Ad Hoc Autonomous Agent Teams: Collaboration without Pre-Coordination

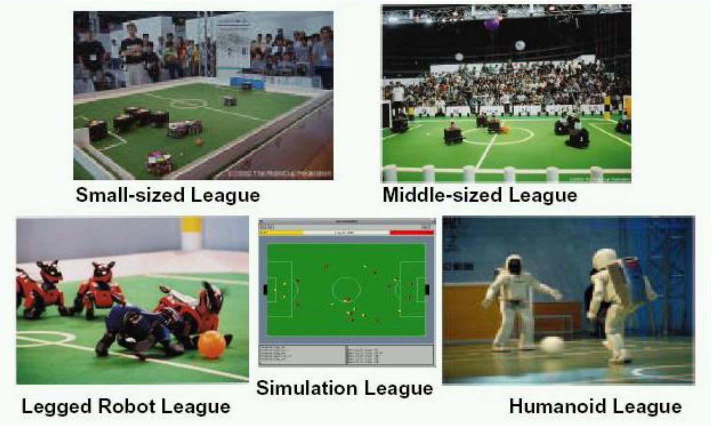
Peter Stone
Director, Learning Agents Research Group
Department of Computer Science
The University of Texas at Austin

Joint work with
Gal A. Kaminka, **Sarit Kraus**, Bar Ilan University
Jeffrey S. Rosenschein, Hebrew University

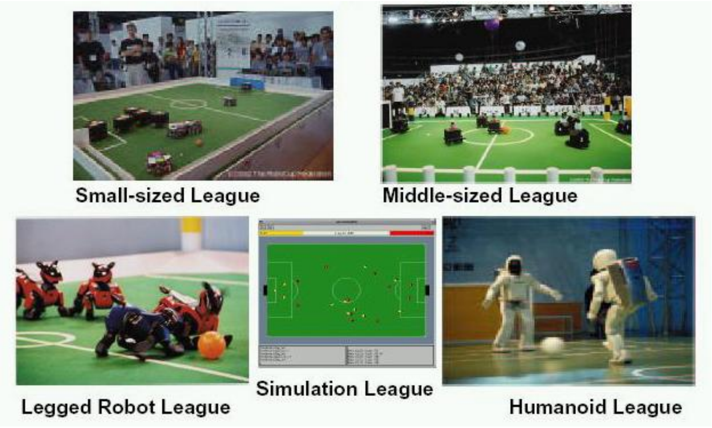
Teamwork

Teamwork


2005/10/ 1 16:55


Small-sized League
Middle-sized League
Legged Robot League
Simulation League
Humanoid League

Teamwork



- Typical scenario: pre-coordination
  - People practice together
  - Robots given coordination languages, protocols
  - "Locker room agreement" (Stone & Veloso, '99)

Ad Hoc Teams

- Ad hoc team player is an individual
  - Unknown teammates (programmed by others)

Ad Hoc Teams

- Ad hoc team player is an individual
  - Unknown teammates (programmed by others)
- May or may not be able to communicate

Ad Hoc Teams

- Ad hoc team player is an individual
  - Unknown teammates (programmed by others)
- May or may not be able to communicate
- Teammates likely sub-optimal: no control

Ad Hoc Teams

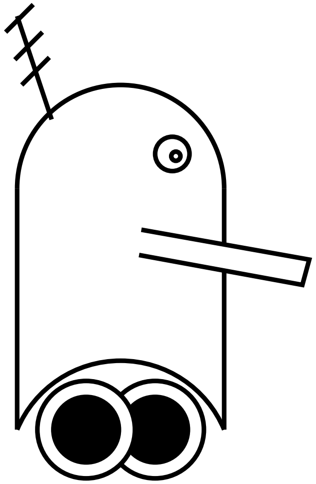
- Ad hoc team player is an individual
  - Unknown teammates (programmed by others)
- May or may not be able to communicate
- Teammates likely sub-optimal: no control

Ad Hoc Teams

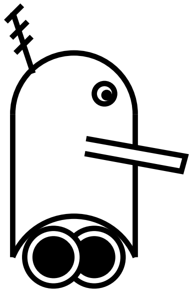
- Ad hoc team player is an individual
  - Unknown teammates (programmed by others)
- May or may not be able to communicate
- Teammates likely sub-optimal: no control

