2013 IEEE 27th International Conference on Advanced Information Networking and Applications

Using Intelligent Multi-Agent Systems to Model and Foster Self-Regulated Learning: A Theoretically-Based Approach using Markov Decision Process

Babak Khosravifar, Francois Bouchet, Reza Feyzi-Behnagh, Roger Azevedo, and Jason M. Harley
Department of Educational and Counselling Psychology
McGill University, Montreal, Canada
babak.khosravifar@mcgill.ca

***Abstract*—In self-regulated learning [1] concept, Intelligent Tutoring Systems (ITS) can be designed to foster learning behaviors through pedagogical agents (PAs) that are used for interactions and exchange information with the human learner. These agents are intelligent and follow rational behaviors, but in the case of multi-agent environments they need to be systematically and specifically designed, however in order to follow a common goal, different self-regulatory systems have been designed that use pedagogical agents, but they fail to constrain the decision making of the agents and maintain a sequential decision making process during learning interactions with human learners. In this paper, we provide a new theoretical model for agent-learner interactions in MetaTutor, a multi-agent hypermedia learning environment, using Markov decision processes. We theoretically define the agents' Markov decisions and their influence on MetaTutor's performance as a whole. First, we formally define the Markov architecture and its parameters. We then link these characteristics to the pedagogical agents we use in MetaTutor and define different versions of MetaTutor agents equipped with Markov decision mechanism. Furthermore, we explore additional details about agents' sequential decision making and how reward functions influence their acting strategies with learners in the learning environment. We introduce the optimization problem in which we aim to maximize the expected return of the overall agents' acts in a self-regulatory system. What specifically distinguishes this work from the previous proposals in the same domain is its novelty in continuous decision making mechanism investigation and performance analysis that improve the applicability of the proposed adaptive model in a multi-agent ITS like MetaTutor.**

***Keywords*-Pedagogical Agents; Self-Regulated Learning; Intelligent Tutoring Systems; Markov Decision Process; Multi-Agent Systems;**

I. Introduction

ITS is developed to foster Self-Regulated Learning (SRL) systems as they are getting increasingly popular with improvements in advanced learning technologies using pedagogical agents. PAs are rational and effectively interact with human learners in order to help them better self-regulate their Cognitive, Affective, Metacognitive, and Motivational (CAMM [2], [10]) learning processes over time. In a typical learning environment, the intention of the human learner is unknown to a certain extent. To this end, PAs have to try different strategies to effectively deal with the variety of learner-system interaction patterns. Solving this ambiguity is the main challenge of these agents whose objective is to optimize the use of self-regulated learning processes by the students.

In the past few years, there have been a number of theoretical proposals that model SRL [14], [15] that consider complex nature of CAMM processes during learning with intelligent multi-agent learning environments, but they fail to effectively interact with human learners with uncertain intentions. The uncertainty is due to the fact that learners encounter a challenging science topic and they need to regulate their CAMM processes throughout the learning task, but they lack the required knowledge and skills to effectively perform those crucial monitoring and regulation processes. It can be shown that uncertain systems could be modelled as Markov Decision Processes (MDPs). Using MDP framework helps PAs to incorporate uncertainty in the ITS, and allows actions to be chosen based on an optimization criteria.

In this paper, we provide a new theoretical model of agent-learner interaction in MetaTutor [1], [2], a multi-agent hypermedia learning environment, using MDPs. Specially, we will define the multi-agent architecture using MDPs to convert PAs to sequential decision makers that are aimed at effectively interacting with uncertain human learners and optimizing their CAMM learning processes. In this regard, we use the Reinforcement Learning (RL [12]) model as an effective intellectual learning technique to enhance the ability of the strategy analysis of PAs. This paper focuses on how we can define a system designed to scaffold SRL using MDPs and RL technique to obtain good decisions while interacting with human learners with a vast range of knowledge, characteristics and attributes. In the proposed MDP, we present details regarding how the intentions of participants with MetaTutor can be broken down into different states and how these states are correlated.

The remainder of this paper is organized as follows. In

[1]Authors are affiliated to the Laboratory for the Study of Metacognition and Advanced Learning Technologies at McGill University, Montreal, Canada. Their interdisciplinary research focuses on designing and testing the effectiveness of adaptive learning environments that are capable of detecting, tracking, modeling, and fostering complex learning processes in the biological and medical sciences.

