Characterizing Human Planning on Large, Real-World Conceptual Networks

Denis C. L. Lan (denis.lan@psy.ox.ac.uk)
Department of Experimental Psychology, University of Oxford
Oxford, United Kingdom

Marcelo G. Mattar (marcelo.mattar@nyu.edu)
Department of Psychology, New York University
New York City, NY, United States

Abstract

Planning in the real world involves navigating vast spaces of possibilities, from finding a route through a city to searching for information online. Yet our understanding of human planning has largely come from studies involving small, simplified environments. To bridge this gap, we explored human planning in the context of the *Wiki Game*, where players start on a random Wikipedia article and are tasked with clicking on hyperlinks to reach a target article with minimal steps. We hypothesized that human planners reduce the computational cost of search by employing heuristic-guided and hierarchical search strategies. Analyzing a dataset of over 75,000 games, we discovered several behavioral signatures of heuristic-guided, hierarchical search. We formalized these insights using computational models, including tree search and hierarchical tree search algorithms. We found that our hierarchical tree search model mimicked these behavioral aspects of human navigation. Moreover, the patterns in human thinking times appeared to resemble patterns in the number of search iterations in the hierarchical tree search model. Collectively, these results suggest that humans use a combination of heuristic-guided search and hierarchical decomposition to efficiently plan in large, complex conceptual spaces.

Keywords: planning; navigation; tree search; hierarchical RL

Introduction

Traditional laboratory tasks have shed much light on the computational mechanisms underlying human planning (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Huys et al., 2012). This work suggests that, when making decisions, people mentally simulate the potential consequences of different action sequences. However, while traditional tasks involve small, simplified state spaces, realistic planning involves large, high-dimensional state spaces where it is difficult to learn the transition model and to simulate the full range of possibilities. To better understand human planning in large state spaces, cognitive scientists have recently turned towards games as a potential means for studying human planning in large state spaces that are more complex and ecological (van Opheusden et al., 2023; Allen et al., 2024). In addition to better capturing the complexity of real-life planning, gameplay datasets often include large numbers of players who voluntarily play the games, hence providing rich, large-scale datasets with which we can probe complex cognitive processes.

In both the machine learning and psychology literatures, numerous proposals have been made about how agents may simplify the search process in large, computationally intractable state spaces (Mattar & Lengyel, 2022). Firstly, agents may rely on heuristics (i.e., a rough estimate of cost or distance to goal) to guide and constrain their search, hence focusing computational resources on the most promising action sequences. For example, when planning on spatial networks, a suitable heuristic might be the Euclidean distance to the goal. An agent planning a train route from London to Athens might hence consider intermediate stations that bring them closer to the goal (e.g., Brussels), while neglecting stations that take them further away (e.g., Edinburgh). In AI, researchers have long relied on "informed" planning algorithms, such as best-first or A* search, that utilize heuristic functions to guide tree search (Dechter & Pearl, 1985; Hart, Nilsson, & Raphael, 1968). In psychology, evidence suggests that humans may plan in a way consistent with heuristic-guided search, considering promising-looking actions first (Newell & Simon, 1972) and pruning actions that look unpromising (Huys et al., 2012).

Additionally, agents may reduce the computational cost of search by hierarchically decomposing large state spaces into smaller sub-spaces. They may then form a high-level plan over these abstract sub-spaces before planning out the lower-level actions within each sub-space. For example, an agent planning a train route from London to Athens may first plan the sequence of countries they might need to pass through (e.g., Belgium, then Germany, then Austria, etc.), before planning the exact route they might take within each country. AI researchers indeed rely on hierarchical algorithms, especially in tasks with large, high-dimensional state spaces such as motor control (Kaelbling & Lozano-Pérez, 2011), and evidence suggests that humans also decompose complex state spaces into smaller sub-spaces for planning (Correa, Ho, Callaway, Daw, & Griffiths, 2023; Tomov, Yagati, Kumar, Yang, & Gershman, 2020; Balaguer, Spiers, Hassabis, & Summerfield, 2016).

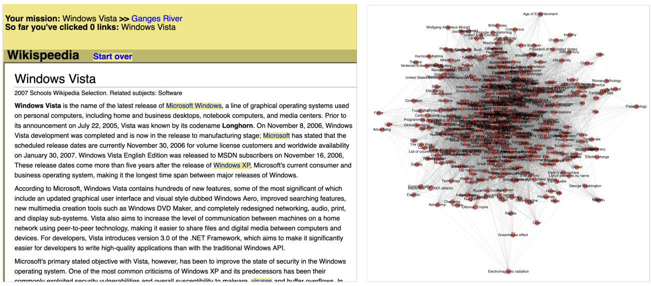
Task and Dataset

Here, we present our analyses of player choices on a popular online game, the *Wiki Game*. In this game, players start at a random Wikipedia article and are tasked with navigating to a random goal article by clicking on as few hyperlinks as possible (Figure 1). This presents us with a unique opportunity to study human planning in complex state spaces. Given the huge number of articles, players likely resort to strategies that help reduce the computational cost of search. More-

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

over, the hierarchical nature of semantic knowledge (Collins & Quillian, 1969), along with the presence of well-connected “hub” articles that might serve well as sub-goals, allows us to probe whether humans use hierarchical strategies to reduce computational costs. Importantly, unlike other laboratory experiments and games that often involve arbitrary actions and state spaces, players in the Wiki Game interact with real-life Wikipedia articles. The game hence emulates the real-world scenario of searching for information on the Internet, allowing us to study planning in a naturalistic setting.

