory. For example, stock exchanges are communities that surely are unidimensional in the modern world. Nevertheless, they exercise a good degree of self-government. This fact can be explained within the basic model by observing that when a single dimension has great importance to the members of a community, it can dominate other dimensions that tie the member to other communities. Alternatively, one can reason that an extremely important single dimension (like the economic interests of a broker in her membership in a stock exchange) spills over into other dimensions: A broker expelled from a stock exchange may be unable to send her children to college and may lose her spouse, thus implicating social, familial, and other noneconomic dimensions.

177. *See* Perritt, *Community Regained*, *supra* note 7, at 1009.

178. *See infra* Part IV.B.1-2.

1. *IMPOSING SANCTIONS WITHOUT VIOLATING THE ANTITRUST LAWS*[179]

As detailed in the preceding sections, contract law provides a promising framework for the proscriptive and most of the institutional prerequisites to a complete legal system for electronic communities. However, contract law may fail to provide any effective coercive sanctions for rule violators. The most likely sanction for violating electronic community rules is exclusion from the community. The problem is that the contract providing for exclusion—or providing the mechanisms through which exclusion is imposed—potentially runs afoul of the Sherman Act.[180] A contract providing for exclusion from the community is a restraint of trade within the meaning of the Sherman Act and, depending upon the specific facts and circumstances, may be categorized as either a "concerted refusal to deal" or an "exclusive dealing arrangement." A concerted refusal to deal arises when two or more persons agree not to deal with a third party.[181] An exclusive dealing arrangement arises when a buyer agrees to purchase all of its requirements from a particular seller.[182] Although excluding a member

---

179. A thorough analysis of the antitrust implications of Internet self-governance is beyond the scope of this paper. This paper will briefly outline the major considerations. For a more complete analysis, *see generally* PERRITT, INFORMATION SUPERHIGHWAY, *supra* note 12. Of course, antitrust laws differ among countries around the world. While the basic rules are similar to those applied in the United States, the details vary considerably. Because of the similarities in the laws, this discussion will focus on the antitrust laws of the United States. However, it is important to note that the self-governance of electronic communities must be effective globally. Ultimately, an international agreement may be necessary to give requisite certainty. In order to be effective, this international agreement should be self-executing so that legislative implementation by state parliaments is not necessary.

180. Section 1 of the Sherman Act declares that "[e]very contract, combination in the form of trust or otherwise, or conspiracy in restraint of trade or commerce among the several States, or with foreign nations, is hereby declared to be illegal." 15 U.S.C. § 1 (1994).

181. *See, e.g.*, Eastern States Retail Lumber Dealers' Ass'n v. United States, 234 U.S. 600, 606-14 (1914); Klor's, Inc. v. Broadway-Hale Stores, Inc., 359 U.S. 207, 209-14 (1959).

182. *See, e.g.*, Standard Oil Co. v. United States, 337 U.S. 293 (1949); Tampa Electric Co. v. Nashville Coal Co., 365 U.S. 320 (1961). Exclusive dealing arrangements may also violate section 3 of the Clayton Act. *See* 15 U.S.C. § 14 (1994). However, the Clayton Act only applies to the sale of goods. Moreover, exclusive dealing arrangements may violate section 5 of the Federal Trade Commission Act. *See* 15 U.S.C. § 45(a)(1) (1994) (declaring

from the community is a restraint of trade, it only violates the Sherman Act if the restraint is unreasonable.

Most restraints are judged under a rule of reason analysis in which the anticompetitive effects of the restraint are weighed against the procompetitive effects.[183] However, this rule of reason analysis entails a fact-intensive inquiry that produces significant societal costs in terms of business certainty and litigation efficiency.[184] Therefore, "there are certain agreements or practices which because of their pernicious effect on competition and lack of any redeeming virtue are conclusively presumed to be unreasonable and therefore illegal without elaborate inquiry as to the precise harm they have caused or the business excuse for their use."[185] This is a per se antitrust analysis. Exclusive dealing arrangements are evaluated under the rule of reason analysis.[186] However, the per se analysis is applied to certain concerted refusals to deal.[187]

Nonetheless, "not every cooperative activity involving a restraint or exclusion will share with the per se [concerted refusals to deal] the likelihood of predominantly anticompetitive consequences."[188] The per se approach is most often utilized when there are joint efforts to disadvantage competitors by denying relationships that the competitors need in the competitive struggle, the dominant parties have market power, and the practices are not justified by any plausible arguments that they were intended to enhance overall efficiency.[189] In cyberspace, there is no anticompetitive effect when the person excluded is not a producer. However, in other situations, the excluded individual may be a producer. For example, a packet routing consortium may decline to handle packets

---

unlawful, any "[u]nfair methods of competition in or affecting commerce, and unfair or deceptive acts or practices in or affecting commerce.").

183. In making this determination, courts consider a number of factors including the natural and probable consequences of the restraint, the history of the restraint, the evil believed to exist, the purpose of the restraint, the market power of the participants, and any other less restrictive alternatives. *See, e.g.*, Chicago Board of Trade v. United States, 246 U.S. 231, 238-39 (1918); National Collegiate Athletic Ass'n v. Board of Regents, 468 U.S. 85, 104-13 (1984).

184. *See, e.g.*, Northern Pacific Railway v. United States, 356 U.S. 1, 5 (1958); Northwest Wholesale Stationers, Inc. v. Pacific Stationery & Printing Co, 472 U.S. 284, 289 (1985).

185. *Northern Pacific*, 356 U.S. at 5.

186. *See e.g.*, Bar. Lab., Inc. v. Abbott Lab., 978 F.2d 98, 110 (3d Cir. 1992).

187. *See, e.g.*, Eastern States Retail Lumber Dealers' Ass'n v. United States, 234 U.S. 600, 606-14 (1914); Klor's, Inc. v. Broadway-Hale Stores, Inc., 359 U.S. 207, 209-14 (1959).

188. *Northwest Wholesale Stationers*, 472 U.S. at 295.

189. *See id.* at 294.

belonging to a network service that fails to apply the rules agreed upon by the consortium. In these circumstances, the bodies of self-governance must be prepared to explain how the sanction of exclusion enhances competition.

Health Care Peer Review[190] is a particularly pertinent area of antitrust analysis of self-governance[191] because health care peer review, like cyberspace adjudication and enforcement is a form of specialized self-government. The result of peer review often is exclusion from a particular facility or specialty, just as the result of cyberspace adjudication may be exclusion from all or parts of cyberspace.[192] "Although revocation of doctor's privileges may, perforce, eliminate competition by decreasing the number of doctors in a given specialty, this alone will not give rise to an antitrust violation."[193] An essential element of a section 1 violation is proof of an unlawful objective, and "[c]orrective action against a physician does not violate the antitrust laws if the physician's peer reviewers had legitimate medical reasons to believe that the physician provided substandard care." That is so because monitoring the competence of physicians through peer review is clearly in the public interest.[194] Actual support for the peer review decision enters into the analysis because if "the peer group's conclusions are so baseless that no reasonable medical practitioner could have reached those conclusions

---

190. Health care peer review is a system through which health care professionals, usually physicians, review the conduct of another member of their profession to determine if it satisfies the applicable norms of practice. When the answer is "no," the result often is exclusion from practice in a particular facility such as a hospital or expulsion from the profession altogether.

191. Other examples of antitrust immunity for self-governing communities are considered in the review of models for self-governance outside the cyberspace context. *See supra* Part VI.

192. The dimensions of the antitrust liability of Health Care Peer Reviews have been altered by Congress' enactment of the Federal Health Care Quality Improvement Act, 42 U.S.C. § 11112 *et seq.* (1994), which immunizes from antitrust liability peer review actions meeting certain criteria: being based on a reasonable belief that the action furthered quality health care, appropriate fact gathering, notice and hearing, and reasonable belief resulting from the fact gathering and hearing that the action taken was warranted. The health care peer review act requires the opportunity for a hearing either before an arbitrator or before a hearing officer or panel not in direct competition with the involved physician. *See* 42 U.S.C. § 11112(b)(3)(A)(iii) (1994). The federal act permits states to opt in or opt out. However, even before the enactment of the new legislation, not all Health Care Peer Reviews were subject to antitrust liability.

