Towards Using Similarity Measure for Automatic Detection of Significant Behaviors from Continuous Data

Ben-Manson Toussaint
Informatics Laboratory of Grenoble
11, rue des Mathématiques
38400 St-Martin d'Hères
+33 4 76 57 48 61
ben-manson.toussaint@imag.fr

Vanda Luengo
Informatics Laboratory of Grenoble
11, rue des Mathématiques
38400 St-Martin d'Hères
+33 4 76 57 47 75
vanda.luengo@imag.fr

Jérôme Tonetti
University Hospital of Grenoble
Boulevard de la Chantourne
38700 La Tronche
+33 4 76 76 66 06
j.tonetti@chu-grenoble.fr

ABSTRACT
This paper presents our method based on similarity measure between contiguous pairs of sequences to yield automatic detection of significant behaviors from raw and continuous traces. The traces, produced by a simulation-based Intelligent Tutoring System dedicated to percutaneous orthopedic surgery, are related to perceptual-gestural behavior and ill-defined tasks involved in this domain. Preliminary qualitative evaluations have been conducted on real data from five simulation sessions and showed the relevancy of our method and adjustments that need to be realized for further experiments.

Keywords
Sequences similarity, Perceptual-gestural behavior, Ill-defined task, Simulation-based ITS, Learners modeling.

1. INTRODUCTION
The learning process of orthopedic surgery is composed of two parts: a theoretical part involving declarative knowledge and a practical part involving perceptual-gestural knowledge related to surgical gestures. This knowledge is qualified as perceptual-gestural because it is tacit and mostly accessible empirically through repeated practices. Tasks related to this knowledge are ill-defined as different strategy patterns can be applied to execute a given operation and no precise way can be defined in advance to satisfy their validation criteria. As demonstrated in [6], there is a gap in the learning process that can hardly be bridged by traditional teaching methods. TELEOS learning environment aims at providing the missed intermediate phase of apprenticeship.

For offering tutoring services in adequacy with perceptual-gestural and ill-defined knowledge, some constraints must be considered like the impossibility to define an exhaustive theoretical framework, the importance of designing an opened knowledge model and the difficulty to assess perceptual-gestural knowledge in the diagnosis process. To overcome these constraints, we want to set up a hybrid approach [2] combining a data-driven paradigm including automatic acquisition of knowledge from traces, with the existing expert-oriented paradigm. The purpose is to keep the knowledge model opened and incremental. To achieve this, we need to capture and model learners' strategies in the execution of simulated operations. That requires on first hand that we foster automatic detection of significant execution behaviors from the continuous raw traces recorded by the simulator.

2. BACKGROUND
The most recent related work reported in the literature is CanadarmTutor, a simulation-based ITS for training astronauts for the handling of an articulated robotic arm [4]. It provides a 3D simulated environment where leaners train in moving the robotic arm from an initial configuration to another predetermined one. As explained in [1] this task is complex and ill-defined.
For offering convenient tutoring services considering these constraints, a hybrid approach combining expert system, model-tracing and automatic acquisition of partial task from experts has been proposed [5]. Like in our case, this work seeks to extract parts of solution paths that are frequently applied to be reused later for supporting key tutoring services. One of the main differences between this work and ours is the importance in our case to link extracted resolution patterns with the phase in which they lie as some actions give different performance insights depending on the phase in which they were executed.

3. METHOD
To capture perceptual-gestural behavior in TELEOS learning environment, we use two complementary devices with the simulator: an eye-tracker for tracing perceptual behavior, that is, points and areas of interest gazed during the execution actions [3] and a haptic arm to capture gestures-related actions executed with the trocar [4]. (Traces from the three tools are recorded independently. They are heterogeneous regarding their content types, their content format and their time granularities. To link each sequence of action to the associated sequences from the complementary devices, we merge their parameters so that each sequence from one source contains the parameters of sequences from the two other sources at the moment it occurred. After this treatment, an action is represented by a subset of sequences that defines its continuum until the next action is executed.)

3.1 Characterizing Significant Behaviors
Our case study is centered on vertebroplasty[1]. This surgical operation is conducted in three phases: the patient preparation, the drawing of the cutaneous marks and the trocar insertion. As opposed to classical open-heart surgery, surgeons are guided all along the operation by X-rays. Validation criteria of each executed action are evaluated by visual analyses of these latter. The first phase of a vertebroplasty is validated if the X-ray appliance (the fluoroscope) is positioned as to generate both face and profile X-rays that render properly the position of the targeted vertebra. In the second phase, the cutaneous marks are validated if their drawing overhangs properly the targeted vertebra on the X-rays. The last phase is validated if the X-rays confirm the correct trajectory of insertion of the trocar.

[1] Vertebroplasty is a percutaneous orthopedic surgery that is practiced to treat fractured spine bones with cement injected with a trocar inserted through small incisions in the skin.

However, validated actions can need to be revised if not executed correctly. These corrections can take place within the same phase or can require that the intern returns to a previous phase. They can be the consequence of different behaviors. For example, the intern may not take enough time to analyze generated X-rays or not enough X-rays to guide his or her actions. On the other hand, another intern can often ask for visual guidance if his or her strategy is to progress with slight and prudent adjustments where another one can ask for very little visual guidance but take more time to analyze each generated X-ray.

