A User Modeling Framework for Exploring Creative Problem-Solving Ability

Hao-Chuan WANG[1], Tsai-Yen LI[2], and Chun-Yen CHANG[3]

Institute of Information Science, Academia Sinica, Taiwan[1]
Department of Computer Science, National Chengchi University, Taiwan[2]
Department of Earth Sciences, National Taiwan Normal University, Taiwan[3]
haochuan@iis.sinica.edu.tw, li@nccu.edu.tw, changcy@cc.ntnu.edu.tw

Abstract. This research proposes a user modeling framework which aims to assess and model users' creative problem-solving ability from their self-explained ideas for a specific scenario of problem-solving. The proposed framework, User Problem-Solving Ability Modeler (UPSAM), is mainly designed to accommodate to the needs of studying students' Creative Problem-Solving (CPS) abilities in the area of science education. The use of open-ended essay-question-type instrument and bipartite graph-based modeling technique together provides a potential solution of user model elicitation for CPS. The computational model has several potential applications in educational research and practice, including automated scoring, buggy concepts diagnosis, novel ideas detection, and supporting advanced studies of human creativity.

1. Introduction

Problem-solving has consistently been an attractive topic in psychological and educational research for years. It is still a vital research field nowadays, and its role is believed to be much more important than it used to be, in alignment with the trends of putting stronger emphasis on students' problem-solving process in educational practices.

User Modeling (UM) for problem-solving ability is an alluring and long-going research topic. Previous works in the area of Intelligent Tutoring Systems (ITS) have endeavoured substantially to model problem-solving process for well defined problem contexts, such as planning a solution path in proving mathematical theorems or practicing Newtonian physics exercises [3]. However, we think the classical ITS paradigm *cannot* well describe the process of *divergent* and *convergent thinking* in the human Creative Problem-Solving (CPS) tasks [1][5]. In other words, the classical approach lacks the functionality to support advanced educational research on the topic of CPS.

In this paper, we propose a user modeling framework, named UPSAM (User Problem Solving Ability Modeler), by exploiting *open-ended essay-question-type instrument* and *bipartite graph-based representation* to capture and model the creative perspective of human problem-solving. UPSAM is designed to be flexible and can have several potential advantageous applications, including: 1) offering functionalities to support educational studies on human creativity, such as automated scoring of open-ended instruments for CPS, and 2) detecting students' alternative conception on a particular problem-solving task for enabling meta-cognitive concerns in building adaptive educational systems.

2. UPSAM: User Problem Solving Ability Modeler

A bird's eye view of the UPSAM framework is abstractly depicted in Figure 1. The grey box labelled *Agent* refers to the core software module implemented several functionalities to perform each process of user modeling as described in [4], including:

1) Perceiving the raw data from the user (the process of eliciting user information),
2) Summarizing the raw data as the structured user model (the process of monitoring/modeling), and
3) Making decisions based on the summarized user model (the process of reasoning).

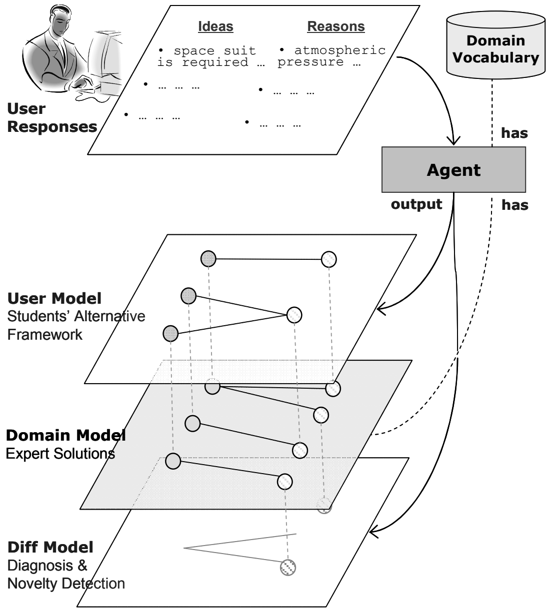
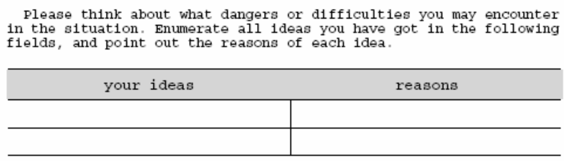


Figure 1. Overview of the UPSAM framework.

Note that the source data for UPSAM are users' free-text responses in natural language toward an open-ended essay-question-type instrument. However, although users' responses are open-ended, they are *not* of no structure by themselves. With the help of a controlled *domain vocabulary* which increases the consistency between users' and the expert's wording, as well as the *pair-wise semi-structured nature* of the instrument which help identify the context of users' answers, it becomes much more tractable to perform the operation of user model summarization from such open-ended answers.