1550-445X/13 $26.00 © 2013 IEEE
DOI 10.1109/AINA.2013.70

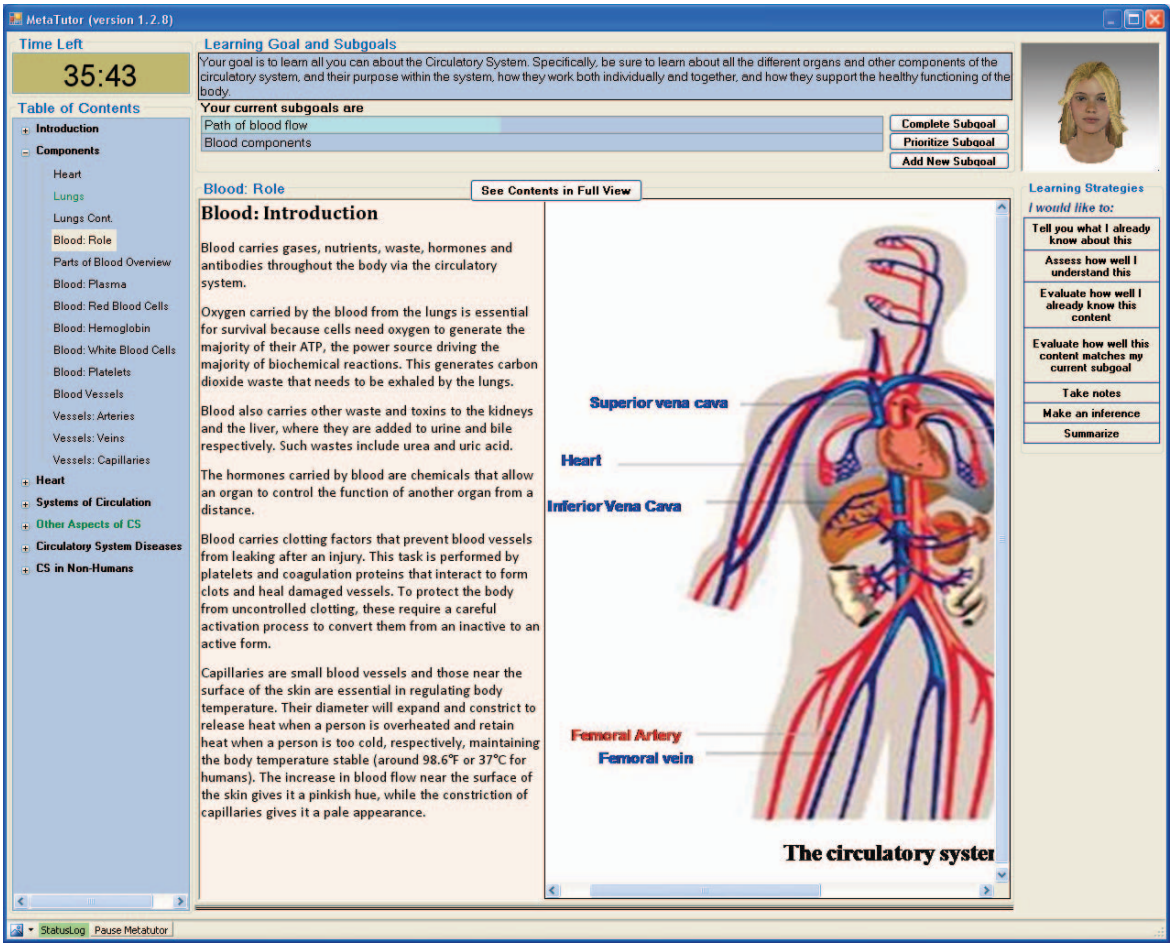

Figure 1. Screenshot of MetaTutor interface.

Section II, we present the related work and specification of different frameworks compared with the one proposed in this paper. Section III explores preliminary details about the MetaTutor system that we use as the basis to equip its developed pedagogical agents with the proposed model. Section IV takes the discussion to theoretical design and formation of MDP that is used with PAs. In Section V, we investigate the log-files and their compatibility to apply MDP structure to enhance SRL overall performance, and finally Section VI concludes the paper and discusses future directions.

II. Related Work

We break down the related work into two different sets where proposed models follow different objectives. In the first set, we highlight SRL-related papers where the objective is mainly enhancing ITS's performance. In [14], [15], authors propose a theoretical framework of SRL where self-regulated learning is explained in deep details. In [3], [5] authors consider SRL as an event and impose different strategic planning to PAs in order to enhance SRL performance. In [6], [18] the authors provide evidence that learners of all ages engage in self-regulation during learning. In [2] authors propose MetaTutor as an example of a multi-agent system designed to foster learning of science topics. In [13], [18] authors use PAs to constrain high performance in maintaining CAMM processes. All the models in the first set are aimed at optimizing and scaffolding students' SRL. They have different assumptions and utilize different environments and techniques to put the objectives into actions. However, they fail to develop a complete theoretical model to provide a systematic computational analysis of the SRL processes using PAs. We would like to advance the state-of-the-art by following an expert approach that defines a systematic infrastructure for MAS-based environments such as MetaTutor to allow applying different strategic intellectual techniques to enhance learners' overall effective use of SRL strategies and performance. We define the system in this paper and provide directions for future work to address the uncertainty problem and relative issues in a systematic sequential fashion.

