In Proceedings of the Sixth International Conference on Cognitive Modeling, 320-325.
Mahwah, NJ: Lawrence Earlbaum.

Modeling Psychological Refractory Period (PRP) and Practice Effect on PRP with Queuing Networks and Reinforcement Learning Algorithms

Changxu Wu (changxuw@umich.edu)
Department of Industrial & Operations Engineering, University of Michigan
1205 Beal Avenue, Ann Arbor, MI, 48109 USA

Yili Liu (yililiu@umich.edu)
Department of Industrial & Operations Engineering, University of Michigan
1205 Beal Avenue, Ann Arbor, MI, 48109 USA

Abstract

PRP (Psychological Refractory Period) is a basic but important form of human information processing in dual-task situations. This article describes a queuing network model of PRP that successfully modeled PRP without the need of setting up complex lock/unlock performance strategies employed in the EPIC model of PRP or drawing complex scheduling charts employed in the ACT-R/PM model of PRP. Further, by integrating queuing networks with reinforcement learning algorithms, the queuing network model successfully simulated practice effect on PRP, which has not been modeled in existing PRP models. The current research indicates that depending on an individual's degree of practice, cognition can be either serial or parallel at the level of production or response selection. Extensions of queuing network model in modeling other tasks and its easiness in modeling concurrent tasks in practice are also discussed.

Introduction

PRP (Psychological Refractory Period) is one of the most basic and simple forms of dual task situations, and has been studied extensively in the laboratory for half a century (Meyer & Kieras, 1997a). In the basic PRP paradigm, two stimuli are presented to subjects in rapid succession and each requires a quick response. Typically, responses to the first stimuli (Task 1) are unimpaired, but responses to the second stimuli (Task 2) are slowed by 300 ms or more (Ruthruff et al., 2001). In the PRP paradigm of Van Selst's study (1999), in task 1, subjects were asked to discriminate tones into high or low tones by vocal responses (audio-vocal responses); in task 2, subjects watched visually presented characters and performed choice reaction task by manual responses (visual-motor responses). They found that practice dramatically reduced dual-task interference in PRP.

The basic PRP has been modeled by several major computational cognitive models based on production rules, notably EPIC (Meyer & Kieras, 1997a) and ACT-R/PM (Byrne & Anderson, 2001). Based on the major assumption that production rules can fire in parallel, EPIC successfully modeled the basic PRP effect by using complex lock and unlock strategies in central processes to solve the time conflicts between perceptual, cognitive and motor processing (Meyer & Kieras, 1997a). In contrast to the parallel cognitive processing assumption of EPIC, ACT-R/PM assumes serial firing of production rules in cognition. It modeled the basic PRP effect by relying on drawing scheduling charts manually to quantify the quick switching of central cognition in information processing, which shares certain similarities with the lock and unlock strategies in EPIC (Byrne & Anderson, 2001). All of these complex lock/unlock strategies or scheduling charts came from the arrangements by the researchers rather than the natural interactions among the processors or modules. Moreover, because of the complexity of dual tasks in practical situations, it may be difficult for a system designer or human factor specialist to figure out these complex strategies or scheduling charts to arrange the activities of the central cognitive processor.

Furthermore, neither EPIC nor ACT-R/PM modeled the practice effect on PRP (Ruthruff et al., 2001).

The current paper describes a model of PRP that integrates queuing network theory (Liu, 1996, 1997) and reinforcement learning algorithms (Sutton & Barto, 1998) and was developed on the basis of neuroscience findings. Model simulation results were compared with experimental results of both the basic PRP paradigm and the practice effect on PRP (Van Selst et al., 1999). All of the simulated human performance came from the natural interactions among servers and entities in the queuing network without setting up lock and unlock strategies or drawing complex scheduling charts.

Modeling the Basic PRP and the Practice Effect on PRP with Queuing Networks

Modeling Human Performance with Queuing Networks

A queuing network is a network of servers and queues that allow two or more servers to act serially, in parallel, or in any network arrangement. Computational models based on queuing networks have successfully integrated a large number of mathematical models in response time (Liu, 1996) and in multitask performance (Liu, 1997) as special cases of queuing networks. A queuing network modeling architecture called the Queuing Network – Model Human Processor (QN-MHP) has been developed and used to generate behavior in real time (Liu, Feyen & Tsimhoni, 2004), including simple and choice reaction time (Feyen & Liu, 2001), driver performance (Tsimhoni & Liu, 2003) and transcription typing (Wu & Liu, 2004a, 2004b). The model in this paper extends QN-MHP by integrating reinforcement learning algorithms and strengthening its long-term memory. The simulation model is easy to use and is implemented with Promodel®, one of the most popular simulation software programs in industry.

