Cognitive Performance Modeling

Chapter · July 2023

DOI: 10.1002/9781119863663.ch24

2 authors:

Maryam Zahabi
Texas A&M University

Junho Park
Santa Clara University

24

Cognitive Performance Modeling

Maryam Zahabi and Junho Park

Wm Michael Barnes '64 Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX, USA

24.1 Introduction

Cognitive performance models (CPM) are computational models that represent humans' performance as they interact with interfaces and provide information on user intentions and information processing. These models can analyze the tasks in high detail, predict operator task performance and cognitive workload, and identify serial and parallel operations.

Use of CPM provides several advantages compared to conducting human subject experiments. These advantages include (i) quantification and prediction of human behavior in natural tasks based on human information processing models; (ii) the model generation does not require human subject involvement and therefore can save time and cost in early stages of interface design and evaluation process; (iii) these computational models can be easily modified and do not require the analyst to have substantial programming background [43]; and (iv) the method is nonobtrusive as compared to physiological measures of workload [45].

The CPM approach has been applied in different human factors applications, including the analysis of aerospace systems [33], augmented cognition [11], computer systems [37], human-artificial intelligent (AI)-robot teaming [10], perception and performance [4], surface transportation [40], and user testing and evaluation [32]. Furthermore, a number of reviews have been conducted on CPM approaches for understanding design patterns [34], decision-making process [28], cognitive architecture [23], application of CPM in aviation safety [26], and error prediction [27]. However, there has been no comprehensive review on application of CPM approaches for understanding human–system interaction (HSI), which is the main focus of this chapter.

24.2 Background

Early models were used to quantify HSI in military applications during World War II [28]. These models included Fitts' law [12], Hick and Hyman's selective response model [15], and the signal detection model [38]. The Command Language Grammar developed by Moran [31] could be considered as the first human–computer interaction (HCI) model [1]. The top-down approach decomposed the tasks and gave a conceptual view of the interface during the design process.

Handbook of Human-Machine Systems, First Edition. Edited by Giancarlo Fortino, David Kaber, Andreas Nürnberger, and David Mendonça.

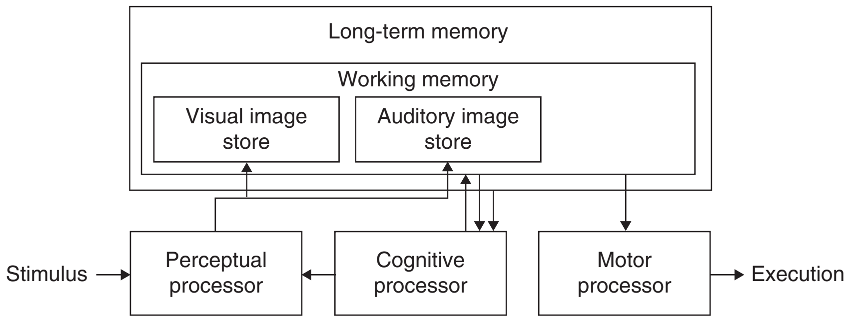

Figure 24.1 Model Human-Processor. Source: Modified from Card et al. [8].

However, it mainly focused on modeling human motor activities and did not include perceptual and cognitive operators.

A basic architecture of human performance model is illustrated in Figure 24.1 [8]. After the initial models, various CPMs emerged applying cognitive science theory [34]. These models included Goals, Operators, Methods, and Selection (GOMS) rules [7], Adaptive Control of Thought-Rational (ACT-R) [2], Executive-Process Interactive Control (EPIC) [21], State, Operator, and Result (SOAR) [25], and Queuing Network-Model Human Processor (QN-MHP) [30].