Specifically, we analyze a public dataset of human players voluntarily playing a simplified version of the *Wiki Game* (West, Pineau, & Precup, 2009; West & Leskovec, 2012). The *Wikispeedia* dataset involves a reduced version of Wikipedia consisting of around 4,600 articles and 120,000 links between these articles. Players with about 7,000 unique IP addresses contributed more than 75,000 games in total, including around 50,000 games where players reached the goal article and around 25,000 where they terminated the game before the goal was reached.



Your mission: Windows Vista **>>** Ganges River
So far you've clicked 0 links: Windows Vista

Wikispeedia **Start over**

Windows Vista

2007 Schools Wikipedia Selection. Related subjects: Software

Windows Vista is the name of the latest release of Microsoft Windows, a line of graphical operating systems used on personal computers, including home and business desktops, notebook computers, and media centers. Prior to its announcement on July 22, 2005, Vista was known by its codename **Longhorn**. On November 8, 2006, Windows Vista development was completed and is now in the release to manufacturing stage; Microsoft has stated that the scheduled release dates are currently November 30, 2006 for volume license customers and worldwide availability on January 30, 2007. Windows Vista English Edition was released to MSDN subscribers on November 16, 2006, These release dates come more than five years after the release of Windows XP, Microsoft's current consumer and business operating system, making it the longest time span between major releases of Windows.

According to Microsoft, Windows Vista contains hundreds of new features, some of the most significant of which include an updated graphical user interface and visual style dubbed Windows Aero, improved searching features, new multimedia creation tools such as Windows DVD Maker, and completely redesigned networking, audio, print, and display sub-systems. Vista also aims to increase the level of communication between machines on a home network using peer-to-peer technology, making it easier to share files and digital media between computers and devices. For developers, Vista introduces version 3.0 of the .NET Framework, which aims to make it significantly easier for developers to write high-quality applications than with the traditional Windows API.

Microsoft's primary stated objective with Vista, however, has been to improve the state of security in the Windows operating system. One of the most common criticisms of Windows XP and its predecessors has been their commonly exploited security vulnerabilities and overall susceptibility to malware, viruses and buffer overflows. In

Figure 1: Left: screenshot of a game where the player is required to navigate from “Windows Vista” to “Ganges River”. Right: Graph structure of the top 500 most-visited articles.

Results

Characteristics of Human Paths

We first examined human decisions on one of the most common trial types in the dataset (“Asteroid” to “Viking”; Figure 2A). Human trajectories frequently involved waypoints that were not on the optimal paths. Players tended to choose non-optimal waypoints that represented general or geographical topics (e.g.,“Earth” or “Europe”), especially when they were semantically related to the target (e.g., “Norse mythology”). Conversely, optimal waypoints were less frequently picked when they represented specific concepts (e.g., “1 Ceres”), and/or were semantically distant from the target article (e.g., “Weather”). These qualitative observations suggest that players chose waypoints based on (1) how semantically similar they were to the goal and (2) how well-connected each article was. Indeed, looking across all games with an optimal path length of 3, players appeared to choose nodes that were more semantically close to the goal (Figure 2B) and more well-connected (Figure 2C) than the average optimal next node. Here, semantic distance was computed by taking the embeddings for each article's title from a pre-trained BERT language model (Reimers & Gurevych, 2019) and computing the cosine similarity between each article and the goal article.

Descriptive Choice Models

To quantify these observations, we fit descriptive choice models to participants' decisions in the game. In the model, choices were governed by three factors: (1) the shortest path length to the goal, (2) the semantic distance to the goal, and (3) the outdegree of the potential next article. The “utility” of each article was hence given by:

$$\begin{aligned} U = {} & \beta_0 * (1 - PathDistance) + \beta_1 * (1 - SemanticDistance) \\ & + \beta_2 * Outdegree \end{aligned} \tag{1}$$

Given these utilities, models then chose successive articles using a softmax choice rule. We also fit models that variously excluded path distance, semantic distance and/or outdegree as possible factors influencing player choices. We fit model parameters (β_0, β_1, and/or β_2) for each participant on half of the participant's choices and tested model fits on the other half. If players' choices were solely governed by the optimal path lengths, cross-validated log likelihoods should be best for the models without weights for semantic distance or outdegree. However, Bayesian Model Comparison using random effects analysis (with VBAToolbox (Daunizeau, Adam, & Rigoux, 2014)) suggested that the full model was the most common model across participants (exceedance probability > 0.999, estimated model frequency = 77.62%; Figure 2D), confirming our hypothesis that participant trajectories were biased towards articles that were well-connected and semantically similar to the goal.

