Research on Automatic Assessment of Transferable Skills

Jinbiao Li and Ming Gu
School of Software, Tsinghua University, Haidian District, Beijing, China

Keywords: Computer Assisted Assessment (CAA), Automatic Assessment (AA), User Behavior Modeling, Soft Sensor, Rule Matching.

Abstract: Automatic Assessment is an important research subject in Computer Assisted Assessment. However, for transferable skills, which have become important talent criteria in the talent standard of modern service industry and higher education, there are few universal and effective automatic assessment methods. In order to improve the efficiency of the assessment of transferable skills, and provide methodology foundation for automatic assessment of transferable skills, this paper combines the automatic assessment methods based on operation result and operation sequence, and proposes an automatic assessment method for transferable skills. This method includes four parts: definition of user behavior model, collection mechanism of user operation sequence, rule matching algorithm, and weighted score summary. In addition, this paper introduces an instantiated application in a virtual simulation environment to evaluate the proposed method.

1 INTRODUCTION

Assessment is a very important link in the teaching process. It can provide an intuitional way for the teachers and students to know the teaching process, and provide feedback for teaching and learning as well. As is known, assessment is a repetitive work, can be defined accurately, and has strong timeliness. Moreover, sometimes people may not be the best valuator, because they may have different understanding of the same subjective item. In this case, Computer Assisted Assessment (CAA) has become a hot topic in the field of computer supported education, because it has advantages such as high efficiency and timely feedback, and it is almost unlimited in users' time and region.

At present, in the research field of CAA, automatic assessment (AA) of personal knowledge and some professional skills has well-developed theories, methods and technologies. Especially in IT skills, there are a large number of relatively mature systems, covering many aspects of the basic IT skills assessment, including programming languages, such as Java (e.g. RoboCode (O'Kelly and Gibson, 2006)), C/C++, VHDL (e.g. CTPracticals (Gutiérrez et al., 2010)), and operating system application skills, such as Linux (e.g. Linuxgym (Solomon et al., 2006)). These systems can be roughly divided into two categories: 1) AA systems for programming competitions and 2) AA systems for (introductory) programming education (Ihantola et al., 2010). And they are playing a great role in promoting the teaching and learning of IT major.

However, with booming and growing rapidly, modern service industry has changed its talent standards, which means the transferable skills have been paid more attention than professional skills gradually. Transferable skills are skills can be applied either or both: (i) across different cognitive domains or subject areas; (ii) across a variety of social, and in particular employment, situations (Bridges, 1993), such as planning capability, teamwork, interpersonal influence, etc. The assessment of transferable skills is also an important issue in talents cultivation and selection of higher education. However, the assessment still mainly uses artificial ways, and lacks suitable automatic methods. There are two primary reasons: the AA of transferable skills needs appropriate simulation environment; there are great differences between the evaluation standards of transferable skills and standards of knowledge or IT skills, which means the former needs to synthesize users' operation information, operation steps, and operation results to make assessment, instead of simply using the result data as the only standard. In order to improve the efficiency of the transferable skills assessment, and provide methodology foundation for AA of

Li J. and Gu M..
Research on Automatic Assessment of Transferable Skills.
DOI: 10.5220/0004410004430448
In *Proceedings of the 5th International Conference on Computer Supported Education* (CSEDU-2013), pages 443-448
ISBN: 978-989-8565-53-2

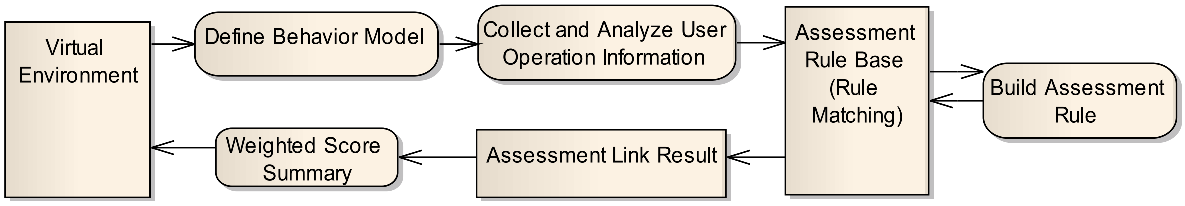

Figure 1: Framework of AA method for transferable skills.

transferable skills, this paper combines the AA methods based on the operation result and operation sequence, and proposes an AA method for the transferable skills. This paper also shows how to apply this method by introducing an application of the method in a virtual simulation environment (SIP).