These CPMs have been extended for specific purposes, especially GOMS and ACT-R models, as illustrated in Figure 24.2. Although the original GOMS models only supported serial activities, this limitation was resolved in later models such as the Critical Path Model-GOMS method. In addition, some variants of keystroke-level models (KLM) were developed for assessing HSI in touchscreen (i.e. touch-level model, TLM) and gesture-based interfaces (i.e. gesture-level model, GLM). Other variants of these models were generated for cockpit evaluation, reliability analysis, and modeling of social interactions (e.g. Enhanced-GOMS Language [13], GOMS-Human Reliability Analysis [5], sociotechnical GOMS [42]). ACT-R models have also been advanced for specific applications such as image processing and analyzing mobile applications, or by integration with other models and different programming languages. However, there have not been many advances in the applications of other methods such as QN-MHP, EPIC, or SOAR.

24.3 State-of-the-Art

In this review, we provided a general description of CPMs with their capabilities and limitations to further identify the challenges associated with their use and provide directions for future research. Modeling using the GOMS method starts with constructing a hierarchical structure of goals from top-level goals and is processed until the unit tasks or subgoal levels cannot be further decomposed [41]. To achieve a goal at the unit-task level, methods (or interactive routines) are required to specify what operators need to be executed to perform a specific action. GOMS method is based on the Model Human Processor (MHP) theory [8]. Goals are symbolic structures that establish the state to be achieved. Operators are fundamental perceptual, motor, or cognitive acts whose execution is needed to change any aspect of the user's mental state or affect the task environment. Methods describe a procedure for a goal. Finally, selection rules are required if more than one method is available for a user to accomplish the goal. There are numerous extensions of GOMS, including KLMs [6], Cognitive-Perceptual-Motor GOMS [16], Goals, Operators, Methods, and Selection Language (GOMSL) [17], Natural Goals, Operators, Methods, and Selection Language

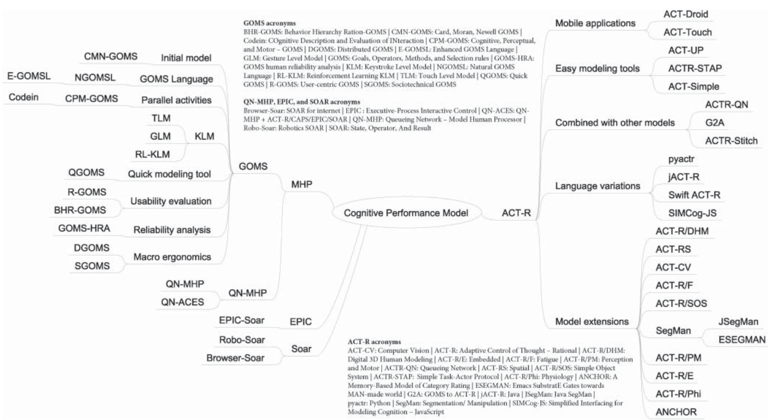

Figure 24.2 Major CPMs and their extensions.

(NGOMSL) [18], and Enhanced GOMSL [13]. KLM is a method that predicts task completion time for experts to accomplish a routine task without errors. CPM-GOMS is an advanced method in that it can model parallel processes. NGOMSL is a high-level (natural language) syntax for GOMS representation. GOMSL is an executable form of NGOMSL and a computationally realized version of MHP, which is a basis of GOMS. Although GOMS is a fast and easy model to develop and includes all processes to access task performance, it mainly represents the behavior of expert users without errors, and is deterministic. Also, adding mental operators should be consistent following the rules in [20].

ACT-R provides models of elementary and irreducible cognitive and perceptual operations that enable human information processing. In theory, each task that humans can perform consists of a series of these discrete operations [3]. ACT-R's primary time unit is 50 ms which can describe human information processing in a fine-grained resolution. In addition, ACT-R can generate essential outcomes such as time to perform a task and accuracy. ACT-R's specific characteristics include connection between human cognition and perception, detailed modeling algorithm for memory process including memory storage/update/decay/retrieval estimates based on mathematical equations, and use of standard symbolic variables. In this method, chunks and production rules are discrete, and their operations are syntactical, which means that they do not refer to the semantic content of the representations but only to the properties that deem them appropriate for calculations. However, ACT-R models have some limitations, including (i) the models only allow serial access of cognitive operators, (ii) they have been mainly used in academic research, and (iii) it takes a long time to model (at least several days to weeks of using the system) and takes months to years to become an expert in its use [35].