193. Willman v. Heartland Hosp. East, 34 F.3d 605, 610 (8th Cir. 1994).

194. *See id.* at 610-611.

after reviewing the same set of facts," a fact finder may infer the existence of an illegitimate motive.[195]

Generally, antitrust scrutiny of competitive collaboration to impose and enforce rules should focus on whether any restraints on competition are (1) ancillary, that is truly necessary for legitimate purposes, and (2) crafted to minimize the risk of anticompetitive effects.[196] On the other hand, restrictions on competition cannot be defended successfully by mere claims that they are inspired by pure or public spirited motives; instead, the actions must be justified as not incompatible with maintenance of effective competition.[197] "Coercive boycotts" of unapproved providers are "almost certainly unlawful regardless of their arguably worthy purpose," and that antitrust immunity depends on the peer review organization simply making a report to others like public licensing authorities, hospitals, insurers, referring physicians, and patients themselves who decide for themselves whether to act on the advice provided by the peer reviewers.[198]

The case law and commentary on physician peer review is directly applicable to "peer review" by competitors in cyberspace. The public policy in favor of self-regulation of cyberspace is similar to the public policy in favor of self-regulation in the medical profession. Market structures are similar, and the utility of due process in deflecting claims of anti-competitive motivations is the same in both industries. The crucial question is whether public policy is stronger in the case of physician self-regulation because it is useful to go beyond the external standards, and because it is clear to everyone that physicians have a profession that outsiders are hard-pressed to analyze. Advocates of similar treatment for cyberspace must show how the criteria for autonomy[199] are satisfied as strongly for cyberspace as for medicine. They probably are. Specialized rules and adjudication are needed as much for cyberspace as for medicine. Traditional communities are probably more indifferent to the content of most cyberspace rules than to most medical practitioner rules because the latter are almost all likely to have effects on nonmembers of the medical professions. The inherent likelihood that a specialized legal system will be more efficient, that it will induce greater voluntary compliance, and that it will regulate

---

195. *See id.* at 611.

196. *See* Clark C. Havighurst, *Professional Peer Review and the Antitrust Laws*, 36 CASE W. RES. L. REV. 1117, 1119 (1986).

197. *See id.* at 1120.

198. *See id.* at 1129.

199. *See supra* Part III.

behavior that otherwise would escape regulation tilt the political balance in favor of autonomy in both areas.

In order to facilitate Internet self-governance, it is important to formulate a more extensive antitrust immunity. First, proof of an anticompetitive purpose that is not legitimated by some plausible need for standardization would defeat the immunity: only those decisions that could be related to a legitimate private government objective would be within the revised immunity. Private governance regimes such as those proposed by the IAHC clearly have a purpose other than restricting competition; indeed they were developed for the purposes of increasing competition in the market for domain name administration services. Second, due process should accompany both rulemaking and adjudicatory and enforcement decisions. Assuring due process would militating in favor of accountability, access to decision-makers, and rationality of decision-making.

2. *TORT IMMUNITY*

Also important is the availability of a tort privilege or immunity so that accusations and findings of fact can be communicated without giving rise to liability for defamation. The present formulation of privilege in the Restatement (Second) of Torts[200] appears to be broad enough to afford the requisite tort privilege. Because the common law is uncertain, however, and because the Restatement only purports to synthesize American common law, it would be desirable ultimately to express the tort privilege in an international agreement that articulates the competition-law immunity.

There also are potential problems with contractual liability when entities covered by the IAHC machinery implement decisions to exclude malefactors. The IAHC machinery cannot be implemented without standardizing contracts of service through the full range of Internet Service Providers. Such standardized contracts not only should present the arbitration alternative for domain name disputes; they also should wave any liability for breach of contract for the enforcement of decisions reached through arbitration.

---

200. *See* RESTATEMENT (SECOND) OF TORTS, §§ 585-590 (absolute privilege to make accusations as a part of legal proceedings). *See generally* PERRITT, INFORMATION SUPERHIGHWAY, *supra* note 12 (discussing tort privileges).

3. *RECONCILING "CONVERSATIONAL" MODES OF GOVERNANCE WITH DUE PROCESS*

The possibility of self-governance in electronic communities is a particularization of a broader set of issues arising from the growing use of digital technologies to conduct social, commercial, and political relations. Many commentators have observed that the growing use of such technologies tends to make human interaction more fluid—more conversational—and to erode formalities. There are, however, important questions presented if this assessment is correct.[201] Then one must address the tension between conversational modes of decision-making and the legal role of formalities. Conversational modes of decision-making may be antithetical to the kind of due process necessary to assure antitrust immunity.

Legal formalities such as signature and writing requirements and witness and attestation requirements in the law of contracts and wills serve three functions: cautionary, evidentiary, and channeling.[202] As digital technologies reduce formality, one must ask whether the need for these functions has been reduced, or whether the need still exists, but they can be performed in other ways with new technologies.

A tension exists between informal decision-making in electronic community self-governance on the one hand and the concepts of procedural due process on the other. As a particular example, if government decision-making becomes a kind of ongoing conversation instead of being manifested in discrete decisional documents like final rules, statutes, and judicial decisions, one must question whether the traditional procedural due process requirement that one have notice of a rule that one is obligated to obey is present. The only way one has notice of the current version of the rule is to participate continuously in the conversation over it. Even if one participates, there is no certainty that the rules will be the same next week as it is today. This kind of uncertainty traditionally is viewed with alarm by advocates of the rule of law.

---

201. It also may be questioned whether the use of digital technologies does tend to make things less formal and more fluid. It may be that the increased scope of participation made possible by digital technologies will increase formality as a means of coping with the disorder and anarchy that otherwise would result.

202. *See generally* PERRITT, INFORMATION SUPERHIGHWAY, *supra* note 12 (explaining purposes of formalities in contracting).

C. International treaty

Given the need for tort and antitrust immunity, as well as the need to recognize private government institutions, the regular states of the world should negotiate an international understanding that implements the principles the Clinton Administration and the European Union announced in July 1997.[203] This legal framework would set the ground rules for private Internet governance in terms of transparency, opportunities to participate, and other due process issues in rulemaking, adjudication, and enforcement. When the private institutions reach decisions under these criteria, signatory states would obligate themselves to respect those decisions. The framework document need not specify in any detail what "respect" means. Any action taken within the appropriate governance mechanism that satisfies the criteria would be immune from antitrust and tort liability under international and national law.

The multilateral international framework also should reduce uncertainty by specifically empowering certain existing multilateral institutions (such as the World Intellectual Property Organization, the International Telecommunications Union, and the World Trade Organization) with certain ministerial powers to support the private Internet governance institutions. Of the existing multilateral organizations, the World Trade Organization is especially desirable because of its commitment to open competition and its recent negotiation of a telecommunications agreement.

VIII. CONCLUSION

Computer networking technologies enable new communities to arise that are not limited by traditional boundaries of time and geography. Some of these new communities may be strong enough or sufficiently specialized that they seek autonomy from surrounding legal institutions. A variety of models for relatively autonomous, self-governing communities exist, and contract law provides the mechanism for beginning the process of self-governance. An international arbitration agreement is a particularly strong mechanism for defining self-governance across international boundaries. Ultimately, however, certain kinds of disputes between community members and outsiders will remain within the jurisdiction of traditional rulemaking and adjudicatory institutions.

The Internet functions through bits and bytes being routed through the Internet protocol to autonomous nodes and networks throughout the

---

203. *See supra* note 1 and accompanying text.

world. Thus understood, the Internet is a prime candidate for self-regulation and private governance. But the Internet also functions through real people, corporations, and non-profit organizations. It functions through hardware, software, and communications channels owned by real people and organizations. Those people, organizations, and their tangible property are currently, and will remain for the foreseeable future, subject to outside legal institutions. Unless appropriate steps are taken to harmonize regular law with new forms of private Internet self-governance, self-governance of the Internet will be frustrated when more than 200 legislatures and thousands of administrative agencies around the world develop their own rules. People will second-guess the decisions of expert Internet adjudicatory bodies. Further, losing parties in the self-governance institutions will ignore decisions they do not like because they need not fear enforcement from the regular police and army. DNS servers, routers, firewalls, and web servers that comply with the private regulatory regime, nevertheless, will be punished and put out of business for failing to comply with traditional law.

This is not a positive scenario. Policy-makers can prevent it only if they by take action designed to develop a comprehensive contractual framework for self-governance. This development should draw particularly on the foundation suggested by the IAHC. Traditional sovereigns should shield it with an over-arching treaty framework of forbearance to assure adequate breathing room to new private self-governance within the Internet.

Regardless of the particular aspects of self-governance that might apply, the concept of self-governance is not helpful unless some electronic communities proceed to take the first few steps. Those steps involve the development of principles, codes of good practice, and even stronger forms of rules. The community should develop them through conventional contractual mechanisms, and actually apply them through some form of arbitration or contractual fact-finding. If an electronic community cannot get this far with self-governance, it will not get further; nor will traditional legal systems accord it the deference or immunity it desires.

