

- [Contents](#) |
- [Author index](#) |
- [Subject index](#) |
- [Search](#) |
- [Home](#)

Reinforcement learning in information searching

[Yonghua Cen](#)

Department of Information Management, School of Economics and Management, Nanjing University of Science and Technology, Jiangsu, 210094, China *and* Advanced Analytics Institute, University of Technology Sydney, PO Box 123, Broadway, NSW, 2007, Australia

[Liren Gan](#) and [Chen Bai](#)

Department of Information Management, School of Economics and Management, Nanjing University of Science and Technology, Jiangsu, 210094, China

Abstract

Introduction. The study seeks to answer two questions: How do university students learn to use correct strategies to conduct scholarly information searches without instructions? and, What are the differences in learning mechanisms between users at different cognitive levels?

Method. Two groups of users, thirteen first year undergraduate students (freshmen) and thirty-four final year undergraduate students (seniors), were recruited into our experimental study and executed ten different search tasks independently. Five reinforcement learning models were introduced to quantitatively simulate the micro process of users' self-regulated learning of search expertise by trial and error.

Analysis. The experimental data were divided into two parts. The first 70% of the data was used to estimate the parameters of each model. The remaining 30% was fitted by the estimated models. The model best fitting the data of users in each group was used to explain their learning behaviour.

Results. Most undergraduates tended to repeat the strategies that brought success in their earlier experiences. Freshmen's learning behaviour manifested remarkable Markov properties. Their strategy selection was always made according to the feedback obtained in the last search activity. Seniors' strategy adjustment depended on the accumulated effect of past strategy adoptions. They displayed strong characteristics of rational thinking.

Conclusions. In the process of learning searching expertise, users demonstrate reinforcement characteristics. Moreover, users at different cognitive levels exhibit different reinforcement patterns. Theoretical and practical implications were proposed from the perspectives of

training programme design, adaptive information retrieval system design and information behaviour model development.

CHANGE FONT

Introduction

The present study is designed to investigate how university students learn to use the search functions provided by scholarly databases and adjust their searching strategies without instructions. The focus of this research is on learning of information searching skills in practice. Here information searching means '*a potential sub-stage in the information-seeking process*' ([Wilson 1999](#): 258) and '*the micro-level of behaviour employed by the searcher in interacting with information systems*' ([Wilson 2000](#): 49).

Since Belkin in the [1980s](#), information science has attempted to bring its information seeking perspective into information search (e.g., [Bates 2002](#); [Belkin et al. 1996](#); [Ingwersen 1996](#); [Saracevic 1996](#); [Sutcliffe et al. 2000](#); [Spink 1997](#); [Spink et al. 2002a](#); [Spink et al. 2002b](#); [Wilson 1999](#), [Wilson 2000](#); [Wilson et al. 2002](#)). However, pertinent literature bridging learning and information searching sheds more lights on *learning of knowledge by searching* ([Colvin and Keene 2004](#); [Ford et al. 2003](#); [Laxman 2010](#); [Marchionini 2006](#); [Puustinen and Rouet 2009](#); [Zhu et al. 2011](#)), rather than *learning of searching by practice*. By *learning of knowledge by searching*, we mean that users acquire knowledge for sense-making or problem-solving purposes through information searching, while *learning of searching by practice* refers to that users improve the level of search skills through practising searches. Several longitudinal studies ([Chu and Law 2007](#); [Vakkari 2001](#); [Warwick et al. 2009](#)) examined users' experiences of academic information seeking and the development of their search expertise. However, few attempts have been made to disclose the behavioural evolution and cognitive dynamics during users' self-regulated learning of search skills by trial and error.

In this study, it is assumed that there is an autonomous reinforcement learning process during academic users' information searching and that different users demonstrate different reinforcement patterns of learning. Specifically, mathematical reinforcement learning models are brought in to fit the data from user experiments. By model fitting and analysis, this study aims to discover the characteristics of users' searching behaviour, and the learning mechanisms controlling users' adjustments of search strategies.

The rest of the paper proceeds as follows: related research is reviewed; then the research questions and assumptions are proposed, followed by the description of the quasi-experiment approach employed in this study, and the process of model estimation and validation; the results and discussion are presented and concluded afterwards; finally, implications and further research are discussed.

Literature review

Learning in information searching

Learning of searching by practice versus learning of knowledge by searching

As commented by Jansen *et al.* ([2009](#)), in many studies concerning information seeking behaviour, the learning aspect is assimilated into other frameworks, such as sense-making and problem-solving ([Brand-Gruwel et al. 2009](#); [Eisenberg and Berkowitz 1990](#); [Kuhlthau 1993](#); [Savolainen 1993](#)). Most of the research linking information seeking or searching with learning emphasizes *learning of knowledge by searching* ([Colvin and](#)

[Keene 2004](#); [Ford et al. 2003](#); [Laxman 2010](#); [Marchionini 2006](#); [Puustinen and Rouet 2009](#); [Zhu et al. 2011](#)), rather than *learning of searching by practice*, although there are commonalities between these two kinds of learning process. Nevertheless, from previous studies, learning as a means to develop searching skills can still be found. For example, some studies underline users' learning and understanding of search tasks, information needs ([Cole et al. 2007](#); [Kelly and Fu 2007](#)), and search strategies ([Halttunen 2003](#); [He et al. 2008](#); [Saito and Miwa 2007](#)). Cole et al. (2007) conducted a field study to examine how domain novices learned to represent the topic spaces of the search tasks. Kelly and Fu (2007) employed online elicitation forms to collect users' descriptions of the search topics. The forms were distributed to users in later experiments, and significantly helped them formulate better queries. Saito and Miwa (2007) carried out controlled experiments to evaluate the educational potentials of a deliberately constructed search-process feedback system in facilitating reflective activities for online searching. Their findings confirm that the performance of the participants supported by the feedback system improved substantially. He et al. (2008) examined the effects of two different training approaches, referred to as conceptual description and search practice, on users' learning and understanding of using a case-based reasoning retrieval system. Halttunen (2003) investigated students' interpretations of information retrieval know-how and summarized the principles of designing constructive learning environments for information retrieval.

Learning process

Studies regarding the process of users' learning of searching expertise can be classified into two categories: self-regulated learning ([Jansen et al. 2009](#); [Kuhlthau 1993](#); [Xie 2000](#); [Xie 2007](#)) and instruction-assisted learning ([Cole et al. 2007](#); [Gerjets and Hellenthal-Schorr 2008](#); [Halttunen 2003](#); [Kelly and Fu 2007](#); [Kuhlthau et al. 2007](#); [Saito and Miwa 2007](#)). Besides, in the view that learning of searching expertise is a dynamic process, researchers ([Chu and Law 2007](#); [Vakkari 2001](#); [Warwick et al. 2009](#)) conducted longitudinal investigations to track the change of users' searching expertise over time.

By self-regulated learning, we mean that users finish searching all by themselves, without guides from others or systems. Grounded in the constructivist view of learning, Kuhlthau (1993) presented a six-stage model of information search process: initiation, selection, exploration, formulation, collection and presentation. The whole process involves 'the total person incorporating thinking, feeling and acting in the dynamic process of learning' (Kuhlthau 1993: 348), in which users move from uncertainty to understanding. Xie (2000; 2007) investigated how the interplay between plans and situations lead to users' shifts of strategies and interactive intentions within an information seeking session. The twofold shifts in Xie's study are essentially the results of users' self-assisted reflective learning. However, in existing research, the learning process in information seeking is aimed at problem solving, rather than search skill acquiring. In response to this, Jansen et al. (2009: 643) called for a learning theory, which 'may better describe the information searching process than more commonly used paradigms of decision making or problem solving'. Their research indicates that different learning levels relate to particular searching characteristics. The results partially support that searching episodes are learning events.

In recent years, instruction-assisted learning including social learning and training-based learning has been much stressed and the influences of external intervention on users' learning extensively analysed. Kuhlthau et al. (2007) elaborated guided inquiry as 'a dynamic, innovative way of developing information literacy'. Cole et al. (2007) claimed that instructive intervention helps novices bridge the gap between their mental models and the thesaurus's hierarchical syndetic representation of the search topic. According to studies of Kelly and Fu (2007) and Saito and Miwa (2007), when provided with analogous information, such as keyword description of similar search topics and information about other participants' search process, participants greatly improve their search effectiveness. Halttunen (2003) maintained that information retrieval instruction should be integrated with constructive learning. Attempting to design constructive learning environments, Halttunen summarized five different aspects of participants' interpretations of information retrieval, and examined their relationship with learning styles and academic backgrounds. Gerjets and Hellenthal-Schorr (2008) proposed a user-oriented

Web training based on a conceptual decomposition of the sub-competencies of media literacy and the sub-processes of information retrieval, and a task analysis of information problems. Their study shows this training approach is more beneficial to develop high school students' declarative knowledge of the Web and facilitate their searching, than conventional technique-oriented trainings.

Taking a long-term view, Chu and Law (2007) investigated twelve postgraduate students' growing understanding of searching skills over a one-year period. They collected data from surveys, interviews, students' search statements and think-aloud protocols. Their findings reveal that, in the beginning, students conducted more questionable subject searches, with little attention paid to keyword searching; later, as they learn more about the capabilities of keyword searching, they prefer keyword searching to subject searching, and at the same time they proceed from simple keyword searches to more complex keyword searches. Vakkari (2001) observed eleven master's students' information searching processes during a period of four months when they were preparing their research proposals. The research corroborates that students' exhibited searching characteristics (including information needs, search tactics, term choices, relevance evaluation and use of obtained information) which correlate highly to their problem-solving stage and their mental model. Based on a two-year investigation of the growth of information seeking skills in a group of undergraduate students, Warwick *et al.* (2009) found that the demands of students' undertakings act as the major factor leading to the progress of their information seeking; students follow the law of minimum effort to retain established information-seeking strategies or seek new methods. Whereas studies by Chu and Law (2007) and Vakkari (2001) provide little evidence on how users acquire the knowledge, research done by Warwick *et al.* (2009) draws a more detailed picture of users' development of searching expertise.

