

Semantic virtual patients: using semantic web technology to improve virtual patients for medical education

Jean-Rémy Duboc
PhD Student, School of Computer
Science
University of Southampton
SO17 1BJ Southampton, UK
+44 (0)7777675458

jmr1c09@ecs.soton.ac.uk

Sunhea Choi
School of Medicine
University of Southampton
Southampton General Hospital,
Tremona Road
SO16 6YD Southampton, UK
+44 (0)2380 796586

s.choi@southampton.ac.uk

Mark J. Weal
School of Computer Science
University of Southampton
SO17 1BJ Southampton, UK
+44 (0)23 8059 6000

mjw@ecs.soton.ac.uk

ABSTRACT

This poster presents an interdisciplinary project on virtual patients. Virtual patients are systems designed to help medical students practice their clinical skills in a safe environment, using feedback provided by the system to reflect on their clinical decisions. A review of existing virtual patient systems has been conducted, and limitations in terms of feedback have been identified in existing systems. This paper proposes that semantic web technologies will help to alleviate some of these limitations.

A new virtual patient system has been designed, and semantic web technologies are used in order to benefit from existing semantic data already available on the web, thus facilitating the virtual case editing process.

Semantic data is also used to generate automated feedback according to each student's choice of interview questions and examinations.

Categories and Subject Descriptors

D.3.3 [Semantic Web]: use of semantic web technologies in education

General Terms

Algorithms, Design, Human Factors.

Keywords

Medical Education, eLearning, Feedback

1. INTRODUCTION

A variety of computer systems called virtual patients are available in medical education today. Virtual patients are designed to emulate realistic clinical cases on a computer, and help students to practice diagnosis and clinical reasoning. They are used as an integral part of the curriculum in many medical schools.

However, the technologies currently used to build virtual patients present limitations. Feedback has to be edited manually by

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WebSci '11, June 14-17, 2011, Koblenz, Germany.
Copyright held by the authors.

medical experts, and the feedback provided is often not adapted to each student's interactions with the virtual patient. This makes creating and editing a virtual patient time-consuming, and limits its pedagogical impact. Indeed, relevant feedback is crucial to help students assess and reflect on their performance, and improve their clinical reasoning skills.

An interdisciplinary project centered on semantic virtual patients is now taking place at the School of Medicine and the School of Computer Science in the University of Southampton. The aim of the project is to design a virtual patient adapted to the needs of medical students in their clinical training (students in year 3, 4 and final year). In particular, discussions with year 3 students revealed a need for practice in history taking. Students in their clinical training years spend most of their time in clinical attachment; therefore the virtual patient system will be designed for individual self-paced study, as opposed to supervised coursework and group work on linear virtual patients typically performed in year 1 and 2.

To design an appropriate data model for a semantic virtual patient, it is crucial to understand how students think through a virtual case. This understanding helps to identify the most common mistakes and the most useful type of feedback needed to help students reflect and change their patterns of thinking. This type of feedback can help students to reason more like experts, and improve their clinical skills.

This paper supports the notion that semantic web technologies offer the potential to provide automated and individualised feedback to medical students within a virtual patient, using a semantic model of patient-related data and students interactions. It is expected that the resulting feedback would allow students to reflect on their assessment of the virtual patient's condition, their clinical decisions and their diagnosis skills. It is also expected that the resulting model will provide a wider interoperability between various types of virtual patient systems (web-based systems, virtual worlds, etc.).

2. ASSESSING EXISTING VIRTUAL PATIENT SYSTEMS

A review of current virtual patient systems used in medical education has been conducted. Many virtual patient systems have been developed in recent years [1,2]. Common types of virtual patient interactions and feedback have been identified during the review, and a review of the technologies used for virtual patient data has been conducted.

Three main types of interactions can be identified from the review: linear virtual cases, branching virtual cases, and exploratory virtual cases.

Linear virtual cases carry the students through a clinical scenario in a pre-defined sequence. The sequence unfolds screen by screen, and each screen presents either an element of the storyline or an activity to complete for the student. Activities include multiple choice questions, ordering lab tests or choosing examinations to perform on the virtual patient. Feedback is given after each activity, and a global feedback is provided at the end of the case.