The related work in the second set are theoreti-

cal/experimental models that use MDPs, POMDPs (Partially Observable Markov Decision Processes), and RL to address similar problems. In [11] the authors survey a considerable amount of research in machine learning and pattern recognition. This work consists of valuable classification of learning methods considering different aspects. In [4], [8] authors express great interest in a number of real-world applications that could be advanced using semi-supervised learning technique. In [9] authors present a number of margin-based online learning algorithms for predicting tasks. Authors consider a sequence of predictions and update steps of different proposed algorithms based on analytical solutions to simple constrained optimization problems. In [16] authors propose an online algorithm for planning uncertainty in multi-agent setting using decentralized-POMDPs. Authors investigate the coordination with little or no communication between agents. The goal of coordination is to ensure that individual decisions of the agents result in (near-)optimal decisions for the group as a whole. Most of the proposed frameworks use MDP and different learning techniques to address a predefined problem. However, the problem might be not realistic, or results might not turn out to be promising. In this paper, we model an actual system, MetaTutor, as MDP and analyse/investigate real data obtained while PAs are interacting with human learners that are voluntarily participating in a two-hour learning session with MetaTutor. We aim to optimize the sequential decision makings PAs conduct while interacting with learners with different characteristics. We explain details of MetaTutor architecture in the next Section as a background required to comprehend the rest of this paper.

III. METATUTOR

MetaTutor [2], [3] is a multi-agent intelligent hypermedia tutoring environment which contains 41 pages of text and static diagrams of the human circulatory system organized by a table of contents displayed in the left pane of the environment (see Figure 1 for an illustration of the learning environment). The underlying assumption of MetaTutor is that students should regulate key cognitive, metacognitive, and affective processes in order to learn about complex and challenging science topics. The non-linear, self-paced structure of the environment allows participants to access content and navigate to new pages by selecting a subtopic from headings located in the table of contents or by using the navigational buttons at the bottom of the screen to progress forward or backward. A timer, located at the top left-hand corner of the environment, displays the amount of time remaining in the session. Four pedagogical agents (Gavin, Pam, Mary, and Sam) are displayed in the upper right-hand corner of the environment. These agents provide varying degrees of prompting and feedback throughout the learning session to scaffold participants SRL skills and content understanding. Each agent serves a different purpose. For example, Gavin the Guide helps participants to navigate through the system; Pam the Planner guides participants in setting appropriate goals; Mary the Monitor helps participants to monitor their progress toward their learning goals; and Sam the Strategizer helps participants to deploy SRL learning strategies, such as summarizing and note-taking. Participants can initiate interactions with these agents and enact specific SRL learning processes by selecting any feature of the SRL palette, which is displayed at the right-hand side of the interface throughout the learning session. Participants can click on buttons on this palette to deploy planning, monitoring, or learning strategies. For instance, by using the SRL palette, participants can take notes or write a summary of the content, test their understanding of the content by assessing their understanding and completing a quiz, activate their prior knowledge of the content, evaluate the relevancy of the content, or make an inference.

Participants overall learning goal and sub-goals are displayed at the top of the interface, and they can select to manage or prioritize their self-set sub-goals. MetaTutor tracks all participant interactions and records user and system actions in a log file. These log files are uploaded to a database, which is then mined for information about participant interactions, including number of times visiting relevant pages, duration of visitation to each page, duration of time viewing images, navigational path, and responses to quizzes and tests. Thus, log files allow for a rich analysis of SRL processes, participant behavior, and comprehension while users interact with MetaTutor. In the next Section, we discuss the formation and use of MDP with PAs where the obtained information from the learner could be used as a means to develop the optimum strategy. In this system, PAs investigate different strategies to effectively cope with learners with a range of knowledge and different attitudes. Using Markov sequential decision makings, PAs follow the best strategy while considering learner's path through progress of SRL.