Because major brain areas with certain information processing time and capacity are localized and connected with each other in brain cortex via neural pathways (Bear & Connor, 2001; Smith et al., 1998; Roland, 1993, see Figure 1 and Table 1), it is assumed that they form a queuing network. One or several brain areas are regarded as servers of the network with neural pathways regarded as routes and information to be processed in these brain areas regarded as the entities. Cognitive performance is the outcome of processing the entities in the queuing network.

Modeling the Basic PRP and Practice Effect on PRP with Queuing Networks and Reinforcement Learning Algorithms

Because the PRP effect prior to or at beginning of learning (the basic PRP) is a special case of the PRP effect during the learning process, the two phenomena of PRP (basic and learning) are modeled with the same mechanism in our queuing network model. The experimental tasks and data of Van Selst et al. (1999) were used to test the model.

Brain areas (servers) and their routes related to the two PRP tasks in Van Selst's study were identified within the general queuing network structure based on recent neuroscience findings (Mitz et al., 1991; Fletcher et al., 2001; Bear & Connor 2001, see Figure 1).

Entities of task 1 (audio-vocal responses) cannot bypass the Hicog server because the function of phonological judgment is mainly located at the Hicog server, and thus there is only one possible route for the entities of task 1 (see the dotted thick line in Figure 1). However, the function of movement selection in task 2 (visual-motor responses) is located not only in the Hicog server but also in PM server. Therefore, there are two possible routes for the entities of task 2 starting at Visk server (see the gray and black solid lines in Figure 1).

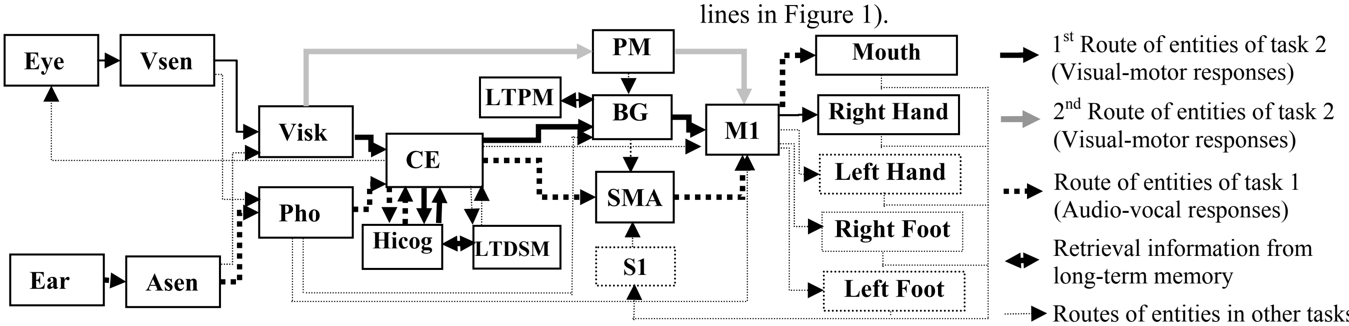

Figure 1: The general structure of the queuing network model (QN-MHP) with servers and routes involved in the PRP task highlighted (server names, brain structures, major functions and processing time are shown in Table 1)

Table 1: Server name, major function and brain structure (Capacity is mainly based on Card et al., 1986; Liu, 1997. Processing time is based on Rudell & Hu, 2001; Feyen, 2001; Romero, 2000; Nakahara et al., 2001)

