Automatic Feedback Generation in Virtual Patients Using Semantic Web Technologies

Jean-Rémy Duboc, Prof. Hugh Davis, Dr. Sunhea Choi, Dr. Mark J. Weal

University of Southampton, School of Electronic and Computer Science and School of Medicine

Abstract

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 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, reflect on their decisions and improve their clinical reasoning skills.

This paper presents research on automatic feedback generation for virtual patients, using semantic web technologies. To generate feedback, a computer model has been designed to represent virtual patients and students' interactions, using semantic web technologies. The use of semantic web technologies allows a computer readable connection between medical conditions, their symptoms and the corresponding examinations. Some of these connections can be pulled from existing linked data available on the web, which would facilitate the creation and maintenance of virtual patient data.

A survey has been conducted to determine the most useful types of feedback for medical students. Relating this encoded knowledge to data describing the student's choices of examinations allows the automatic generation of such feedback in virtual patients.

Exposure to many different clinical scenarios is a crucial part of medical education. Indeed, medical students acquire their skills and knowledge by accumulating a great deal of experience meeting patients with a wide variety of presentations and pathologies (Gartmeier, Bauer, Gruber, & H. Heid, 2008). Existing research suggests that medical expertise is not only the result of superior reasoning abilities or deeper theoretical knowledge, but rather stems from contextualised information about prototypical or actual patients (G. Norman, 2005; Schmidt, G. R. Norman, & Boshuizen, 1990). As a result, it is of crucial importance for medical students to study many different medical cases in order to improve their clinical reasoning skills.

However, several factors limit the time and frequency of medical students' encounters with real patients: difficulty to obtain patients' consent, budgetary limits, health and safety precautions, rarity of certain conditions, etc. To enhance students' performances despite these constraints, many medical schools use computer systems designed to emulate clinical cases and expose students to as many clinical scenarios as possible in an interactive manner (Cook & Triola, 2009; Kenny, Parsons, Gratch, Leuski, & Rizzo, 2007; Poulton, Conradia, Kavia, Rounda, & Hilton, 2009). Additionally, these virtual patient systems allow students to practice their diagnostic reasoning skills in a safe environment, using the feedback provided by the system as a learning tool (Larsen, Butler, & Roediger, 2008).

This research explores the technological potential of the semantic web for the design of virtual patient systems aimed at medical students in their early clinical years. At that stage of the medical curriculum, students express the need for self-study tools allowing them to safely test their knowledge and reasoning skills on a variety of clinical cases.

The first section of this paper presents the results of a survey conducted among students to determine the most useful types of feedback in a virtual patient. A classification of seven types of feedback has been established, and students were asked to evaluate each type of feedback on a Likert scale.

The second section describes the ontology of virtual patients designed to support automatic feedback generation, and explains the benefits of semantic web techniques over relational databases and XML. These benefits include easy re-use of existing data available on the web, and sound knowledge representation based on formal logic, allowing reasoning and querying across a variety of data sources in order to generate the types of feedback considered most useful by students.
Such a model allows the generation of feedback regarding the chosen questions and examinations, the diagnoses proposed by the students, and the evolution of their differential diagnosis.

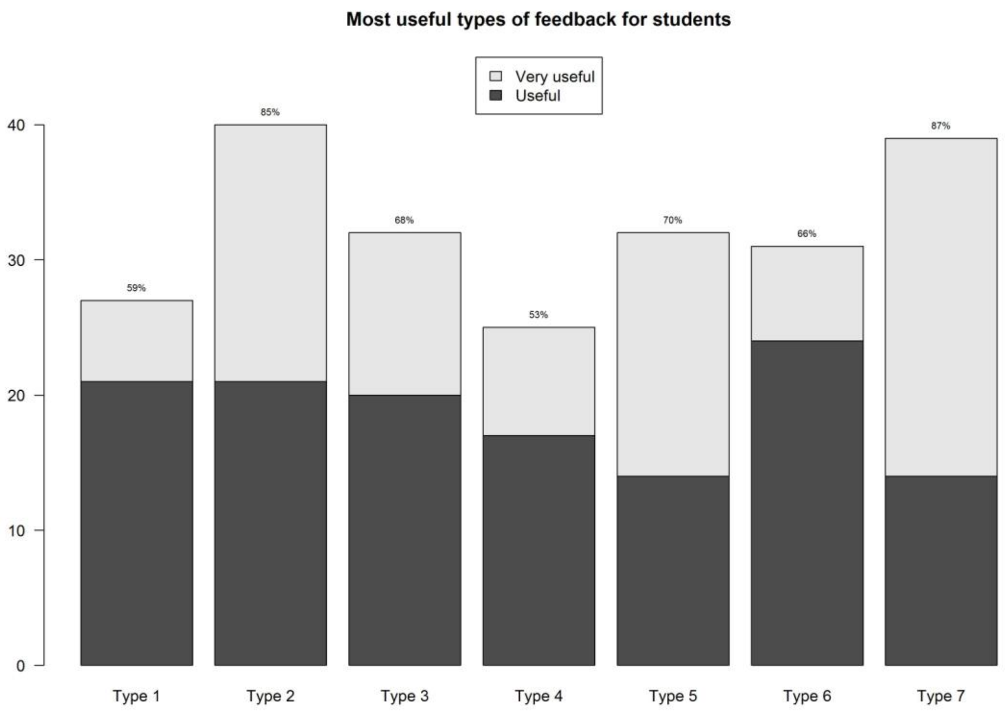
Survey Results: Most Useful Types of Feedback

