What To Do Next? A Comparative Study of Human and Rational Decision-Making*

Bram Swaanen and Tomas Klos

Utrecht University, Heidelberglaan 8, 3584 CS Utrecht, The Netherlands
{b.j.g.swaanen,t.b.klos}@uu.nl

Abstract. This study aims to investigate how human decision-making compares to fully rational decision-making in concrete situations. In the second half of the twentieth century it became more and more apparent that human decision-making is not entirely rational [4]. On the other hand, this time period also witnessed the development of computational systems able to reason rationally in limited domains.
In our study, human participants were asked to play a reasoning game and their reasoning behaviour was recorded and compared to that of a fully rational, computational agent. In 84% of cases, participants' decision-making corresponded to the agent's. The other 16% of decisions constituted deviations due to the use of (sub-optimal) heuristics or a failure to observe all logical consequences of the game's rules. We interpret this as confirmation that humans are boundedly rational, as proposed by Simon in [16]. They generally act rationally, but at times fail to process all necessary information for a rational decision, due to computational constraints. By leaving the experimental setting open as much as possible, we were able to observe which specific heuristics emerged, and propose explanations relating them to the characteristics of the task environment.

Keywords: Knowledge representation and Reasoning · Logics and Normative systems · AI and Social sciences · Bounded rationality · Wumpus World.

1 Introduction

As human beings, we are curious to know how we make decisions. This is not surprising; on many occasions it can be useful to predict what decision another human being is going to make. It is therefore not a coincidence that questions concerning the structure of human decision-making play a role in almost every field of social science, whether it be economics, cognitive science or psychology.

Traditionally, human beings are assumed to make those decisions that maximise their personal satisfaction given all information available to them. We call this idea rationality [10]. For example, when a person finds two *identical* dresses d_1 and d_2 on different websites where d_1 costs €40,- and d_2 costs €41,- then, all other things being equal, they will prefer dress d_1. Research has shown that this economic view of human beings being rational is too simplistic. Humans commonly base their decisions on other factors than what might be best given some optimisation criterion. In the example, the person might consider their previous experiences with the two websites. A person's environment, emotions and cognitive limitations are all sources of other than strictly rational behaviour [10]. Another cause of other than rational behaviour is our inability to evoke all the knowledge that is relevant for a task [17], preventing us from deriving the full implications of our knowledge.

Part of the traditional view of human beings' decision-making is that we derive information deductively, using general rules of inference. Already since the ancient Greeks, when Aristotle published his *Prior analytics*, philosophers and mathematicians have tried to define formal concepts that capture the essence of valid reasoning, i.e. rules of what is necessarily true given some knowledge. An example is the rule of *modus ponens*, which states that if it is true that *if A then B* and it is true that A, then it is true that B, where A and B are sentences that can be true or

* We would like to sincerely thank our colleagues that took the time to review this paper and provided many valuable suggestions for improvement.

false. For example, if it is true that "if you are a person, then you have a brain" and it is true that "you are a person" then it is true that "you have a brain."

As we sometimes seem to reason in this way, it might be tempting to assume we always do. However, there is a large body of psychological research indicating that humans often fail to select the right logical rule for reasoning tasks [3]. This suggests that human reasoning is at best only partially deductively valid.

The upsurge of artificial intelligence in the twentieth century has provided us with new tools to study how human decision-making compares to purely rational strategies. The computer systems that have since been developed can solve many logical problems flawlessly, by always evoking the right logical rules and applying them correctly. These systems make use of algorithms that can prove all and only those sentences that are necessarily true given some knowledge. The proof that these algorithms work was largely provided by logicians around the 1960s (see [14]). On top of that, systems have been developed that can weigh actions perfectly, given a clear objective and score function [14]. The result are systems that can perform rational decision-making.

In this study we hope to use some of these tools to shed light on how human decision-making compares to a fully rational approach. We will compare the performance of human beings in a game that requires logical reasoning to that of a computational logic-based agent.

The following sections provide a detailed description of the conducted research. In Section 2 we will provide the necessary preliminaries on human and rational decision-making. Then, in Section 3 the research will be presented, including the research question and hypothesis. Subsequently, the experimental setup is discussed in Section 4, followed by an analysis of the results in Section 5. Finally, in Section 6, we discuss the study's results and give suggestions for further research.

2 Preliminaries

This section provides preliminaries in human rational decision-making. Some preliminary knowledge of these fields is necessary to formulate our research question and hypothesis.

2.1 Human Decision-Making

In economics, rational action is understood as maximising one's satisfaction given one's (1) budgetary constraints and (2) preferences [10]. This would mean that in a fully observable environment, human reasoners would take all relevant information into account when deciding on an action. In this paper we will use the term *all-information-rational* to refer to this type of rationality. However, there is extensive literature on human decision-making strategies that are not in accord with this model [3]. In some fully observable environments, humans seem to reason employing strategies that ignore "part of the information, with the goal of making decisions more quickly, frugally, and/or accurately than more complex methods" and are called heuristics [6, p. 454]. Two answers have been proposed to the question why people employ heuristics. The first theory states that we have limited memory and processing resources [8], which prevents us from considering every situation and action fully. This makes it rational to reduce the information used in a decision problem by ignoring irrelevant information. The idea of "bounded rationality" proposed by Herbert Simon attempts to conceptualise this point. It is the idea "that the choices people make are determined (...) by the knowledge that decision makers do and don't have of the world, their ability or inability to evoke that knowledge when it is relevant, to work out the consequences of their actions, to conjure up possible courses of action, to cope with uncertainty (...) and to adjudicate among their many competing wants" [17, p. 25]. Note that the first five of these limitations prevent humans from fully analysing which decisions maximise their satisfaction and the last point limits the idea that humans have clear and unambiguous preferences. This results in sub-optimal behaviour in the mathematical sense.

