To Scenarize the Assessment of an Educational Activity

Guillaume Durand
Université de Savoie
Chambéry, Savoie, France
Guillaume.durand@etu.univ-savoie.fr

Christian Martel
Université de Savoie
Chambéry, Savoie, France
Christian.Martel@univ-savoie.fr

Abstract: In this paper, we present our work which deals with learner assessment in the context of a scenarized learning activity. We propose in this context to scenarize assessment in a specific scenario. Assessment is then considered as an activity being able to be scenarized. But scenario languages and infrastructures are not designed for assessment scenarios. Based on LDL (Learning Design Language), a scenario language, and its infrastructure, we set out the way in which we have dealt with this problem. Indeed the scenarization of an assessment requires improvements in order to allow the expression of the results and their communication between scenarized activities. To express the results, a specific results model is proposed for the scenario infrastructure. This model makes it possible to exchange, store and of course to evaluate the results obtained from a scenarized learning activity.

Introduction

Scenario infrastructures (LDI (Martel *et al.* 2006), RELOAD[1], LAMS[2]) today offer great freedom to design learning activities according to the wishes of the teachers (Burgos 2005). Traditionally, assessment plays a fundamental role in learning activities. One can thus expect, that within the framework of activities modelled by scenario languages, teachers also need to evaluate and to observe their pupils. We have to admit that in scenario languages and infrastructures, assessment does not take its fundamental place. For the designers of the scenario infrastructure, assessment is generally associated with multiple-choice questionnaire tests. The diversity of the Technology Enhancing Learning (TEL) used in classes shows that assessment goes far beyond this framework (Mitchell 1992). After a quick summary of the existing types of assessment in TEL, we show initially how to scenarize assessment. For this a scenario language developed within the SysCom team of the University of Savoy (Martel *et al.* 2006) to express learning activities is used. Then a proposal to express results obtained from the execution of the learning scenarios and their use in the activity of assessment defined by an assessment scenario are described. Last but not least, we present a conclusion on ongoing experiments and future works.

The diversity of assessment in TEL

In TEL, there are several ways to assess learners. We present here some well-known types of assessment.

Skills Assessment

The assessment of skills aims at taking a “photograph” of a learner’s skills level. Most of the time this type of assessment uses Multiple Choice Questions (MCQ) organised in a course. This is a summative assessment (Angelo 1993) which generally provides a final score. The course can be highly sophisticated, in particular in Computer Adaptive Tests (CAT) (Green *et al* 1984). In CAT, the course depends on the response given by the learner, and it adapts to the learner. The TOEFL[3] test is certainly the most well-known. But there are alternatives to this assessment of skills, such as diagnostic

[1] RELOAD : http://www.reload.ac.uk/
[2] LAMS : http://www.lamsinternational.com/
[3] ETS : www.ets.org

assessment (Kellough 1999), which does not provide a score but detail on the learner's strengths and weaknesses (Jean *et al*, 1999). This type of assessment is more adapted to a formative scope favourable with a remediation than the other skills assessment methods.

Self-assessment

In many cases, it can be useful to offer the learner self-assessment in order to support the automatic regulation of his/her training courses. It is a formative assessment (Shunk 1990). The device proposed by Eyssautier (Eyssautier 2004) is based on a list of questions. For each question, the learner can benefit or not from clues to the answer. After having answered the questions and seen the correction, the learner evaluates his/her understanding of the concepts used in the questionnaire by grading it on a scale of 1 to 5. But this self-assessment can also be collective. In peer assessment, the assessment is carried out by a group of learners on a production created by another group. The difficulty consists among other things in the members of the group who are evaluated agreeing on the right answers (Juwah 2003).

Assessment of Participation

Although common to all group activities, the assessment of the participation is mostly within activities based on a communication tool. The assessment relates to the quality of the contributions. By using techniques of "text mining" and being interested in the relationship between key words (those expected by the teacher) and number exchanges, it is possible to define the quality of contributions. This, for example, is the aim of the DIAS software (Bratitsis et al 2005).

Scenarization of Learning Activities

The scenarization of learning activities is relatively recent. It results from the need to allow teachers to organize the learning activities they desire.

