*Review*

# Biological and artificial cognition: what can we learn about mechanisms by modelling physical cognition problems using artificial intelligence planning techniques?

Jackie Chappell[1],* and Nick Hawes[2]

[1]*School of Biosciences, and* [2]*School of Computer Science, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*

Do we fully understand the structure of the problems we present to our subjects in experiments on animal cognition, and the information required to solve them? While we currently have a good understanding of the behavioural and neurobiological mechanisms underlying associative learning processes, we understand much less about the mechanisms underlying more complex forms of cognition in animals. In this study, we present a proposal for a new way of thinking about animal cognition experiments. We describe a process in which a physical cognition task domain can be decomposed into its component parts, and models constructed to represent both the causal events of the domain and the information available to the agent. We then implement a simple set of models, using the planning language MAPL within the MAPSIM simulation environment, and applying it to a puzzle tube task previously presented to orangutans. We discuss the results of the models and compare them with the results from the experiments with orangutans, describing the advantages of this approach, and the ways in which it could be extended.

**Keywords:** comparative cognition; mechanism; function; modelling; planning; artificial intelligence

## 1. INTRODUCTION

In recent years, increasingly elegantly designed experiments testing a range of non-human animal species in both the wild and captivity have shown that many species are capable of behaviours of complexity that surprised investigators (see recent studies [1–3] for reviews). However, these experiments have also revealed that in many cases, we (animal cognition researchers) lack appropriate analytical tools to enable us to de-construct those behaviours, compare the competences within or between species, or to tentatively assign biological mechanisms. In this study, we suggest a novel approach to this problem, using techniques from the field of artificial intelligence (hereafter, AI) as an analytical tool to help researchers understand the domain being tested, to plan appropriate experiments and to facilitate quantitative and qualitative analyses of animal behaviour.

### (a) *The current situation and its problems*

As animal cognition researchers, we design tests for our chosen species in order to determine whether individuals of that species possess a particular hypothesized cognitive capability. These capabilities are usually given broad, functional labels (such as causal reasoning,

planning or theory of mind), each of which describes a suite of related behaviours that may have a number of possible levels of complexity. In addition, the behaviour of animals on a particular task may span the functional labels given earlier. For example, an experiment designed to test physical causal reasoning through tool use might also involve planning if the subject has the opportunity to choose among different strategies before starting to act. Mindful of this, researchers try to design experiments in such a way that the cognitive capability of interest is isolated [4], and that the subjects' responses in the experiment are diagnostic of their level. Because these kinds of capabilities are defined in such a way that they could not be achieved through associative learning alone (though this tends to be a controversial assertion), researchers must also ensure that they can exclude the possibility of associative learning by limiting the number of trials, presenting novel tasks or requiring the subject to abstract general principles from learned examples in order to solve the task [5].

However, there are a number of difficulties with this approach.

(i) Excluding all possibility of associative learning is extremely difficult in practice. Even if the number of trials is limited, animals may still learn rapidly within the first few trials. It is possible to analyse *only* the first trial of a particular kind of test, but this can make conventional statistical analyses difficult (see also recent studies

[4,6]). If novel tasks or stimuli are used, one faces the problem of defining what is novel for that individual. How dissimilar does a stimulus or task need to be before it can be regarded as novel? Do we know, or can we make firm assumptions about, the subject's prior experience? These are all difficult questions to answer, particularly when working with wild animals where the animal's history may be unknown.

(ii) There is a persistent problem in animal cognition [5,7] and related fields [8] of attempting to apply dichotomous or binary labels to cognitive capabilities (see Chittkka *et al.* [9] for an extended discussion on this topic). For example, researchers might claim that performance on a particular task does or does not involve associative learning, or that a subject has shown that it does or does not plan. This is clearly inadequate to describe the richness, complexity and variety of cognitive abilities in animals. First, as mentioned earlier, capabilities such as causal reasoning or planning have multiple levels of complexity, and while researchers may aim to differentiate between these levels experimentally, it is not always possible to do so. Subjects' responses on the task may fall between two levels, or they may show behaviours relating to multiple levels of the capability. In addition, components or sub-components of these kinds of complex behaviours may be generated by different mechanisms. Thus, it is generally unjustifiable to claim (on the basis of observed behaviours) that the animal's responses can be explained or cannot be explained by associative learning (for example), because a complex internal process may use a mixture of mechanisms of different kinds, some associative and some not. Alternatively, the same observed behaviours could be generated by different internal mechanisms.

(iii) The variety of experimental protocols used and the difficulty of explaining performance makes it very difficult to compare the capabilities of interest between and within species (see Thornton & Lukas [6] for an extended discussion on this topic). Because it is often not made explicit what the task presented in the experiment involves (i.e. what the components of the task are, what information is required to solve the task, the different ways in which the task can be solved, etc.), it is difficult to assess whether tasks posed in different experiments are equivalent in complexity, or even targeting the same cognitive ability.

(iv) Well-controlled experimental protocols testing physical cognition are very difficult to design because of the problems outlined in points (i) and (ii), and because, in practice, subjects often show solutions to the task that were not anticipated when the experiment was designed. This tends to happen when researchers are focussed on testing a hypothesis about a particular functional label (and therefore neglect the possibility that other processes may be at work), or because the problem posed to the animal has not been thoroughly analysed in advance to discover the *possible* ways in which it might be solved.

(v) Once designed, experiments can also be difficult to analyse because they involve limited numbers of trials or first-trial performance only, because there is often substantial inter-individual variation (see recent studies [4,6,10,11] among many others), or because animals show unanticipated behaviours (see point iv). Quantitative analysis is desirable, but it would also be useful to compare *patterns* of behaviour qualitatively against a set of models describing the possible ways in which the task might be solved.

In summary, the problems with the current state of the art in animal cognition experiments can be grouped into two broad sets of related issues: (i) imprecision in the predictions made by the underlying theories; (ii) lack of understanding about what the tasks we set for our subjects actually require of them. Imprecision in the predictions made arises (in part) because associative learning and the various broad functional labels mentioned earlier do not have the same status. While we have a good understanding of associative learning at both the behavioural and neurobiological levels [12], we do not understand the biological mechanisms underlying cognitive capabilities such as planning. Thus, when we assign a label such as 'planning' to a behaviour, we are not specifying a mechanism (see earlier studies [9,13] for an extended discussion). In addition, because we often do not fully understand the problem domain (i.e. all the potential actions that can be taken in the course of solving the problem and their consequences) in which we test our subjects, we cannot accurately map putative cognitive capabilities on to performance, and we risk missing possibilities that we may have overlooked.