Self-governance for the Internet is desirable for several reasons: self-governance may be more efficient; electronic network communities need different rules and procedures; open networks escape enforcement of conventional rules; and self-governance promotes voluntary compliance. Self-governance for the Internet is legally feasible within contractual frameworks and already exists in certain parts of cyberspace. These contractual models, properly supplemented by aspects of other models for private autonomous communities, will provide a complete system for private rulemaking, adjudication, and coercive enforcement of

community decisions. Antitrust and tort immunity is necessary to permit such a system to function effectively. Fortunately, there is much precedent for such immunities, and they can be limited by criteria for open participation, due process, and protection for traditional norms.

IX. APPENDIX: CRITERIA FOR AUTONOMY

On October 8, 1997, a number of Internet stakeholders met in Washington to define the boundary between Internet self-governance and the governments of sovereign countries. This author convened the meeting in response to declarations by the United States and European governments that called for private sector leadership and self-regulation of the Internet. Participants recognized that no system of self-governance can exist independently of national systems of law and that the degree of connection between private regulatory bodies and traditional legal institutions varies by issue. In any system of self-regulation, it is necessary to ask what can be done to heighten confidence that a particular issue will be handled in a way that will be fair, legitimate, and efficient.

Self-regulatory systems meeting certain criteria can inspire that confidence. The participants in the October 8th meeting reached agreement in principle on five such criteria, which are set forth below. The strength of agreement was greater for the first three criteria than the fourth and fifth, and greater on the text of each criterion than on the explanatory notes that follow the statement of each criterion. The explanatory notes are examples and limitations to explain the intended operation of the criteria. Not every participant on October 8th agreed with every word of the principles and the explanatory notes, but the following statement fairly reflects the judgment of the group taken as a whole.

These criteria are intended for use by the designers of self-regulatory systems, by government policy-makers, and by judges who must determine the degree of deference to accord the decisions of private self-regulatory bodies for the Internet. When a self-regulatory system meets all the criteria, its private decisions made consistent with its constitutional documents are entitled to judicial deference and to some insulation from antitrust and tort law.

A. Any private system of Internet domain name administration and any other aspects of self-regulation must be transparent.

Explanatory notes:

- Rules and agreements should be disseminated and published widely on the Net, in an understandable and complete form.
- The process for amending and setting rules should be fully disclosed.
- Rules should be able to be created and changed only after an adequate notice period.

- Initiation and results of adjudications should be fully disclosed, including the factual and legal basis for the decision.
- Enforcement procedures and decisions should be fully disclosed.
- Who is making decisions and how they were selected should be publicly disclosed.

B. Rule making and adjudication within a private governance body must provide due process.

Explanatory notes:

- Decisions should be expressed in writing (including electronic formats).
- Adjudicatory decisions should be preceded by some form of hearing appropriate to the factual issues, and to the magnitude of the interests at stake.
- Decisions on rules and adjudications should be preceded by notice.
- Review of self-government decisions should be available, but should be confined to whether due process was made available not to the correctness of the decision on the merits; exceptions to this limitation on review should be reserved to cases implicating the protective principle below.

C. The actions of a private system of Internet domain name administration must be accountable.

Explanatory notes:

- The market provides a substantial degree of accountability, insofar as registrants may choose freely (in a free market) among a number of different registrars and registries offering diverse terms, conditions and policies.
- Additional accountability stems from the felt duty of all industry providers to assure that the net continues to work smoothly.
- Policy-making should be centralized only for issues as to which there is a need for a single, central rule, such as the policy of concurrence or interoperation.
- Each registrar is accountable to registrants according to the terms of the registration contract, and vice versa, provided that the registrar does not engage in fraud.
- Countries may or may not choose to require that actions within a country code comply with, and are thus accountable to, the law or policy established by that local government. In

any event, the relationship between any particular country code domain and the law or institutions of a particular country should be disclosed to registrants, who should be free to decide whether or not to contract to register within such domains.

- Registries, which set policies for any particular domain and the corresponding set of registrars, should promise each other that they will enforce their own stated policies, and should be accountable to each other for doing so.
- Registries, individually and in groups, should appoint or elect appropriate bodies to resolve disputes and make rules with respect to registrations within their domains.
- One or more new entities, constituted as membership organizations or non-profit corporate entities (perhaps with multiple classes of stock), membership in (or ownership of) which is open to all in exchange for appropriate fees, should establish or oversee policies for various domains.
- Entities governing particular domains may appoint or elect a centralized entity to coordinate their actions and/or play the centralized roles previously performed by IANA.
- The decisions of such domain policy setting entities should be entitled to deference by local courts under doctrines similar to the business judgment rule, and under the criteria expressed in this document.
- Insofar as the officers or trustees of entities exercising policy oversight over domains are elected on the basis of membership or stock ownership, individual persons or corporations should not be allowed to accumulate or vote multiple or duplicative memberships or ownership interests. Such memberships or stock interests may have multiple classes, reflecting appropriately the relative economic stake or representative reach of the institutions eligible to hold such classes of membership or stock.

D. An open opportunity must exist for anyone meeting stated qualifications to participate.

Explanatory notes:

- Openness must operate on four levels:
  - Cooperative agreements among sovereigns (treaties).
  - Composition and deployment of policy oversight entity.
  - Freedom of entry among registrars (multiple business models).
  - Consumer choice (portability and variety).

- Freedom of entry for registries should be tempered by:
  - Assurances of continuous and accurate resolution of domain name requests by way of a shared database.
  - Insurance against private failure leading to collapse of system by way of surety bonds and maintenance of "slave" servers.
- ISP's should subsidize the root server infrastructure.

E. Acceptable criteria must exist to avoid contract overreaching and for intellectual property protection and protection of the interests of third parties.

Explanatory notes:

- Inter-registrar agreements should recognize intellectual property rights
- There must be some recourse to national sovereignty.
- Dispute policy must come from a source other than registrars.
- It may be desirable for all registrars to follow the same dispute policy.
- Adjudicators (dispute resolvers) should be empowered to set aside overreaching contract provisions. "Overreaching" must be carefully defined but, for example, the agreement that "anyone with a trademark registration wins" is an example of overreaching.
- Domain name holders (but not holders of e-mail addresses) must be known; anonymity is not permitted.
- Some guidance should be provided on jurisdictional issues.

COMMENT

COMMUNICATIONS TOWER SITINGS: THE TELECOMMUNICATIONS ACT OF 1996 AND THE BATTLE FOR COMMUNITY CONTROL

SUSAN LORDE MARTIN †

TABLE OF CONTENTS

I. INTRODUCTION 483
II. LOCAL COMMUNITY RESIDENTS OPPOSE CELLULAR PHONE TOWERS 485
III. THE TELECOMMUNICATIONS ACT OF 1996 AND ITS EFFECT ON LOCAL REGULATION OF CELLULAR TOWER FACILITIES 488
IV. JUDICIAL RESOLUTION OF TELECOMMUNICATIONS TOWER SITING DISPUTES SINCE THE ENACTMENT OF THE TELECOMMUNICATIONS ACT 492
V. DISCUSSION AND CONCLUSIONS 498

I. INTRODUCTION

The Telecommunications Act of 1996[1] (Act) was enacted by Congress on February 8, 1996, primarily to promote a pro-competitive, deregulatory environment for telecommunications providers that would secure lower prices, better service, and faster access to new technologies for consumers.[2] Universal service is also a cornerstone of the congressional plan.[3] The Act's chief method of accomplishing these goals

---

† Associate Professor of Business Law, Frank G. Zarb School of Business, Hofstra University; J.D., 1987, Hofstra University; A.B. Barnard College.

1. Pub. L. No. 104-104, 110 Stat. 56 (codified as enacted and amended in scattered sections of 15 U.S.C., 18 U.S.C. and 47 U.S.C.).

2. 110 Stat. at 56; 141 CONG. REC. H8269 (daily ed. Aug. 2, 1995) (statement of Rep. Linder); F.C.C. Proposed Rules: Implementation of the Local Competition Provisions in the Telecommunications Act of 1996, 61 Fed. Reg. 18,311 (1996) (to be codified at 47 C.F.R. ch. I).

3. S. REP. NO. 104-23, at § 103 (1995).

is the "removal of barriers to entry"[4] into the businesses of telecommunications services, including those provided by local, and long distance telephone companies and video, cable, and wireless companies.[5] This plan sounds laudable and seems to be one to which most consumers would subscribe. Nevertheless, Congress recognized that difficulties might arise in its implementation if state and local governments attempted to exert their jurisdiction in ways that would erect or maintain barriers to telecommunications facilities.[6]

One such problem involves the siting of telecommunication towers and antennas. This problem existed before the new Act became law and continues to create rancor and litigation. From one end of the country to the other, communities have been fighting against telecommunications companies that want to put facilities in their neighborhoods.[7] The new law, rather than solving the problem, exacerbates it by providing ammunition for both sides of the controversy. On one hand, the Act

---

4. Telecommunications Act of 1996, 47 U.S.C. § 253.

5. *See, e.g.*, F.C.C., THE HARD ROAD AHEAD -- AN AGENDA FOR THE FCC IN 1997 (Dec. 26, 1996).

