Capturing Multi-Perspective Knowledge of Job Activities for Training

Dimoklis Despotakis[1]

School of Computing, University of Leeds, UK
scdd@leeds.ac.uk

Abstract. Using simulated environments for experiential learning gains a growing popularity in professional training. However, simulated context cannot capture the complexity of real world activities, hindering the adaptation to individual learning needs and real world experiences. On the other hand, there is a vast amount of user contributed content about real world activity which represents different viewpoints and contexts. Although this content can be a useful source for enriching the experiential learning experience in simulated environments, it has not been exploited to date. The main limitations are the poor structure and the lack of approaches to retrieve the knowledge nuggets embedded in the existing digital content (e.g. videos or user stories) and to relate them to the simulated context. The aim of this doctoral project is to develop a novel approach for capturing multi-perspective knowledge of job related activities from existing digital records and personal experiences. The proposed approach aims at extracting an advanced context model which will augment digital records of job activities with semantics, and will provide intelligent search to augment simulated context with real life experiences.

Keywords: Activity, Context, Context-awareness, Activity Modelling

1 Motivation and Description of the Project

This PhD is motivated by two recent trends in professional training. Firstly, research and development in simulated activities and virtual environments is becoming widely used for training and has significant influence on future learning technologies. However, the learning experience in a simulated environment is disconnected from real-world job experience, reducing the effectiveness of learning. Simulated environments embed predefined interactive scenarios which include fixed parameters, whereas real world activities are affected by dynamic conditions and complex situations which are hard to capture in simulated world. Hence, there is limited contextual alignment between the simulated experience and the learner's real job activities. The representation of and adaptation to the real world context in simulated environments is a major challenge, which can play a crucial role in affecting the quality of learning, especially in the area of adult training.

[1] Supervisor: Dr. Vania Dimitrova

On the other hand, there is a remarkable amount of resources available over the web, which offer rich content related to job activities and every day practice that can be used as a source for capturing contextual knowledge. Moreover, people tend to share and search for training material and exchange opinions and personal experiences that mirror their real world job context. This abundance of user generated content offers exciting opportunities to capture contextual knowledge about job activities and use it for training. Such contextual knowledge, however, is currently disconnected from the user interaction in simulated environments where it could offer new means of adaptation to individual needs. The key challenge for making this connection is the lack of effective knowledge elicitation mechanisms to derive a structure that enables intelligent content retrieval for experiential learning.

The key challenge addressed in this project is to find a way to link simulated experiences with real world job-related experience records. *Our main goal is to augment digital content related to job activities with multi-perspective contextual information based on real world experiences in order to improve training and enable context-aware intelligent search.*

The pedagogical underpinning for this PhD is based on the Experiential Learning theory [1] especially applicable to job-based training where people learn from experience, both their own and other peoples'. The research aims at finding a way to capture this experience and use it as a knowledge source for training by relating it to individual experiences and context in simulated environments.

This PhD addresses the following research questions:

RQ1: How to capture multi-perspective contextual knowledge embedded in digital content and personal experience? This includes: deciding how to elicit knowledge from human descriptions and comments related to job-related activities; identifying the main actions of a job-related activity and the important aspects associated with an action; identifying what connections may exist between actions; identifying how these actions are connected with different individual experiences and context, i.e. different perspectives; and how to represent multi-perspective contextual knowledge to augment a reflective digital object.

RQ2: How to use contextual knowledge to retrieve digital content related to a specific situation? This includes discovering structural connections between different pieces of contextual knowledge and finding similarities between context models.

To deal with the above questions, we propose a conceptual framework for contextual knowledge capturing and retrieval, which consists of three layers, as shown in Figure 1. The **Acquisition Layer** deals with the development of a model to capture multi-perspective knowledge of job-related activities**.** This utilizes digital records of a specific **Activity** (e.g. job interview videos) and user input (e.g. comments/stories about interviews) to capture personal experiences and descriptions of the activity presented in the digital content. This will enable us to extract a **Context Rich Activity Model (CRAM)** (see Section 3) in the **Modelling Layer**. The **Application Layer** will provide a retrieval mechanism to map contextual representation of digital content to representation of a specific simulated context and suggest relevant digital content.

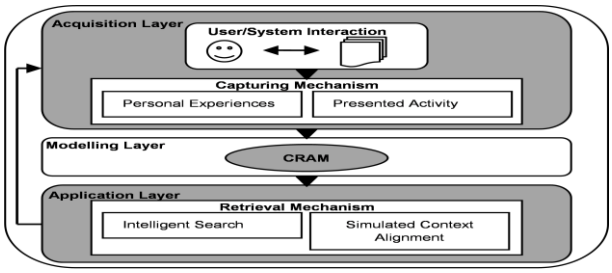

Figure 1 Conceptual framework for capturing multi-perspective knowledge of job-related activities and individual experiences.