Challenge: Create a good team player

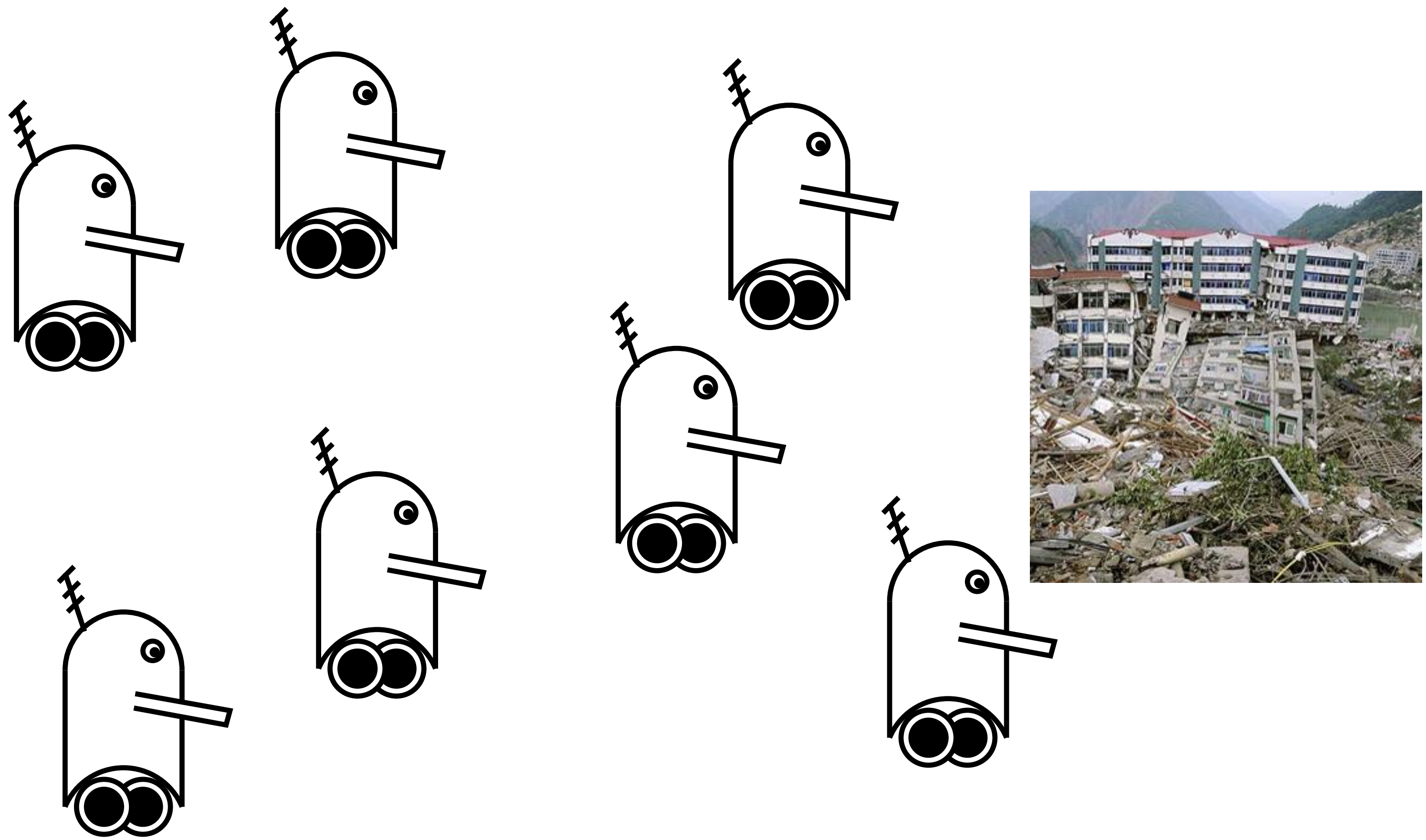
Illustration



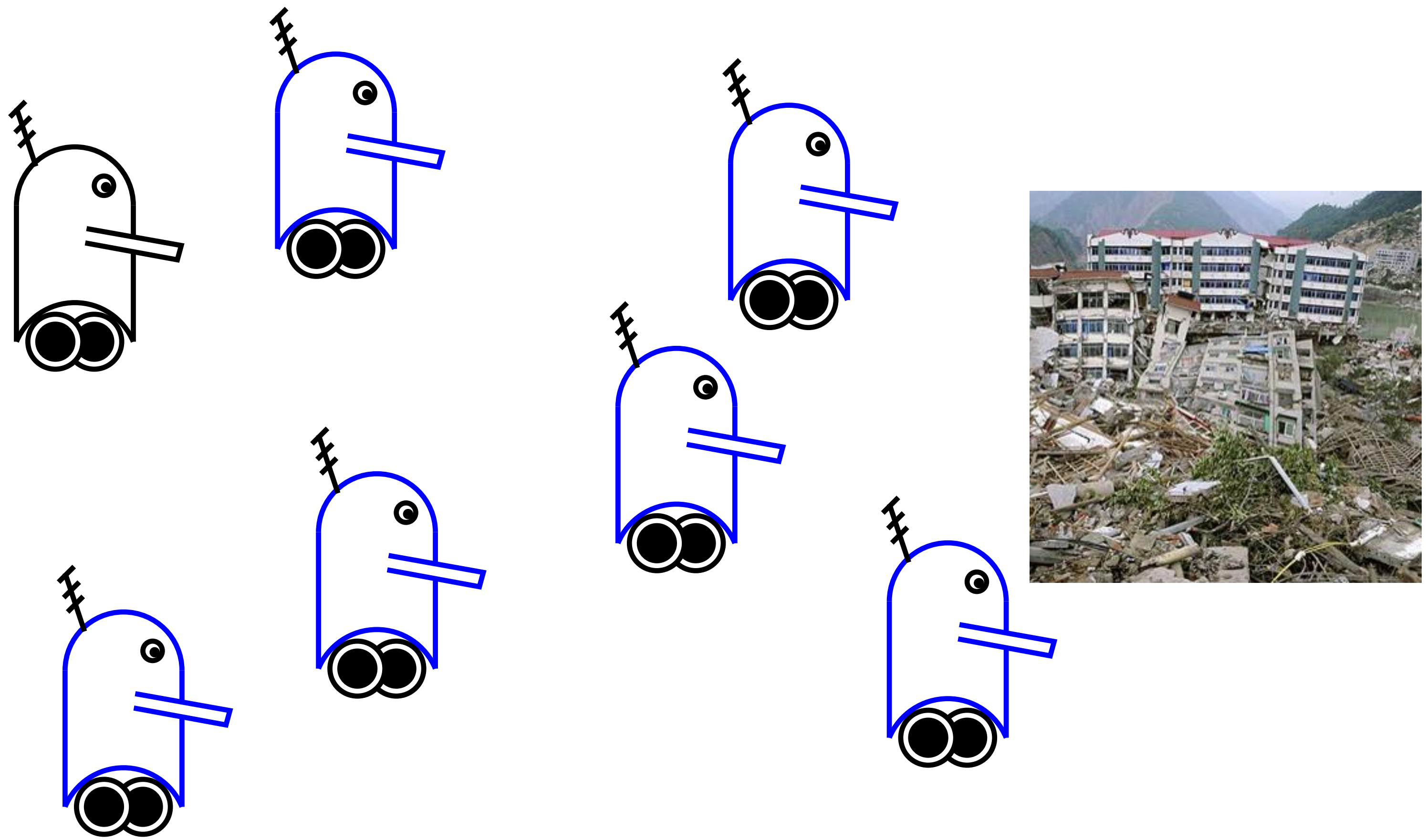
An Individual



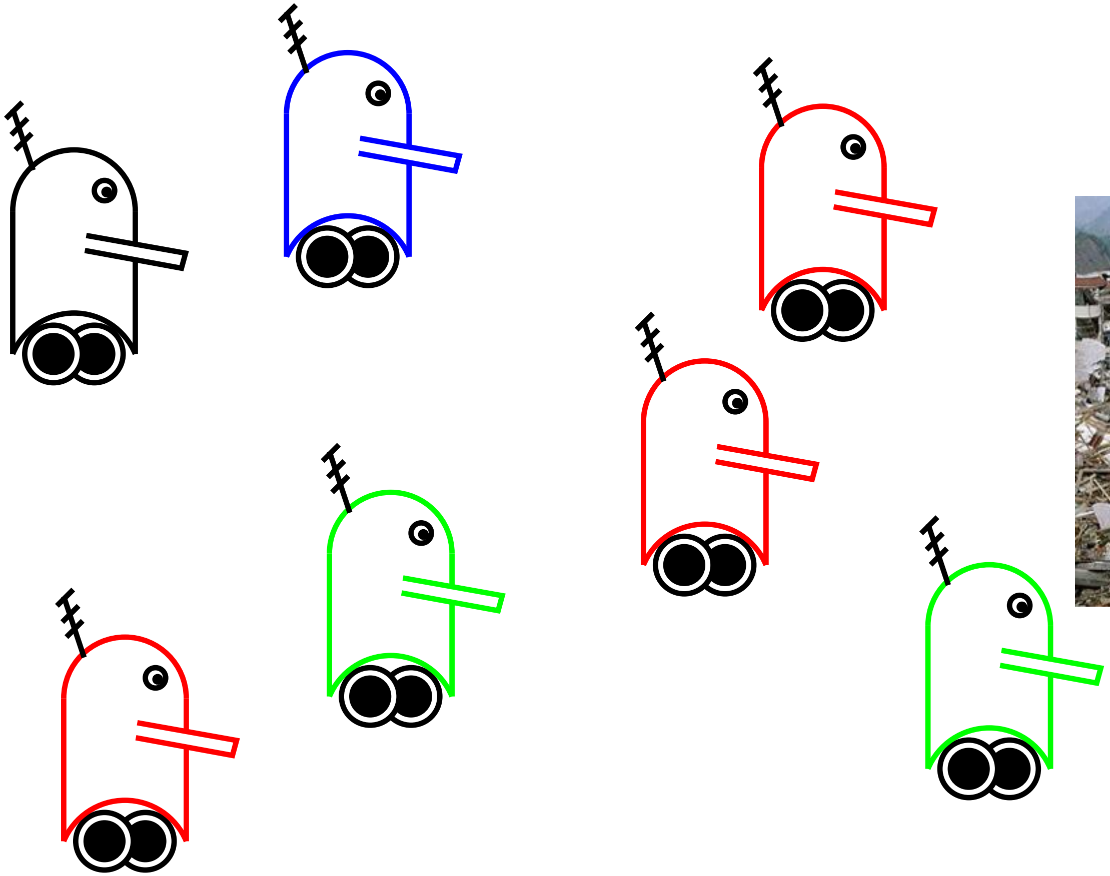
With Teammates



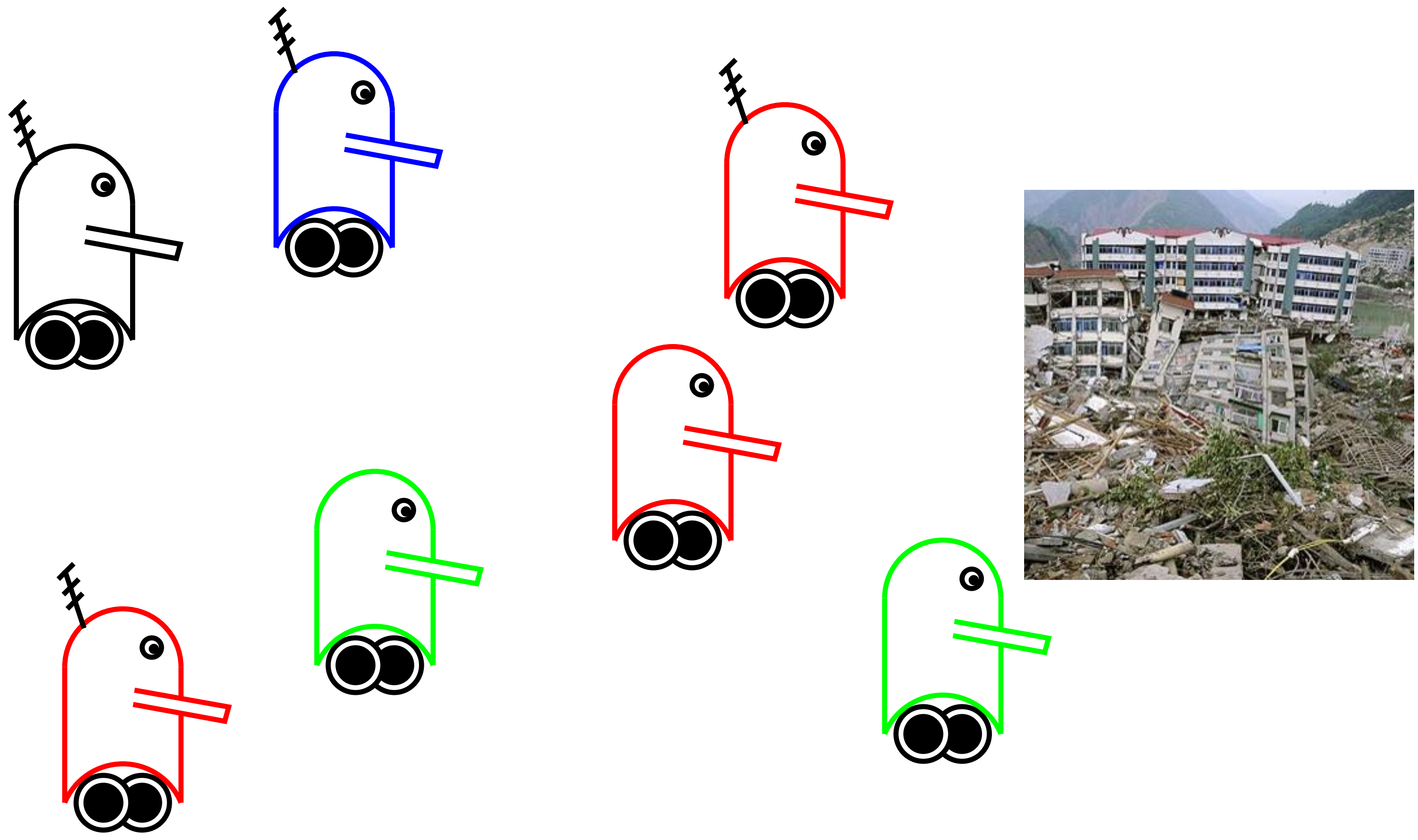
Made by Others



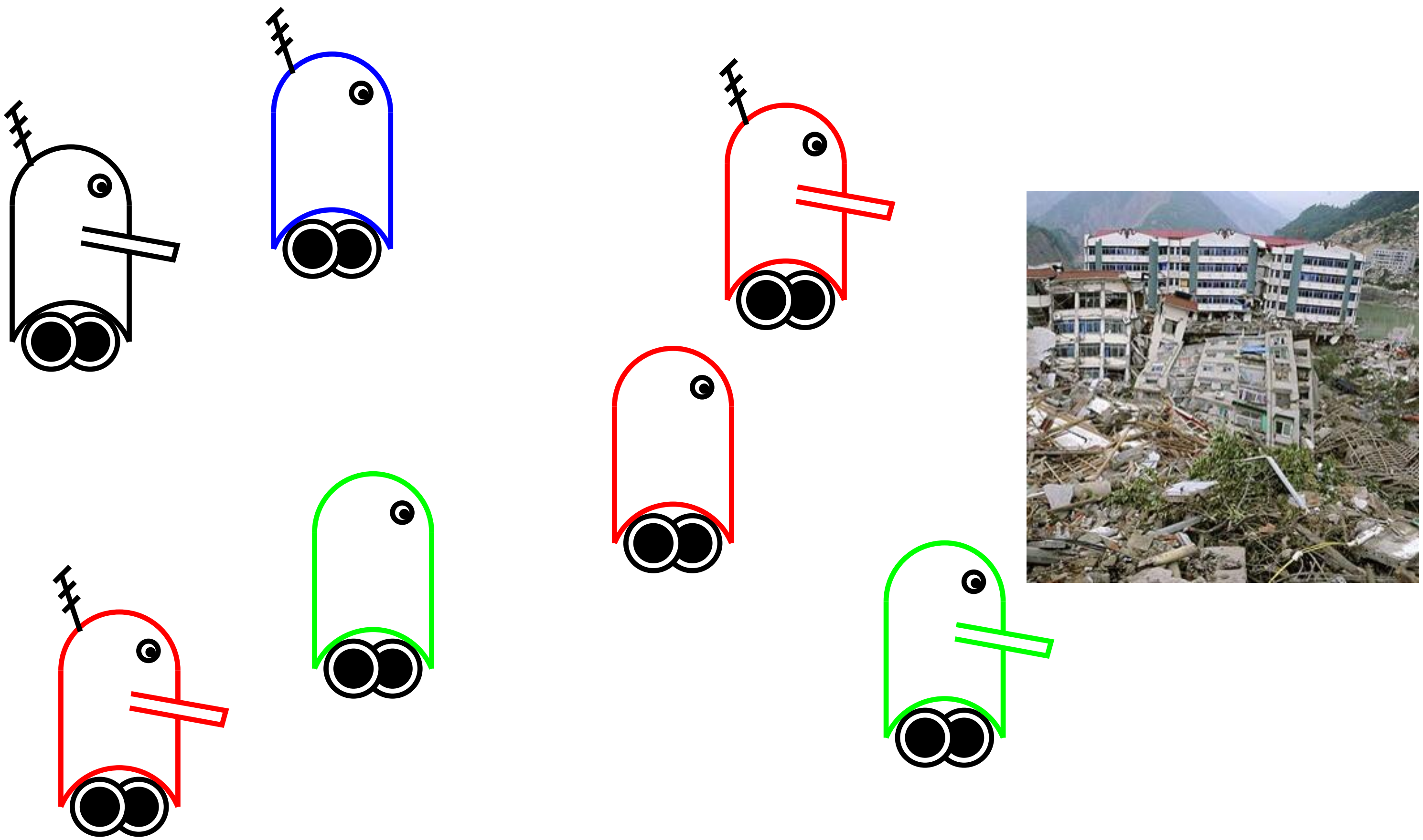
Heterogeneous



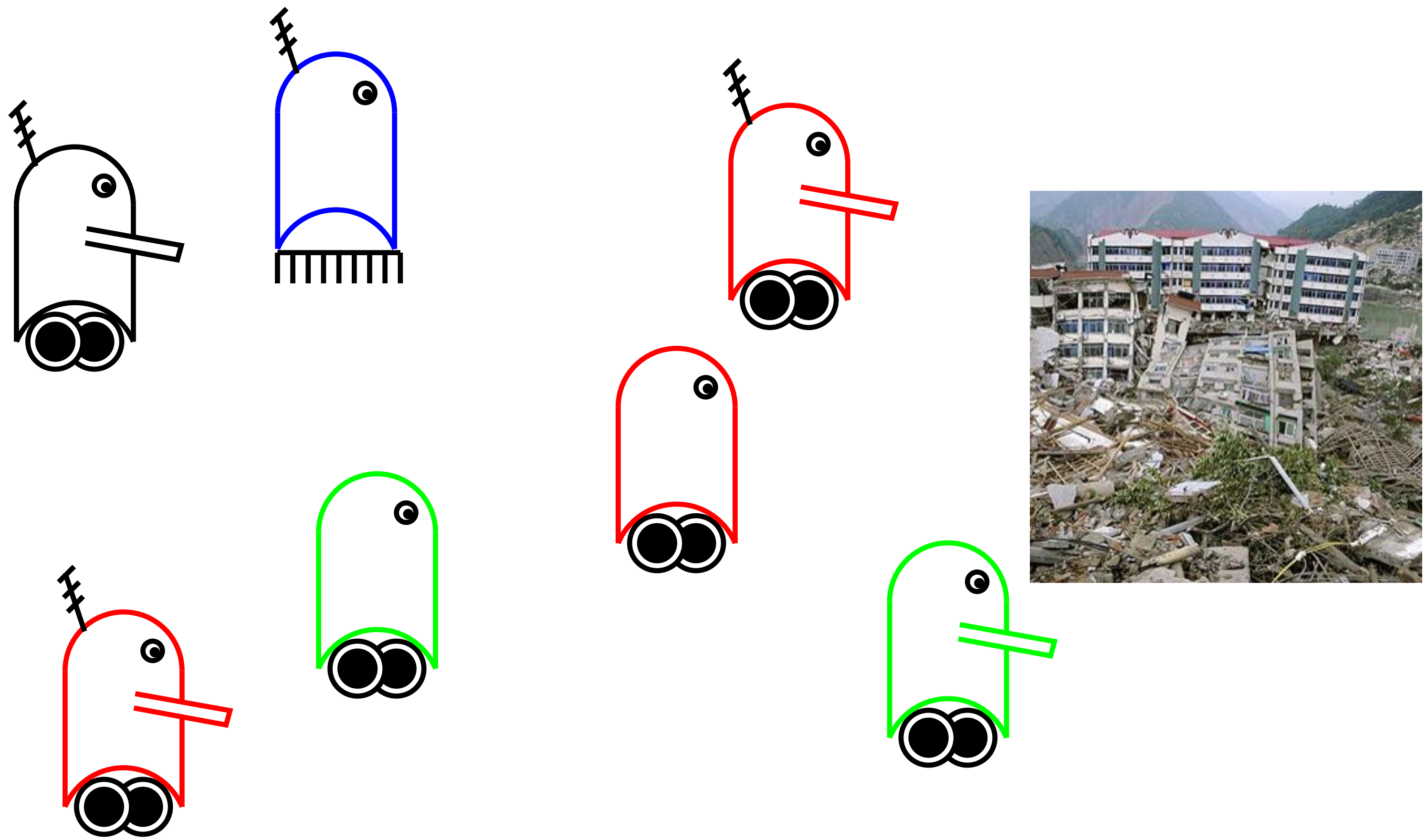
May not Communicate



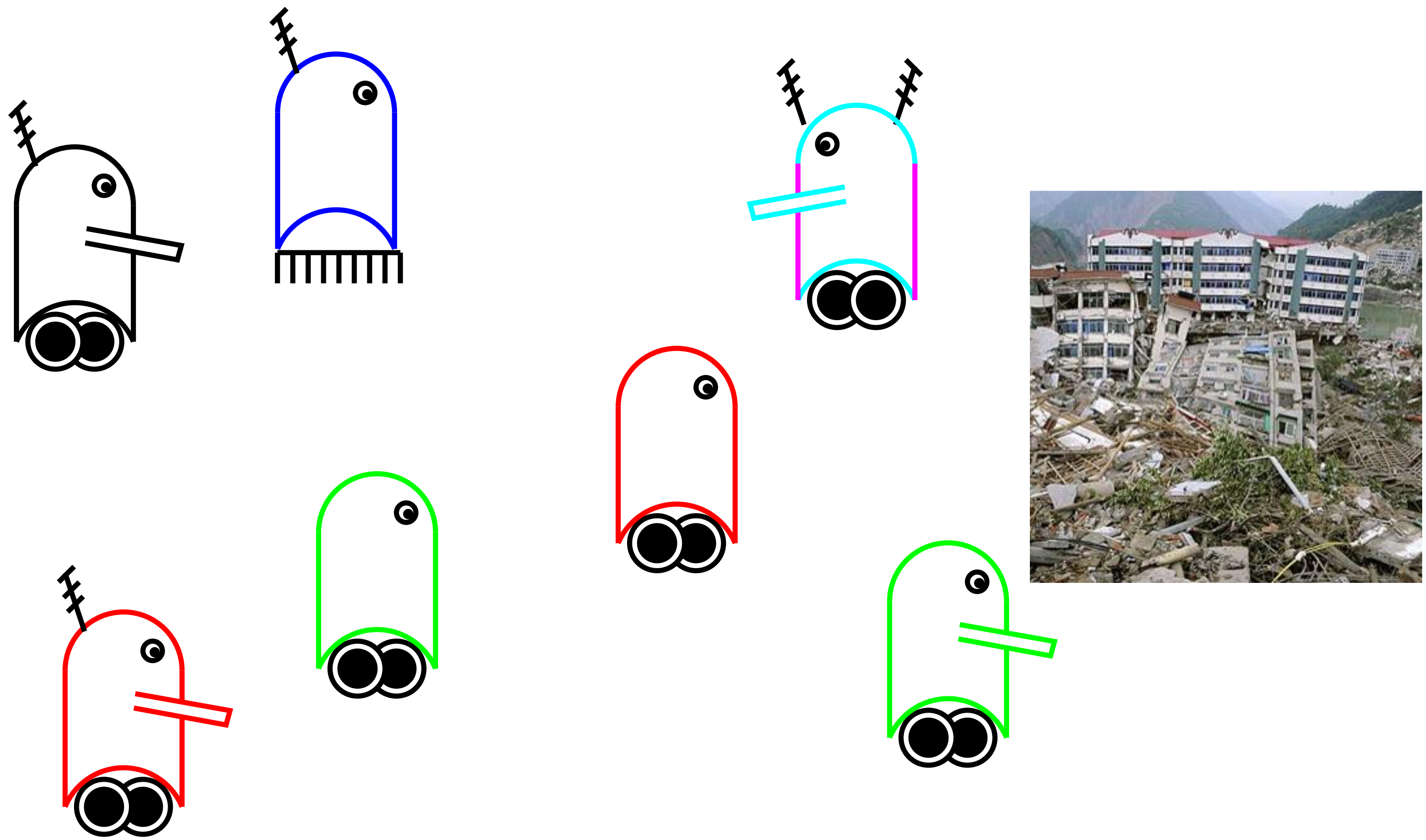
May Have Different Capabilities



And/Or Maneuverability



May be a Previously Unknown Type



Human Ad Hoc Teams

- Military and industrial settings

Human Ad Hoc Teams

- Military and industrial settings