Temporal Dynamics

However, outdegree and semantic distance did not influence player choices uniformly throughout their trajectories. In Figures 2B and 2C, it is apparent that while outdegree spiked at the start of players' trajectories and subsequently declined, semantic distance decreased most rapidly as players approached the goal. This suggests that outdegree influenced player choice more in the opening, while semantic distance influenced player choice more near the end. Indeed, when we fit the full choice model separately on player choices in the early vs the late game (i.e., before or after the midpoint in each trajectory), we found that the fitted choice weights for outdegree were higher in the early game ($t(963) = 19.30$, $p < .0001$), while the fitted choice weights for semantic distance were higher in the late game ($t(963) = -16.22$, $p < .0001$; Figure 1E). This temporal pattern suggests that participants used a strategy where they first headed towards a well-connected sub-goal before using semantic distance to head towards the goal.

Within-Category Bias

One hierarchical strategy that humans might use is to decompose large graphs into sub-graphs and form lower-level plans

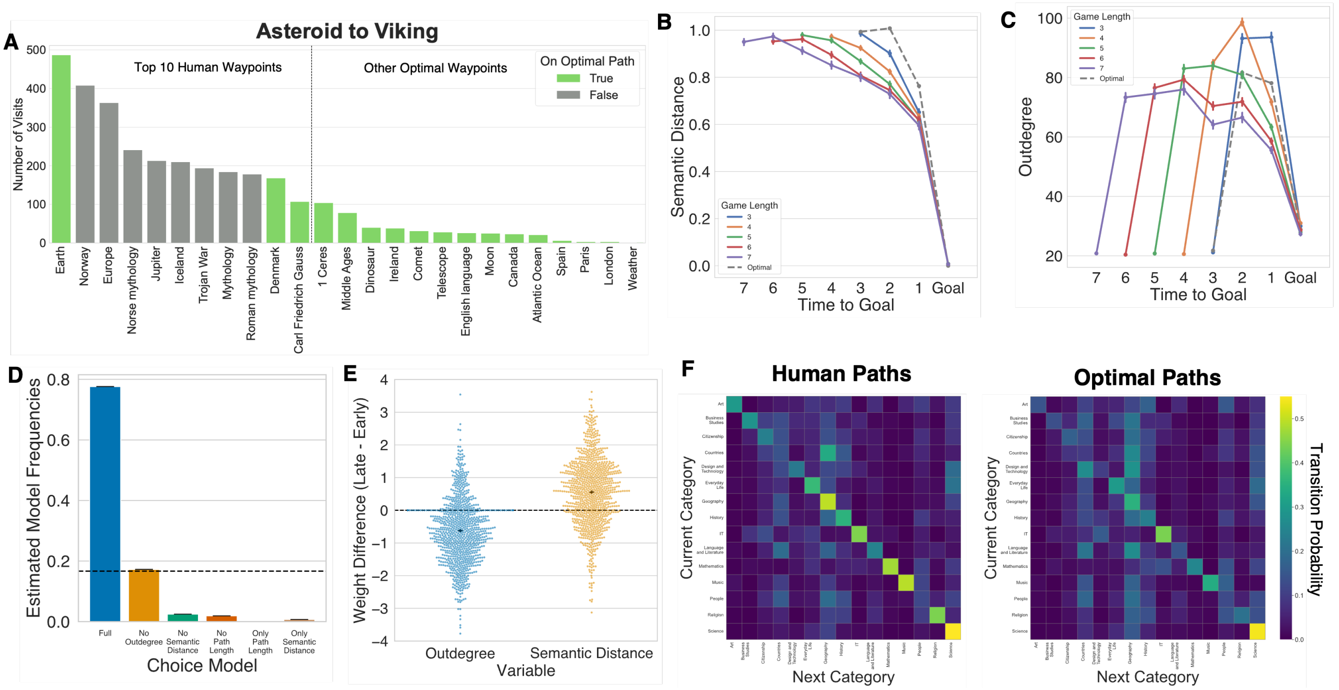


Figure 2: A: Histogram of human-selected waypoints for traveling from Asteroid to Viking. Hue indicates whether the waypoint lies on an optimal path between Asteroid and Viking. B: Temporal evolution of semantic distance to goal in human trajectories, time-locked to reaching the goal. Each hue represents human trajectories of a different length. Gray dashed line represents optimal paths. C: Temporal evolution of outdegree in human trajectories, time-locked to reaching the goal, replicated from West and Leskovec (2012). D: Estimated model frequencies for model comparison of choice models. E: Difference in choice weights for outdegree and semantic distance from early to late gate. F: Transition probabilities between high-level categories for humans and optimal agents. All error bars represent ±1 SE.

within each sub-graph. Planning within each sub-graph reduces computational cost by constraining the search space (e.g., searching through only articles about Science when the goal is "Radium"), but comes at the cost of neglecting potentially shorter routes that require traversals through multiple sub-graphs (e.g., going through "Poland" and "Marie Curie" to get to "Radium"). To examine whether humans used such strategies, we constructed transition probability matrices between high-level categories for both humans and optimal agents (Figure 2F). By focusing on the diagonals of these transition matrices, we observed that human paths had a higher probability of within-category transitions than optimal paths. To control for the tendency of articles within each category to link more frequently to others in the same category, we ran linear regression models using random walk transition probabilities and a categorical "within-category" variable as predictors of human or model transition probabilities. This revealed that humans had significantly higher within-category transitions even after controlling for random-walk transition probabilities ($\beta = 0.13$, $p < .0001$), while optimal paths actually had significantly less within-category transitions ($\beta = -0.02$, $p = .029$). This is consistent with the idea that players constrained their search to smaller sub-graphs at expense of neglecting routes that traversed multiple sub-graphs.