6. *See, e.g.*, 47 U.S.C. §§ 224(c), 253(a), 253(d) (1996).

7. *See, e.g.*, Lisa Buie, *Cellular Towers Bedevil Board*, HERNANDO TIMES, Dec. 3, 1996, at 1 (Hernando County, Florida); Lisa Frederick, *Towers Raise Ire as Cellular Phone Structures Spring Up, Some Residents Are Voicing Concerns*, ATLANTA J. CONST., Jan. 2, 1997, at 1 (Covington, Georgia); Joe Gose, *Cities in Search of Ways to Regulate Cellular Towers - Johnson Countians Want to Limit Construction, Fearing the Impact on Property Values*, KANSAS CITY STAR, Nov. 20, 1996, at B3 (Johnson County, Kansas); Helen Bennett Harvey, *Plans for Tower Anger Residents*, NEW HAVEN REGISTER, Sept. 20, 1996 (Orange, Connecticut); Mark Larson, *And Takes on Phone Towers*, BUS. J. - SACRAMENTO, Sept. 9, 1996, at 2 (Sacramento County, California); Mike Maller, *Cell-Site Doesn't Ring Well with Neighbors*, IDAHO STATESMAN, July 17, 1996 (Cloverdale, Idaho); Jonathan Marshall, *Where's the Antenna?*, THE COLUMBIAN, Dec. 18, 1996, at E1 (San Francisco, California); *Medina Applies Another Hold on Cell Towers*, SEATTLE TIMES, Oct. 29, 1996 (Medina, Washington); James A. Merolla, *Zoning Board Gets an Earful from Tower Opponents*, PROVIDENCE J. BULL., Oct. 30, 1996, at 3C (Richmond, Rhode Island); Barbara Miller, *57 Residents Sign a Petition Opposing Cell Telephone Tower*, HARRISBURG PATRIOT, Jan. 28, 1997, at 7 (Palmyra, Pennsylvania); Doug Nurse, *Towers Loom as Upcoming Problem*, TAMPA TRIB., Jan. 5, 1997, at 1 (Lakeland, Florida); Tom O'Neill, *Here's the Church, Here's the (Fake) Steeple - It's a Cell Tower, and Much Debated*, CINCINNATI ENQUIRER, Jan. 20, 1997, at B1 (Cincinnati, Ohio); Paul Rogers, *Emerson Considers Cell Tower Ordinance Law Would Limit Height, Placement*, THE RECORD, Jan. 14, 1997, at L01 (Emerson, New Jersey); Shaun Sutner, *Towers Loom Large - Cell Service Riles Neighbors*, SUNDAY TELEGRAM, Dec. 29, 1996, at B1 (Worcester, Massachusetts); *Towers of Power*, KNOXVILLE NEWS-SENTINEL, June 16, 1996 (Knoxville, Tennessee); Pat Wiedenkeller, *In Hempstead: Bloc Aims to Trim Plan for Tower*, NEWSDAY, Sept. 15, 1996 (Malverne, New York).

states that "[n]o State or local statute or regulation ... may prohibit ... the ability of any entity to provide any interstate or intrastate telecommunications service;"[8] while, on the other hand, the Act provides that "[n]othing in this section shall affect the ability of a State to impose ... requirements necessary to ... protect the public safety and welfare, ... and safeguard the rights of consumers."[9] These provisions make it reasonable for telecommunications companies to argue that a local zoning ordinance cannot prohibit the construction of a tower in the location and of the dimensions necessary for seamless cellular phone service. Local residents, however, can also make a compelling argument that zoning rules limiting the size and placement of telecommunications facilities protect their economic and emotional welfare.

This article first describes the problem that arises when telecommunications companies seek to erect towers in order to provide cellular phone service. It then discusses the relevant provisions of the Telecommunications Act of 1996 and the role of the Federal Communications Commission (FCC) in implementing them. Next, the article surveys the cases that have dealt with the cellular tower issue. Finally, the article concludes that Congress should amend the Act to define acceptable methods of state and local regulation of communication facilities and to require that, in support of their applications for variances, communication service providers demonstrate that they have taken into consideration the interests of local residents in siting their facilities. Congress should also amend the Act to allow states and local governments to rely on evidence of health and environmental effects when making decisions about the location of communications facilities, even when that evidence contradicts FCC standards. In the meantime, the FCC and the courts should use their power to preempt state or local requirements only after giving due consideration to the rights and interests of affected local residents.

II. LOCAL COMMUNITY RESIDENTS OPPOSE CELLULAR PHONE TOWERS

Cellular phone service was first offered in the United States in 1983.[10] Since then, telecommunications businesses have been attempting to erect towers with antennas in or near almost every local community in order to provide service that reaches every area of the country.[11] A few years ago there were several thousand telecommunications towers in the

8. 47 U.S.C. § 253(a) (1996).

9. 47 U.S.C. § 253(b) (1996).

10. CONTEL CELLULAR, INC., 1994 ANNUAL REPORT 6.

11. *Id.* at 6-9.

nation.[12] Today, there are about 25,000.[13] Experts estimate that by 2002, there will be 100,000 towers.[14] Although cellular phones have become very popular, and people want service with good sound quality, most are unwilling to obtain it if the price is living next to, or within viewing distance of, a tower.[15]

There are two primary objections raised to the proximity of telecommunications towers to residential neighborhoods. First, people are concerned about the health risks associated with electromagnetic fields generated by cellular phone facilities.[16] Even though there is no conclusive evidence that electromagnetic fields are cancer-causing, particularly at the low levels emitted by cellular phone transmitters,[17] there is also no conclusive evidence that they are not. In fact, many studies have found a correlation between exposure to electromagnetic fields and cancer.[18] Therefore, with twenty-two countries still studying

---

12. *See, e.g.*, Clarke Canfield, *Analysts Say the Number of Telecommunications Towers in Maine and the Nation Could Quadruple in the Next Five Years. That Explosive Growth Challenges the Abilities of Some Towns to Balance Land Use Concerns with the Needs of the Industry, Creating a High-Tech, High-Wire Act*, PORTLAND PRESS HERALD, Oct. 16, 1997, at 4E.

13. *See id.*

14. *See id.*

15. *See supra* note 7.

16. *See* Susan Lorde Martin, *Communities and Telecommunications Corporations: Rethinking the Rules for Zoning Variances*, 33 AM. BUS. L.J. 235, 241-44 (1995).

17. Antennas commonly have 100 watts of power or less, compared to the 5000 watts of an AM radio station, the 50,000 watts of an FM station or more than 300,000 watts of a television station. *See* Candice Millard, *Mending Fences*, CELLULAR BUS., Dec. 1, 1996 at 40; Barbara Miller, *57 Residents Sign a Petition Opposing Cell Telephone Tower*, HARRISBURG PATRIOT, Jan. 28, 1997, at 7. The strength of the magnetic fields created by these power sources varies with the distance from the source so that, for example, a hair dryer could have a magnetic field of 60 to 20,000 milliGauss when it is 1.2 inches away, but only 1 to 70 milliGauss when it is 12 inches away. *See* Sharon Tomecek, *What Are Electromagnetic Fields?*, REAL ESTATE TODAY, Nov./Dec. 1992, at 19.

18. *See, e.g.*, Nancy Wertheimer & Edward Leeper, *Electrical Wiring Configurations and Childhood Cancer*, 109 AMER. J. OF EPIDEMIOLOGY 273-84 (1979); L. Tomenius, *50-Hz Electromagnetic Environment and the Incidence of Childhood Tumors in Stockholm County*, 7 BIOELECTROMAGNETICS 191-207 (1986); D.A. Savitz et al., *Magnetic Field Exposure From Electric Appliances and Childhood Cancer*, 131 AMER. J. OF EPIDEMIOLOGY 763-73 (1990); J.R. Wilkins & Ruth Koutras, *Paternal Occupation and Brain Cancer in Offspring: A Mortality-Based Case-Control Study*, 14 AMER. J. OF IND. MED. 299-318 (1988); K.T.S. Yao, *Microwave Radiation-Induced Chromosomal Aberrations in Corneal Epithelium of Chinese Hamsters*, 69 J. OF HEREDITY 409-12 (1978). *See also* ELLEN SUGARMAN, WARNING: THE

the health effects of exposure to electromagnetic fields, people remain afraid.[19] Second, people are concerned that proximity to a tower will lower their property values.[20] The manager of a real estate brokerage office in New York has called the towers "the kiss of death," claiming that a home with a tower in its backyard can sell for twenty-five percent less than a comparable home without a tower.[21] Homeowners are also concerned for their own visual comfort, because of the poor aesthetics of the tower facilities.[22]

ELECTRICITY AROUND YOU MAY BE HAZARDOUS TO YOUR HEALTH, app. A (1992) (containing extensive list of major studies).