  - Outsourcing

Human Ad Hoc Teams

- Military and industrial settings
  - Outsourcing
- Agents support human ad hoc team formation

(Just et al., 2004; Kildare, 2004)

Human Ad Hoc Teams

- Military and industrial settings
  - Outsourcing
- Agents support human ad hoc team formation

(Just et al., 2004; Kildare, 2004)

- Autonomous agents (robots) deployed for short times
  - Teams developed as cohesive groups
  - Tuned to interact well together

Challenge Statement

Create an autonomous agent that is able to efficiently and robustly collaborate with previously unknown teammates on tasks to which they are all individually capable of contributing as team members.

Challenge Statement

Create an autonomous agent that is able to efficiently and robustly collaborate with previously unknown teammates on tasks to which they are all individually capable of contributing as team members.

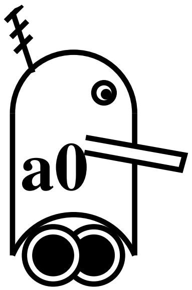
- Aspects can be approached theoretically

Challenge Statement

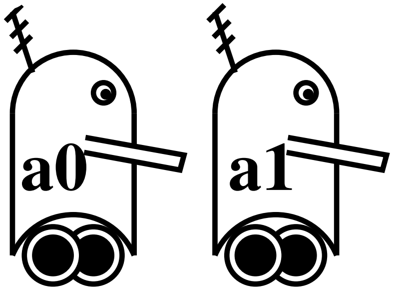
Create an autonomous agent that is able to efficiently and robustly collaborate with previously unknown teammates on tasks to which they are all individually capable of contributing as team members.