Branching virtual cases allow students to choose an action from a list of possible options at each stage of the clinical scenario. Each decision brings the student to the corresponding outcome. The feedback is given when the student reaches a final outcome (final "nodes" of the branching structure), which can be either successful or not.

Exploratory virtual cases are designed to allow students to take the initiative of their own decisions within the virtual case. Students can choose from a variety of actions (interview questions, examinations and lab tests), and use information from the outcomes of these actions to choose a final diagnosis and treatment. Exploratory systems differ from branching systems in that they do not provide a narrative path, but merely a list of actions to perform and a final feedback. All possible actions in the system are available at all times, when in branching cases each node offers a specific, limited range of actions. In existing exploratory systems, the feedback generated focuses mainly on indicating if the chosen diagnosis is correct or not. The software resulting from this research is intended to go further and to provide feedback on users' choices of actions.

3. PEDAGOGICAL IMPACT OF VIRTUAL PATIENTS

The traditional Bayesian model of hypothetico-deductive clinical reasoning, although accurate and useful in principle, seems to show limitations, in particular when it comes to identifying what makes experts' diagnosis process superior to students' reasoning. In particular, experts tend to generate a smaller number of diagnosis hypotheses, but with more accuracy [3]. The hypothetico-deductive model of clinical reasoning falls short when it comes to helping students generate better hypothesis and testing them. In fact, in some cases conclusions reached using the Bayes's theorem might conflict with clinical intuition emerging from clinicians' experience [4]

Moreover, evidence shows that explaining the clinical reasoning process as a general process, independent of expert knowledge, leads to limitations. The effect of clinical experience and medical knowledge, a crucial factor in the acquisition of medical expertise, does not fit into the hypothetico-deductive model.

Current theories of clinical reasoning don't focus only on the process of generating and testing hypothesis, nor do they rely solely on medical knowledge itself to describe and assess the clinical reasoning process. Rather, the structure of medical knowledge is examined and used to assess the clinical reasoning process. Indeed, the way information is stored and available for processing is considered to be a determinant factor in the clinician's ability to generate and test relevant and accurate diagnoses. Script theory, from cognitive psychology, provides a framework to explain this process [5] and to assess it [6]. Schemas theory can be used for the same purpose [7,8]. As

described by Marshal[9], schemas are mental structures containing both conceptual and procedural knowledge. This means that when concepts are stored using schemas, they are not just retained in abstract terms, but are usable to make decisions. Scripts are typically acquired and modified through experience, and provide specific knowledge from memory of specific events, but also contain a sufficient level of abstraction that allows adaption to a new problem, or a new presentation of an existing problem.

Feedback is crucial for the acquisition of knowledge. Negative knowledge acquired by experience helps students to understand actions to avoid in real clinical situations [10]. Marshall [9] also emphasises the impact of repetition on schema building. The repetition of several similar scenarios through experience fosters the assimilation and adjustment of accurate and usable schemas related to a variety of conditions and presentations.

To help students acquire appropriate scripts for accurate diagnoses, it is essential to provide feedback on their actions on a specific case. The feedback on the relevance of their choices during the cases helps students to create and alter clinical schemas based on their mistakes. Providing a system to help students establish conceptual and procedural knowledge on a number of different medical cases is the main pedagogical benefit of this study.

Semantic web data is structured as graphs representing simple statements in the form **Object, Predicate, Object** [11]. This basic structure, easily translatable in formal logic terms, allows a natural way of representing knowledge regarding the virtual patient and everything surrounding it. In addition, complex ontologies built in OWL [12], based on description logic, can enrich semantic web data. Relational databases and XML, in comparison, do not allow the representation of knowledge in such a form.

The semantic web is designed to allow the re-use of data across various sources on the web, through the use of URIs (Uniform Resource Identifiers). Through this mechanism, it becomes possible to exploit existing medical data and existing models to build a virtual patient, as shown in the following section. Integration with many other data sources can be imagined and implemented.

4. AN ONTOLOGY OF VIRTUAL PATIENTS

A review of existing medicine-related semantic data sources and web ontologies has been conducted. The objective of this review is to determine how semantic medical data already available can be used to represent virtual cases; in order to facilitate the virtual case editing process and to generate automated feedback using a model of students' interactions with the virtual patient. Many biomedical ontologies and knowledge bases are available for re-use, describing medical data from a clinical point of view, or from a life science perspective [13-17]. The OpenGalen biomedical ontology [18] has been selected as a basic building block for an ontology of virtual patients. Ontology design patterns [19] were also used to solve specific data modeling problems.

