Title
Genetic Programming for Developing Simple Cognitive Models

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Bartlett, Laura
Pirrone, Angelo
Javed, Noman
et al.

Publication Date
2023

Peer reviewed

Genetic Programming for Developing Simple Cognitive Models

Laura K. Bartlett (l.bartlett@lse.ac.uk)
Centre for Philosophy of Natural and Social Science, London School of Economics, UK

Angelo Pirrone (a.pirrone@lse.ac.uk)
Centre for Philosophy of Natural and Social Science, London School of Economics, UK

Noman Javed (n.javed3@lse.ac.uk)
Centre for Philosophy of Natural and Social Science, London School of Economics, UK

Peter C. R. Lane (p.c.lane@herts.ac.uk)
School of Physics, Engineering and Computer Science, University of Hertfordshire, UK

Fernand Gobet (f.gobet@lse.ac.uk)
Centre for Philosophy of Natural and Social Science, London School of Economics, UK

Abstract

Frequently in psychology, simple tasks that are designed to tap a particular feature of cognition are used without considering the other mechanisms that might be at play. For example, the delayed-match-to-sample (DMTS) task is often used to examine short-term memory; however, a number of cognitive mechanisms interact to produce the observed behaviour, such as decision-making and attention processes. As these simple tasks form the basis of more complex psychological experiments and theories, it is critical to understand what strategies might be producing the recorded behaviour. The current paper uses the GEMS methodology, a system that generates models of cognition using genetic programming, and applies it to differing DMTS experimental conditions. We investigate the strategies that participants might be using, while looking at similarities and differences in strategy depending on task variations; in this case, changes to the interval between study and recall affected the strategies used by the generated models.

Keywords: delayed-match-to-sample; genetic programming; memory; psychology

Introduction

In the field of psychology, a number of assumptions are regularly made with regard to the stimuli, experimental paradigms and cognitive mechanisms at play in experimental settings. For example, the popular delayed-match-to-sample (DMTS) task is often used to measure the memory abilities of a participant, usually without justifying a number of potentially critical experimental choices, such as the duration that each stimuli is presented for, the type of stimuli, and the time delay between stimuli presentation. As this paradigm has been prevalent in memory research for a number of years, it is important to explore the interplay between these factors and the interacting cognitive mechanisms that give rise to behaviour on this task.

Importantly, this is just one example of a seemingly simple task that is consistently used in psychology without critical consideration. This paper will outline how a genetic programming approach (in particular using the GEMS methodology – genetically evolving models in science, Addis, Gobet, Lane, and Sozou (2019); Frias-Martinez and Gobet (2007); Lane, Sozou, Gobet, and Addis (2016)) can allow greater insights into the cognitive mechanisms underlying behaviour on the DMTS task. The methodology has shown some success in other domains (e.g., decision-making, Pirrone and Gobet (2020)), and is applied here to two experiments with different timings and stimuli. A discussion of the models and strategies generated by GEMS will then propose some directions of research for aiding future researchers in their use of the DMTS task. Further, we will discuss the ways in which this methodology can be used flexibly for different tasks and different domains of study.

The delayed-match-to-sample task

The delayed-match-to-sample (DMTS) task is an important and popular paradigm in cognitive psychology. Typically, the task begins with a brief presentation of a stimulus (study phase; e.g. a visually presented circle; see Figure 1). After a delay, two comparison stimuli are presented (recall phase), where the participant must choose which item matched the initial study phase stimulus. Participant accuracy and response time (RT) are most commonly recorded. In the literature, this basic methodology has been varied, allowing us to tease apart elements of the task and better understand the underlying cognitive mechanisms. For example, variations in the number of stimuli to be remembered (e.g., Olsen et al. (2009); Raabe, Fischer, Bernhardt, and Greenlee (2013)), differences in the type of stimuli (e.g., shapes, faces, arbitrary symbols), as well as alterations in the presentation duration of stimuli and the delay between study and recall phases. An alternative non-match to sample paradigm has also been developed, where the novel stimulus must be selected, providing further insight into human cognition and the abstract concepts of 'same' and 'different' (Hochmann, Mody, & Carey, 2016).