Computational Models

To formalize our hypotheses about humans strategies, we developed several computational models and simulated their performance on the game. We hypothesized that human behavior could be explained by (1) the use of heuristics such as semantic distance and outdegree to guide search and (2) hierarchical planning by decomposing the entire graph into sub-graphs.

The first model we simulated was an A*-like search algorithm with a heuristic function based on a combination of semantic distance and outdegree. To model humans' limited computational resources, we augmented the algorithm with an early stopping (stopping search early on each iteration with some probability) and a pruning mechanism (considering only the top nodes based on the heuristic function). We simulated human's imperfect knowledge of the graph by removing 10% of the connections and re-attaching these connections between random nodes. Instead of formulating and following a plan from the start, agents re-planned on every node and moved to the next node in the path it found (or, if a path was not found, the next node in the best sequence of nodes as implied by estimated total distances to goal f). To determine whether heuristics were necessary for capturing human behavioral patterns, we also simulated a version of the model with a flat heuristic function, essentially amounting to a breadth-first search. Conversely, to determine whether tree search was a necessary component, we also simulated a version which simply picked the connected article with the highest heuristic value.

We then designed a hierarchical version of A* which decomposes the Wikipedia graph into sub-graphs, formulates a high-level plan over these sub-graphs, and then successively forms plans within each sub-graph (Figure 3, Algorithm 1):

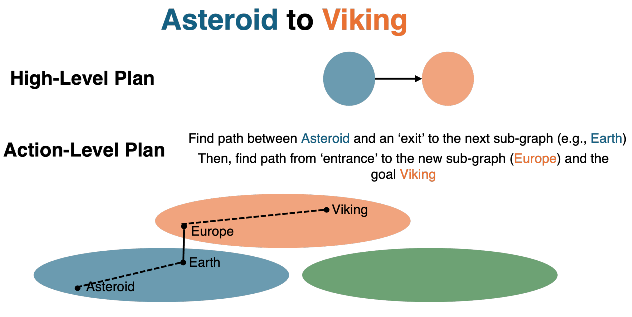


Figure 3: Schematic diagram of Hierarchical A*

Algorithm 1 Hierarchical A*

```
$path' \leftarrow A^*(G_s, G_g, H)$
$G_{\text{current}} \leftarrow G_s$
$G_{\text{next}} \leftarrow path'[1]$
while $G_{\text{current}} \neq G_g$ do
    $g' \leftarrow \{h \mid h \in G_{\text{current}}$ and $\exists h' \in G_{\text{target}} : h$ is connected to $h'\}$
    $path \leftarrow A^*(s, g', G_{\text{current}})$
    $s \leftarrow path[1]$
    if $s \in g'$ then
        $G_{\text{current}} \leftarrow G_{\text{next}}$
        $path' \leftarrow path'[1:]$
        $G_{\text{next}} \leftarrow path'[1]$
    end if
end while
while $s \neq g$ do
    $path \leftarrow A^*(s, g, G_g)$
    $s \leftarrow path[1]$
    $path \leftarrow path[1:]$
end while
```

Specifically, we initialized the agent by decomposing the agent's knowledge network into sub-graphs G using the Louvain community detection method. This is an algorithm for identifying clusters of nodes that are more tightly connected to each other than to the rest of the network. We defined the local 'hubs' h within each sub-graph by identifying the top-50 outdegree nodes within each sub-graph. We then defined the high-level graph H between sub-graphs by connecting sub-graphs whenever the local hub in one sub-graph was connected by a link to a local hub in the other. Hierarchical planning proceeded as follows: the agent first formed a high-level plan over sub-graphs by finding the shortest path between the starting sub-graph and the goal sub-graph. Then, within each successive sub-graph on the high-level plan, the agent planned from the current node to one of the local hubs in the current sub-graph that was connected to a hub in the next target sub-graph. Finally, when the agent arrived at the sub-graph containing the goal, it searched for a path from the current node to the goal node. In other words, the model formulated a high-level plan over broad clusters (using A*) before formulating a lower-level plan for navigating between clusters (again, using A*). Hence, if all nodes belonged to the same cluster (or conversely, if each node belonged to its own cluster), hierarchical A* reduces back to A*.

Model Performance

We tested 30 iterations of each model on 100 games of optimal path length 3. The hierarchical A* model performed the best out of the four models, reaching the goal within 10 steps at a higher rate than the A*, breadth-first, and heuristic-only models (Figure 4A). The hierarchical model also performed less search iterations per move then the other three models (Figure 4B), suggesting that it was planning more efficiently. Interestingly, when we simulated model performance on games (i.e., start-goal pairs) that humans had also played, there was a strong relationship between human performance and performance for the hierarchical model (Figure 4C). When included as predictors in a linear regression model, hierarchical A* performance ($\beta = 1.96, p < .0001$), but not the performance of the three other models, positively predicted human performance. This was despite all goals being 3 steps away, suggesting that the hierarchical model was successfully capturing how difficult each game was for humans.