An illustration of bounded rationality is provide by the Wason Selection Task (WST) experiment [18], illustrated in Figure 1. This experiment has shown that human beings often fail to grasp the full meaning of the sentence "if A then B", a sentence which is crucial in the idea of modus ponens (Section 1). The fact that participants in the WST recognise they have made a mistake when it is pointed out to them [11], suggests that participants fail to evoke the appropriate logical rule for the task, even though they do know it. This results in a wrong judgement of the consequences of their actions.

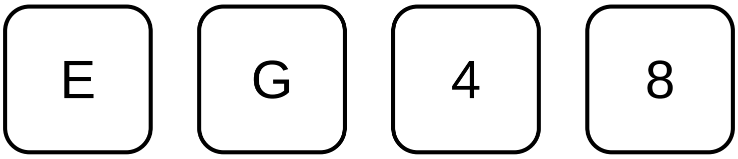


Fig. 1: The Wason Selection Task (WST). Each card has a letter on one side and a number on the other. Cards which have an E on one side must have a 4 on the other side. Which card(s) do you need to turn over to check for errors? [5].

The second theory states that in environments that are dominated by uncertainty, it might be rational to ignore information as a way to prevent overfitting one's predictive model of the results of actions. The set of action-result experience a person can use to learn a action-result model is often very small, resulting in a high risk that models results from actions fit on features that all experiential samples accidentally share. These features are hence perceived to be pointing to a certain result, even though in reality they do not. This can hurt predictive accuracy. Hence, in these cases it might be rational to ignore features. This does, however, require the reasoner to judge what type of environment they are reasoning about [6]. Hence, in this case mathematical sub-optimality might result from a reasoner's wrong estimation of the type of environment it is reasoning about. This would lead to an alternative explanation for people's trouble with the WST: due to the fact that they see the numbers and the letters, they could misjudge the question to be about mathematical equivalence rather than implication.

Gigerenzer and Gaissmaier discuss four classes of heuristics: "the first class exploits recognition memory, the second relies on one good reason only (and ignores all other reasons), the third weights all cues or alternatives equally, and the fourth relies on social information," [6, p. 459]. Exploiting recognition in this case means relying on the theory that recognition is a different cognitive process than recollection, appearing earlier in consciousness.

Heuristics from the first class are of particular interest to the current study. The main heuristic from the first class is the *recognition heuristic*. It prescribes: "if one of two alternatives is recognised and the other is not, then infer that the recognised alternative has the higher value with respect to the criterion" [6, p. 460]. Here, the recognition of alternative A refers to the match-up of A with some piece of information already in one's memory [7]. Two specifications of this heuristic are the *fluency heuristic* and the *take-the-first heuristic*. The fluency heuristic is applicable in situations where "both alternatives are recognized" [6, p. 462] but one is recognised earlier than the other. It prescribes: "If both alternatives are recognized but one is recognized faster, then infer that this alternative has the higher value with respect to the criterion" [6, p. 462]. The take-the-first heuristic presupposes a serial exploration of alternatives and prescribes: "Choose the first alternative that comes to mind" [6, p. 462]. In the context of decision-making, applying this rule would result in one action in a decision instance being recognised earlier than another and hence being selected. Application of these heuristics from the recognition class is rational in an environment where the probability of recognising one out of two alternatives is positively correlated with the probability of the first being the better alternative.

It is important to note here that even though these rules of thumb may often lead you to satisfactory outcomes, they do not ensure that you absolutely maximise your satisfaction in the mathematical sense. Imagine, for example, the *recognition* heuristic in the context where you are cooking dinner and realise you have not bought rice. In a hurry you run to the supermarket, where you find two brands of rice, the one you always buy and a new one. You do not have time to look at the price tags and hence you buy the familiar brand. It could well be the case that the rice you did not buy was cheaper, so because you applied the *recognition* heuristic you have not maximised your utility.

2.2 Computational Decision-making

There exist artificially intelligent systems that, in fully observable domains, exhibit the all-information-rationality associated with economic models, in particular by relying on logical formalisation and complete search algorithms. Logic-based AI aims to use "logical techniques to formalize the reasoning problems that AI needs to solve" [12, para. 34]. Logic entails formally representing and reasoning about information [14]. To formally represent knowledge means to define (strict rules for) the language you allow yourself to represent that knowledge in. This stands in contrast to natural language, where the meaning of words is often ambiguous.

For this paper we look specifically at first-order logic, a formal representation devised to express properties of and relations between objects. A first-order language is designed as a vocabulary that contains "**constant symbols**, which stand for objects; **predicate symbols**, which stand for relations; and **function symbols**, which stand for functions" [14, p. 275]. Predicates can also have one argument, in which case they are referred to as **properties**. In addition to objects, relations and functions, any first-order language also contains a logical vocabulary that expresses how the truth of a certain sentence is related to the truth of its sub-sentences. Table 1 outlines the symbols in this logical vocabulary and their semantics, along with the natural language expressions these symbols express.

Symbol	Meaning	Semantics
$\neg$	"not"	$\neg A$ is true iff A is not true
$=$	"equals"	$(s = t)$ is true iff terms s and t refer to the same object
$\wedge$	"and"	$(A \wedge B)$ is true iff A is true and B is true
$\vee$	"or"	$(A \vee B)$ is true iff A is true or B is true
$\rightarrow$	"if ... then ..."	$(A \rightarrow B)$ is true iff in all cases where A is true, B is true
$\leftrightarrow$	"... if and only if ..."	$(A \leftrightarrow B)$ is true iff A and B are both true or both false
$\forall$	"for all x ... holds"	$\forall x A$ is true iff A is true for all assignments to x
$\exists$	"for some x ... holds"	$\exists x A$ is true iff there is an assignment to x for which A is true

Table 1: An overview of the logical operators for first-order logic.