IV. PEDAGOGICAL AGENTS USING MARKOV DECISION PROCESSES

A. MDP Architecture

Markov Decision Processes (MDPs) provide a mathematical framework for modeling decision making in situations where outcomes are somehow random and partly under the control of a decision maker. MDPs have proved very useful for planning and learning under uncertainty because they study a wide range of optimization problems solved via dynamic programming. MDPs offer a natural extension of these frameworks for operative multi-agent systems. In this paper, we adopt an MDP framework to model MetaTutor PAs to constrain an adaptive dynamic sequential decision making procedure. In this regard, we define PAs as sequential decision makers using MDPs as a mathematical framework to address the decision making processes where outcomes

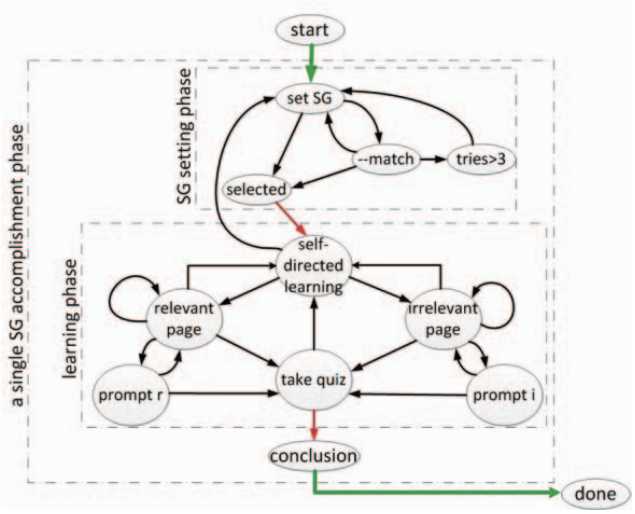

Figure 2. MetaTutor's designed state diagram in accomplishment of a sub-goal (SG).

are partly under control [7]. In fact system components could adopt any unpredictable strategy and are not completely controlled by PAs. In MetaTutor, PAs are effective in helping students to plan and monitor their learning, and use effective learning strategies. However, they do not have control regarding learners' overall progress through the content of the learning environment. Therefore, MDPs are extremely useful in optimizing such complex systems via the use of reinforcement learning techniques [12]. More precisely, MDP is a discrete time stochastic control process. That means at each time step, the process is in some state , and the pedagogical decision maker may choose any action that is available in the current state. The process responds at the next time step by randomly moving into a new state, and giving the decision maker a corresponding reward. The assigned reward is influenced with a range of parameters that are defined in the multi-agent environment (*e.g.* MetaTutor). The probability that the process moves into its new state is influenced by the chosen action. Specifically, it is given by the state transition function. Thus, the next state exclusively depends on the current state and the decision maker's action. But given and, it is conditionally independent of all previous states and actions; in other words, the state transitions of an MDP possess the Markov property (that we represent in the Definition 1). In the following, we formally define a MDP.

Definition 1 (MDP). A Markov decision process is a tuple $< I, S, \{A_i\}, \{\Omega_i\}, P, R, b^o >$ where

- I is a finite set of agents indexed $1, ..., n$. Notice that when $n = 1$, a MDP is equivalent to a single-agent MDP. In MetaTutor we have four agents which with the learner makes $n = 5$.
- S is a finite set of system states. A state summarizes all the relevant features of the dynamical system and satisfies the Markov property. That is the probability of the next state depends only on the current state and the joint action, not on the previous states and the joint actions:

$$P(s^{t+1}|s^0, a^0, ..., s^{t-1}, a^{t-1}, s^t, a^t) = P(s^{t+1}|s^t, a^t)$$

We consider the Markov property to omit error might come through when analysing the previous histories of interactions. In fact, we use previous data while proposing the best action in the current state. That means, over time PAs might change their action while being influenced with learner's characteristics, prior knowledge, and attitude. Figure 2 provides a preliminary state diagram that is used to assist a learner to accomplish a sub-goal (SG) set earlier in the learning session with MetaTutor. This could be carried out using a sequence of SGs. The state diagram mainly consists of two separate parts associated with the SG setting phase and the learning phase. When finished learning about a SG, the learner takes the generated quiz to conclude the SG and moves on to the next one.

- A_i is the finite set of actions available to pedagogical agent i and $A = X_{i \in I} A_i$ is the set of joint actions, where $a = (a_1, ..., a_n)$ denotes a joint action. We assume agents do not observe which actions are taken by others at each time step. In fact, we abstract the set of actions just to the joint action to avoid losing the generality of MetaTutor and steps in state transitions. We are mainly interested in managing the optimum SRL process, and improving the coordination and cooperation of the PAs. In Table I, we list the set of actions on the top row. The Table represents the initial (no observation yet) transition probability that PAs (as a group) might use while they are in different states.
- Ω_i is the finite set of observations available to agent i and $\Omega = X_{i \in I} \Omega_i$ is the set of joint observations, where $o = (o_1, ..., o_n)$ denotes a joint observation. We assume each agent observes its own component at each time step.
- P is the transition probability table. $P(s'|s, a)$ denotes the probability that taking joint action a in state s results in transition to state s'. P describes the stochastic influence of actions on the environment. We assume that the transition probabilities are stationary, which means that they are independent of the time step.
- $R : S \times A \to R$ is a reward function. $R(s, a)$ denotes the reward obtained from taking joint action a in state s. R specifies the agents' goal. It is an immediate reward for agents taking the joint action in some state. The agents do not generally observe the immediate reward at each time step, although their local observation may determine that information.
- $b^o \in \Delta(S)$ is the initial belief state distribution. It is a

Table I
METATUTOR'S INITIALIZED STATE/ACTION TRANSITION PROBABILITY.