Server(Capacity)	Brain Structure	Major function (Processing logic)	Processing Time
Eye (∞)	Eye, LGN, SC, visual pathway	Visual sampling & signal transmission	Eye movement: 50ms
VSen (17 letters)	Distributed parallel area, SFS, dorsal & ventral system	Visual sensory memory & perception	Normal distribution: M=263ms, SD=11ms
Visk (3 chunks)	Right posterior parietal cortex	Visuospatial sketchpad to store the graphic information	Included in CE server
Ear (∞)	Ear, auditory pathway	Converts sound waves to neuron signals	10 ms
Asen (5 entities of tones)	Primary auditory cortex & planum temporale	Auditoral sensory memory and perception	50 ms
Pho (3 chunks)	Left posterior parietal cortex, inferior parietal lobe	Phonological loop to store auditoria & textual information	Included in CE server
CE (3 chunks)	Dorsal lateral prefrontal cortex & ACC	Mental response inhibition & selection, improve processing rate of premotor cortex via parallel learning mechanism	Exponential distribution: Mean=70ms, Min=25ms
Hicog (1 chunk)	Left interior partial cortex, IPS and VLFC	Phonological judgment, visuomotor choices by retrieving production rules at LTDSM	[a] $A_i+B_i \operatorname{Exp}(-\alpha_i N_i)$
LTDSM (∞)	Hippocampus	Long-term declarative memory (e.g. production rules in judgments and choices) and spatial memory	Included in Hicog
PM (1 chunk)	Premotor cortex (BA 6)	Select movement in learning visuomotor association	[a] $A_i+B_i \operatorname{Exp}(-\alpha_i N_i)$
BG (1 chunk)	Basal ganglia	Motor program retrieval	[a] $A_i+B_i \operatorname{Exp}(-\alpha_i N_i)$
LTPM (∞)	Striatal and cerebellar systems	Long term procedural knowledge storage	Included in BG server
SMA (2 letters)	Supplementary motor area & pre-SMA	Motor program assembly, error detection	188 ms
M1 (2 letters)	Primary motor cortex	Addressing spinal motorneourons	70 ms
Mouth (1 letter)	-	Converts neuron signal to movement of vocal organ	10 ms
Hand (1 letter)	-	Execution of motor movement	[b] $I_m \log_2(D/S+0.5)$

a. See the part "learning process of individual servers" in this paper. b. Fitt's Law: I_m=26.3,D=1.3cm, S: movement distance.

However, how do the entities of task 2 choose one of the two alterative routes in the network? What is the behavioral impact of this choice on PRP and the practice effect on PRP? This can be answered by integrating queuing networks with reinforcement learning algorithms.

Before exploring the mechanism with which entities of task 2 select from the 2 routes, it is necessary to understand the learning process of individual brain areas. It was discovered that each individual brain area reorganizes itself during the learning process and increases its processing speed (Ungerleider, 2002). For example, for the simplest network with 2 routes (see Figure 2), if servers 2 and 3 change their processing speeds, different routes chosen by an entity (1→3→4 or 1→2→4) will lead to different performance. Without considering the effect of error, entities will choose the optimal route with the shortest processing time if they want to maximize the reward of performance.

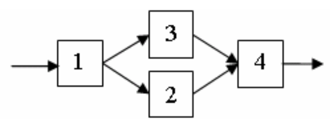

Figure 2: The simplest queuing network with 2 routes

Consequently, to model learning, it is first necessary to quantify the learning process in individual servers. Based on that, the condition under which an entity switches between the 2 routes shown in Figure 2 can be established and proved by integrating queuing network with reinforcement learning. Finally, this quantitative condition of route switching can be applied to the more complex model of 18 servers with 2 routes (see Figure 1) to generate the basic PRP and reduction of PRP during the learning process.

Learning Process in Individual Servers Based on the function of the servers in Table 1, the two long-term memory servers (LTDSM and LTPM) play the major roles in learning phonological judgments (task 1) and choice reaction (task 2) (Bear & Connor, 2001). Because the learning effects of long-term memory are represented as speed of retrieval of production rules and motor programs from the two long-term memory servers at Hicog and BG servers, it is important to quantify the processing time of Hicog and BG servers. In addition, because premotor cortex (PM) server is activated in learning visuomotor association (Mitz et al., 1991), change of the processing speed of PM server is also to be considered in the learning process of the model.

Because exponential function fits the learning processes in memory search, motor learning, visual search, and mathematic operation tasks better than power law (Heathcote et al., 2000), it is applied to model the learning process in the individual servers (Equation 1).

$$1/\mu_i = A_i + B_i \operatorname{Exp}(-\alpha_i N_i) \qquad (1)$$

μ_i: processing speed of the server i; ($1/\mu_i$) is its processing time; A_i: The minimal of processing time of server i after intensive practice; B_i: The change of expected value of processing time of server i from the beginning to the end of practice; α_i: learning rate of server i; N_i: number of customers processed by server i.

For BG server, $1/\mu_{BG}$: motor program retrieving time; A_{BG}: the minimal of processing time of BG server after practice (314 ms, Rektor et al., 2003); B_{BG}: the change of expected value of processing time from the beginning to the end of practice (2*314=628 ms, assumed). α_{BG}: the learning rate of server BG (0.00142, Heathcote et al., 2000); N_{BG}: number of entities processed by server BG which is implemented as a matrix of frequency recorded in LTPM server.

