

Tracking and Annotation in Skills-Based Learning Environments

Mark J. Weal, Danius T. Michaelides, Kevin R. Page and David C. De Roure

IAM Group, University of Southampton, UK

{mjw}{dtm}{krp}{dder}@ecs.soton.ac.uk

Mary Gobbi and Eloise Monger

School of Nursing and Midwifery, University of Southampton, UK.

{m.o.gobbi}{em}@soton.ac.uk

Fernando Martinez

Mixed Reality Lab, University of Nottingham, UK.

fxm@cs.nott.ac.uk

Abstract

This paper outlines a series of experiments looking at the annotation and subsequent analysis of skills-based learning and teaching in the domain of Nursing. The experiments used Semantic Web technologies to create and store annotations of the student's activities. A number of techniques were used to create these annotations ranging from manual textual annotations by observers through to automatic spatial annotations from a location tracking system. Combining these annotations sets points towards new methods for analysis of both student learning and educational viewpoints on such skills-based labs. Initial observations from these proof of concept studies will be presented along with future directions the work may take.

1 Introduction

This paper sets out a series of proof of concept deployments of semantic annotation systems within a skills-based learning environment in the domain of Nursing. In the UK, skills-based learning has formed a part of a government agenda to help ensure that practitioners are “fit for practice”. Such techniques can help improve competence levels and working practices [10]. We have previously extracted annotations from collaborative tools (instant messaging, knowledge mapping, and intelligent task lists) to create replay structure for distributed meetings [11].

The experiments were centred on scenario-driven skills-based learning in hospital ward simulations [4]. In the School of Nursing and Midwifery (SONAM), Southampton, students take part in skills-based learning scenarios in ward simulations located within the teaching facilities.

A typical skills-based session involves a group of three or four students carrying out a scenario lasting half an hour. Teaching staff take the role of doctors and consultants, and

the clinical scenarios are constructed from real patient data to provide convincing and realistic contexts and responses. The simulated wards (shown in Figure 1) contain up to a number of beds, sinks, a fully equipped nurses station and equipment that would be expected in a normal hospital ward, crash cart, oxygen, etc. The patients are portrayed by mannequins. These are sometimes instrumented, as in the case of SimMan¹ (see Figure 5), allowing the students to assess the patients by performing patient assessment tasks such as taking their pulse [1]. The SimMan mannequins provide numerous multi-sensory physical outputs (e.g. blood pressure, heart rate, and voice/respiratory sounds) as well as event logging. It contains a loudspeaker through which facilitators can pretend to speak as the patient and the patient's condition can be altered during the course of the scenario programmatically.



Figure 1. The ward setup with mannequins and equipment.

The simulations are designed to promote the acquisition

¹<http://www.laerdal.com/SimMan>

of practical skills as well as decision-making, team working, communication and problem-solving etc. [3]. They are incorporated into assessment of student performance [12] requiring that the approaches for assessment and feedback need to be sound, valid, reliable, feasible, educational, and importantly acceptable to practitioners [7]. Although simulated, the student experiences are exactly as they would experience in the workplace in real time. Hence, when the students and mentors are ‘immersed’ in the simulation and behave ‘as in real practice’ the video captured data provides important information about their performance.

The virtual ward is instrumented for recording with six cameras. The cameras are controllable from the central control room with a 360° viewing angle, enabling them to be pointed at the beds or perhaps the nursing station depending on where the activity of the scenario is taking place. Often multiple cameras might be used to record activity around a single bed in order to gain clear views of activities. Microphones suspended above the beds are used to record audio. The video feeds are currently recorded using a bank of DVD recorders. The cameras and microphones provide the basic mechanism for lecturers and researchers to observe the students carrying out the learning activities, see Figure 2.



Figure 2. The students in a scenario.

The trials presented in the paper have sought to augment this basic setup with semantic annotation of the physical space through a combination of textual and audio annotations, annotations derived automatically from the SimMan mannequins, and spatial annotations obtained through the deployment of an indoor location tracking system. It is hoped that the subsequent analysis of these skills-based learning activities and gain new insights into what actions students perform during these tasks, how they move physically about the space, how students learn in such scenarios and in addition, what the mentors are seeing when they observe and evaluate students as they perform their

tasks. These last question is of particular interest as it is recognised that mentors whilst observing such activities can quickly identify ‘able’ and ‘less able’ students, but it is less clear what specifically they might be observing that leads them to make these distinctions.

