Modelling insight: The case of the nine-dot problem

Thomas C. Ormerod (t.ormerod@sussex.ac.uk)
School of Psychology, University of Sussex, Falmer, BN1 9QH UK

Patrice Rusconi (p.rusconi@surrey.ac.uk) & Adrian Banks (a.banks@surrey.ac.uk)
School of Psychology, University of Surrey, Guildford, GU2 7XH UK

James N. MacGregor (jmacgreg@uvic.ca)
School of Public Administration, University of Victoria, Victoria BC, Canada

Abstract

A number of frameworks for capturing insight phenomena have been proposed, but there are no executable models of knowledge-lean insight problem-solving. Here, an ACT-R model is presented for the nine-dot problem, which implements the Criterion for Satisfactory Progress theory for this problem. The model has two main components: a mechanism for searching for possible moves in the problem representation, and a mechanism for expanding the search to discover new moves not immediately available in the initial problem representation. The model accounts for key phenomena including impasse, fixation and the 'aha' moment, as well as predicting the relative difficulty of different problem variants.

Keywords: Insight; ACT-R; Problem-solving; 9-dot problem.

Introduction

Recent theories of insight are of two kinds: knowledge-based accounts such as Representational Change Theory (RCT: Knoblich et al., 1999), in which problem difficulty is mediated by inappropriate knowledge; and strategic accounts such as Criterion for Satisfactory Progress theory (CSP: MacGregor, Ormerod, & Chronicle, 2001), in which problem difficulty is mediated by search for moves that maximize progress towards a goal. Most researchers agree both knowledge and strategy are essential for explaining insight (e.g., Kershaw & Ohlsson, 2004), and integrated frameworks have been proposed (e.g., Hélie & Sun, 2010). However, progress is hampered by a lack of executable models of knowledge or strategy mechanisms.

An ACT-R model of 9-dot problem-solving

Here we present an ACT-R implementation of CSP for the 9-dot problem ("Draw four connected straight lines to cancel 9 dots arranged in a 3x3 grid"). The problem is notoriously difficult, with solution rates < 5%. Although knowledge-based accounts predict that a given first line extending beyond the array should serve as a solution cue, a first line remaining within the square leads to higher solution rates. An internal first line leads to earlier criterion failure, which motivates change of search strategy (MacGregor et al, 2001, Expts. 4-5). Our ACT-R model implements two heuristics: *maximisation* and *minimisation* (Chronicle, MacGregor, & Ormerod, 2004; MacGregor et al., 2001; Ormerod et al, 2013) to solve the problem.

Search through maximisation

Under maximisation, individuals select moves that appear most promising to achieve a hypothesised goal. Progress is monitored against a criterion derived from the problem statement. With the nine-dot problem, an initial line connecting three dots represents an implementation of a maximising heuristic because individuals cancel the most dots in a single move. Progress made with this move is evaluated against a criterion equal to the number of remaining dots divided by the number of remaining lines.

To implement maximization, the model searches for previously unattended and uncancelled dots at random and tests how many are cancelled by each move between dots. If it cancels more than the current best move, this move is stored in the imaginal buffer (where problem representations are stored). Then the cycle repeats until all unattended dots are inspected, when search for another one fails. This triggers a reset of all uncancelled dots to 'unattended'.

The line stored in the imaginal buffer represents the move that maximises progress. This line is checked against the progress-monitoring criterion. This criterion derives from two main sources of information in the initial representation: the number of dots and number of lines to be drawn. The criterion is equal to the number of remaining dots divided by the number of remaining lines. In the production (P PROMISING), if the number of cancelled dots is greater than the criterion, then the move is labelled as 'promising' (status slot of the imaginal chunk). Otherwise, in the production (P EXHAUSTED), there is criterion failure and the move is categorized as 'exhausted' in the 'status' slot of the imaginal chunk, and another best move is looked for. If the move has a promising status, the model draws a line. This is the first move. After a line is drawn, the model begins again the first cycle selecting previously unattended and uncancelled dots at random. The first stage stops either when there are no more dots to be cancelled or when the move count has reached the value of four and thus four moves are completed: except that it never does, without stage 2, the relaxation of the minimisation heuristic.

Discovery through minimisation

According to the minimisation heuristic, people limit a problem representation to the minimum required to achieve satisfactory progress toward a goal. In the 9-dot problem, minimisation constrains the initial representation to the dot array presented in the initial problem (a grid of 3 x 3 dots). Relaxation of minimisation is triggered by criterion failure. Once relaxed, parsing the properties of previously explored moves, to identify invariants (cf. Kaplan & Simon, 1990) or unique move properties, derives new knowledge that can be used to infer possibilities for the discovery of new moves.