EPIC is a general framework, represented as a simulation modeling environment, in which models of human performance in specific tasks may be constructed [21]. Since EPIC was mainly developed for modeling perception and motor functions [39], it influenced the development of ACT-R/Perceptual and Motor and SOAR that combined detailed perceptual and motor components into their models [34]. Similar to ACT-R and SOAR, EPIC models encompass a production-rule system (a "cognitive processor") that provides procedural knowledge. "Perceptual processors" also process different sensory information including visual, tactile, and auditory information. The outcomes of the perceptual processors are delivered to working memory. There are two types of working memory that are not related to the sensory-motor information: one is to store current goals and steps to reach those goals (i.e. "control store"), and the other is a general working memory for miscellaneous information. Similar to ACT-R, modeling work of EPIC is complex, and therefore, it has been mainly used for academic research. Also, the model does not consider changes in human behavior as a result of learning a system. For example, users may perform tasks in a serial manner at the beginning, while they can perform the same tasks in parallel once they learn the system and change their interaction style.

SOAR is a functional model to understand cognitive mechanisms as a basis of intelligent human behavior [25]. Also, it is an architecture for human cognition expressed in the form of a production system. SOAR can represent extensive and complex rule sets. Its primary use is in artificial intelligence (AI) and cognitive modeling. In addition, it has been combined with EPIC's perceptual-motor processors. AI agents in SOAR can be developed based on different types of knowledge, whether programmed or learned by the system. In addition, in SOAR, cognitive tasks can be processed in parallel, which is the main distinction between this model and models such as ACT-R and CPM-GOMS.

Lastly, QN-MHP is a computational cognitive architecture that integrates the mathematical framework of queuing network theory with MHP [30]. Based on a network structure of 20 process

Table 24.1 Comparison of cognitive performance models.

Feature	GOMS	ACT-R	EPIC	SOAR	QN-MHP
Capability	Estimate task performance and cognitive workload	Estimate task performance and cognitive workload, outcomes similar to neurological data	Estimate task performance, cognitive workload	Estimate task performance, cognitive workload	Estimate task performance, cognitive workload
Characteristics	Fast and easy modeling, can see the detailed codes	Integration of perception, cognition, and motor functions, detailed modeling of memory, symbolic language	Detailed perception and motion algorithm	Multiple problem spaces, unlimited working access	Based on human brain structure and functions, use of queuing theory, real-time visualization
Limitations	Modeling of skilled user behavior, deterministic operator time	Serial cognitive process, mainly used in academia, complex modeling	Does not consider human learning capabilities, mainly used in academia, complex modeling	Not subject to forgetting, learning is tied to impasse, complex modeling	Not publicly available[a)]
Major application	Human–computer interaction	Memory, attention, education	Visual search, auditory tasks	Expert system, Autonomous agents	Surface Transportation
Language or Tool	Cogulator	CogTool, Common Lisp, Python, Java	Common Lisp, C++	C, Java	ProModel, Matlab, Eclipse

a) The model can be accessed by contacting the model authors.

units (e.g. visual recognition server, phonological loop server, or sensorimotor integration server), different cortical areas of human brain and corresponding functional modules of human information acquisition, processing, and implementation are simulated. Because of this “brain-like” structure, QN-MHP can visualize internal information flows during the simulation of related activities. However, its inability to generate or model complex cognition such as language comprehension or problem-solving requires creating new rules rather than only using the rules preprogrammed by a model developer [29]. In addition, QN-MHP codes and software application for creating these models are not publicly available. A comparison of these models is provided in Table 24.1.