Influencing factors

Besides measuring the impacts of external instructions, the majority of previous work concerning learning behaviour in information searching highlights the influences on users' learning process of users' personalities (including individual experience, knowledge, cognitive style, learning style, and so on) (Bilal and Kirby 2002; Jansen *et al.* 2009; Tabatabai and Shore 2005; Tenopir *et al.* 2008; Thatcher 2008; Wildemuth 2004; Zhang 2008), task complexity (which is associated with users' familiarity with the search task) (Jansen *et al.* 2009; Kim 2002; Zhang 2008) or system characteristics (Wilson *et al.* 2009).

For instance, using comparative studies, Bilal and Kirby (2002), Tabatabai and Shore (2005) and Thatcher (2008) reported that users with different knowledge backgrounds or cognitive capacity (such as novices and experts, children and adults) exhibit different behavioural characteristics in information searching. Wildemuth (2004) conjectured that domain knowledge affects the adjustments of search tactics: insufficient domain knowledge is accompanied with awkward concept representations and erroneous reformulations of search patterns. Zhang (2008) explored the effects of mental models on undergraduate students' online searching. The researcher concluded from experimental studies that students' familiarity with the task significantly influences their ways to initiate interaction, query constructions, and search tactics. Recently, Jansen *et al.* (2009) examined the learning characteristics of users with different cognitive levels in completing search tasks of different complexities. Their study substantiates the differences in exhibited searching characteristics among users of different learning styles. Tenopir *et al.* (2008) examined the affective and cognitive dimensions of searching behaviour and included learning styles as an influencing factor. They administered 41 participants into experiments and used audio/video devices to capture and record their interactions with ScienceDirect. The researchers reported the associations between engineering graduate students' learning styles (converging vs. assimilating) and the characteristics of their search sessions. Kim (2002) confirmed that cognitive style (field dependence vs. field independence), search experience (novice vs. experienced searchers), and task type (known-item vs. subject search tasks) are variables impacting users' search performance and navigational style on the Web. Wilson *et al.* (2009) quantified the strengths and weaknesses of three advanced search interfaces in scaffolding user-system interactions by integrating existing research models of users, needs, and behaviour.

In summary, prior research has attempted to connect information searching with learning; however, limited efforts have been made to model the underlying process of users' learning in information searching. This is preliminarily examined in our work.

Reinforcement learning models

Humans share with other animals a simple way of learning, which is usually called reinforcement learning. This reinforcement learning seems to be biologically inherent. If an action leads to a disadvantageous outcome (also refers to a negative *payoff* or punishment), this action will be avoided in the future; otherwise, if an action leads to a favourable outcome (a positive *payoff* or reward), it will reoccur ([Brenner 2006](#); [Sutton and Barto 1998](#)). Here, the word *action* can also be understood as *strategy*.

In the spirit of reinforcement learning, a variety of reinforcement learning models have been established in psychology, economics and computer science to quantitatively analyse different learning behaviour in different contexts ([Börgers and Sarin 2000](#); [Bush and Mosteller 1953](#); [Cross 1973](#); [Erev and Roth 1996](#); [Fu and Anderson 2006](#); [Izquierdo et al. 2007](#); [Roth and Erev 1995](#); [Shimokawa et al. 2009](#)). Among them, Bush and Mosteller's model ([Bush and Mosteller 1953](#)), Cross's model ([Cross 1973](#)), Börgers and Sarin's model ([Börgers and Sarin 2000](#)) and Roth and Erev's two models ([Roth and Erev 1995](#); [Erev and Roth 1996](#)) can be regarded as the five most typical ones, and are employed to fit the experimental data in our study. These models are briefly compared in Table 1. More detailed mathematical descriptions regarding these models can be found in the [Appendix](#).

Table 1: Comparison of the five typical reinforcement learning models

Model	Mechanism by which payoffs affect strategy adjustments	Measure of the extent to which payoffs affect strategy adjustments	Basic ideas
Bush and Mosteller's model		A fixed constant	When a certain strategy leads to a positive payoff, the probability of this strategy being chosen again increases and the probability of it being avoided decreases. Otherwise, the probability of the strategy being further adopted decreases and the probability of it being avoided increases.
Börgers and Sarin's model	Payoff of the last strategy adoption	Difference between the actual payoff and the expected one	If the actual payoff of a strategy exceeds the expectation, the probability of this strategy being further selected increases; if the payoff is smaller than the expectation, the probability of the strategy being further adopted decreases.
Cross's model		A monotonic function of the payoff	The attraction of a strategy is defined as a linear function of the payoff, by configuring the reinforcement strength as a variable correlated to the payoff.
Roth and Erev's model		Accumulated payoff from adopting a strategy	Decision makers choose a strategy based on their experiential expectations for all strategies. These expectations result from the accumulated effect of their past strategy adoptions, not only the last one.
	Accumulated effects of all	Accumulated payoff from adopting a	A forgetting parameter is incorporated into the basic model of Roth and Erev to measure the attenuation degree of users'

Roth and Erev's modified model	the previous strategy adoptions	strategy (taking forgetting, subjective cognition and neighbour strategies into account)	experiences influencing their strategy selections. A transferring parameter is added to determine the extent of the reinforcement strength being transferred to the unemployed strategies. At the same time, different individuals make different subjective evaluations to a strategy even when the payoffs from applying the strategy are equal.
--------------------------------	---------------------------------	--	--

The process of information searching is also a process of decision-making or action-taking ([Du and Spink 2011](#); [Kuhlthau 1993](#); [Savolainen 1993](#)). Users exhibit similar reinforcement learning characteristics in this process. Reinforcement learning models can be adopted or revised to disclose the mechanisms dominating users' learning of searching knowledge. This is further studied in our research.

Research questions and assumptions

Research questions

The focus of this study is on *learning of searching by practice*, instead of *learning of knowledge by searching*. It also concerns the effects of personal traits (e.g., information seeking experience and academic backgrounds) on users' learning of search strategies in information searching. However, it is not to provide evidence for or against these effects by qualitative or quantitative analysis of data gathered from experiments, questionnaires, interviews or observations. Rather, this study brings in several reinforcement learning models to examine the micro process of users' self-regulated learning of search expertise by trial and error. It aims at mining the mechanisms underlying users' behaviour adjustments and discovering their learning characteristics and cognitive dynamics during information searching.

The specific research questions are as follows:

1. How do university students learn to use correct strategies to conduct scholarly information searches without instructions? In other words, are there learning rules controlling their strategy adjustments during searching? If so what are the rules?
2. What are the differences in learning mechanisms between users at different cognitive levels?

Assumptions

The research question design, experiment design, model application and explanation in this study are founded on the following assumptions:

(1) In the process of self-regulated learning of searching expertise, users demonstrate reinforcement characteristics.

When a user completes a search task by a certain strategy, the user may evaluate this process in terms of time cost, quantity of relevant results, and so on. Depending on this evaluation, the user will form a tendency to retain this strategy or reject it by switching to other strategies for next tasks. In other words, users adjust their behaviour by referring to their experience in database using and based on their knowledge about the available strategies. This process of dynamic alignment tallies with the core conception of reinforcement learning ([Sutton and Barto 1998](#)). Figure 1 describes this process of strategy reformulation.

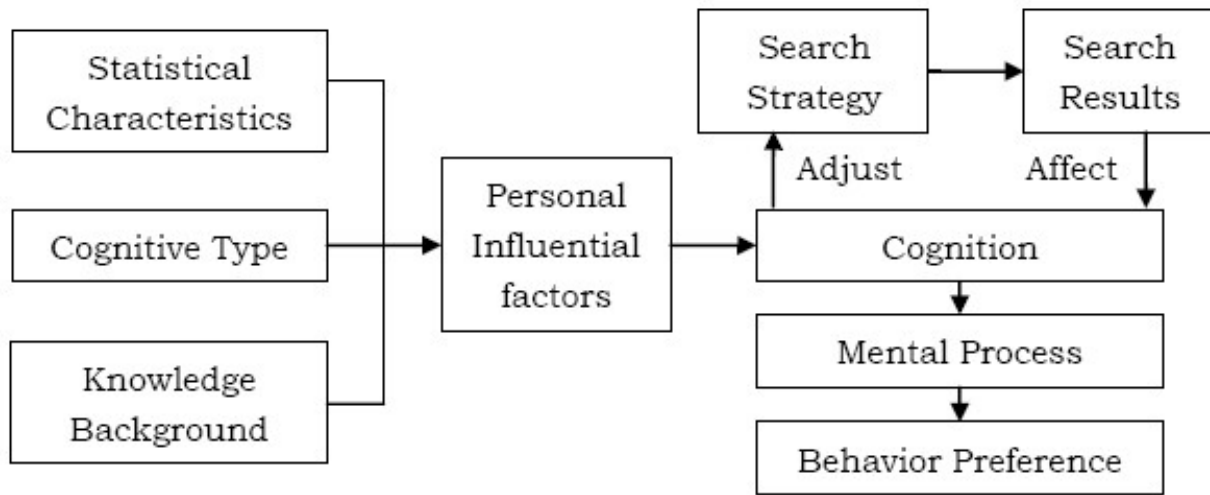


Figure 1: Reinforcement learning mechanism in search strategy formulation

The above process of reinforcement learning and search strategy adjustments is also consistent with the information search process proposed by Ellis (1989) and Wilson (1997), in which a user first defines information needs, and then formulates or selects a search strategy, performs searching or browsing, obtains and evaluates the search results.