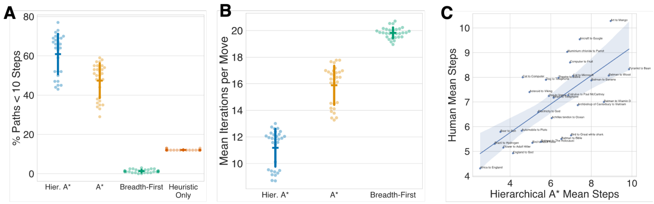


Figure 4: A: Proportion of paths completed under 10 steps for the four models. B: Mean search iterations per move for all models (except the heuristic-only model). C: Relationship between human and hierarchical A* performance on games with different starts and goals.

Descriptive Choice Models

We then fit the descriptive choice models to the simulated trajectories from the four models. As before, we fit the models on half of the choices and evaluated the cross-validated log likelihood on the other half. Of the four models, the full model (with weights for path length, semantic distance, and outdegree) fit best for the A* (exceedance probability $> .999$, estimated frequency = 97.54%) and hierarchical A* models (exceedance probability $> .999$, estimated frequency = 97.54%). However, trajectories from the breadth-first search model were best fit by the model with only a parameter for path length (exceedance probability = .85, estimated frequency = 57.63%), suggesting that heuristic-guided search was necessary for producing biases towards semantic distance and outdegree. On the other hand, trajectories from the pure heuristic model were fit roughly equally well by the full model (exceedance probability = .5072, estimated frequency = 49.07%) and the model without path length as

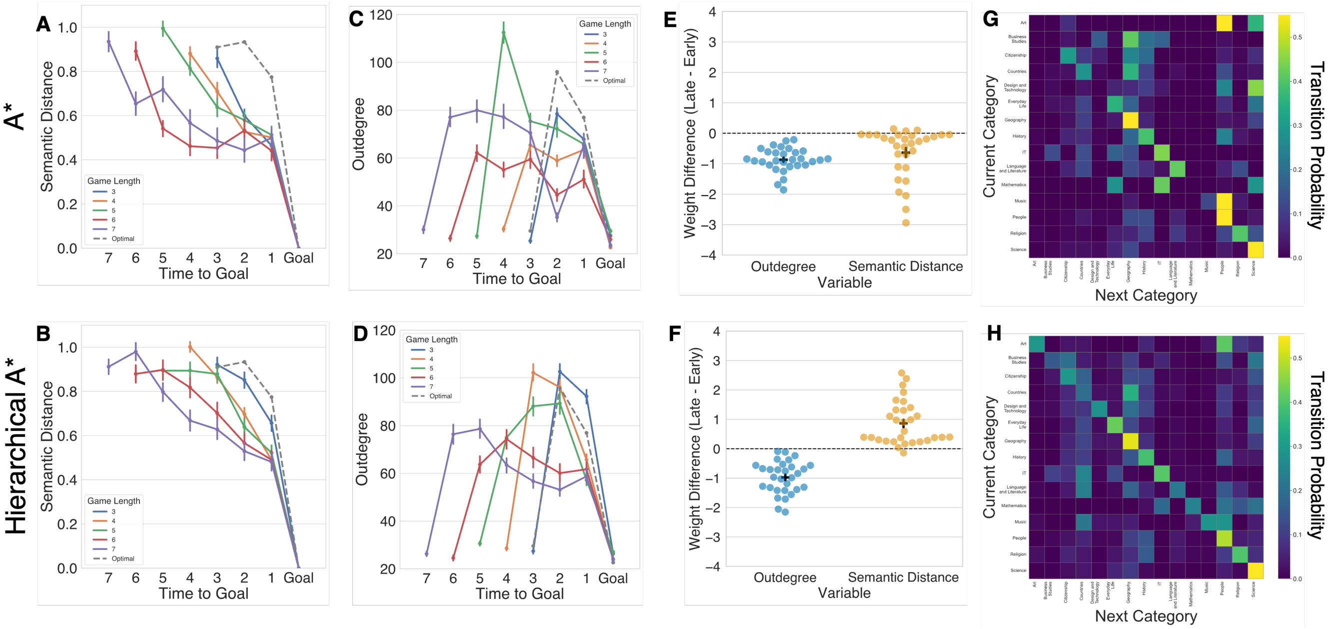


Figure 5: A/B: Temporal evolution of semantic distance to goal, time-locked to reaching the goal for the A* and hierarchical A* models. Hue represents trajectories of different lengths. Gray dashed lines represent optimal paths. C/D: Same as A/B but for outdegree. E/F: Difference in choice weights for outdegree and semantic distance from early to late gate for the A* and hierarchical A* models. G/H: Transition probabilities between high-level categories for A* and hierarchical A* models. All error bars represent ±1 SE.

a parameter (exceedance probability = .4928, estimated frequency = 48.78%). The fitted parameters for path length for the heuristic-only model were also close to zero in the full model ($M = 1.37 * 10^{-7}$, $SD = 6.83 * 10^{-7}$). This suggests that tree search might be a necessary component for capturing human abilities to find accurate routes above and beyond simply using heuristics.