The simplicity of this experimental paradigm has led to its widespread use with a variety of populations, including patients with schizophrenia (e.g., Lencz et al. (2003); Minzenberg, Laird, Thelen, Carter, and Glahn (2009)), infants (Hochmann et al., 2016) and animals (see Lind, Enquist, and Ghirlanda (2015) for a review), and the task is included in the Cambridge Neuropsychological Test Automated Battery (CANTAB). The DMTS has also been consistently used

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

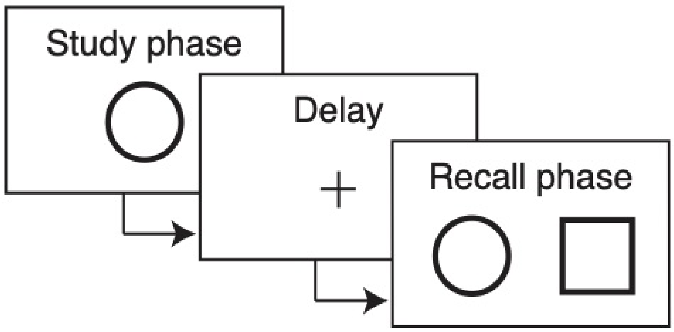

Figure 1: An example trial for the DMTS experiment

in working-memory research, particularly for neuroimaging experiments (see Daniel, Katz, and Robinson (2016) for a meta-analysis).

The wealth of experimental data makes a simple representative model of the DMTS difficult to develop under a single unifying framework. Despite consistent use of the task, the basic cognitive mechanisms and strategies used are not often explored. Rather, the task is most commonly used to tap specific underlying mechanisms. The DMTS paradigm in fact represents the complex interplay between psychological constructs often researched in a modular fashion, for example experiments focusing on working memory might fail to consider the decision-making elements of the task when interpreting data, while others may only focus on decision making. It is important to bring increasingly specialised areas of research together to form a more unified representation of human cognition. The current research uses symbolic cognitive modelling and genetic programming techniques to generate novel models of behaviour that can inform future research and expand upon existing models.

While the DMTS task has been modelled previously, the focus has primarily been on low-level brain functionality and neural-net modelling. In particular, the DMTS task has been used to demonstrate the efficacy of a neural model with the focus of addressing issues surrounding the interpretation of functional imaging data (Tagamets & Horwitz, 1998). For example, Corbitt, Ulloa, and Horwitz (2018) refined the large scale neural model developed by Tagamets and Horwitz (1998), simulating laminar fMRI activity and demonstrating that it can be utilised for layer-specific fMRI interpretation. Neural network models have been used to perform many variations of the DMTS, including visual (Tagamets & Horwitz, 1998), spatial (Lee Moody, Wise, Pellegrino, & Zipser, 1998) and auditory tasks (Wen, Ulloa, Husain, Horwitz, & Contreras-Vidal, 2008).

Alongside neural models, ACT-R (Anderson et al., 2004) has been applied to the DMTS task. ACT-R is a cognitive architecture concerned with modelling human behaviour, striving for a unified model of cognition as advocated by Newell (1990). This architecture consists of central production systems, which respond to information made available in modular buffers; for example, a visual buffer accessing information from the visual field and a retrieval buffer accessing information from long-term memory. To explore visual rehearsal in ACT-R, Cebulski and Somers (2014) used the DMTS paradigm to compare two models, finding that minimising interference from previous trials led to a better fit to the behavioural data. This result is clearly a useful advancement on DMTS task knowledge. However, it only addresses one component of the task.