- Aspects can be approached theoretically
- Ultimately an empirical challenge

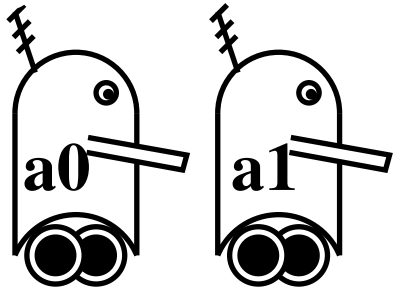
Empirical Evaluation



Evaluation: A Metric

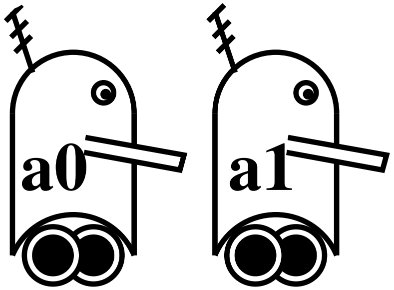


Evaluation: A Metric

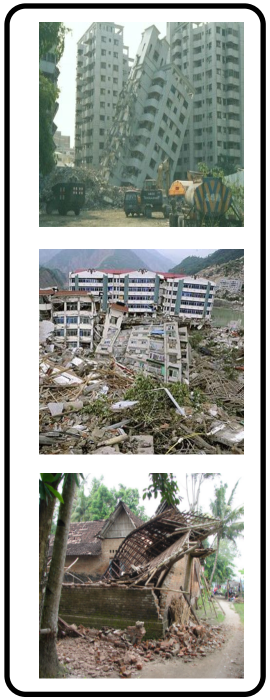


- Most meaningful when a0 and a1 have similar individual competencies

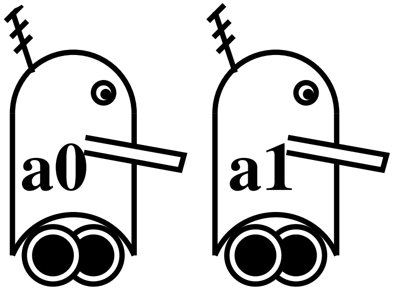
Evaluation: Domain Consisting of Tasks



D

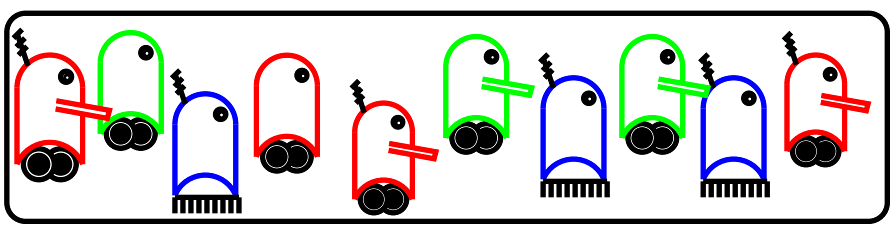


Evaluation: Set of Possible Teammates

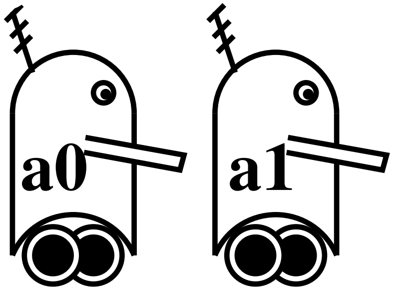


D

A

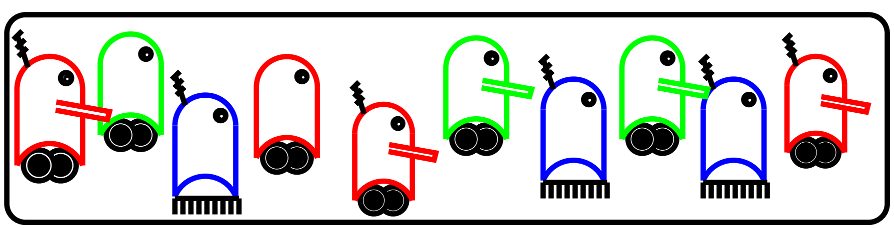


Evaluation: Draw a Random Task

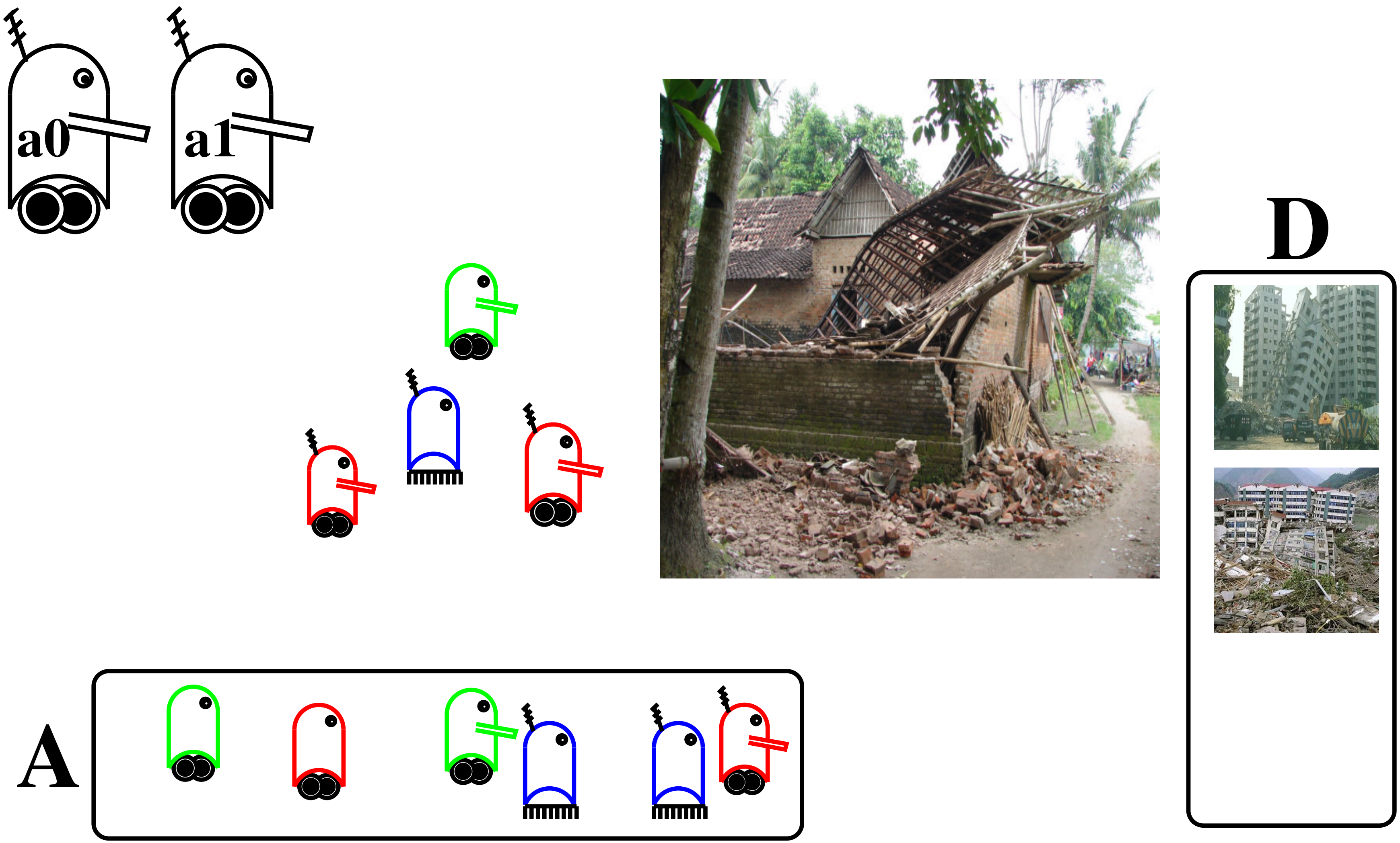


D

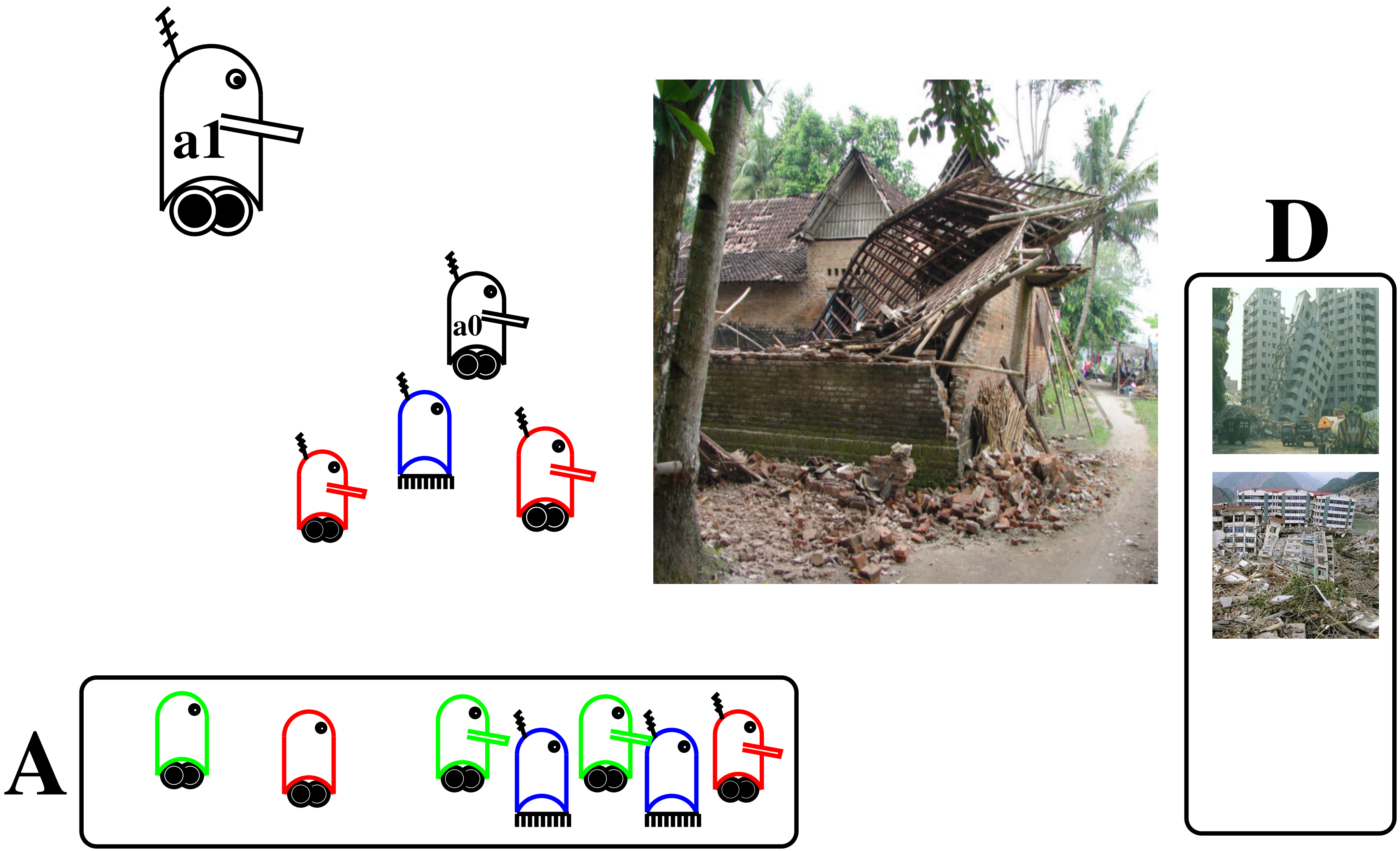
A



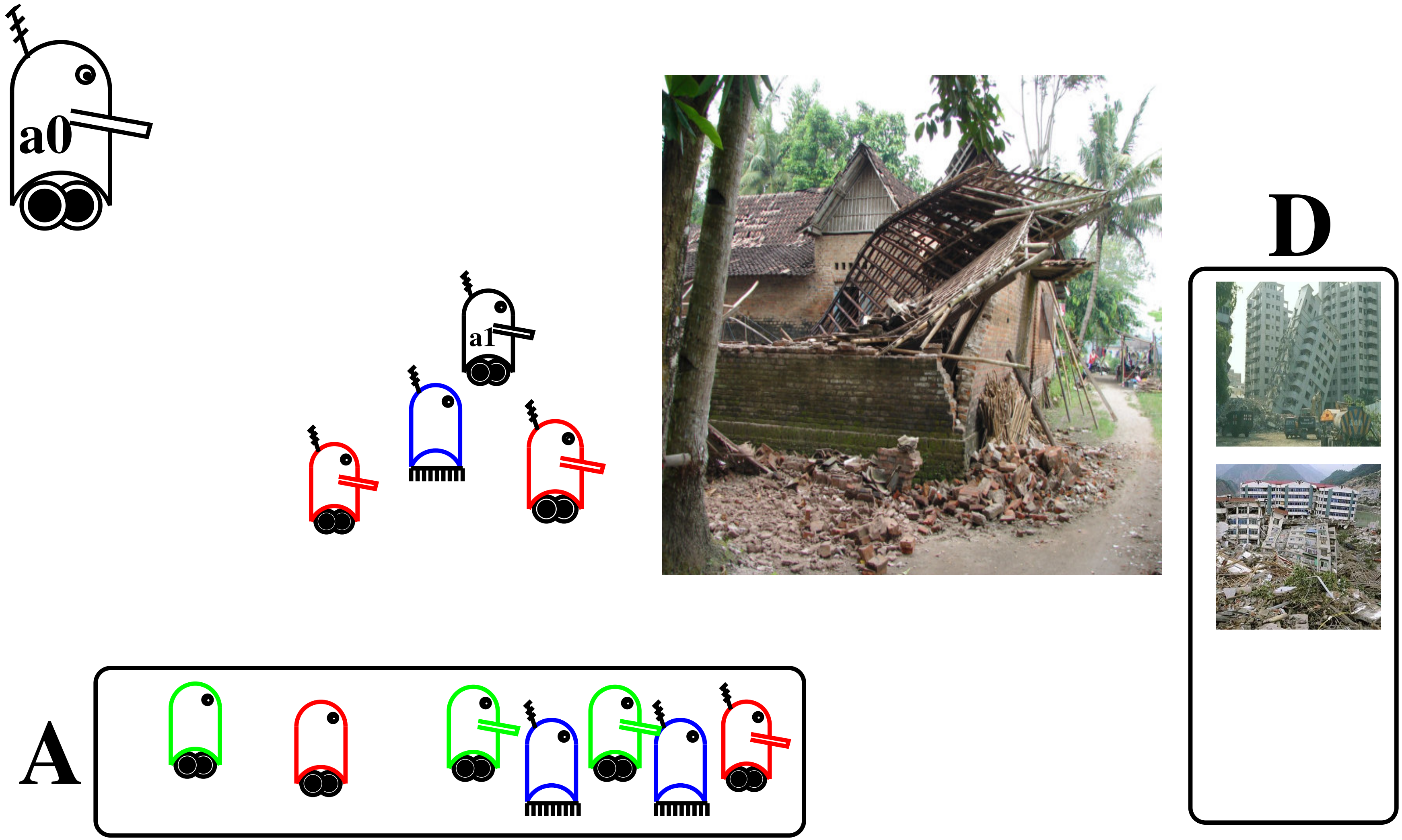
Evaluation: Random Team, Check Comp



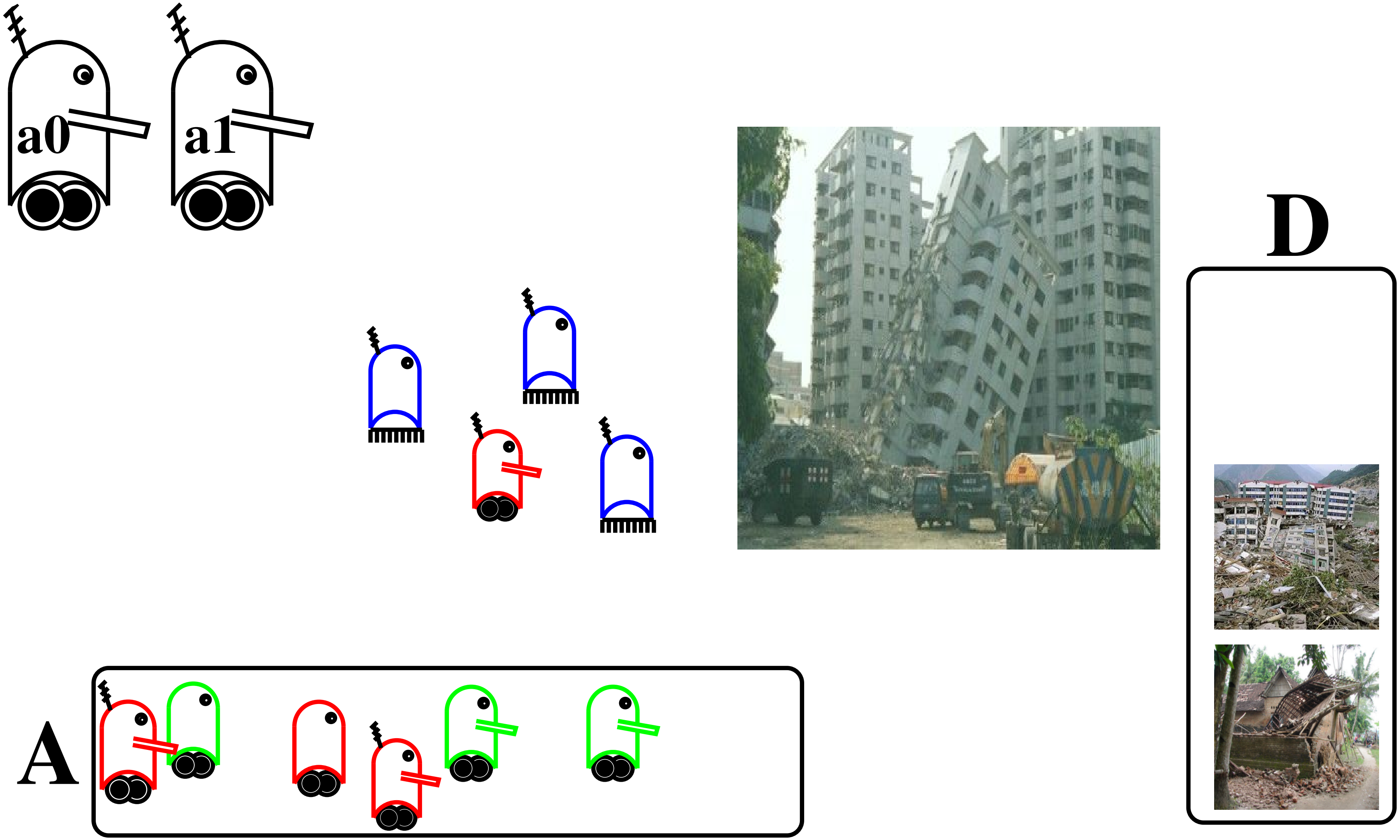
Evalution: Replace Random with a0



Evaluation: Then a1 — Evaluate Diff



Evaluation: Repeat



Evaluate(a_0, a_1, A, D)

- Initialize performance (reward) counters r_0 and r_1 for agents a_0 and a_1 respectively to $r_0 = r_1 = 0$.

- Repeat:
  - Sample a task d from D.
  - Randomly draw a subset of agents B, $|B| \geq 2$, from A such that $E[s(B,d)] \geq s_{min}$.
  - Randomly select one agent $b \in B$ to remove from the team to create the team B^-.
  - increment r_0 by $s(\{a_0\} \cup B^-, d)$
  - increment r_1 by $s(\{a_1\} \cup B^-, d)$

- If $r_0 > r_1$ then we conclude that a_0 is a better ad-hoc team player than a_1 in domain D over the set of possible teammates A.

Technical Requirements

- Assess capabilities of other agents (teammate modeling)

Technical Requirements

- Assess capabilities of other agents (teammate modeling)
- Assess the other agents' knowledge states

Technical Requirements