Temporal Dynamics

While descriptive choice model fits suggested that both the A* and hierarchical A* models captured humans' bias towards articles that were semantically similar to the goal or high in outdegree, closer examination suggests different temporal dynamics in how semantic distance and outdegree shaped model trajectories. Firstly, semantic distance in hierarchical A* model trajectories appeared to decrease more gradually at the start and more rapidly at the tail end of trajectories (Figure 5A), just like in humans, while semantic distance in A* model trajectories appeared to decrease rapidly at the start and slowly near the end (Figure 5B). Moreover, only the hierarchical A* model, but not the A* model, consistently mimicked human tendencies to choose waypoints that were of higher outdegree than optimal waypoints at the start of their trajectories (Figures 5C and D). Indeed, when choice weights were fit separately for the early and late game, we observed human-like increases in semantic distance weights from early to late game only in the hierarchical model ($t(29) = -6.38$, $p < 0.0001$), while the non-hierarchical model actually saw a decrease ($t(29) = 4.20$, $p = 0.0002$). Overall, only the hierarchical model captured the temporal pattern of human choices, where outdegree was more influential in the early game but semantic distance was more influential in the late game.

Within-Category Bias

We examined category transition probability matrices of the two A* models (Figures 5E and F). We found that the hierarchical model had significantly higher within-category transition probabilities after controlling for random-walk transition probabilities ($\beta = 0.14$, $p < .0001$), but the non-hierarchical model did not ($\beta = 0.06$, $p = 0.10$). In a regression model predicting human transition probabilities from that of random, hierarchical, non-hierarchical, and optimal agents, the hierarchical transition probabilities positively predicted human transition probabilities ($\beta = 0.33$, $p < .0001$), but the non-hierarchical probabilities actually negatively predicted human transition probabilities ($\beta = -0.08$, $p = .005$). This suggests that only the hierarchical model captured human tendencies to constrain search within a subset of the graph.

Thinking Times

Lastly, we probed players' thinking times for further insights into the cognitive mechanisms underlying planning. For each choice, we computed outdegree and semantic distance scores based on the rank of the chosen article within all possible connected articles (i.e., 1 indicates the top outdegree and 0 indicated the lowest outdegree out of all possible connected articles). Since the dataset only included completion times for the whole trajectory, we looked at whether the mean time taken per step varied with the mean outdegree and semantic distance scores for the whole trajectory. If players used outdegree and semantic distance to guide their search, we would expect thinking time to be shorter on trajectories where they mainly traversed through articles that were high in outdegree and semantic similarity to the goal, but longer when they ultimately chose lower-outdegree or semantically distant articles, after having already considered the high-outdegree and semantically similar options. This prediction turned out to be true (Figure 6A), and a mixed effect model revealed that thinking times were significantly predicted by outdegree, semantic distance, and their interaction (outdegree: $\beta = -0.02$, $p < .0001$; distance: $\beta = 0.10$, $p < .0001$; interaction: $\beta = 0.02$, $p < .0001$). When we looked at the number of search iterations for the computational models, we observed a similar relationship in the hierarchical model (Figure 6B; outdegree: $\beta = -0.26$, $p < .0001$; distance: $\beta = 0.18$, $p < .0001$; interaction: $\beta = 0.06$, $p = .006$), but not the non-hierarchical model, where there was only an effect of semantic distance (Figure 6C). These results suggest that hierarchical models mimicked patterns of how long humans planned before they decide on different choices. Notably, the number of search iterations in the breadth-first search model were correlated with neither semantic distance nor outdegree (Figure 6D), providing further support for the suggestion that heuristic use is important for capturing patterns of human behavior.

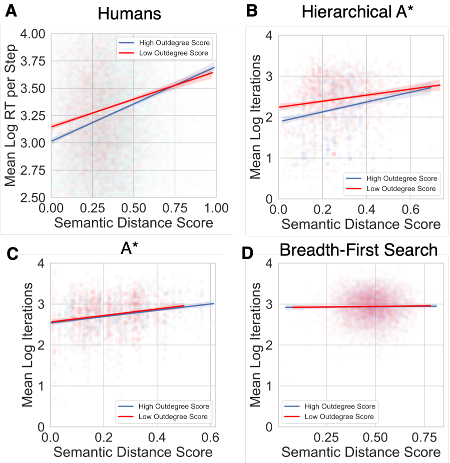


Figure 6: Mean reaction times (A) or search iterations per step (B-D) as a function of mean semantic distance score (x-axis) and outdegree score (median split, hue) across each trajectory, for humans and models

Discussion

While cognitive science has recently turned towards large-scale gameplay datasets to probe the cognitive mechanisms underlying human planning, these games often involve arbitrary rules and state spaces that bear little resemblance with real-life planning. Here, we analyzed a dataset where players interacted with real-world Wikipedia articles, emulating the naturalistic scenario of information search on vast online information networks. Our results indicate that humans (1) have a preference for waypoints that are high in outdegree and semantically close to the goal, (2) are more strongly influenced by outdegree in the early game and by semantic distance in the late game, (3) 'think' longer before making low-outdegree and semantically distant choices, and (4) are more likely to transition to articles within the same high-level category. These behaviors are emulated by our hierarchical A* model, suggesting that humans were relying on heuristic-guided, hierarchical planning strategies.