Many virtual patient systems have been developed in recent years (Cook & Triola, 2009; Huwendiek, Haider, Toenshoff, & de Leng, 2009). Common types of virtual patient interactions and feedback have been identified, and a review of the technologies used for virtual patient data has been conducted.
Using this literature, in addition to literature on assessment in medical education, seven main types of virtual patient feedback have been identified. A classification of these types of feedback has been listed as follows:

- **Feedback Type 1: "A list of interview questions and examinations and tests the student should *NOT* have chosen, and the justification (not appropriate, irrelevant, redundant, etc.)".**
  This includes all actions that can be considered irrelevant to the case, and constitute a waste of time and money in a clinical setting, or represent a danger or an unnecessary discomfort for the patient.
- **Feedback Type 2: "A list of the interview questions and examinations students *should* have chosen, and the justification (type of disease to consider, related symptom to check, possible conditions to rule out, etc.)".**
  Missing relevant interview questions and examinations can lead students to an incorrect diagnosis.
- **Feedback Type 3: "If the diagnosis is wrong, feedback telling the student if the chosen diagnosis is still coherent with the results of the chosen interview questions and examinations".**
  Even though students might miss relevant examinations and questions (feedback type 2), the incomplete assessment of the student might still lead to a coherent diagnosis, even though the diagnosis might be incorrect. This feedback type provides useful insight when combined with feedback types 1 and 2.
- **Feedback Type 4: "Feedback about the order in which student performs specific actions".**
  This includes feedback related to sequences of choices that represent a logical progression. For instance, if the student takes the patient's temperature and observes a fever, it is logical to investigate various types of infections that could affect the patient afterwards. However, checking for infections before confirming that the patient has a temperature might not be a logical course of action.
- **Feedback Type 5: "A sequence of the "ideal" history taking and examination process that an expert would use, with the rational for each step".**
  This feedback type is derived from feedback type 4, and provides an example to follow for the clinical case presented in the virtual patient system.
- **Feedback Type 6: "If student chooses an inadequate action, a narrative description of the consequences on the patient, if applicable".**
  This is also closely related to feedback type 4, and provides important information concerning critical actions with potentially dangerous consequences. Although sometimes applicable for examination and interview questions, this type of feedback is most useful to provide feedback regarding the student's choice of treatment.
- **Feedback Type 7: "A list of all diagnoses the student should have tested and ruled out, given the initial presentation of the patient".**
  This is an important type of feedback that can help students to generate hypotheses when trying to establish a diagnosis. Indeed (Friedman, Connell, Olthoff, Sinacore, & Bordage, 1998) show that experts tend to generate a smaller number of diagnosis hypotheses, but with more accuracy. Students need practice to generate more accurate diagnoses.

A survey was conducted among 50 medical students in the School of Medicine in the University of Southampton. Most students were in year 3, and others were in year 4 and 5. Students in year 1 and 2 were not interrogated, because they are not in clinical practice yet. Students were asked to rate each feedback type on a 6-point Likert scale ranging from "Completely useless" to "Very Useful". Most feedback types were most considered useful or very useful by most participating students. Feedback 2 and 7 seem to be the most useful types of feedback, with respectively 85% and 87% of student considering them useful of very useful. The results of the survey are displayed on figure 1.

Figure 1. Survey result: evaluation of seven virtual patient feedback types by students in year 3, 4 and 5.