Thus, we want to automatically detect, from the simulation traces, phase changes, corrections within the same phase and taken step back to decide on next action or to validate passed action. To achieve this, we need to identify the amplitude of displacements of the simulation environment tools step by step, that is, from one sequence to the next. We need also to identify the elapsed time between these sequences. In fact, similarity between sequences of the same action continuum is supposed to be high and important changes, marked by contiguous sequences with low similarity. The elapsed time between contiguous sequences gives also insights on the learners' behavior as it can point out the time taken for modifying an action, for thinking on the next action to execute or on the validation check of a passed action.

Based on observations of simulation sessions realized by interns at the university hospital of Grenoble, we made the assumption that small temporal gaps coupled with low similarities are more likely to represent phase changes; large temporal gaps coupled with low similarities are more likely to represent corrections within a phase and large temporal gaps coupled with high similarity are more likely to represent step back for deciding on next action or to check the validation criteria of passed action.

3.2 Computing similarity

We used the cosine similarity measure to calculate the similarity between pairs of contiguous sequences. This measure is given by the following formula:

$$\mathrm{sim}(A,B) = \cos(\theta) = \frac{A \cdot B}{||A||\,||B||} = \frac{\sum_{i=1}^{n} A_i \times B_i}{\sqrt{\sum_{i=1}^{n} A_i^2} \times \sqrt{\sum_{i=1}^{n} B_i^2}} \qquad (1)$$

This metric excerpt the proximity of the orientations of the vectors A and B based on the angle θ that they form and consequently, the level of similarity of the sequences that they represent. Its outcomes are bounded in [0,1]: 1, representing a perfect similarity between the elements of the two vectors and 0, a strong dissimilarity between them.

4. EXPERIMENTS

We conducted a preliminary qualitative experimentation based on traces from five simulation sessions of vertebroplasty. Each session was executed by one different intern surgeon at the University Hospital of Grenoble and was screen video recorded. We proceeded to the comparison of behaviors listed from videos with the list of automatically detected behaviors from the traces as to identify accurate, missed and false detections. As reported in Table 1, the automatic detection method demonstrated good performance for the detection of phase changes. Indeed, all of those that it reported were relevant, bringing the precision of detections for this category of behavior to 1.00. However, 60% of these changes were missed. This explains the poor recall score (0.40) for this category of behavior, as for the detections of corrections within phases for which the recall is only of 0.33. High precision (0.91) and recall (0.76) are recorded for the detections of taken step back for the five sessions.

Behaviors \ **F-Score**	**Precision**	**Recall**	**F_1-score**
Corrections within phases	0.60	0.33	0.42
Phase changes	1.00	0.40	0.57
Step back	0.91	0.76	0.83

Table 1. Measures of Precision and Recall of the automatic method compared to observations from video recorded simulation sessions.

5. CONCLUSION

These evaluations, specially the obtained recall scores, reveal the sensitivity of sequences similarity outcomes in presence of other factors that were not considered in this experiment like the level of experience or competence of the interns. The choice of similarity and temporal gap thresholds should be adapted regarding these factors in future evaluations. This work is the first step in achieving more fine-grained diagnosis by integrating in the process learners' simulation execution strategies along with evaluation of their single actions. The planned next step is to yield automatic recognition and categorization of significant behaviors signatures in addition to the mere detection of their occurrences.

6. REFERENCES

[1] Fournier-Viger, P., Nkambou, R., Mephu Nguifo, E. 2010. Building Intelligent Tutoring Systems for Ill-Defined Domains. In *Advances in Intelligent Tutoring Systems*, Springer, Heidelberg. 81–101

[2] Fournier-Viger, P., Nkambou, R., Mayers, A., Mephu Nguifo, E., Faghihi, U. 2011. A Hybrid Expertise Model to Support Tutoring Services in Robotic Arm Manipulations. In *Proceedings of the 10th Mexican International Conference on Artificial Intelligence*. MICAI 2011. LNAI 7094, Springer, 478–489

[3] Jambon, F., Luengo, V. 2012. Analyse oculométrique « on-line » avec zones d'intérêt dynamiques : application aux environnements d'apprentissage sur simulateur. In *Actes de la Conférence Ergo'IHM sur les Nouvelles Interactions, Créativité et Usages*, Biarritz, France

[4] Luengo, V., Larcher, A., Tonetti, J. 2011. Design and implementation of a visual and haptic simulator in a platform for a TEL system in percutaneous orthopedic surgery. In *Medecine Meets Virtual Reality 18*, Westwood J.D., Vestwood, S.W., Ed. MMVR 2011. 324–328

[5] Minh Chieu V., Luengo V., Vadcard L. 2010. Student Modeling in Orthopedic Surgery Training: Exploiting Symbiosis between Temporal Bayesian Networks and Fine-grained Didactical Analysis. In *International Journal of Artificial Intelligence in Education 20.* IJAIED 2010. IOS Press. 269-301

[6] Tonetti, J., Vadcard, L., Girard, P., Dubois, M., Merloz, P. Troccaz, J. 2009. Assessment of a percutaneous iliosacral screw insertion simulator. In *Proceedings of the Conference Clinical Orthopaedics and Related Research.* CoRR 2009. 471-477