The Basis of Scenarization

A scenario is the specification of a future learning activity which becomes a "scenarized activity" expressed in a scenario language such as IMS-LD[4] or LDL (Martel *et al.* 2006). To create a corresponding activity in a targeted environment, a scenario has to be "operationalized". This consists firstly in choosing the participants, then attributing roles foreseen by the scenario to the proper participants, and finally selecting the services and contents required by the scenario. The execution of the operationalized scenario will provide the learners, teachers, tutors, etc. with the means (resources, services, tools, etc.) to take part in the activity. Other activities may take place simultaneously within the scenarized activity; we call these "spontaneous activities". Given their unpredictable nature, they cannot be specified in any scenario, neither can they be controlled or followed.

Assessment and Scenarization

Practices of assessments are relatively rich in TEL. However there are few types of assessments within the framework of scenarized activities. For example in the case of the IMS-LD language, assessment is not part of the description of the learning activity. There is no specific object which allows one to describe the assessment desired within the scenario. Assessment is left to the tools used during the learning activity. IMS-LD makes it possible to describe the results obtained from these tools. In other words, to describe the grade obtained by a student in an IMS-QTI[5] multiple-choice questionnaire. This score is expressed in the IMS-LD language by a property. The value of this property (property-value) can be used under conditions to redirect learning towards one activity or another, as described by the scenario. However, as stated previously, TEL student assessment does not provide only a summative test of learner(s) (Mitchell 1992).

---

[4] IMS-LD : www.imsglobal.org/learningdesign
[5] IMS-QTI : www.imsglobal.org/question

To Scenarize Assessment

Proposition

The proposition made to allow the taking into account of the practices of assessment in TEL, by the means of a scenario, is to scenarize assessment. Indeed, assessment is a natural and social activity (Sutherland 1996). It must be possible, in the way that it exists in TEL, to scenarize assessment activity just like a learning activity. It must then be possible to express assessment in a scenario formalism. Moreover, assessment is a reusable activity. It must be possible to reuse some of the assessment of one learning activity with another (Williamson *et al.* 2003). Assessment must thus have a relative independence with respect to the learning activity. Therefore, in a functional and practical vision, we have directed our research towards a description of assessment in a specific scenario. We propose an example of an assessment scenario.

Example

The chosen assessment scenario describes an individual diagnostic assessment. This assessment activity proceeds simultaneously with an answering activity (the learning activity) in which a student answers a list of questions, each of which has the same learning objective. Each question has only one solution, and a list of errors with the possible origin of each error. It is supposed that these metadata are accessible. The assessment scenario describes the correction of each of the learner's answers (thanks to the expected answer), but also, in the event of error, the determination of the possible origin of the error as well as a proposal for a remediation activity (*cf* Figure1). This remediation activity is also scenarized; it offers the learner a course explaining the concept which has been misunderstood. During the remediation activity, the answering activity is on standby. Once remediation activity is finished (3), the student continues the answering activity (4). If learner does not make an error, he/she proceeds to the following question envisaged by the answer scenario (2).