For Hicog and PM servers, to avoid building an ad-hoc model and using the result of the experiment to be simulated directly, nine parameters in Hicog and PM servers were calculated based on previous studies (see Appendix 1).

Learning Process in the Simplest Queuing Network with 2 routes Based on the learning process in individual servers, the condition under which an entity switches between the 2 routes in the simplest form of queuing networks with 2 routes (each capacity equals 1) (from route 1→2→4 to route 1→3→4, see Figure 2) was quantified and proved by the following mathematic deduction.

1) Q Online Learning Equation (Sutton & Barto, 1998).

$$Q^{t+1}(i,j) \leftarrow Q^t(i,j) + \varepsilon \{r_t + \gamma \max_k [Q^t(j,k)] - Q^t(i,j)\} \qquad (2)$$

- $Q^{t+1}(i,j)$: online Q value if entity routes from server i to server j in t+1 th transition
- $\max_k[Q(j,k)]$: maximum Q value routing from server j to the next k server(s) (k>=1)
- $r_t=\mu_{j,t}$: reward is the processing speed of the server j if entity enters it at t th transition
- N_{jt}: number of entities goes to server j at t th transition;
- ε: learning rate of Q online learning ($0< \varepsilon<1$)
- γ: discount parameter of routing to next server ($0<\gamma<1$)
- p: probability of entity routes from server 1 to server 3 do not follow the Q online learning rule if Q (1,3)> Q (1,2). For example, if p=0.1, then 10% of entity will go from server 1 to server 2 even though Q (1,3)> Q (1,2).

State is the status that an entity is in server i; Transition is defined as an entity routed from server i to j. Equation 2 updates a Q value of a backup choice of routes ($Q^{t+1}(i,j)$) based on the Q value which maximizes over all those routes possible in the next state ($\max_k[Q(j,k)]$). In each transition, entities will choose the next server according to the updated $Q^t(i,j)$. If Q (1,3)>Q (1,2), more entity will go from server 1 to server 3 rather than go to server 2.

2) Assumption

- ε is a constant which do not change in the current learning process ($0< \varepsilon<1$)
- Processing speed of server 4 (μ_4) is constant

3) Lemma 1. At any transition state t ($t\neq0$), if $1/\mu_{2,t}< 1/\mu_{3,t}$, then $Q^{t+1}(1, 2)>Q^{t+1}(1, 3)$

Proof of Lemma 1 (see Appendix 2).

Based on Lemma 1 and equation 1, we got Lemma 2:

4) Lemma 2. At any transition state t ($t\neq0$), if $A_2+B_2\operatorname{Exp}(-\alpha_2 N_{2t}) < A_3+B_3\operatorname{Exp}(-\alpha_3 N_{3t})$, then $Q^{t+1}(1,2)>Q^{t+1}(1,3)$

Prediction of the Basic PRP and the Practice Effect on PRP with the Queuing Network Model Based on equation 1 and Lemma 1 and 2, we can predict the possible simulation result of PRP and PRP practice effect.

For the entities in task 2 (see Figure 1), at the beginning of the practice phrase, because the visual-motor mapping is not ready in PM (Mitz et al., 1991), PM takes a longer time to process the entities than the CE and Hicog. Thus, the Q value from Visk to PM (Q (1,3)) is lower than the Q value from Visk to CE (Q (1,2)). According to Lemma 1, the majority of the entities will go to the CE and Hicog server at the beginning of the learning process in dual tasks. Consequently, because entities from task 1 also go through the CE and Hicog server, a bottleneck at the Hicog server would form at this stage which produces the basic PRP effect.

During the learning process, CE will send dummy entities which increase the processing speed of PM based on the parallel learning mechanism between visual loop (including CE) and motor loop (including PM) (Nakahara et al., 2001, see Table 1). Therefore, when Q value of the 2^{nd} route of task 2 increases, more and more entities of task 2 will travel in the 2^{nd} route, forming an automatic process and two parallel routes in this dual-task situation. However, because the learning rate of PM server (1/16000) is lower than that of Hicog server for the entities in task 2 (1/4000), the majority of the entities will still go through the Hicog server.

Simulation Result

Figure 3 shows the simulation result of the basic PRP effect compared with the experiment data (Van Selst et al., 1999). The linear regression function relating the simulation and experiment results is: Y =1.057X-58 (Y: Experiment Result; X: Simulated Result; R square =.984, p<.001; sig. of constant > .05).