As well as allowing the repurposing of the video material for subsequent teaching, the annotations can be used to help researchers more quickly hone in on areas of interest in the data and readily identify patterns across large datasets that would be prohibitively time consuming to analyse through more traditional transcription or video review methods.

The next section details the annotation approach taken along with the variety of annotations created during the trials. Subsequently replaying and analysis of the annotations will be discussed before conclusions are drawn and planned future work is described.

2 Annotating activities

In order to construct annotations of the activities taking place within the space Semantic Web technologies were chosen as the structuring and storage methods. An ontology was constructed through a co-design process.

In constructing the ontology the aim was to identify what observable characteristics and features of the students’ performance could form valid and reliable criteria for formative and/or summative assessment purposes. In this field, little or no benchmarking of ‘natural’ performance capability for students or practitioners exists. To some extent this is because current technologies have not been developed to handle the vast volume of data generated by a single student.

2.1 Textual annotations

Video annotation systems for other domains exist, for example news production [9] where the focus is on more explicit description of content or for collaborative annotation of video [2]. In the case of our textual annotations, the authoring process can occur both in real-time or post-exercise, with the annotations potentially reused throughout this process. An example of this would be coarse annotations made during the exercise being used as a basis for creating more detailed annotations about specific activities and events at a later point.

For the purposes of the trials we used naturalistic time sequenced observation, followed by clustering of themes according to discipline specific relationships, for example ‘taking a pulse’ was clustered under a heading of ‘taking and recording vital signs’. The individual activity of the pulse then breaks down into further components.

Researchers are able to view the scenario on a monitor in the control room. Using a simple click to add annotation

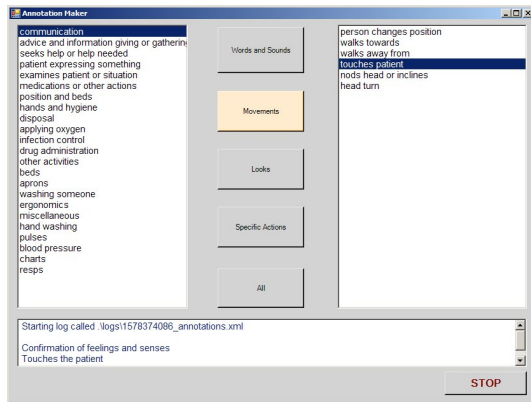


Figure 3. Adding Textual Annotations.

authoring tool (shown in Figure 3) researchers can annotate what they see on the video monitor. The interface is kept simple, with double clicking on the annotation name creating a new time stamped annotation. This allows the lecturer to spend the maximum time watching the video monitor and the minimum time interacting with the interface. Due to the cognitive load of annotating in real-time, simple scoping is provided to narrow the range of annotation types to speed up selection. A proposed extension to the system is to use knowledge from the ontology as to sequentially related annotations, to enable the interface to provide prompts for likely subsequent annotations. If an annotation has been made referring to the blood pressure cuff being put on the patient, subsequent common actions such as inflating the cuff and removing the cuff can be suggested as likely to occur.

Figure 4 contains a snippet of an XML log file generated as part of a real-time annotation session.

More comprehensive detail of the text annotation system, the phases of annotation creation, and replay of the annotations has been covered elsewhere [13].

One of the main problems with such simple annotations system however, is the inability to easily identify specific students and attach annotations to them. The annotation ‘student looks at monitor’, might apply to more than one student visible in the video at that moment in time. A different annotation approach was used to address this issue.

2.2 Automatically derived annotations

Some annotations can be derived directly from the SimMan control system as it logs all activity concerning the mannequin. These will include events initiated by the programmed scenario, for example the scenario may contain a phase where the patient goes into a steady ‘deterioration’. Interactions with the mannequin itself are also recorded (it can detect when an oxygen mask is placed over the mouth).

```
<annotationfile date="2007-06-20 12:46:09Z"
  timestamp="12707973">
<videofile name="Sessions_20-6-07.avi"
  startindex="3804"/>
<annotation id="annotation_12717707"
  type="look" timestamp="10">
  <name>at paperwork charts</name>
  <description>Looking at paperwork charts
  </description>
</annotation>
<annotation id="annotation_8723513"
  type="words" timestamp="407"
  communication="true"
  involves_others="true">
  <name>solicited information</name>
  <description>Solicited information
  </description>
</annotation>
```

Figure 4. A fragment of annotation file (in XML).