One of the earliest proponents of logic-based AI was John McCarthy in his article "Programs With Common Sense" [9]. His idea was that a machine should be programmed such that it can be "told" facts and rules about its environment in a formal manner and deduce the consequences of this information for itself, something he denoted by the term "common sense" [9]. In the article, McCarthy proposes a model, the *Advice taker*, as a general illustration of what a system implementing this idea should look like. This program would, for example, be able to tell you what to do if you are seated at your desk and wish to go to the airport [9]. In the '60s and '70s considerable effort has been put into the objective to develop actual systems that can achieve this goal. These systems are able to make all-information-rational decisions in observable domains.

To meet McCarthy's objective, systems should, apart from deriving logically necessary knowledge, be able to weigh actions based on their knowledge and their desired goal. A way of realising this is by using a search algorithm on the graph induced by such a planning problem, defined by (a) an initial state, (b) a set of actions available to the agent in each state, (c) a description of how each actions turns a given state into a resulting state, (d) a set of goal states that define desired results and (e) a path-cost function that assigns a numeric cost to each action [14]. A shortest path to the goal state can then be found by breadth-first search from the initial state, by applying all possible actions to the initial state and listing the resulting states. Then we do the same for each of the resulting states, after first checking whether they are a goal state (Figure 2). If there is a solution to the search problem, we are bound to find it if we search all states in the graph, as the graph is an enumeration of all states (nodes) accessible from the start state by applying actions (edges).

However, the first goal state the agent finds should have the *cheapest* path, not the shortest. This means we have to be selective about which state we choose to expand first. We achieve this by using a best-first search algorithm, which we abbreviate as BFS in this paper. This algorithm selects the cheapest of all not expanded states to expand next (Figure 3). As the algorithm checks a state to be a goal state when it is selected to be expanded, this ensures that the goal state with the lowest cost will be found first [14]. Combining logical deduction and a BFS-algorithm in a computer program produces an agent that always applies all relevant knowledge and is bound to select the cheapest set of actions to acquire a desired outcome and therefore necessarily minimises costs.

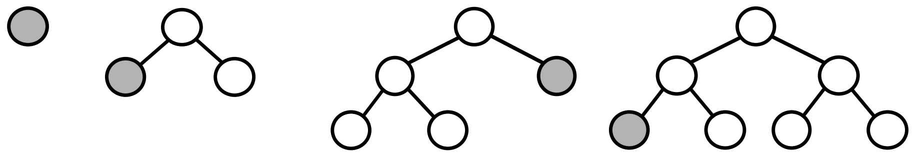


Fig. 2: Example of a search problem with two possible actions in each state. The grey node indicates the state that will be expanded next. This node is checked to be a goal state and if it is not, its successor states are generated by applying all possible actions [14].

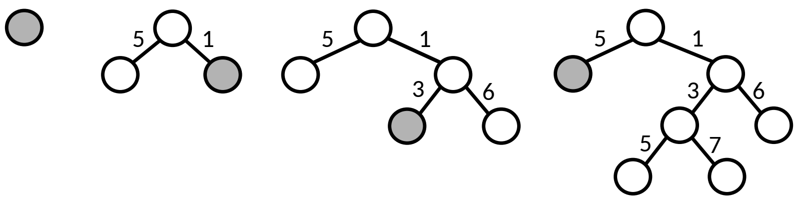


Fig. 3: Example of a Best-First Search. Labels indicate the cost of the path towards each state. States are expanded in a cheapest-first fashion.

3 Research

Given the above exposition, the question naturally arises to what extent there is correspondence or deviation between human and all-information-rational problem-solving strategies in fully observable domains. More specifically, it is interesting to observe specific situations in which such correspondence or differences arise. Therefore, our research seeks to compare human decision-making to an all-information-rational strategy in a concrete situation. We aim to answer the question: *how does human decision-making compare to that of an all-information-rational agent in a concrete decision-making environment?* Based on the preliminary information we hypothesize that human decision-making is mostly all-information-rational, but at times deviates from such a strategy, due to humans' use of heuristics.

To test this hypothesis we require an agent to implement all-information-rational decision making, human subjects, and a task environment that allows for an adequate and precise determination of the all-information-rationality of choices, while still being complicated enough to make employment of heuristics either necessary or a possible strategical candidate. We have chosen to implement the artificial agent in the programming language Prolog because it implements a deductively valid reasoning method (resolution), as argued in Section 2. The task environment is an adaption of Gregory Yub's *Hunt the Wumpus* game [19], to which we add action costs to allow for the determination of all-information-rational decisions given the explicit objective of minimising costs. In Section 4 we describe the experiment's setup, including the specifics of the game *Hunt the Wumpus*, the Prolog agent, the human experiment and our predictions. We present and analyse the results from our experiment in Section 5.

4 Experimental setup

To test our hypothesis we designed and implemented a all-information-rational computational agent and let it play (that is, reason about) three different levels of the game *Hunt the Wumpus* [19]. We have also let human participants play these levels and compared their performance to that of the agent. The game is explained in section 4.1 and the design of the computational agent in section 4.2. We describe how we set up the experiments with our human subjects in section 4.3 and make predictions of the outcomes in an attempt to falsify our hypothesis (section 4.4).

4.1 The Wumpus World

Both the human and the agents played several 'levels' of our adaption of the game *Hunt the Wumpus* that was created by Gregory Yub [19]. The game consists of a square grid of squares (see Figure 4a), of which in the initial state only the bottom-left square $(0, 0)$ can be observed by the agent—at least this is the initial situation in Figure 4a, where $(0, 0)$ is labeled "Start". In our experiments, for completeness' sake, the starting point was always the top-left square $(0, 3)$ (see Figure 7). The goal is to find a square containing the 'gold' by navigating the grid, while avoiding falling into 'pits' and being eaten by the Wumpus, a monster lurking in the grid. The agent moves by turning in the desired direction and going forward. In every square the agent may perceive sensations that hint at a pit or the Wumpus being nearby, as follows. Whenever there is a pit in a square (orthogonally) adjacent to the one the agent is on, the agent will perceive a breeze. When the Wumpus is in an adjacent square the agent will perceive a stench in their own square (Figure 4). The agent also carries one arrow that kills the Wumpus if the agent turns in the direction of the Wumpus and fires the arrow from any number of squares away.

