

## Exploring cognitive machines—Neural models of reasoning, illustrated through the two-sticks paradigm

John G. Taylor, Matthew Hartley\*

Department of Mathematics, King's College London, London WC1R 2LS, UK

### ARTICLE INFO

Available online 10 May 2008

#### Keywords:

Reasoning  
Cognition  
Cognitive machines  
Animal reasoning

### ABSTRACT

Cognition is a complex and wide-ranging field of study. We suggest a possible approach to modelling reasoning as a particular aspect of cognition involving the use of forward and inverse internal models to allow the effect of actions on the world to be considered. The field of animal reasoning provides an excellent source of manageable problems. We describe one of these, the two-sticks paradigm, in detail and construct a model of how the reasoning process involved in solving this paradigm might operate. Results from this simulation are presented, and we discuss important features of the model, in particular the use of spatially invariant goal representations and fast movements of reward.

© 2008 Elsevier B.V. All rights reserved.

### 1. Introduction

Understanding the mechanisms underlying human cognition is an extremely complex problem, constructing models of this cognition even more difficult. One possible way to approach the construction of cognitive machines is to examine cognitive processes in animals, where simpler paradigms exist compared to human cognition.

#### 1.1. Defining cognition

Typical definitions of cognition are:

That operation of the mind by which one becomes aware of objects of thought or perception; it includes all aspects of perceiving, thinking, and remembering.

Mental functions such as the ability to think, reason, and remember.

High level functions carried out by the human brain, including comprehension and use of speech, visual perception and construction, calculation ability, attention (information processing), memory, and executive functions such as planning, problem-solving, and self-monitoring.

All of these definitions indicate that cognition is complex. We wish to concentrate here on reasoning, and planning as advanced components of cognition; we leave out speech as not to be

discussed at any deep level, nor consciousness/awareness. We should note here, however, that consciousness has been proposed as a feature arising out of working memory (WM) activations marshalled by attention, for example as through the CODAM approach [14].

Various approaches have been developed for reasoning and planning. Among them we single out three that have recently proved important:

- Symbolic, using logical inference embedded in linguistic/symbolic structures.
- Probabilistic, using cognition as defined as probabilistic inference.
- Connectionist, with numerous discussions of how inference can be obtained from low-level neural network structures.

We will discuss here only the third of the above approaches since it fits in most naturally to the neural systems relevant to the consideration of animal cognition, more specifically animal reasoning. Cues on higher cognition from brain-based systems possessed by animals could be important in helping us better understand how such feats are achieved in humans.

#### 1.2. Components of modelling cognition

The most important components of our approach to cognition will be based on (1) forward models, to be used as a basis of encoding the causality of the world; (2) WM modules, for imagining future events when combined with forward models (this process is termed “prospection” in [5]), (3) attention control, enabling the selection of only one of numerous distracters in

\* Corresponding author. Tel.: +44 2085300441.

E-mail address: [mhartley@cantab.net](mailto:mhartley@cantab.net) (M. Hartley).

lower level cortices to be evaluated or transformed to enable certain goals to be attained. These components have already been included in the CODAM model [13], although we must be careful to differentiate between a state estimator and a forward model. Here we need to do so: a state estimator is one that estimates the state of the plant being controlled at a given time, whereas a forward model makes a prediction of a future state of the plant. Such a prediction can arise from building a state estimator not for the state now but for that 1 s (or whenever, after recurrent running of the predictor) ahead. This would therefore require an efference copy of the control signal to update the state estimate to that for the next time step.

That such forward models occur in the brain in the parallel case of the motor control system has been demonstrated experimentally by numerous authors. For example in [1] it has been shown how adaptation to novel force fields by humans is only explicable in terms of both an inverse controller and a learnable forward model. More recent work has proposed methods by which such forward models can be used in planning (where motor action is inhibited during the running of the forward model) or in developing a model of the actions of another person [9]. Such planning has been analysed in those references and numerous other publications for motor control and actions, but not for more general thinking, especially including reasoning. Nor has the increasingly extensive literature on imagining motor actions been appealed to: it is important to incorporate how motor actions are imagined as taking place on imagined objects, so as to reason as to what will be optimally rewarded possible actions.

The schemata of the PFC are intrinsically forward models of sensory/response sequences: starting at the initial state of the sequence, the further ones will be generated sequentially, using if necessary further buffer capacity to become aware of the various states of the sequence, or the final one. Thinking through the consequences of an action on a stimulus could thus be achieved in that way, by running through a given schemata from an initial state. Planning would require some final goal state being active, and comparison with the various states in a sequence generated in the above manner made to check if the goal was yet reached. In the process all response patterns would be inhibited from making any real actions (as observed in over-activity of striatum when imagining motor actions). Various strategies for good planning (such as back-tracking, etc.) would need to be used; however, these appear only to be second order features of the basic architecture, so will be presently neglected. This overall architecture provides, then, a principled approach to reasoning and planning in neural systems (and it is suggested the approach in the brain). We add finally that the forward models provided by the ACTION-network style of the PFC may be an extension of more time-limited forward models in parietal cortex, as proposed in terms of parietal deficits.

It has to be recognised that there is presently some controversy over the level of cognitive powers possessed by animals [2]. It is in particular as to the “understanding” of a particular tool relevant to a reasoning task that is at especially at issue. Numerous animals (such as macaques) when challenged, in what looks like a reasoning task, by making the task more difficult, tend to fail the harder task. This was so in the “trapping tube” task: a hole is placed on the underside of a transparent tube, and the animal has to push a reward inside the tube out of one end and avoid the trapping hole. This can be done by a macaque if the hole is at the centre, but it was found difficult if the hole was off-centre. This implied that the macaque was very likely using some simple rule to gain the reward (such as “push a stick into the end of the tube farthest away from the reward”); the more difficult off-centre hole task needed a more complex rule. The difficulty is thus

that animals may appear to be using rationality in solving tasks but may only be appealing to simple associatively based rules, so not possess any real and flexible understanding of the task and the tool use needed to solve it.