The model described in the first stage fails to find a criterion-satisfying fourth move, because the only places in the initial problem array correspond to dot coordinates. To 'learn' new places to look for moves, the minimisation heuristic needs to be relaxed. Relaxing the minimisation heuristic invokes a move parser that analyses the best moves produced to date and extracts properties that may enable discovery of new move types. In the four-line 9-dot problem, properties include space between known points, line lengths, and angles between lines. The model then uses these properties, in an order ranked according to principles of commonality, to discover new options to the current problem space based on inferences drawn from this knowledge (e.g., "if the most maximizing move currently has an average unit distance between cancelled dots of 1 unit, extend the line by 1 unit as a putative new move").

In this second stage, the model compares the properties of the lines drawn at the first stage. In the production (P COMPARE) it notices differences and invariants in terms of X and Y coordinates among the 'best moves' drawn. Based on these detected units of invariance among moves, the model, through the production (P EXTEND), uses the extracted units of invariance to extend the length of the first drawn line. In this way, the knowledge about properties extracted by comparing lines allows the problem space to be expanded to include (non-dot) spaces.

Phenomena captured by the model

Runs of the model provide ordinal differences between problem variants that are consistent with the published empirical literature on the problem. Like human solvers, it struggles to solve the problem (demonstrating impasse): in trials invoking 50 runs of the two-stage model, solution rates are less than 5%. It also returns, after attempts that extend beyond the 3x3 dot array, to exploring moves within the array (demonstrating fixation). However, it does solve on occasion (demonstrating the 'Aha' experience).

Also like human solvers, it easily solves (within 2 runs) the 13-dot variant in which the complete problem space is available in the initial representation, and finds solutions to 12- and 11-dot variants, where non-dot gaps within the dot array must be discovered, with increasing complexity but in runs < 10 (McGregor et al, 2001, Expt. 2). Finally, the implementation captures the difference between variants in which the first line is given, extending outside or within the initial dot array (McGregor et al, 2001, Expts. 4 and 5), with significantly fewer runs required for the latter than the former to discover solution, $p < .01$.

Discussion

The ACT-R implementation of CSP theory for the 9-dot problem demonstrates basic phenomena of insight captured by two simple heuristics governing search and expansion of an initial problem representation. Maximisation is a hill-climbing heuristic, while minimisation is a forcing function for discovering new problem knowledge based on recent discoveries of solution attempt properties. No additional knowledge is required, suggesting knowledge-rich accounts of insight (e.g., Knoblich et al., 1999; Kershaw & Ohlsson, 2004) may be overly elaborate for this particular problem.

Much remains to be done to provide a full implementation of knowledge-lean insight problem solving. Critically, the properties of the initial problem representation are hard-wired. Our hope is that the mechanism for minimisation can also be applied to parse the problem statement to build an initial representation. Building the ACT-R implementation raised new questions, such as whether maximization should be optimal (finding the very best move each run) or satisficing (finding the first criterion-satisficing move). These questions remain to be answered, but the growing ACT-R implementation provides a vehicle for doing so. In future work, we aim to extend the same principles to modeling other knowledge rich problems, such as the six-coin problem (Chronicle et al, 2004).

References

Chronicle, E. P., MacGregor, J.N., & Ormerod, T.C. (2004). What makes an insight problem? The roles of heuristics, goal conception, and solution recoding in knowledge-lean problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 14-27.

Hélie, S., & Sun, R. (2010). Incubation, insight, and creative problem solving: A unified theory and a connectionist model. *Psychological Review, 117*, 994-1024.

Kaplan, C. A., and Simon H.A. (1990). In search of insight. *Cognitive Psychology,* 22, 374-419.

Kershaw, T. C., & Ohlsson, S. (2004). Multiple causes of difficulty in insight: The case of the nine-dot problem. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 3-13.

Knoblich, G., Ohlsson, S., Haider, H., & Rhenius,D. (1999). Constraint relaxation and chunk decomposition in insight problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1534-1555.

MacGregor, J. N., Ormerod, T. C., & Chronicle, E. P. (2001). Information processing and insight: A process model of performance on the nine-dot and related problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 176-201.

Ormerod, T. C., MacGregor, J. N., Chronicle, E. P., Dewald, A. D., & Chu, Y. (2013). Act first, think later: The presence and absence of inferential planning in problem solving. *Memory & Cognition, 41,* 1096-1108.