- Assess capabilities of other agents (teammate modeling)
- Assess the other agents' knowledge states
- Estimate effects of actions on teammates

Technical Requirements

- Assess capabilities of other agents (teammate modeling)
- Assess the other agents' knowledge states
- Estimate effects of actions on teammates
- Be prepared to interact with many types of teammates:
  - May or may not be able to communicate
  - May be more or less mobile
  - May be better or worse at sensing

Technical Requirements

- Assess capabilities of other agents (teammate modeling)
- Assess the other agents' knowledge states
- Estimate effects of actions on teammates
- Be prepared to interact with many types of teammates:
  - May or may not be able to communicate
  - May be more or less mobile
  - May be better or worse at sensing

A good team player's best actions will differ depending on its teammates' characteristics.

Preliminary Theoretical Progress

- Aspects can be approached theoretically
- Ultimately an empirical challenge

Preliminary Theoretical Progress

- Aspects can be approached theoretically
- Ultimately an empirical challenge

Be prepared to interact with many types of teammates

Preliminary Theoretical Progress

- Aspects can be approached theoretically
- Ultimately an empirical challenge

Be prepared to interact with many types of teammates

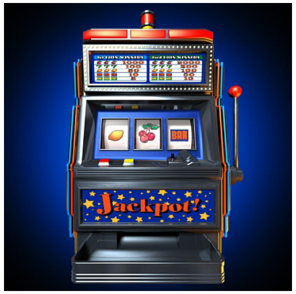
- Minimal representative scenarios
  - One teammate, no communication
  - Fixed and known behavior

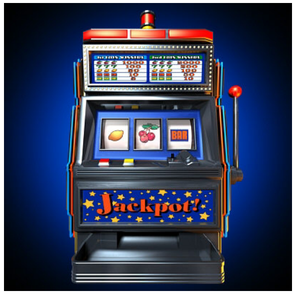
Scenarios

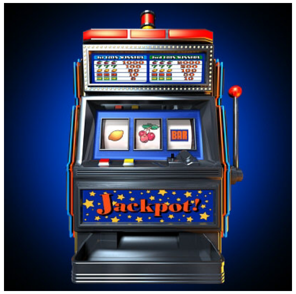
- Cooperative iterated normal form game (w/ Kaminka & Rosenschein—AMEC'09)

$M1$	b_0	b_1	b_2
a_0	25	1	0
a_1	10	30	10
a_2	0	33	40

- Cooperative k-armed bandit (w/ Kraus—AAMAS'10)