### 1.3. Animal reasoning foundations and Betty the crow

While human cognition is exceedingly complex, we can consider the somewhat simpler problems involved in animal cognition. That an animal reasons, as opposed to solving problems purely by trial and error, is supported by the fact that they have been observed remaining still while assumedly internal processes are going on in their brains as they view apparatus in the environment involved with their obtaining a food reward. A sudden sophisticated move is then made by the animal so that it gains its goal (usually food). There can be a period before an action is made when an animal can be interpreted as “sizing up” the situation set by some difficult task, such as a squirrel pausing and looking in detail at the environment when it is trying to obtain nuts from a nut dispenser placed inaccessibly high above the ground at the top of a long pole that the squirrel ultimately assesses is unassailable. It then turns away.

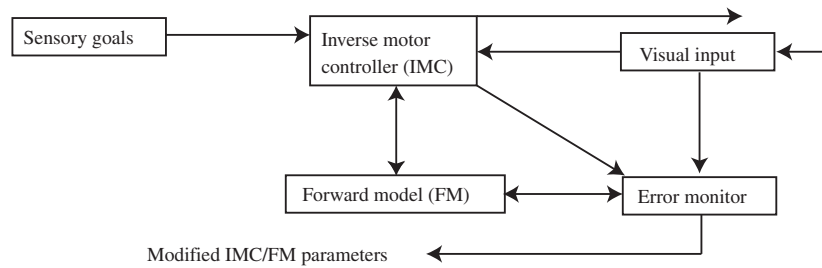
It is possible that such a stationary animal is thinking of other things than how to get the food goal. But it would appear to have poor survival value to sit and think of one's income tax if one is very hungry and can see food, but not immediately obtain the visible food reward. Nor would it have much survival value to rush around, using energy, in random or near-random efforts to get the food. Periods of inactivity are thus, we suggest, not spent in day-dreaming or thinking of unrelated topics by the animal but are of interest for further probing as to possible ongoing brain processes relevant to the animal achieving its task: obtain the food. Even in a more complex animal such as a human being, functional brain imaging, for example, is crucially based on the fact that humans are indeed thinking about relevant strategies for solving the tasks set them by the experimenter rather than subjects letting their minds wander into areas completely irrelevant to the task.

One of the clearest demonstrations of animal reasoning is that of Betty the New Caledonian crow (who lived in Oxford: [16]). She was able to extract her lunch-basket from a vertical transparent tube when the apparatus was presented to her. She was always presented with both a straight piece of wire and a bent one. According to report [16] she always picked up in her beak the bent wire. Such a choice very likely had a genetic basis, since she and all other New Caledonian crows come from a forest environment in which bent sticks are prevalent, and young naive (cage-reared) crows, who had never had any prior experience or coaching in stick use, were noted to prefer bent sticks to straight ones.

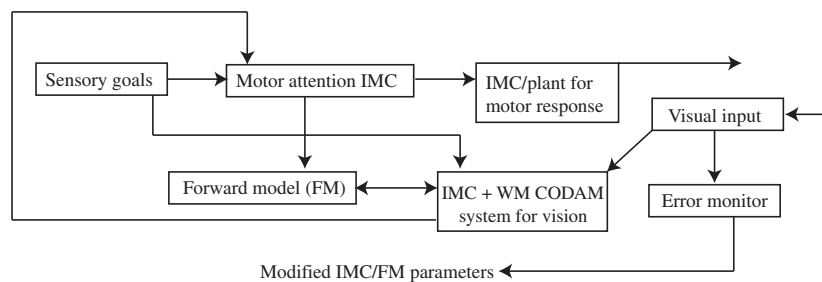
## 2. An architecture for reasoning

A general form of a reasoning system (extending [9]) is given in Fig. 1: This shows the visual input (from parietal lobe) proceeding to an inverse model controller (IMC) to generate a suitable action to achieve the desired (goal) from the visual state arising from the visual input. There is a corollary discharge of the IMC motor control signal (whose main output will generate motor responses in motor cortex/cerebellum/spinal cord) fed to a forward model, and so allow a prediction to be made of the next sensory (visual) input. This process can be speeded up by use of prefrontal chunking of sequences of actions (such as from the initial position of the bent stick at the top of the tube to its touching the handle of the food bucket at the bottom of the tube).

There is no attention control system present in the system of Fig. 1, so leaving the system vulnerable to distracters. This can be



**Fig. 1.** Basic architecture for reasoning. This is a standard motor reasoning system, where the input to the inverse model controller (IMC) in the recurrent mode is solely that arising from the FM from a previous step, with no input being used from the visual input module. The output (the dashed arrow) of the error monitor (which compares visual input to that predicted by the FM) can be used to train the FM and IMC. The output from the IMC is sent not only to the FM (as an efference copy) but mainly to the motor system (motor cortex, cerebellum, basal ganglia, spinal chord) to move the muscles, and thence cause action on the stimulus generating the visual input.



**Fig. 2.** Attention-controlled forward/inverse model Paris for reasoning. Note that the visual input to the forward model is assumed to be under the control of the attended visual output, as from the visual working memory (WM), and the FM feeds back to update the WM, as occurs in mental imagery involving transformations of imagined objects. The goals module has been labelled .visual., to emphasise that it involves the final visual states of objects to be manipulated. See text for further explanation.

remedied by including a CODAM form of attention control, as shown in Fig. 2.