	Reload	Complete	Refuse SG	Accept SG	Encourage Reselect	Warn Low Performance	Accept Reading	Refuse Reading
Start	40%	60%	0	0	0	0	0	0
Set SG	20%	0	35%	35%	0	10%	0	0
Not Matched	20%	0	20%	20%	20%	20%	0	0
tries > 3	20%	0	0	0	50%	30%	0	0
SG Selected	10%	80%	5%	0	0	5%	0	0
S-D Learning	10%	0	0	0	5%	5%	40%	40%
Relevant Page	30%	10%	0	0	0	5%	40%	15%
Irrelevant Page	30%	0	0	0	10%	15%	15%	30%
Prompts	20%	40%	0	0	5%	5%	15%	15%
Quiz	10%	50%	0	0	30%	10%	0	0
Conclusion	50%	0	0	0	40%	10%	0	0

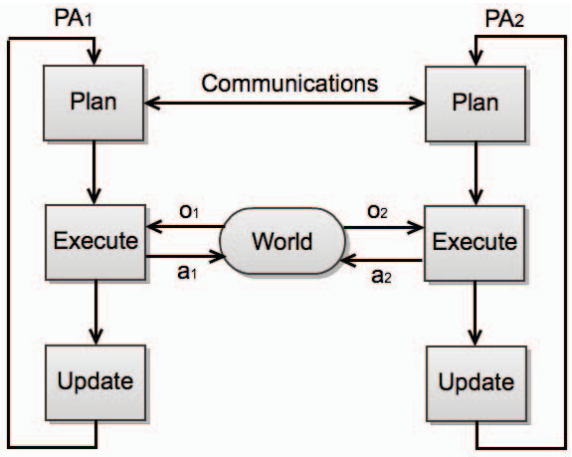

Figure 3. MetaTutor's online planning framework for two PAs.

vector that specifies a discrete probability distribution over S that captures the agents' common knowledge about the starting state.

In MetaTutor, PAs have full source of information about other PAs in the learning environment. However, they have limited source of information regarding the observations that other agents had during their interaction with the learner. To this end, PAs must reason about the histories of other PAs and how these histories affect their own interactions. Agents' communication can resolve this problem by sharing private information such as sensory data. Communication is used to improve the performance of PAs as a group. In MetaTutor, we consider direct communication and use the sync communication model [17] where each pedagogical agent broadcasts a part of its local information (mainly consist of its new observations) to other agents. The communication language simply allows transmission of the agents' actions-observation histories. Designing a general communication strategy, we determine when to communicate, with whom to communicate, and what to send. MetaTutor's detailed features are compatible to sync communication language. The communication is started by Gavin (the pedagogical guide agent who starts interacting with the learner and transfers the observations to the others for better decision makings). The communication is assumed to be instantaneous. That means a message is received without any delay. In Figure 3, we explore more details about how the two agents can instantaneously communicate and plan while each one is using its own decision making procedure. Online planning is executed in parallel by all PAs, interleaving planning and execution.

B. Sequential Decision Making

In the previous Subsection, we formally defined the Markov decision process that is compatible with MetaTutor's infrastructure with pedagogical agents. In this Subsection, we explore more details about the sequential decision making that PAs maintain over the MetaTutor learning session. So far, we know the set of states S and actions A. We also know about the state/action probability transitions. We now need to define the reward function that evaluates the reward R associated with state transition (s to s') via a particular action (a). In MetaTutor, PAs' interactions with human learners and states/actions/transitions could be compared with the ones associated to the ideal learner representing the case where the learner navigates through system states with the best (*e.g.*, most effective in time and learning management) strategy. The reward R (see Equation 1) would be always the portion of current progress pg_L over the ideal process obtained from the ideal learner pg_I^*. The progress pg_L is a function of current time t, learner's knowledge k, his/her goal of learning g, and the efficiency e towards reaching the goal. These parameters are continuously updated throughout the learning session. In function f, we compare the goal/knowledge status proportional to the remaining time where T denotes the total SRL session time. The parameter e is rated via PAs during the learning session and denotes the efficiency in learning the content.