But see William J. Broad, *Cancer Fear Is Unfounded, Physicists Say*, N.Y. TIMES, May 14, 1995, § 1, at 19 (reporting that American Physical Society, world's largest group of physicists, asserts that it can find no evidence that EMFs from power lines cause cancer).

The most recent reported study, conducted over a five-year period by the National Cancer Institute and the University of Minnesota, found no evidence that electromagnetic fields increase the risk of acquiring childhood leukemia. *See* Robert Langreth, *No Evidence Is Found Linking Leukemia In Children and Electromagnetic Fields*, WALL ST. J., July 3, 1997, at B6. There are, however, some scientists who still think there may be a relationship between electromagnetic fields and some kinds of cancer. *Id.*

19. *See* Soraya Sarhaddi Nelson, *AT&T Antennae a Tough Cell*, NEWSDAY, Jan. 22, 1997, at A23.

20. *See* Joe Catalano, *Similar Houses, Different Prices? It's Time to Look at the Externals*, NEWSDAY, Mar. 14, 1997, at D2; *see also* Martin, *supra* note 16, at note 59 and accompanying text.

21. Catalano, *supra* note 20, at D2; *see also* Martin, *supra* note 16, at note 59 and accompanying text.

22. *See e.g.* Evans v. Shore Communications, Inc., 685 A.2d 454, 462 (Md. Ct. Spec. App. 1996); Westinghouse Elec. Corp. v. Council of Township, 686 A.2d 905, 908 (Pa. Commw. Ct. 1996).

The towers can range in height from 55 feet to 500 feet. *See, e.g.*, Tina Allen, *Ice Skating Center Gets Planning Commission OK*, LAS VEGAS REV.-J., June 25, 1997, at 3AA (55 feet); Jerry Fallstrom, *Officials to Take Up Tower Plan Some Residents Say the Telecommunications Structures Make Surrounding Property Values Plummet*, ORLANDO SENTINEL, June 24, 1997, at 1 (200 feet); Helen Bennett Harvey, *Plans for Tower Anger Residents*, NEW HAVEN REGISTER, Sept. 20, 1996, at C1 (120 feet, 180 feet, 185 feet); John J. Keller, *Bad Reception - With Cellular Towers Sprouting All Over, Towns Begin to Rebel*, WALL STREET JOURNAL, July 2, 1996, at A1 (125 feet); Jeff Ostrowski, *Rentin' the Roof Market for Cell Towers Has Nowhere to Go but Up*, S. FLA. BUS. J., Aug. 29, 1997, at 1 (500 feet); Peter Pochna, *Freeport Planners Continue Work on Tower Limits A Moratorium on New Towers Will Be Extended While Restrictions Are Being Written*, PORTLAND PRESS HERALD, Sept. 4, 1997, at 1B (300 feet). Towers are usually four-sided, lattice-style towers or monopoles, that is, solid single poles. *See, e.g.* Fallstrom at 1.

The conflict between the goals of telecommunications companies and those of residents of local communities has created disputes that end up being resolved by courts. After the companies select sites that maximize communication distance and quality, local zoning ordinances frequently require them to obtain variances for non-conforming uses.[23] It is not unusual for the community zoning board to respond to citizens' complaints and deny the application for a variance. The telecommunications companies are prepared for this result and appeal the denial in court, where they frequently win.[24] If the zoning board grants the application for the variance, it is likely that community residents will not appeal the decision because they lack the financial resources; if they do appeal, they usually lose.[25]

The playing field is not level when local citizens, attempting to protect their physical, emotional and economic health, are required to battle in court against large telecommunications corporations with vast financial resources and experience in litigating these kinds of cases. Unfortunately, the Telecommunications Act of 1996 does nothing to reduce the need for or likelihood of litigation when these corporations decide to erect cellular phone towers in residential neighborhoods.

III. THE TELECOMMUNICATIONS ACT OF 1996 AND ITS EFFECT ON LOCAL REGULATION OF CELLULAR TOWER FACILITIES

The Telecommunications Act of 1996 describes itself as "[a]n Act to promote competition and reduce regulation in order to secure lower prices

---

23. *See* Martin, *supra* note 16, at 245-46.

24. *See, e.g.*, Oldham County Planning & Zoning Comm'n v. Courier Communications Corp., 722 S.W.2d 904 (Ky. Ct. App. 1987); Nynex Mobile Communications Co. v. Hazlet Township Zoning Bd. of Adjustment, 648 A.2d 724 (N.J. Sup. Ct. App. Div. 1994); Kingwood Township Volunteer Fire Co. Number One v. Board of Adjustment, 640 A.2d 356 (N.J. Sup. Ct. Law Div. 1993); New Brunswick Cellular Telephone Co. v. Old Bridge Township Planning Bd., 636 A.2d 588 (N.J. Sup. Ct. Law Div. 1993); Cellular Telephone Co. v. Rosenberg, 624 N.E.2d 990 (N.Y. 1993); Cellular Telephone Co. v. Village of Tarrytown, 624 N.Y.S.2d 170 (App. Div. 1995); Cellular Telephone Co. v. Meyer, 607 N.Y.S.2d 81 (App. Div. 1994).

But see Awacs, Inc. v. Warwick Township Zoning Hearing Bd., 656 A.2d 608 (Pa. 1995) (affirming Zoning Hearing Board's denial of telecommunications company's application for variance).

25. *See, e.g.*, Payne v. Taylor, 578 N.Y.S.2d 327 (App. Div. 1991); Jaffee v. RCI Corp., 500 N.Y.S.2d 427 (App. Div. 1986); Hawk v. Zoning Hearing Bd., 618 A.2d 1087 (Pa. Commw. Ct. 1992); Hilltop Terrace Homeowner's Assoc. v. Island County, 891 P.2d 29 (Wash. 1995).

and higher quality services for American telecommunications consumers and encourage the rapid deployment of new telecommunications technologies."[26] To accomplish those goals, the Act provides in subsection 253(a) that "[i]n general—[n]o State or local statute or regulation, or other State or local legal requirement, may prohibit or have the effect of prohibiting the ability of any entity to provide any interstate or intrastate telecommunications service."[27] When the House of Representatives was debating a version of the Act, some members, while agreeing that local communities should not be able to prohibit access to new communications facilities, expressed concern that the foregoing language might have the undesirable result of keeping counties, cities, and towns from enforcing their zoning and building codes.[28] One member declared that nothing in the Act should "preempt[] the ability of local officials to determine the placement and construction of ...new [cellular phone] towers. Land use has always been, and ... should continue to be, in the domain of the authorities in the areas directly affected."[29]

The Act does go on to say in subsection 253(b), that states shall maintain their ability "to impose ... requirements necessary to ... protect the public safety and welfare ...."[30] However, that language is followed, in section 253(d), by the warning that if the FCC[31] "determines that a State or local government has permitted or imposed any statute, regulation, or legal requirement that violates subsection (a) ... the Commission shall preempt the enforcement of such statute, regulation, or legal requirement to the extent necessary to correct such violation or inconsistency."[32]

In its instructions to the FCC regarding the regulation of mobile communications services, Congress directed the Commission to consider "safety of life and property," "efficiency," "competition," and the provision of services to the "largest number of feasible users."[33] Congress also specified that states and local governments could not keep companies from providing mobile services or regulate the rates they could

---

26. Telecommunications Act of 1996, Pub. L. No. 104-104, 110 Stat. 56.

27. 47 U.S.C. § 253(a) (1996).

28. 141 CONG. REC. H8269 (daily ed. Aug. 2, 1995) (statements of Reps. Moran, Clyburn, and Goss).

29. *Id.* (statement of Rep. Goss)

30. 47 U.S.C. § 253(b) (1996).

31. The FCC was created to execute and enforce federal law related to communications in order to make "communication by wire and radio" available to everyone in the nation rapidly, efficiently, and at reasonable prices. 47 U.S.C. § 151 (1934) (amended 1996).

32. 47 U.S.C. § 253(d) (1996).