The extension from Fig. 1 to 2 is by the addition of two further modules. One of these is a motor attention IMC module, applying an attention signal to the motor .plant. (consisting of the IMC for motor responses, supposedly sited in motor CX, Cb, BG, etc.) to pick out the desired action signal; this motor attention IMC is known present in the left angular gyrus [10,11]. It uses as input the goal position and the actual visual input from the WM (visual) module in the visual CODAM system. Thus only the attended visual state is used to guide the motor attention IMC, and is fed back from it to update it.

The other extra module is a set of CODAM-like modules for visual attention, consisting of a visual attention IMC, a WM (visual), a WM (corollary discharge) and an error monitor; this latter is different from the explicit monitor in Fig. 2 imported from Fig. 1, which is purely for training the sensory FM and the motor attention IMC. Thus we have a full CODAM-type model for visual processing (so as to avoid visual distracters by the full use of attention, and for which there is experimental evidence for each component) whilst we have taken only a ballistic attention model for motor attention (although there may also be the CODAM-type extensions, these only complicate an already complicated figure).

We note that the goal module in Figs. 1 and 2 have been taken as that of the final sensory state to be achieved by the movement. This is different from suggestions in [15], for example, where actual motor actions were taken to be represented in prefrontal goal modules. However, there is considerable simplification achieved by considering prefrontal goal states only as desired states of external stimuli, such as objects or components of the body. This choice is consistent with the usage of goal states considered in motor planning in [9] and in motor control in [3], who define the goal as a particular desired state of a given external stimulus. This is also consistent with more general

remarks on goals in motor control as corresponding to effector final states [17]. There is also direct experimental evidence of this assumption [7]. Thus goal states in general are not specific actions themselves, but are in the sensory domain. They generate the requisite actions by use of the control apparatus (motor attention IMC and lower level motor IMC). This is also consistent with coding in the highest areas of PFC which are sensory in content, as compared to motor codes in the lower non-primary motor areas (PMC, SMA, and related areas).

The crucial mode of action of the visual FM is to accept input from the motor attention planner and from the attended visual output of the WM (visual) site, possibly accompanied by the lower level codes for object and feature maps if necessary (these also arise from the CODAM-type visual module in Fig. 2). These two sets of inputs provide an update of the visual activation by the FM, as a new sensory state arrived at by the action input to the FM. The new sensory state then can lead, in combination with the desired state goal in the goal module, to an action produced by the motor IMC that will cause the new visual state to be transformed to the desired goal state if the specific action is taken. The resulting motor action is then, in the reasoning mode, not taken in actuality but is fed back to the FM (visual), to be either used for further updating of the FM (as in continued planning or imagining of a sequence of visual stimuli) or (in the acting mode) to provide a specific external motor action by updating the IMC (motor attention). Also, in the reasoning mode there is associated inhibition of the lower level motor apparatus when the overall motor attention IMC and the visual FM are used in reasoning (imagining sequences of actions on sequences of visual stimuli).

From several lines of evidence the visual FM of Fig. 2 may be regarded as a point of contact between the two CODAM networks, one for motor attention and the other for visual attention, that have been proposed in [15,12]. It has been suggested as being in the posterior parietal lobe and/or cerebellum [3].

We will now consider in more detail another animal reasoning paradigm—the “two-sticks” paradigm.

### 3. The two-sticks paradigm—constructing a model of simple cognition

This is a paradigm employed by experimenters working with chimpanzees [8]. The task involves two sorts of sticks: S1 (short) and S2 (long), one of each being present on a given trial. A chimpanzee in the experiment wants to reach the food, but this is further away than can be reached by use of the stick S1 alone. The food can be reached by using S2, but S2 can only be reached by use of S1 (since the chimpanzee is sequestered in a cage). We can see a cartoon of the paradigm in Fig. 3.

The paradigm is interesting because, although it can be solved by a trial and error method involving attempting to take every reachable object and use it on every other object until the reward is reached, this does not seem to happen in practice. Instead, the chimpanzees can take the necessary reasoning steps to deduce that it is necessary to take S1, use S1 to reach S2 then use S2 to reach the reward, then carry out these actions, without extensive trial and error.

We can consider a possible sequence of mental steps that a reasoner might take to solve the problem.

- (I) Try to reach the button directly. This is: “NOGO”, leading to further search.
- (II) See the stick S1, and try virtually to pick it up and use it. Again this leads to lack of success (since it is not long enough to reach the button), leading to further search.
- (III) The stick S2 is observed, and if it were picked up virtually, it is found that S2 would be able to reach the button. Then S2 becomes rewarded as a subgoal (with a new but temporary value map entry for S2).
- (IV) With S2 as a subgoal, use S1 virtually to reach S2, which is achieved; S1 now becomes a subgoal (the most temporarily rewarded of all the goals).

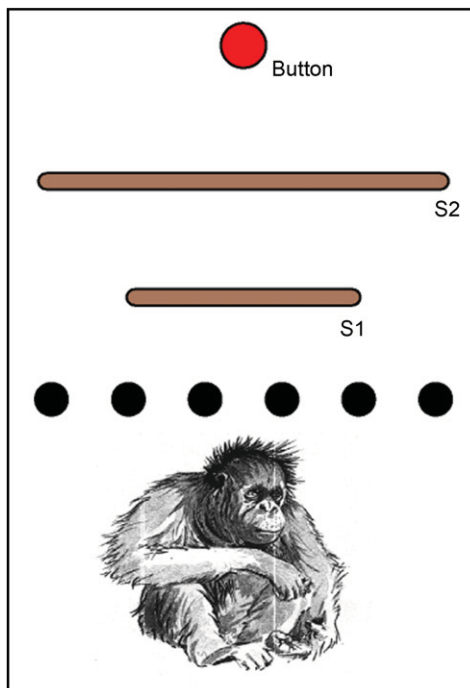


Fig. 3. Cartoon of the two-sticks paradigm.