$$R = R_L = \frac{pg_L}{pg_I^*} \tag{1}$$

where

$$pg_L = f_L(t, k, g, e) = \frac{g - k}{T - t} \times e$$

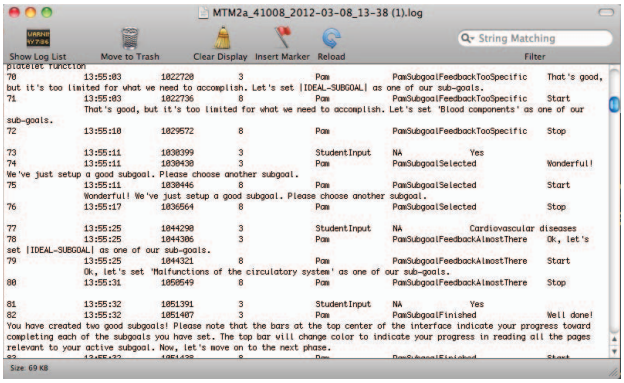

Figure 4. A sample log-file of a learner while interacting with PAs. In this part, a SG is ideally selected.

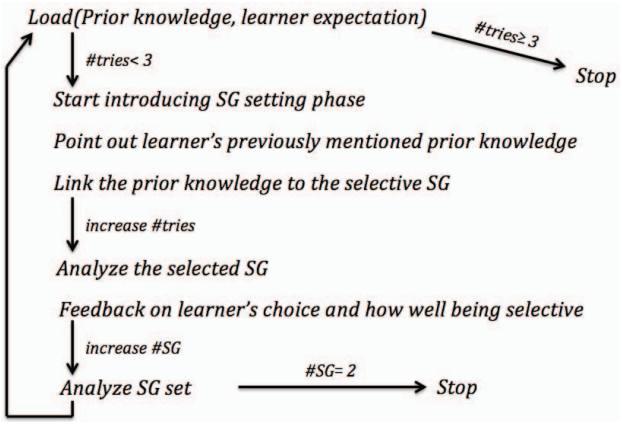

Figure 5. A solution plan for SG setting phase.

Considering the reward function obtained from Equation 1, we can reward the state/action transition to reward the state transition to s' from state s while taking action a ($R(s, a, s')$). This reward rates the movement maintained towards the predefined objective in the learning session. So the path ($pt(s_1, a_1, s_2, a_2, ..., s_t, a_t, s_{t+1})$) with the length of t could be also rewarded. Equation 2 evaluates the reward associated to a particular path of state/action transition in SRL session. Parameter μ denote the time discount factor ($\mu \in [0, 1)$) to promote the recent state/transition reward. The idea is to prevent equally considering rewards associated to different times of SRL session.

$$R(pt) = \sum_{i=1}^{t} \mu^{i-1} \times R(s_i, a_i, s_{i+1}) \quad (2)$$

In sequential decision making, PAs follow a policy Θ that maximizes the rewards of the navigation path following Θ. Equation 3 evaluates the expected return dedicated to policy Θ.

$$E(\Theta) = \sum_{pt \in PATH} p(pt|\Theta) \times R(pt) \quad (3)$$

where

$$p(pt|\Theta) = p(s_1) \prod_{i=1}^{t} p_\Theta(s_{i+1}|s_i, a_i)$$

The objective is to find an optimum policy Θ^* that maximizes the expected return ($E(\Theta)$). This is maintained via different strategy analysis and is a form of optimization problem with its corresponding constraints. In MetaTutor, PAs investigate different policies that influence the state/action navigation path (pt) and enhance their adopted policy expected value over time. We investigate different strategies and let the optimization problem find the best reward over the sequential decision making processes maintained by PAs. In the following Section, we extend more details about the solution plans we provided for MetaTutor for optimum decision making procedures.

V. Observatios

The analyses of the log-file data explores details in different aspects of individual learner's interactions with MetaTutor. We map this file (see Figure 4) to the proposed state diagram and highlight the state transitions. We use the log-file data as an additional resource to the reinforcement learning technique that PAs are equipped with. We anticipate that results will show how effective MDP would be in enhancing the quality of learning in multi-agent environments.

The results and ensuing proposed architecture will demonstrate the advantage of using intelligent, multi-agent computerized learning environments to advance our understanding of what, how, when, and why students' know and feel during complex learning episodes. The data sources will provide evidence that has the potential to advance current conceptual, theoretical, methodological, and analytical frameworks related to SRL processes. These advances will in turn allow researchers to design more effective multi-agent learning environments that are sensitive and responsive to students' cognitive, metacognitive, affective, and motivational needs during learning. In addition, the proposed architecture has the potential to advance current systems by: emphasizing a coordinated effort in understanding and sharing CAMM processes across agents; coordinating efforts in sharing data about individual students as key in; fostering the right amounts and types of scaffolding, feedback, and adaptivity (for each CAMM process and across CAMM processes); facilitating the development of a more accurate student model; leading to more accurate inferences about students' understanding of the content and ability to regulate certain aspects of their learning.