The log files that are created, and the scenario files driving the activities, can both be parsed and added to the annotations created manually.



Figure 5. Annotations from SimMan.

2.3 Location annotation

Ubisense² is a commercially available ultra-wideband (UWB) radio frequency real-time location system providing “15cm 3D accuracy, real-time sub-second response and proven robustness in challenging [...] environments”.

As with the existing capture technology in the skills-based lab, the Ubisense sensors were deployed at ceiling height in such a way as to minimise interference with the

²<http://www.ubisense.net/>

‘normal’ running of the hospital ward or distract from student immersion in the scenarios (Fig. 6). The sensors are connected using ethernet cabling – which in a more permanent installation could utilise existing trunks – back to a PC in the control room where the trial was monitored and the location data logged.

The battery powered Ubisense tags are small and light enough to be worn with an acceptably minor impact on the nurses actions and behaviour. In previous deployments we found that tight grouping of people can cause significant degradation to the radio signal; the teaching scenarios meant bunching of students around a patient was almost inevitable, so we attached the tags to the epaulettes on the shoulders of the student uniform (positioning of the sensors at ceiling height also helped minimise interference). Once a session had started it would not be appropriate to interrupt for technical adjustments, so we instrumented two of the students with a tag on both shoulders – both for redundancy, and to expand the data set to enable the exploration of location consistency (two sensors moving together a fixed distance apart) and orientation of the student. Other actors in the scenario – the ward sister and doctor – wore the tags more conventionally using a lanyard around the neck.



Figure 6. Sensors positioned around the room.

Two pieces of mobile equipment in the lab – the dressing trolley and crash trolley – were also tagged, as it was known that these should be moved by the students during the exercise. While the Ubisense sensors must be calibrated before use, this only fixes the position of the sensors in relation to each other, and the tags in relation to the sensors. In addition to measuring the physical layout of the room, we also used tags to ‘trace’ around the ward and objects while the video cameras were recording, ensuring a record of the room layout in relation to the sensor network.

The Ubisense software logs a time-stamped record of updated tag positions; the sensor cell allocates timeslots for



Figure 7. Location tags worn by the participants.

the tags it can ‘see’ based upon their motion/activity up to a maximum of 10 updates per second. It is, of course, essential to record a mapping between the tag IDs that appear in the logs and the tag(s) attached to participants and objects.

With a record of tag location, we can consider how we might relate this to student activity in the scenario and other available annotations:

- if the cameras are set with fixed fields of view for the duration of the session (e.g. one camera pointing at each bed) then it will be known which video stream(s) show a particular student at any one time.
- proximity to an area of the ward with a specific function implies the tagged student is performing a particular task or class of tasks – e.g. washing hands at a sink, using the phone at the nursing station.
- given that the condition of the patient in each bed is pre-planned, the student a text annotation is referring to can be narrowed down by proximity to a particular bed, and therefore activity.
- this can be further narrowed if the patient is a SimMan as there will be a log of the symptoms and senses experienced at that bed.
- movement of mobile equipment indicates at least the expectation of a specific class of activity.
- position around a bed (at the head or foot) and in relation to other participants may imply which student is performing an annotated task.

Since time is the key axis against which we wish to align our annotations and video, having a fixed point common to all media and annotations is necessary for synchronisation

of playback. For this we used the buttons on a Ubisense tag: the button press is registered in the Ubisense log and can be synchronised to a verbal countdown on the audio recording (in the spirit of a clapperboard). For a permanent installation one might expect to synchronise all the capture machines to a common reliable clock source.

The nursing students found the Ubisense tags easy to wear and were able to ignore their presence; similarly the sensors and knowledge of data collection didn't cause any apparent distraction beyond that caused by the cameras and microphones. It helps, of course, that the students are engrossed in the scenario and engaged as if it were 'real life' – the additional technology quickly recedes into the background.