While we consider our reliance on a large, naturalistic dataset a strength, the lack of tight experimental controls necessarily introduces some ambiguity in interpreting the results. For example, articles of high outdegree are also likely more well-known and familiar to human players. This prevents us from ruling out the possibility that participants' choices are driven not solely by hierarchical strategies, but also by familiarity or prior knowledge of certain topics. Nevertheless, a uniform bias towards familiar articles likely cannot account for the fact that the bias towards outdegree is stronger earlier on in the participants' navigation trajectories. In future work, we intend to validate our findings in more experimentally controlled settings. Moreover, our analyses of thinking times were limited by the lack of fine-grained reaction times, preventing us from analyzing how long players contemplated for each individual decision. Reaction times have increasingly been used as an additional behavioral index of human planning (Fernandez Velasco et al., 2025; Jensen, Hennequin, & Mattar, 2024), and collecting fine-grained reaction times in a future dataset would help us further validate our theories of human planning. For example, if participants were planning hierarchically, one prediction would be that thinking times would be longer at subgoals, when participants might be formulating lower-level plans towards the next subgoal.

The mechanisms we illuminate likely apply not just to planning on conceptual networks, but also more broadly to planning on all vast, high-dimensional state spaces. Indeed, our findings are reminiscent of those from studies in different domains, such as spatial navigation. Specifically, just like how human paths in conceptual networks are biased by conceptual distance, human spatial navigation is strongly shaped by Euclidean geometry (Garvert, Saanum, Schulz, Schuck, & Doeller, 2023; Bongiorno et al., 2021; Lan, Hunt, & Summerfield, 2024), consistent with the use of Euclidean distance as a heuristic. Moreover, just like how humans appear to use hierarchical strategies for conceptual planning, recent findings also suggest that humans use hierarchical strategies to reduce the computational cost of planning in real-world spatial navigation (Fernandez Velasco et al., 2025). These parallels align with previous suggestions of a common mechanism underlying both spatial and conceptual maps (Epstein, Patai, Julian, & Spiers, 2017; Constantinescu, O'Reilly, & Behrens, 2016), providing further support for domain-general mechanisms in planning and navigation.

References

Allen, K., Brändle, F., Botvinick, M., Fan, J. E., Gershman, S. J., Gopnik, A., … Schulz, E. (2024, June). Using games to understand the mind. *Nature Human Behaviour*, *8*(6), 1035–1043. Retrieved 2025-01-06, from https://www.nature.com/articles/s41562-024-01878-9 (Publisher: Nature Publishing Group) doi: 10.1038/s41562-024-01878-9

Balaguer, J., Spiers, H., Hassabis, D., & Summerfield, C. (2016, May). Neural Mechanisms of Hierarchical Planning in a Virtual Subway Network. *Neuron*, *90*(4), 893–903. Retrieved 2025-01-06, from https://www.cell.com/neuron/abstract/S0896-6273(16)30057-5 (Publisher: Elsevier) doi: 10.1016/j.neuron.2016.03.037

Bongiorno, C., Zhou, Y., Kryven, M., Theurel, D., Rizzo, A., Santi, P., … Ratti, C. (2021, October). Vector-based pedestrian navigation in cities. *Nature Computational Science*, *1*(10), 678–685. Retrieved 2022-11-23, from https://www.nature.com/articles/s43588-021-00130-y (Number: 10 Publisher: Nature Publishing Group) doi: 10.1038/s43588-021-00130-y

Collins, A. M., & Quillian, M. R. (1969, April). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 240–247. Retrieved 2025-01-12, from https://www.sciencedirect.com/science/article/pii/S0022537169800691 doi: 10.1016/S0022-5371(69)80069-1

Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. J. (2016, June). Organizing conceptual knowledge in humans with a gridlike code. *Science*, *352*(6292), 1464–1468. Retrieved 2022-11-25, from https://www.science.org/doi/abs/10.1126/science.aaf0941 (Publisher: American Association for the Advancement of Science) doi: 10.1126/science.aaf0941

Correa, C. G., Ho, M. K., Callaway, F., Daw, N. D., & Griffiths, T. L. (2023, June). Humans decompose tasks by trading off utility and computational cost. *PLOS Computational Biology*, *19*(6), e1011087. Retrieved 2025-01-06, from https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1011087 (Publisher: Public Library of Science) doi: 10.1371/journal.pcbi.1011087

Daunizeau, J., Adam, V., & Rigoux, L. (2014, January). VBA: a probabilistic treatment of nonlinear models for neurobiological and behavioural data. *PLoS computational biology*, *10*(1), e1003441. doi: 10.1371/journal.pcbi.1003441

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011, March). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, *69*(6), 1204–1215. Retrieved 2024-05-03, from https://www.cell.com/neuron/abstract/S0896-6273(11)00125-5 (Publisher: Elsevier) doi: 10.1016/j.neuron.2011.02.027

Dechter, R., & Pearl, J. (1985, July). Generalized best-first search strategies and the optimality of A*. *J. ACM*, *32*(3), 505–536. Retrieved 2025-01-06, from https://dl.acm.org/doi/10.1145/3828.3830 doi: 10.1145/3828.3830