The AA method of transferable skills is presented in Section 2. The feature of the virtual simulation environment (SIP) and the application of this AA method in SIP are introduced in Section 3. Conclusion and future work are discussed in Section 4.

2 AUTOMATIC ASSESSMENT METHOD OF TRANSFERABLE SKILLS

From point of view of the implementation, AA methods can be roughly divided into two categories: methods based on operation result and methods based on operation sequence, which respectively correspond to the Summative Assessment and Formative Assessment (Harlen and James, 1997). AA methods based on operation result are relatively simple and intuitive, as both the acquisition and assessment of operation result are easy to achieve. Nevertheless, because operation result does not include operation process information of students, the assessment based on operation result is one-sided in a certain degree (Xuan-hua and Ling, 2012). There are few mature AA methods based on operation sequence, and exist following problems: 1) the operation sequence is various in forms and lacks a unified formalization method; 2) user operation is uncertain, so it is difficult to ensure that the operation information collected is valid, which also causes difficulties in analyzing result; 3) the final result of the operation sequence is usually not unique, due to the difference in the timing and repetition of some operations, which makes it difficult to achieve a high accuracy rate in automatic assessment.

In order to achieve a more comprehensive and accurate assessment result, it is necessary to integrate these two methods, that means both the operation result and operation sequence are used as standards for evaluation. Therefore, this paper combined AA methods based on the operation result and operation sequence to propose an automatic assessment method for transferable skills. The method could efficiently solve the problems mentioned in the end of last paragraph. Its process framework is shown in Figure 1, including following four steps: define user behavior model, collect and analyze user operation information, rule matching and weighted score summary.

2.1 Define Behavior Model

In order to describe user behavior in a unified way and make it easier to collect user operation information, this paper defined a user behavior model according to the characteristics of user behavior in simulation environment. This model described the detail information of user operation, including action, operation time, parent node of the current operation, result of the current operation. And in order to make it easily implemented on computer, this model was defined as a four-tuple type E(A,T,P,R):

- A: Action, including operation object and brief event description.
- T: Time, the operation time of the action, including starting and ending time. It is very important to record the operation time, for the assessment of transferable skills is usually related to the completion time.
- P: Parent, the parent event node of the current operation, also called pre-node. There is more than one operation in an assessment link, and they are connected by pre-nodes.
- R: Result, the result caused by the current operation. It can be a piece of result data or change of system state.

This formalized form could record relatively complete user operation information and had good versatility. When applied to certain environment, the

E(A,T,P,R) could be adjusted depending on the simulation environment. Using a serial of associated E, the user operation information would be represented clearly.

2.2 Collect and Analyze User Operation Information

Because the user operation in simulation environment is uncertain, not all user operation information has effective semantic. Traditional methods in user behavior mining (Baglioni et al., 2003) (e.g. web log mining, web usage mining), would gain a lot of useless user operation information, which could not be analyzed before the preprocessing steps such as data cleaning, user identification, session identification and so on. These methods were not suitable for automatic assessment technology, due to high complexity of algorithms. In order to ensure that user operation information collected was effective as far as possible, this paper designed an information collection mechanism, called "Soft Sensor", which meant it was similar to the sensor and could be “inserted” into the simulation environment in right place to collect user operation sequence. “Soft Sensor” was based on the user behavior model E(A,T,P,R), and mainly included three parts: Event Listener, Event Buffer, Event Handler.