Applying the GEMS methodology

The current paper describes new models explaining cognitive behaviour and operations during the DMTS task, generated using genetic programming techniques (Koza, 1992). This method generates and evolves a population of models over a number of generations, altering the models through different operations (such as crossover, where some components of models are swapped) to find the best models against a given fitness function. (Note that genetic programming is different from *genetic algorithms*. While both belong to the family of evolutionary computation, genetic programming evolves entire programs, which can be of variable size, whilst genetic algorithms represent solutions as fixed-length strings of bits.) Models developed with genetic programming can complement and advance the insights provided by previous DMTS models. Firstly, it reduces the potential bias involved in generating a theory of behaviour – researchers are often focused on one area of research and so may overlook other influential cognitive mechanisms at play. Secondly, models are generated and fit to different behavioural datasets, compared to modelling protocols where a model may be generated to account for only one source of data. Similarly, the models generated in the current paper are evolved for different versions of the DMTS task, so that models are able to account for a range of task variations. Finally, parameters are not adjusted to derive the best fit to the data, as is common for modelling in psychology; rather, the fundamental cognitive operators driving the effects are explored. The advantage of generating multiple models in this way is that they can indicate novel research directions and inform future research.

Models were generated using GEMS, a genetic programming approach for developing models of human cognitive behaviour. GEMS is a semi-automated system designed to discover potential computational models based on patterns in experimental data, to better characterise cognitive operations underlying tasks, and as such addresses some of the difficulties of developing psychological models (for more details regarding GEMS, see Lane et al. (2016); Frias-Martinez and Gobet (2007).

The GEMS system architecture includes (i) a clock, (ii) a short-term memory store consisting of three items, (iii) an indication of where attention is focused, (iv) a 'salient' buffer that encodes what is perceived as salient in the visual field, (v) a 'current' buffer where items that have been attended are stored, and (vi) a response buffer where a response is held until the trial ends. Generated models consist of sequences of operators, which are basic cognitive operations (see Table 1 for a list of the operators included in the current GEMS research) that interact with the cognitive architecture to provide potential strategies for the task. The timings for each operator type were derived from Card, Moran, and Newell (1983). GEMS also includes a short-term memory decay function, such that old items are forgotten unless rehearsed.

Models were generated and compared against human experimental data, with a fitness value computed for each model, weighting response accuracy the highest (weight = 0.7), followed by response time (weight = 0.2), and model size (valuing smaller models, weight = 0.1). Values were chosen to reflect the importance of accuracy, considering the typically high accuracy levels found for human participants in DMTS tasks. These weights can be altered to optimise for different goals, which could result in different strategies being selected. For the current experiment, these values produced more models that found an accurate solution with a good fit. Lower fitness values indicate a better fit to the data. A phased approach to fitness was used, whereby models were optimised for accuracy first, and upon reaching an accuracy threshold proceeded to be optimised for response time, and then model size. This is because previous research with GEMS found that non-phased evolution does not always converge well (Lane, Bartlett, Javed, Pirrone, & Gobet, 2022). Best fitting models were then retained for the next generation of models.

DMTS modelling

The GEMS system ran against data from two published research papers that featured the DMTS task (Chao, Haxby, and Martin (1999); Edwards, Boyer, Bell, and Sturz (2016), see Table 2 for experimental details). While only two experiments were included for the current paper, GEMS can be applied to many more datasets and account for multiple experiments. Both experiments followed the trial structure outlined in Figure 1, but critically differed in the duration of delay between the study and recall stimuli. While the type of stimuli image also differed between the experiments (i.e., faces, tools, or simple shapes), for the current paper they were coded symbolically without exploring these differences. Future work will assess the influence of stimuli type on the strategies outputted by GEMS, for example by encoding stimulus complexity. To determine an appropriate threshold for short-term memory decay, GEMS was initially run with a series of threshold values using both datasets, with a population of 15,000 over 1,500 generations. The best model fits were found with a threshold value of 0.1. As such, this value was used for subsequent runs.

As seen in Figure 2, Dataset-1 reached a good fit quicker than Dataset-2; however, both found stable, good fitting models within the first 250 generations. The sudden improvement in fitness merits some comments. As genetic programming searches for the best solution, the fitness landscape can shift abruptly due to the interaction of mutation and cross-over operators. This is because GP manipulates discrete structures (programs) and progress in the search space is not incremental, as for example with gradient descent in neural networks. This is reinforced by the relatively simple problems GEMS deals with, where a certain composition of operators might reach 50% correct, while a small change, e.g., putting the correct items in STM, leads to 100% correct without intermediary partial solutions.