## 2. OUTLINE OF THE MODELLING PROCESS

In this section, we broadly outline a modelling process that we claim can improve upon the current situation by allowing researchers to understand the domain being tested, and to help them plan experiments and analyse the results qualitatively and quantitatively. The modelling technique is based on *AI planning* [14, ch. 10], but the general concepts and workflow (figure 1) could be implemented in a number of different ways, using different modelling techniques. In §4, we use this approach to model a task previously presented to orangutans (*Pongo pygmaeus*), in which an orangutan must push a nut through a horizontal tube to an opening, while avoiding allowing the nut to fall into inaccessible traps [10].

AI planning is a technique that searches for a series of *actions* that an agent can execute to achieve a *goal* (see also Shanahan [15] for a detailed example of how searches of action-trees can be implemented). The resulting series of actions is called a *plan*. A goal for an agent is a description of a specific future *state* of the world. States are described using a list of *facts* that are currently true, with facts usually being represented in a variant of predicate logic. The actions an agent can perform are represented using *preconditions* and *effects*. Preconditions describe the facts that must be true in
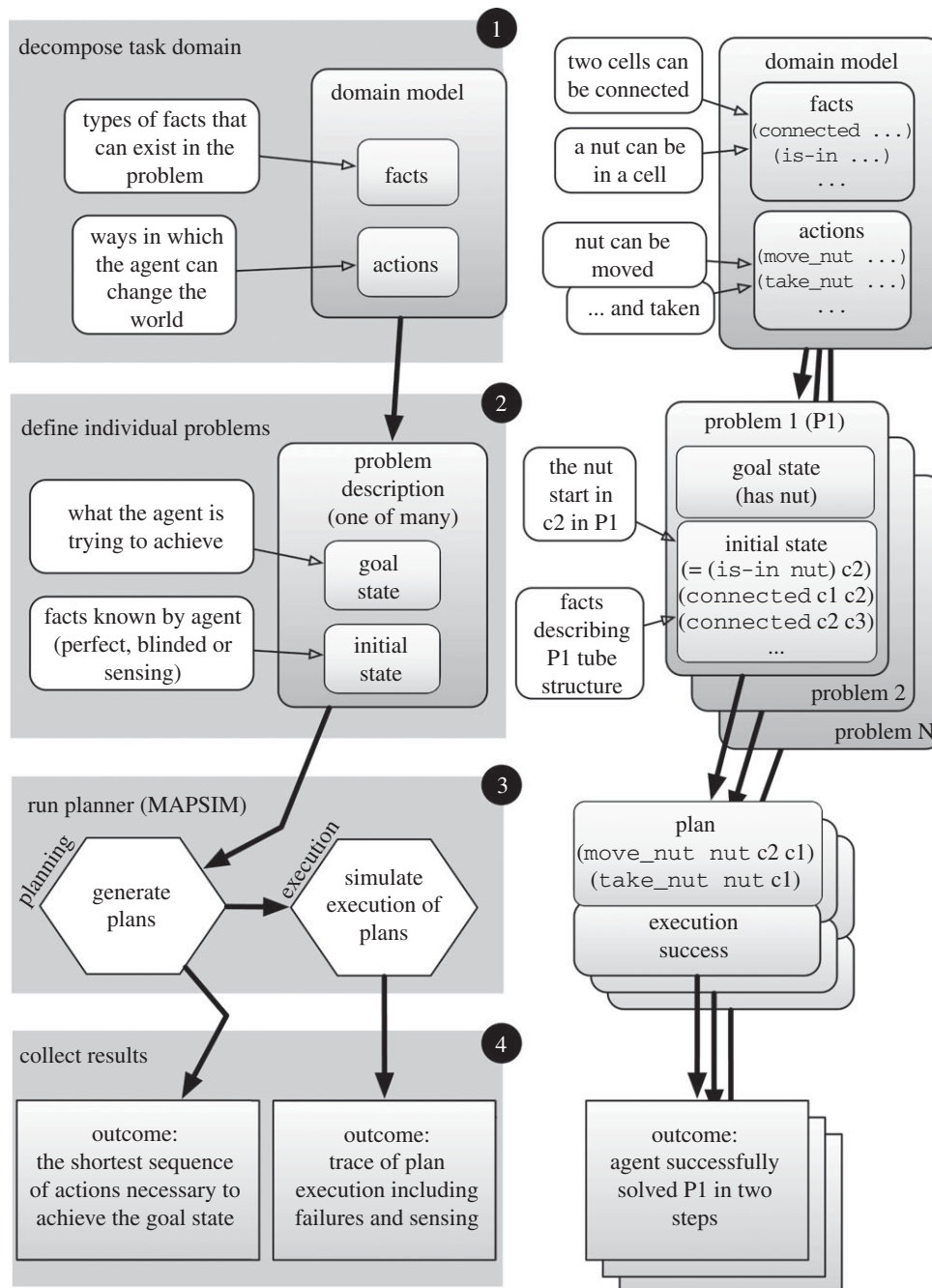
Figure 1. A flowchart outlining the general process of modelling using artificial intelligence planners. The numbered panels in the left-hand column outline the main steps in the process, whereas the right-hand column provides a simple example.

the current state before the action can be used, and effects describe the changes to the state that occur owing to action execution. A classic AI problem that can be solved by planning is the Blocks World [16]. In this, an agent is presented with towers of blocks and must rearrange the blocks to build a different tower. In this world, the *actions* available to the agent are 'pick-up a block' or 'put down a block'. The *state* is a list of *facts*, each one describing the position of a block in a tower relative to the block directly below it. The *preconditions* of picking a block up are that the agent's hand is empty and that nothing is stacked on top of the block in question, and the *effect* of putting a block down on top of another block is that the *state* now contains an additional *fact* describing their relationship. Given a *goal state* describing a particular tower, a

planner will search for the shortest sequence of actions to create that tower from the *initial state* it has been given.