A semantic model of virtual patient has been designed to allow the automatic generation of feedback, in particular of feedback type 2 and 7.

Semantic Virtual Patient Model

Overview of the Virtual Patient Ontology

An ontology, in computer science, is a formal representation of knowledge as a set of concepts within a domain, and the relationships between those concepts.
An ontology representing the domain of a virtual case has been designed using OWL (Web Ontology Language) (McGuinness & van Harmelen, 2004). The virtual patient ontology makes use of many concepts (or classes) and properties from OpenGalen, an extensive biomedical ontology (Rector, Rogers, Zanstra, & Van Der Haring, 2003). Re-using concepts and properties from external ontologies is straightforward in OWL, as each concept or property has a unique URI to reference it.

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 (simplified notation for semantic web data):

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.

Generating Feedback from the Virtual Patient 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.

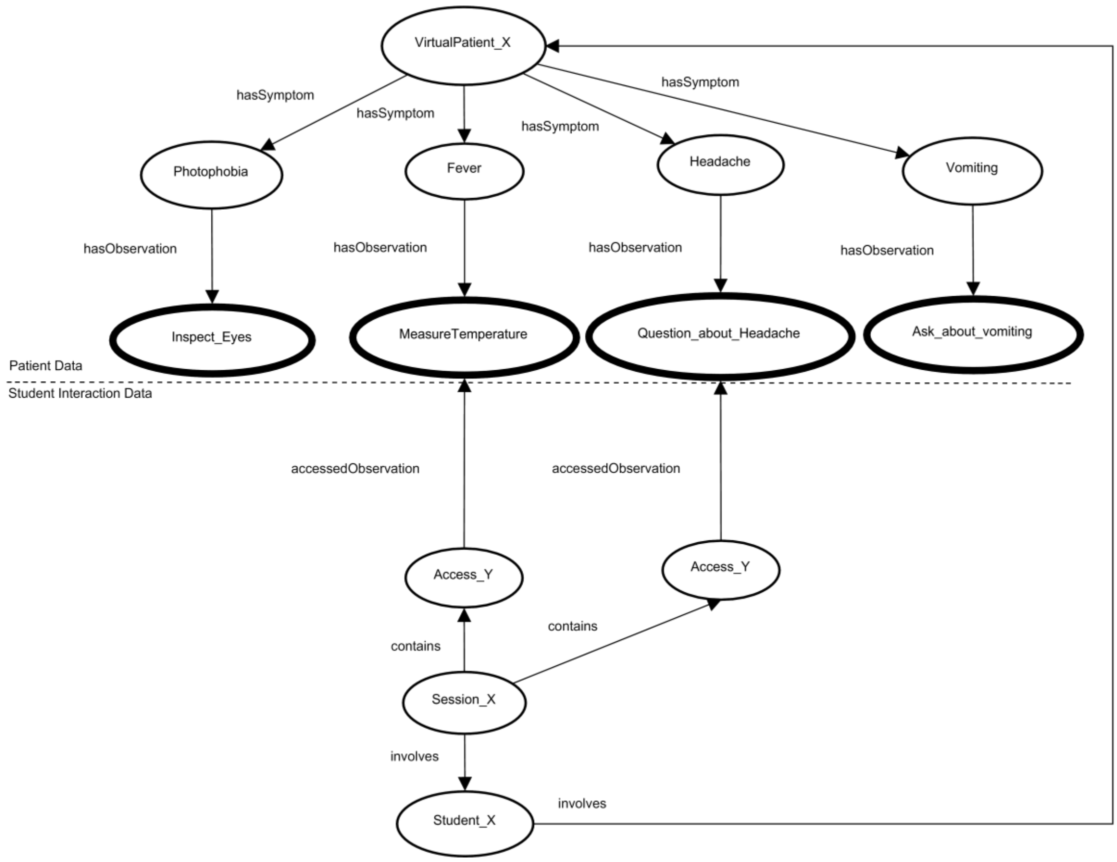
Feedback Regarding Chosen Examinations and Questions

Students can select the questions and examinations that they want to perform. It is possible to show students the important questions and examinations they neglected to choose during their investigation (feedback type 2). On a more complex level, it is also possible to determine if the student missed an important symptom that they should have identified through the result of an examination or the answer to a question.

In the virtual patient ontology, a variety of symptoms are identified, and linked to the observations needed for students to access 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 will be explicitly marked as important by the virtual patient author, and connected to the condition they confirm or rule out.