2 Background Research and Related Work

This Section outlines the key sources from the background research done so far (this PhD is approaching the end of first year), which built the foundational knowledge for addressing the research questions listed above. The references used here do not intend to give comprehensive description but to show the direction followed to date.

Related to RQ1 an extended literature review on Context, Context-Aware systems [2] and Activity Theory [3-4] has been undertaken to help understand the potential dimensions (aspects) of information that can describe human activity and approaches to model them. To examine knowledge elicitation and representation (i.e. RQ1) the Semantic MediaWiki (SMW) [5] framework was reviewed, which is a technology widely used to facilitate knowledge access and reuse as well as consistency of content. This work will be benefited by semantic technologies to tackle the problem, hoping to conclude with a dynamic ontology to represent Context-Rich Activity. However, the SMW appears to have some technical limitations, which in turn reduce the potential to serve for our purpose. Regarding RQ2 (retrieval of knowledge given particular context), the limitations of SMW include the lack of user profiling mechanisms and the inflexibility of querying algorithms. The potential use of information extraction technologies, e.g. GATE [6], has been examined to derive semantic categories from semi-structured user comments. In addition, Dialogue Agents for knowledge capturing have also been investigated in order to provide structure and guidance for capturing multi-perspective knowledge.

There are several projects that developed approaches which have some similarity with our approach. APOSDLE [7] promotes capturing of job-related experiences into job-related (task) objects, focusing on computer-based tasks. In contrast, our approach considers real world activities which are not performed with a computer (e.g. job interviews, advising students). MATURE [8] aims at capturing organizational knowledge from experiences, considering broad job activities. AWESOME [9] deals with capturing student experiences in academic writing. There is a close similarity of our goal with goals of MATURE and AWESOME. Though, there is a fundamental difference in the methods. Both MATURE and AWESOME consider capturing tags/semantic annotations and accumulating them in a lightweight ontology. Distinctively, we consider more expressive ontologies, look at different activity dimensions (following Activity Theory, see below), and explicitly represent

different individual perspectives. Finally, KP-LAB [10] used records of job-related activities to create pedagogical scenarios for experiential learning where people work in groups and reflect on job activities. The approach of the present project was inspired by the findings in KP-LAB and will take it further by implementing smartness in a system to enable capturing the different aspects of an activity and utilizing this to empower simulated settings. In particular, this work aims to distinguish from KP-LAB in three points: include multi-perceptiveness (i.e. individual knowledge views) in the knowledge base; advance the knowledge elicitation process by implementing methods to provide user-awareness of context and related activities; and provide more expressive models to augment digital content.

3 Research Methodology and Progress to Date

This Section describes the methodology for addressing the research questions and summarizes the progress to date. We will use the term **contextual knowledge** as a holistic description of peoples' job-related experiences. The notion of **experience** refers to the activity that takes place in a real-world job environment. To scope the project we will focus on job interview activities, where a wealth of contextual knowledge can be captured based on peoples' experiences (e.g. university students preparing for jobs or placements; career advisers, interviewers). There is a great amount of digital content about job interviews (e.g. videos with interview examples or personal stories are available in social web sites).

3.1 Capturing Multi-perspective Contextual Knowledge (RQ1)

Contextual knowledge elicitation. Based on a review of existing techniques for contextual knowledge elicitation (such as semantic wikis, information extraction and dialogue agents), we will compose a novel framework for capturing human experiences. To define the type of resources that will be used as records of real job-related activities, we have collected a sample of video files and user comments. Four categories of interview records have been identified: guides (explanations of best practices), interviewees' stories, interviewers' stories and interview examples. We decided to focus on examples and personal stories, as these resources can be closely connected to real world context. For the user input, two interaction dimensions have been identified: users can capture actions in the video; and comment on the actions by providing contextual descriptions and personal experiences. Here we define user as a learner or trainee. Knowledge elicitation mechanism will provide opportunity for reflection and knowledge awareness. A prototype to test the hypothesis of capturing multi-perspective knowledge and start collecting a corpus of resources and user input data has been developed. We will then use information extraction techniques to identify knowledge statements related to key aspects of interview actions and to capture them in an ontology. We also envisage the use of simple dialogue templates to guide the user knowledge acquisition process.

Identifying the main actions of an activity, the important aspects associated with an action and connections that may exist. Due to the complexity of job-related

activities it was decided to exploit Activity Theory to generate the core dimensions related to an activity and the important factors affecting different actions. We utilize Context and relate the activity dimensions to contextual descriptions of actions following the approach in [11]. Another aspect of the problem is to discover relations between activity actions. So far, four types of relations have been considered: time link (e.g. sequence of actions); similarity of actions (e.g. emotional state of the interviewee); connected actions (i.e. derive patterns of annotations from user comments on actions); and connections between different digital content.

