A process model for human problem solving in puzzles

Jeroen Olieslagers (jo2229@nyu.edu)
Center for Neural Science, New York University
New York, United States

Zahy Bnaya (zahy.bnaya@gmail.com)
Center for Neural Science, New York University
New York, United States

Wei Ji Ma (weijima@nyu.edu)
Center for Neural Science and Department of Psychology, New York University
New York, United States

Abstract

What goes on in your head when you are solving a problem? Advances towards this question will allow us to better understand and predict complex human behaviour as well as obtain estimates for latent abilities such as how deep a person plans. We choose to contribute to this question by studying problem solving in puzzles: well-controlled and tractable environments, in particular the game of Rush Hour™. We present three distinct process models for solving puzzles (models describing people's reasoning), fit these on subjects' data and discuss the strengths and shortcomings of each. We find that models which assume a constant depth of planning fail to capture some important trends in summary statistics. We hypothesize that people dynamically change their depth of planning and propose models to incorporate this. This work outlines a path to connect the field of human problem solving using rigorously studied puzzles to the normative models used to study human behaviour in games (e.g. (van Opheusden et al., 2021)).

Keywords: planning; games; process models; computational modelling; Rush Hour; puzzles

Introduction

Problem solving is an ability all living organisms are capable of. Whether it is a bacteria moving up a chemical gradient for nutrients (Palma, Gutiérrez, Vargas, Parthasarathy, & Navarrete, 2022), a younger chimpanzee deceiving the alpha male for extra food (Byrne & Whiten, 1992), or a person taking out a first aid kit to treat a wound, problem solving provides an evolutionary benefit for all. Studying how problems are solved could help explain otherwise strange behaviour and could allow us to infer latent thoughts and abilities just by looking at behaviour. While the field has relied on retrospective studies in the past (Newell & Simon, 1972; de Groot, 1966), computational models have proven to be a great tool to analyse larger populations and provide more rigorous inferences (Kolling, Scholl, Chekroud, Trier, & Rushworth, 2018; Snider, Lee, Poizner, & Gepshtein, 2015).

We chose to study Rush Hour™ for three reasons: it strikes a good balance between complexity and tractability, it is abstract enough such that information search and perceptual effects are small, and finally because it presents attractive cause-effect relationships (a useful property to study construals (Ho et al., 2021)). To solve a Rush Hour™ puzzle, you have to move the red car all the way to the right out of the board. To do this, you have to first move the cars that are in the way such that a path exists for the red car to leave. Horizontally oriented cars may only make horizontal moves, and vertically oriented cars may only make vertical moves. Cars may never overlap, and are not allowed to be lifted off the board. Figure 1 shows an example puzzle and its associated state space. In this state space, every node is a unique board position and the blue nodes show the optimal solution.

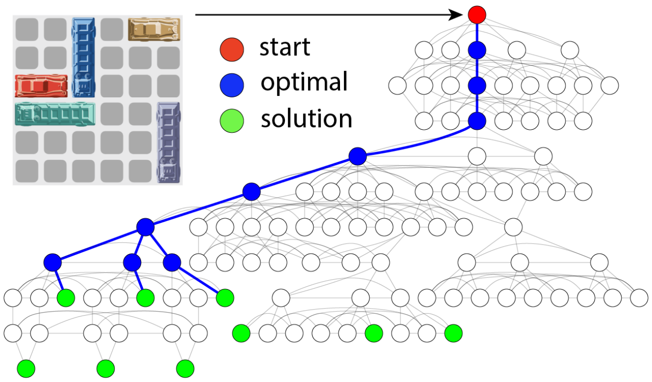

Figure 1: Rush Hour™ example puzzle and associated state space. The red node is the starting position, the blue nodes are states on the optimal paths to the optimal solutions and the green nodes are solved states (goal states).

Results

We collected data by recruiting subjects (N=42) to play Rush Hour™ in an online web-based experiment. Subjects completed as many puzzles as they could in an hour. They were allowed to skip and restart puzzles and were awarded a bonus based on how many puzzles they had completed to incentivize completing as many puzzles as possible. Throughout this section, we will be referring to two main components of a Rush Hour™ state: the distance (in number of moves) to the nearest solution d_{goal} and the number of possible actions n_A. Parameters are fit on a per-subject basis.