A large number of different representations and algorithms for planning exist [17–19]. In our work, we have chosen to use the MAPL language to represent our problems, and the associated MAPSIM simulator to generate plans and simulate their execution [20]. A number of alternatives exist to search-based planning, including Markov decision processes [21], reinforcement learning [22] and reactive behaviour generation systems [23]. While these systems use different algorithms and assumptions to generate behaviour, they all require problems to be formulated as states and actions in a way similar to AI planning. Other approaches such as pure behaviour-based systems [24] or neural networks require less designer-provided structure but may
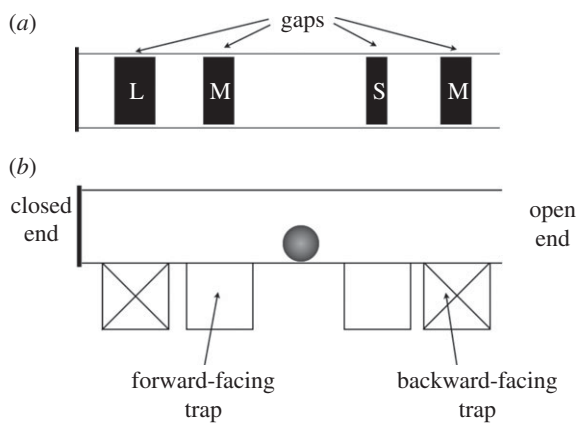
Figure 2. An example of one configuration of the puzzle tube. (*a*) Top view of the tube, showing the gaps in the floor of the tube which provide or deny access to each of four traps beneath. (*b*) Side view of puzzle tube, showing end and trap configurations. This particular configuration would be coded as OCRE in figure 3, indicating that one end is open and one closed, and that the nut can be obtained from the open end of the tube, which is on the right.

exploit or learn structure that is present in the task or environment. While these alternatives to AI planning all have features that might make them interesting for modelling, we have chosen to use AI planning because it supports the separation of the behaviour of the world from the knowledge about the world held by the agent, and because the general approach is comprehensible with limited or no understanding of computer science.

## 3. AN EXAMPLE: PUZZLE TUBE TASK
### (a) *Description of the task*
We have used AI planning to explore a range of models for a puzzle tube task previously used with orangutans [10], in a way that allows us to compare our simulated results with those obtained from real animals. The goal for the orangutan was to roll a nut through a horizontal tube to a point at which it could access the nut. The tube had a number of interchangeable components: each end of the tube could be open or closed, and there were four gaps in the floor of the tube, each leading to a vertical trap (see figure 2 for an example of the puzzle tube in one configuration). The gaps could be small, medium or large, and the nut fitted only through the large gap. The traps could be forward-facing (towards the orangutan) or backward-facing (away from the orangutan). The nut was placed in the middle of the tube (between the two sets of gaps/traps) at the start of the trial, and the task for the orangutan was to move the nut either left or right to attempt to gain access to it. It could get access to the nut at one of the ends of the tube if one was open, or through a forward facing trap, but only if the trap had a large gap above it, allowing the nut to drop through. The orangutan could 'lose' the nut temporarily by rolling it against a closed tube end, or lose it permanently by rolling the nut over a large gap above a backward-facing trap. In the configuration illustrated in figure 2, the correct response would be to move the nut towards the right, because the nut would roll over

both the small and medium gaps and could be accessed from the right end. Subjects were allowed to correct their errors if the nut was not lost to the trap, but the orangutans' initial choice of direction was scored. In comparing the results from the orangutans and the models in this study, the same conventions were used.

The orangutans were each tested on 64 unique tube configurations, and we have run each of our models on the same 64 configurations, allowing us to compare actions on a trial-by-trial basis. The configurations were structured in such a way that on each trial, moving the nut in one direction would result in obtaining the nut, and moving it in the other would result in losing the nut. Thus, for each configuration, there is a correct and incorrect direction in which to move the nut, and all trials allowed the nut to be obtained. In §4, we describe in more detail the processes needed to model this problem using the MAPSIM AI planner.

## 4. MODELLING THE PUZZLE TUBE DOMAIN
### (a) *Decomposing the problem*
First, the experimental problem we intend to set for the subject must be decomposed into a 'domain model' containing the types of facts that can be used to describe states and the actions used to change the states. In particular, the decomposition should identify the physical components of the problem (in our case, gaps, ends, traps, etc.), the ways they can change (e.g. tube ends can be open or closed) and the possible relationships between the physical components. To create a domain model, suitable facts should be created to represent all of these things. For example, the tube must be divided into physical 'cells' and facts created describing whether or not any two cells are con-nected. Given these fact representations, the next step is to encode the actions that change the state. These actions can define both what the subject can do (e.g. push the nut right), and what happens as a consequence of the subject's action (e.g. if the subject pushes the nut over a large gap, then the nut falls through the gap into the trap below). Some planning approaches, including MAPL, are also able to represent sensing actions that the subject may use to gather knowledge about the task as it acts (e.g. sensing the width of a particular gap).