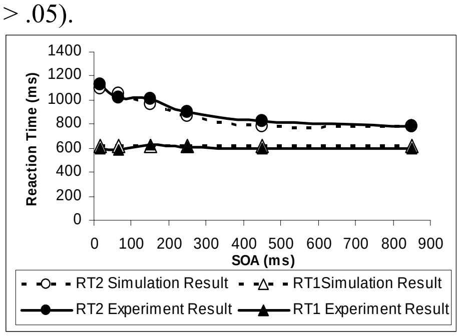

Figure 3: Comparison of simulation and experiment results at the beginning of practice (basic PRP effect)

Figure 4 compares of simulation and experiment result of PRP effect at the end of practice (after 7200's trials). The linear regression function relating the simulated results and experiment results is: Y=1.03X+105 (R square =.891, p<.001; sig. of constant > 0.05).

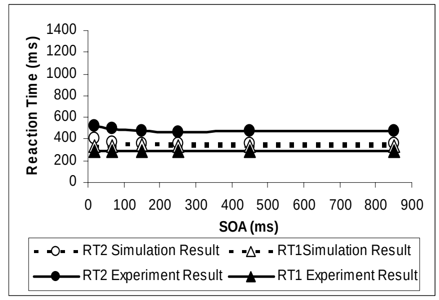

Figure 4: Comparison of the simulation and experiment results at the end of practice

Figure 5 shows the comparison of simulation and experiment results during the practice process (7200 trials). The linear regression function relating the simulated results and experiment results is: Y=0.965X +10 (R square =.781, p<.001; sig. of constant > 0.05). Moreover, it was found that the Q value of the 2^{nd} route of task 2 never exceeded that of the 1^{st} route of task 2 during the practice process and the majority of entities of task 2 went through the 1st route rather than the 2^{nd} route.

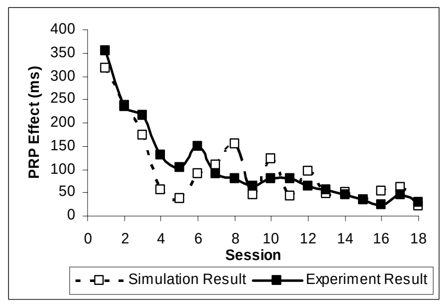

Figure 5: Comparison of simulation and experiment results during the practice process (7200 trials)

Discussion

Integrating queuing networks and reinforcement learning algorithms, the extended version of QN-MHP simulated both basic PRP effect and the practice effect on PRP in Van Selst (1999)'s study without setting up lock and unlock strategy or drawing complex scheduling chart.

With the formation of an automatic process during learning, two parallel routes were formed in the dual-task situation, which partially eliminated the bottleneck at the Hicog server. The PRP effect is reduced greatly with the decrease of the processing time in both the Hicog and the PM server. However, since the majority of the entities of the two tasks still went through the Hicog server, the effect of the automatic process on PRP reduction does not exceed the effect of the reduction of RT 1 on the PRP. This is consistent with the result of Van Selst (1999) that automatic process does grow from weak to strong but only contributes small PRP reduction.

The current model indicates that depending on an individual's degree of practice, cognition can be either serial or parallel at the level of production or response selection. At the beginning of practice, a bottleneck at the Hicog server within the central cognition is critical for the generation of PRP effect, which is consistent with the assumption of ACT-R/PM that cognition must be serial without considering the effect of practice. However, the formation of a parallel process in production selection after practice indicates that there are two response or production selection processors running in parallel in the cognitive system in the well-learned situation, which is different from the basic assumption in ACT-R/PM. This parallel processing in the well-learned situation in the current model is supported by Van Selst (1999)'s study and the review of fMRI study by Collette (2001) who found that in dual task situations not only the brain area of CE server (BA9/46: DLPFC & ACC) but also the area of well-learned visuomotor control (BA 6: premotor area, PM server) were activated simultaneously. In addition, it is also important to note that the formation of two parallel processing routes in the current model is different from the assumption and mechanism of EPIC because the two routes were formed with reinforcement learning naturally without using any unlock and unlock strategies.

Because the queuing network model was built in a general structure of human information processing, it can be easily transformed to model other task situations involving concurrent activities in practical situations, e.g. dual tasks in transcription typing (Wu & Liu, 2004b). In contrast to complex unlock and lock strategies in EPIC or drawing complex scheduling charts in ACT-R/PM, the current model can model concurrent tasks more easily by adding a route for the secondary task in the model, feeding the stimuli of the secondary task to the model, and coding the task description into an easy-to-use Excel sheet. This unique feature offers great potential of the model for easy application and learning by researchers and engineers in practice.