Heuristic model

The first model we consider is a heuristic model, where for a given state $s \in \mathcal{S}$, we assign a Q-value to each action we could take $a \in \mathcal{A}(s)$, resulting in state $s_i' \in \mathcal{S}$. Formally:

$$p(a|s,\lambda,\beta) = \frac{\lambda}{|\mathcal{A}(s)|} + (1-\lambda)\frac{e^{\beta Q(s,a)}}{\sum_{a_i \in \mathcal{A}(s)} e^{\beta Q(s,a_i)}} \quad (1)$$

where $Q(s,a) = -\text{heuristic}(s')$ and out of the many heuristics we tried, the "mean-leave-one-out" heuristic performed best. In this heuristic, we remove one car (except the red car) and compute d_{goal}. This is repeated for all cars and the average value of d_{goal} is returned. The β parameter controls the softmax temperature (noise in picking maximum Q value) and λ is the lapse rate. The rationale for this model is that people often don't look at the full puzzle, but rather consider a subset of cars to plan with. Figure 2 (red area) shows the summary statistics of model simulations from all states visited by the subjects, which is considerably better than chance but still struggles at the lower end of d_{goal}. This is a result of the β parameter being constant for all states.

Eureka model

To allow for the differentiation of "hard" and "easy" states and inspired by patterns seen in the data, the second model we investigate called the "Eureka" model, works by specifying a depth d, and if the value of d_{goal} in a state is less than or

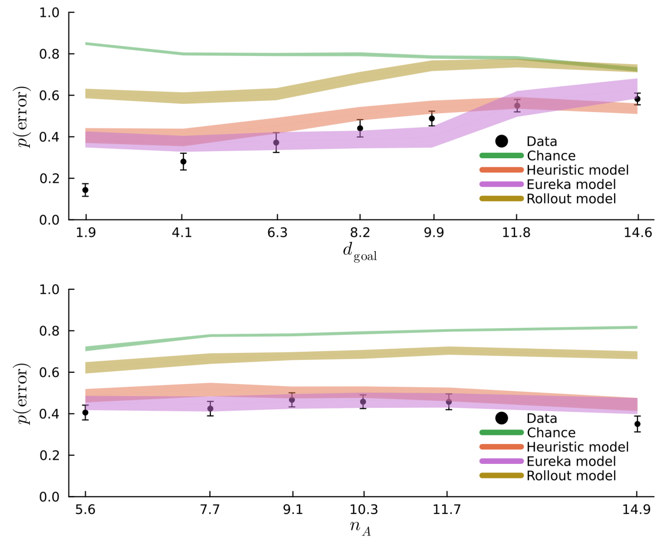

Figure 2: Probability of choosing an optimal move for subjects (mean $\pm 2 \times$ SEM in black), a random agent (green), as well as the three models. The top panel has distance to the nearest solution binned on the x-axis (average bin value displayed) and the bottom panel has the number of possible actions on the x-axis.

equal to d, the subject takes an optimal action, otherwise they act randomly. Formally:

$$p(a|s,\lambda,d) = \begin{cases} \frac{1}{\mathcal{A}(s)}, & d_{\text{goal}}(s) > d \\ \frac{\lambda}{\mathcal{A}(s)} + (1-\lambda)\frac{\mathbb{1}_{\mathcal{A}_{\text{opt}}(s)}(a)}{\mathcal{A}_{\text{opt}}(s)}, & d_{\text{goal}}(s) \leq d \end{cases} \quad (2)$$

where $\mathcal{A}_{\text{opt}}(s)$ is the set of optimal actions in state s (optimal meaning that they reduce the value of d_{goal} by one). From the data, it looks like subjects perform something akin to a random walk, followed by a strong drift towards the solution ($d_{\text{goal}} = 0$), which is why we believed this model would improve upon the heuristic model. The "transition" point where subjects go from random search to their "eureka" moment where they found the solution is parameterized by d. In figure 2 we again see that this model cannot capture the low end of d_{goal} (purple area). The reason for this is due to the high value of λ (changing lambda only shifts this flat line up or down, but does not make the fit better). An explanation could be that subjects don't have a fixed value of d for puzzles or maybe even states.