The domain modelling is perhaps the most important step in the modelling process as it unambiguously defines the kinds of the things the subject must need to know about and do in order to solve a particular problem. This process relies on the designer's ability to produce appropriate logical representations by abstracting away from the real situation. This includes discretizing continuous values (such the distance a nut can move) into distinct relevant intervals (e.g. the cells in the tube domain). AI planning provides no guidance on how to best do this. When producing domains for artificial systems, it is usually best to abstract as much as possible while still capturing the essential structure of the problem. This is because more abstract domains produce smaller search problems when looking for plans. It is not clear that this approach is desirable when modelling animal behaviour, as the abstraction
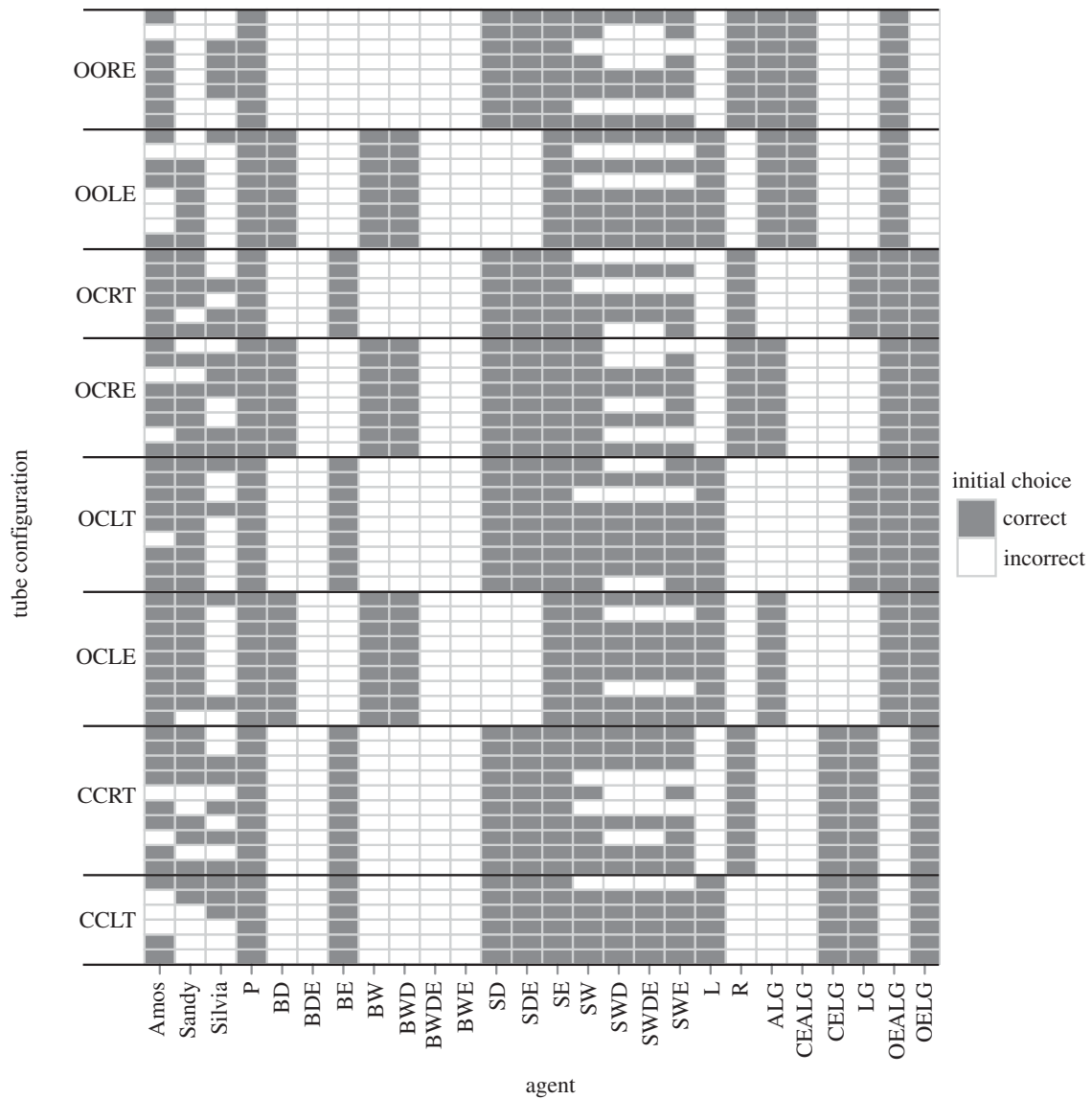
Figure 3. Output grid comparing correct or incorrect initial choice on each tube configuration for the following: orangutans (Amos, Sandy and Silvia); the perfect information system (P); blinded facts models (starting with B, where the additional letters indicate that they are blind to direction of trap, width of gap or state of the ends); sensing facts models (starting with S, where the additional letters have the same meaning as for the blinded facts models); manually produced simulations (using a spreadsheet to manually calculate the outcome of each of 64 trials) from the original paper (feature is chosen unless 'avoid' is specified: L, always move left; R, always move right; ALG, avoid large gap; CEALG, closed end/avoid large gap; CELG, closed end/large gap; LG, large gap; OEALG, open end/avoid large gap; OELG, open end/large gap). Note that simulations involving random choices are not shown here as they do not produce determinate outcomes on each trial. Tube configurations are grouped and coded according to the open/closed state of each end (OO/OC/CC) and whether the nut can be obtained from the open tube end at the left or right (LE/RE) or the large trap on the left or right (LT/RT).

process may discard parts of the problem that are important for modelling the subject. For example, as we describe later, the planning agent used in our models has two actions, one of which is 'take-nut'. In our model (in the interests of producing a clear and simple explanation of our approach), this is a very simple action that can be executed when the nut is either in a forward-facing trap, or at an open end of the tube. However, for a real orangutan, there may be a functional difference between using the action at the two locations: one may be physically easier to achieve, or may entail a smaller risk of the nut rolling on the floor and being taken by a competitor. Furthermore, we do not attempt to model the motor activity involved in grasping the nut with the digits and removing it from the trap. A full

decomposition of the problem would have to consider all possible actions that might be used to obtain the nut, such as pushing the nut hard to make it jump over the gap (a tactic that our orangutans occasionally used). In addition, one would need to be more thorough in specifying all the ways in which information might be obtained about the task, such as tactile exploration and so on. Nevertheless, modelling is an iterative process, and one can start with abstract representations, then gradually make them more concrete to capture the important parts of the world for the animal being modelled.

In modelling the puzzle tube domain, we started by discretizing the tube into individual cells that can be connected. Each cell can be empty, contain a trap or

be a tube end. Traps can be one of three widths (small, medium or large) and face one of the two directions (forward or backward). Tube ends can be either open or closed. The planning agent has only two actions: `move_nut` and `take_nut`. The former moves the nut between cells, and the latter allows the agent to retrieve the food reward from the tube (from a forward-facing trap, or from an open end). The position of the nut in the tube is represented with the function `is-in`, so that for every cell in the tube, the function `is-in` will return true or false, depending on the position of the nut. The domain model features the `derived` predicate `(is-lost? nut)`, which is automatically made true if the nut is pushed into a large, backward-facing trap. The `move_nut` action has `(not(is-lost? nut))` (i.e. the nut is not lost) as a precondition. Taken together, these create the causal logic that a nut pushed over a large, backward-facing trap can no longer be moved. The `take_nut` action has the effect that `(has-nut? nut)` is true. This predicate indicates that the agent has obtained the foodreward. The preconditions of `take_nut` are that the nut is either in a forward-facing trap with a large gap above it, or in the open end of the tube.