In the version of the game we used for our experiment, actions have costs. Moving one square costs 10 points and shooting the arrow costs 30 points. Finding the gold 'costs' -1000 points and both falling into a pit and being eaten by the Wumpus costs 1000 points. Turning does not have associated costs. The goal is to finish the level with as high a score as possible.

When completing the Wumpus task, it is highly advantageous to have the ability of logical reasoning and cost-based planning. Information about the presence of a Wumpus or pit on a certain square can in many cases be deduced from the sensations in previously visited squares. First-order logic is particularly useful for this task since the quantifiers ($\forall$, $\exists$) enable us to formulate general rules about the squares in the game in a concise manner, without having to address every square individually. For example, we can formulate a rule like

$$\forall s, r((\text{Adjacent}(r, s) \wedge \text{Stench}(r)) \rightarrow \text{PossiblyWumpus}(s)),$$

to express that whenever we have observed a stench in a square r adjacent to some square s it might be that case that there is a Wumpus on square s. Adjacency can similarly be defined in a concise way.

4.2 Prolog Agent

The agent was constructed in Prolog, a logical programming language which can be thought of as mimicking very closely the process of deductive reasoning, allowing for rules and facts to be stated in first-order-like terms. Reasoning steps that the agent should take to successfully optimise its behaviour can therefore be encoded in Prolog in a way that very closely resembles first-order logic formulas. For example, one of the rules of the game is that if a square contains a pit, all adjacent squares exhibit a breeze, or in other words (formulated as the contrapositive): a square contains no pit if there is some square adjacent to it that contains no breeze. This rule can be expressed in first-order logic as

$$\forall s(\exists q(\text{Adjacent}(q, s) \wedge \neg\text{Breeze}(q)) \rightarrow \text{agentKnowsNoPit}(s)),$$

and then immediately in Prolog syntax as the following rule (which is slightly more elaborate to fit in with the rest of our framework, where we need to register which squares have been visited).

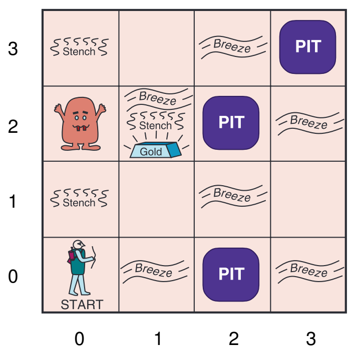
```
agentKnowsNoPit([X,Y],Visited):-
    setof([A,B],(
      adjacent([A,B],[X,Y]),
        member([A,B],Visited)
        ),Z),
    atLeastOneNoBreeze(Z),
    assert(agentKnowsNoPit([X,Y])).
```

All the rules the agent needs, to successfully optimise its behaviour are encoded as Prolog code in this way. We designed the agent to perform a best-first search (as described above), making use of the information that is deducible from these rules, supplemented with some datastructures to allow the BFS algorithm to work smoothly. The source code can be downloaded and studied at https://www.whattodonext.nl/downloads.

Prolog deviates from standard first-order logic to make programming and execution faster. Mainly, Prolog uses database semantics, which entails making three assumptions. The unique-names assumption means that different constants refer to different objects. Domain closure means that no objects exist in the domain than the ones identified by the constants in the language. The closed-world assumption states that atomic sentences that are not mentioned as true are false.

Execution of Prolog programs depends on (a variant of) the resolution method [13]. This mechanism is *sound*, which means all sentences it can derive from a Prolog program are indeed logical consequences of the knowledge in the program. It is also *complete*, which means that it can be used to prove all sentences that follow logically from the program. The specific implementation of the resolution mechanism in Prolog preserves the *completeness* entirely and the *soundness* almost entirely [14].

The state of affairs in a particular Wumpus World can also be encoded in a Prolog source file. We can encode where the Wumpus, the pits and the gold are located and we can construct rules for the percepts that follow from this state of affairs. See Figure 4 for an example.



(a) A configuration of a Wumpus World (source: [14]).

```
% specific world setup
pit([2,0]).
pit([2,2]).
pit([3,3]).
gold([1,2]).
wumpus([0,2]).

wall([X,Y]):- X < 0.
wall([X,Y]):- Y < 0.
wall([X,Y]):- X > 3.
wall([X,Y]):- Y > 3.
```

(b) The Prolog specification of this world.

```
%general rules for
    every world
stench(S):-
  adjacent(R,S),
  wumpus(R).
breeze(S):-
  adjacent(R,S),
  pit(R).
glitter(S):-
  gold(S).
bump(S):- wall(S).
```

(c) Prolog specifications for the general dynamics of worlds.

Fig. 4: A configuration of a Wumpus World due to Russell and Norvig [14] (left) along with its specification in Prolog (right).

Given a certain Wumpus world like the one in Figure 4, the agent can be told to find the optimal strategy with the query `play([0,0],Path,n)`. This in turn calls upon the `bfs`-predicate, which searches for squares that (i) have not been visited yet and are known to be safe, (ii) are sure to contain a Wumpus or (iii) are sure to contain gold. When such a square `EndPos` is found the `Path` towards that square is returned. In case (i) and (iii) this will be a path consisting of only moves, in case (ii) it will consist of moves and end with a shoot action. Then `play` is called again but now with `EndPos` as starting point. This process continues until the agent reaches the square which contains the gold, in which case `play([0,0],Path,n)` succeeds with `Path`, the path that is the optimal strategy. If from some square all possible paths have been considered and none of these reach a square that is safe and not visited, contains the Wumpus or contains the gold, the agent returns "No solution found" followed by the path that has been assembled up to that point.