Models from the final generation were selected for further analysis. From these models, those with the best fit were processed and cleaned, where operators that did not affect the models (rather, they were there to take up time) were removed and replaced with a relevant 'wait-' operator. Duplicate models were then removed, leaving a number of unique models (9 remained for Dataset-1, 10 remained for Dataset-2).

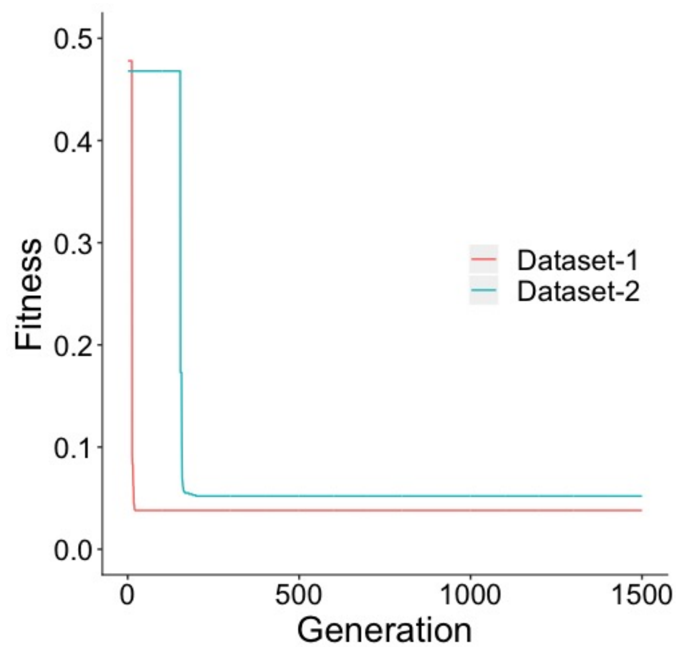

Figure 2: Model fitness by generation. Lower values indicate a better fit to the experimental data.

The models generated for each dataset showed a fairly similar strategy, mainly differing in the order that operations were conducted. The heat maps shown in Figures 3 and 4 confirm this similarity for the final models in each experiment; models were compared against each other (with each model represented by a number in the Figures), and the syntactic similarity of the components determined the similarity statistic. Lower numbers indicate more dissimilar components, and a value of 1 indicates the models are the same (as shown in the heat maps when a model is compared against itself). All models began by using the detect-attend-putstm operator, placing the study phase stimulus into STM slot-1. For Dataset-2 (see Figure 6 for an example tree), attention was then shifted to the left, and a 'right' response was given. After using the wait operators, the recall phase stimuli on the left side was

Table 1: Overview of the operators used by GEMS. Each operator type had a set time (in milliseconds, ms) as follows: input (100 ms), output (140 ms), cognitive (70 ms), STM (50 ms), syntax (0 ms).

