Discovering Models in Cognitive Psychology Using Experimental Data

Fernand Gobet[1], Laura K. Bartlett[1], Noman Javed[1], Peter C. Lane[2], & Angelo Pirrone[1]

[1] Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science

[2] Department of Computer Science, University of Hertfordshire

This multidisciplinary project (experimental psychology, computer science, philosophy, and cognitive neuroscience) aims to develop a novel methodology for generating and testing scientific models semi-automatically. While the current domain of application is cognitive psychology, the methodology is domain-general and can be applied to any scientific fields where (a) experimental data are available and (b) basic processes are postulated. The methodology can be used with sparse data sets.

The Original Discovery Problem

To develop models explaining how people perform standard experiments in cognitive psychology. For example, a model of how a person memorises a list of items or how they select one of two items.

How Is the Problem Formulated in Computational Terms?

The problem is formulated as a heuristic search through a problem space of discrete structures. The structures/models are programs for a high-level, domain-specific cognitive architecture, implemented as an interpreter for a virtual machine. The search is carried out using genetic programming. The system is currently implemented in ANSI Common Lisp.

What Data and Knowledge are Provided to the System?

The system is provided with three types of data and knowledge:

1. Description of the experimental protocol, including: what are the manipulated variables? What are the stimuli? What is the timeline of the experiment?
2. The results obtained in specific experiments, for each of the experimental conditions. Typical dependent variables include percentage correct and response times.
3. Description of the architecture and the operators to be used by the system. The **architecture** specifies the non-changeable structures of the models: for example, short-term memory and long-term memory, and whether items are subject to activation and decay. The **operators** specify the basic cognitive operations that can be carried out by the model; for example: “put an item into short-term memory,” “compare two items,” or “decrease the activation level of an item.” The scientific literature in psychology is used to specify the architecture and the operators, and many options are possible for both.

How Are the System’s Inputs and Outputs Represented?

Experimental inputs. The experimental protocol and the timeline consist of Lisp code; the stimuli are symbols (typically number or letters); the experiments results are vectors of real numbers. The **architecture** is specified by a virtual machine with **operators** supporting a simple interpreted language. The **outputs** are models with measures of goodness of fit.

The Space of Candidate Models that the System Searches

This (infinite) space consists of all potential programs that can be generated from the operators.

What Criteria are Used to Evaluate the Candidate Models?

A fitness function computes the match between the predictions of a model (e.g. percentage correct, response time) and the human data. Different variables can be given different weights, and criteria such as parsimony (size of the program/model) can also be used.

How are the Results Generated by the System Interpreted?

The models are expressed as operators for a high-level, symbolic cognitive architecture, and so are easily interpretable - either as abstract-syntax trees (in a Lisp-like symbolic-expression format) or as pseudo-code. However, one difficulty is that the models can be long and complicated, despite methods used to simplify them and reduce the number of similar models. Models can however be visualised as clusters based on syntactic similarity, with these clusters representing semantically different solutions to the task.

References

Addis, M., Lane, P. C. R., Sozou, P., & Gobet, F., (2019). Semi-automatic generation of cognitive science theories. In M. Addis et al., *Scientific discovery in the social sciences* (pp. 155-171). New York: Springer. [Link]

Bartlett, L. K., Pirrone, A., Javed, N., & Gobet, F. (2023). Computational scientific discovery in psychology. *Perspectives on Psychologogical Science, 18*, 178-189. [Link]

Lane, P. C. R., Bartlett, L., Javed, N., Pirrone, A., & Gobet, F. (2022). Evolving understandable cognitive models. *International Conference on Cognitive Modeling 2022*. [Link]

Lane, P. C. R., Sozou, P. D., Gobet, F., & Addis, M. (2016). Analysing psychological data by evolving computational models. In A.F.X. Wilhelm, H.A. Kestler (eds.), *Analysis of Large and Complex Data, Studies in Classification, Data Analysis, and Knowledge Organization, Proceedings of ECDA 2014* (pp. 587-597). New York: Springer. [Link]