(2) Users at different cognitive levels demonstrate different reinforcement patterns.

It is assumed that users' personal traits have impacts on their information behaviour, and there are differences in the reinforcement characteristics between different users during their learning of searching expertise. This assumption is justified in the present study by introducing different reinforcement learning models to fit the experimental data collected from different user groups, and evaluating the applicability of the models to the data.

Research design

Overview

A quasi-experiment approach was designed according to the requirements of data analysis and model inference and fitting. Two groups of undergraduates at different cognitive levels participated in the experiments in January 2009. They were asked to execute set search tasks in a specified academic database system independently. The process of their strategy adjustments by trial and error was observed and recorded by questionnaires and a screen-tracking software. The gathered experimental data were quantitatively fitted by different reinforcement learning models. The fitness of the models to the data was checked and the best model to explain the learning behaviour of users in each group was chosen. By doing this, the dynamic learning mechanisms behind users' explicit strategy formulations were analysed and the differences in learning characteristics between different user groups were examined.

Participants

In the first experiment, thirteen first-year undergraduate students (freshmen) who had little knowledge of academic information searching were organised into our laboratory, while in the second experiment, thirty-four fourth-year undergraduate students (seniors) who did have experience of academic information seeking were administered together. All students had experience of using [Google](#) or [Baidu](#) (a well-known local search engine in China).

It is supposed that there are discrepancies in the level of cognitive processing between freshmen and senior

students, considering the differences in their information seeking experience, knowledge and capability of comprehension, application, analysis, synthesis and evaluation ([Bloom *et al.* 1956](#)). The cognitive level of participants is the independent variable in this study. It is assumed to affect the dependent variable, i.e., users' reinforcement learning behaviour.

Experiment settings

All participants were required to log in the search page of [CNKI](#), a well-known scholarly database system in China, and perform ten different search tasks without extra instructions.

The same search tasks were assigned to all participants. These tasks were designed before the experiments by the researchers. The tasks relate to different subjects. A task form giving descriptions for each task was handed out to participants before they started the tasks. The descriptions include the task title and several keywords associated with the task topic, which removed the chance participants would misunderstand the task.

For each task, the researchers had done a test search in the database system beforehand, and labelled all the relevant search results. These results served as standard ones. Once participants finished a task, the standard results were presented to them to check the correctness of their search performances.

A questionnaire was devised to solicit the perceptions of a participant with regard to the formulated search strategy for each task. The perceptions include:

1. The description of the search strategy, including the search function, the keywords, the way the keywords were input, and additional details;
2. The participant's expectation of the strategy bringing desired results;
3. The satisfaction of the participant with the strategy after applying it and comparing the results with the standard ones.

An incentive mechanism was designed to avoid the possible insufficiency of users' motivation to complete the tasks: those who got better search results would be rewarded with delicate and attractive presents.

Besides, participants were told by the researchers that for each task:

1. All keywords that represent the task topic must occur in each title of the search results. To this end, participants must learn to use multiple search boxes and logical AND connector, so that they could input each keyword in each box and formulate a correct query to fulfil the task.
2. Search results totally consistent with the standard results would be considered satisfactory, and presents would be awarded to those who reached the satisfactory results.

The participants' interactions with the database system were recorded by a screen-tracking software to provide extra information for data analysis.

The above experimental design provides a quasi-experiment approach. The variables such as experimental environments, search tasks, information need understanding, and external stimulations were controlled to be consistent between each participant. As for information need understanding, it was not necessary for participants to figure out what keywords should be used for each task, since standard keywords were offered in the task form. With respect to external stimulations, there was no instruction supplied to participants, and the same incentive mechanism was applied to each of them.

By controlling the above interventions, the effects of factors other than participants' cognitive levels were excluded from the experiments to the maximum extent practicable, and therefore the process of participants' strategy adjustments in performing the search tasks could be more accurately observed.

Search strategies

In relation to search strategies, Bates ([1979](#)) defined twenty-nine tactics in four categories: monitoring, file structure, search formulation and term. In Bates's model, search formulation tactics are the moves that searchers make to design or redesign search formulation, while term tactics are the actions searchers take in selecting and revising terms within the search formulation. Likewise, Belkin *et al.* ([1996](#)) proposed a classification scheme of search strategies. In Belkin's taxonomy, strategies encompass term strategies, database strategies, interaction strategies, and search strategies. Search strategies or tactics in these studies are conceptualised to describe the possible actions a user can take from initiating a search task to concluding it.

In the present research, a search strategy refers to the action that a participant takes to carry out a search task, by selecting one of the search functions offered by the search system and formulating a search query. The optional search functions include the basic search, the advanced search and the expert search. To facilitate model inference and fitting, the search strategies that a participant could apply to construct a query were categorised into three types:

1. The first type, the *simple-search strategy*, refers to when a participant inputs all the keywords in a single textbox either in the basic search page or the advanced page. Since in the experiment system, those input keywords without any Boolean operator are processed according to default 'OR' logic, this strategy may incur much irrelevant feedback. In other words, the search results may be of high recall but of low accuracy.
2. The second type, the *unsuccessful multiple-textbox strategy*, refers to when a participant selects the advanced search, inputs keywords in multiple textboxes as per one word in one box, but does not specify any Boolean operator to logically connect the keywords. In the same way to the simple-search strategy, the system processes the keywords under 'OR' logic, and the user may not get the exact feedback up to the standard results. However, from the perspective of learning, when participants apply this strategy, they somewhat get the conception of the advanced search, which is supposed more effective than the simple search.
3. The third type, the *logic-AND-search strategy*, is the target strategy for the experiments in our study. When applying this strategy, a participant selects the advanced search, inputs keywords in multiple textboxes with one word in one box, and uses 'AND' operators to organise the keywords into a meaningful query. If all the required keywords associated with a search task are input, this strategy is expected to lead to correct search results.

From the collected experimental data, it was found that no student ever made attempts at the expert search.

Procedure

Given a search task, a participant was asked to carry out the following process:

1. Understand the task by examining the required keywords listed in the task form;
2. Figure out a strategy, including the search function and the keyword inputting scheme;
3. Depict the search strategy on the questionnaire;
4. Write down an expectation score (i.e., the participant's confidence of the strategy bringing desired results) on the questionnaire;
5. Execute the search (namely apply the formulated strategy);
6. Evaluate the search results by comparing them with the standard results presented by the organisers;
7. Write down a satisfaction score on the questionnaire;
8. Continue the next search task until all tasks are completed.

Each participant's learning process was observed by tracking their strategy adjustments in executing all the

search tasks in sequence.

Data analysis

For each of the two student groups, the collected experimental data were divided into two parts: (1) The first 70% of the data (associated with the first seven search tasks) were used to infer the parameters of each model; (2) The remaining 30% (regarding the last three tasks) were fitted by the estimated models. The model best fitting the data was used to explain the learning behaviour of the users in the corresponding group.

Estimation of model parameters

The maximum likelihood method was used to estimate the parameters of each model with regard to the experiment data of each group. The likelihood function for the g -th group and k -th model is defined as:

$$LL_{g,k}(\Theta) = \prod_{i=1}^{N_g} (\prod_{t=1}^T P_{i,k}^j(t)) \quad (\text{Equation 1})$$

where Θ denotes the parameters, $T=7$ is the number of training tasks, and N_g is the number of participants in group g . $P_{i,k}^j(t)$ stands for the attraction of strategy j adopted by user i for task t , and is computed under the updating rules of model k .

Table 2 details the parameter estimates.

Table 2: Parameters estimates

Student Group	Bush and Mosteller's model	Börger's and Sarin's model	Cross's model	Roth and Erev's modified model
Freshmen	$\alpha^{BM}=0.2; \beta^{BM}=0.1$	$\beta^{BS}=0.100$	$\alpha^{CR}=0.1;$ $\beta^{CR}=0.1$	$\phi=0; \varepsilon=0.3428$
Seniors	$\alpha^{BM}=0.1; \beta^{BM}=0.1$	$\beta^{BS}=0.258$	$\alpha^{CR}=0.1;$ $\beta^{CR}=0.4$	$\phi=0.4; \varepsilon=0.2407$

Note there is no parameter in Roth and Erev's basic model. The parameter X_{min} in Roth and Erev's modified model can be directly derived from questionnaire data. It is the minimum expectation per participant for all strategies.

Model fitting and verification

The final models were obtained by replacing the parameters with the estimates. The models were then applied to the experimental data associated with the last three search tasks: given a participant and a task, the probabilities of the participant choosing different search strategies were computed, and the strategy with the maximum probability was ticked as the predicted strategy. This process is referred to as model fitting, or in this study, strategy simulation.

The effectiveness of model fitting was evaluated by measuring the difference between the simulated strategies derived from each model and the actual strategies that participants took. This difference was gauged by the mean squared distance in the present study. The mean squared distance for the i -th participant and the k -th model is computed as follows:

$$MSD_{i,k} = \sum_{t=0.7 \cdot T+1}^T \sum_{j=1}^m \left(\left(P_{i,k}^j(t) - I(j, d_i(t)) \right)^2 / (0.3 \cdot T \cdot m) \right) \quad (\text{Equation 2})$$

where $T=10$ is the total number of search tasks, m denotes the size of strategy set, $P_{i,k}^j(t)$ is the probability of participant i taking strategy j to fulfil task t predicted by model k , $d_i(t)$ denotes the actual strategy chosen by participant i in period t , and $I(j, d_i(t))$ is a contingent decision function whose value is 0 when $j \neq d_i(t)$ or 1 when $j = d_i(t)$.