When specifying an individual puzzle tube planning problem, it is possible to omit facts that describe parts of the setup. This can be carried out to represent the subject's lack of information about this fact at the task's outset. To allow the subject to overcome this omission, we extend the domain model with one or more sensing actions to allow the agent to sense: the width of the trap; its direction; or whether an end is open or not (see §4b,c for details). We included sensing as an explicit action to highlight the fact that one should consider how animals obtain information about the physical structure and relationships between parts of the task. By generating all possible combinations of these three sensing actions, we created seven different domain models: one that could sense no facts, one that could sense all facts and five different domains that each provide sensing actions for different combinations of facts (e.g. open ends but not trap direction or width).

While this domain model abstracts away a great deal of complexity from the problem (including spatial reasoning, and continuous time and space), it produces a representation of the problem that is easy to understand, and its structure has a direct mapping to the representations used to specify and interpret the original animal experiments.

### (b) *Defining individual problems*

The domain model defines the space of possible problems a subject may face. Within this space, a researcher must create concrete 'problem descriptions' to reflect the different specific configurations or treatments of the experimental problem (figure 1). For AI planning, each problem description will yield a different initial state for the planner, defined in the language of the domain model. The problem description must also include the goal for the current problem. Given

these, the planner can then reason about the actions necessary to achieve the goal from the initial state.

The problem description describes exactly what the agent knows at the start of a trial. Separating problem-specific knowledge (the problem description) from problem-general knowledge (the domain model) allows the researcher to test different hypotheses about the knowledge held by the subject. For example, if the subject did not know that the nut would fall through the large gap, how would that alter the patterns of solutions produced? Note that (at least initially), no assumptions are made about the mechanisms by which the subject might have gained such knowledge, only that it has or does not have a particular piece of knowledge. More complex models than we demonstrate here could extend this by allowing the agent to learn over time. For example, the agent in the model would lack knowledge that the nut would fall through the large gap, but could update its knowledge following its first experience of that type of event. In this way, one could construct a principled method of predicting the outcome in all trials *if* a particular component of the subject's behaviour (e.g. knowledge of the consequences of moving the nut over the large gap) is supported by a certain biological mechanism (e.g. associative learning). Again, it is useful to undertake this exercise before testing subjects to get a detailed picture of the scope of the proposed task, and perhaps to modify the design if the cognitive capacity of interest is not isolated cleanly.

We created 64 different problem descriptions using the 64 different tube configurations from the original animal experiments. To examine the use of sensing, we also created copies of these descriptions by systematically omitting all possible combinations of the facts potentially obtainable by sensing: trap width, trap direction and end state (open or closed). We therefore had seven different problem sets of 64 configurations: one containing all the facts, and one with all the earlier-mentioned three fact types omitted, and five with different partial of omissions of these types.

### (c) *Running the model*

At this point, the model can be run to generate predictions of subject behaviour. In the case of AI planning, running the model entails running the planner to create a plan, then executing this plan in a simulation of the world to determine its consequences. As the subject may not have complete or accurate knowledge of the world (as in our **B** and **S** systems described below), the execution step is crucial in determining what actual behaviour the model produces. The fact that MAPSIM is able to simulate plan execution, including replanning when a subject's expectations are violated, was one of the reasons we chose it for our modelling tool.

If a MAPSIM planning agent has complete knowledge of its environment, then it will find a plan if one exists. MAPSIM is designed in such a way that this plan will be the shortest possible plan in terms of the number of actions. MAPSIM and other planners are also able to reason about plan *cost* as distinct from length. For example, one could assign a cost to sensing the environment, switching direction or any other kind of action, which could be thought of as representing

the time or energy costs incurred by the real animal being modelled. For simplicity, we have not used cost measures in the example in this paper (even though they simplify some domain modelling tasks). Different planning techniques may also provide different guarantees on whether they find a plan if one exists and whether they always return the best plan according to some cost measure.

When using MAPSIM (or any planning approach), the result of running the model is the sequence of actions the subject performed, plus a report on whether it achieved its goal or not. If the agent is able to sense facts during plan execution, it may decide to create a new plan when it gains relevant information (this is known as replanning). This means that the results of running such models may show direction switches in strategy mid-trial.

If a model is run for each problem description *before* real subjects are tested, potential problems with the experimental design may become apparent through patterns in the output from the model that might not have been immediately obvious. For example, it may become apparent that the subject can perform significantly better than chance by attending to a single simple visual cue, because the relationship between that cue and the reward position has not been properly balanced between the different configurations of the task. This process of 'pre-modelling' can be particularly helpful when the experimental design is complex and potential problems are consequently harder to detect unaided.

We ran our seven different domain models on selected problem sets from the possible seven described earlier. This process produced three classes of model/problem combination (see electronic supplementary material, figure S1 for a full list)[1]:

—Perfect information: this agent has all the facts about the world when it starts planning and thus does not need to perform sensing. This is our **P** system.
—Sensing facts: these agents are given problems in which facts are removed from the problem description, but they are able to sense them during execution. These are our **S** systems. The letters after the **S** refer to the type of facts they are required to sense, e.g. **SWDE** is model/problem combination in which width, direction and end-state facts are removed from the initial state and must be sensed.
—Blinded facts: these agents are given problems in which facts are removed from the problem description and they not able to sense them during execution. These are our **B** systems. The letters after the **B** refer to the type of facts they cannot sense, e.g. **BW** is model/problem combination in which width facts are removed from the initial state and cannot be sensed.

These combinations produced the following types of behaviour (figure 3). The **P** system was able to successfully obtain the nut in every configuration, and the unconstrained access to information allowed it to do so always using the shortest possible plan. As the **S** system had access to all information via sensing, these models were also ultimately successful in all configurations (as they would never choose to lose the nut after sensing a large backwards trap). However, MAPSIM is an optimistic planner, meaning that it assumes that it will always get the sensing results that will allow it to execute the shortest possible plan. In cases where the sensing results are not as the agent expects, it must replan. This happened to the **S** agents when the optimistic plan did not reflect reality, e.g. when a large trap proved to be in the backward-facing configuration rather than in the optimistically assumed forward-facing configuration. In these cases, the **S** agent would replan, changing direction of nut movement as appropriate (see electronic supplementary material, figure S2). In other words, just like real orangutans, the **S** agents could make initial errors in choosing the direction in which to push the nut, unlike the **P** system. The **B** agents behaved in the same optimistic manner as the **S** agents, but were unable to use sensing to verify that their expectations were correct. In the cases where this meant that they pushed the nut over a backward-facing trap, they would fail on the task.