In figure 2, data from the student's session shows that the student neglected to ask the patient about any vomiting, and to inspect the patient's eyes for photophobia. Both of these symptoms confirm the final diagnosis, in this case meningitis.

Figure 2. Data representing important observations for a patient affected by Meningitis. The nodes in bold represent important observations.



The following set of rules, written here in natural language, identifies important observations using 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 important 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. After a list of all important observations has been generated using the rules above, a simple query generates a list of all important actions that the student missed. The query could be described as follows:

*"SELECT all nodes **?observations***
*WHERE **?observation** belongs to the class **ImportantObservation** MINUS all **?observations** that the student has selected."*

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 is then displayed to the student in a human-readable format.

Feedback Regarding Proposed Diagnoses

Students can propose several diagnoses for the case, and associate a level of certainty for each proposed diagnosis, at any given time, before proposing a final diagnosis at the end of the case.
Using data describing the student's choice of questions and data describing the diagnoses they proposed, it is possible to determine if the student has not correctly interpreted the result of the observations they chose (feedback types 3 and 7).

If a student proposes an incorrect final diagnosis, either the student has not chosen the right observations, or the right observations have been chosen, but they have been interpreted incorrectly.
The problem of generating feedback in the first situation I largely solved by the mechanism described in the previous section: Feedback regarding chosen examinations and questions. However, it is possible to go further, by indicating which examinations are coherent with each of the diagnoses proposed by the student during the case.

Figure 3. Data showing a student session with a choice of observation and two diagnosis suggestions.

The following query returns variable bindings containing the node representing a condition, and an observation associated to that condition.

"SELECT all nodes **?condition**, **?observation** and **?symptom**
WHERE **?symptom** is a symptom of **?condition,**
?symptom has observation **?observation**,
The student's session contains a diagnosis proposal,
The diagnosis proposal has **?condition** as its proposed diagnosis.
ORDER the result by **?condition**"

The resulting bindings are sorted by condition. The semantic links between conditions and symptoms can be extracted from the Freebase knowledge base (http://www.freebase.com/).

The example data illustrated in figure 3 would return the following data:

Table 1. Data returned by a query searching for observations chosen by the student with associated condition and symptom.

condition	observation	symptom
Meningitis	"Measure Temperature"	"Fever"
Influenza	"Measure Temperature"	"Fever"

This table gives indications to the student regarding the relevance of the chosen examinations, in relation to the proposed diagnoses. The condition "Sepsis" is not listed, for it has not been proposed by the student in this example.

Feedback can also be generated by starting from the final diagnosis suggested, and determining if the observations chosen are consistent with that diagnosis. Using again the data illustrated on figure 3, and assuming that the final diagnosis "Urinary tract infection" has been suggested, the following query will return all observations chosen by the students that or consistent with this diagnosis:

"Select all nodes **?observation** and **?symptom**
WERE **?symptom** is a symptom of Urinary infection,
?symptom has **?observation** as an observation."

The resulting data is shown on table 2. As fever is indeed a symptom of a urinary tract infection, measuring temperature is a coherent choice of observation. In this case, however, the diagnosis is wrong because other observations have to be chosen to investigate the patient further and differentiate all possible diagnoses. A list of these important investigations is displayed to the student using the mechanism described above, in Feedback regarding proposed diagnoses.

Table 2. Data returned by a query returning all observations chosen by the student that are consistent with the chosen final diagnosis

observation	symptom
"Measure Temperature"	"Fever"

A Note on Feedback Regarding Prescribed Treatments

Feedback on the treatment options prescribed by students important to students in their last years of training, when they must be able to propose suitable management solutions and treatments once they have established a diagnosis.

Treatment mistakes come in two forms: either providing the wrong treatment, with eventual undesirable side effects, or failing to provide the correct treatment. This can have various consequences ranging from discomfort to further complications and death. These consequences are described directly in the virtual patient data. Secondary effects of certain drugs can also be pulled out of the SIDER ontology

(Kuhn, Campillos, Letunic, Jensen, & Bork, 2010).
By linking conditions to their appropriate treatment and to their potential side effects, generating feedback using queries similar to those proposed above should be a simple matter. However, this study is focused on feedback regarding history taking skills, in particular for year 3 students who are not yet required to prescribe treatments. Thus, feedback regarding treatment options will be treated at a later time.

Conclusion and Future Work

To determine if the important types of feedback are indeed provided adequately using the model presented on this paper, a prototype virtual patient system is currently under implementation.
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 on a larger sample.
This will inform the design of future virtual patients to be used in the School of Medicine.

