

CAUSAL LEARNING BY A ROBOT WITH SEMANTIC-EPISODIC MEMORY IN AN AESOP’S FABLE EXPERIMENT

Ajaz A. Bhat

School of Psychology
University of East Anglia
Norwich, NR47TJ, UK
a.bhat@uea.ac.uk

Vishwanathan Mohan

School of Computer Science and Electronic Engineering
University of Essex
Wivenhoe Park Colchester, CO4 3SQ, UK
vishwanathan.mohan@essex.ac.uk

ABSTRACT

Corvids, apes, and children solve The Crow and The Pitcher task (from Aesop’s Fables) indicating a causal understanding of the task. By cumulatively interacting with different objects, how can cognitive agents abstract the underlying cause-effect relations to predict affordances of novel objects? We address this question by re-enacting the Aesop’s Fable task on a robot and present a) a brain-guided neural model of semantic-episodic memory; with b) four task-agnostic learning rules that compare expectations from recalled past episodes with the current scenario to progressively extract the hidden causal relations. The ensuing robot behaviours illustrate causal learning; and predictions for novel objects converge to Archimedes’ principle, independent of both the objects explored during learning and the order of their cumulative exploration.

1 INTRODUCTION

The ability to learn causal regularities and object affordances allows organisms to exploit objects in the service of their needs and wants, an obvious adaptive advantage in the race for survival. Experiments exploring paradigms like floating peanut task (Hanus et al., 2011), trap-tube task (Martin-Ordas et al., 2008) and more recently the Crow and the Pitcher task from Aesop’s Fable (Jelbert et al., 2014; Cheke et al., 2012) provide evidence that children, primates and corvids can reason to varying degrees about the causally task-relevant properties of objects. A pertinent question therefore is how through cumulative and explorative interactions with different objects in the world, cognitive agents learn task-relevant physical and causal relations and then exploit these flexibly in novel contexts. Accordingly, theoretical works have investigated causal learning in cognitive psychology (Gopnik et al., 2004; Griffiths & Tenenbaum, 2005), computer science (Pearl, 2009; Shimizu et al., 2006), recently in machine learning (Battaglia et al., 2016; Iten et al., 2020; Baradel et al., 2020) but less so in robotics (Xiong et al., 2016). We explore this question in the context of robotics to show how a semantic-episodic memory system along with a set of four learning rules endows a robot, *iCub* with the capability to cumulatively extract and infer causal relations between objects and actions in “the Crow and the Pitcher” task. In designing this model, we connect some major trends from neuroscience on distributed hub-based semantic memory representation (Kiefer & Pulvermüller, 2012; Ralph et al., 2017), small-word properties (Sporns, 2011), episodic memory circuitry (Allen & Fortin, 2013; Lee et al., 2015) to hypothesize possible mechanisms of causal learning.

1.1 CAUSAL LEARNING TASK

The task is inspired from an Aesop’s fable (see analogous empirical works (Jelbert et al., 2014; Bird & Emery, 2009)), in which a thirsty crow drops pebbles into a half-filled water pitcher, raising the water level high enough to drink. In a comparable scenario (Figure 1 A), *iCub* has a set of objects (on a reachable stand) and a jar of water containing a floating target (green ball) in front. With the goal to reach the green ball (which as such is unreachable), *iCub* explores if use of available objects