- (V) Proceed with actual actions and so obtain S1 first, then drag S2 to the reasoner by use of S1, and thence the food reward is obtained by use of S2 on the button.

#### 3.1. Simulation modules

We can consider what modules might be necessary for a model of the reasoning process involved in the two-sticks paradigm, extending our earlier consideration of cognition.

**Drives:** These should consist of the basic drives that cause the system to attempt actions. In the present paradigm's case the sole basic drive is hunger, to be satisfied by pressing the button on a suitably distant wall, that will cause a food reward to be given (the association between the button and the food reward having forming previously at some point).

**Goal list:** This is composed of the goals which are available to the system (independent of the available actions within the simulated world). In our approach goals are represented by stimuli. For this paradigm there are three goals: button/S1/S2, representing attempting to press the button, or grasp the sticks.

**Vision:** The vision system provides the simulation with information about the current state of the world and the reasoner within it. This allows an IMC to calculate movements necessary to achieve selected goals, or to return “NOGO” if they are not achievable.

**Motor IMC:** The IMC components allow the simulation to determine whether goals are achievable or not, given the current state of the world. They also allow the use of “virtual” actions, whereby the results of actions currently physically impossible can be considered. These can then modify internal parameters that determine what actions are and are not possible. For example, the IMC associated with stick S2 allows consideration of what actions could be performed while stick S2 is grasped even if the stick is not currently physically reachable. We discuss these “spatially invariant” representations of objects in further detail later.

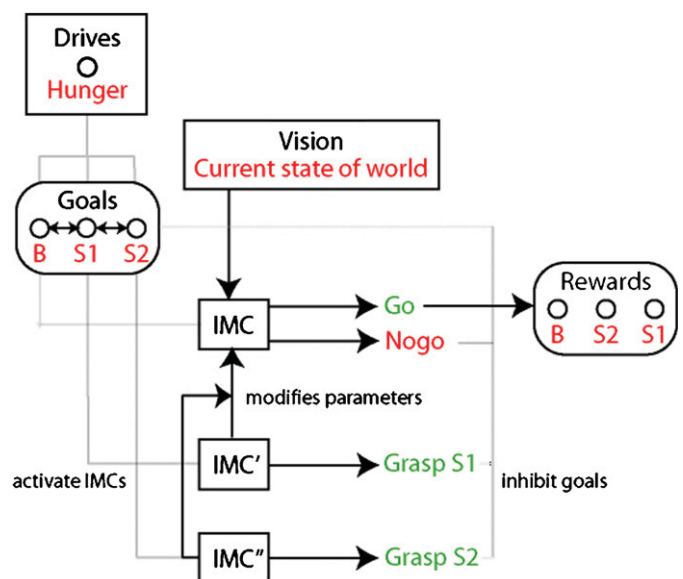


Fig. 4. Details of the simple model for cognitive reasoning in the two-sticks paradigm. The drives module primes the goals (which have a competitive interaction such that only one can be active at the same time). Goal activations prime the IMC modules, which determine whether actions are possible, and affect goals and rewards appropriately. Vision keeps the current state of the world updated.

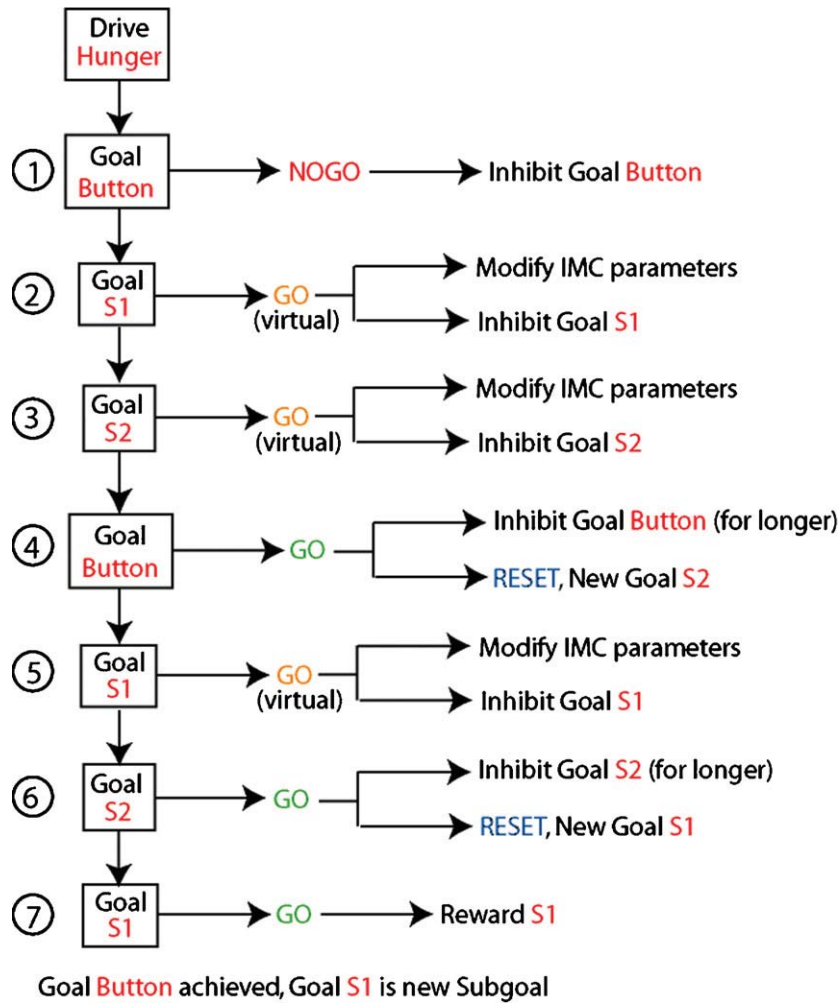


Fig. 5. Suggested reasoning sequence of goal usages, showing the order in which goals are attempted, and the “virtual” actions that modify system parameters to allow the chain of reasoning to be carried through.