Capturing multi-perspectives. We aim to capture the relation between the context presented in a digital resource (i.e. internal descriptions from user comments on the presented activity) and the individuals' perspectives (external descriptions of personal experiences). This will enable us to produce different knowledge views and provide new means of adaption to individual learning needs.

3.2 Representation of multi-perspective contextual Knowledge: The Context Rich Activity Model.

The main constructive component of CRAM is the Context Rich Activity Object (CRAO). A CRAO is a digital object that consists of a semantic multi-perspective knowledge wrapper of a digital record to represent job-related activities and individual experiences. Potential relations between CRAOs will be explored to build a semantic graph of objects and provide a classification framework (Figure 2, left). Figure 2 (right) presents the internal structure of CRAO. We refer to the digital records of real job-related activities as **Activity Resources (AR)**. The Activity is segmented to Actions. Each Action is represented by Engestrom's **Extended Activity Theory** framework [4] and is described by two types of contextual representations: (a) the description of the action as presented in the Activity Resource and (b) the external individual experience representation. Both representations come from the user's comments. The Actions are linked together with various types of relations that we cited in this Section.

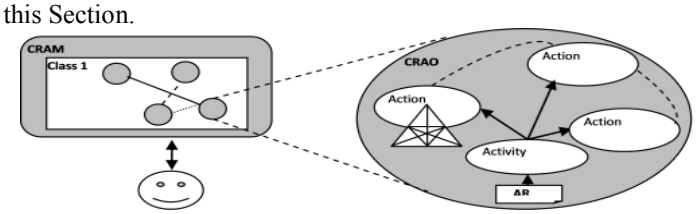

Figure 2 Context Rich Activity Model (CRAM)

3.3 Multi-perspective Contextual Knowledge Retrieval (RQ2) and Evaluation

Example intelligent searches to retrieve knowledge **adapted to particular context** will be conducted. The context dimensions for the queries (i.e. the input fields) will be derived from users-learners and simulated settings. Knowledge retrieval methods will include extensive reasoning on the CRAM ontology and SPARQL queries application.

CRAM will be evaluated in two stages: firstly, the context model to semantically augment the digital record with contextual knowledge; and later, the multi-perspective context model to provide individual knowledge views of the activity. In both stages, CRAM will be evaluated iteratively in collaboration with domain experts, e.g. career advisers. A preliminary evaluation schema for CRAM includes: context dimensions coverage; correctness; and knowledge structure. For the knowledge retrieval process, the functionality of the model will be evaluated not only with individual's context but also with simulated settings to align the simulated experiences with reality aspects.

4 Contribution

The major contribution of this PhD project is the development of a holistic real world activity model that will enable intelligent content search based on personal human-oriented experiences and the adaptation of reality aspects to simulated settings for experiential learning. This process will involve the deployment of a novel system with intelligent services for knowledge elicitation and retrieval and will promote the semantic analysis of real-world activity in ill-defined domains [12], such as interpersonal communications exemplified with job interview activities.

References

1. Kolb, A.Y. and D.A. Kolb, *Experiential Learning Theory: A Dynamic, Holistic Approach to Management Learning, Education and Development*, S.J. Armstrong and C.V. Fukami, Editors. 2009, Sage: London. p. 42-68.
2. Dey, A.K., *Understanding and Using Context.* Personal Ubiquitous Computing, 2001. 5(1): p. 4-7.
3. Bakhurst, D., *Reflections on activity theory.* Educational Review, 2009. 61(2): p. 197 - 210.
4. Engestrom, Y., *Activity theory as a framework for analyzing and redesigning work.* Ergonomics, 2000. 43(7): p. 960 - 974.
5. Krötzsch, M., et al., *Semantic Wikipedia.* Web Semantics: Science, Services and Agents on the World Wide Web, 2007. 5(4): p. 251-261.
6. Cunningham, H., et al. *GATE: A framework and graphical development environment for robust NLP tools and applications*. in *Proceedings of the 40th Annual Meeting of the ACL*. 2002.
7. *APOSDLE*. Available from: http://www.aposdle.tugraz.at/.
8. *MATURE*. Available from: http://mature-ip.eu/.
9. Bajanki, S., et al., *Use of Semantics to Build an Academic Writing Community Environment*, in *Proceeding of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling*. 2009, IOS Press. p. 357-364.
10. *KP- LAB*. Available from: http://www.kp-lab.org/.
11. Kofod-Petersen, A. and J.r. Cassens. *Using Activity Theory to Model Context Awareness*. in *Lecture Notes on Artificial Intelligence*. 2006: Springer.
12. Mitrovic, A. and A. Weerasinghe, *Revisiting Ill-Definedness and the Consequences for ITSs*, in *Proceeding of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling*. 2009, IOS Press. p. 375-382.