24.4 Current Research Issues

Some of the developed models (e.g. SOAR and EPIC) are difficult to use for practitioners or experts in different domains even though there are several documents explaining the methods in detail (e.g. [19]). There were a few cases of using CPMs in the military domain such as Air Force [24], or Office of Naval Research [22]. This was mainly possible because the authors of these models (i.e. SOAR and EPIC) were engaged in those projects. Although all of the models argue that they can be used to evaluate human performance while interacting with a commercially available product or service such as mobile phones, in-vehicle displays, or vehicles, applying CPM to commercial applications still seems far away [44]. Finding, learning, and applying models in other commercial domains still demand practitioners to consider the merits of using models in their work.

Another issue with CPMs is model validation. Some of these models have been found to be inadequate or outdated after conducting validation studies [1]. This issue and the limitation above regarding the difficulty of using these models could be why so many model extensions in the family tree are not currently used. In addition, without validation, use of these models in commercial applications is not justifiable.

Finally, moderator effects have not been included in the current models. The models assume a fixed set of parameters and do not reflect users’ interaction style change in the middle of the task (e.g. serial to parallel task performance). Although a few studies investigated the effects of stress, fatigue, sleep deprivation, drugs, or emotions on human performance, they did not implement how the model mechanism could change as a result of these effects [14]. Also, group behavior modeling has not been incorporated in these models (e.g. drones, ground weapon systems, or robots).

24.5 Future Research Directions Dealing with the Current Issues

In order to enhance the use of CPM approaches in future, we recommend the following steps:

1. A technical group which has the expertise and authority to manage models is needed (e.g. Human Performance Modeling-Management Group [HPM-MG]). HPM-MG includes steps for model development and releases and protocols for model verification, validation, and accreditation (VVA) [36]. In HPM-MG, a simplified version of the VVA process should be established and applied to the models. In the verification stage, researchers first propose a new methodology with possible data. This is what current studies are doing. Next, in the validation stage, the researchers need to validate the model with additional data and compare with other models. One of the possible ways for validation is using a genetic algorithm to automate the model fitting

work [34]. Finally, the researchers should develop a user-friendly interface (e.g. graphical user interface), manual, and tutorials for practitioners in the accreditation stage. HPM-MG oversees each stage and finally accredits the model. Furthermore, this process can improve the transparency and reproducibility of the model. A common repository or portal can also be provided to researchers to share each model's latest status to avoid redundant work or using outdated information.

2. A user-friendly modeling interface is needed. CogTool is one of the examples which enables modeling without coding. However, it is mainly used for webpage usability analysis, and the tool needs to be updated. The latest version on Github is from 2014 (https://github.com/cogtool/cogtool). Cogulator is another CPM interface, which requires coding and does not allow loops (https://cogulator.io/). Distract-R is a GUI-based model for driving (https://www.cs.drexel.edu/~salvucci/cog/distract-r/). However, the model has specific routes and can only model dial-pad based secondary task devices.
3. Analysts who are interested in using CPM but are not familiar with modeling work should refer to Figure 24.1 and Table 24.1 first to decide the appropriate model and its extension. Then, they need to select and use a GUI-based tool. Finally, once they develop a particular model and if need additional functions, they can look into a coding-based approach that enables more detailed functionalities. For example, an analyst can implement or modify specific arguments of a function in the original code.
4. Dynamic parameter adjustment or mechanism of calculation is required for users' skill development and changes in performance due to external stimulus (e.g. fatigue or stress) in the middle of the task. For example, Cognitive Jack (CoJACK) project enabled an agent to learn knowledge from the environment and generate new behaviors [34]. Recently, ACT-R/phi has been developed to model moderators of cognition and combined that with human physiology. This is a more straightforward way as some of the physiological modulating processes entail bottom-up effects on cognitive processes [9].

References

1 Al Seraj, M.S., Pastel, R., and Al-Hasan, Md. (2018). A survey on user modeling in HCI. *Computer Applications: An International Journal (CAIJ)* 5 (1): 1–8.

2 Anderson, J.R. (1993). *Rules of the Mind*, 319p. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

3 Anderson, J.R., Matessa, M., and Lebiere, C. (1997). ACT-R: a theory of higher level cognition and its relation to visual attention. *Human–Computer Interaction* 12 (4): 439–462.