The Ubisense product performed adequately, though not perfectly: on numerous occasions there was some lag before the sensors reported a tag in motion, particularly after the tag had been at rest, or when a nurse moved quickly across the room; in some instances it seemed that a tag had become 'stuck', although no single tag was motionless for the entire session.

While waiting to perform further trials it is not clear whether these discrepancies were caused by problems with the tags, with the particular sensor arrangement and calibration in the ward, or a more general weakness of the technology. A comparative analysis of the tag movement data has not yet been performed, but we expect this to be an interesting test in dealing with incomplete and unreliable data – after all, it is unlikely that sensor data will ever be perfect, and inaccuracies will always be present in human sources such as the observational text annotations.

3 Replaying and analysing the annotations

3.1 Simple replay of the annotations

The textual annotations recorded could be played back to the students as overlay captions on the video (see Figure 8.) More complete descriptions of the textual replay tool and scenarios can be found in [13] and [8].

3.2 The Digital Replay System (DRS)

The Digital Replay System (DRS) is a software tool to support the coordinated replay, annotation and analysis of combinations of video, audio, transcripts, images and system log files [6]. DRS enables time-based data – i.e., system recordings and audio/visual recordings – to be combined and replayed side-by-side and for annotations to be added to create new representations.

Amongst other things, DRS is a re-engineering of the VidGrid ReplayTool to use the Resource Description

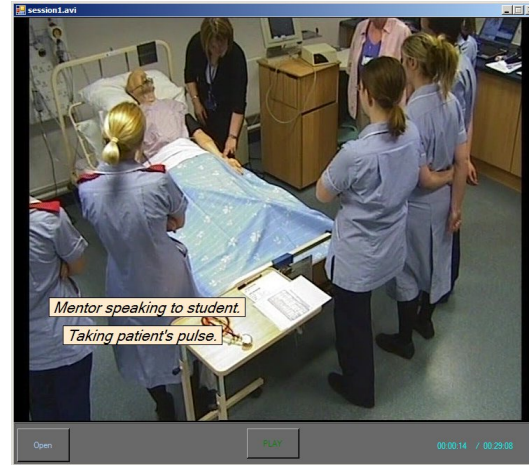


Figure 8. Replaying the video for the students.

Framework (RDF) and Web Ontology Language (OWL) for its core data model [5]. DRS provides:

- Generalised support for project and data management and data overview.
- Complex synchronisation between multiple related media and log files.
- Complex and structured annotations.
- Support for log-file processing, storage and visualisation within the tool-supported environment.

DRS has been extended to view and analyse the data from the nursing trial. Log file importers have been written for the text-based observational annotations and the Ubisense location logs; a data viewer has also been written to visualise the Ubisense logs. These annotation sources are combined in a project with the audio/video recordings from the ward and analyses can then be constructed that allow the user to navigate through the data from different standpoints (Fig. 9).

Having analysed the data set in DRS, future work involves working with the nursing staff and students to create interfaces tailored to their needs, either through specialised DRS data viewers or standalone Semantic Web applications.

4 Conclusions

In our recent work we have trialled the use of location tracking technology to capture and augment semantic annotations in the context of a skills-based teaching lab for

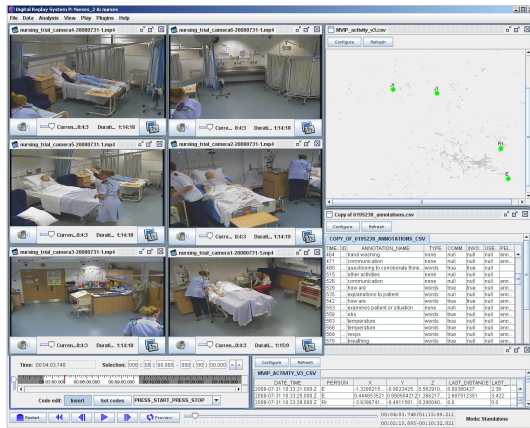


Figure 9. Replaying the scenario in the Digital Replay System.

nurses. After an initial analysis, we believe the location data - albeit imperfect - to be a useful bridge between observational text annotations and the full video record of the session. By using extensible ontologies we expect to also integrate annotations from the SimMan mannequins, and extend capture to other pervasive sensors should they be installed (telemetry from other equipment, light switches, sensors on soap dispensers etc.).