Table 3 reports the mean and standard deviation of the mean squared distances with regard to each student group and each model.

Table 3: Results of model verification

Mean and standard deviation of the mean squared distances per student group per model		Bush and Mosteller's model	Cross's model	Börger's and Sarin's model	Roth and Erev's model	Roth and Erev's modified model
Freshmen students	Mean	0.05278	0.004631	0.022475	0.024239	0.016259
	Standard deviation	0.013887	0.001122	0.010654	0.006058	0.002018
Senior students	Mean	0.017591	0.009455	0.034181	0.006506	0.006587
	Standard deviation	0.006063	0.019643	0.150697	0.002358	0.001466

For each group of students, the model with the smallest mean and standard deviation was chosen as the optimal model to fit their behaviour data. Consequently, based on data in Table 3, for freshmen, Cross's model fits best, while for seniors, Roth and Erev's modified model is the best.

Results

Freshmen's learning: Cross's model

From the data in [Table 3](#), it can be inferred that freshmen's search strategy adjustments comply more with Cross's model.

(1) Freshmen showed insistence and inertia towards earlier strategy preferences.

According to the updating rules of strategy attraction in Cross's model (Equations [8](#) and [9](#), see [Appendix](#)), freshmen (first year students) are more inclined to continue the search strategies employed in their last task.

Table 4 presents the statistics of users' behaviour obtained from the experiment data. It can be seen that freshmen were more likely to choose the simple search as the initial strategy and input keywords in a single search box. They did so based on their former experience of general search engine using.

Table 4: Statistics of users' learning behaviour in information searching

Indicators	Freshmen Students	Senior Students
Percentage of users with the initial strategy being the simple search	92.31	82.35
Average tasks after which users switched to the advanced search page	5.62	4.44

Average tasks after which users started to use the logic-AND-search strategy	7.61	5.74
--	------	------

The average tasks after which users first switched to the advanced search page and the average tasks after which users started the logic-AND-search are also reported in Table 4. The results tell that freshmen took more time to leave the simple search, learn to use new search functions and take new strategies. Their behaviour followed a Markov process, and they were somewhat insistent to their earlier strategy preferences.

(2) Freshmen could finally give up experiential preferences and comprehend new strategies by learning.

The parameter estimates of Cross's model for freshmen are: $\alpha^{CR}=0.1$; $\beta^{CR}=0.1$ (see Table 2). It implies that freshmen held insistence and inertia to the established strategies, but the extent was not so remarkable. As shown in Table 4, averagely after 6 to 8 tasks, freshmen gave up their preference of the simple search. They learned to use the advanced search and took the logic-AND-search strategy through trial and error. Most freshmen finally found out and used the logic-AND-search strategy, which was more possible to bring search results consistent with the standard ones.

Seniors' learning: Roth and Erev's modified model

The data in Table 3 indicate that for seniors (final year students), Roth and Erev's modified model is more ideal to fit their learning behaviour. They depended on their past experiences to align search strategies. At the same time, they developed strategies through rational thinking.

(1) Seniors were ready to make comprehensive decisions based on recent experiences.

The estimate of the forgetting parameter ϕ in Roth and Erev's modified model for seniors is 0.4 (see Table 2). According to Equation 12-15, this means, to a non-negligible extent, seniors would like to make comprehensive decisions based on their recent experiences. Basically, the more recently a search experience happens, the greater impact it has on the current decision making.

(2) Seniors showed strong subjectivity when evaluating the feedback from adopting a certain strategy.

According to Equation 14 (see Appendix), $R(\pi(t))=\pi(t)-X_{min}$, when making decisions, seniors demonstrated strong cognitive subjectivity. Different seniors might make different evaluations towards equal strategy payoffs.

Figures 2 and 3 depict the perceptions of the students who adopted the logic-AND-search strategy. Figure 2 portrays the average expectation per task of the freshmen and the seniors. Figure 3 illustrates the changes of their satisfactions. It can be inferred that, the freshmen held high expectations before applying the logic-AND-search strategy, and consistently scored high satisfactions with the feedback. In contrast, the seniors' expectations and satisfactions in different tasks were quite unsteady, and were almost lower than those of the freshmen.

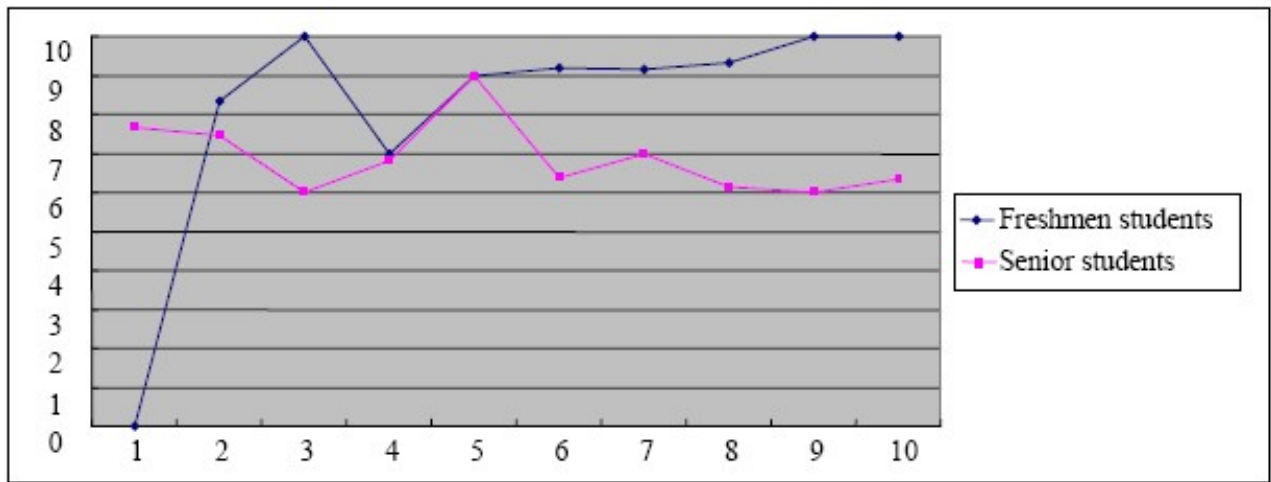


Figure 2: Average expectation per task of those students who adopted the logic-AND-search strategy

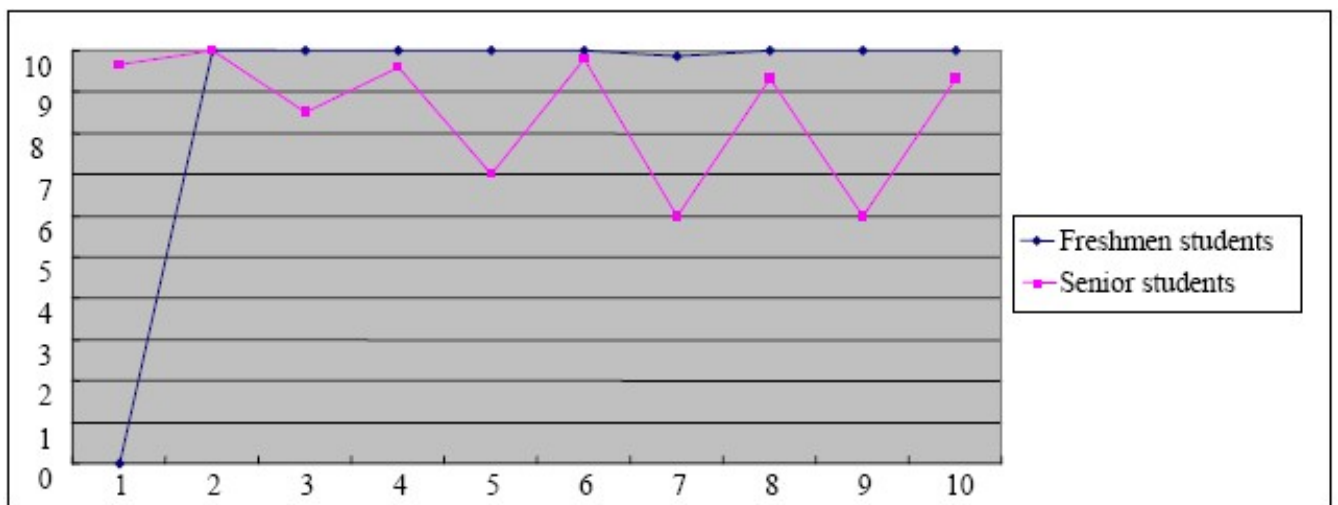


Figure 3: Average satisfaction per task of those students who adopted the logic-AND-search strategy

(3) Seniors paid attention to neighbour strategies.

The estimate of the transferring parameter ε in Roth and Erev's modified model for seniors is 0.2407 (see [Table 2](#)). According to [Equation 12-13](#), this means when adjusting their strategy, seniors were not completely affected by the information of the strategy adopted in the last search, but also concerned about the unemployed strategies. The strength of the unemployed strategies influencing their current strategy selection is 24.07%. In other words, seniors paid attention to neighbouring strategies.