## (d) *Comparing output with results from real animals*

Once the experimental design is finalized and animals have been tested, the generated models can be used to assist in analysing the results, both quantitatively and qualitatively. If one of the models (or a group of similar models) tends to fit the results shown by the subjects better than others, it should allow researchers some insight into the knowledge possessed or gained by the subjects, and perhaps potential mechanisms involved. If there is inter-individual variability in performance, the models that best fit each subject's performance can be compared to try to determine the sources of inter-individual variability. Interestingly, Brown *et al.* [25] used a similar rationale 35 years ago to model subtraction errors made by school children doing arithmetic. In addition to examining the number of trials that lead to success and failure and comparing with actual results, one can also examine the pattern of behaviour within a trial. For example, do animals (like models that support replanning) always correct their action at a certain point? If we have constructed a model in such a way that it can only sense a particular feature (the width of a gap in the floor of the apparatus, for example) when the target object is moved close to that feature, the fact that both the model and the animals tend to correct their action immediately after they have moved the object close to the gap might suggest that the animals share this constraint. Caution is obviously needed in drawing these kinds of conclusions through analogy alone, but such observations allow us to generate new hypotheses to be tested.

The grid in figure 3 compares the pattern of correct and incorrect initial choices for the three orangutans tested in the original paper [10] along with some of the previously simulated outcomes presented in that paper, and with the new models outlined in this study. Each row of the grid identifies a particular configuration of the puzzle tube, with each column representing one

Table 1. Percentage of trials on which the initial directional choice of the orangutan was the same as that of the model, compared on a trial-by-trial basis for a selection of the models. The results from two models are given on one row where the pattern of output from both models was identical (e.g. BW/BWD). The coefficient of agreement was also calculated for all pairs of models using Cohen's kappa. None of the coefficients suggested a significant match except, *$p < 0.05$ for a positive match, **$p < 0.05$ for a negative match (i.e. the opposite pattern).

| model | Amos | Sandy | Silvia |
|---|---|---|---|
| blind to direction of trap (BD) | 40.63 | 56.25 | 46.88 |
| blind to end state (BE) | 50.00 | 56.25 | 53.13 |
| blind to width of gap or width and direction (BW/BWD) | 40.63 | 56.25 | 46.88 |
| senses direction or direction and end state (SD/SDE) | 65.63 | 50.00 | 56.25* |
| senses width and direction or width, direction and end state (SWD/SWDE) | 59.38 | 56.25 | 40.63 |
| senses width and end state (SWE) | 64.06 | 60.94 | 42.19 |
| senses width (SW) | 67.19 | 64.06 | 39.06 |
| always go left (L) | 45.31 | 60.94 | 35.94** |
| always go right (R) | 54.69 | 39.06 | 64.06* |
| avoid large gap (ALG) | 50.00 | 43.75 | 46.88 |
| choose closed end/avoid large gap (CEALG) | 29.69 | 26.56** | 51.56 |
| choose closed end/choose large gap (CELG) | 29.69 | 35.94 | 60.94 |
| choose large gap (LG) | 50.00 | 56.25 | 53.13 |
| choose open end/avoid large gap (OEALG) | 70.31 | 64.06 | 39.06 |
| choose open end/choose large gap (OELG) | 70.31 | 73.44* | 48.44 |

agent (orangutan or model). Thus, if two columns display matching colours for a particular row of the grid, both agents made the same initial directional choice for that configuration. Similar grids analysing instances of switching of direction and number of moves are available as electronic supplementary material, figures S2 and S3.

We also calculated percentage agreement between the orangutan results and a selection of the models that appeared to fit best from initial inspection of figure 3 (table 1). The percentage agreement was calculated by summing the number of rows on which a given pair of agents matched their initial directional choice and then dividing by the total number of trials. This figure is easy to appreciate intuitively, but it does not indicate how likely it is that this level of agreement would have occurred by chance, because this depends on the base rate of responses by each agent. Thus, we also calculated Cohen's kappa (more commonly used to estimate inter-observer agreement on scores) to derive a *p*-value for the agreement. Only three of the 45 comparisons yielded agreement significantly greater than would be expected by chance: Sandy and the open-end-large-gap simulation (OELG), and Silvia and the move-right simulation (R), and senses direction or direction and end state (SD/SDE). The inverse of the first two of these simulations (closed-end-avoid-large-gap CEALG, and move-left L) both showed a significant

negative match, as would be expected. However, if the appropriate Bonferroni correction is applied for 45 comparisons, alpha reduces to 0.001, and none of the comparisons approach significance.

Taking the patterns in the figure and the percentage agreement scores together, we can see a number of trends. It is clear that none of the blinded facts (B) models fit the pattern of initial choices of the orangutans well. Those blind to end state and either width or direction of the traps (BDE, BWDE and BWE) fail completely because they do not have (and cannot gain, unlike the sensing models) enough information to construct a plan. The model blind to end state (BE, 50–56.25% agreement) seems to have approximately the opposite pattern to that shown by the orangutans. For example, the orangutans (except Silvia, because of her right-side bias) performed well on the range of configurations in which the nut is obtained from the open left end (OCLE), where BE consistently made incorrect initial choices. On the configurations with both ends closed where the nut is obtained via the trap (CCRT and CCLT), most of the orangutans initially chose incorrectly on many of the trials, whereas BE consistently made correct choices. The models blind to width or to both width and direction of the trap (BW, BWD) showed a different pattern, but do not have a greater percentage agreement than BE (40.63–56.25%).

This confirms our intuition (based on observing the behaviour of the subjects during our experiments) that orangutans might find it easier to perceive the state of the ends of the tube and the subsequent consequences than they would to perceive and understand the consequences of the width or direction of the traps. It also broadly agrees with the findings in Tecwyn *et al.* [10] that the simulated models involving the configuration of the open ends (OEALG and OELG in figure 3) fitted the behaviour of Amos and Sandy better quantitatively (percentage agreement ranges from 64.06 per cent to 73.44 per cent for Amos and Sandy compared with OEALG and OELG) than the alternative simulated models.