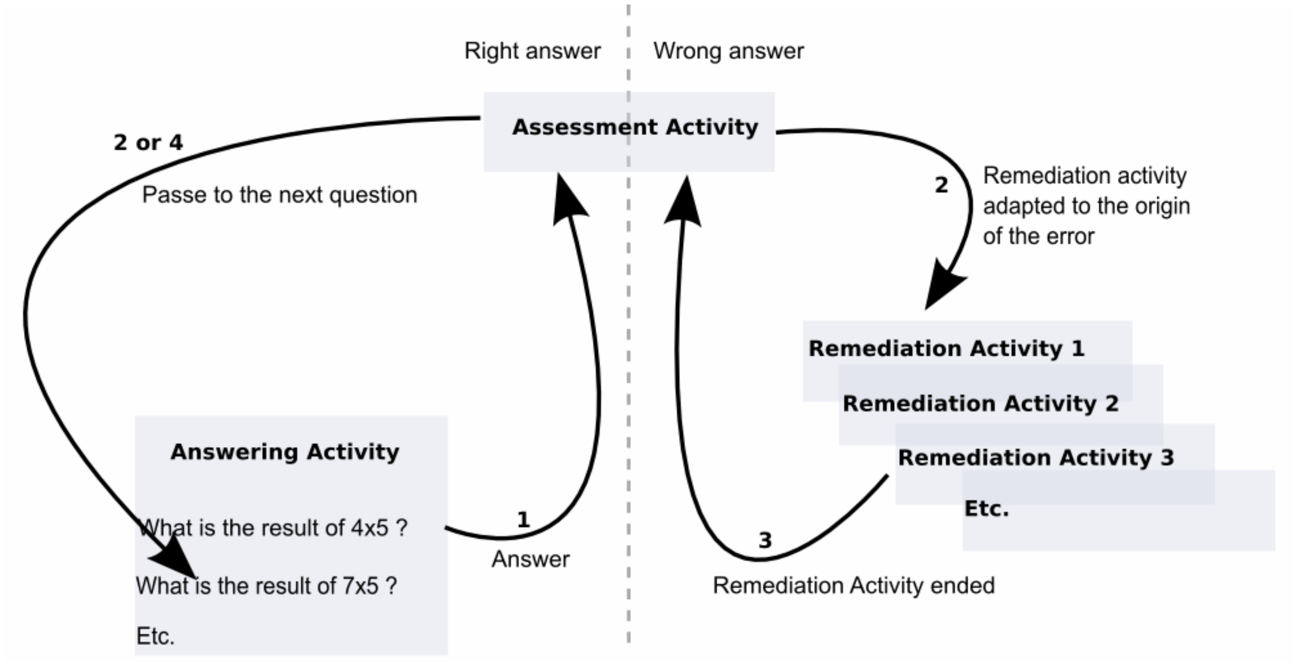

Figure 1: Example of an assessment activity

Implementation

These assessment scenarios, as well as the other learning scenarios which describe the example activity, have been implemented in order to check the feasibility of our proposal. As the implementation context we have used the LDL scenario language and its LDI infrastructure (Martel *et al.* 2006).

LDL/LDI

The LDL language makes it possible to describe a learning activity in the XML format. LDL is a scenario language which expresses learning models of activities. LDL organizes a learning activity into interactions which can proceed in a sequential or parallel way inside structures. An interaction is carried out by one or more participants having a role, in a place called an arena (chatroom, file, etc...). Rules control the course of the activity. These rules have conditions which relate to positions. Positions are objects of the LDL model whose aim is to describe and qualify the relationship between each actor in an arena. They correspond to a qualitative point of view expressed by an actor on him/herself, about another actor or about the arena. When a learner gives an answer, he/she takes a position on the arena question.
LDI is a software infrastructure which allows one to operationalize a learning scenario expressed in the LDL language. This infrastructure allows the loading of an XML scenario, its instantiation and its unfolding.

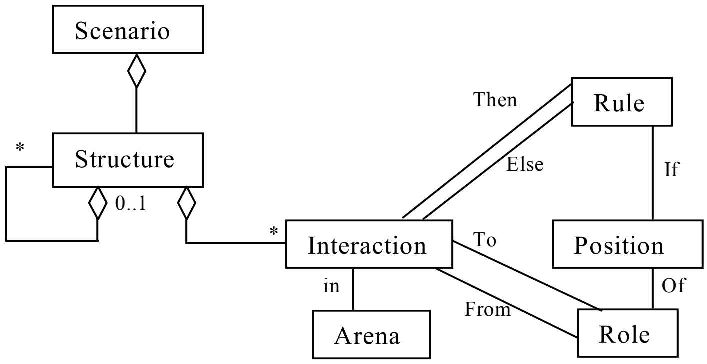

Figure 2 : LDL simplified representation

What is Required to Scenarize Assessment

During the answering activity of our example, the assessment activity must obtain results which are the answers given by the learner. But this induces problems:

- First of all, how to obtain these results i.e. how to express in LDL the observation of the results?
- Then, how to communicate the results between the two activities i.e. which mechanism can permit the scenario infrastructure to allow communication between activities?
- And finally how to express a result?