Figure 4 describes the percentages of students who adopted the unsuccessful multiple-textbox strategy in each task. Figure 5 presents the percentages of students who correctly tried the logic-AND-search in each task. Interestingly, more seniors used logic-AND-search in the fourth task than in the fifth task. Correspondingly, fewer seniors took the unsuccessful multiple-textbox strategy in the fourth task than in the fifth task. That means some of the seniors who chose the correct strategy in one task returned to incorrect strategies in later tasks. This kind of phenomenon occurs several times (see Figures 4 and 5). After tracing back to the screen videos, the researchers found that a few seniors who had successfully employed the logic-AND-search started to explore other search options such as document type, year range, and so on. These options probably confused them and made them fail to use logic AND operators in subsequent tasks. Undoubtedly, those seniors displayed strong characteristics of rational thinking. This point is exactly what Roth and Erev's models try to reveal.

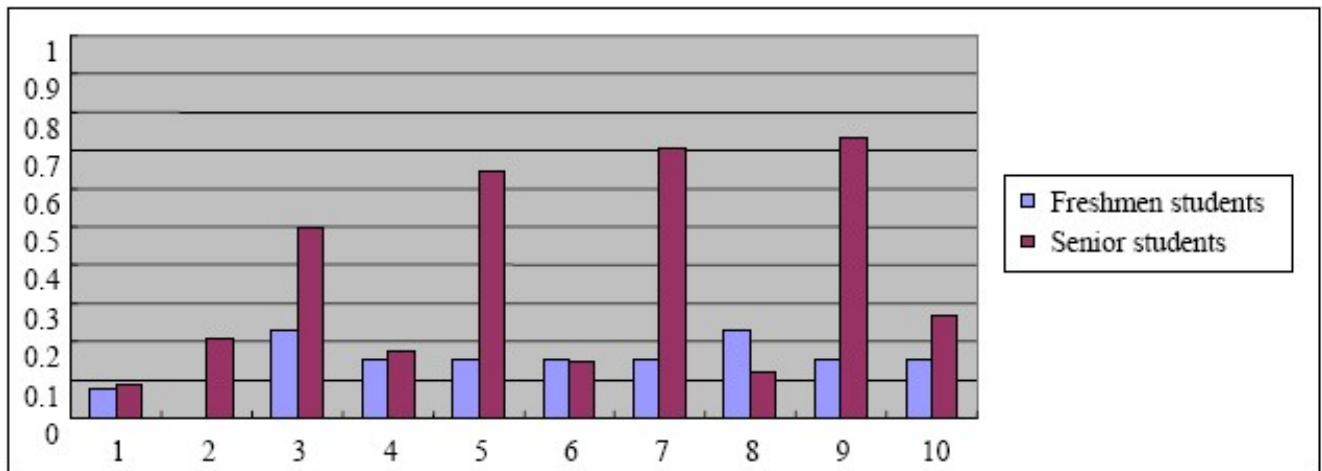


Figure 4: Percentages of students who followed the unsuccessful multiple textbox strategy

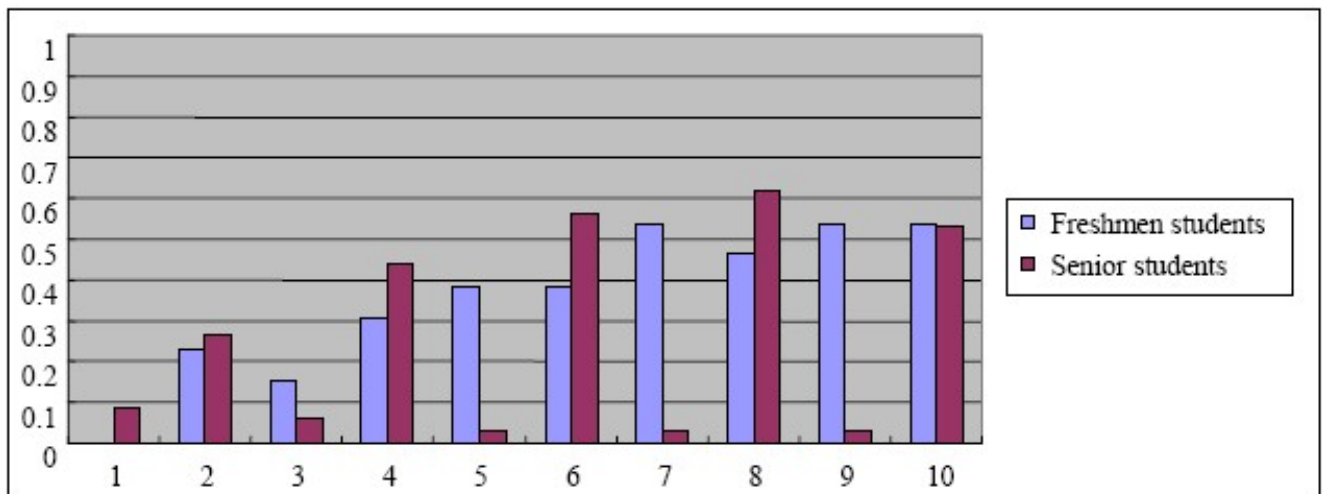


Figure 5: Percentages of students who adopted the logic-AND-search strategy

Summary

The above findings give substantial answers to the research questions, and confirm the theoretical assumptions.

- *Question:* How do users learn to use correct strategies to conduct scholarly information searches without instructions? In other words, are there learning rules controlling their strategy adjustments? If so what are the rules?

Answer: In a scheme of things, users demonstrated reinforcement learning characteristics. The strategies that brought success in their earlier experiences would be repeated with a higher probability. Through learning by trial and error, both freshmen and seniors could finally comprehend new search strategies. Answer to this question justifies the first assumption of this study.

- *Question:* What are the differences in learning mechanisms between users at different cognitive levels?

Answer: Users at different cognitive levels demonstrated different reinforcement patterns. The learning behaviour of freshmen showed remarkable Markov properties. Their strategy selection was determined by the feedback obtained in the last search activity. Cross's model better explains their learning mechanisms. For seniors, their strategy selection depended on the accumulated effect of past strategy adoptions. They displayed strong characteristics of rational thinking. Roth and Erev's modified model better describes their learning behaviour. Answer to this question substantiates the second research

assumption.

Discussion

Characteristics of reinforcement learning

It was found that most undergraduates preferred to repeat the strategies that bring success in their earlier experiences. This is highly consistent with the findings of Warwick *et al.* (2009: 2402) that undergraduate students

used their growing expertise to justify a conservative information strategy, retaining established strategies as far as possible and completing tasks with minimum information-seeking effort.

Specifically, according to this study, in the first task, 85% of undergraduates (92.3% of freshmen and 82.4% of seniors, See [Table 4](#)) chose the simple search as the initial strategy. It was supposed that the studied students were influenced by their former experience of general search engine using ([Du and Evans 2011](#); [Fast and Campbell 2004](#); [George et al. 2006](#); [Haglund and Olsson 2008](#); [Malliari et al. 2011](#)).

There were differences in the reinforcement learning process between freshmen and seniors, as previously claimed. Freshmen can be considered to be novices with little perception of scholarly information seeking, while seniors are users with more expertise. From this point of view, the differences in the reinforcement learning patterns between freshmen and seniors can be expanded by findings of Warwick *et al.* (2009: 2413), as follows:

Reflection on the learning theories of Kolb (1984) ... learners will often resist acquiring new skills because rejecting existing skill causes negative emotions (e.g., confusion, anger, upset). Existing skill is guarded zealously and adapted repeatedly until it finally fails ... Expert searchers therefore are not only differentiated by their existing skills but also potentially by their attitude to acquiring new ones.

Warwick *et al.* grounded the above point by referring to Kolb's (1984) learning theories, which are congruous with the assumptions of this study.

Effectiveness of reinforcement learning

Consider the average number of tasks it took participants to change from the simple search to the advanced search and start the logic-AND-search (See [Table 4](#)). It can be concluded that the learning effectiveness of academic users through self-regulated trial and error was not so satisfying. Especially, freshmen spent more time to learn the correct search strategy; the average tasks it took them to use the logic-AND-search were 7.61 out of 10. This highlights the necessity of external instructions to improve the effectiveness of user's learning of information seeking, especially for novices. Although this declaration should be further justified, the researchers are still positive with it by referring to other studies ([Colvin and Keene 2004](#); [Halttunen and Jarvelin 2005](#); [Ren 2000](#)).

Besides, seniors learned the correct search strategy more quickly than freshmen, as described in [Table 4](#). This is in agreement with the studies of Chen (2009), Eshet-Alkalai and Chajut (2009), Hsieh-Yee (1993), Korobili *et al.* (2011), and Thatcher (2008). Specifically, this study to some extent confirmed the findings of the recent work done by Korobili *et al.* (2011), that there are statistical significant relationships between users' experience in databases or e-journals and the variables: *more than one keyword*, *Boolean operators* as search

techniques, *change strategy*, *different keywords* as techniques to modify the initial strategy, and so on.

Conclusions and implications

The study observed the strategy adjustments of thirteen first-year undergraduates and thirty-four fourth-year undergraduates in carrying out ten search tasks in a specified database system independently. It was assumed that there are discrepancies in the level of cognitive processing between the two groups of users. The impacts of cognitive levels on learning of searching skills were examined by excluding the effects of other factors through quasi-experimental settings. When executing a search task, a user was asked to write down: (1) the description of the formulated search strategy; (2) the expectation of the strategy bringing desired results; and (3) the satisfaction with the strategy. The dynamics of search strategies, expectations and satisfactions of each user across different tasks were simulated through five reinforcement learning models. These dynamics were supposed to be the outcomes of participants' learning and reflection.

It is found that undergraduates prefer to retain established strategies. It takes them a long time to change from the simple search to the advanced search and learn to use the most effective strategy. Generally, in the process of searching expertise learning, users demonstrate reinforcement characteristics. If a search strategy leads to satisfactory results, this strategy will be more likely to be repeated with high expectation later; if a strategy leads to unsatisfactory results, it will be more likely to be avoided afterwards. Specifically, users at different cognitive levels demonstrate different reinforcement patterns. Freshmen's strategy selection is always made according to the feedback obtained in the last search activity, whereas seniors rely on their search experiences and rational thinking to make comprehensive decisions.