The agent also has a *guess-mode*. This can be activated by calling `play([0,0],Path,y)`. The process is identical to `play([0,0],Path,n)`, except that when all options for a path have been exhausted it does not return "No solution found" but goes into a new BFS called `guessBfs`. This BFS allows squares that are not visited and `possiblySafe` to be added to its path. In this way it 'guesses' to go to a square that could be safe without knowing for sure that it is. When the discovered square is indeed safe, `play` is called again with `EndPos` as its starting point. When the agent winds up on a square with a pit or the Wumpus, the `play`-predicate returns "Game over" followed by the path that has been assembled up to that point. When all options for the `guessBfs` have been exhausted as well, the agent will return "No solution found" followed by the path that has been assembled up to that point. The agent has been extensively tested on twenty 4×4 worlds with randomly generated positions for the pits.

4.3 Human Experiment

Participants in the human experiment were asked to complete three configurations of the Wumpus World, increasing in logical complexity, and instructed to report their thought process while playing.

Ethical Considerations Participation in this study was completely voluntary, each participant explicitly providing informed consent. Participants' age and education level were recorded as well as their voices. As these can reasonably be categorised as personal data according to the GDPR, all video recordings were transferred to UU's 2FA secure OneDrive server and deleted from the recording device. Furthermore, participants were informed that their consent could be retracted and their data deleted at any time when requested.

Participants In total, 13 participants took part in this experiment. Most of them (9) were aged between 20 and 26 years old, one aged 17, one aged 18, one aged 46 and one aged 48. Of these participants, 12 enjoyed secondary education and are either enjoying or have enjoyed higher education.

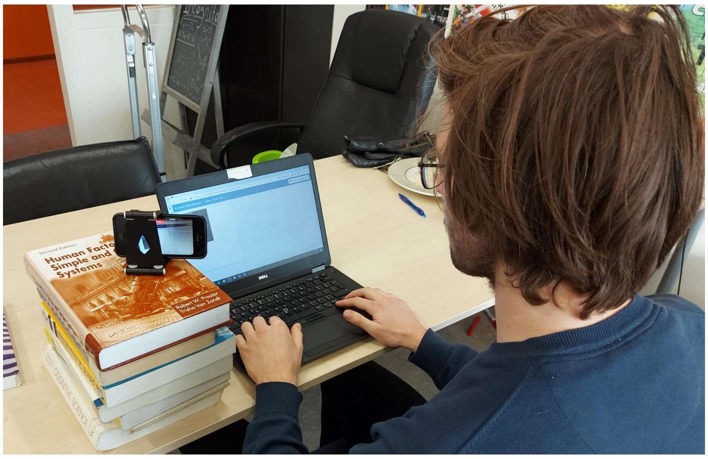


Fig. 5: Participants were seated behind a laptop during the experiment. The screen was filmed and their voices recorded as they performed the task and reported on their thought process.

Apparatus The experiments were conducted using laptops. 12 of the experiments were conducted using a Dell Latitude E5470 laptop with Intel(R) Core(TM) i5-6300U CPU @ 2.40GHz processor, 8.00 GB RAM and a 14 inch screen with 1366x768 resolution. The other experiment was performed on a laptop owned by the participant themself. The experiment can be accessed via https://www.whattodonext.nl/wumpus. During the experiment the participants were asked to play the Wumpus World game and instructed to report on their reasoning steps out loud. The screen was being filmed and their voice recorded. The filming and recording was done using an iPhone SE with 12 MP camera (Figure 5).

Design and procedure The participants took part in the experiment, sitting at a desk, behind a laptop. A researcher was present during the procedure to monitor whether the explication of the participant's thought process was explicit enough and to repeat the rules of the game in case the participant had forgotten one. The experiment started with a welcome screen (see Figure 6). The participant could click a link when they had read the welcome message. Then, on the next screen, the participant could read the instructions to the game. Then they all received an identical small demonstration of how one navigates through the world in a practice level. After that, the participant was free to try their hand at it themselves, still in the practice level. Then, when the participant either found the gold or was game over, they could click on to the next screen. This screen announced that the real levels were about to start. By clicking a link again the real levels started. This is also the moment the supervisor started the video and audio recording. The participants played three levels after which they were redirected to the thank you screen.

The worlds were deliberately constructed in such a way that for each new level the complexity of reasoning required to solve it increases. In all levels it is necessary to identify which paths were cheapest. Then, for world 1 (Figure 7a) it suffices to use one logical rule to solve it: the rule "if a square is empty then all adjacent squares are safe". For world 2 (Figure 7b) there are two necessary rules: "if a square is empty then all adjacent squares are safe" and "when a square has a stench and all adjacent squares are safe except some square [X,Y] then [X,Y] contains the Wumpus". It also requires the insight that shooting the Wumpus when there are no safe squares to be visited is one's only sure chance to visit a new square. World 3 (Figure 7c) requires the participant to identify that no surely safe actions can be made and hence has to identify which squares are possibly safe using the rule "if the perceptions in some square [X,Y] can be explained by other adjacent squares than some adjacent square [A,B], then [A,B] is possibly safe". For some game-plays, the participant should also be able to deduce that if there is some square [X,Y] for which some adjacent square [A,B] does not contain a stench and some adjacent square [C,D] does not contain a breeze then square [X,Y] is safe.