The Virtual Patient Ontology contains classes to identify the virtual patient. These classes include, for instance, **FemalePatient** and **MalePatient**. The data representing the fact that the virtual patient is a male patient (belongs to the **MalePatient** class) would read as follows in the N3 notation [20]:

```
virtual_Patient_x rdf:type MalePatient
```

Other classes are used to identify the patient's body parts. These classes are used associated with the property **isLocatedOn**, which links a virtual patient to each of its body part. The following statements in N3 indicate that an object belonging to the class **Head** "is located on" a patient called **virtual_patient_X**:

```
_:head_patient_x rdf:type Head.
_:head_patient_x isLocatedOn virtual_patient_X.
```

The state of the patient as a whole or the state of each individual body part is characterised using numerous classes designed to describe the patient's morphology, the dimensions of each body part, etc. Classes to represent the symptoms and conditions affecting the patient are also available.

Each of these patient's features is accessible to students using another class, called **Observation**. An observation represents a question, an examination, or a lab test, and can be presented to the patient as a text, picture, video or any media file. Observations are selected by students, and the choices of each student are recorded. Feedback is then generated using the virtual patient data and the data describing the student's choices of observations.

4.1 Feedback Generation using Semantic Data

In the semantic virtual patient, feedback is generated from two types of data: virtual patient data (entered mainly by the virtual case author, and completed by data extracted from external linked data sources), and interaction data entered by each student trying to solve the case. Using querying and reasoning mechanisms, it is possible to generate feedback regarding the questions and examinations that the students have performed, the diagnoses they propose, and the treatments they prescribe. This section describes the types of feedback that can be generated, and how the semantic virtual patient system generates feedback.

For instance, students can select the questions and examination they want to perform. It is possible to show students the important questions and examinations they neglected to choose during their investigation, and which chosen questions or examinations are irrelevant.

In the virtual patient ontology, a variety of symptoms are identified, and linked to the corresponding observations needed to observe them. Observations linked to symptoms affecting the patient are therefore easily identified. Other observations not linked to a specific symptom can still be important to the case, as they can confirm or rule out a given condition. These observations can be explicitly marked as important by the virtual patient author, and connected to the condition they confirm or rule out.

The following set of rules, written in natural language, identifies important observations by inference by finding observations linked the patient's symptoms:

- If the patient has a symptom, and if that symptom has an observation, then the observation is important (belongs to the class "**importantObservation**").
- If the patient has a symptom, and the symptom is located on an anatomical entity, and the anatomical entity has an observation, then the observation is important (belongs to the class "**importantObservation**").

After applying these rules, all the observations which have not been explicitly marked as important by the virtual case author are now identified.

When a student investigates the patient, she selects several observations (questions, examinations, and lab tests). As a list of all important observations has been generated earlier, this SPARQL query generates a list of all important actions that the student missed:

```
SELECT ?observation WERE{
    ?observation                                rdf:type
    virtual_patient_ontology:ImportantObservation;
    MINUS {
        <session_X> virtual_patient_ontology:contains
    ?observation
    }
}
```

This query lists all important observations, minus all observations chosen by the student. The result is a list of all important observations that the student has neglected to choose. This list can then be displayed to the student in a human-readable format.

5. CONCLUSION

To determine the relevance of automated feedback and the desired types of interaction in virtual patients, a survey was conducted among medical students and the clinicians who teach them in year 3, 4 and 5 of the University of Southampton School of Medicine. In this pilot study, students and clinicians indicated that the types of feedback described on this poster are indeed useful in virtual patients used for personal study.

To confirm this result, a fully functional prototype of the semantic virtual patient system is currently under development, using input from a small group of students for the interaction design. The prototype will then be tested by students in year 3, 4 and 5, and various types of feedback will randomly be presented to each student. Each piece of generated feedback will be evaluated by the students receiving it, in order to identify the most relevant and useful types of feedback in virtual patients. This will inform the design of future virtual patients to be used within the School of Medicine.