Through observing and quantitatively simulating the micro process of academic users' learning of searching expertise, the current research enhances our understanding of users' experience of scholarly information seeking. Besides, based on the research outcomes and discussion, implications can be proposed from the perspectives of training programme design, adaptive information retrieval system design and theoretical development.

As formerly discussed, learning through self-regulated trials is not the most effective and economic way for academic users to develop searching expertise. Extra instructions are needed to improve their learning performance. Instructions can be imparted through training curriculums offered by librarians, as well as online learning or help features incorporated into information retrieval systems. Rather than just a 'list of skills' of information literacy ([Maybee 2006](#)), the instructions should be tailored to the learning patterns of different users. This deserves further investigation by librarians.

By monitoring users' searching behaviour and identifying users' learning characteristics, information retrieval systems can offer personalised supports to suit the users and their search tasks, and assist them to complete the tasks, as suggested by Li and Belkin ([2008](#)), Stelmaszewska *et al.* ([2005](#)) and Xie and Cool ([2009](#)), and technically practiced by de la Chica *et al.* ([2008](#)), Frias-Martinez *et al.* ([2007](#); [2008](#)), Hurst *et al.* ([2007](#)), Jansen ([2005](#)), Stelmaszewska *et al.* ([2005](#)) and Tsuji and Yamamoto ([2001](#)). This kind of adaptive feature is expected to facilitate users' learning of searching expertise and improve the effectiveness of their interactions with the search systems. The present research provides understanding of observational variables (e.g., initial search strategy, strategy adjustments, behavioural pathway, combination of Boolean operators, and so on) for automatically identifying users' learning characteristics in the development of such adaptive systems.

Due to the small sample size, the findings reported in this paper are considered to be exploratory and preliminary. Further efforts can be dedicated to develop a comprehensive quantitative research framework. This research framework synthesises learning theories and information-searching paradigms, as partly described by [Figure 1](#). It is expected to '*better describe the information searching process than more commonly used paradigms of decision making or problem solving*' ([Jansen et al. 2009](#): 643). According to Kuhlthau ([1993](#):

342), the whole information search process '*incorporates three realms of human experience: the affective (feelings), the cognitive (thoughts) and the physical (actions)*'. The complexities of affective, cognitive and physical interactions within this process require deliberate design of learning parameters and reinforcement adjustment functions. Besides, the effects of contextual elements including instructional variables (e.g., search tips, anchored helps, graphic or video demos, result faceting, clustering or visualisation, and so forth) on the performance of users' learning and information searching should be included to establish a more meaningful learning model.

Acknowledgements

This research was supported by the National Natural Science Foundation of China under contract No. 70773054, No. 71001052 and No.71003049. The authors wish to give special thanks to Prof. T.D. Wilson, Prof. A. Smith, Dr. J.T. Du and all the anonymous reviewers, for their kind suggestions and comments for improving this particular research.

About the authors

Yonghua Cen is an Associate Professor in the Department of Information Management, School of Economics and Management, at Nanjing University of Science and Technology in China. He is also a researcher in Advanced Analytics Institute at University of Technology Sydney in Australia. His current scientific interests are in the fields of behaviour informatics, dynamic social network analysis and scientometrics. He can be contacted at: justin.cen@gmail.com.

Liren Gan, is a Professor in the Department of Information Management, School of Economics and Management, at Nanjing University of Science and Technology. She is a PhD supervisor. Her principal research interests concern analysis of user cognition, mental models and behaviour in online environments. She can be contacted at: gan5707@vip.sina.com.

Chen Bai, is a PhD student under the supervision of Professor Gan. Her research topics cover digital library and user analysis. She can be contacted at: flyluo77@sina.com.

References

- Bates, M.J. (1979). Information search tactics. *Journal of the American Society for Information Science*, **30**(4), 205-214.
- Bates, M.J. (2002). Toward an integrated model of information seeking and searching. *The New Review of Information Behaviour Research*, **3**, 1-15.
- Belkin, N.J. (1980). Anomalous states of knowledge as a basis for information retrieval. *Canadian Journal of Information Science*, **5**(2), 133-143.
- Belkin, N.J., Cool, C., Koenmman, J., Ng, K.B. & Park, S. (1996). [Using relevance feedback and ranking in interactive searching](#). *Proceedings of the Fourth Text Retrieval Conference (TREC-4)*. (pp. 181-209). Gaithersburg, MD: National Institute of Standards and Technology. Retrieved 7 March, 2013 from http://comminfo.rutgers.edu/~belkin/articles/trec4_paper.pdf (Archived by WebCite® at <http://www.webcitation.org/6EwtGjxT6>)
- Bilal, D., Kirby, J. (2002). Differences and similarities in information seeking: children and adults as Web users. *Information Processing & Management*, **38**(5), 649-670.
- Bloom, B.S., Englehard, E., Furst, W. & Krathwohl, D.R. (1956). *Taxonomy of educational objectives: the classification of educational goals*. New York, NY: McKay.
- Börgers, T. & Sarin, R. (2000). Naive reinforcement learning with endogenous aspirations. *International Economic Review*, **41**(4), 921-950.
- Brand-Gruwel, S., Wopereis, I. & Walraven, A. (2009). A descriptive model of information

- problem solving while using internet. *Computers & Education*, **53**(4), 1207-1217.
- Brenner, T. (2006). Agent learning representation: advice in modelling economic learning. In L. Tesfatsion & K.L. Judd (Eds.), *Handbook of computational economics* Volume 2. (pp. 895-947). Elsevier.
 - Bush, R.R. & Mosteller, F. (1953). A stochastic model with applications to learning. *The Annals of Mathematical Statistics*, **24**(4), 559-585.
 - Chen, H.L. (2009). An analysis of undergraduate students' search behaviors in an information literacy class. *Journal of Web Librarianship*, **3**(4), 333-347.
 - Chu, S.K., & Law, N. (2007). Development of information search expertise: postgraduates' knowledge of searching skills. *portal: Libraries & the Academy*, **7**(3), 295-316.
 - Cole, C., Lin, Y. & Leide, J. (2007). A classification of mental models of undergraduates seeking information for a course essay in history and psychology: preliminary investigations into aligning their mental models with online thesauri. *Journal of the American Society for Information Science & Technology*, **58**(13), 2092-2104.
 - Colvin, J. & Keene, J. (2004). [Supporting undergraduate learning through the collaborative promotion of e-journals by library and academic departments](http://InformationR.net/ir/9-2/paper173.html). *Information Research*, **9**(2), paper 173 Retrieved 7 March, 2013 from <http://InformationR.net/ir/9-2/paper173.html> (Archived by WebCite® at <http://www.webcitation.org/6CiDyWv6S>)
 - Cross, J.G.A. (1973). Stochastic learning model of economic behavior. *Quarterly Journal of Economics*, **87**, 239-266.
 - de la Chica, S., Ahmad, F., Sumner, T., Martin, J.H. & Butcher, K. (2008). Computational foundations for personalizing instruction with digital libraries. *International Journal of Digital Libraries*, **9**, 3-18.
 - Du, J.T. & Evans, N. (2011). Academic users' information searching on research topics: characteristics of research tasks and search strategies. *The Journal of Academic Librarianship*, **37**(4), 299-306.
 - Du, J.T. & Spink, A. (2011). Towards a Web search model: integrating multitasking, cognitive coordination and cognitive shifts. *Journal of the American Society for Information Science and Technology*, **62**(8), 1446-1472.
 - Eisenberg, M.B. & Berkowitz, R.E. (1990). *Information problem-solving: the Big Six skills approach to library and information skills instruction*. Norwood, NJ: Ablex.
 - Ellis, D. (1989). A behavioural approach to information retrieval system design. *Journal of Documentation*, **45**, 171-212.
 - Erev, I. & Roth, A.E. (1996). On the need for low rationality, cognitive game theory: reinforcement learning in experimental games with unique, mixed strategy equilibria. Pittsburgh, PA: University of Pittsburgh. [Mimeographed manuscript.]
 - Eshet-Alkalai, Y. & Chajut, E. (2009). Changes over time in digital literacy. *CyberPsychology & Behavior*, **12**(6), 713-715.
 - Fast, K.V., & Campbell, D.G. (2004). 'I still like Google': university student perceptions of searching OPACs and the Web. *Proceedings of the American Society for Information Science & Technology*, **41**(1), 138-146.
 - Ford, N., Wilson, T.D., Foster, A., Ellis, D. & Spink, A. (2002). Information seeking and mediated searching. Part 4, Cognitive styles in information seeking. *Journal of the American Society for Information Science & Technology*, **53**(9), 728-735.
 - Ford, N., Miller, D. & Moss, N. (2003). Web search strategies and approaches to studying. *Journal of the American Society for Information Science & Technology*, **54**(6), 473-489.
 - Frias-Martinez, E., Chen, S.Y. & Liu, X. (2007). Automatic cognitive style identification of digital library users for personalization. *Journal of the American Society for Information Science & Technology*, **58**(2), 237-251.
 - Frias-Martinez, E., Chen, S.Y. & Liu, X. (2008). Investigation of behavior and perception of digital