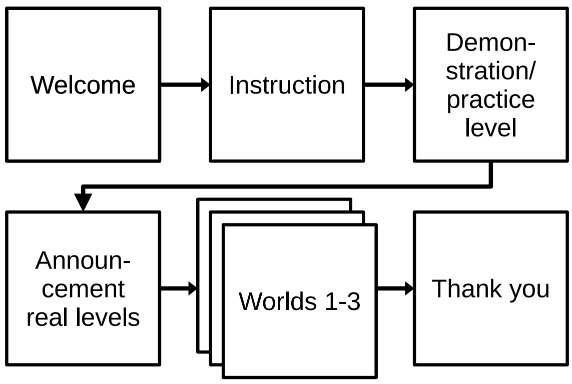
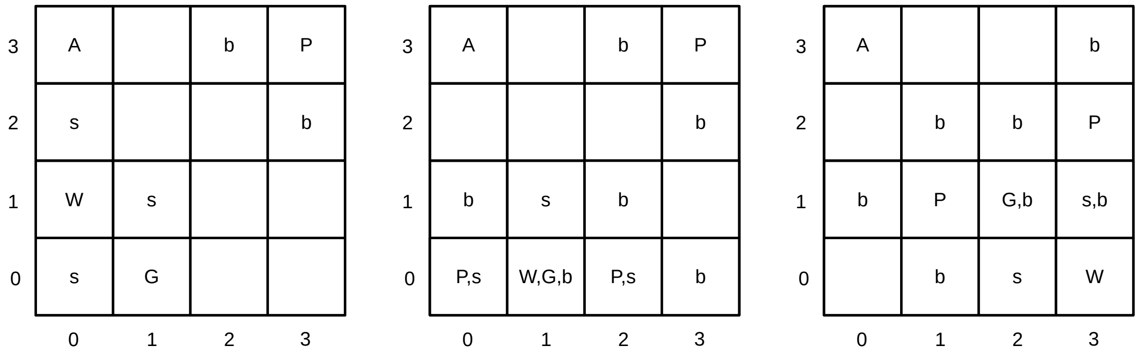


Fig. 6: The setup of the experiments.

Analyses The audiovisual data of the experiment have been coded (see for example [1]) in order to compare them to the performance of the computational agent in a history-independent matter. This was necessary because the order in which the computational agent happens to select actions may exclude all-information-rational strategies that participants might have selected. Firstly, the participants' reports were transcribed. Then, based on the transcripts as well as the video data,



(a) Level 1. (b) Level 2. (c) Level 3.

Fig. 7: A schematic representation of the experiment's levels. The agent is represented by the letter 'A', breezes by the letter 'b' and stenches by the letter 's'.

they were freely coded, indicating inference rules applied by the participants. After that codes were standardised to prevent pieces with the same meaning being indicated by different words.

For example, participant 8's utterance "Geen beginwaarschuwing dus ik ga rondlopen," [English: "No start warnings so I will walk around"] was coded with the code **NPNS**, meaning: "If this square shows no percepts then all adjacent squares are safe." Another example is participant 9's utterance "Dus dat betekent dat de Wumpus onder het stench vakje zit, dus die kunnen we dan schieten," [English: "So that means the Wumpus is under the stench square, so we can shoot it,"] which was initially coded with the code **Weet W → schieten** [English: **Knows W → shoot**] which was after the standisation step replaced by **W → S** meaning "If you know where the Wumpus is, you shoot it." An exampe of a coded transcript can be found in the Supplementary Information (SI) Figure 1 and a list of all standardized codes can be found in SI Figure 2.

Then, codes were grouped in higher-level groups of meaning corresponding to the all-information-rationality of the rule applied. Rules could be either *rational, irrational, conditionally rational* or *neither rational nor irrational.* "Rational" hence be understood as all-information-rational in this context. For example, the rule indicated by the code **NPNS** from the above example is all-information-rational, since this rule follows logically from the rules of the game and the code **W → S** is conditionally all-information-rational, since it is only all-information-rational when there are no squares left that are known to be safe and cost less to reach than it costs to shoot your arrow. Observe that to determine whether the condition for all-information-rationality is met here, the researcher must have access to the current state of the game via the video material. *Irrational* rules are those that are (i) deductively invalid given the rules of the game or (ii) not cost-efficient. An example of case (i) is concluding from observing a breeze that all adjacent square must contain a pit (code **BAAP**) and an example of case (ii) is choosing a systematic approach over cost-efficiency (code **O>C**). *Neither rational nor irrational* rules are those that were sometimes applied to decide between equally all-information-rational choices. For example, when a participant would start level 2 going down from the initial square because they went right in level 1 (an example of the rule with the code **explore**, in contrast to the code **exploit**). The full grouping can be found in SI Figure 3.

Finally the participants' optimal and sub-optimal decisions were counted. In this final step, *optimal* was taken to mean *following rules classified as rational or as conditionally rational provided the condition is met.* In the case of sub-optimal choices it was also recorded which code corresponded with their reason for that choice. The full counting of optimal and suboptimal choices can be found in SI Figure 4.

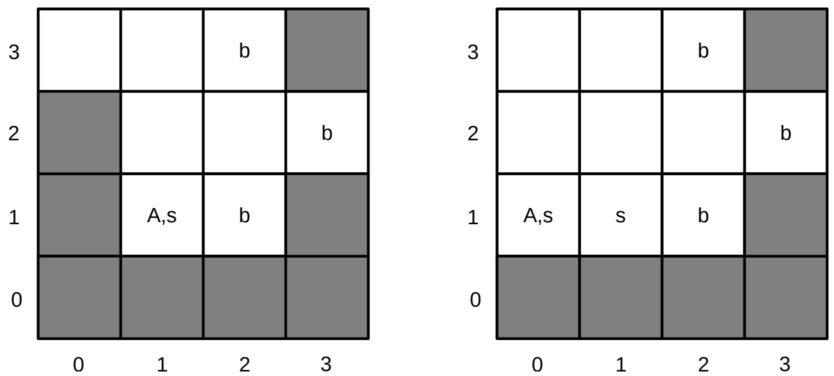
4.4 Predictions

Our hypothesis is that human strategies are 'more than 50% all-information-rational' but, because they apply heuristics, will deviate from those of a all-information-rational strategy more than 0% of the time. Therefore, we expect that in more than half of the cases, participants would select action paths that corresponded to the agent's strategy. In the other cases we expect participants to select actions paths that are not as cost-efficient as possible. We expect sub-optimal decisions to stem from a disregard of some logical rule for the situation or a disregard of some cheaper possibility.