33. 47 U.S.C. § 332(a) (1996).

charge, but states could regulate other terms and conditions of mobile communications services.[34] Specifically, states and local governments can regulate "the placement, construction, and modification" of service facilities with the following limitations.[35] State and local regulation may not "unreasonably discriminate among providers" or "prohibit or have the effect of prohibiting the provision of personal wireless services."[36] Furthermore, when a communications service provider requests authorization to construct facilities, the state or local government must act on the request "within a reasonable period of time" and must support any decision to deny a request with "substantial evidence contained in a written record."[37] Any provider issued such a denial or adversely affected by a failure to respond to such a request may, within thirty days, commence an action in any court with jurisdiction, and the court must hear and decide the case "on an expedited basis."[38] The Act also specifically prohibits states and local governments from regulating the placement and construction of communications facilities, like antennas and towers, on the basis of the environmental effects of electromagnetic fields if the facilities meet FCC standards for emissions.[39] If states or local governments ignore this prohibition, then any provider adversely affected may petition the FCC for relief.[40]

---

34. 47 U.S.C. § 332(c)(3)(A) (1996).

35. 47 U.S.C. § 332(c)(7)(A) (1996).

36. 47 U.S.C. § 332(c)(7)(B)(i)(I) & (II) (1996).

37. 47 U.S.C. § 332(c)(7)(B)(ii) & (iii) (1996).

38. 47 U.S.C. § 332(c)(7)(B)(v) (1996).

39. 47 U.S.C. § 332(c)(7)(B)(iv) (1996). FCC limits on exposure to electromagnetic emissions are based on a measure of the rate of radiofrequency energy absorption, the specific absorption rate (SAR). The agency has set those limits at four watts per kilogram based on American National Standards Institute (ANSI) guidelines set in 1982. 47 C.F.R. pts. 1, 2, 15, 24, 97 (1996).

One complaint from critics is that the FCC allows providers to "self-certify" that they meet the standards. Evelyn Gilbert, *Lethal Lampposts? Cell Phone Antennas May Threaten Your Health*, VILLAGE VOICE, Aug. 26, 1997, at 24. Another complaint is that the standards have been set too low. *See generally*, PAUL BRODEUR, THE GREAT POWER-LINE COVER-UP: HOW THE UTILITIES AND THE GOVERNMENT ARE TRYING TO HIDE THE CANCER HAZARD POSED BY ELECTROMAGNETIC FIELDS (1993).

40. 47 U.S.C. § 332(c)(7)(B)(v) (1996). The Cellular Telecommunications Industry Association (CTIA) has petitioned the FCC for a declaratory ruling that prohibits local governments from creating moratoria on the siting of telecommunications facilities. Comments on the petition were due on September 11, 1997 and replies were due on September 26, 1997. 16 FCC DAILY DIG. 145 (July 29, 1997).

Both cellular phone service providers and local community zoning boards opposing proposed tower facilities can claim some support in the Act for their positions. The Act gives the latter the right to use zoning regulations to protect the welfare of citizens threatened by towers; it imposes limitations, however, such that the advantage is clearly with communications corporations. Allowing states and local governments to regulate the placement of cellular phone towers, except when such regulation will have the effect of prohibiting the provision of cellular phone service, will give the service providers a very easy argument for having any regulation voided: if they are denied a variance to use the site of their choice, the service providers will assert that any other site would not be as cost-effective and, therefore, either they must be given permission to use their chosen site or they will not bring service to the local area. Moreover, requiring the expeditious resolution of these disputes gives a distinct advantage to the corporations that have staffs of lawyers and engineers, previously prepared research, and litigation experience with similar cases. Local residents have none of these, and very limited financial resources with which to try to match the corporations. To require that they quickly catch up to their opponent's levels in research and expert support renders the residents position untenable in most cases.

The Act pays lip service to the importance of local zoning regulation, sufficient to encourage litigation, but without any genuine recognition of the importance of a homeowner's property values, peace of mind, and, particularly, health concerns. The Act denigrates health concerns by assuming that FCC standards for electromagnetic emissions will protect the public health. That assumption is premature, given the large amount of ongoing scientific research on the subject and the lack of clear conclusions. The Congressional Conference Report indicates that the Act preempts state and local regulation of the environmental effects of electromagnetic emissions when it has requirements beyond those of FCC rules.[41] This preemption discourages states from doing their own research on the health effects of these emissions because they cannot rely on the results in formulating regulations.[42] That result does a disservice to the public. The FCC, in promulgating its rules setting a specific absorption rate limit for electromagnetic emissions at four watts per kilogram, noted that research in this area related to human health and

41. S. Rep. No. 104-230, at 208 (1996).

42. Some states have indicated a desire to do such research and to regulate facilities that emit electromagnetic radiation based on the results of that research. *See, e.g.*, Letter from Michele C. Farquhar, Chief, Wireless Telecommunications Bureau of the FCC, to Thomas E. Wheeler, President and CEO, Cellular Telecommunications Industry Association (Jan. 13, 1997).

safety is ongoing and that changes to recommended exposure limits are possible in the future.[43] With that admitted uncertainty, it is unreasonable to limit what states and local governments may do to protect their residents.

Some local governments have imposed temporary moratoria on the issuance of such permits, to allow themselves time to study the impact of cellular communications antennas and towers before granting permission for their construction.[44] In early 1997, the Cellular Telecommunications Industry Association responded by filing a petition with the FCC for a declaratory ruling seeking preemption of such moratoria on the grounds that they violate the Telecommunications Act of 1996 and that the Act authorizes FCC preemption.[45] The FCC should deny the petition using the reasoning articulated by the United States District Court for the Western District of Washington in one of the few cases concerning the siting of telecommunications towers decided since the Act went into effect.[46]

IV. JUDICIAL RESOLUTION OF TELECOMMUNICATIONS TOWER SITING DISPUTES SINCE THE ENACTMENT OF THE TELECOMMUNICATIONS ACT

Three months after the Act was signed into law, a federal district court in Washington State decided a case challenging a six-month moratorium on issuing permits for new telecommunications facilities established by the City of Medina.[47] Medina has about 3,000 residents,

---

43. *See* Guidelines for Evaluating the Environmental Effects of Radiofrequency Radiation, 61 Fed. Reg. 41,006, 41,007 (1996) (to be codified at 47 C.F.R. pts. 1, 2, 15, 24, 97).

44. *See, e.g.*, Farquhar, *supra* note 42; Sprint Spectrum v. City of Medina, 924 F. Supp. 1036 (W.D. Wash. 1996).

45. *See* Public Comment Invited; Commission Seeks Comment on Petition for Declaratory Ruling of the Cellular Telecommunications Industry Association, 62 Fed. Reg. 4047 (Jan. 28, 1997).

46. Sprint Spectrum v. City of Medina, 924 F. Supp. 1036 (W.D. Wash. 1996). Other cases involving disputes about the siting of telecommunications towers and referring to the Telecommunications Act of 1996 include Paging, Inc. v. Board of Zoning Appeal, 957 F. Supp. 805 (W.D. Va. 1997); BellSouth Mobility, Inc. v. Gwinnett County, Georgia, 944 F. Supp. 923 (N.D. Ga. 1996); Crown Communications v. Zoning Hearing Bd., 679 A.2d 271 (Pa. Commw. Ct. 1996); Westel-Milwaukee Co., Inc. v. Walworth County, 556 N.W.2d 107 (Wis. Ct. App. 1996).

47. *City of Medina*, 924 F. Supp. at 1040. Enacting a temporary moratorium on the grant of permission to erect telecommunications towers is a technique that is growing in popularity with local governments. CTIA reports that nationwide there were 226

and is approximately two and one-half square miles in area, zoned entirely for low-density residential use. It is a prime location for cellular phone towers, however, because of its proximity to a state highway and a bridge.[48] For several years, Medina has had cellular phone facilities belonging to two service providers, but after the Act became effective, the city expected additional applications for tower construction permits and feared becoming an "antenna farm."[49] Five days after the effective date of the Act, Medina's moratorium went into effect in order to give the city time to study the allocation of suitable sites.[50] One month later, Sprint filed a lawsuit alleging that the moratorium violates the Act because any delay in its obtaining full cellular phone coverage in the region would cause it to lose a great deal of money resulting in irreparable harm to the company.[51]

The court noted that Medina citizens were concerned about the health hazards and negative aesthetic effects associated with cellular phone towers, but emphasized that if the city did not have time to study the appropriate siting of facilities, there may not be adequate sites for competing providers.[52] Thus, without the careful allocation of sites, beneficial services might be rendered unavailable.[53]

The court provided an instructive analysis of the relevant portions of the Act. To Sprint's claim that the moratorium "'prohibit[s] or [has] the effect of prohibiting the provision of personal wireless services,'" the court responded that the moratorium was not a prohibition, but merely a short-term suspension.[54] To Sprint's contention that the moratorium kept the city from "'act[ing] on' its application 'within a reasonable period of

---

moratoria in effect at the end of June, 1997, a thirty-four percent increase from April. Karissa Booney, *Getting Over It*, WIRELESS WORLD, Oct. 30, 1997, at 1. *See, e.g.*, Ted Cohen, *Kennebunk OKs Rules for Cellular Towers - A Unanimous Vote at a Special Town Meeting Finally Gives the Town Control Over the Location and Height of the Towers*, PORTLAND PRESS HERALD, Oct. 22, 1997, at 1B (voters in Arundel, Maine considering six-month moratorium); Gary Gerew, *Board Towers Over Antenna Moratorium - A Proposal to Ban Installing Cellular Phone Towers Will Be the Topic of a Public Hearing Nov. 10*, THE POST-STANDARD (Syracuse, N.Y.), Oct. 23, 1997, at 3; Catherine Kozak, *Colington Tower Fails to Win OK*, VIRGINIAN-PILOT & LEDGER-STAR, Oct. 15, 1997, at B1 (moratorium to be proposed in Dare County, Va. in Nov.; moratorium imposed in Currituck County, Va.).

