CRISTAL: Adapting Workplace Training to the Real World Context with an Intelligent Simulator for Radiology Trainees

Hope Lee[1], Amali Weerasinghe[1], Jayden Barnes[1], Luke Oakden-Rayner[2], William Gale[1], Gustavo Carneiro[1]

[1] Department of Computer Science, University of Adelaide, SA, Australia
[amali.weerasinghe, gustavo.carneiro]@adelaide.edu.au
[2] Department of Radiology, Royal Adelaide Hospital, SA, Australia
lukeoakdenrayner@health.sa.gov.au

Abstract. Intelligent learning environments based on interactions within the digital world are increasingly popular as they provide mechanisms for interactive and adaptive learning, but learners find it difficult to transfer this to real world tasks. We present the initial development stages of CRISTAL, an intelligent simulator targeted at trainee radiologists which enhances the learning experience by enabling the virtual environment to adapt according to their real world experiences. Our system design has been influenced by feedback from trainees, and allows them to practice their reporting skills by writing freeform reports in natural language. This has the potential to be expanded to other areas such as short-form journalism and legal document drafting.

Keywords: Adult Learning · Self-Regulated Learning · Simulated Environments for Learning · Radiology Training · Natural Language Processing

1 Introduction

Learning by doing: this is the underlying concept of intelligent tutoring systems (ITSs), serious games and immersive activities, and learning environments such as these are becoming increasingly popular. Their ability to generate an individualized environment tailored to each user's learning needs is one of their main strengths. These adaptations are generally based on a learner's interactions within the digital world, and as a result, learners find it difficult to transfer their knowledge to real world tasks [1]. Consequently, there has been a strong interest in linking learners' real world behaviors with the digital world to facilitate a more integrated learning experience. On-the-job training is one example of this. The focus of the systems ImREAL [1], MIRROR [2], PORML [3], ALPS [4] and KP-Lab [5] is to assist informal workplace training by supporting learners to transform experience into knowledge through socio-pedagogical models. Out of these, only ImREAL [2], MIRROR [3] and PORML [4] link both worlds by giving learners simulated tasks that correlate with real world activities. However, these systems rely on the learners to actively partici-

adfa, p. 1, 2011.

pate in knowledge transfer between the workplace and the simulated environment. The PORML [4] framework supports reflection in a digital environment for emergency service workers immediately after performing a job activity. However, the adaptation is specific to that job activity and does not influence the system's behavior in subsequent activities.

As the domain used in this research is radiology, we have also explored how ITSs have supported learning in this area. VIA-RAD [6], RadTutor [7] and MR Tutor [8] present a constrained format where users are asked to describe radiology images by selecting options from a pre-determined list. These systems focus entirely on the virtual world, and real world behaviors are not considered within the simulation. In contrast, a radiology ITS called GIMI (Generic Infrastructure for Medical Informatics) uses real world experiences to modify the learner model by allowing manual input of personal performance data [9]. In contrast to GIMI we plan to automate the integration of real world experiences into our simulator with data mining techniques.

In our research towards developing CRISTAL (Clinical Radiology Intelligent Simulation Tool with Adaptive Learning) we explore the following objectives:

1. How can we use real world behaviors to adapt the virtual learning experience?
2. How can we automate the process of integrating real world behaviors with the simulated environment without active intervention from the user?
3. How can we provide intelligent support for radiology trainees' learning that is relevant to the requirements of their job?

Using real world behaviors to tailor the simulated learning experience, we can enable trainees to seamlessly transition between the roles of worker and learner. The overall goal of our research is to explore how an intelligent simulator that links both real and virtual environments can support individuals as they progress through different phases of their training: starting as workers, progressing to learners and finally becoming experts.

2 Radiology Training Practices

Trainee radiologists are qualified doctors who are enrolled in the Royal Australian and New Zealand College of Radiologists' five-year training program to become consultant radiologists. A survey conducted by the Royal Australian and New Zealand College of Radiologists (RANZCR) in 2012 found that the majority (89%) of trainees spend at least 36 hours per week on clinical work, the majority of which is spent writing reports on radiology images [10].