*Rewards:* Reward values can be attached to all of the goals available to the system. These reward values can be modified by the results of the mental simulation, and this modification allows the correct actions to be executed to solve the paradigm.

*Forward models:* We have not included these explicitly in the architecture of Fig. 4. We consider that each IMC has an associated FM which is easier to train than the IMC, since the FM has an immediately generated error as a comparison of the predicted next state (after an action) and the actual state as determined by sensors. This error can be used to train both the FM and IMC; it may arise most powerfully from the cerebellum but may also involve a dopamine component in learning. The forward models are needed if extrapolation through a sequence of states and actions is needed in a reasoning process. The processes in the two-sticks paradigm are simple enough to assume that no forward extrapolation is needed, since the actions are assumed taken in one ballistic movement to attempt to touch the button or to grasp a stick.

We can see these components together in the model, illustrated in Fig. 4. There is a goals module, coding for the three stimuli of button, sticks S1 and S2 (to press the button, pick up the stick S1, and pick up the stick S2, respectively). There are three IMC modules, the first (denoted IMC) being for pushing the button by the gripper, the second is for grasping stick S1 (and denoted by IMC', with consequent alteration of the length of the gripper in IMC to correspond to carrying stick S1), and the third (denoted

IMC'') is for performing a similar action with stick S2 (and consequent change of parameters used in the gripper IMC). There is also a reward module in which there is modifiable steady activity corresponding to the current reward value of either the button B or the sticks S1 and S2 (all the observable objects in the environment).

Given these defined goals, we might then expect the flow of reasoning described earlier to occur in terms of the model components as in Fig. 5.

At each stage of the simulation, the system is presented with a range of possible actions and must choose an action to perform. The IMC and FM (the latter not shown in Fig. 2) allow the system both to mentally simulate these actions (with no external actions) or alternatively to instantiate the actions (the former by inhibition of any output from the IMC, the latter when this input is allowed to activate the effectors by switching off the inhibition).

### 3.2. Simulation mechanics

Parameters are given in Table A1.

Goal nodes are modelled as simple graded cells with the equation:

$$C \frac{dV}{dt} = g_{leak}(V - V_{leak}) + I, \tag{1}$$



where  $I$  is the neuron's input current. Their output is sigmoidal in the form

$$\frac{1}{1 + \exp(-V/V_{\text{scale}})} \quad (2)$$

They are mutually interconnected with inhibitory connections of weights  $-0.8$ . Each goal node is driven by an object node from the visual system (with constant current outputs), with weights varying from  $0.8$  to  $1.2$ . The drives module provides additional input to the goals (being connected to each goal with weights from  $0$  to  $0.4$ ). The competition between goal nodes then selects a goal, and when the activity passes a threshold  $T_{\text{goal}}$ , activates an IMC module.

The IMC modules each raise their activity to  $Act_{\text{IMC-NOGO}}$ , when performing a calculation that results in a NOGO signal, and  $Act_{\text{IMC-GO}}$  when the result is a GO signal. Activity decays exponentially for  $0.1$  s, then returns to zero. If an IMC returns a NOGO signal, it provides a negative feedback signal of magnitude  $I_{\text{NOGO}}$  to the goals module for a time equal to  $t_{\text{feedback}}$ , if it returns a GO signal, it provides a negative feedback signal  $I_{\text{GO}}$  until the simulation ends. Whenever a GO signal is produced, the IMC activates the corresponding reward node as described below (each IMC connects to the corresponding reward node).

Reward nodes raise their activity to a value of  $Act_{\text{rwd}}$  when a successful transfer of reward occurs, and their activity decays such that

$$\frac{dA}{dt} = -k_A A, \quad (3)$$

where  $A$  is the activation level of the reward node. Reward node activations essentially provide the “readout” of the system's results, if an instantiated system (such as a robot) were used, these rewards would drive movement to complete the task. It is likely that the real brain has a more complex operation than simply carrying out tasks in order of reward, such that the modified reward values would affect goal sites in the brain (such as dorsolateral prefrontal cortex or parietal lobe), as part of a more complicated system for generating action schema. However, in this very simple case, the list of reward values serves as the output.

### 3.3. Simulation results—goals, IMCs, rewards

The simulation performs as we expect from our suggested reasoning flow. After an initial attempt to press the button, a NOGO result is obtained. After this, but with the hunger drive still activating motor activity, S1 is attempted to be gripped, which having been achieved still fails to allow the button to be reached. After S2 is attempted, however, the button is now reachable. This causes S2 to be rewarded, and moves the goal backward to achieving S2. This requires S1 to be gripped, so once that is obtained, S1 is rewarded and the simulation has correctly rewarded all of the potential goals and hold their relative activities with  $r(S1) > r(S2) > r(\text{button})$ .

We can see the temporal flow of the goal activations in Fig. 6, showing the competition between goals, and how inhibition allows a suitable sequence of goal activations to occur. We can also see the corresponding IMC activations in Fig. 7 and the reward values in Fig. 8. Of particular note is that the final relative valuations of the rewards allows the problem to be solved (we assume that as each action was physically performed, as prompted by the rewards, these rewards would be extinguished in turn).

### 3.4. Further consideration of the model—importance of modules and parameters

To aid in understanding the function of the model, we can consider the importance of the various modules by imagining what functionality would be lost by their removal.