4 Bolkhovsky, J., Ritter, F.E., Chon, K.H., and Qin, M. (2018). Performance trends during sleep deprivation.

5 Boring, R.L. and Rasmussen, M. (2016). GOMS-HRA: a method for treating subtasks in dynamic human reliability analysis. *Proceedings of the 2016 European Safety and Reliability Conference*, 956–963.

6 Card, S.K., Moran, T.P., and Newell, A. (1980). The keystroke-level model for user performance time with interactive systems. *Communications of the ACM* 23 (7): 396–410.

7 Card, S.K., Moran, T.P., and Newell, A. (1983). *The Psychology of Human–Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

8 Card, S., Moran, T., and Newell, A. (1986). The model human processor: an engineering model of human performance. In: *Handbook of Perception and Human Performance, Cognitive Processes and Performance*, vol. 2 (ed. K.R. Boff, L. Kaufman, and J.P. Thomas), 1–35. Wiley.

9 Dancy, C.L. (2019). A hybrid cognitive architecture with primal affect and physiology. *IEEE Transactions on Affective Computing* 12 (2): 318–328.

10 Dudzik, K.A.T. (2019). Cognitive modeling as a method for agent development in artificial intelligence. PhD thesis. Carleton University.

11 Fincham, J.M. (2005). Cognitive modeling and fMRI: an integrated approach toward understanding mechanisms of complex skill performance. PhD thesis. Carnegie Mellon University.

12 Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology* 47 (6): 381.

13 Gil, G.-H. (2010). An accessible cognitive modeling tool for evaluation of human–automation interaction in the systems design process. PhD thesis. North Carolina State University.

14 Gunzelmann, G. and Gluck, K.A. (2008). Approaches to modeling the effects of fatigue on cognitive performance. *Proceedings of the 17th Conference on Behavior Representation in Modeling and Simulation*, 136–145. Citeseer.

15 Hick, W.E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology* 4 (1): 11–26.

16 John, B.E. and Gray, W.D. (1995). CPM-GOMS: an analysis method for tasks with parallel activities. *Conference Companion on Human Factors in Computing Systems*, 393–394.

17 Kieras, D.E. (1988). Towards a practical GOMS model methodology for user interface design. In: *Handbook of Human–Computer Interaction*, 135–157. Elsevier.

18 Kieras, D. (1994). *A Guide to GOMS Task Analysis*. University of Michigan.

19 Kieras, D.E. (1999). *A Guide to GOMS Model Usability Evaluation Using GOMSL and GLEAN3*, 313. University of Michigan.

20 Kieras, D. (2001). *Using the Keystroke-Level Model to Estimate Execution Times*, 555. University of Michigan.

21 Kieras, D.E. and Meyer, D.E. (1995). Predicting human performance in dual-task tracking and decision making with computational models using the epic architecture. *Proceedings of the 1st International Symposium on Command and Control Research and Technology*, June 1995. Washington, DC: National Defense University, Citeseer.

22 Kieras, D.E. and Wakefield, G.H. (2016). Extending and Applying the Epic Architecture for Human Cognition and Performance: Auditory and Spatial Components. *Technical report*. FR-11/ONR-EPIC-17. University of Michigan Division of Research Development Ann Arbor United States. https://apps.dtic.mil/sti/pdfs/ADA535789.pdf.

23 Kotseruba, I. and Tsotsos, J.K. (2020). 40 years of cognitive architectures: core cognitive abilities and practical applications. *Artificial Intelligence Review* 53 (1): 17–94.

24 Laird, J.E. and Nielsen, E. (1994). Coordinated behavior of computer generated forces in TacAir-SOAR. *AD-A280* 63 (1001): 57.

25 Laird, J., Hucka, M., Huffman, S., and Rosenbloom, P. (1991). An analysis of SOAR as an integrated architecture. *ACM SIGART Bulletin* 2 (4): 98–103.

