Suboptimality and Error-Monitoring in Human Sequential Planning

Zahy Bnaya (zahy.bnaya@nyu.edu) Wei Ji Ma (weiji.ma@nyu.edu)
Center For Neural Science and Department of Psychology, New York University

Abstract

People are naturally good, yet suboptimal, at various sequential planning tasks. While human planning is still not well understood, computational approaches for suboptimal planning, such as *agent-centered* search, are extensively studied. In *agent-centered* search the agent searches for an incomplete strategy and re-plans when necessary. Here we present data collected from 11 subjects who played a natural planning game. Our contribution is twofold. First, we provide initial evidence that subjects follow incomplete strategies, as in *agent-centered* search. Subjects tend to perform a slow move, followed by a series of rapid moves, which indicates that on some steps subjects devise a plan and execute it on consecutive steps. Second, our results show evidence for a metacognitive process of error-monitoring. We show that subjects react slower when they are about to make mistakes and tend to forfeit a puzzle only when they drifted far away from the goal. In this work, we characterize human suboptimal behavior and also provide important evidence that people relay on partial strategies and error-monitoring in a sequential planning task, which is natural, a planning domain that is rarely studied.

Keywords: decision making; planning; heuristic search; error-monitoring

Introduction

The goal in sequential planning is to find a strategy that reaches a desired state, given a description of the initial state of the world and the set of possible actions. Previous studies show that humans are not optimal but perform well in various planning tasks (?, ?, ?, ?, ?). However, human planning is still not well understood. In Artificial Intelligence planning, on the other hand, the common approach, which is to perform *search* in the state space (?, ?), is extensively studied. Planning tasks can quickly become intractable due to combinatorial explosion, making *optimal* search methods, which guarantee the shortest strategy (?, ?, ?), less attractive than suboptimal methods. A common approach for suboptimal planning is *real-time* or *agent-centered* search (?, ?, ?) which finds only a limited or partial strategy and revise when necessary. In this paper we present preliminary results that characterize human behavior and study the relation between human behavior and *agent-centered* search. We found that although there is no explicit feedback in our task, our subjects had some indication of how well they are doing. This implicit awareness of mistakes, also known as *error-monitoring* (?, ?), might play a role in human planning.

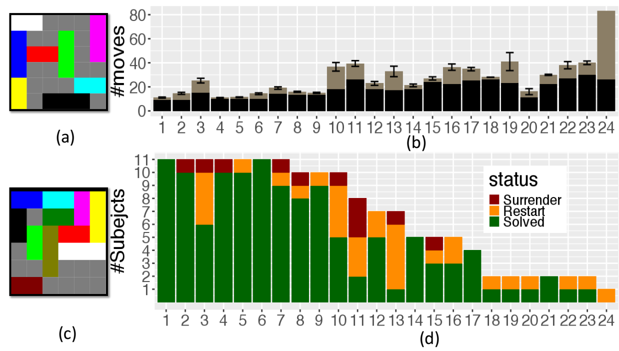


Figure 1: Two example puzzles: (a) puzzle-1 (easy) and (c) puzzle-16 (difficult); (b) mean number of moves to solve each puzzle (minimal solution length in black). (d) number of subjects who decided to surrender, restart (at least once), or solved the puzzle on the first trial

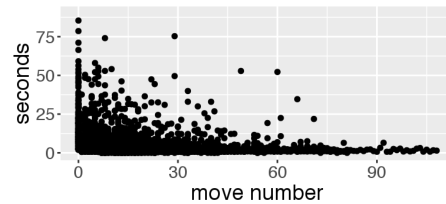


Figure 2: Response time in relation to move number

Methods

Eleven subjects solved up to 40 puzzles of *rush-hour*[1]. Rush-hour is a planning game, shown to be in PSPACE-complete (?, ?). The puzzle (Figure **??**) consists of a set of cars located on a 6 x 6 grid. Cars are oriented either horizontally or vertically and can move only in the direction of orientation. The goal is to find a sequence of moves, such that the red car moves to the right edge of the screen. Subjects used a mouse and a keyboard to move the cars around. We imposed a time limit of 30 minutes. The order of puzzles was chosen arbitrary, but with a tendency to increase in difficulty (Figure **??**(b)). At any point, subjects were allowed to *surrender* or *restart* the puzzle by pressing the relevant key on the keyboard. We recorded response times.

