EPAM/CHREST Tutorial: Fifty Years of Simulating Human Learning

Peter C. R. Lane (peter.lane@bcs.org.uk)
School of Computer Science, University of Hertfordshire, College Lane, Hatfield AL10 9AB, UK

Fernand Gobet (fernand.gobet@brunel.ac.uk)
School of Social Sciences, Brunel University, Cleveland Road, Uxbridge, Middlesex, UB8 3PH, UK

Keywords: CHREST; EPAM; cognitive modelling; chunking

Overview

Generating quantitative predictions for complex cognitive phenomena requires precise implementations of the underlying cognitive theory. The history of computational modelling includes many diverse approaches to this problem of creating process models. These include models of single phenomena (such as Sternberg's model of STM; (Sternberg, 1966)), to integrated models covering a wide range of different phenomena (such as Soar (Newell, 1990) and ACT-R (Anderson & Lebière, 1998)), to over-arching principles, which guide the development of models in disparate domains (such as connectionist approaches (McLeod, Plunkett, & Rolls, 1998), or embodied cognition (Pfeifer & Scheier, 1999)).

This tutorial focuses on the EPAM/CHREST tradition, which has been providing significant models of human behaviour for 50 years. The first implementation of EPAM (Elementary Perceiver and Memoriser) was developed by Edward Feigenbaum in 1959. Early models of EPAM provided the impetus to develop the chunking theory (Chase & Simon, 1973; Gobet et al., 2001), which has been an important component of theories of human cognition ever since. With an emphasis on learning phenomena, EPAM and CHREST (Chunk Hierarchy and REtrieval STructures) are sensitive to the ways in which a model's information is built up through interactions with an external environment. Thus, EPAM/CHREST models are typically developed from large quantities of naturalistic input. For example, in modelling expert perception of chess players, actual chess games are used (Gobet & Simon, 2000). Similarly, in modelling the acquisition of syntax, large corpora of mother-child interactions are employed to develop the model's long-term memory (Freudenthal, Pine, Aguado-Orea, & Gobet, 2007).

The tutorial is structured so that participants will:

1. Acquire a complete understanding of the EPAM and CHREST approach to computational modelling, and their relation to the chunking and template theories of cognition;
2. Explore some key learning phenomena supporting the chunking theory, based around experiments in verbal-learning, categorisation and the acquisition of expertise;
3. Be introduced to an implementation of CHREST which can be used for constructing models of their own data.

Further information about CHREST, supporting publications and implementations can be found at: `http://chrest.info`

Chunking and Template Theories

A *chunk* is a 'familiar pattern', an item stored in long-term memory. Chunks collect together more basic elements which have strong associations with each other, but weak associations with other elements (Chase & Simon, 1973; Cowan, 2001). The chunking theory is based on Miller's observation (Miller, 1956) that short-term memory typically contains a limited number of pieces of information, but that the size of these pieces varies with context. Chase and Simon (1973) confirmed the presence of chunks in the recall of chess positions, and the EPAM model provides a means of learning, storing and retrieving such chunks.

The *template theory* (Gobet & Simon, 1996, 2000) extends the chunking theory by adding mechanisms to create retrieval structures, using specific retrieval cues to store and obtain information rapidly. The template is a form of slotted schema, containing a *core*, of stable information, and *slots*, containing variable information. Where the chunking theory captures much of how the average person learns in tasks such as verbal-learning, the template theory further captures the way in which highly-trained human experts perceive and identify patterns in their domain of expertise.

A more detailed overview of the chunking and template theories is contained in Gobet et al. (2001).

Implementation

CHREST comprises three basic modules:

- Input/output module, which is responsible for feature extraction, passing the features to the long-term memory for sorting, and guiding the eye movements;
- Long-term memory, which holds information in the form a discrimination network; and
- Short-term memories, which hold pointers to nodes in the long-term memory.

The key feature which distinguishes EPAM/CHREST models is the discrimination network for storing and retrieving information in long-term memory. Information is assumed to form a list of subobjects, each of which is either a further list of subobjects or else a primitive. Tests in the discrimination network check for the presence of individual primitive objects, or a list of subobjects. The discrimination network is trained by exposing CHREST to a large set of naturalistic data. Typical sizes of network for an expert in a complex domain is of the order of 100,000 nodes.