We asked a group of trainees for their input regarding the type of experience they wanted from a simulated tutoring system. They told us they want a system that makes it easy to identify important weaknesses, access relevant cases, gives high-quality but targeted feedback and suggests (rather than demands and tests on) relevant study material. They disliked systems that impose set exercises and learning content, and responded negatively to the idea of adaptive dialogues. Their response was: "We know how to learn, and we have our own preferred resources." The trainees' responses echo some of Malcolm Knowles' well-known assumptions regarding adult learners: they have a clear preference for self-directed learning, a strong internal motivation to learn, and are oriented towards learning tasks that have immediate relevance to their societal roles [11]. This emphasis on andragogy (the theory of adult learning) is echoed by the authors of the ImREAL project [2].

3 System Overview

As per the extended self-regulated learning (SRL) model described by Hetzner et al. [1], the architecture of CRISTAL spans across the real world and the simulated environment (Fig. 1). We will first describe the real world environment. In the workplace, the majority of images that a trainee reports on are subsequently sent to a consultant radiologist, who then types an addendum containing any necessary corrections. We plan to collect and store all trainees' reports and their corrections in our database. This information will be used to update the learner model in the simulated environment, enabling the training module to adapt the learning task based on a trainee's weaknesses in the real world. The training module will then request relevant problems from the report and image database. This database consists of real radiology images (such as x-rays and CT scans) and their corresponding reports. These reports have been written on-the-job by domain experts in the past.

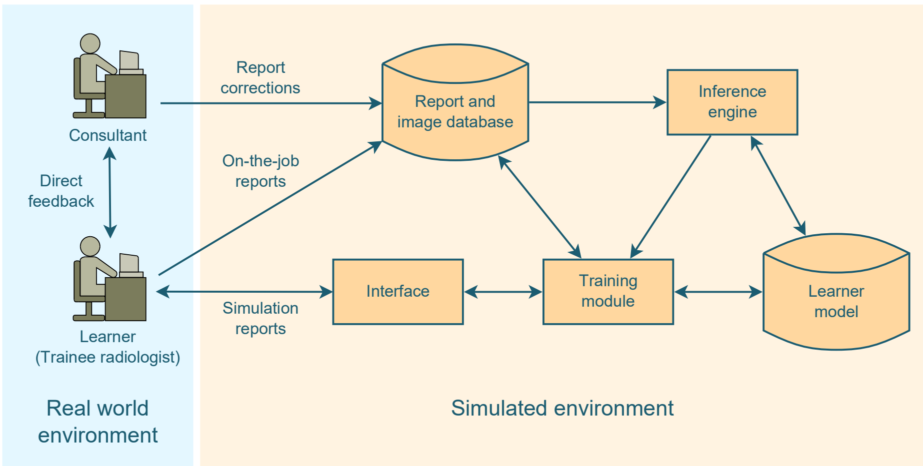

Fig. 1. Diagram of the system architecture

Mirroring the real world environment, the simulator interface will present one of these images to the learner (trainee radiologist), who will be prompted to write a report on the image in freeform natural language. Once complete, the report will be sent to the inference engine to be compared with the original report. The quality of the trainee's report will be determined both by the presence of the correct diagnosis and the completeness of the report.

We are currently using a latent semantic indexing (LSI) model to characterise each sentence in our report corpus. To determine the completeness of the trainee's report, each sentence is compared with sentences from the original report and matched with the one with which it has the greatest cosine similarity. If a sentence pair's similarity is above a pre-defined threshold the trainee's sentence is considered to contain appropriate meaning. Below this threshold, the trainee's sentence is considered to be incorrect. The training module will also detect missing sentences: important sentences in the original report that were not identified in the trainee's report, via the same thresholding approach. The learner model will be updated with data regarding the correctness and completeness of the trainee's report.