26 Leiden, K. and Best, B. (2005). *A Cross-Model Comparison of Human Performance Modeling Tools Applied to Aviation Safety*. Boulder, CO: Micro Analysis & Design, Inc.

27 Leiden, K., Laughery, K.R., Keller, J. et al. (2001). A review of human performance models for the prediction of human error. *Ann Arbor* 1001: 48105.

28 Li, N., Huang, J., and Feng, Y. (2020). Human performance modeling and its uncertainty factors affecting decision making: a survey. *Soft Computing* 24 (4): 2851–2871.

29 Liu, Y. (2009). QN-ACES: integrating queuing network and ACT-R, CAPS, EPIC, and SOAR architectures for multitask cognitive modeling. *International Journal of Human–Computer Interaction* 25 (6): 554–581.

30 Liu, Y., Feyen, R., and Tsimhoni, O. (2006). Queuing Network-Model Human Processor (QN-MHP) a computational architecture for multitask performance in human–machine systems. *ACM Transactions on Computer–Human Interaction (TOCHI)* 13 (1): 37–70.

31 Moran, T.P. (1981). The command language grammar: a representation for the user interface of interactive computer systems. *International Journal of Man-Machine Studies* 15 (1): 3–50.

32 Oyewole, S.A., Farde, A.M., Haight, J.M., and Okareh, O.T. (2011). Evaluation of complex and dynamic safety tasks in human learning using the ACT-R and SOAR skill acquisition theories. *Computers in Human Behavior* 27 (5): 1984–1995.

33 Redding, R.E. (1992). Cognitive Task Analysis of Prioritization in Air Traffic Control.

34 Ritter, F.E., Tehranchi, F., Dancy, C.L., and Kase, S.E. (2020). Some futures for cognitive modeling and architectures: design patterns for including better interaction with the world, moderators, and improved model to data fits (and so can you). *Computational and Mathematical Organization Theory* 26 (3): 278–306.

35 Salvucci, D.D. and Lee, F.J. (2003). Simple cognitive modeling in a complex cognitive architecture. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 265–272.

36 Sanders, P. (1996). DoD Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A). *Technical Report*. NSN 7540-01-280-5500. Office of the Under Secretary of Defense for Acquisition and Technology. https://apps.dtic.mil/sti/pdfs/ADA315867.pdf.

37 St. Amant, R., Horton, T.E., and Ritter, F.E. (2004). Model-based evaluation of cell phone menu interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 343–350.

38 Swets, J.A. (1964). *Signal Detection and Recognition in Human Observers: Contemporary Readings*. Wiley.

39 Taatgen, N. and Anderson, J.R. (2010). The past, present, and future of cognitive architectures. *Topics in Cognitive Science* 2 (4): 693–704.

40 Tsimhoni, O. and Reed, M.P. (2007). The virtual driver: integrating task planning and cognitive simulation with human movement models. *SAE Transactions* 116: 1525–1531.

41 Van Rijn, H., Johnson, A., and Taatgen, N. (2011). Cognitive user modeling. In: *Handbook of Human Factors in Web Design*, (eds. A. Johnson, R. Proctor), 527–542. CRC Press.

42 West, R.L. and Nagy, G. (2000). Situating GOMS models within complex, sociotechnical systems. *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 22.

43 Wiendahl, M., Wierling, P.S., Nielsen, J. et al. (2008). High throughput screening for the design and optimization of chromatographic processes–miniaturization, automation and parallelization of breakthrough and elution studies. *Chemical Engineering & Technology: Industrial Chemistry-Plant Equipment-Process Engineering-Biotechnology* 31 (6): 893–903.

44 Wilson, M.D., Boag, R.J., and Strickland, L. (2019). All models are wrong, some are useful, but are they reproducible? Commentary on Lee et al.(2019). *Computational Brain & Behavior* 2 (3): 239–241.

45 Zahabi, M., White, M.M., Zhang, W. et al. (2019). Application of cognitive task performance modeling for assessing usability of transradial prostheses. *IEEE Transactions on Human–Machine Systems* 49 (4): 381–387.