In both EPAM and CHREST, chunks are stored as nodes within the discrimination network. CHREST also includes mechanisms to bring together chunks when an internal node meets specific criteria relating to its connections with other nodes within memory. A template is then formed from the common information in the linked chunks, with slots created for the variable information. Just as EPAM was the computational embodiment of key aspects of the chunking theory, CHREST implements essential aspects of the template theory.

Input can be provided to CHREST in one of two ways. As a single pattern, which is assumed to be perceived in a single glance. These patterns are input to the network and stored directly. The second way is to use the in-built attentional mechanism, by which CHREST scans an input array, such as a chess board, and stores parts of the input array into memory. Short-term memory will then hold a set of chunks, each of which may hold information about a different part of the chess board, and collectively holding information about most of the board. The attention mechanism in CHREST is described in Lane, Gobet, and Ll. Smith (2009).

CHREST is implemented in Lisp, and uses Tk to provide a graphical interface. A graphical environment enables users to create simple CHREST models by providing data within an input data file. The implementation also supports more complex tailored models which may be developed by writing special-purpose code using the packages within CHREST. Within the tutorial we will introduce participants to the graphical environment, walk them through a number of provided examples which will illustrate the workings of the architecture and some samples of successful applications, and finally describe the input data format for applying the environment to new domains. A library and manual is provided to assist users wishing to write more complex models.

Applications

A variety of experimental data will be covered to illustrate the theory and processes. The EPAM learning system itself was initially developed as a simulation of human verbal learning processes. EPAM's memory structure is constructed through an interlinked set of learning operations which alternately extend and elaborate information in the network. We illustrate these key learning processes using applications from verbal learning (Feigenbaum, 1959; Feigenbaum & Simon, 1984). Further properties of the chunking network will be described with reference to results from categorisation (Gobet, Richman, Staszewski, & Simon, 1997), implicit learning and language learning (Freudenthal et al., 2007; Jones, Gobet, & Pine, 2007).

More elaborate models of expertise require an understanding of how information is retrieved from the external environment. To illustrate this aspect of the theory, we focus on chess expertise, particularly the challenge of the recall task. This application is used to describe CHREST's attention mechanisms (Lane et al., 2009) and how they relate to training the discrimination network.

References

Anderson, J. R., & Lebière, C. (Eds.). (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55-81.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*, 87–114.

Feigenbaum, E. A. (1959). *An information processing theory of verbal learning.* The RAND Corporation Mathematics Division, P-1817.

Feigenbaum, E. A., & Simon, H. A. (1984). EPAM-like models of recognition and learning. *Cognitive Science*, *8*, 305–336.

Freudenthal, D., Pine, J. M., Aguado-Orea, J., & Gobet, F. (2007). Modelling the developmental patterning of finiteness marking in English, Dutch, German and Spanish using MOSAIC. *Cognitive Science*, *31*, 311–341.

Gobet, F., Lane, P. C. R., Croker, S. J., Cheng, P. C.-H., Jones, G., Oliver, I., et al. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, *5*, 236–243.

Gobet, F., Richman, H., Staszewski, J., & Simon, H. A. (1997). Goals, representations, and strategies in a concept attainment task: The EPAM model. *The Psychology of Learning and Motivation*, *37*, 265–290.

Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology*, *31*, 1–40.

Gobet, F., & Simon, H. A. (2000). Five seconds or sixty? Presentation time in expert memory. *Cognitive Science*, *24*, 651–82.

Jones, G. A., Gobet, F., & Pine, J. M. (2007). Linking working memory and long-term memory: A computational model of the learning of new words. *Developmental Science*, *10*, 853–873.

Lane, P. C. R., Gobet, F., & Ll. Smith, R. (2009). Attention mechanisms in the CHREST cognitive architecture. In L. Paletta & J. K. Tsotsos (Eds.), *Proceedings of the fifth international workshop on attention in cognitive science* (Vol. LNAI 5395, pp. 183–196). Berlin: Springer-Verlag, GmbH.

McLeod, P., Plunkett, K., & Rolls, E. T. (1998). *Introduction to connectionist modelling of cognitive processes*. Oxford, UK: Oxford University Press.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Pfeifer, R., & Scheier, C. (1999). *Understanding intelligence*. MIT Press.

Sternberg, S. (1966). High speed scanning in human memory. *Science*, *153*, 652–4.