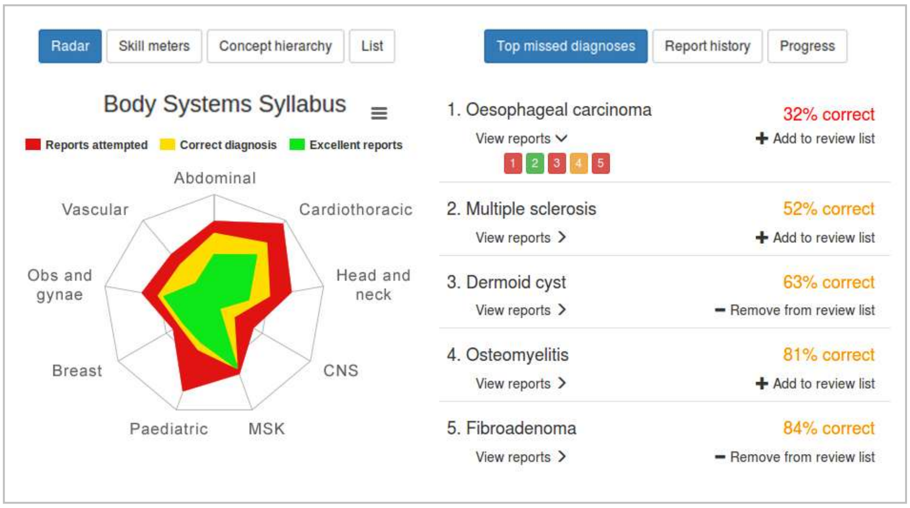

Fig. 2. Screenshot of the open learner model

Trainees will receive feedback directly after report submission, however they will also be able to review their overall performance through an open learner model (Fig. 2). Diagnoses are grouped according to the Body Systems Syllabus of the RANZCR Radiodiagnosis Training Program Curriculum [12], and multiple views are provided to enable trainees to select their preferred format. The importance of diagnoses will be determined based on their categorization within the Curriculum. The Curriculum will also influence problem selection: for example, first year trainees will be provided with images related to critical diagnoses (known as Key Conditions in the Curriculum), with a specific emphasis on conditions they have not seen in practice and those they have reported incorrectly. Senior trainees will instead be exposed to more esoteric conditions (Categories 2-3 in the Curriculum) selected from areas they have had limited exposure to, or reported with high rates of error. To respect their autonomy, trainees will also have the option to select specific learning topics.

4 Preliminary Testing

We have conducted some initial basic tests to assess the quality of our language model using LSI. To simulate missing and incorrect sentences, whole sentences were removed at random, or swapped in from reports containing different diagnoses. We tested how many of these abnormal sentences were detected by our system relative to the number of alterations. As expected, the variation in missing and incorrect sentences is directly proportional to the number of alterations made in the reports (Fig. 3), suggesting we are able to identify unmatched sentences. We do note however that the LSI model is unlikely to achieve acceptable performance in all of the required tasks. This training module will be improved by the implementation of a recurrent neural network language model to overcome the more difficult challenges: the identification of diagnostic sentences (which are most important for the teaching process), and the discrimination between positive and negative diagnostic sentences (as negations are not well captured with naive "bag-of-words" models like LSI [13]).

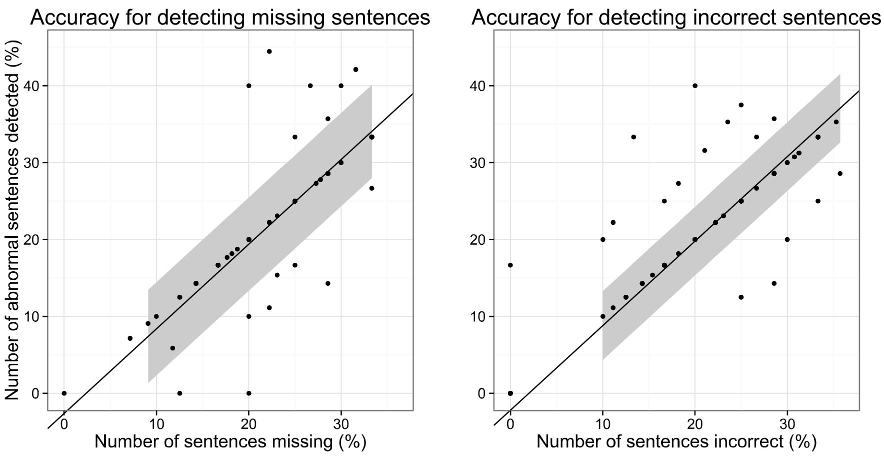

Fig. 3. Preliminary tests of the accuracy of missing and incorrect sentence detection