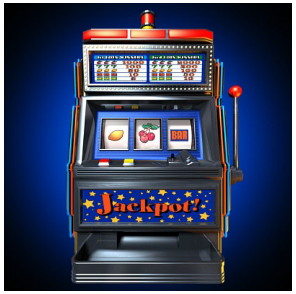


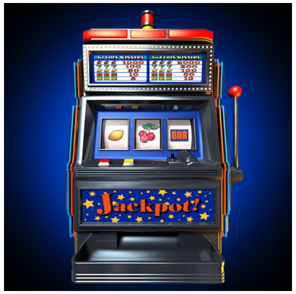
Scenarios

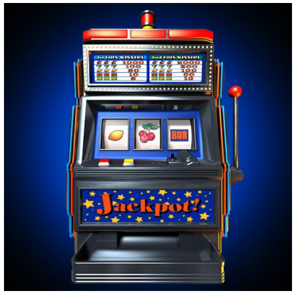
- Cooperative normal form game

$M1$	b_0	b_1	b_2
a_0	25	1	0
a_1	10	30	10
a_2	0	33	40

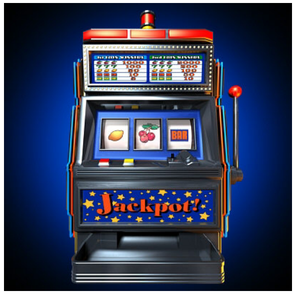
- Cooperative k-armed bandit







3-armed bandit

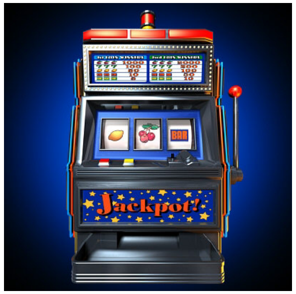


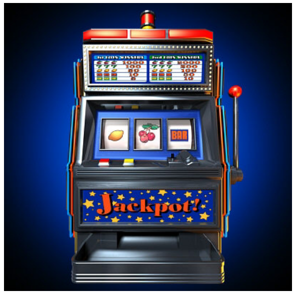
$\Longrightarrow$

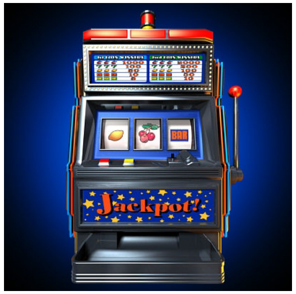
- Random value from a distribution
- Expected value μ

3-armed bandit

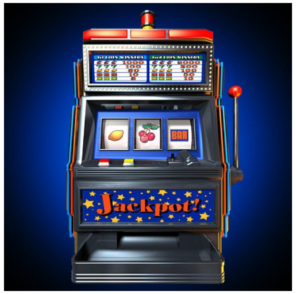
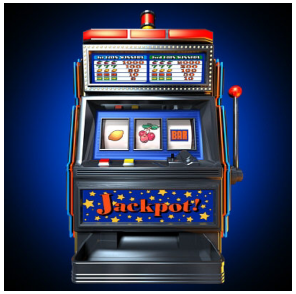
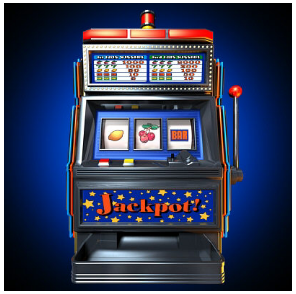
Arm$_*$ Arm$_1$ Arm$_2$





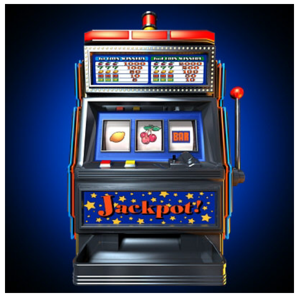
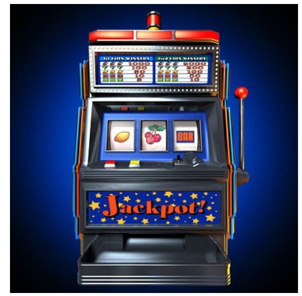


3-armed bandit

Arm_*		Arm_1		Arm_2
				
μ_*	$>$	μ_1	$>$	μ_2

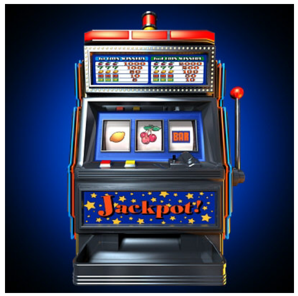
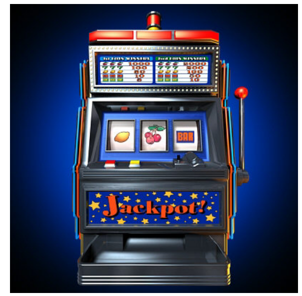
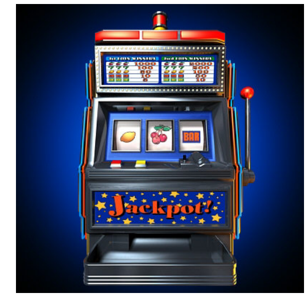
- Agent A: teacher
  - Knows payoff distributions
  - Objective: maximize expected sum of payoffs

3-armed bandit

Arm_*		Arm_1		Arm_2
				
μ_*	$>$	μ_1	$>$	μ_2

- Agent A: teacher
  - Knows payoff distributions
  - Objective: maximize expected sum of payoffs
  - If alone, always Arm_*

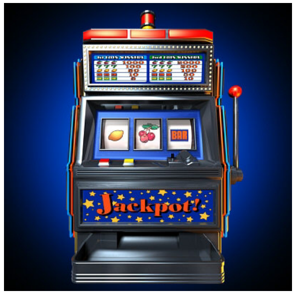
3-armed bandit

Arm_*		Arm_1		Arm_2
				
μ_*	$>$	μ_1	$>$	μ_2

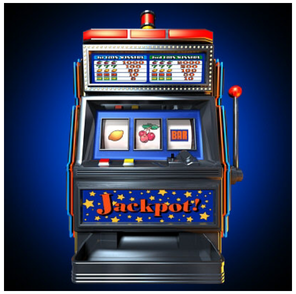
- Agent A: teacher
  - Knows payoff distributions
  - Objective: maximize expected sum of payoffs
  - If alone, always Arm_*
- Agent B: learner
  - Can only pull Arm_1 or Arm_2

3-armed bandit

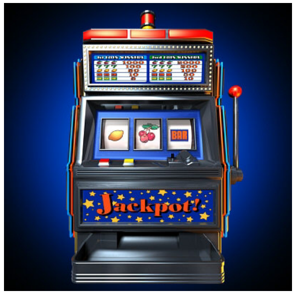
Arm$_*$



Arm$_1$



Arm$_2$



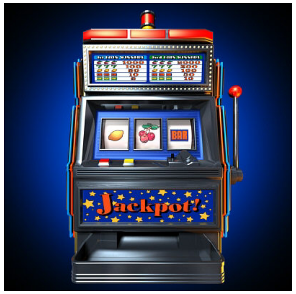
$\mu_* \quad > \quad \mu_1 \quad > \quad \mu_2$

- Agent A: teacher
  - Knows payoff distributions
  - Objective: maximize expected sum of payoffs
  - If alone, always Arm$_*$

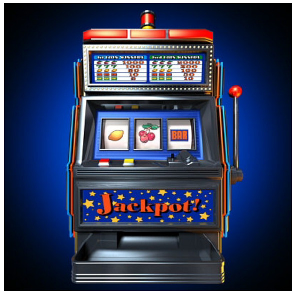
- Agent B: learner
  - Can only pull Arm$_1$ or Arm$_2$
  - Selects arm with highest observed sample average

Assumptions

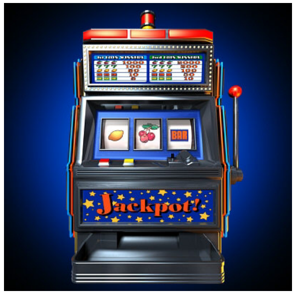
Arm_*



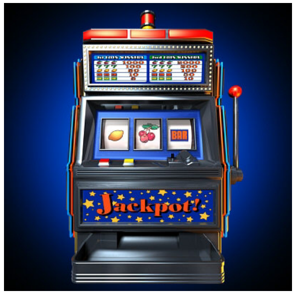
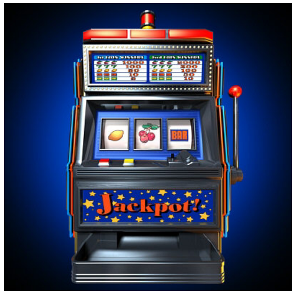
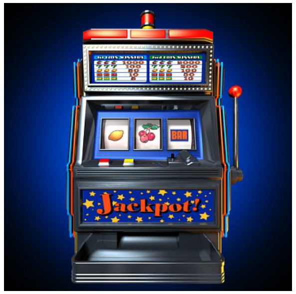
Arm_1



Arm_2

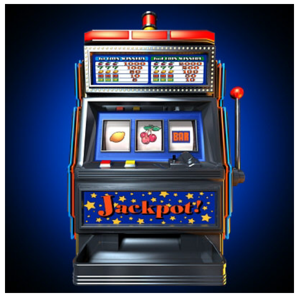
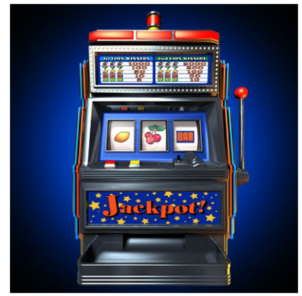


Assumptions

Arm$_*$		Arm$_1$		Arm$_2$
				
μ_*	$>$	μ_1	$>$	μ_2

- Alternate actions (teacher first)
- Results of all actions fully observable (to both)