Acknowledgments

This article is based upon work supported by the National Science Foundation under Grant No. NSF 0308000. However, any opinions, findings and conclusions or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation (NSF).

Reference

Bear, M., & Connor, B. (2001). *Neuroscience: Exploring the Brain (2nd edition)*. Lippincott Williams& Wilkins Publisher

Byrne, M.D., & Anderson, J.R. (2001). Serial Modules in Parallel: The Psychological Refractory Period and Perfect Time Sharing. *Psychology Review*, 108, 4, 847-869

Card, S., Moran, T., & Newell, A. (1986). The Model Human Processor: An engineering model of human performance. In Boff, K. (Eds.) *Handbook of Perception and human performance*, New York: Wiley

Collette, F. & Linden, M.V. (2002). Brain Imaging of the Central Executive Component of Working Memory. *Neuroscience and Biobehavioral Review*, 26, 105-125

Feyen, R.G. & Liu, Y. (2001). Modeling Task Performance Using the Queuing Network Model Human Processor (QN-MHP). *Proceedings of the Fourth International Conference on Cognitive Modeling*, 73-78

Fletcher.P.C., & Henson, R.N.A. (2001). Frontal Lobes and Human Memory: Insights from Functional Neuroimaging. *Brain*, 124, 849-881

Flynn, B.M. (1943). Pitch discrimination: the form of the psychometric function and simple reaction time to liminal differences. *Archives of Psychology (Columbia University)*, 280, 41

Heathcote, A., Brown, S., & Mewhort, D.J.K (2000). The power law repealed: the case for an exponential law of practice. *Psychological Bulletin,* 7, 2, 185-207

Liu, Y. (1996). Queuing network modeling of elementary mental processes. *Psychological Review*, 103, 1, 116-136

Liu, Y. (1997). Queuing network modeling of human performance of concurrent spatial and verbal tasks. *IEEE Transactions on Systems, Man, Cybernetics*, 27, 195-207

Liu, Y., Feyen, R., and Tsimhoni, O. (2004). *Queuing Network – Model Human Processor (QN-MHP): A Computational Architecture for Multitask Performance in Human Machine Systems.* Technical Report 04-05. Department of Industrial and Operations Engineering, University of Michigan

Meyer, D.E & Kieras, D.E. (1997a). A computational theory of executive cognitive. *Psychology Review*, 104, 1, 3-65

Mitz, A. R., Godshalk, M., & Wise, S.P. (1991). Learning – dependent activity in the premotor cortex: activity during the acquisition of the conditional motor assassinations. *The Journal of Neuroscience*, 11, 6, 1855-1872

Mowbray, G. H., & Rhoades, M.V. (1959). On the reduction of choice reaction time with practice. *Quarterly Journal of Experimental Psychology*, 14, 1-36

Mustovic. H., Scheffler, K., Di Salle, F. (2003). Temporal integration of sequential auditory events: silent period in sound pattern activates human planum temporale. *NeuroImage*, 20, 429–434

Nakahara, H., Doya, K., & Hikosaka, O. (2001). Parallel Cortico-Basal Ganglia Mechanisms for Acquisition and Execution of Visuomotor Sequences-A Computational Approach. *Journal of Cognitive Neuroscience,* 13, 5, 626–647

Rektor, I., Kanovsky, P., Bares., M. (2003). A SEEG study of ERP in motor and premotor cortices and in the basal ganglia. *Clinical Neurophysiology,* 114, 463–471

Roland, P.E. (1993). *Brain Activation*. John Wiley &Sons

Romero, D.H, Lacourse, M.G, Lawrence,K.E. (2000). Event-related potentials as a function of movement parameter variations during motor imagery and isometric action. *Behavioral Brain Research*, 117, 83–96

Rudell, A.P. & Hu, B. (2001). Does a warning signal accelerate the processing of sensory information? Evidence from recognition potential responses to high and low frequency words. *International Journal of Psychophysiology*, 41, 31-42

Ruthruff, E., Johnston, J.C & Selst, M.V. (2001). Why Practice Reduces Dual-Task Interference. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 1, 3-21

Schmidt, R.A. (1988). *Motor control and learning* Human Kinetics Publishers, 82-298

Smith, E.E, & Jonides, J. (1998). Neuroimaging analyses of human working memory. *Proc. Natl. Acad. Sci. USA*, 95, 12061–12068

Sutton, R.S. & Barto, A.G. (1998). *Reinforcement Learning: An Introduction.* MIT Press: Cambridge, Massachusetts.