The drives module provides the impetus for the system to attempt tasks. Without its presence there would be no actions carried out, since no goals would be primed. The goals module is critical to the simulation's operation since it provides the list of possible actions, the system could not operate in its absence. The rewards module is the system's output—removing it would not interfere with the process of reasoning, but would not allow any actions based on that reasoning to be taken.

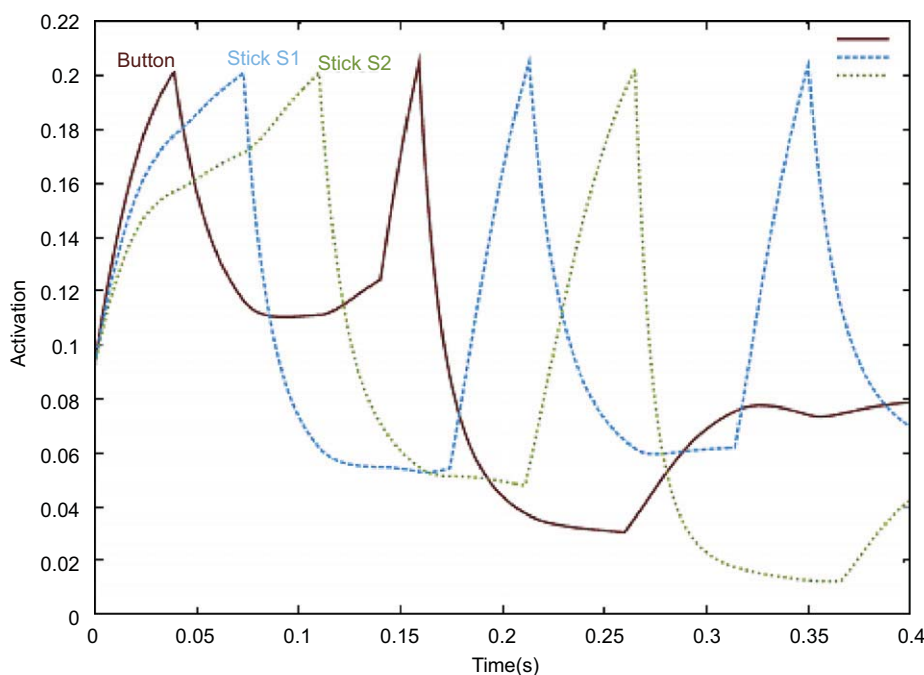


Fig. 6. Temporal flow of goal node activations, showing activation (vertical axis) against time in seconds (horizontal axis).

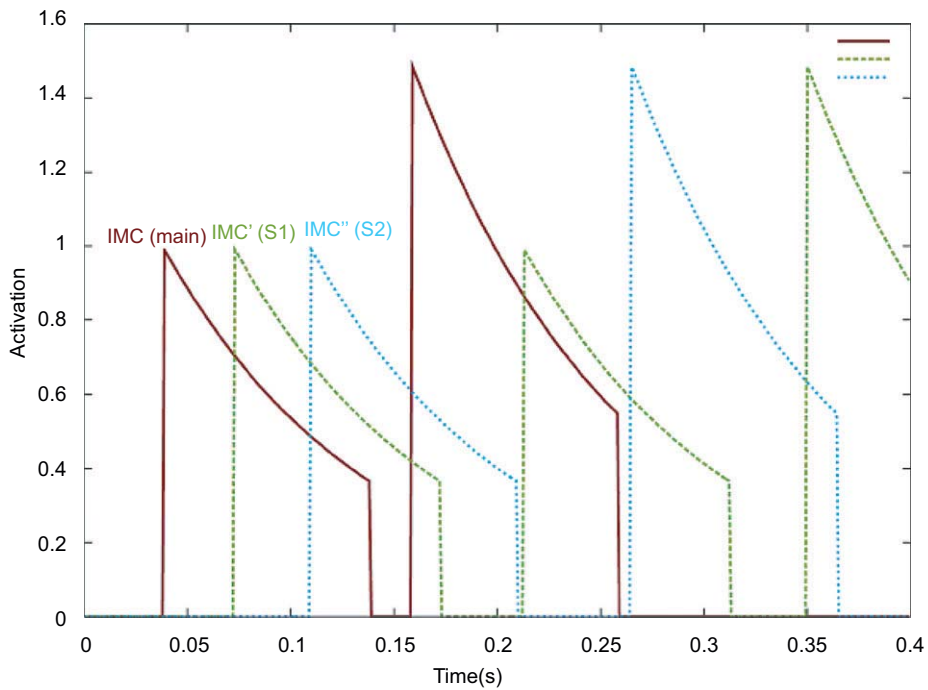


Fig. 7. Temporal flow of IMC activations, showing activation (vertical axis) against time in seconds (horizontal axis).