The location tracking data has further potential as a primary data source. When synchronised with the activity data we expect to extract and infer from the combined semantic annotations, it could provide valuable insight into streamlining ward layout. The same activity data could be fed back into the text-based annotation tool on a real-time basis, automatically informing the choice of annotations and further enhancing the process.

We have used digital replay technology to integrate the annotations (both manually authored and automatically gathered) with the video streams for both teaching purposes and for research analysis both of the students and the mentors in the scenarios. It is hoped that as we continue to construct these richly annotated learning environments we can begin to use inference over the semantic annotations in order to speed up the identification of interest points in the data and therefore potentially enable richer forms of data analysis and novel research into the way nursing students learn and the way mentors observe their behaviour.

Acknowledgments

This work was funded by NCeSS grant RES-149-25-1056 and constitutes a case study within EP/C010078/1. This work was also funded by the Engineering and Physical Sciences Research Council “Semantic Media” grant (EP/C010078/1).

References

- [1] T. Blackburn and C. Sadler. The role of human simulators in health-care training. *British Journal of Hospital Medicine*, 64(11):677, 2004.
- [2] M. Fraser, J. Hindmarsh, K. Best, C. Heath, G. Biegell, C. Greenhalgh, and S. Reeves. Remote collaboration over video data: Towards real-time e-social science. *Computer Supported Cooperative Work (CSCW)*, 15(4):257–279, Aug. 2006.
- [3] M. Gobbi, E. Monger, G. Watkinson, A. Spencer, M. Weaver, J. Lathlean, and B. S. Virtual interactive practiceTM: A strategy to enhance learning and competence in health care students. In M. e. a. Fieschi, editor, *Medinfo 2004*, pages 874–878, 2004.
- [4] R. Greenberg, G. Loyd, and G. Wesley. Integrated simulation experiences to enhance clinical education. *Medical Education*, 36:1109–1110, 2002.
- [5] C. Greenhalgh, A. French, J. Humble, and P. Tennant. Engineering a replay application based on rdf and owl. In *Online Proceedings of the UK e-Science All Hands Meeting 2007*, Sept. 2007.
- [6] C. Greenhalgh, A. French, P. Tennant, J. Humble, and A. Crabtree. From replaytool to digital replay system. In *3rd International Conference on e-Social Science*, Oct. 2007.
- [7] R. Hays, H. Davies, J. Beard, L. Caldon, E. Famer, and P. Fincane. Selecting performance assessment methods for experienced physicians. *Medical Education*, 36:910–917, 2002.
- [8] J. W. McDonald, M. O. Gobbi, D. Michaelides, E. Monger, M. J. Weal, and D. De Roure. Grid-enabled data collection and analysis : semantic annotation in skills-based learning. In *4th International Conference on e-Social Science*, June 2008.
- [9] F. Nack and W. Putz. Designing annotation before it’s needed. In *MULTIMEDIA ’01: Proceedings of the ninth ACM international conference on Multimedia*, pages 251–260, New York, NY, USA, 2001. ACM Press.
- [10] T. D. of Health. Delivery 21st century it support for the nhs: national strategic programme. Technical report, The Stationary Office, London, June 2002.
- [11] K. R. Page, D. T. Michaelides, S. J. B. Shum, Y.-H. Chen-Burger, J. Dalton, D. C. De Roure, M. Eisenstadt, S. Potter, N. R. Shadbolt, A. Tate, M. Bachler, and J. Komzak. Collaboration in the semantic grid: a basis for e-learning. *Applied Artificial Intelligence*, 19(9-10):881–904, Nov. 2005.
- [12] M. Seropian, K. Brown, J. S. Gavilanes, and B. Driggers. Simulation: not just a manikin. *Journal of Nursing Education*, 43(4):164–169, 2004.
- [13] M. J. Weal, D. T. Michaelides, D. C. De Roure, M. Gobbi, E. Monger, and J. W. McDonald. Semantic annotation in ubiquitous healthcare skills-based learning environments. In *Workshop on Semantic Web in Ubiquitous Healthcare in conjunction with ISWC2007*, Nov. 2007.