48. *City of Medina*, 924 F. Supp. at 1037.

49. *Id.*

50. *Id.* at 1037-38.

51. *Id.*

52. *Id.* at 1038.

53. *Id.*

54. *Id.* at 1040.

time,'" the court averred that the language in the Act did not suggest that Congress "intended to force local government procedures onto a rigid timetable where the circumstances call for study, deliberation, and decision-making among competing applicants."[55] The court concluded that the Act's legislative history indicated that Congress did not intend to give preferential treatment to the telecommunications industry in the processing of zoning applications.[56]

Finally, the court held that Medina's six-month moratorium on issuing new permits for telecommunications facilities for the purpose of information-gathering did not violate any provisions of the Telecommunications Act.[57] Thus, the court interpreted the Act's provisions in a light most favorable to the retention of some local control over the environment in which residents live, keeping the profit-making motives of telecommunications corporations from being the ultimate value in the regulation of telecommunications facilities.

In *BellSouth Mobility, Inc. v. Gwinnett County, Georgia*,[58] the United States District Court for the Northern District of Georgia interpreted section 332(c)(7)(B)(iii) of the Act, which requires that a denial of a telecommunications service provider's application to construct facilities be "supported by substantial evidence contained in a written record."[59] In this case, BellSouth applied to the Gwinnett County Board of Commissioners for a permit to erect a 197-foot monopole that would improve the quality of its cellular telephone service.[60] In support of its application, BellSouth provided the following documents: a report by the Airspace Safety Analysis Corporation showing that the monopole was not hazardous to aircraft, a certified appraiser's report concluding that monopoles did not decrease property values, a line of sight survey, prepared by Aerial Instrument Research Systems, showing the visibility of a red balloon floated to varying heights at the proposed site, and a list of BellSouth's unsuccessful efforts to find other suitable sites.[61] Residents, on the other hand, submitted no documents, relying merely on a representative who attended a Board hearing and made conclusory expressions of concern regarding the monopole's safety, health, aesthetic,

55. *Id.*

56. *Id.*

57. *Id.*

58. 944 F. Supp. 923 (N.D. Ga. 1996).

59. *Id.* at 928.

60. *Id.* at 924.

61. *Id.* at 925-26.

and economic threats.[62] Based on this record the court held that the Board violated the "substantial evidence" provision in the Act.[63]

The court then had to decide on the appropriate remedy.[64] Although the Act allows anyone denied a permit for telecommunications facilities to seek relief "in any court of 'competent jurisdiction'," it does not specify the remedy for violation of the Act.[65] The choices available to the district court were to remand the matter to the Board for it to make a decision supported by substantial evidence or to order the Board to issue the permit for the monopole. The court did the latter, explaining that the Act requires the court to "hear and decide such action on an expedited basis" and, therefore, a mere remand would thwart the intent of the Act to encourage the expeditious installation of new telecommunications facilities.[66]

This case illustrates why, in fairness and concern for citizens' ability to exert their rights, the Act should give some deference to local governments in their disputes with telecommunications corporations regarding the location of telecommunications towers and antennas. BellSouth had knowledge, experience, legal counsel, and the financial resources to have experts prepare reports in support of its application for a permit to erect a 197-foot monopole. Local residents had none of those resources. That lack does not necessarily indicate that there was no substantial evidence to support their position, but perhaps merely that they did not know they needed it, did not know where to get it, or did not have the financial resources to pay for it. Moreover, in this kind of situation, government representatives may not be of much help, because they are also lay people with budgetary constraints and, therefore, they are no match for business adversaries.[67]

Illinois RSA No. 3, Inc. v. County of Peoria[68] also illustrates the poor preparation of the residents who opposed the construction of a 140-foot cellular transmission monopole. The United States District Court for the

---

62. *Id.* at 926.

63. *Id.* at 928.

64. *Id.* at 929.

65. *Id.*

66. *Id.*

67. For a description of a rural community's dispute with a large telecommunications corporation over the latter's construction of a 300-foot tower, *see* Martin, *supra* note 16, at 250-55 (occurring before the enactment of the Telecommunications Act). The community reached a successful compromise with the company probably because it mounted a more sophisticated campaign (that included written supporting materials from experts), *id.* at 252, than those described in the Georgia case and the Illinois, and New Mexico cases, *infra*.

68. 963 F. Supp. 732 (C.D. Ill. 1997).

Central District of Illinois noted that the local zoning board received a petition signed by 200 people opposing the monopole, but that there was no indication of the basis for their opposition.[69] A realtor with twenty-four years experience objected to the monopole because it would cause a decrease in property values, but offered no analysis, studies or examples to support the reasonableness of the objection.[70] Lastly, the residents presented a survey that was meant to show that potential home buyers would not buy a home near a telecommunications tower.[71] The court concluded that there was no evidence of the survey's statistical or scientific merit, however, because there was no information on how the survey was conducted or how the respondents were chosen.[72]

On the other hand, Illinois RSA, the telecommunications provider, presented evidence from three certified real estate appraisers indicating that cellular transmission towers do not cause real estate prices to fall.[73] One presented an analysis of similar tower sitings at other locations that indicated that towers did not have an adverse effect on property values.[74] Illinois RSA had an engineer and surveyor present line of sight drawings demonstrating that the tower would not be visible from nearby residences.[75]

The Peoria residents also stated their concerns about the health effects of living close to telecommunications transmission facilities, but the court held that under the Telecommunications Act, health effects could not be considered as long as emissions were within the standard set by the FCC.[76] Thus, the court concluded that there was no substantial evidence, as required by the Telecommunications Act,[77] to deny Illinois RSA's request to construct its tower.[78] It also concluded that the county of Peoria had violated the Act "in the most basic way" by not issuing a written statement containing the reasons for its denial.[79] In deciding on a remedy, the Illinois district court, citing the *Gwinnett County* case, rejected the option of remanding the case to the county zoning board for

---

69. *Id.* at 737-38.

70. *Id.* at 738.

71. *Id.* at 739.

72. *Id.* at 745.

73. *Id.*

74. *Id.* at 739.

75. *Id.*

76. *Id.* at 745 (citing 47 U.S.C. § 332(c)(7)(B)(iv)).

77. 47 U.S.C. § 332(c)(7)(B)(iii) (1996).

78. 963 F. Supp. at 743.

79. *Id.*

reconsideration and a written decision.[80] The court concluded that such a course "would be a waste of time and would frustrate the Telecom Act's direction to expedite these proceedings."[81] Instead, the court issued an injunction directing the county to issue a permit for the tower and to remove any obstacles to its construction.[82]

The United States District Court in New Mexico also cited *Gwinnett County* in providing mandamus relief for Western PCS II Corporation, a telecommunications company that had been denied a special exception request to mount antennas on an existing water tank by the zoning authority for Santa Fe.[83] In this case, the Santa Fe zoning authority failed to comply with what the district court deemed the "most basic of the Telecommunications Act's requirements," a written record supporting its denial of the company's request.[84] This led the judge to resist remanding the matter, because the court could not find the "substantial evidence" upon which the zoning authority must rely to sustain its denial of a permit.[85] The only evidence submitted by those opposed to the antennas was the expression of "generalized concerns" by several neighbors.[86] Moreover, those concerns centered on a "visual blight in the neighborhood," even though the antennas were going to be no higher than the already-existing water tank, they were going to be painted to match the color of the tank, and Western PCS was going to remove graffiti from the water tank.[87] As presented by the court, the facts of this case make the objectors' case seem very weak, but it is hard to know whether it was objectively weak or just poorly presented.

In contrast to the federal district courts in Georgia, Illinois, and Mew Mexico, the state court of appeals in Wisconsin held that, in light of the Act, a remand to the local zoning authority for reconsideration of its decision to deny a permit for the construction of a 200-foot

---

80. *Id.* at 747.