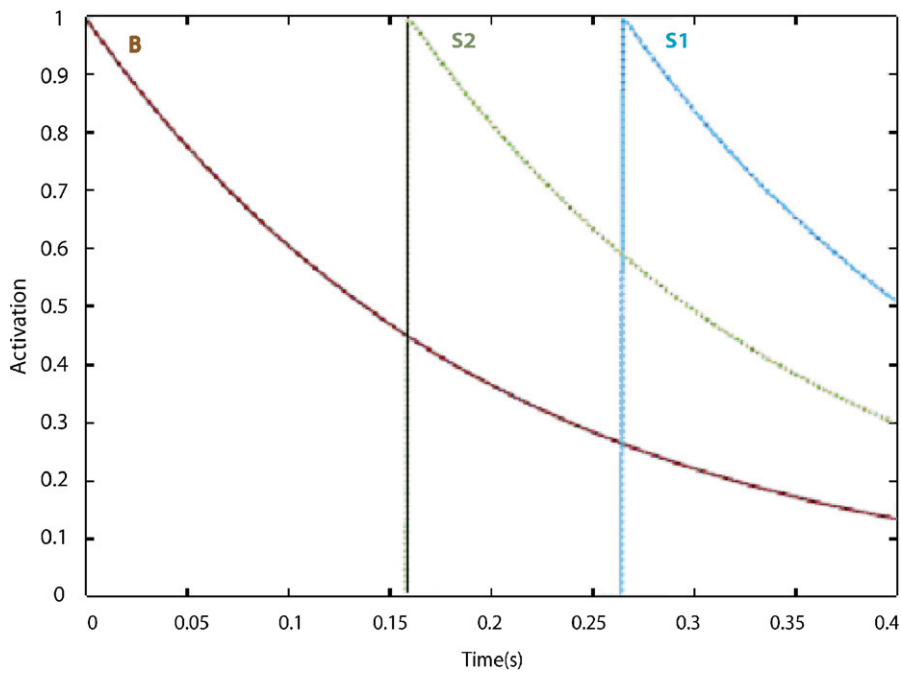


Fig. 8. Temporal flow of reward node activations, showing activation (vertical axis) against time in seconds (horizontal axis).

The model IMCs allow the system to determine whether goals are achievable given the current state. Without the module, it would be impossible to determine which goals were achievable, and thus to reason the correct goal order.

We can also consider the effects of altering the model parameters. The speed of neural response of the goal neurons is controlled by their leak currents and time constants. Altering

these changes the rate at which goals become activated, potentially allowing goals to be more or less rapidly considered. The speed of operation of the IMCs controls how quickly the system is able to consider actions, and their consequences. Altering the reward decay parameters controls the period for which the model output is usable—this is linked to the timescale of reward usage in the brain.

#### 4. Discussion

The results show that the model successfully reproduces the suggested sequence of goal activations in Fig. 5, and produces the correct relative reward values necessary to complete the paradigm successfully. We can also consider some key features of the model that enable it to simulate the reasoning process of the two-sticks paradigm.

##### 4.1. Spatial versus spatially invariant goals

When reasoning, we need to be able to consider the consequences of actions we cannot currently perform. We can imagine manipulation of an object that is currently out of reach in some way, and use this to solve problems. For example, when realising that we need to fetch a set of keys from another room to obtain a goal beyond a locked door, we can imagine unlocking the door without currently possessing the keys.

In our system this is possible through the use of both spatial and non-spatial goal representations. The non-spatial version allows virtual manipulation of the goal objects without the requirement of possessing the object physically, while spatial goals come into play when considering how to achieve the goal itself, or even achieving them virtually by making what would be physically impossible movements such as flying up to the ceiling. Some form of spatially invariant representations are clearly necessary to be able to perform complex reasoning tasks without having to move through a time-consuming process whereby objects must be held before their uses can be considered.

##### 4.2. Movement of reward

In any system whereby intermediate goals must be used to reach an ultimate goal, some mechanism must exist to make the reasoner perform the intermediate goals. We suggest that rapid movement of reward can do this, by attaching these rewards to the subgoals which then gives reason to perform them. Without this movement of reward, some other mechanism must fulfil the same function—otherwise, the original maximally rewarded task (such as pushing the button, in our example paradigm) would be continually attempted, neglecting all subtasks. It is possible that the brain can do this in some other way; however, rewarding these intermediate actions provides a relatively simple mechanism.

##### 4.3. Comparison to existing models

There is some theoretical literature on reasoning in animal reasoning, for example various collected writings in the book of Hurley and Nudds [6], and various reviews of animal reasoning [2]. However, there are few direct simulation models of reasoning processes in animals, models of reasoning tend to be aimed at artificial reasoning systems.

An additional comparison that may be made is to models of learning automata. These are systems with a set of possible actions and probabilities of generating those actions. Feedback from the environment modifies the action probabilities based on teaching data from the environment. These systems bear some similarity to the reasoning model; however, the model's primary method of operation is that of attaching reward values to certain actions based on internal calculations which is somewhat different from permanent variation in probabilities of action generation. Additionally the model is focussed on providing biologically inspired mechanisms to solve the problems of reasoning rather than a permanent learning system.

Thus there are the differences of our model to that of learning automata:

- (a) The latter generally have no separate reward or stimulus-value module.
- (b) Nor have they such reward modules used so as to modify goal activities and change goals.
- (c) Nor do they involve dynamics described by the biologically based neurons present in the modules, as displayed in Eqs. (1) and (2).
- (d) Nor do they possess the faculty of attention, such as described in the general architecture of Section 2, neither of a sensory nor a motor form (as discussed there). However, there are definitions of automata, such as “a self-operating machine” which cover any autonomous system, and so cover the general model we are presenting. Certainly such definitions do not relate to the more specific biological bases of the models we are seeking.

##### 4.4. Model predictions

We have shown how our model solves the reasoning problem presented in [8], we can consider what predictions are made by the model. Particular predictions relate to the already discussed issues of spatial and spatially invariant goals, and movement of reward.

We suggest movement of reward as a mechanism to assist subgoal completion. Recent experiments in distractor devaluation during an attentional blink paradigm [4] have shown that a lasting reduction in value can be attached to certain stimuli (pictures of faces in the paradigm). It may be possible to construct reasoning tasks that could then be used to measure the value attached to components of the task. This would provide an indication as to whether the process of reward movement has lasting effects on the perception of value.

It may also be possible to examine this hypothesis experimentally by observing neurotransmitters and neuromodulators known to be related to reward (such as dopamine) within the brain during reasoning (as in orbito-frontal cortex or amygdala), or by seeing how agonists/antagonists to these neurochemicals affect reasoning.