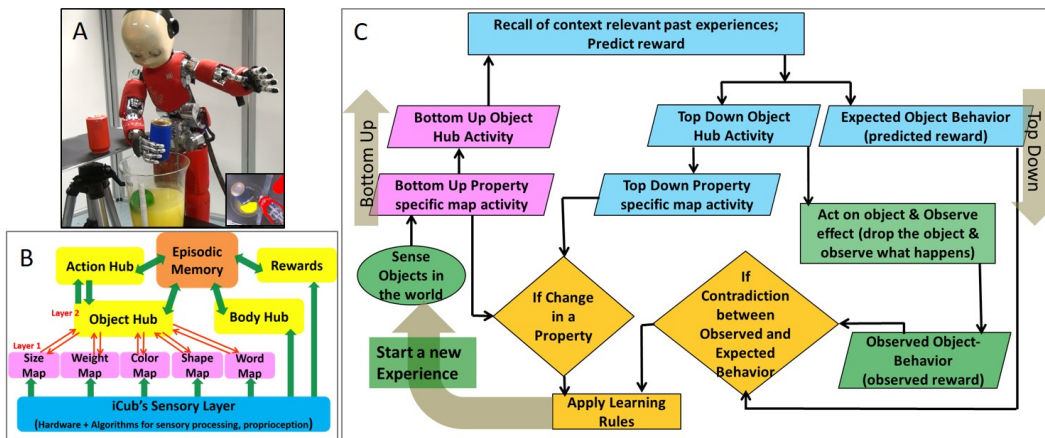


Figure 1: Panel A shows the setup. Each episode involves robot either dropping a single object or making a choice between multiple objects. Objects vary in their physical properties: color, size, shape and weight. Panel B shows the block diagram of the proposed model. Panel C is a flowchart of how the learning rules are applied.

can help realize its (otherwise unrealizable) goal. A priori, there is no previous experience with any of the objects. Hence, iCub knows nothing a priori concerning the causal nature of the task.

2 MODEL DESCRIPTION

Hub-based Semantic Memory: Figure 1 B shows the proposed model. At the bottom is the sensory layer to analyze object properties: colour, shape, size and weight. Word labels are provided by the teacher either to issue goals to iCub or teach names of new objects. Output from sensory/physical layer is passed upwards to a bunch of *property-specific* self-organizing maps (SOMs) that encode object properties as concept features. Neural connectivity between the sensory layer and the property-specific maps is learnt using standard SOM procedure (see Bhat et al. (2016)). There onwards, perceptual analysis of an object (e.g. a *large heavy blue cylinder*) leads to activations in different property-specific SOMs coding for color, shape, weight and size respectively. In this sense, layer 1 emulates the distributed property-specific organization of perceptual information in brain (Patterson et al., 2007; Martin, 2016). Activity in these maps forms the bottom up input to the layer 2 SOM i.e. the object hub. The learning rule for dual dyad (Park & Friston, 2013) connectivity matrix W from SOMs to the object hub (and its counterpart W' backwards) is: if the net activity due to a neuron i and a neuron j winning in the maps manages to activate a neuron k in the object hub, set $W_{ik} = 1$ and $W'_{jk} = 1$. Thereby, the object hub facilitates integration and multimodal object representation, as evidenced neuro-scientifically (Ralph et al., 2017; Kiefer & Pulvermüller, 2012). This small-world network (Sporns, 2011) of maps and the object hub is complemented with a dynamics to let the neural activity in one map retro-activate other members of the network, hence allowing information to move *top down*, *bottom up* or in *cross modal* fashion. So, a word like blue cylinder activates word-map that forwards activations to the object hub, which thereafter retro-activates the shape and colour maps, as if expecting top down what a blue cylinder might be perceptually. In an embodied robot, *objects* in the environment are employed via *actions*. Here, actions are represented at an abstract level (what can be done with an object) separated from the action planning details (how to do). While the former relates to the motor affordance of an object, the latter relates to motion planning details, and interested reader may refer to Bhat et al. (2017) for details related to motion planning. In layer 2, the abstract representation corresponds to the action hub and consists of single neurons coding for different action goals such as reach, grasp, etc. In this sense, action hub neurons are similar to *canonical* neurons in the pre-motor cortex (Rizzolatti & Craighero, 2004) that are activated at the sight of objects to which specific actions are applicable. Finally, the consequences of both perceptions and actions alter the state of the body itself. The body hub, another layer 2 neural map explicitly encodes these states of the body (like failing to reach an object etc.). Reward of an experience is either given by the user or evaluated by the robot itself through observation. In this task, reward is the volume/level of water raised by dropping an object into the jar of water.