This last question will be developed in the Result Model part.

To Obtain Results

In the last example, an answer given by the student to a question from the answering activity is observable. This observable has been defined in the answer scenario. The independence between the activity model which is the scenario and the infrastructure in which it is played out has been preserved (as we can see in the XML example at the end of the paper). To the LDL language, the definition of a type of a particular position has been added which describes the observation of the answer given by the learner to the question. The observation position is defined in the scenario. The infrastructure creates a position object at the beginning of the scenarized activity. The position object is notified as soon as the learner has given his /her answer and the data of the position becomes the value of the answer. We use the observer-observable design pattern. The position is the observer, and the question the observable.

Communication of the Results

As can be seen in figure 1, the answering activity must transmit the answers given by the learner to the assessment activity. To communicate between activities, observation positions are shared by the scenarios. There are particular positions which are known as "shared". They are defined in several scenarios, and when a value is changed in one activity, the modification has repercussions on all the others in a cascade effect. If in the answering activity, the learner gives an answer to the first question, defined by the position "first_answer" in the answer scenario, the position "first_answer" in the assessment activity is also updated.

A Result Model

The last problem to overcome was the expression of the results exchanged between each activity.

Existing Result Formalisms

There are a certain number of result formalisms, even if the result part not always does represent the principal object of the specifications. Indeed, in the IMS-LIP[6] model, the objective is to specify an XML formalism to express the profile of a learner. IMS-LIP deals with results only in the part which is dedicated to the results of assessment. IMS-QTI Result Reporting[7] specifies a formalism of results obtained by a learner during an IMS-QTI test session. As well as the scenario devices allow only one restricted vision of assessment based on MCQ, result formalism expresses almost only results resulting from MCQ. This had to be improved, taking into account the diversity of the results able to be obtained from a scenarized learning activity.

The Diversity of the Results

In order to illustrate this diversity (i.e. the range of results that a scenarized learning activity can produce), we focused on various activities that we scenarized and implemented. We will briefly present them.

A Remediation Activity
This remediation activity is similar to the example scenario. It is made up of three activities, each described by a scenario: an answer scenario, an assessment scenario and a remediation scenario. In the answering activity, the learner answers a succession of questions and carries out simulations of electrical circuits. The assessment activity evaluates the answers to the questions and in the event of an error, it returns to a remediation activity made up of a lecture and questions. Once the remediation activity is finished the answering activity carries on.

An Activity of Long Duration
A “long life” scenario , i.e. a scenario of long duration, can scenarize an activity of a one year distance learning course such as those proposed by the CNED (French national center of distance learning). During this type of course the learner has access to resources. He/she must complete his/her work and return it as mail, text files, images, etc.

A Group Search
In our treasure hunt, groups of learners search for the name of a famous person. Clues are hidden in various documents. The members of the same group are physically distant; they organize their group research and communicate within a chatroom.

Typology of Results
Taken as a whole, these activities provide our typology of possible results:

- The remediation activity offers results of simulation, the results coming from a QTI-like questionnaire.
- The “longlife” activity provides results of the deliverable type which can be files, and/or files deposited in a shared space.
- The group search, introduces the results which are a product of the communication and the coordination of the participants. There are results such as their interventions in a forum. Results which make it possible to evaluate their participation in the activity.

All these scenarios have results relating to the learners’ navigation the activity. For example the number of accesses to an

[6] IMS-LIP : www.imsglobal.org/profiles
[7] IMS-QTI Result Reporting : www.imsglobal.org/question/qtiv1p2 /imsqti_res_infov1p2.html

arena, or the duration of the reading of an instruction can also be the subject of an assessment (for example the assessment of a statement reading).

Proposition of a New Result Model

We do not claim that our model is complete as regards the types of results which can be obtained from a scenarized learning activity. The ambition of this model is to allow the expression of the most common results, such as those seen previously.