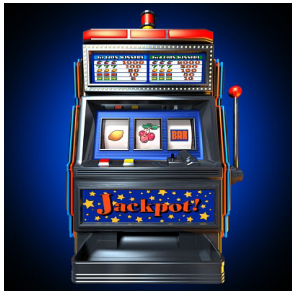
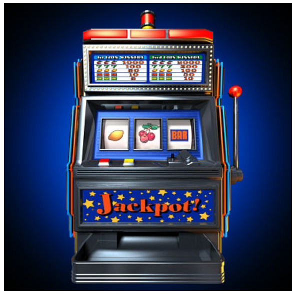
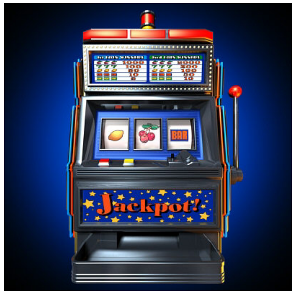
Assumptions

Arm$_*$	Arm$_1$	Arm$_2$
		

$$\mu_* \quad > \quad \mu_1 \quad > \quad \mu_2$$

- Alternate actions (teacher first)
- Results of all actions fully observable (to both)
- Number of rounds remaining finite, known to teacher

Assumptions

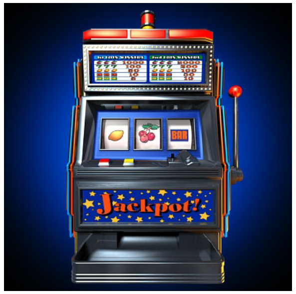
Arm$_*$		Arm$_1$		Arm$_2$
				
μ_*	$>$	μ_1	$>$	μ_2

- Alternate actions (teacher first)
- Results of all actions fully observable (to both)
- Number of rounds remaining finite, known to teacher

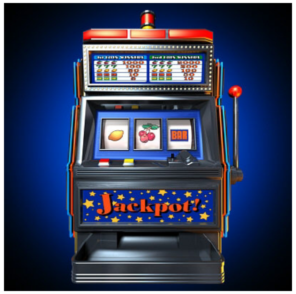
Objective: maximize expected sum of payoffs

Summary of Findings

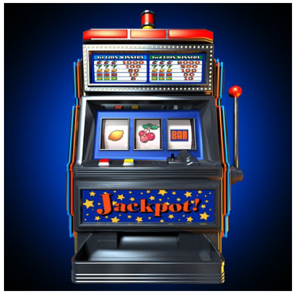
Arm_*



Arm_1

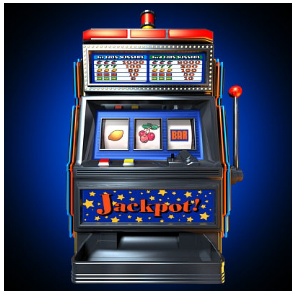
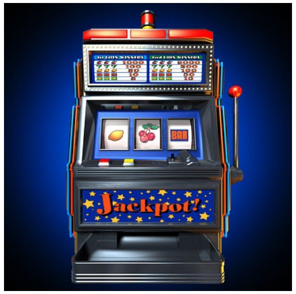
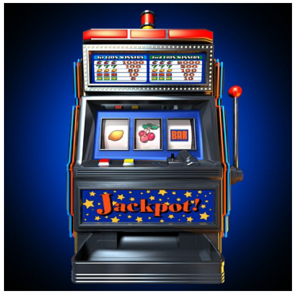


Arm_2



Summary of Findings

Arm_* Arm_1 Arm_2

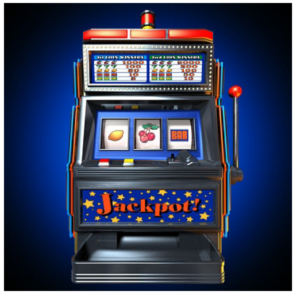
  

$$\mu_* > \mu_1 > \mu_2$$

- Arm_1 is sometimes optimal
- Arm_2 is never optimal

Summary of Findings

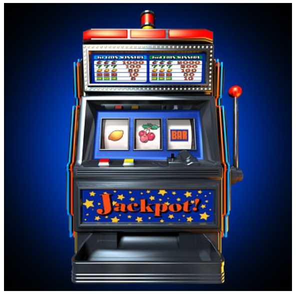
Arm$_*$



μ_*

$>$

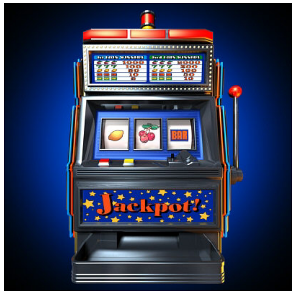
Arm$_1$



μ_1

$>$

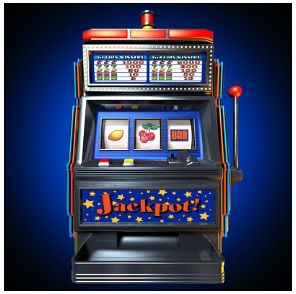
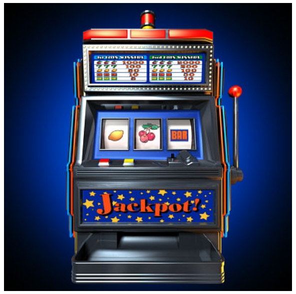
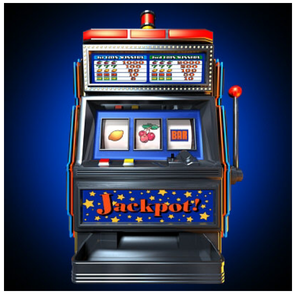
Arm$_2$



μ_2

- Arm$_1$ is sometimes optimal
- Arm$_2$ is never optimal
- Optimal solution when arms have discrete distribution
- Interesting patterns in optimal action
- Extensions to more arms

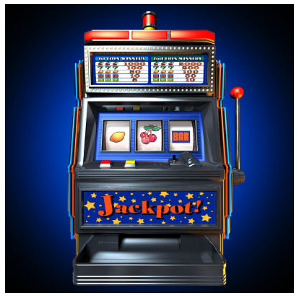
Summary of Findings

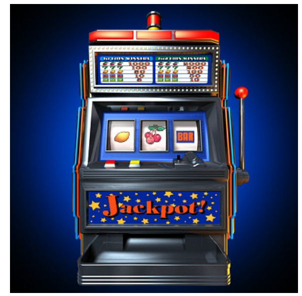
Arm$_*$		Arm$_1$		Arm$_2$
				
μ_*	$>$	μ_1	$>$	μ_2

- Arm$_1$ is sometimes optimal
- Arm$_2$ is never optimal
- Optimal solution when arms have discrete distribution
- Interesting patterns in optimal action
- Extensions to more arms
- Exploitation vs.

Summary of Findings

Arm_* | Arm_1 | Arm_2





$$\mu_* > \mu_1 > \mu_2$$

- Arm_1 is sometimes optimal
- Arm_2 is never optimal
- Optimal solution when arms have discrete distribution
- Interesting patterns in optimal action
- Extensions to more arms
- Exploitation vs. vs. teaching

Challenge Statement

Create an autonomous agent that is able to efficiently and robustly collaborate with previously unknown teammates on tasks to which they are all individually capable of contributing as team members.

Suggested Research Plan

1. Identify the full range of possible teamwork situations that a complete ad hoc team player needs to be capable of addressing (D and A).

Suggested Research Plan

1. Identify the full range of possible teamwork situations that a complete ad hoc team player needs to be capable of addressing (D and A).

2. For each such situation, find theoretically optimal and/or empirically effective algorithms for behavior.

Suggested Research Plan

1. Identify the full range of possible teamwork situations that a complete ad hoc team player needs to be capable of addressing (D and A).

2. For each such situation, find theoretically optimal and/or empirically effective algorithms for behavior.

3. Develop methods for identifying which type of teamwork situation the agent is currently in, in an online fashion.

Suggested Research Plan

1. Identify the full range of possible teamwork situations that a complete ad hoc team player needs to be capable of addressing (D and A).

2. For each such situation, find theoretically optimal and/or empirically effective algorithms for behavior.

3. Develop methods for identifying which type of teamwork situation the agent is currently in, in an online fashion.

- 2 and 3: the core technical challenges

Suggested Research Plan

1. Identify the full range of possible teamwork situations that a complete ad hoc team player needs to be capable of addressing (D and A).

2. For each such situation, find theoretically optimal and/or empirically effective algorithms for behavior.

3. Develop methods for identifying which type of teamwork situation the agent is currently in, in an online fashion.

- 2 and 3: the core technical challenges

- 1 : a knob to incrementally increase difficulty

Related Work

Multiagent learning (Claus & Boutilier, '98),(Littman, '01), (Conitzer & Sandholm, '03),(Powers & Shoham, '05),(Chakraborty & Stone, '08)

Opponent Modeling

- Intended plan recognition (Sidner, '85),(Lochbaum,'91),(Carberry, '01)
- SharedPlans (Grosz & Kraus, '96)
- Recursive Modeling (Vidal & Durfee, '95)

Human-Robot-Agent Teams

- Overlapping but different challenges, including HRI (Klein, '04)
- Out of scope

Much More pertaining to specific teammate characteristics

Acknowledgements

- Fulbright and Guggenheim Foundations
- Israel Science Foundation

Ad Hoc Teams

- Ad hoc team player is an individual
  - Unknown teammates (programmed by others)
- May or may not be able to communicate
- Teammates likely sub-optimal: no control

Challenge: Create a good team player