- library users: a cognitive style perspective. *International Journal of Information Management*, **28**(5), 355-365.
- Fu, W.T. & Anderson, J.R. (2006). From recurrent choice to skill learning: a reinforcement-learning model. *Experimental Psychology: General*, **135**(2), 184-206.
 - George, C., Bright, A., Hurlbert, T., Linke, E.C., St. Clair, G. & Stein, J. (2006). [Scholarly use of information: graduate students' information seeking behaviour](#). *Information Research*, **11**(4), paper 272 Retrieved 7 March, 2013 from <http://InformationR.net/ir/11-4/paper272.html> (Archived by WebCite® at <http://www.webcitation.org/6CiESmEj2>)
 - Gerjets, P. & Hellenthal-Schorr, T. (2008). Competent information search in the World Wide Web: development and evaluation of a web training for pupils. *Computers in Human Behavior*, **24**(3), 693-715.
 - Haglund, L. & Olsson, P. (2008). The impact on university libraries of changes in information behavior among academic researchers: a multiple case study. *The Journal of Academic Librarianship*, **34**(1), 52-59.
 - Halttunen, K. (2003). Students' conceptions of information retrieval: implications for the design of learning environments. *Library & Information Science Research*, **25**(3), 307-332.
 - Halttunen, K. & Järvelin, K. (2005). Assessing learning outcomes in two information retrieval learning environments. *Information Processing & Management*, **41**(4), 949-972.
 - He, W., Erdelez, S., Wang, F. & Shyu, C. (2008). The effects of conceptual description and search practice on users' mental models and information seeking in a case-based reasoning retrieval system. *Information Processing & Management*, **44** (1), 294-309.
 - Hsieh-Yee, I. (1993). Effects of search experience and subject knowledge on the search tactics of novice and experienced searchers. *Journal of the American Society for Information Science*, **44**(3), 161-174.
 - Hurst, A., Hudson, S.E. & Mankoff, J. (2007). Dynamic detection of novice vs. skilled use without a task model. In *CHI'07: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, April 28-May 3, 2007* (pp. 271-280). New York, NY: ACM Press.
 - Ingwersen, P. (1996). Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory. *Journal of Documentation*, **52**(1), 3-50.
 - Izquierdo, L.R., Izquierdo, S.S., Gotts, N.M. & Polhill, J.G. (2007). Transient and asymptotic dynamics of reinforcement learning in games. *Games and Economic Behavior*, **61**(2), 259-276.
 - Jansen, B.J. (2005). [Seeking and implementing automated assistance during the search process](#). *Information Processing & Management*, **41**(4), 909-928. Retrieved 7 March, 2013 from <http://www.interruptions.net/literature/Jansen-InformProcessManag05.pdf> (Archived by WebCite® at <http://www.webcitation.org/6EwrdV8SZ>)
 - Jansen, B.J., Booth, D. & Smith, B. (2009). Using the taxonomy of cognitive learning to model online searching. *Information Processing & Management*, **45**(6), 643-663.
 - Kelly, D. & Fu, X. (2007). Eliciting better information need descriptions from users of information search systems. *Information Processing & Management*, **43**(1), 30-46.
 - Kim, K. (2002). Information-seeking on the Web: effects of user and task variables. *Library & Information Science Research*, **23**(3), 233-255.
 - Kolb, D.A. (1984). *Experiential learning experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.
 - Korobili, S., Malliari, A. & Zapounidou, S. (2011). Factors that influence information-seeking behavior: the case of Greek graduate students. *The Journal of Academic Librarianship*, **37**(2), 155-165.
 - Kuhlthau, C.C. (1993). A principle of uncertainty for information seeking. *Journal of Documentation*, **49**(4), 339-355.
 - Kuhlthau, C.C., Caspari A.K. & Maniotes, L.K. (2007). *Guided inquiry: learning in the 21st century*. Westport, CT: Libraries Unlimited.

- Laxman, K. (2010). A conceptual framework mapping the application of information search strategies to well and ill-structured problem solving. *Computers & Education*, **55**, 513-526.
- Li, Y. & Belkin, N.J. (2008). A faceted approach to conceptualizing tasks in information seeking. *Information Processing & Management*, **44**(6), 1822-1837.
- Marchionini, G. (2006). Exploratory search: from finding to understanding. *Communication of the ACM*, **49**(4), 41-47.
- Malliari, A., Korobili, S. & Zapounidou, S. (2011). Exploring the information seeking behavior of Greek graduate students: a case study set in the University of Macedonia. *The International Information & Library Review*, **43**(2), 79-91.
- Maybee, C. (2006). Undergraduate perceptions of information use: the basis for creating user-centered student information literacy instruction. *The Journal of Academic Librarianship*, **32**(1), 79-85.
- Puustinen, M. & Rouet, J.F. (2009). Learning with new technologies: help seeking and information searching revisited. *Computers & Education*, **53**, 1014-1019.
- Ren, W. (2000). Library instruction and college student self-efficacy in electronic information searching. *The Journal of Academic Librarianship*, **26**(5), 323-328.
- Roth, A.E. & Erev, I. (1995). Learning in extensive form games: experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, **8**(1), 164-212.
- Saito, H. & Miwa, K. (2007). Construction of a learning environment supporting learners' reflection: a case of information seeking on the Web. *Computers & Education*, **49**(2), 214-229.
- Saracevic, T. (1996). Modeling interaction in information retrieval (IR): a review and proposal. *Proceedings of the ASIS Annual Meeting*, **33**, 3-9
- Savolainen, R. (1993). The sense-making theory: reviewing the interests of a user-centred approach to information seeking and use. *Information Processing & Management*, **29**(1), 13-28.
- Shimokawa, T., Suzuki, K., Misawa, T. & Okano, Y. (2009). Predicting investment behaviour: an augmented reinforcement learning model. *Neurocomputing*, **72** (16-18), 3447-3461.
- Spink, A. (1997). Study of interactive feedback during mediated information retrieval. *Journal of the American Society for Information Science*, **48**(5): 382-394.
- Spink, A., Wilson, T.D., Ford, N., Foster, A. & Ellis, D. (2002a). Information seeking and mediated searching. Part 1, Theoretical framework and research design. *Journal of the American Society for Information Science & Technology*, **53**(9), 695-703.
- Spink, A., Wilson, T.D., Ford, N., Foster, A. & Ellis, D. (2002b). Information seeking and mediated searching. Part 3, Successive searching. *Journal of the American Society for Information Science & Technology*, **53**(9), 716-727.
- Stelmaszewska, H., Blandford, A. & Buchanan, G. (2005). Designing to change users' information seeking behaviour: a case study. In S.Y. Chen & G.D. Margoulas (Eds.), *Adaptable and adaptive hypermedia systems* (pp. 1-18). Hershey, PA: Idea Group.
- Sutcliffe, A., Ennis, M., & Watkinson, S.J. (2000). Empirical studies of end-user information searching. *Journal of the American Society for Information Science*, **51**, 1211-1231.
- Sutton, R.S. & Barto, A.G. (1998). *Reinforcement learning: an introduction*. Cambridge, MA: MIT Press.
- Tabatabai, D. & Shore, B.M. (2005). How experts and novices search the Web. *Library & Information Science Research*, **27**(2), 222-248.
- Tenopir, C., Wang, P. & Zhang, Y. (2008). Academic users' interactions with ScienceDirect in search tasks: affective and cognitive behaviours. *Information Processing & Management*, **44**(1), 105-121.
- Thatcher, A. (2008). Web search strategies: the influence of Web experience and task type. *Information Processing & Management*, **44**(3), 1308-1329.
- Tsuji, S. & Yamamoto, Y. (2001). A framework to provide integrated online documentation. In: *Proceedings of the 19th Annual International Conference on Computer Documentation*

(SIGDOC'01), October 21-24, Santa Fe, New Mexico, USA, (pp. 185-192). New York, NY: ACM Press.

- Vakkari, P. (2001). A theory of the task-based information retrieval process: a summary and generalisation of a longitudinal study. *Journal of Documentation*, **57**(1), 44-60.
- Warwick, C., Rimmer, J., Blandford, A., Gow, J. & Buchanan, G. (2009). Cognitive economy and satisficing in information seeking: a longitudinal study of undergraduate information behavior. *Journal of the American Society for Information Science & Technology*, **60**(12), 2402-2415.
- Wildemuth, B.M. (2004). The effects of domain knowledge on search tactic formulation. *Journal of the American Society for Information Science & Technology*, **55**(3), 246-258.
- Wilson, M.L., Schraefel, M.C. & White, R.W. (2009). Evaluating advanced search interfaces using established information-seeking models. *Journal of the American Society for Information Science & Technology*, **60**(7), 1407-1422.
- Wilson, T.D. (1997). Information behaviour: an interdisciplinary perspective. *Information Processing & Management*, **33**(4), 551-572.
- Wilson, T.D. (1999). [Models in information behaviour research](http://informationr.net/tdw/publ/papers/1999JDoc.html). *Journal of Documentation*, **55**(3) 249-270. Retrieved 7 March, 2013 from <http://informationr.net/tdw/publ/papers/1999JDoc.html> (Archived by WebCite® at <http://www.webcitation.org/6EwqakJoK>)
- Wilson, T.D. (2000). [Human information behavior](http://www.inform.nu/Articles/Vol3/v3n2p49-56.pdf). *Informing Science*, **3**(2), 49-55. Retrieved 7 March, 2013 from <http://www.inform.nu/Articles/Vol3/v3n2p49-56.pdf> (Archived by WebCite® at <http://www.webcitation.org/6EwqTIg0B>)
- Wilson, T.D., Ford, N., Ellis, D., Foster, A.E. & Spink, A. (2002). Information seeking and mediated searching. Part 2, Uncertainty and its correlates. *Journal of the American Society for Information Science & Technology*, **53**(9), 704-715.
- Xie, H. (2000). Shifts of interactive intentions and information-seeking strategies in interactive information retrieval. *Journal of the American Society for Information Science*, **51**(9), 841-857.
- Xie, H. (2007). [Shifts in information-seeking strategies in information retrieval in the digital age: planned-situational model](http://InformationR.net/ir/12-4/colis/colis22.html). *Information Research*, **12**(4), paper colis22. Retrieved 7 March, 2013 from <http://InformationR.net/ir/12-4/colis/colis22.html> (Archived by WebCite® at <http://www.webcitation.org/6CiEeOu3a>)
- Xie, I. & Cool, C. (2009). Understanding help seeking within the context of searching digital libraries. *Journal of the American Society for Information Science & Technology*, **60**(3), 477-494.
- Zhang, Y. (2008). The influence of mental models on undergraduate students' searching behavior on the Web. *Information Processing & Management*, **44**(3), 1330-1345.
- Zhu, Y., Chen, L., Chen, H. & Chern, C. (2011). How does Internet information seeking help academic performance? The moderating and mediating roles of academic self-efficacy. *Computers & Education*, **57**, 2476-2484.