The sensing facts (S models) show more interesting patterns of initially correct choices. SE has perfect information (i.e. it has knowledge of width and direction of traps from the world, and can sense end direction, and is thus equivalent to P and can be discounted. The models that can sense the direction of the traps but have information about the width (SD, SDE) have a similar lack of fit in the range of CCRT/CCLT tube configurations on which the orangutans often make errors (50–65.63% agreement). However, the models that can sense width (SW, SWD, SWDE, SWE) get closest (out of the models generated in this paper) to matching the pattern shown by the orangutans. SW shows the highest percentage agreement for the planner-based models for Amos and Sandy (64.06–67.19%), but does not fit Sandy's behaviour well (39.06%). This is interesting because we included these models to test a hypothesis that the orangutans might not be able to accurately determine whether a gap would allow the nut to pass through unless the nut was immediately adjacent to that gap. The fit between these models and the real

behaviour is not precise; so we have clearly not quite captured the complexity of the problem. However, it is a good example of how the process of specifying the domain model using hypotheses about real animals can allow us insights that would otherwise be difficult to grasp, particularly with a more complex domain than this puzzle tube.

### (e) *Iterate and refine hypotheses*
Modelling is intended to be an iterative process. It is possible that none of the models constructed before the experiment adequately capture what the subjects do in the real experiment. In this case, the models could be refined, made more complex or more detailed, and compared against the results again in an iterative process. The important principle here is that increasing the complexity of the model should only be carried out if it fails to fit the behaviour, so that the simplest possible model that adequately captures the results (quantitatively and qualitatively) is accepted.

## 5. CONCLUSIONS
We are far from being the first to use AI-inspired models to try to understand animal behaviour (see Webb [26] for a detailed review). Previous applications of AI-inspired techniques and models range from constructing decision trees to understand the complexity of the hierarchical food-processing behaviour of mountain gorillas (*Gorilla gorilla beringei* [27]), to constructing detailed, neurobiologically based models of cricket phonotaxis [28]. Bayesian modelling is increasingly popular with researchers trying to understand the development of cognition in humans, and in particular, how children are able to make inferences on the basis of sparse information when learning languages (see earlier studies [8,29] for recent reviews). Others are even re-evaluating Pavlovian conditioning using Bayesian models [30].

The AI-inspired models mentioned earlier were all focused on attempting to understand the *mechanisms* underlying the behaviour of organisms. However, our proposal for using AI techniques outlined in this study had a slightly different motivation. The problem we face in studying cognition in animals is that we do not yet understand enough about the *problems* animals face and how they might solve them to assign mechanisms to particular cognitive abilities. Thus, rather than using models to directly test or propose candidates for biological mechanisms, we argue that such techniques can be used as part of a design-based approach [13] for systematically analysing and—eventually—understanding the problems animals face. Once we understand the structure of the problem, we can use the same tools to explore systematically the space of possible solutions, again without needing to assign biological mechanisms to those solutions before they are properly understood.

It also encourages taking an information-centred approach, in which one can consider to what information the agent has access and how it can obtain that information. If preliminary modelling suggests that having information about certain features or components can substantially alter performance on the task, the experimental design can be altered to systematically provide or withhold that information. For example, in the puzzle tube experiment, our initial modelling suggested that information about the direction of traps or the width of the gaps was important. We could therefore have re-designed the apparatus in such a way that an orangutan had to remove a barrier to observe the state of the traps and/or gaps, allowing us to control and monitor how information was gained, and how that might alter performance. Of course, as animal cognition researchers, we already try to consider these issues, but it can be difficult to anticipate them all in advance of designing and running the experiment without using these kinds of techniques.

Another advantage of the techniques outlined in this paper is that it provides an alternative method of analysing the resulting experimental data in more qualitative ways. This might allow us to design more complex experiments with a larger range of alternative actions open to the subject, and multiple routes by which the goal can be reached. Indeed, this kind of experiment would be better suited to the planner we used in this study, as there would be a greater range of ways in which it could solve the problem, generating more variation. Thus experiments can be richer, and we avoid the frustration of having to categorise subjects' responses as either 'solving' or 'not solving' the task. In addition, we might get closer to capturing the way in which animals are using their cognitive abilities in the wild. Importantly, the combination of de-constructing the problem to its component parts and being able to analyse responses in more qualitative ways might also help us to compare cognitive abilities between species, by providing a more abstract formulation of the problem and any patterns apparent in the possible solutions.

Given the limited space available here, we could only begin to explore the possibilities of modelling the puzzle tube problem, as an example of how AI techniques might be used. However, there are a number of obvious ways in which it might be extended, which might require different tools. The variant of MAPSIM that we used did not allow for any stochasticity or uncertainty in the agent, and we did not use any information about the order of trials or any effects of experience. A model that allowed an agent to gain knowledge about events after experiencing them (i.e. learning) would add further richness and biological plausibility. However, we hope that we have provided enough detail to show the potential of this approach, and stimulated animal researchers to learn more about AI themselves, or initiate collaborations with AI researchers. Indeed, even if researchers are not interested in taking this approach themselves, we would strongly encourage them to include detailed trial-by-trial results for the training and testing phases for all of their subjects (even those 'failing' a particular stage) in their papers, as this would provide an invaluable resource for those interested in exploring these kinds of modelling problems on a large and varied set of studies. Such data could be provided within the published paper itself, the supplementary information or in data repositories. This practice is also advocated by Seed *et al.* [4] and Thornton & Lukas [6], and allows the research

community as a whole to make the most out of precious data that is time-consuming and difficult to collect.

## ENDNOTE
[1]This ignores computationally identical performance produced by different domain model/problem combinations.