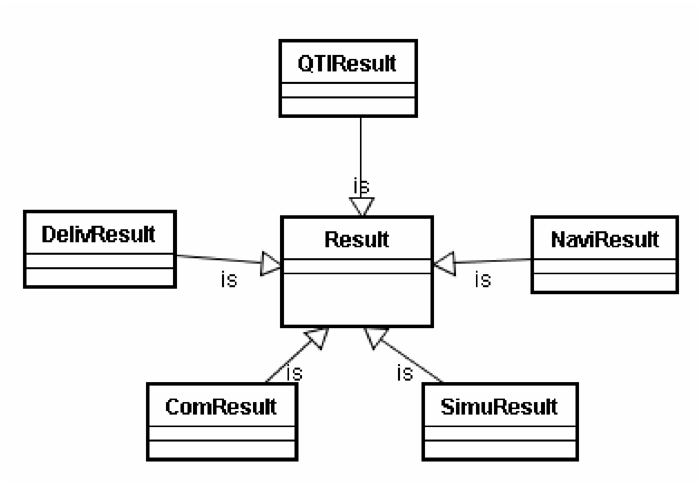

Figure 3 : Types of results

In figure 3 we find the diversity of the results produced by the various activities which we scenarized. We have the deliverables (DelivResult), the results of communication (ComResult), the results produced by a simulation (SimResult), the navigation of learning in the activity (NaviResult), and the results which come from MCQ (QTIResult).

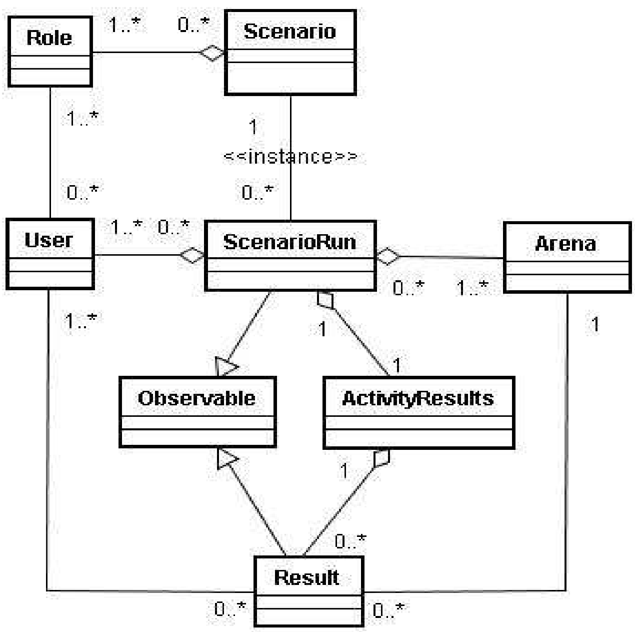

Figure 4 : Global view of the model

Each of these results is observable by the infrastructure (observer-observable design pattern). This characteristic is usable within the framework of an assessment scenario within a position. A result is produced in an arena which is specified at the instantiation of the scenario (Scenario) and used during the activity (ScenarioRun). During an activity a great number of results can be generated (ActivityResult). A result is produced by participants who have a role. The result model suggested re-uses takes again the concept of arena resulting from LDL. This concept is re-used taken again with different names in the other scenario languages. IMS-LD speaks about "learning Object" and service. The model is therefore not specific to LDL.

Use of the Result Model

Expression of a Result Coming From a Scenarized Activity

In order to express results the model has been carried out in the XML format. The results are stored by the scenario infrastructure in an XML database. As we have seen, these results come from the arena handled during the learning activity.

```
<ScenarioRun>
   <Scenario>……</Scenario>
   <Arenas>…….</Arenas>
   <Users>…...</Users>
   <ActivityResults id="1">
           <QTIResult id="r1">
                       <date>22/10/05 15H06</date>
                       <description>Answer first question</description>
                       <arenaref>Question1</arenaref>
                       <userref>user1</userref>
                       <result>12</result>
                       <solution>20</solution>
                       <score>0</score>
                       <scoremax>10</scoremax>
                       <scoremin>0</scoremin>
           </QTIResult>
   </ActivityResults>
</ScenarioRun>
```

The highest element in the XML hierarchy of the model is the ScenarioRun object which integrates all the others: Scenario, Arena, Users, and ActivityResults. In the example the learner "user1" gave the answer 12 to the question “Question1” whereas the right answer was 20. This result is usable in an assessment activity.