How to cite this paper

Cen Y., Gan L. & Bai C. (2013). Reinforcement learning in information searching *Information Research*, **18**(1) paper 569. [Available at <http://InformationR.net/ir/18-1/paper569.html>]

Find other papers on this subject

Scholar Search

Google Search

Bing

Check for citations, [using Google Scholar](https://scholar.google.com/citations?view_op=view_citation&hl=en&user=informationr.net/ir/18-1/paper569.html#UId__dKAquM)

Like

0

Tweet

0

Appendix: Reinforcement learning models

The basic ideas of the models listed in [Table 1](#) are further explained as follows:

(1) Bush and Mosteller's model

In Bush and Mosteller's model ([Bush and Mosteller 1953](#)), a probability variable $P(i)$ is used to define the attraction of a strategy to a certain user (denoted as u). Let $d(t)$ denote the strategy which is chosen by user u in period t , and $\pi(t)$ stand for the reward or punishment fed back to the user in period t . A nonnegative $\pi(t)$ means the user gets a reward, otherwise a punishment. Suppose in period t , user u chooses the j -th strategy from the strategy set, i.e. $j=d(t)$. Then for u , the attraction of strategy j is updated under the following rules:

$$P(j, t+1) = \begin{cases} P(j, t) + \alpha^{BM} \cdot (1 - P(j, t)) & j = d(t) \wedge \pi(t) \geq 0 \\ P(j, t) - \beta^{BM} \cdot P(j, t) & j = d(t) \wedge \pi(t) < 0 \end{cases} \quad (\text{Equation 3})$$

For each strategy k other than j (namely those unemployed strategies), the attraction value is updated according to the following rules:

$$P(k, t+1) = \begin{cases} P(k, t) - \alpha^{BM} \cdot P(k, t) & k \neq d(t) \wedge \pi(t) \geq 0 \\ P(k, t) + \beta^{BM} \cdot (1 - P(k, t)) & k \neq d(t) \wedge \pi(t) < 0 \end{cases} \quad (\text{Equation 4})$$

In the above adjusting rules, α^{BM} and β^{BM} are two parameters to be estimated. $\alpha^{BM} \in [0, 1]$ is the weight factor assigned to a nonnegative payoff, while $\beta^{BM} \in [0, 1]$ the weight factor to a negative payoff. A smaller α^{BM} means that a nonnegative payoff plays a slighter part in the strategy selection, while a smaller β^{BM} means that a negative payoff plays a minor role in the strategy selection.

More intuitively, the learning rules that Bush and Mosteller's model describes can be interpreted as: when a certain strategy leads to a positive payoff, the probability of this strategy being chosen again increases and the probability of it being avoided decreases; otherwise, the probability of the strategy being further adopted decreases and the probability of it being rejected increases.

(2) Börgers and Sarin's model

Compared to Bush and Mosteller's model, Börgers and Sarin's model ([Börgers and Sarin 2000](#)) details the information for evaluating the payoff of a strategy adoption. It assumes that the evaluation of a strategy does not directly rely on the absolute value of the actual payoff, but on the difference between the actual payoff and the expected one. Let $A(t) \in [0, 1]$ denote the payoff expectation of a user before employing a strategy in period t and $A(1)$ be the initial expectation for the user before decision-making.

If $\pi(t) \geq A(t)$, the attraction value of each strategy after period t is updated by:

$$P(j, t+1) = \begin{cases} P(j, t) + \alpha^{BS} \cdot (1 - P(j, t)) & j = d(t) \\ P(j, t) - \alpha^{BS} \cdot P(j, t) & j \neq d(t) \end{cases} \quad (\text{Equation 5})$$

Otherwise, the attraction values are updated as follows:

$$P(j, t + 1) = \begin{cases} P(j, t) - \alpha^{BS} \cdot P(j, t) & j = d(t) \\ P(j, t) + \alpha^{BS} \cdot (1 - P(j, t)) & j \neq d(t) \end{cases} \quad (\text{Equation 6})$$

The payoff expectation is updated as follows:

$$A(t + 1) = (1 - \beta^{BS}) \cdot A(t) + \beta^{BS} \cdot \pi(t) \quad (\text{Equation 7})$$

The parameter α^{BS} is regarded as the reinforcement strength, whose value is the absolute difference between the actual payoff and the expected one, i.e., $\alpha^{BS} = |\pi(t) - A(t)|$. The parameter β^{BS} is set fixed, which stands for the adjustment speed of payoff expectation. The bigger β^{BS} is, the more greatly the current payoff influences the further strategy selection.

Similarly, Börgers and Sarin's model can be summarised as: if the actual payoff exceeds the expectation of an individual after a strategy is settled, then the probability of this strategy being further selected increases. On the contrary, if the actual payoff is smaller than the expectation, the probability of the strategy being adopted in future decreases. The expected payoff changes dynamically according to the actual payoff of the previous strategy adoption.

(3) Cross's model

As a modification to Bush and Mosteller's model, Cross's model ([Cross 1973](#)) is one of the most acknowledged reinforcement learning models.

Let $R(\pi(t))$ be the reinforcement strength, which is a monotonic function of the payoff $\pi(t)$. The attraction value of each strategy after period t is updated by:

$$P(j, t + 1) = \begin{cases} P(j, t) + R(\pi(t)) \cdot (1 - P(j, t)) & j = d(t) \\ P(j, t) - R(\pi(t)) \cdot P(j, t) & j \neq d(t) \end{cases} \quad (\text{Equation 8})$$

$$R(\pi(t)) = \alpha^{CR} \cdot \pi(t) + \beta^{CR} \quad (\text{Equation 9})$$

In the above rules, $\alpha^{CR} \in [0, 1]$ and $\beta^{CR} \in [0, 1]$ are two parameters that control the updating mechanism of the attraction of each strategy.

In Cross's model, the attraction of a strategy is defined as a linear function of the payoff by configuring the reinforcement strength as a variable correlated to the payoff, whereas in Bush and Mosteller's model, the reinforcement strength factors, α^{BM} and β^{BM} , are fixed and independent to payoffs.

(4) Roth and Erev's model

Both Cross's model and Börgers and Sarin's model are essentially modifications of Bush and Mosteller's model. All these models place emphasis on the Markov characteristics of players' strategy selection. In other words, when making a decision, an individual prefers to choose a strategy in terms of the payoff gained from the last strategy adoption. In contrast, Roth and Erev's models ([Roth and Erev 1995](#)) underline users' prior experience. That is to say, decision makers select a strategy based on their experiential expectations for all strategies. These expectations result from the accumulated effect of their past strategy adoptions, not only the last one.

In Roth and Erev's model, the attraction value of each strategy after period t is updated under the following

linear rules:

$$A_k(t+1) = \begin{cases} A_k(t) + \pi(t) & k = j \\ A_k(t) & k \neq j \end{cases} \quad (\text{Equation 10})$$

$$P_k(t+1) = A_k(t+1) / \sum_i A_i(t+1) \quad (\text{Equation 11})$$

Here, $A_k(t)$ is the accumulated payoff from adopting the k -th strategy before and in period t .

(5) Roth and Erev's modified model

In Roth and Erev's modified model ([Erev and Roth 1996](#)), the attraction values of the strategies after period t are updated:

$$A_k(t+1) = (1 - \varphi) \cdot A_k(t) + E_j(k, R(\pi(t))) \quad (\text{Equation 12})$$

$$E_j(k, R(\pi(t))) = \begin{cases} R(\pi(t)) \cdot (1 - \varepsilon) & k = j \\ R(\pi(t)) \cdot \varepsilon & k \neq j \end{cases} \quad (\text{Equation 13})$$

$$R(\pi(t)) = \pi(t) - X_{min} \quad (\text{Equation 14})$$

$$P_k(t+1) = A_k(t+1) / \sum_i A_i(t+1) \quad (\text{Equation 15})$$

where φ is a forgetting parameter measuring the attenuation degree of users' experiences influencing their strategy selection, and X_{min} is the minimum expectation of a user for all the strategies. Through φ and X_{min} , different users may make different subjective evaluations to a strategy even when the payoffs from applying the strategy are equal. $E_j(k, R(\pi(t)))$ is a function controlling how the payoff $\pi(t)$ from implementing strategy j updates the reinforcement strength $A_k(t+1)$, and ε is a transferring parameter that determines the extent of the reinforcement strength transferring to the unemployed strategies.

484

© the authors, 2013.

Last updated: 8 March, 2013

- [Contents](#) |
 - [Author index](#) |
 - [Subject index](#) |
 - [Search](#) |
 - [Home](#)
-