5 Results Analysis

The humans participants' and the agent's strategies were mostly similar. They deduced, from the rules of the game, which squares were necessarily safe. Then they considered which squares were easiest to reach by the fewest steps. They often made use of the rules "if a square contains a breeze and three surrounding squares are known not to contain a pit then the remaining adjacent square contains a pit" (SI Figure 2: PBS) and "if a square contains a stench and three surrounding squares are known not to contain the Wumpus then the remaining adjacent square contains the Wumpus" (SI Figure 2: WSS) to determine the location of pits or the Wumpus. This resulted in an average of 86% choices in correspondence to the agent's in level 1, 81% in level 2 and 86% in level 3.

Interestingly, participants even used methods that were different from the agent's. By that we mean all-information-rational strategies that follow from the rules of the game. For example, participants used the fact that if they perceived a percept in a square, at least one of the adjacent squares had to contain the corresponding danger. They remembered this set of squares. Then, in next stages, by excluding options that became impossible, they eliminated elements of the set until only one was left. Then, they deduced that square should contain the danger. This is an equally fruitful strategy as the one the agent applies, since the moment all but one possibility have been excluded by resolution is exactly the moment when the agent can deduce where the danger is by rule PBS or WSS (Figure 8). However, the way the information is stored in memory and the temporal distribution of the calculation is different.



(a) Initialisation of resolution reasoning. (b) Moment of final conclusion.

Fig. 8: Schematic illustration of a participant applying resolution. Shaded squares are still unknown. In situation 8a the participant remarked: "This square has a stench [(1,1)] (...) so this square [(0,1)] or this square [(1,0)] is the field's only Wumpus." After proceeding and ending up in situation 8b they remarked: "This is the Wumpus [(1,0)]." Note that it is equally fruitful to use WSS in the second instance to deduce that the Wumpus is on square (1,0).

Nevertheless, participants also made choices that were not in accord with the agent's. This could be explained by the employment of heuristics. In particular the rules "if there is no breeze on some square then all adjacent squares do not contain a pit" (NPNB) and "if there is no stench on some square then all adjacent squares do not contain the Wumpus" (NWNS) were not applied, constituting 25% of sub-optimal decisions. It is not surprising that participants failed to observe these rules, which are contra-positives of the quantified rules of the game as presented to the participants. This is exactly in accordance with the evidence found in the Wason Selection Task [18].

The other big cause of deviation was that participants not always made choices with the lowest cost (23%). The Wumpus game is an environment where recognising a strategy from before is not highly correlated with it being the cost-minimising strategy as each level requires a different reasoning rule. Hence, particpants' failure to consistently perform cost-minimising actions might be attributed to the take the first heuristic: actions that resulted in the discovery of safe squares before might be recognised first even though in a new level minimising costs might require a new action.

Other causes of deviation from the agent's strategy were:

1. Participants tended to shoot the Wumpus as soon as they knew where it was.
2. Some participants tended to infer from the breezes the minimal number of pits that could explain them and then (often incorrectly) considered all other squares to be safe (Figure 9).
3. Some participants expressed the desire to maintain overview by working in a structured manner. This structured approach caused them to take extra steps. For example, they would go from left to right, then completely back to the left, then down, then to the right again, etc.

The tendency to shoot the Wumpus as soon as it has been discovered can also be explained by the *take-the-first* heuristic. Since in most video games it is always desirable to kill an opponent when one gets the chance, it is likely that the action of shooting the Wumpus is recognised earlier than the alternative action of moving, resulting in the choice to shoot the Wumpus, even though moving away instead might have been the cost-minimising strategy. The tendency to approach a problem in a structured way, on the other hand, could be accredited to the *recognition* heuristic [6]. In most puzzle-games it is wise (and at zero cost) to choose actions that preserve one's overview of the problem. Hence, recognising the option of a structured action earlier than more haphazard alternatives might result in those actions being applied.

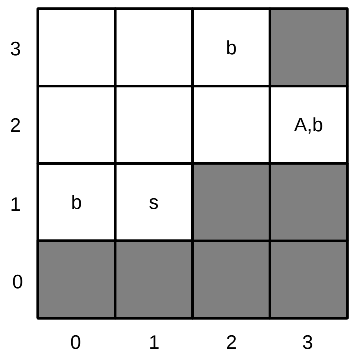


Fig. 9: In this situation participant 5 remarked: "That's a pit, up there [(3,3)] (...) Then I can probably go down here."

Interestingly, the tendency to shoot the Wumpus pre-emptively was also observed in one of the two participants over 40. As these individuals may not have been raised with computer games, a different explanation might be required. An alternative explanation could be that pre-emptive shooting of the Wumpus comes from a desire to minimise the amount of information that has to be memorised. Conscious of their own limited memory capacity, the participants shoot the Wumpus. Then, at least, they do not have to remember where they had deduced the Wumpus to be.

As for comparison among levels, the majority of sub-optimal decisions in level one were due to the pre-emptive shooting of the Wumpus, while in level two and three the majority of mistakes was made due to a mix of uncommon mistakes. It seems that whenever participants were not sure what to do they just acted randomly, thinking up a reason later. This could be a sign of rationalisation. Cushman [2, p. 1] writes that "rationalization occurs when a person has performed an action and then concocts the beliefs and desires that would have made it rational."

The prevalence of pre-emptive shooting of the Wumpus in level one could be explained by the fact that it is not hard to deduce where the Wumpus is in that level. It is only necessary that a participant starts exploring in a downwards fashion to find the Wumpus' location in relatively few steps. In world two and three, finding the Wumpus is considerably harder. Moreover, in world two, in most cases, it is an optimal move to shoot the Wumpus whenever you have derived where it is. For level three it rarely happened that participants found the Wumpus, because it is in the lower right corner. Most participants would not even reach that corner because they would die or find the gold earlier.