Use of a Result in LDL

Indeed, in LDL, it is possible to define positions which point to observables. However, in our result model, any result is observable. The expression of a result and its uses is then possible thanks to the positions. Below is the expression of a position which points to the answer to a QTI question.

```
<Position id="first_answer">
        <title> Answer first question</title>
        <valeur>
           <Observer id="QTI">result</Observer>
        </valeur>
        <sur>Question1</sur>
        <de>learner </de>
</Position>
```

Once the result is described in the scenario, within a position, it is possible to define treatments on this position in the scenario. In our case the treatment is an assessment carried out during an interaction by a human or software participant. The correspondence between the result described by the position, and the result obtained during the activity is managed by the scenario infrastructure.

Conclusion and Future Works

In this paper, we have made the choice to scenarize assessment by re-using scenario devices. We have shown the feasibility of this proposal by implementing it on the LDL language and its LDI infrastructure. At the same time a result model and formalism has been developed for our needs, but which could in the future be used more generally by others. From now on, our efforts will be concentrated on this result model, which will be improved and tested within the framework of a European project. The objective of the experimentation is to check the relevance of the model in the carrying out of assessment. Another aim is to develop an author tool for LDL which allows one to scenarize assessment. Indeed, it is very difficult to write assessment scenarios taking into account the dependences between scenarios due to the positions.

Acknowledgements

The authors wish to thank their collaborators in the Odile Jacob Multimedia and Pentila companies and the Syscom laboratory.

References

Angelo, T. A., Cross, K. P., « Classroom Assessment Techniques: A Handbook for College Teachers (2nd ed.) », San Francisco: Jossey-Bass, 1993.

Bratitsis T., Dimitracopoulou A., « Data recording and usage interaction analysis in asynchronous discussions: The D.I.A.S. System », AIED Workshops (AIED'05), July 2005.

Burgos D., Arnaud M., Neuhausser P., Koper R., « IMS Learning Design : la flexibilité pédagogique au service des besoins de l'e-formation », Revue de l'EPI. December 2005

Eyssautier C., Jean-Daubias S, « Device Helping Learners to Self-Assess Themselves », Proceedings of CALIE 2004, Grenoble, France, 2004.

Green B., Bock R., Humphreys L., Linn R., Reckase M., « Technical guidelines for assessing computerized adaptive tests », Journal of Educational Measurement, vol. 21, session 4, p. 347-360, 1984.

Jean S., Delozanne E., Jacobini P. and Grugeon B., « PÉPITEST, a software to establish the cognitive profile of the students in elementary algebra », WG 3.3 Working Conference "Human Computer Interaction and Educational Tools ", Sozopol, Bulgaria, 1999

Juwah C., « Using peer assessment to develop skills and capabilities », United States Distance Learning Association, vol. January 2003, p. 39-50, January 2003.

Kellough, R.D. and Kellough, N.G., « Secondary School Teaching: A Guide To Methods And Resources; Planning For Competence », Copyright by Prentice Hill, Upper Saddle River, New Jersey, 1999.

Martel C., Vignollet L., Ferraris C., David J.P., Lejeune A., « Modeling collaborative learning activities on e-learning platforms », to be published in the proceedings of ICALT 2006, Kerkrade, The Netherlands, July 2006.

Mitchell, R., « How American education got into the testing trap », In Testing for Learning, New York: The Free Press, 1992.

Schunk, D.H., « Goal setting and self-efficacy during self-regulated learning. Educational Psychologist », 25(1), 1990.

Sutherland, G., «Assessment: Some historical perspectives », in H. Goldstein & T. Lewis (Eds.), Assessment: Problems, developments and statistical issues. Chichester, England: Wiley, 1996.

Williamson D., Bauer M., Mislevy R., Behrens J., « An ECD approach to designing for reusability in innovative assessment », American Educational Research Association (AERA'03), April 2003.