Name	Function	Type
attend	Sets model 'current' to what is stored in model 'salient'	cognitive
move-att-X	Shift 'attention' to a location in the visual display (X $\in \{centre, left, right, clockwise, counterclockwise\}$)	cognitive
if-current-stm-RX	Compare item in model 'current' with those in STM; if find a match, respond in line with X (X $\in \{STMitem/\text{'}current' - item\}$)	cognitive
if-stmN-R	Compares stm item N with other items in STM, if match, respond in line with matching item (N $\in \{1,2,3\}$)	cognitive
compare-M-N-p	Predicate comparing value of STM items M and N (M$\neq$N $\in \{1,2,3\}$)	cognitive
compare-current-N-p	Predicate comparing value of model 'current' and STM items N (N $\in \{1,2,3\}$)	cognitive
current-X-p	Predicate for stimulus type of model 'current' (X $\in \{target, stimulus\}$)	cognitive
detect	Puts the item at the attention location into model 'salient'	input
respond-X	Sets model 'response' to X (X $\in \{left, right, centre, \text{'}current' - item - location\}$)	output
rehearsal-N	Updates item-time in STM item N (N $\in \{1,2,3\}$) to current model 'clock'	stm
retrieve-N	Sets model 'current' to STM item N (N $\in \{1,2,3\}$)	stm
retrieve-X	Sets model 'current' to STM item matching type X, (X $\in \{target, stimulus\}$)	cognitive
nil	Sets model 'current' to nil	cognitive
put-stm	Pushes value in model 'current' and model 'clock' to STM slot 1	cognitive
dotimes-N	Repeats a given expression (N $\in \{2,3,4\}$)	syntax
if	Executes condition, executes one of two expressions depending on the condition	syntax
prog-N	Sequence of expressions (N $\in \{2,3,4\}$)	syntax
wait-N	Advances model clock (N $\in \{25, 50, 100, 200, 1000, 1500,$ 0.5-trial-length, 0.25-trial-length, 0.1-trial-length$\}$)	syntax
while-N	Repeats an expression for a set time in ms (N $\in \{100, 200\}$)	syntax
detect-attend	Executes the detect operator, then the attend operator	combined
detect-attend-putstm	Executes the detect operator, the attend operator and then the put-stm operator	combined

placed into STM slot-1, moving the study stimuli to slot-2. The strategy then makes use of the 'if-current-stm-Rc' operator, which looks through STM and if it finds a match to the item in 'current', will respond with the location of the 'current' item.

Similarly, the Dataset-1 models (see Figure 5) moved attention to the right, gave a 'left' response, and used the 'detect-attend-putstm' operator a second time to move the study phase stimuli to STM slot-2. The recall stimulus located on the right was then placed into STM slot-1, following which the model used the 'if-stm3-R' operator, comparing the item in STM3 with items saved in the other STM slots.

Essentially, models for both experiments are utilising the same strategy – put the item in STM, pick a direction to focus attention on (left or right), and have a response for the opposite side ready if the attended stimulus does not match. As such, rather than having to engage with all three items in the display, the strategy is to only look at one in order to make a decision (this is consistent with a previous implementation of GEMS for the DMTS task that used a simpler setup and only one dataset (Frias-Martinez & Gobet, 2007), suggesting that this is a robust strategy). The key experimental difference between the two datasets is the interval between study and recall phase. The key difference between the strategies generated by GEMS however is whether the models put all items into STM (as with the models of Dataset-1) or instead just put the study item in STM, comparing this against the stimulus currently attended. These strategies differ in their reliance on STM, and suggest that a better test of memory capabilities would require more stimuli. In particular, these models suggest that participants are responding to a 'same' judgement, and doing nothing in response to a 'different' judgement, rather than looking at both stimuli and responding left or right accordingly. This changes the dynamics at play in the experiment, and brings into question how participants might be using different strategies to the ones assumed by researchers.

Discussion

While GEMS can provide insight into the potential strategies that are used in experiments, it has a number of limitations. One limitation is that the evolutionary system currently only works across the syntax trees of the model programs; the numeric parameters, such as the timing of operators, are left as hyper-parameters and were not optimised in this paper. As noted above, these timing parameters are derived from Card et al. (1983). However, as the generation of models in part relies on response time data, any changes to these values could result in considerably different models being generated. One

Table 2: DMTS experiments included in GEMS

Details	**Dataset 1**	**Dataset 2**
Experiment	Chao et al. (1999)	Edwards et al. (2016)
Number of trials	60	60
Delay phase	500ms	5000ms
Study stimulus presentation time	1000ms	1000ms
Recall stimulus presentation time	2000ms	1500ms
Type of stimuli	black-and-white photographs of animals, tools, faces and houses	unfilled simple shapes
Mean response time	767ms	485ms
Mean accuracy	0.957	0.94