6 Discussion

Our results answer our research question *How does human decision-making compare to that of an all-information-rational agent in a concrete decision-making environment?* by confirming our hypothesis that human decision-making mostly corresponds to that of a all-information-rational agent, but that, at times, it deviates. We have seen that an all-information-rational computational agent does not face computational limitations in our versions of the Wumpus World. It tackles the Wumpus World by calculating which game states can be accessed via the current game state and what the associated costs are. In this way it searches for the cheapest path resulting in more information and does not risk to go Game Over, except for when such a risk is absolutely necessary. By implementing a best-first search and logical deduction in this way the agent always finds the optimal decision.

Human beings largely adopt the same strategy, but at times do not work out all possibilities the computational agent does. They do not consider all possibilities, either because they cannot hold everything in their memory at once, fail to deduce relevant information that logically follows from the game state and rules of the game or wrongfully judge the environment to be one where employing heuristics is rational. Specifically contrapositive reasoning seems hard for people, which is in line with the findings of the Wason Selection Task [18]. Both these causes are computational in nature, supporting the idea of human rationality being "bounded" by computational constraints [17]. To overcome these limitations humans use heuristics, but these sometimes result in sub-optimal behaviour as well.

This research's strength lies in it's 'ecological' approach to human heuristic reasoning. Rather than investigating a specific heuristic, we have focused on creating an environment in which the desire to act rationally coupled with computational constraints gives rise to heuristics that were not predefined by the researchers. Hence, this study has given us the opportunity to study how and where heuristics arise in a concrete environment.

On the other hand, this focus on ecology rather than experimentation makes result analysis harder. We have not obtained numerical results, but rather spoken reports and video material that requires interpretation. Hence, the research's analysis cannot be seen as fully separate from the researchers' own ideas about rationality. Let this research's analysis then partially be understood as engaging in the ongoing discussion about how we should understand the concept of rationality.

Another limitation of this study is that we, as the writers of the source code for the agent, had to judge whether the participants acted in accordance with it. The question thus arises whether the agent then was necessary in the first place. The answer to this question is yes and for two reasons. Firstly, by building the agent we have sharpened our vision on what all-information-rational choices are in the Wumpus World. We encountered many edge cases where we were forced to think about what the all-information-rational strategy was at that point, without having thought of it ourselves beforehand. We are sure that we would not have been able to analyse the participants' performance had we not written the agent first. (Note that this is a sign of our own computational limitations.)

Secondly we would argue that judging the performance of the participants' according to our knowledge of what the agent would do, is not any different from having a computational agent do it. That is, we have not at every stage of the game asked ourselves whether the decision the participant made was the all-information-rational action. We simply considered whether the participant's action was in accordance with the rules of the program. One could say that in this way we have become a computational agent ourselves. We are not alone in this thought. Consider John Searle's Chinese Room argument, where he argues that computers can have no consciousness if they simply follow rules without knowing what these rules mean [15]. In following the rules of the program, we have in a way become such a computer.

Lastly, this study would have benefitted from a more diverse group of participants. Unfortunately, time and resource constraints did not allow us to look far for participants. Hence, the participants are mostly students or have enjoyed higher education.

All taken together, the results of this study support an idea of human rationality as heuristic and bounded by computational constraints. When compared to a all-information-rational computational agent, human subjects perform mostly in accordance with it, but at times fail to do so. We hope this study can serve as an example of how artificial intelligence can be employed to study human decision-making. By comparing human behaviour to a computational agent's, we can gain insight into how human beings reason in a concrete problem-solving environment. Further research in this area might be directed at human decision-making in other games to enable us to filter which human decision patterns are game-independent. Conversely, the outcomes of studies like this can help improve AI-systems that interact with human beings, like self-driving cars and care robots. In this way we can help ourselves, and AI-systems, to better understand how we think.

References

1. Auerbach, C., Silverstein, L.: Qualitative data: An introduction to coding and analysis. New York University Press (2003)
2. Cushman, F.: Rationalization is rational. Behavioral and Brain Sciences **43** (2020)
3. Evans, J.: Logic and human reasoning: An assessment of the deduction paradigm. Psychology bulletin **128** (2002)
4. Evans, L., Wilding, E.: Recollection and familiarity make independent contributions to memory judgments. Journal of Neuroscience **32** (2012)
5. Friedenberg, J., Silverman, G., Spivey, M.: Cognitive science: An Introduction to the study of mind. Sage Publications (2022)
6. Gigerenzer, G., Gaissmaier, W.: Heuristic decision making. Annual Review of Psychology **62** (2011)
7. Goldstein, D., Gigerenzer, G.: Models of ecological rationality: The recognition heuristic. Psychological review **109** (2002)
8. Jones, B.: Bounded rationality. Annual review of political science **2** (1999)
9. McCarthy, J.: Programs with common sense. In: Proc. Symposium on Mechanisation of Thought Processes (1958)
10. Pindyck, R., Rubinfield, D.: Microeconomics. Pearson (2018)
11. Priest, G.: Revising logic. In: Rush, P. (ed.) The Metaphysics of Logic. Cambridge University Press (2014)
12. Richmond, T.: Logic and artificial intelligence. In: Zalta, E. (ed.) The Stanford Encyclopedia of Philosophy. The Metaphysics Research Lab (2020), https://plato.stanford.edu/archives/sum2020/entries/logic-ai/
13. Robinson, J.: A machine-oriented logic based on the resolution principle. Journal of the ACM **12** (1965)
14. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Pearson (2022)
15. Searle, J.: Minds, brains, and programs. Behavioral and Brain Sciences **3** (1980)
16. Simon, H.: Models of man: Social and rational. Wiley (1957)
17. Simon, H.: Bounded rationality in social science: Today and tomorrow. Mind & Society **1** (2000)
18. Wason, P.: Reasoning about a rule. Quarterly Journal of Experimental Psychology **20** (1968)
19. Yub, G.: Hunt the Wumpus. Peoples Computer Company (1973)