Tsimhoni,O., & Liu, Y.(2003). Modeling steering with the Queuing Network-Model Human Processor (QN-MHP). *Proceedings of the 47th Annual Conference of the Human Factors and Ergonomics Society*, 81-85

Ungerleider, L.G., Doyon, J., & Karni, A. (2002). Imaging Brain Plasticity during Motor Skill Learning. *Neurobiology of Learning and Memory*, 78, 553–564

Van Selst, M., Ruthruff, E. & Johnston, J.C. (1999). Can Practice Eliminate the Psychological Refractory Period Effect? *Journal of Experimental Psychology: Human* Perception and Performance, 25, 5, 1268-1283

Wu, C. & Liu, Y. (2004a). Modeling behavioral and brain imaging phenomena in transcription typing with queuing networks and reinforcement learning algorithms. *Proceeding of Sixth International Conference on Cognitive Modeling*, Pittsburgh, PA, USA (In press).

Wu, C. & Liu, Y. (2004b). Modeling human transcription typing with queuing network-model human processor. *Proceedings of the 48th Annual Meeting of Human Factors and Ergonomics Society*, New Orleans, Louisiana, USA (In press).

Appendix

1. Parameters Setting at Hicog and PM server

1) $A_{Hicog\text{-}symbol}$: minimal value of the processing time of task 2 entity in Hicog server. Since choice reaction time (RT) of 4 alternatives can be reduced to RT of 2 alternatives with practice (Mowbray et al., 1959), after intensive practice, RT of 8 alternative choices in Van Selst's experiment will reduce to RT of 4 alternatives without intensive practice. $A_{Hicog\text{-}symbol}$ equals the RT of 4 alternatives (Hick's Law, intercept:150ms, slope:170ms/bit, Schmidt, 1988) minus 1 average perception cycle (100ms), 2 cognitive cycles (2*70 ms) and 1 motor cycle (70ms) (Card et al., 1986). Therefore, $A_{Hicog\text{-}symbol}=150+170*\mathrm{Log}_2(4)-100-2*70-70=180$ ms.

2) $B_{Hicog\text{-}symbol}$: change of processing time of task2 entity in Hicog server at the beginning and end of practice. At the beginning of the practice in single task 2, RT of the 8 alternatives (Hick's Law, intercept:150ms, slope:170ms/bit, Schmidt, 1988) is composed of 1 perception cycle (100ms), maximum processing time at Hicog ($A_{Hicog\text{-}symbol}+B_{Hicog\text{-}symbol}$), and 1 motor cycle (70ms) (Card et al., 1986). Therefore, $B_{Hicog\text{-}symbol}=150+170*\mathrm{Log}_2(8)-100-A_{Hicog\text{-}symbol}-70=170$ms.

3) $\alpha_{Hicog\text{-}symbol}$, $\alpha_{Hicog\text{-}tone}$: learning rate of Hicog server in processing the task 2 and task 1 entities. Based on $\alpha=0.001$ approximately in Heathcote et al. (2000)'s study, learning difficulty increased 4 times because of the 4 incompatible alternatives. Thus, $\alpha_{Hicog\text{-}symbol}=\alpha_{Hicog\text{-}tone}=0.001/4=1/4000$.

4) $A_{Hicog\text{-}tone}$: minimal value of the processing time of task1 entity in central executive. After intensive practice, the discrimination task of the 2 classes of tones in Van Selst (1999) 's experiment can be simplified into a choice reaction time of two alternatives, requiring the minimum value of 1 cognitive cycle (25ms) (Card et al., 1986).

5) $B_{Hicog\text{-}tone}$: change of processing time of task1 entity in Hicog at the beginning and end of practice. At the beginning of the single task 1, the reaction time to discriminate the 2 classes of tone is 642ms (Flynn, 1943), which is composed of 1 perception cycle (100ms), 2 cognitive cycles (70*2 ms), ($A_{Hicog\text{-}tone}+B_{Hicog\text{-}tone}$) and 1 motor cycle (70ms). Therefore, $B_{Hicog\text{-}tone}=642-100-2*70-A_{Hicog\text{-}tone}-70=307$ms.

6) $A_{PM\text{-}symbol}$: minimal value of the processing time of task 2 entity in PM. After intensive practice, RT of the 8 alternative choices in Van Selst's experiment will transform to RT of 8 most compatible alternatives (RT=217ms, Schmidt, 1988) which is composed 1 perception cycle and 1 motor cycle. Therefore, $A_{PM\text{-}symbol}=217-100-70=47$ms.