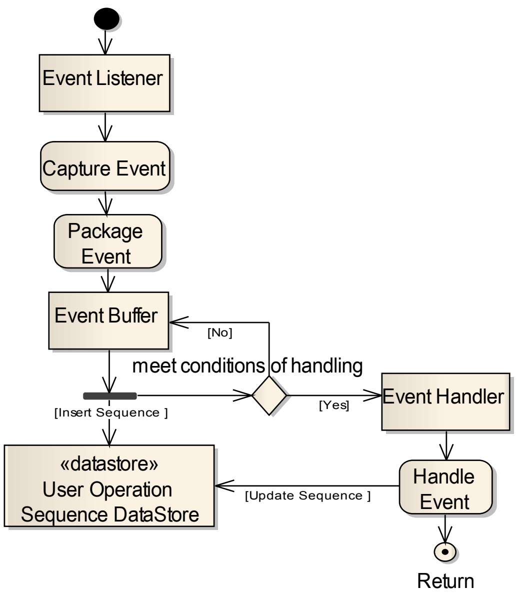

Figure 2: Soft Sensor Workflow

The workflow of “Soft Sensor” is shown as Figure 2. Event Listener was used to capture user operations that were meaningful to transferable skills assessment. The object to which the listener monitored might be a key user operation, system state, or data. When the operation was called or the state/data was changed, the Event Listener would capture this event, and package it into a four-tuple E (A,T,P,R), and sent to the Event Buffer. In the Event Buffer, two things need to be done: insert E(A,T,P,R) into the user operation sequence data store; send those events that meet their processing conditions to the Event Handler, while keeping others waiting in Event Buffer. Depending on the types of operation events, event processing conditions could be defined as three types: immediately, wait for activation (activated by other events), setting time. In Event Handler, the event will return to the original processing operation in system after completing its process flow. Meanwhile, the T(Time) and R(Result) in the event E(A,T,P,R) had probably been changed, so it was necessary to update them in the user operation sequence data store.

2.3 Rule Matching

Due to the difference in the timing and repetition of some operations, the final result of the operation sequence was usually not unique. In order to increase the accuracy of operation sequence assessment, this paper proposed a fuzzy matching algorithm.

An orderly linked list of user operation sequence was generated after the collection and analysis mentioned in Section 2.2. The element of the linked list was the four-tuple E(A,T,P,R), and links between elements were maintained by P(Parent). Before the linked list was processed, a series of standard rules should be established as reference standards for assessment. Inference rules were divided into two kinds, rules based on the operation result and rules based on operation sequence.

Inference rules based on operation result were relative simple: for results of numeral type, corresponding scores could be derived directly compared to reference answers; for results of string type, Levenshtein Distance could be used to calculate evaluation score.

Inference rules based on operation sequence were more complicated. Since operation sequences of different users in the same simulation link might be different in timing or repetition of some operations (e.g. the student failed in a sub-link, and retried several times before success), therefore user operation sequences were diverse. In order to illustrate inference rules based on operation sequence, we defined those indispensable events in one simulation link as "critical node" $E_k(A_k,T,P,R_k)$, and operation sequences made of critical nodes as "critical path", which could also be considered as reference answer. As the proper operation path of one simulation link might be more than one, the corresponding matching rules should be composed of one or more critical paths as well. Fuzzy matching was used to process user operation sequence, which meant using regular expressions to filter out non-critical node event. For example, one critical path was $E_{k1} \rightarrow E_{k2} \rightarrow \cdots \rightarrow E_{kn}$, and its regular form used for fuzzy matching should be $E_{k1} \rightarrow E_x\{0-N\} \rightarrow E_{k2} \rightarrow E_x\{0-N\} \dots \rightarrow E_{kn}$, in which E_x meant non-critical node and N valued depending on the complexity of simulation event.

The rule matching algorithm for operation information is shown in Figure 3.

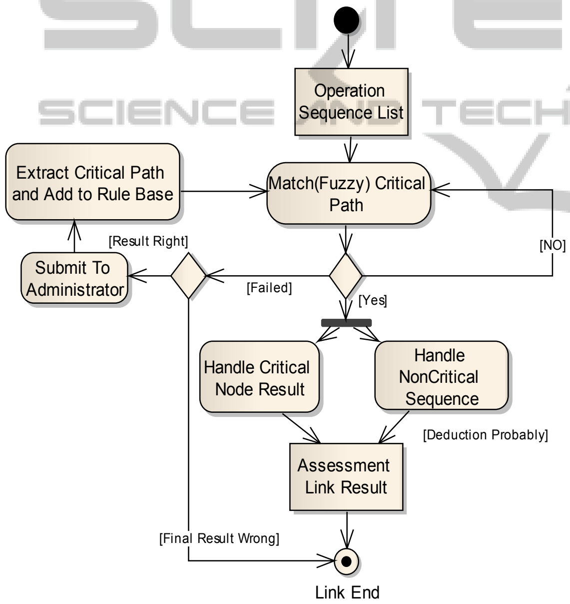

Figure 3: Rule Matching Workflow

1. Use fuzzy matching method to match the user operation sequence list with the critical paths in rule base;
2. Matching successful:
   a) Handle result data of critical nodes using inference rules based on operation result, and send the scores to the assessment link result;
   b) Handle non-critical sequences that are filtered out. Minus scores will be evaluated, according to the time consumption and complexity of the non-critical sequences. That ensures students those spend less time and energy on the non-critical paths to get higher scores.
3. Matching failed:
   a) The result of final critical node is correct:
      i. Submit the user operation sequences to administrator, because it may be correct but does not exist in the current rule base.
      ii. Extract a new critical path from the sequences, and add it into the rule base (by administrator).
      iii. Return to step 1.
   b) The result is not correct, end.