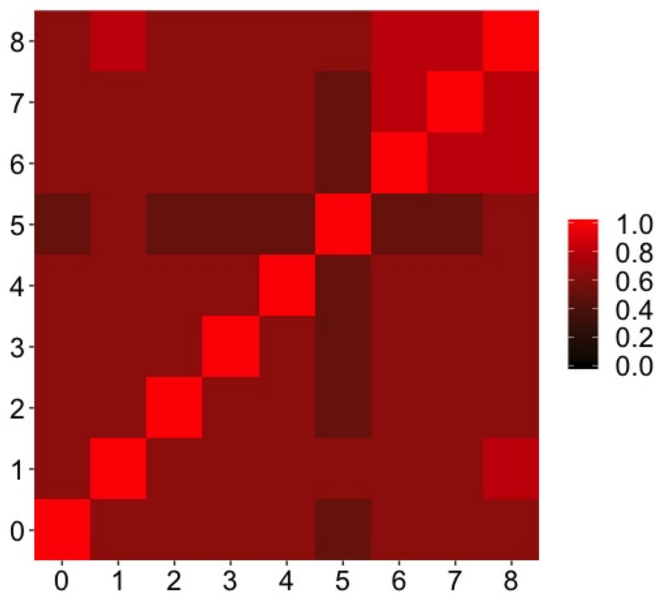

Figure 3: Syntactic similarity of components for the final models generated for Dataset-1. Each number on each axis represents a model.

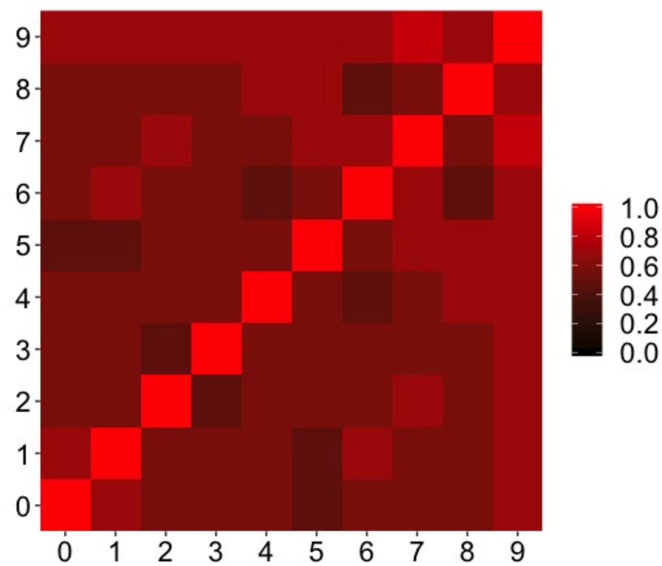

Figure 4: Syntactic similarity of components for the final models generated for Dataset-2. Each number on each axis represents a model.

likely area of future work within GEMS is to integrate mechanisms to optimise hyper-parameters such as the operator timings, in order to ascertain differences in the models generated. A second limitation, inherent to all cognitive modelling efforts, is that the operators used by the system are created by researchers, and as such they are liable to any biases researchers may have. For example, if a researcher has a focus on memory research, they might be more likely to provide the system with a disproportionately large amount of memory operators. Further, the way the operators are coded could be a source of bias, where some researchers might disagree over the mechanisms for different processes. However, the GEMS methodology can actually turn this weakness into a strength: by combining operators from different research domains, it makes possible cross-domain theoretical contributions, which in turn could provide a powerful means to combine current psychological knowledge.

Future analyses can look at how additional changes in the experimental conditions, as well as changes to the model architecture and operator set used with GEMS in this paper, can impact the models that are generated. For example, limiting the models to one response (rather than the current models, where different responses can be made throughout the trial, and only the final response is accepted) would require substantively different models to solve the task. Further, different experimental dynamics have been explored, for example giving two study phase stimuli, followed by two recall stimuli, one of which matches one of the study stimuli (Olsen et al., 2009). These questions will continue to be explored by the GEMS methodology, alongside simple experiments across different experimental domains.