Episodic memory and its link to the Hubs: Practically, when a robot interacts with its environment, it is the ongoing sequences of actions on various objects, the ensuing consequences, internal body state, and rewards received that mainly form the content of its experiences. Thus, in our model,

it is the temporal sequence of activations in the different hubs (object, action, reward, body) during an episode that make up the episodic memory content. The episodic memory is realized using a excitatory-inhibitory neural network of auto-associative memory (Mohan et al., 2014; Hopfield, 2008). It consists of 1000 neurons, organized in a sheet like structure with 20 rows each containing 50 neurons. Every row is an event in time (indicating activation in object hub, action hub, body hub or reward) and the complete memory as an episode of experience. For example, being unable to reach a floating object in a jar of water (Body hub state), perceiving a red cylinder (Object hub state), dropping it in water (Action hub state), fetching a reward of 100 (end reward). In future, if the robot perceives the red cylinder, the object hub state serves as a partial cue to reconstruct the full experience. Importantly, in the memory network of 1000 neurons, multiple episodic memories (≈ 230) can be represented and retrieved (Hopfield, 2008). See Bhat et al. (2016) for methods to encode new experiences in episodic memory and recall past ones from partial cues.

Learning Rules for Causal Abstraction: As Figure 1 C shows, these rules compare *what the robot has experienced in the past* against *what is happening in the present situation*. Let $\Delta Property$ be the difference in activity in a property-specific map when activated bottom up (through sensory layer) and when activated top down through recall of past event from episodic memory. Let $\Delta Contradiction$ be the difference between the robots anticipation of how an object might behave (expected reward due to recalled past experience) and the real observed behaviour. Then the learning rules are as follows:

(E)limination rule : If $\Delta Property \wedge \neg \Delta Contradiction$, then that property is not causally dominant and hence drastically reduce the connection strength between the object hub and the associated property-specific map.

(G)rowth rule : If $\Delta Property \wedge \Delta Contradiction$, then that property is causally dominant. Hence connectivity between the object hub and the associated property-specific is strengthened, and encode the experience in episodic memory. Contradiction in the robots anticipation implies that there is something new to learn.

(U)ncertainty rule : If $\neg \Delta Property \wedge \Delta Contradiction$, the connection strength between the object hub and the map is marginally reduced. In this condition it is not possible to infer whether the property is causally dominant or not, unless further experience is gained.

(S)tatus Quo rule : If $\neg \Delta Property \wedge \neg \Delta Contradiction$, nothing new to learn, so no change in connectivity.

3 RESULTS

Given that the robot is learning cumulatively, this section presents successive episodes of learning (a video playlist of these experiments is available *online*). All episodes share a set of computational processes, 1) bottom up experience/interaction with the world and hence activation of various maps 2) recall of past experiences from memory (if any); 3) Use of recalled past experiences to anticipate and 4) application of the learning rules.

Learning that colour of objects is causally irrelevant to the Aesops fable task

Episode 1: In the first episode, iCub is given the goal to reach the green ball (in the jar of water). The motion planning system of iCub provides the information that the goal is unreachable. A *large heavy red cylinder* is available and detected (see Figure 2 left, *Object 1*). Bottom up sensory streams activate property-specific maps related to (red) color, (cylinder) shape, (11.5cm) size and (420g) weight properties leading to a distributed representation of the object in the maps and object hub. The object hub activity leads to generation of a partial cue for the recall of any related past episodes. Since there is no previous experience in the episodic memory, nothing is recalled. So, there is no top down activity in the object hub nor any reward expected. With only option to explore, the robot picks and drops the object into the jar of water. The object sinks in the water displacing a volume of water of about $365cm^3$ enough to make the floating green sphere reachable. This experience is encoded into the episodic memory: an unreachable goal (body hub state), dropping a large heavy red cylinder (object hub activity), a volume of water displaced $365cm^3$ (reward) and goal realized successfully (body hub state). Note, this is a rapid one-shot event encoding into memory. **Episode 2:** iCub is presented with a *blue cylinder* of the same size and weight as one in episode 1. Object hub activity (with partial similarity to *Object 1* generates a partial cue that) leads to recall of the only one past experience (i.e. episode 1: Figure 2 left, *Object 2*). So iCub anticipates that large heavy *blue cylinder* would displace $365cm^3$ of water and this turns out to be the case once the robot actually drops the object into water. Comparing the expected behaviour of object (reward hub activity in