2.4 Weighted Score Summary

It is not comprehensive to evaluate one transferable skill by a single simulation link, because user's transferable skill may be affected by suitability of simulation environment. Therefore, one skill should be assessed through a serial of related situation, and use the summary of weighted score from each assessment situation.

3 APPLICATION OF AA METHOD OF TRANSFERABLE SKILLS IN SIMULATION PLATFORM

In Section 2, this paper proposed a common AA method for transferable skills, and this method need to rely on a virtual simulation environment. This section introduces an existing virtual internship simulation platform (SIP, Service Industry Perception and Virtual Enterprise Practice), and describes how to apply the AA method into SIP and basically achieve the automatic assessment of transferable skills.

3.1 Brief Introduction of SIP

The SIP platform is a CSCL platform based on virtual business environment, and its major objects are senior undergraduate students. Students participate in the SIP in several teams, and complete the following tasks: team building, founding enterprise, virtual business operation and competition, etc. Thus, on the one hand, this virtual internship can deepen students' awareness of the modern service industry; on the other hand, it can help students improve their transferable skills, such as teamwork, planning capability, interpersonal communication skill, etc.

The SIP platform has the following features:

1. It is a multi-disciplinary platform, and has low requirements for students' professional

knowledge, but high requirements for students' transferable skills such as learning ability, planning capability, teamwork and collaboration.

2. It is interesting and attractive to the students, because the SIP is more like a virtual business game than an assignment or test. The students finish their tasks in "non-examination condition", which also means the assessment result has higher reliability.
3. The operation links of the SIP are distinctly separated, and the phased target of each link is also defined clearly.

According to the above characteristics, this paper applied the AA method based on both operation result and operation sequence into the SIP platform, to provide assessment results for the students' transferable.

3.2 AA Method in SIP

This section mainly introduces the implementation of Soft Sensor and Rule Matching in SIP.

The SIP platform was developed using SSI (Spring + Struts + iBATIS), so Soft Sensor was realized by adopting Spring AOP mechanism. The "probe" of Event Listener used the custom java annotation (e.g. @interface SoftSenorListener {String eventID = "";}), thus it could be easily inserted into the code or action that needed to be monitored. A pointcut advisor was implemented to monitor the code which had been appended the annotation "@SoftSensorListener". And a method interceptor was defined to act as the Event Buffer. The Event Handler was a serial of service which implemented the same interface. When the processing conditions were met, the Event Handler would be called by the method interceptor (Event Buffer) using "eventID", which was also the id of corresponding service.

The students play different roles in the SIP platform, and their tasks and transferable skills required in each link are also different. Therefore, this paper chose a typical role – the team leader (also called "CEO" in virtual enterprise) as the assessment object to expound the application of AA method in SIP.

The "CEO" in the SIP platform has a lot of independent operation links, and one of them named "Team Building" is chosen as an assessment example.

Team building, refers to the process that CEOs simulate personnel recruitment of their enterprises, including creating department and position (authority distribution included) as E1, releasing recruitment notice (according to position) as E2, receiving resumes as E3, screening resumes as E4, providing offers as E5, receiving applicant feedback as E6, completing recruitment as E7. The major transferable skills tested in this link are planning capability ("can the CEO organize his/her enterprise and make plan well?") and interpersonal influence ("can the CEO attract other to join his/her enterprise, and is the CEO popular in class?"). Make Ek0 the start event of team building, thus the most perfect critical path is shown as (1), which means CEO smoothly completes team in the shortest way.

$$E_{k0} \rightarrow E_{k1} \rightarrow E_{k2} \rightarrow E_{k4} \rightarrow E_{k5} \rightarrow E_{k6} \rightarrow E_{k7} \quad (1)$$