Conclusions

There is a need to not take for granted the usual theoretical interpretations of experimental paradigms and to consider other

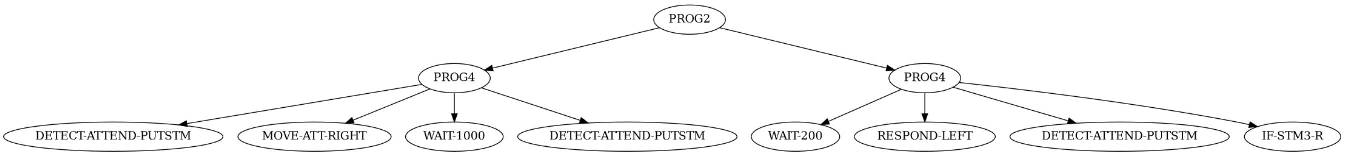

Figure 5: Typical example of best-fitting model for Dataset-1. Fitness = 0.038; Accuracy = 1; Response Time = 770ms

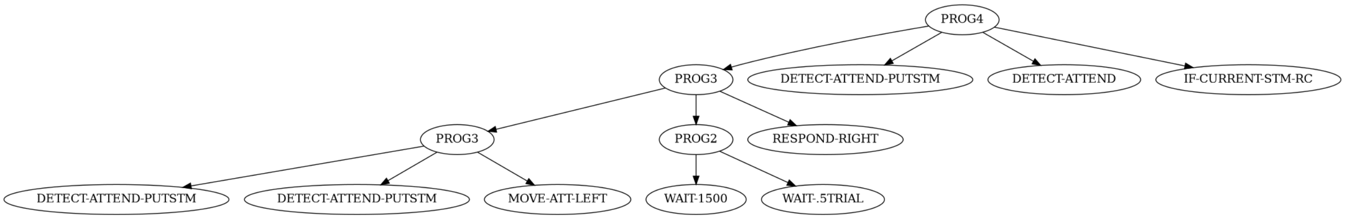

Figure 6: Typical example of best-fitting model for Dataset-2. Fitness = 0.052; Accuracy = 1, Response Time = 490ms

factors that might affect behavioural results. While new and compelling experiments are being designed for more interesting research questions, it is still important to make time for these foundational paradigms. GEMS can provide a simple tool to search the vast model space, which includes the combination of all possible models with the given operators, a considerably larger space than the parametrisation of a single model. Further, GEMS could function as a collaborative project between experts in different research domains – for example researchers focusing on decision-making contributing operators and testing their theories, alongside memory researchers, and even extending outside of psychology, allowing interdisciplinary collaboration.

Acknowledgements

Funding from the European Research Council (ERC-ADG-835002—GEMS) is gratefully acknowledged.

References

Addis, M., Gobet, F., Lane, P. C., & Sozou, P. D. (2019). Semi-automatic generation of cognitive science theories. In M. Addis, P. C. Lane, P. D. Sozou, & F. Gobet (Eds.), *Scientific discovery in the social sciences.* Springer.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060. doi: 10.1037/0033-295X.111.4.1036

Card, S. K., Moran, T. P., & Newell, A. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Erlbaum.

Cebulski, S., & Somers, S. (2014). ACT-R models of a delayed match-to sample task. *Proceedings of the Annual Meeting of the Cognitive Sciences Society*, *36*(36), 302–306.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, *2*(10), 913–919. doi: 10.1038/13217

Corbitt, P. T., Ulloa, A., & Horwitz, B. (2018). Simulating laminar neuroimaging data for a visual delayed match-to-sample task. *NeuroImage*, *173*, 199–222. doi: 10.1016/j.neuroimage.2018.02.037

Daniel, T. A., Katz, J. S., & Robinson, J. L. (2016). Delayed match-to-sample in working memory: A BrainMap meta-analysis. *Biological Psychology*, *120*, 10–20. doi: 10.1016/j.biopsycho.2016.07.015