Please think about what dangers or difficulties you may encounter in the situation. Enumerate all ideas you have got in the following fields, and point out the reasons of each idea.

your ideas	reasons

Figure 2. A snapshot of the answer sheet showing the pair-wise relation among ideas and reasons.

Figure 2 depicts the format of the instrument for eliciting user information, which is based on the structure of the CPS test proposed by Wu *et al.* in [5]. Users are required to express their ideas (cf. the production of divergent thinking in CPS) in the problem-solving context described by the instrument, and then explain/validate each idea with reasons (cf. convergent thinking in CPS).

3. Bipartite Graph-based Model

In UPSAM, an important feature to capture users' CPS ability is to structure the domain and user models (see Figure 1) as bipartite graphs. Actually, a domain model is simply a special case of user model summarized from domain experts with a different building process. Domain models are now authored by human experts manually, while user models are built by UPSAM automatically. Therefore, the fundamental formalism of the domain and user models is identical.

One of the most important features in CPS is the relation bewteen divergent thinking and convergent thinking. The bipartite graph in the graph theory is considered appropraite to represent this feature. A bipartite graph is one whose vertex set can be partitioned into two disjoint subsets such that the two ends of each edge are from different subsets [2]. In this case, given a set of ideas $A=\{a_1, a_2, \ldots, a_n\}$ and a set of reasons $B=\{b_1, b_2, \ldots, b_m\}$, the domain model can be represented as an undirected bipartite graph $G=(V, E)$ where $V=A \cup B$ and

$A \cap B = \phi$. The connections between ideas and reasons are represented as $E=\{e_{ij}\}$, and each single edge e_{ij} represents a link between idea a_i and reason b_j.

Different ideas, reasons, and combinations of the *(idea, reason)* pairs should be given different scores indicating the quality of answers. The scoring functions are assigned to *A*, *B*, and *E*, respectively:

$$Sc = \{good\ answer,\ regular,\ no\ credit\},\ f_A : A \to Sc, f_B : B \to Sc, \text{and } f_E : E \to Sc$$

where S_C denotes the range of these scoring functions, and each ordinal value (ex. "*regular*") is connected to a corresponding numeric value. Then the total score of a model $G=(A \cup B, E)$ can be computed as the weighted summarization of individual part of scores:

$$f_{total}(G) = (w_A f_A(A) + w_B f_B(B) + w_E f_E(E)) / (w_A + w_B + w_E)$$

w_A, w_B, and w_E are weighting coefficients that can be tuned according to the needs of each application. Therefore, the score for a user *U* can be reasonably defined as the ratio of the user model's (G_U) total score to the domain model's (G_D) total score. That is, $Score(U)=f_{total}(G_U)/f_{total}(G_D)$. An automated scorer for grading semi-structured responses can then be realized accordingly. Moreover, a fine grained analysis of users' cognitive status is possible by considering the difference between the domain and user models. The *Diff Model* representing the difference is defined as $G_{diff}=(G_U \cup G_D) - (G_U \cap G_D)$. Its properties and applications deserve further exploration.

The process of building the bipartite graph-based user models from users' answers is computationally tenable. The kernel idea is to employ techniques of Information Retrieval (IR) to identify the similarity between users' open-ended entries and the descriptions associated to each vertex in the domain model. As mentioned in Section 2, the incorporation of a controlled vocabulary and the structure of the instrument are considered helpful to the process. A prototypical automated user modeling and scoring system has been implemented, and more details will be reported soon.

4. Conclusion

In this paper, we have briefly described a user modeling framework for CPS ability, UPSAM. Empirical evaluations, full-fledged details, and applications of the framework are our current and future works. We also expect that the computational model can be of contribution to the study of human creativity in the long run.

References

[1] Basadur, M. (1995) Optimal Ideation-Evaluation Ratios. *Creativity Research Journal*, Vol. 8, No. 1, pp.63-75.

[2] Boundy, J., Murty, U.S.R. (1976) *Graph theory with applications*, American Elsevier, New York.

[3] Conati, C., Gertner, A.S., VanLehn, K., and Druzdzel, M.J. (1997) On-Line Student-Modeling for Coached Problem Solving Using Bayesian Network. *Proceedings of 6th International Conference on User Modeling,* Italy.

[4] Kay, J. (2001) User Modeling for Adaptation. *User Interface for All: Concepts, Methods, and Tools*, Lawrence Erlbaum Associates, pp. 271-294.

[5] Wu, C-L., Chang, C-Y. (2002) Exploring the Interrelationship Between Tenth-Graders' Problem-Solving Abilities and Their Prior Knowledge and Reasoning Skills in Earth Science. *Chinese Journal of Science Education,* Vol. 10, No. 2, pp. 135-156.