But in fact, in most cases, it is difficult for a CEO to finish the task in such ideal way like (1). Situations may occur during the team building: after releasing the recruitment notice, CEO realizes that the position setting is unreasonable so that he/she has to adjust the positions and accordingly release the recruitment notice again, as in (2); when one round recruitment is finished, there are some positions still vacant so CEO need to repeat the recruitment again.

$$E_{k0} \rightarrow E_{k1} \rightarrow E_{k2} \rightarrow E_{k1}' \rightarrow E_{k2}' \rightarrow E_{k4} \rightarrow E_{k5} \rightarrow E_{k6} \rightarrow E_{k7} \quad (2)$$

Under the circumstances, to ensure that a critical node in the linked list is latest, it is necessary to search the last critical node of the same event and update it to a non-critical node, when a new critical node is inserted into the list. At this point, Equation (2) should be changed to Equation (3).

$$E_{k0} \rightarrow E_1 \rightarrow E_1 \rightarrow E_{k1} \rightarrow E_{k2} \rightarrow E_{k4} \rightarrow E_{k5} \rightarrow E_{k6} \rightarrow E_{k7} \quad (3)$$

Using fuzzy matching strategy to match the operation information list of CEO with the critical path, as in (1), could determine whether the CEO had completed the entire team building process, then judged the complexity of the non-critical sub-paths, including time spent and repeat times. If CEO had finished this link, the "length" of his/her operation list was shorter, the better score he/she would get in planning capability and interpersonal influence. Conversely, if the complexity of non-critical paths was high, that meant the CEO was under performance of these two abilities. In addition, result data was also used as assessment standard, such as the quantity of resumes received by CEO, the number of positions recruited successfully in the same round recruitment, and regular data like clicks and glance time of the notice, etc.

The assessment of CEOs' planning capability and interpersonal influence reflected in team building was achieved using the AA method. But it was just part of the overall assessment of these two abilities, with summary of weighted assessment of each link, we could acquire relatively objective and comprehensive assessment data.

4 CONCLUSIONS & FUTURE WORK

This paper integrates the advantages of the AA methods based on operation result and operation sequence, and put forward a new AA method for the transferable skills.

The method proposes a standardized way for user behavior in simulation environment, and has greatly improved the efficiency in collecting valid user operation information. The method uses both operation result and operation sequence as assessment criteria so that it can achieve more comprehensive and objective assessment results.

Moreover, this AA method has good versatility, and is flexible to make appropriate adjustments according to requirements of the automatic assessment environment. This paper describes an instantiated application of the method in a simulation platform, and realizes simple automatic assessment of transaction skills. Although the method may not be mature enough, it provides methodology foundation for the automatic assessment technology of transferable skills.

In future work, this method could be further improved in the following ways: the perfection of rule base mentioned in Section 2.3 is artificial, so we will introduce machine learning strategy into the method to achieve the automatic improvement of the rule base; we will also make further application of the method to more assessment environment, and collect application feedback and result data to improve this AA framework and method.

REFERENCES

O'kelly, J. & Gibson, J. P., 2006. RoboCode & problem-based learning: a non-prescriptive approach to teaching programming. *ACM SIGCSE Bulletin*, 217-221.

Gutiérrez, E., Rrenas, M. A., Ramos, J., Corbera, F. & Romero, S., 2010. A new Moodle module supporting automatic verification of VHDL-based assignments. *Computers & Education*, 54, 562-577.

Solomon, A., Santamaria, D. & Lister, R., 2006. Automated testing of unix command-line and scripting skills. In *ITHET'06, 7th International Conference on Information Technology Based Higher Education and Training*, IEEE, 120-125.

Ihantola, P., Ahoniemi, T., Karavirta, V. & Seppälä, O., 2010. Review of recent systems for automatic assessment of programming assignments. In *Koli Calling'11, 10th Koli Calling International Conference on Computing Education Research*, 86-93.

Bridges, D., 1993. Transferable skills: a philosophical perspective. *Studies in Higher Education*, 18, 43-51.

Harlen, W. & James, M., 1997. Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education*, 4, 365-379.

Xuan-hua, C. & Ling, Y., 2012. Research on key technology of automatic evaluation facing simulation system. *Computer and Modernization*, 59-61.

Baglioni, M., Ferrara, U., Romei, A., Ruggieri, S. & Turini, F., 2003. Preprocessing and mining web log data for web personalization. In *AI*IA'03, 8th Advances in Artificial Intelligenc*e, 237-249.