Using the proposed architecture, and with the study on the log data, we provide solution plans in algorithmic forms to enhance the performance while using Markov decision processes. As shown in Figure 5, in sub-goal setting phase, there are a number of states where we might encounter. This depends on learner's selected actions and strategic behaviors.

Dealing with this various state transitions, we impose some restrictions to narrow down the choices. In this algorithm, we mainly highlight the role of number of tries, which represent learner's attempts in selecting an appropriate sub-goal. We lead the state navigator path to visit to states where the learner needs more directions and assistance, while the number of attempts in sub-goal selection has exceeded the predefined threshold. In this solution plan, learner's selected sub-goal is analysed and the corresponding feedback is provided via the interfering pedagogical agent. This phase is finalized when the learner has successfully selected two relative sub-goals.

The solution plan proposed in Figure 5 has limitations with respect to unexpected issues where the learner might by chance has selected an appropriate sub-goal and has difficulty matching another relevant sub-goal. We have extended the proposed solution plan in Figure 6 where we provide more flexible state navigation system. In the extended algorithm, the interfering pedagogical agent starts by highlighting sub-goal setting strategies and upon recipe of a proposal, the agent provides detailed feedback on the proposed sub-goal and the selecting strategy. In case the proposal was not relevant, the agent would explain more about the sub-goal strategies of selection, and leads the learner to either reselect a sub-goal or to the suggested sub-goal. The same idea would be applied to the second proposed sub-goal until learner can successfully finalize two relative sub-goal setting phase. The provided algorithms enhance the decision making performance while each action is analysed with respect to the associated reward. We essentially obtain an optimum navigation path where states associated to different parts of the system are visited.

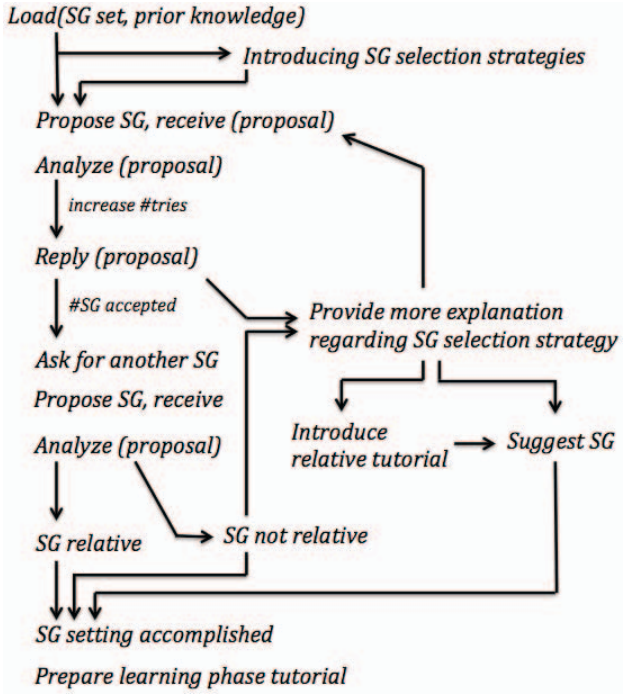

Figure 6. Extended solution plan for SG setting phase.

VI. Conclusion

This paper provides a proposition for a novel multi-agent system design for MetaTutor as an agent-based learning environment where pedagogical agents coordinate with one another to facilitate SRL processes. The proposed model points out Markov decision processes and how they are implemented in MetaTutor with four pedagogical agents. Our objective is to enhance the efficiency of the system as a whole, in terms of scalability, complexity, and data analysis. We apply reinforcement learning technique to capture best features where agents obtain best reward as a pay off to their strategic behaviors. We also extended our discussion of agents' mental state to regulate their inter-communication protocols. Our model proposes the following advantages: (1) the characteristics of a Markov decision processes are stated and implemented in a real world application of MetaTutor to facilitate self-regulated learning; (2) we proposed an agent communication protocol that enhances system performance in analysing large scale data obtained from interacting with different learners over two-hour sessions; and (3) our reinforcement learning algorithm helps the agents' communication protocol to effectively extend their knowledge representation and cooperation between different agents.

In future research, we plan to extend our study towards a performance analysis of MetaTutor, where four pedagogical agents could effectively interact with a human learner and one another. We intend to propose different mechanisms to develop adaptive multi-agent communication and decision making to represent an optimally efficient MetaTutor to facilitate self-regulatory processes. We also intend to apply different machine learning techniques to use previous experience in agents' further decision making procedures.

Acknowledgement

The research presented in this paper has been supported by funding from the Fonds de recherche du Qubec - Nature et technologies (FQRNT) awarded to the first author and National Science Foundation, Institute of Education Sciences, and Social Sciences and Humanities Research Council of Canada awarded to the third author.