Interests

This research is funded by the Strategic Health Authority (SHA). There is no conflict of interest.

References

Azan, B., & Ellaway, R. (2010). MedBiquitous Virtual Patient Specifications and Description Document, (1), 1-128.

Berners-Lee, T. (1998). Notation3 (N3): A readable RDF syntax. W3C. Retrieved from http://www.w3.org/DesignIssues/Notation3.

Cook, D. a, & Triola, M. M. (2009). Virtual patients: a critical literature review and proposed next steps. *Medical education*, *43*(4), 303-311. doi: 10.1111/j.1365-2923.2008.03286.x.

Fischer, M. R. (2000). CASUS - An Authoring and Learning Tool Supporting Diagnostic Reasoning. *Zeitschrift für Hochschuldidaktik*, *1*, 87-98.

Friedman, M. H., Connell, K. J., Olthoff, A. J., Sinacore, J. M., & Bordage, G. (1998). Medical student errors in making a diagnosis. *AAMC Academic Medicine Journal of the Association of American Medical Colleges*, *73*(10), 19-21.

Gartmeier, M., Bauer, J., Gruber, H., & Heid, H. (2008). Negative Knowledge: Understanding Professional Learning and Expertise. *Vocations and Learning: Studies in Vocational and Professional Education*, *1*(30), 87-103.

Huwendiek, S., Haider, H. R., Toenshoff, B., & Leng, B. de. (2009). Evaluation of curricular integration of virtual patients: Development of a student questionnaire and a reviewer checklist within the electronic virtual patient (eVIP) project.

BAMS, *5*(9), 35–44. Retrieved April 12, 2011, from http://149.156.12.42/bams/bams3_pdf/EVALUATION OF CURRICULAR INTEGRATION OF VIRTUAL from BAMS9-17.pdf.

Kenny, P., Parsons, T. D., Gratch, J., Leuski, A., & Rizzo, A. A. (2007). Virtual Patients for Clinical Therapist Skills Training. *IVA 07: Proceedings of the 7th international conference on Intelligent Virtual Agents* (pp. 197-210). Berlin, Heidelberg: Springer-Verlag. doi: http://dx.doi.org/10.1007/978-3-540-74997-4_19.

Kuhn, M., Campillos, M., Letunic, I., Jensen, L. J., & Bork, P. (2010). A side effect resource to capture phenotypic effects of drugs. *Molecular systems biology*, *6*, 343. EMBO and Macmillan Publishers Limited. Retrieved August 24, 2010, from http://dx.doi.org/10.1038/msb.2009.98.

Larsen, D. P., Butler, A. C., & Roediger, H. L. (2008). Test-enhanced learning in medical education. *Medical education*, *42*(10), 959-66. doi: 10.1111/j.1365-2923.2008.03124.x.

Manola, F., & Miller, E. (2004). RDF primer. (M. E., Ed.)*RDF recommendations*, *2009*(12 october). W3C. Retrieved from http://www.w3.org/TR/REC-rdf-syntax/.

McGuinness, D. L., & Harmelen, F. van. (2004). OWL overview, *2009*(12 october). W3C. Retrieved from http://www.w3.org/TR/owl-features/.

Norman, G. (2005). Research in clinical reasoning: past history and current trends. *Medical education*, *39*(4), 418-27. doi: 10.1111/j.1365-2929.2005.02127.x.

Poulton, T., Conradia, E., Kavia, S., Rounda, J., & Hilton, S. (2009). The replacement of "paper" cases by interactive online virtual patients in problem-based learning. *Medical Teacher*, *31*, 752-758. doi: 10.1080/01421590903141082.

Rector, a L., Rogers, J. E., Zanstra, P. E., & Van Der Haring, E. (2003). OpenGALEN: open source medical terminology and tools. *AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium*, 982. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1480228&tool=pmcentrez&rendertype=abstract.

Ruderich, F., Bauch, M., Haag, M., Heid, J., Leven, F. J., Singer, R., et al. (2004). CAMPUS--a flexible, interactive system for web-based, problem-based learning in health care. *Studies in health technology and informatics*, *107(Pt 2)*, 921-925. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15360947.

Schmidt, H. G., Norman, G. R., & Boshuizen, H. P. (1990, October). A cognitive perspective on medical expertise: theory and implication. *Academic medicine : journal of the Association of American Medical Colleges*. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18214707.