Rollout model

To allow d to vary more naturally, we investigate a rollout model (similar to that of (van Opheusden et al., 2021)). According to this model, subjects plan by imagining a set of random moves, terminating with probability γ after each move, and then repeating this rollout procedure k times (figure 3). This model has an explicit planning component, which we believe is key for performing well in Rush Hour™. The fitting procedure in this case is a lot more elaborate (van Opheusden, Acerbi, & Ma, 2020). We found no good combination of γ and k that provides a better fit than the previous two models. As

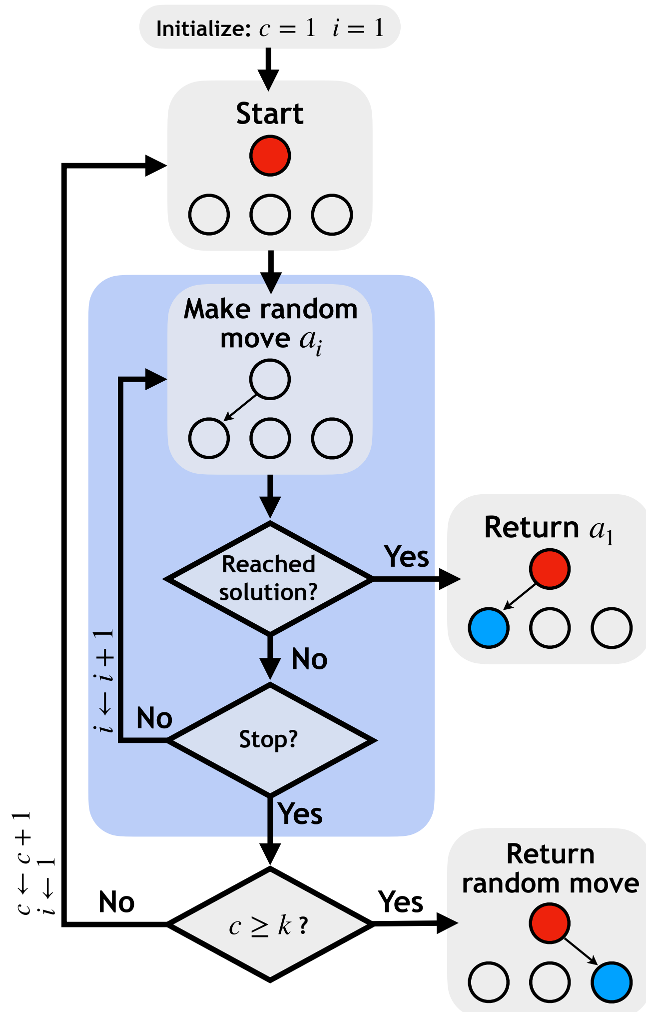

Figure 3: Flowchart for rollout model. In this case, "Stop" is "Yes" with probability γ on each iteration. The red node is the current state and the blue node is the move the model commits to.

seen in figure 2 (yellow area), this model actually performs worse than the previous two. This is due to the random nature of the rollouts, something humans are unlikely to do.

Conclusion

In this work, we came up with and fitted three distinct models of problem solving to individual moves from subjects playing Rush Hour™. These models can recapitulate summary statistics, provide move level predictions and provide us with subject-level parameters such as depth of planning. The next steps are three-fold: first, we would like to come up with additional summary statistics of states besides the two we presently have, since n_A seems fairly uninteresting as well as additional dependent variables besides $p(\text{error})$. Secondly, we would like to pursue the rollout model further, but change the random action selection to be based on heuristics so the search is directed. Finally, we wish to generalize this work to other puzzles.

References

Byrne, R. W., & Whiten, A. (1992). Cognitive evolution in primates: evidence from tactical deception. *Man*, 609–627.

de Groot, A. (1966). Perception and memory versus thought: Some old ideas and recent findings. *Problem solving*, 19–50.

Ho, M. K., Abel, D., Correa, C. G., Littman, M. L., Cohen, J. D., & Griffiths, T. L. (2021). Control of mental representations in human planning. *arXiv e-prints*, arXiv–2105.

Kolling, N., Scholl, J., Chekroud, A., Trier, H. A., & Rushworth, M. F. (2018). Prospection, perseverance, and insight in sequential behavior. *Neuron*, *99*(5), 1069–1082.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Palma, V., Gutiérrez, M. S., Vargas, O., Parthasarathy, R., & Navarrete, P. (2022). Methods to evaluate bacterial motility and its role in bacterial–host interactions. *Microorganisms*, *10*(3), 563.

Snider, J., Lee, D., Poizner, H., & Gepshtein, S. (2015). Prospective optimization with limited resources. *PLoS computational biology*, *11*(9), e1004501.

van Opheusden, B., Acerbi, L., & Ma, W. J. (2020). Unbiased and efficient log-likelihood estimation with inverse binomial sampling. *PLoS computational biology*, *16*(12), e1008483.

van Opheusden, B., Galbiati, G., Kuperwajs, I., Bnaya, Z., Ma, W. J., et al. (2021). Revealing the impact of expertise on human planning with a two-player board game.