References

[1] Azevedo, R., Moos, D., Johnson, A., and Chauncey, A. (2010). Measuring cognitive and metacognitive regulatory processes used during hypermedia learning: Issues and challenges. Educational Psychologist, 45, 210-223.

[2] Azevedo, R., Witherspoon, A., Chauncey, A., Burkett, C., and Fike, A. (2009). MetaTutor: A meta- cognitive tool for enhancing self-regulated learning. In R. Pirrone, R. Azevedo, G. Biswas, (eds.), Proceedings of the AAAI Fall Symposium

on Cognitive and Metacognitive Educational Systems (pp. 14-19). Menlo Park: Association for the Advancement of Artificial Intelligence (AAAI) Press.

[3] Azevedo, R., Behnagh, R., Duffy, M., Harley, J., and Trevors, G. (2012). Metacognition and self-regulated learning in student-centered leaning environments. In D. Jonassen and S. Land (Eds.), Theoretical foundations of student-center learning environments (2nd ed.)(pp. 171-197). New York: Routledge.

[4] Bandos, T., Zhou, D., and Camps-valls, G. (2006). Semi-supervised hyperspectral image classification with graphs. In IEEE International Conference on Geoscience and Remote Sensing Symposium, pp. 3883-3886, 2006.

[5] Biswas, G., Jeong, H., Kinnebrew, J., Sulcer, B., and Roscoe, R. (2010). Measuring self-regulated learning skills through social interactions in a Teachable Agent Environment. Research and Practice in Technology-Enhanced Learning, 5(2), 123-152.

[6] Calvo, R. A. and D'Mello, S. K. (Eds.) (2011). New Perspectives on Affect and Learning Technologies. New York:Springer.

[7] Carlin, A., and Zilberstein. S. (2011). Decentralized monitoring of distributed anytime algorithms. In M.Rovatos (Ed.), Proceedings of the Tenth International Conference on Autonomous Agents and Multi-agent Systems (AAMAS), (pp. 157-164), Taipei, Taiwan: Inter- national Foundation for AAMAS.

[8] Cherniavsky, N., Laptev, I., Sivic, J., and Zisserman, A. (2010). Semi-supervised learning of facial attributes in video. In First International Workshop on Parts and Attributes, in conjunction with European Conference on Computer Vision, 2010.

[9] Crammer, K., Dekel, O., Keshet, J., Shalev-Shwartz, S., and Singer, Y. (2006). Online passive-aggressive algorithms. Journal of Machine Learning Research: 7(2006):551-585.

[10] Feyzi-Behnagh, R., Trevors, G., and Azevedo, R. (2012, September). Metacognition in multimedia: A micro-Analysis of process and judgment data. Paper to be presented at the Biennial Meeting of the European Association for Research on Learning and Instruction (EARLI) Metacognition SIG, Milan, Italy.

[11] Singh, A., Nowak, R.D., and Xhu, X. (2008). Unlabeled data: Now it helps, now it doesn't. In Advances in Neural Information Processing Systems. In Proceedings, pp. 1513-1520, 2008.

[12] Ross, S., Chaib-draa B., and Pineau J. (2008). Bayesian reinforcement learning in continuous POMDPs with application to robot navigation. International Conference on Robotics and Automation (ICRA), pp. 2845-2851. Jochen Triesch: IEEE.

[13] Winne, P.H., and Nesbit, J.C (2009). Supporting self-regulated learning with cognitive tools. In D.J. Hacker, J. Dunlosky, A. C. Graesser (eds.), Handbook of metacognition in education. Mahwah: Erlbaum.

[14] Winne, P. H., and Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, and A. C. Graesser (Eds.), Metacognition in educational theory and practice (pp. 277-304). Mahwah, NJ: Lawrence Erlbaum Associates.

[15] Winne, P. H., and Hadwin, A. F. (2008). The weave of motivation and self-regulated learning. In D. H. Schunk and B. J. Zimmerman (Eds.), Motivation and self-regulated learning: Theory, research, and applications (pp. 297-314). Mahwah, NJ: Lawrence Erlbaum Associates.

[16] Wu, F., Zilberstein, S., and Chen, X. (2011). Online planning for multi-agent systems with bounded communication. Artificial Intelligence, 175(2):487-511, 2011.

[17] Xuan, P., Lesser, V., and Zilberstein, S. (2001). Communication decisions in multi-agent cooperation: Model and experiments. In Proceeding of the 5th International Conference on Autonomous Agents, pp. 616-623.

[18] Zimmerman, B., and Schunk, D. (2011). Handbook of self-regulation of learning and performance. New York: Routledge.