Edwards, J. E., Boyer, T. W., Bell, Z. K., & Sturz, B. R. (2016). Isolated processing of geometric shapes and their corresponding shape words: Evidence from a delayed match-to-sample task. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(8), 1088–1103. doi: 10.1037/xhp0000210

Frias-Martinez, E., & Gobet, F. (2007). Automatic generation of cognitive theories using genetic programming. *Minds and Machines*, *17*(3), 287–309. doi: 10.1007/s11023-007-9070-6

Hochmann, J. R., Mody, S., & Carey, S. (2016). Infants' representations of same and different in match- and non-match-to-sample. *Cognitive Psychology*, *86*, 87–111. doi: 10.1016/j.cogpsych.2016.01.005

Koza, J. R. (1992). *Genetic programming: On the programming of computers by means of natural selection*. MIT Press.

Lane, P. C. R., Bartlett, L., Javed, N., Pirrone, A., & Gobet, F. (2022). Evolving understandable cognitive models. In T. C. Steward (Ed.), *Proceedings of the 20th international conference on cognitive modelling.*

Lane, P. C. R., Sozou, P. D., Gobet, F., & Addis, M. (2016). Analysing psychological data by evolving computational models. In A. Wilhelm & H. Kestler (Eds.), *Analysis of large and complex data. studies in classification, data analysis, and knowledge organization.* Springer, Cham. doi: 10.1007/978-3-319-25226-1

Lee Moody, S., Wise, S. P., Pellegrino, G. D., & Zipser, D. (1998). A model that accounts for activity in primate frontal cortex during a delayed matching-to-sample task. *Journal of Neuroscience*, *18*(1), 399–410. doi: 10.1523/JNEUROSCI.18-01-00399.1998

Lencz, T., Bilder, R. M., Turkel, E., Goldman, R. S., Robinson, D., Kane, J. M., & Lieberman, J. A. (2003). Impairments in perceptual competency and maintenance on a visual delayed match-to-sample test in first-episode schizophrenia. *Archives of General Psychiatry*, *60*(3), 238–243. doi: 10.1001/archpsyc.60.3.238

Lind, J., Enquist, M., & Ghirlanda, S. (2015). Animal memory: A review of delayed matching-to-sample data. *Behavioural Processes*, *117*, 52–58. doi: 10.1016/j.beproc.2014.11.019

Minzenberg, M. J., Laird, A. R., Thelen, S., Carter, C. S., & Glahn, D. C. (2009). Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Archives of General Psychiatry*, *66*(8), 811–822. doi: 10.1001/archgenpsychiatry.2009.91

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

Olsen, R. K., Nichols, E. A., Chen, J., Hunt, J. F., Glover, G. H., Gabrieli, J. D., & Wagner, A. D. (2009). Performance-related sustained and anticipatory activity in human medial temporal lobe during delayed match-to-sample. *Journal of Neuroscience*, *29*(38), 11880–11890. doi: 10.1523/JNEUROSCI.2245-09.2009

Pirrone, A., & Gobet, F. (2020). Modeling value-based decision-making policies using genetic programming: A proof-of-concept study. *Swiss Journal of Psychology*, *79*(3–4), 113–121. doi: 10.1024/1421-0185/a000241

Raabe, M., Fischer, V., Bernhardt, D., & Greenlee, M. W. (2013). Neural correlates of spatial working memory load in a delayed match-to-sample saccade task. *NeuroImage*, *71*, 84–91. doi: 10.1016/j.neuroimage.2013.01.002

Tagamets, M. A., & Horwitz, B. (1998). Integrating electrophysiological and anatomical experimental data to create a large-scale model that simulates a delayed match-to-sample human brain imaging study. *Cerebral Cortex*, *8*(4), 310–320. doi: 10.1093/cercor/8.4.310

Wen, S., Ulloa, A., Husain, F., Horwitz, B., & Contreras-Vidal, J. L. (2008). Simulated neural dynamics of decision-making in an auditory delayed match-to-sample task. *Biological Cybernetics*, *99*(1), 15–27. doi: 10.1007/s00422-008-0234-0