81. *Id.*

82. *Id.*

83. Western PCS II Corp. v. Extraterritorial Zoning Auth. of the City and County of Santa Fe, 957 F. Supp. 1230, 1233-34 (D.N.M. 1997), *notice of appeal filed*, (10th Cir. Mar. 25, 1997). The city of Santa Fe, New Mexico has filed a notice of its intent to appeal the district court's writ of mandamus instructing the Extraterritorial Zoning Authority to approve the special exception request to mount telecommunication antennas on an existing water tank.

84. *Id.* at 1236.

85. *Id.* at 1237

86. *Id.* at 1236.

87. *Id.* at 1234-37.

telecommunications tower was an appropriate remedy.[88] The Wisconsin court considered the Act's language requiring courts to hear these cases on an expedited basis, but did not relate that mandate to the remedies available to courts. The body of case law on the subject is still much too small to draw any general conclusions, but it will be interesting to note whether any pattern emerges of federal district courts construing the Act strictly, or of state courts deferring to local zoning authorities.[89]

These cases suggest that the Act has not sufficiently clarified the role of state and local governing bodies in making decisions about the siting of cellular phone towers to discourage litigation. To the contrary, the statute creates new questions about what constitutes "the effect of prohibiting the provision of personal wireless services," what is "a reasonable period of time" for acting on requests to construct telecommunications facilities, and what kind of regulating is actually left for local governments to do regarding such construction. The latter question includes the specific issue of what health and welfare or safeguarding the rights of "consumers" can mean, particularly when state and local governments cannot consider the possible effects of human exposure to electromagnetic fields.

V. DISCUSSION AND CONCLUSIONS

Congress clearly intended for the 1996 Act to limit state and local regulation of the telecommunications industry.[90] The idea was to eliminate regulatory barriers to promote competition in the industry in order to encourage technological advancement and to give consumers

---

88. Westel-Milwaukee Co. v. Walworth County, 556 N.W.2d 107, 107-110 (Wis. Ct. App. 1996).

89. One case that might support the latter is Crown Communications v. Zoning Hearing Bd., 679 A.2d 271 (Pa. Commw. Ct. 1996). In this case Crown's application to erect a 375-foot telecommunications tower was denied by the local zoning board. *Id.* at 272. Crown appealed and the lower court overturned the denial. *Id.* The appellate court reinstated the denial. *Id.* at 275. The appellate court's decision was based on a state law issue, and the court concluded that Crown had waived its challenge based on the Telecommunications Act. *Id.* at 275 n. 11. The court noted, however, that had it considered Crown's argument that the zoning board violated section 332(c)(7)(B)(i) of the Act, which disallows any state or local regulation that has "the effect of prohibiting the provision of personal wireless services," it still would have upheld the denial of Crown's application. *Id.* The court gave no explanation for its gratuitous remark.

90. *See, e.g.*, Implementation of the Local Competition Provisions in the Telecommunications Act of 1996, 61 Fed. Reg. 45,476, 45,489 (1996) (to be codified at 47 C.F.R. pts. 1, 20, 51, 90).

choices.[91] In its zeal to accomplish these goals, however, Congress neglected to sufficiently consider the interests of local residents, other than the interests they have as consumers of telecommunications services, and the advantages given to the industry vis-à-vis citizens.[92] To rectify this oversight, Congress should amend the Telecommunications Act in four specific ways.

First, the Act should clarify the conditions of mobile communications services that state and local governments can regulate. These conditions should include the siting of facilities and the specific form that the facilities take, although the regulations should not result in the barring of service in the area. It is reasonable for people to be concerned about the effects of proximity to cellular phone towers on health, their property values, and the aesthetics of a home's landscape. It is unfair to dismiss these interests as merely symptoms of the "not-in-my-backyard" (NIMBY) syndrome and, therefore, interests to be ignored when the proliferation of cellular phone sites is at stake. In fact, there is nothing in the Telecommunications Act of 1996 that encourages telecommunications companies to take these interests into consideration in siting their facilities.

Second, Congress should require service providers to include substantial evidence that they are requesting siting permits for the least intrusive facilities available in the least intrusive locations under the circumstances. Such a requirement, in addition to addressing some of the concerns of local residents, would promote the congressional goals of advancing technology and encouraging competition. There are many ways of making communications facilities less intrusive—hiding antennas is one[93]—but they may be more expensive than the installation of a traditional 200 or 300-foot tower.[94] For example, microcells do not have

---

91. *See, e.g.*, 141 CONG. REC. S7881-02, S7886 (daily ed. June 7, 1995) (statement of Sen. Pressler); 142 CONG. REC. H1149 (daily ed. Feb. 1, 1996) (statement of Rep. Fields).

92. This is not the only area in which provisions of the Telecommunications Act of 1996 have resulted in disadvantages to the public that were probably not considered by many members of Congress. For example, the Act's emphasis on competition is causing a substantial increase in residential phone bills, a result probably not foreseen or desired by Congress. *See* 143 CONG. REC. S2048 (daily ed. Mar. 10, 1997) (statement of Sen. Dorgan noting his dismay that consumers are losers because "the major titans in the telecommunications industry battled for advantage under this act" and the result is going to be increases in residential telephone rates); Mark Landler, *Rising Phone Bills Are Likely Result of Deregulation*, N.Y. TIMES, Mar. 30, 1997, at 1.

93. See *Use Creativity, Education to Tackle Tower Issues, Insiders Say*, ADVANCED WIRELESS COMMUNICATIONS, Jan. 4, 1995.

94. *See* Candice S. Millard, *Mending Fences*, CELLULAR BUS., Dec. 1, 1996 at 40 (noting that price of palm tree tower could be twice as much as that of traditional tower).

the same height and power requirements as macrocells, but a larger number of the microcells are needed to provide widespread coverage.[95] Microcells do not have to be located on high towers; they can be installed in church steeples, on rooftops, and even inside offices where they would not be noticed.[96] They can be attached to utility poles and lamp posts with cables running down to equipment located in underground shelters.[97] There are also coverage enhancer systems that can reduce the number of necessary towers by one third to one half depending on the terrain.[98]

Third, the Act should allow state and local governments to rely on scientifically objective evidence of the health risks associated with electromagnetic fields when making decisions regarding the siting of communications towers and antennas. There are clear advantages to having a national policy on telecommunications. Nevertheless, because there are such wide disparities within the worldwide scientific community about the effects of electromagnetic fields (even at low levels) on human health, it should be up to local communities to decide how much risk they are willing to undertake.

Finally, the Act discourages study and planning with its "prohibit or have the effect of prohibiting the provision of personal wireless services" and "within a reasonable period of time" language.[99] This language should be clarified to allow a realistic amount of time for communities to plan for the best use of their resources. For example, companies can be required by zoning boards to share sites (known as co-location)[100] in an effort to reduce the number of towers, but for a local government to be able to create such requirements supported by substantial evidence, however, it would need the time to study and formulate an all-encompassing plan for the community and potential permit applicants. Current language does not, of course, prohibit planning, but it encourages service providers to commence court actions when a permitting agency does not expeditiously grant a permit application.

Legislative clarification is preferable to the ad hoc decision-making that courts will be required to do. Nevertheless, when judges are presented with these cases, they should keep in mind that the Act specifically allows local regulation of the terms and conditions of

---

95. *See* Renee Saunders, *Target: Total Coverage*, TELEPHONY, Feb. 24, 1997 at 14.

96. *See* Rhonda Wickham and Shawn Steward, *Sizzling Products*, CELLULAR BUS., Jan. 1, 1995, at 22.

97. *See* Millard, *supra* note 95.

98. *Id.*

99. 47 U.S.C. § 332(c)(7)(B)(i)(I) & (II), (ii) (1996).

100. Sandra J. Grove, *Developing a Tower Strategy*, CELLULAR BUS., May 1, 1995, at 27.

telecommunications services, and that the other provisions in the Act cannot render that provision meaningless.

Congress could not have meant for the Telecommunications Act to imply that having cellular phone service is more important to a community than having the freedom to decide what health risks are worth undertaking or than maintaining the value of neighborhood homes: the most valuable asset most homeowners have. Nevertheless, as written, the Act does not give corporations that provide cellular phone service any incentive to work cooperatively with the communities they intend to make their customers. Congress has overestimated the role that competition would play in giving local residents input in the siting of telecommunications towers. Residents and cellular phone customers, particularly in more rural areas, have not had a variety of service providers vying for their business. When there is only one provider in the area, it does not have to curry favor with potential customers by being a good neighbor.

With no evidence that Congress intends to amend the Act in the very near future, and because once towers are erected they are probably in place permanently, it will be up to courts to interpret the Act in an even-handed manner according to its language. If courts give local communities the leeway to regulate the terms and conditions of tower sitings in a thoughtful manner that will not prohibit the availability of service, the damaging effects of a poorly designed statute can be controlled.