Epstein, R. A., Patai, E. Z., Julian, J. B., & Spiers, H. J. (2017, November). The cognitive map in humans: spatial navigation and beyond. *Nature Neuroscience*, *20*(11), 1504–1513. Retrieved 2024-01-30, from https://www.nature.com/articles/nn.4656 (Number: 11 Publisher: Nature Publishing Group) doi: 10.1038/nn.4656

Fernandez Velasco, P., Griesbauer, E.-M., Brunec, I. K., Morley, J., Manley, E., McNamee, D. C., & Spiers, H. J. (2025, January). Expert navigators deploy rational complexity–based decision precaching for large-scale real-world planning. *Proceedings of the National Academy of Sciences*, *122*(4), e2407814122. Retrieved 2025-01-27, from https://www.pnas.org/doi/10.1073/pnas.2407814122 (Publisher: Proceedings of the National Academy of Sciences) doi: 10.1073/pnas.2407814122

Garvert, M. M., Saanum, T., Schulz, E., Schuck, N. W., & Doeller, C. F. (2023, April). Hippocampal spatio-predictive cognitive maps adaptively guide reward generalization. *Nature Neuroscience*, *26*(4), 615–626. Retrieved 2024-09-03, from https://www.nature.com/articles/s41593-023-01283-x (Publisher: Nature Publishing Group) doi: 10.1038/s41593-023-01283-x

Hart, P. E., Nilsson, N. J., & Raphael, B. (1968, July). A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Transactions on Systems Science and Cybernetics*, *4*(2), 100–107. Retrieved 2025-01-06, from https://ieeexplore.ieee.org/document/4082128 (Conference Name: IEEE Transactions on Systems Science and Cybernetics) doi: 10.1109/TSSC.1968.300136

Huys, Q. J. M., Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012, March). Bonsai Trees in Your Head: How the Pavlovian System Sculpts Goal-Directed Choices by Pruning Decision Trees. *PLOS Computational Biology*, *8*(3), e1002410. Retrieved 2025-01-06, from https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1002410 (Publisher: Public Library of Science) doi: 10.1371/journal.pcbi.1002410

Jensen, K. T., Hennequin, G., & Mattar, M. G. (2024, July). A recurrent network model of planning explains hippocampal replay and human behavior. *Nature Neuroscience*, *27*(7), 1340–1348. Retrieved 2025-01-27, from https://www.nature.com/articles/s41593-024-01675-7 (Publisher: Nature Publishing Group) doi: 10.1038/s41593-024-01675-7

Kaelbling, L. P., & Lozano-Pérez, T. (2011, May). Hierarchical task and motion planning in the now. In *2011 IEEE International Conference on Robotics and Automation* (pp. 1470–1477). Retrieved 2025-01-06, from https://ieeexplore.ieee.org/document/5980391 (ISSN: 1050-4729) doi: 10.1109/ICRA.2011.5980391

Lan, D. C. L., Hunt, L., & Summerfield, C. (2024, October). *Goal-directed navigation strategies in humans and deep meta-learning agents.* OSF. Retrieved 2025-01-27, from https://osf.io/6x4r2 doi: 10.31234/osf.io/6x4r2

Mattar, M. G., & Lengyel, M. (2022, March). Planning in the brain. *Neuron*, *110*(6), 914–934. Retrieved 2025-01-23, from https://www.sciencedirect.com/science/article/pii/S0896627321010357 doi: 10.1016/j.neuron.2021.12.018

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Oxford, England: Prentice-Hall. (Pages: xiv, 920)

Reimers, N., & Gurevych, I. (2019, August). *Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks.* arXiv. Retrieved 2025-01-13, from http://arxiv.org/abs/1908.10084 (arXiv:1908.10084 [cs]) doi: 10.48550/arXiv.1908.10084

Tomov, M. S., Yagati, S., Kumar, A., Yang, W., & Gershman, S. J. (2020, April). Discovery of hierarchical representations for efficient planning. *PLOS Computational Biology*, *16*(4), e1007594. Retrieved 2025-01-06, from https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1007594 (Publisher: Public Library of Science) doi: 10.1371/journal.pcbi.1007594

van Opheusden, B., Kuperwajs, I., Galbiati, G., Bnaya, Z., Li, Y., & Ma, W. J. (2023, June). Expertise increases planning depth in human gameplay. *Nature*, *618*(7967), 1000–1005. Retrieved 2025-01-06, from https://www.nature.com/articles/s41586-023-06124-2 (Publisher: Nature Publishing Group) doi: 10.1038/s41586-023-06124-2

West, R., & Leskovec, J. (2012, April). Human wayfinding in information networks. In *Proceedings of the 21st international conference on World Wide Web* (pp. 619–628). New York, NY, USA: Association for Computing Machinery. Retrieved 2025-01-06, from https://doi.org/10.1145/2187836.2187920 doi: 10.1145/2187836.2187920

West, R., Pineau, J., & Precup, D. (2009, July). Wikispeedia: an online game for inferring semantic distances between concepts. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence* (pp. 1598–1603). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.