5 Future Work

At this stage we have developed a working model of CRISTAL's simulated learning environment. Our current focus is developing the inference engine for processing trainees' on-the-job reports and their corresponding consultant corrections. This will complete the feedback loop between the training module and the real world environment. Our next step for clinical implementation is preliminary evaluation of the simulation environment. We will present a set of pre-determined training scenarios to trainees and adjust our model based on their feedback.

The strengths of CRISTAL are that it allows automated integration of real world data into the learner model, and it has the ability to analyse and give feedback on freeform reports. Our system has the potential to be extended to ITS research in other domains, if those domains fulfil the following criteria: there is a real-world task in which adult learners could benefit from intelligent feedback and adaptive training, the task is performed frequently and results in text output, and there is a written "ground truth" for this text to be compared against. Examples could include education settings with short answer questions (including in online education), as well as professions such as short-form journalism and legal document drafting, where each document is edited by a senior practitioner. Thus our system of connecting the real-world and the simulated environment can be seen to apply more widely to ITS research, contributing to the impact and relevance of virtual learning environments.

6 References

1. Hetzner, S., Steiner, C.M., Dimitrova, V., Brna, P., Conlan, O.: Adult Self-regulated Learning through Linking Experience in Simulated and Real World: A Holistic Approach. Lecture Notes in Computer Science 6964, 166-180 (2011)
2. Krogstie, B.R., Prilla, M., Wessel, D., Knipfer, K., Pammer, V.: Computer Support for Reflective Learning in the Workplace: A Model. In: Advanced Learning Technologies (ICALT), 2012 IEEE 12th International Conference on, pp. 151-153. IEEE, Italy (2012)
3. Eamsinvattana, W.: Reflective Dialogue for On-the-Job Training in Emergency Services. School of Computing. University of Leeds, United Kingdom (2011)
4. Murphy, A., Laxton, J.: Views of a Structured Assessment Tool for Observing Practice. Social Work Education 33, 190-208 (2014)
5. Kai, H., Hanni, M., Hannu, M.: Design principles for the Knowledge-Practices Laboratory (KP-Lab) project. Proceedings of the 7th international conference on Learning sciences, pp. 934-935. International Society of the Learning Sciences, Bloomington, Indiana (2006)
6. Rogers, E.: VIA-RAD: a blackboard-based system for diagnostic radiology. Artificial Intelligence in Medicine 7, 343-360 (1995)
7. Azevedo, R., Lajoie, S.: The Cognitive Basis for the Design of a Mammography Interpretation Tutor. International Journal of Artificial Intelligence in Education 9, 32-44 (1998)
8. Sharples, M., Jeffery, N.P., du Boulay, B., Teather, B.A., Teather, D., du Boulay, G.H.: Structured Computer-based Training in the Interpretation of Neuroradiological Images. International Journal of Medical Informatics 60, 263-280 (2000)
9. Yap, M.H., Gale, A.G., Scott, H.J.: Generic Infrastructure for Medical Informatics (GIMI): The Development of a Mammographic Training System. Lecture Notes in Computer Science 5116, 577-584 (2008)
10. Munro, P.L., Bradshaw, N., Stephenson, N.: 2012 RANZCR Radiology Workforce Census Report: Trainees. RANZCR, Sydney (2013)
11. Knowles, M.S.: The modern practice of adult education. New York Association Press, New York (1970)
12. The Royal Australian and New Zealand College of Radiologists: Radiodiagnosis Training Program Curriculum. RANZCR, Sydney (2014)
13. Wiegand, M., Balahur, A., Roth, B., Klakow, D., Montoyo, A.: A Survey on the Role of Negation in Sentiment Analysis. Proceedings of the Workshop on Negation and Speculation in Natural Language Processing, pp. 60-68. Association for Computational Linguistics (2010)