Results

Subjects completed on average 14 ± 0.9 puzzles (min=6, max=24; all errors reported are s.e.m across subjects) and performed $42\% \pm 2.4\%$ more moves than the minimal solution. Out of 231 solutions, 45 were optimal. Subjects tended to avoid the restart and surrender buttons. Five subjects surrendered, and only three surrendered more than once. Eight subjects chose to restart at least one puzzle (Figure **??**(c)).

[1] `www.thinkfun.com/play-online/rush-hour/`

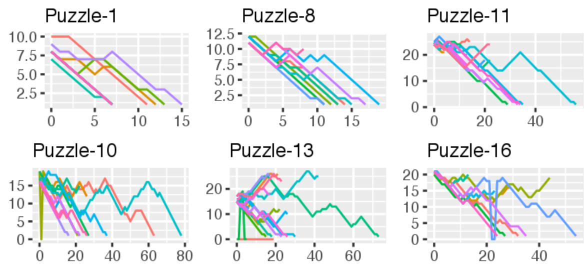
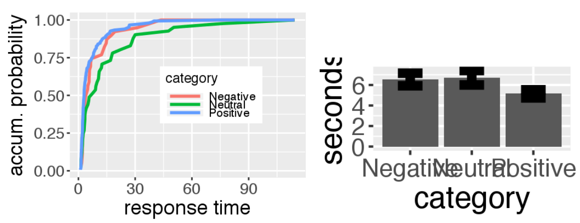


Figure 3: Plan execution progress in example puzzles. A completed plan reaches distance zero. Easier puzzles (top row) show less variability than more difficult puzzles (bottom row). Subjects are generally able to recover from errors and move quickly towards the goal.



(a) **(b)**

Figure 4: (a) CDF of one subject response time per category (b) Median response times for move category

Burst patterns Response times correlated with the step number in the plan (-0.21 ± 0.01, Spearman, Figure **??**(a)). We found that subjects spent considerable amount of time on first decisions (factor of 5.7 ± 0.18 more than the mean) and on other steps within a plan. We examined whether response times follow a pattern. As a preliminary test, we performed a median split on response times for each executed plan, splitting moves into *fast* and *slow* moves. We define a *burst* of length n to be a sequence of one slow move followed by n-1 fast moves. We found that the mean *burst length* in our results was $29\% \pm 0.5\%$ larger than if moves response times did not follow any particular pattern, suggesting that slow decisions tend to be followed by "bursts" of fast decisions. A possible interpretation is that subjects interleaved planning (on the slower steps) and executing (on faster consecutive steps), such as in *agent-centered* search. Bursts repeat within individual plan executions, which further suggests that subjects did not search for a complete plan.

Error-monitoring in planning We track the progress of plan executions by calculating the minimal distance to the goal after each step (Figure **??**). We found that subjects decided to surrender when they were 3.3 ± 0.90 steps farther from the goal than previously in the plan, and restart when they were 2.5 ± 0.31 steps farther. Remarkably, on both restarts and surrenders, subjects were actually closer to the goal than they were in the initial state. This suggests that although neither the path or the distance to the goal were known, subjects had a sense that they made a mistake. Next, we categorized moves into *positive* moves, if the move progressed closer to the goal, *negative* (moved away from the goal), or *neutral* (kept the same distance). The mean number of moves was 263 ± 24.6 (positive), 74.1 ± 8.02 (negative) and 63.8 ± 7.63 (neutral). We tested whether the distribution of response times was different between categories of moves (Figure **??**). We found that the median response time was different between positive and negative moves (Wilcoxon signed-rank test; p< 0.001) and between positive and neutral moves (p$= 0.004$), but not between negative and neutral moves (p$= 0.39$). This provides another indication of error-monitoring, as subjects spend more time before making a mistake.

Discussion and conclusion

We describe human behavior on a natural sequential planning task. We provide initial evidence that human behavior resembles *agent-centered* search, a powerful suboptimal method. Patterns in response times suggest that subjects repeat interleaved periods of fast and slow moves and that they do not search for a complete strategy. Additionally, we find indication of error-monitoring, which is valuable for suboptimal planning. Subjects spent more time before making incorrect decisions and decided to forfeit a task only when they were getting far away from the goal. The relation between suboptimal search, error-monitoring and human behavior provide a fresh look on human sequential planning. Moreover, we used a natural game task, on which especially not much is known.