If reasoning does use both spatial and spatially invariant goals, it is possible that these different types of goal representation might be encoded in different parts of the brain. If so, imaging studies might be able to reveal the use of different representations at different times during the process of reasoning which would provide insight into the temporal flow of reasoning.

#### 5. Conclusions

We have suggested some principles of modelling cognition, and key mental components, particularly the use of forward and inverse models, goals and rewards. Implementing these in a simple neural architecture allows us to demonstrate a possible solution to the two-sticks paradigm from animal reasoning. The model makes particular use of two interesting features—fast movement of reward between goals and the use of both spatial and spatially invariant representations.

Some of these suggested mechanisms may be experimentally testable. Rewards may very well be neurochemical in nature and so administering neuromodulators that relate to possible reward neurochemicals could affect performance in reasoning paradigms. If spatial and spatially invariant representations are separate they may be encoded in different parts of the brain (dorsal and ventral areas for example), and specific imaging studies might be able to



**Table A1**  
Values of constants

Variable name	Value	units
$V_{leak}$	−70	mV
$C$	25	nF
$g_{leak}$	0.025	$\mu$ S
$Act_{IMC-NOGO}$	1	n/a
$Act_{IMC-GO}$	1.5	n/a
$Act_{rwd}$	1	n/a
$k_A$	10	n/a
$I_{GO}$	−50	nA
$I_{NOGO}$	−20	nA
$\tau_{feedback}$	75	ms

reveal these differences. Additionally, behavioural experiments that carefully examine the balance of trial and error, pure reasoning and a mixture of the two might help to reveal both whether the model is accurate, and the limits of internal reasoning capability (Would a three-sticks paradigm be solvable for example?).

This work provides a basic framework for modelling simple cognition—the use of coupled forward and inverse internal models allows predictions about actions on the state of the reasoner and the world to be made, and these predictions are critical for solving problems through internal reasoning rather than trial and error. Development of fuller cognitive machines will require additional understanding of the operation of these internal models, goal representations and rewards, and there is much of interest to be learned here.

## Appendix A. Model parameters

Parameters are given in Table A1.

## References

- [1] N. Bhushan, R. Shadmehr, Computational nature of human adaptive control during learning of reaching movements in force fields, *Biol. Cybern.* 81 (1999) 39–60.
- [2] L.A. Bluff, A.A.S. Weir, C. Rutz, J.H. Wimpenny, A. Kacelnik, Tool-related cognition in new caledonian crows, *Comp. Cognit. Behav. Rev.* 2 (2007) 1–25.
- [3] M. Desmurget, S. Grafton, Forward modeling allowed feedback control for fast reaching movements, *Trends Cognit. Sci.* 4 (2000) 423–431.
- [4] M. Eimer, M. Kiss, Involuntary attentional capture is determined by task set: evidence from event-related brain potentials, *J. Cognit. Neurosci.*, in press.
- [5] N.J. Emery, N.S. Clayton, The mentality of crows: convergent evolution of intelligence in corvids and apes, *Science* 306 (2004) 1903–1907.
- [6] S. Hurley, M. Nudds, *Rationality in Animals?*, Oxford University Press, Oxford, 2006.
- [7] P. Morasso, Spatial control of arm movements, *Exp. Brain Res.* 42 (1981) 223–227.
- [8] N.J. Mulcahy, J. Call, R.I.M. Dunbar, Gorillas and orangutans encode relevant problem features in a tool using task, *J. Comp. Psychol.* 199 (2005) 23–32.
- [9] E. Oztop, D. Wolpert, M. Kawato, Mental state inference using visual control parameters, *Cognit. Brain Res.* 22 (2004) 129–151.
- [10] M.F.S. Rushworth, A. Ellison, V. Walsh, Complementary localization and lateralization of orienting and motor attention, *Nat. Neurosci.* 4 (2001) 656–661.
- [11] M.F.S. Rushworth, H. Johansen-Berg, S.M. Gobel, J.T. Devlin, The left parietal and premotor cortices: motor attention and selection, *Neuroimage* 20 (2003) S89–S100.
- [12] J. Taylor, N. Fragopanagos, Modelling human attention and emotions, in: *Proceedings of the IJCNN, Budapest, 2004*.
- [13] J.G. Taylor, Paying attention to consciousness, *Prog. Neurobiol.* 41 (2003) 305–335.
- [14] J.G. Taylor, Mind and consciousness: towards a final answer?, *Phys. Life Rev.* 2 (2005) 1–45.
- [15] J.G. Taylor, N. Fragopanagos, Simulations of attention control models in sensory and motor paradigms, in: *Proceedings of the International Joint Conference on Neural Networks, IEEE, New York, 2003*.
- [16] A.A.S. Weir, J. Chappell, A. Kacelnik, Shaping of hooks in New Caledonian crows, *Science* 297 (2000) 981–983.
- [17] S.P. Wise, R. Shadmehr, Motor control, in: *Encyclopedia of the Brain*, vol. 3, Elsevier, USA, 2002, pp. 1–21.



**John G. Taylor** is an Emeritus Professor of Mathematics at King's College London. He has an extensive background of research, with over 300 published papers, firstly in theoretical physics and later in wide ranging aspects of neural networks. His current research interests include many ways in which attention operates in the brain, cognitive reasoning and the "relational" theory of consciousness.



**Matthew Hartley** obtained his BA and MSc in Experimental and Theoretical Physics from the University of Cambridge. His PhD from the University of London was entitled "Mechanisms of memory and plasticity in the hippocampus: a simulation approach," and focussed on computational neuroscientific approaches to understanding the hippocampus. Since completion of his PhD, he has worked on projects involving the learning of visual attention and motor skills. His research interests include spike-time-dependent plasticity, the hippocampus and episodic memory, learning by imitation/observation and modelling cognitive reasoning.