## REFERENCES

1 Byrne, R. W. & Bates, L. A. 2011 Cognition in the wild: exploring animal minds with observational evidence. *Biol. Lett.* **7**, 619–622. (doi:10.1098/rsbl.2011.0352)

2 Seed, A., Emery, N. & Clayton, N. 2009 Intelligence in corvids and apes: a case of convergent evolution? *Ethology* **115**, 401–420. (doi:10.1111/j.1439-0310.2009.01644.x)

3 Chittka, L. & Jensen, K. 2011 Animal cognition: concepts from apes to bees. *Curr. Biol.* **21**, R116–R119. (doi:10.1016/j.cub.2010.12.045)

4 Seed, A. M., Seddon, E., Greene, B. & Call, J. 2012 Chimpanzee 'folk physics': bringing failures into focus. *Phil. Trans. R. Soc. B* **367**, 2743–2752. (doi:10.1098/rstb.2012.0222)

5 Kacelnik, A., Chappell, J., Weir, A. & Kenward, B. 2006 Cognitive adaptations for tool-related behaviour in New Caledonian crows. In *Comparative cognition: experimental explorations of animal intelligence* (eds E. Wasserman & T. Zentall), pp. 515–528. Oxford, UK: Oxford University Press.

6 Thornton, A. & Lukas, D. 2012 Individual variation in cognitive performance: developmental and evolutionary perspectives. *Phil. Trans. R. Soc. B* **367**, 2773–2783. (doi:10.1098/rstb.2012.0214)

7 Chappell, J. 2006 Avian cognition: understanding tool use. *Curr. Biol.* **16**, R244–R245. (doi:10.1016/j.cub.2006.03.019)

8 Perfors, A., Tenenbaum, J. B., Griffiths, T. L. & Xu, F. 2011 A tutorial introduction to Bayesian models of cognitive development. *Cognition* **120**, 302–321. (doi:10.1016/j.cognition.2010.11.015)

9 Chittka, L., Rossiter, S. J., Skorupski, P. & Fernando, C. 2012 What is comparable in comparative cognition? *Phil. Trans. R. Soc. B* **367**, 2677–2685. (doi:10.1098/rstb.2012.0215)

10 Tecwyn, E. C., Thorpe, S. K. S. & Chappell, J. 2012 What cognitive strategies do orangutans (*Pongo pygmaeus*) use to solve a trial-unique puzzle-tube task incorporating multiple obstacles? *Anim. Cogn.* **15**, 121–133. (doi:10.1007/s10071-011-0438-x)

11 Seed, A. M., Tebbich, S., Emery, N. J. & Clayton, N. S. 2006 Investigating physical cognition in rooks *Corvus frugilegus*. *Curr. Biol.* **16**, 697–701. (doi:10.1016/j.cub.2006.02.066)

12 Daw, N. & Doya, K. 2006 The computational neurobiology of learning and reward. *Curr. Opin. Neurobiol.* **16**, 199–204. (doi:10.1016/j.conb.2006.03.006)

13 Sloman, A. 2005 The design-based approach to the study of mind (in humans, other animals, and machines), including the study of behaviour involving mental processes. See http://cs.bham.ac.uk/research/projects/cogaff/misc/design-based-approach.html.

14 Russell, S. J. & Norvig, P. 2010 *Artificial intelligence: a modern approach*, 3rd edn. Upper Saddle River, NJ: Pearson Education.

15 Shanahan, M. 2012 The brain's connective core and its role in animal cognition. *Phil. Trans. R. Soc. B* **367**, 2704–2714. (doi:10.1098/rstb.2012.0128)

16 Winograd, T. 1971 Procedures as a representation for data in a computer program for understanding natural language. MIT AI Technical Report 235.

17 Göbelbecker, M., Gretton, C. & Dearden, R. 2011 A switching planner for combined task and observation planning. In *Proc. 25th AAAI Conf. Artificial Intelligence (AAAI 2011), San Francisco, CA, 7–11 August 2011*, pp. 964–970. Menlo Park, CA: AAAI Press.

18 Hoffmann, J. 2001 FF: the fast-forward planning system. *AI Mag.* **22**, 57–62.

19 Fikes, R. E. & Nilsson, N. J. 1971 STRIPS: a new approach to the application of theorem proving to problem solving. *Artif. Intell.* **2**, 189–208. (doi:10.1016/0004-3702(71)90010-5)

20 Brenner, M. & Nebel, B. 2009 Continual planning and acting in dynamic multiagent environments. *Auton. Agents Multi-Agent Syst.* **19**, 297–331. (doi:10.1007/s10458-009-9081-1)

21 Kaelbling, L. P., Littman, M. L. & Cassandra, A. R. 1998 Planning and acting in partially observable stochastic domains. *Artif. Intell.* **101**, 99–134. (doi:10.1016/S0004-3702(98)00023-X)

22 Wyatt, J. 2005 Reinforcement learning: a brief overview. In *Foundations of learning classifier systems* (eds L Bull & T Kovacs), pp. 179–202. Berlin, Germany: Springer.

23 Bryson, J. J. 2003 The behavior-oriented design of modular agent intelligence. In *Agent technologies, infrastructures, tools, and applications for e-services* (eds R. M. Kowalszyk, J. P. M Müller, H. Tianfield & R. Unland), pp. 61–76. Berlin, Germany: Springer.

24 Brooks, R. A. 1986 A robust layered control system for a mobile robot. *IEEE J. Robot. Autom.* **2**, 14–23. (doi:10.1109/JRA.1986.1087032)

25 Brown, J. S., Burton, R. R. & Larkin, K. M. 1977 Representing and using procedural bugs for educational purposes. In *Proc. the 1977 Annual Conf. ACM' 77*, pp. 247–255. New York, NY: ACM.

26 Webb, B. 2001 Can robots make good models of biological behaviour? *Behav. Brain Sci.* **24**, 1033–1050. (doi:10.1017/S0140525X01000127)

27 Byrne, R. W., Corp, N. & Byrne, J. M. E. 2001 Estimating the complexity of animal behaviour: how mountain gorillas eat thistles. *Behaviour* **138**, 525–557. (doi:10.1163/156853901750382142)

28 Webb, B. & Scutt, T. 2000 A simple latency-dependent spiking-neuron model of cricket phonotaxis. *Biol. Cybern.* **82**, 247–269. (doi:10.1007/s004220050024)

29 Chater, N., Tenenbaum, J. B. & Yuille, A. 2006 Probabilistic models of cognition: conceptual foundations. *Trends Cogn. Sci.* **10**, 287–291. (doi:10.1016/j.tics.2006.05.007)

30 Courville, A. C., Daw, N. D. & Touretzky, D. S. 2006 Bayesian theories of conditioning in a changing world. *Trends Cogn. Sci.* **10**, 294–300. (doi:10.1016/j.tics.2006.05.004)