7) $B_{PM\text{-}symbol}$: change of processing time of task2 entity in PM at the beginning and end of practice. At the beginning of practice in single task 2, RT of 8 alternative choice reaction time (Hick's Law: 50ms, slope: 170ms/bit) is composed of 1 average perception cycle (100ms), ($A_{PM\text{-}symbol}+B_{PM\text{-}symbol}$), 1 motor cycle (70ms). Thus, $B_{PM\text{-}symbol}=150+170*\mathrm{Log}_2(8)-100-A_{PM\text{-}symbol}-70=443$ ms.

8) $\alpha_{PM\text{-}symbol}$: learning rate of PM in processing the task2 entity The speed of formation of the automatic process in PM is slower than Hicog because it receive the dummy entities from CE server via the indirect parallel learning mechanism with the 4 incompatible alternatives (Nakahara et al., 2001). Thus, $\alpha_{PM\text{-}symbol}=(0.001/4)/4=1/16000$.

2. Proof of Lemma 1

Lemma 1. At any transition state t ($t\neq 0$), if $1/\mu_{2,t} < 1/\mu_{3,t}$, then $Q^{t+1}(1,2) > Q^{t+1}(1,3)$

Proof. Using mathematic deduction method

i) At t=0: $Q^1(1,3)=Q^1(1,2)=Q^1(2,4)=Q^1(3,4)=0$

ii) At t=1: Using the online Q learning formula:

$Q^2(1,3)=Q^1(1,3)+\varepsilon[r_t+\gamma Q^1(3,4)-Q^1(1,3)]=\varepsilon\mu_{3,1}$

Note: because entity routes to only one server (server 4) $\max_b Q^t(S_t+1,b)=Q(3,4)$, $Q^2(1,2)=\varepsilon\mu_{2,1}$, $Q^2(3,4)=\varepsilon\mu_4$, $Q^2(2,4)=\varepsilon\mu_4$. If $1/\mu_{2,1}<1/\mu_{3,1}$, then $\varepsilon\mu_{3,1}<\varepsilon\mu_{2,1}$ (given $0<\varepsilon<1$), i.e. $Q^2(1,2)>Q^2(1,3)$

Thus, Lemma is proved at t=1.

iii) According to Mathematic Deduction Method， Lemma 1 is correct: i.e. at transition state t=k: if $1/\mu_{2,k}<1/\mu_{3,k}$, then $Q^{k+1}(1,2)>Q^{k+1}(1,3)$. We want to prove at transition state k+1, Lemma is still correct: i.e.At transition state t=k+1:

if $1/\mu_{2,k+1}<1/\mu_{3,k+1}$, then $Q^{k+2}(1,2)>Q^{k+2}(1,3)$

At t=k+1:

$Q^{k+2}(1,2)=Q^{k+1}(1,2)+\varepsilon[\mu_{2,k+1}+\gamma\varepsilon\mu_4-Q^{k+1}(1,2)]$

$Q^{k+2}(1,3)=Q^{k+1}(1,3)+\varepsilon[\mu_{3,k+1}+\gamma\varepsilon\mu_4-Q^{k+1}(1,3)]$

$Q^{k+2}(1,2)-Q^{k+2}(1,3)$

$=Q^{k+1}(1,2)+\varepsilon[\mu_{2,k+1}+\gamma\varepsilon\mu_4-Q^{k+1}(1,2)]-\{Q^{k+1}(1,3)+\varepsilon[\mu_{3,k+1}+\gamma\varepsilon\mu_4-Q^{k+1}(1,3)]\}$

$=(1-\varepsilon)[Q^{k+1}(1,2)-Q^{k+1}(1,3)]+(\varepsilon\mu_{2,k+1}-\varepsilon\mu_{3,k+1})$

With equation 3 and $0<\varepsilon<1$, we have:

$(1-\varepsilon)[Q^{k+1}(1,2)-Q^{k+1}(1,3)]>0$

Given $1/\mu_{2,k+1}<1/\mu_{3,k+1}$ and $0<\varepsilon<1$, then $(\varepsilon\mu_{2,k+1}-\varepsilon\mu_{3,k+1})>0$, i.e. $Q^{k+2}(1,3)-Q^{k+2}(1,2)>0$

Thus, Lemma 1 is correct at t=k+1. Lemma 1 is proved.