6. REFERENCES

- [1] S.R. Huwendiek, B.A.D. Leng, N. Zary, M.R. Fischer, J.G. Ruiz, and R. Ellaway, "Towards a typology of virtual patients," *Medical Teacher*, vol. 31, 2009, pp. 743-748.
- [2] D. a Cook and M.M. Triola, "Virtual patients: a critical literature review and proposed next steps.," *Medical education*, vol. 43, Apr. 2009, pp. 303-311.
- [3] M.H. Friedman, K.J. Connell, A.J. Olthoff, J.M. Sinacore, and G. Bordage, "Medical student errors in making a diagnosis," *AAMC Academic Medicine Journal of the Association of American Medical Colleges*, vol. 73, 1998, pp. 19-21.
- [4] A.S. Elstein, A. Schwartz, and A. Schwarz, "Clinical problem solving and diagnostic decision making:

- selective review of the cognitive literature.," *BMJ (Clinical research ed.)*, vol. 324, Mar. 2002, pp. 729-32.
- [5] B. Charlin, H.P. a Boshuizen, E.J. Custers, and P.J. Feltovich, "Scripts and clinical reasoning.," *Medical education*, vol. 41, Dec. 2007, pp. 1178-1184.
- [6] B. Charlin, C. Brailovsky, L. Roy, and F. Goulet, "RESEARCH BASIC TO MEDICAL EDUCATION The Script Concordance Test : A Tool to Assess the Reflective Clinician.," *Family Medicine*, 1989, pp. 189-195.
- [7] T. van Gog, F. Paas, and J.J.G. van Merriënboer, "Process-Oriented Worked Examples: Improving Transfer Performance Through Enhanced Understanding.," *Instructional Science*, vol. 32, Jan. 2004, pp. 83-98.
- [8] G. Gauthier, L. Naismith, S.P. Lajoie, and J. Wiseman, "Using Expert Decision Maps to Promote Reflection and Self-Assessment in Medical Case-Based Instruction."
- [9] S.P. Marshall, *Schemas in problem solving*, Cambridge Univ Pr, 1995.
- [10] M. Gartmeier, J. Bauer, H. Gruber, and H. Heid, "Negative Knowledge: Understanding Professional Learning and Expertise.," *Vocations and Learning: Studies in Vocational and Professional Education*, vol. 1, 2008, pp. 87-103.
- [11] F. Manola and E. Miller, "RDF primer.," *RDF recommendations*, vol. 2009, 2004.
- [12] D.L. McGuinness and F. van Harmelen, "OWL overview.," vol. 2009, 2004.
- [13] C.J. Wroe, J.J. Cimino, and a L. Rector, "Integrating existing drug formulation terminologies into an HL7 standard classification using OpenGALEN.," *Proceedings / AMIA ... Annual Symposium. AMIA Symposium*, Jan. 2001, pp. 766-70.
- [14] C. Golbreich, M. Horridge, I. Horrocks, B. Motik, and R. Shearer, "OBO and OWL: Leveraging semantic web technologies for the life sciences.," *The Semantic Web*, 2007, p. 169-182.
- [15] M.A. Haendel, F. Neuhaus, D. Osumi-Sutherland, P.M. Mabee, J.L.V. Mejino, C.J. Mungall, and B. Smith, "CARO — The Common Anatomy Reference Ontology.," *Anatomy Ontologies for Bioinformatics Principles and Practice*, vol. 2006, 2007, p. 327-349.
- [16] O. Hassanzadeh, A. Kementsietsidis, L. Lim, R.J. Miller, and M. Wang, "LinkedCT: A Linked Data Space for Clinical Trials.," *Arxiv preprint arXiv09080567*, 2009, p. 5.
- [17] A. Jentzsch, J. Zhao, and O. Hassanzadeh, "Linking open drug data.," *Proceedings of the*, 2009, pp. 3-6.
- [18] a L. Rector, J.E. Rogers, P.E. Zanstra, and E. Van Der Haring, "OpenGALEN: open source medical terminology and tools.," *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium*, Jan. 2003, p. 982.
- [19] A. Gangemi, "Ontology Design Patterns for Semantic Web Content.," *The Semantic Web / ISWC 2005*, Springer Berlin / Heidelberg, 2005, pp. 262-276.
- [20] T. Berners-Lee, "Notation3 (N3): A readable RDF syntax.," 1998.