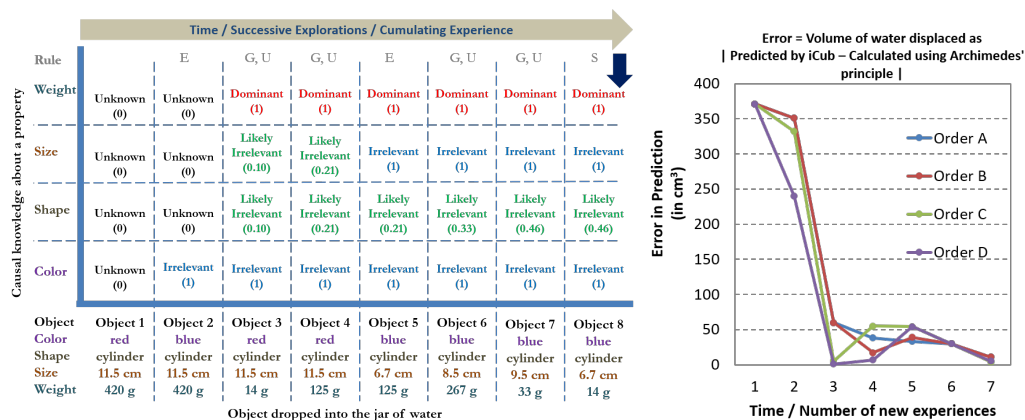


Figure 2: Left panel plots growing causal knowledge as robot explores objects over successive episodes. Causal knowledge regarding a property is either *Unknown* (depicted by 0 in the plot); or has been learnt to be *Dominant* or *Irrelevant* (depicted by 1); or the system expects the property to be *Likely Irrelevant* (a certainty value between 0 and 1). Right panel plots error in iCub’s prediction about the volume of water displaced on dropping an object against actual volume displacement calculated using Archimedes principle. The four curves correspond to four different orders in which the objects were presented to robot for exploration.

the recalled past experience) and the observed behaviour, the robot finds no contradiction. In sum, *there is change in a property (colour), but it did not cause any contradiction in the expected behaviour*. Elimination rule applies here implying colour is not a causally dominant property. The connectivity between the colour map and the object hub is drastically reduced, so they no longer will retro-activate each other. This episode is not encoded into episodic memory.

Learning that weight is a causally dominant property

Episode 3: iCub is presented with a very *light* cylinder (14g) with other properties same as in episode 1 (see Figure 2 left, *Object 3*) Bottom-up activity recalls episode 1. A high reward of 365cm^3 is anticipated. However, only a small amount (24cm^3) of water is displaced after robot’s action, leading to a contradiction between expected and observed behaviours. A comparison of the bottom up activity and the reconstructed top down activity reveals there is a difference in weight map. Growth rule is applied because *a change in the weight property causes contradiction*. The new experience is encoded into the episodic memory which can be recalled next time for better prediction. Furthermore, activity in shape and size maps showed no change even though there was a contradiction between the expected and the observed behaviour. Hence the Uncertainty rule applies too: as the robot still has no experience or complete knowledge of the causal relevance of object-size or shape. The robot partially believes at this point that shape and size of the object may not be relevant in causing water rise.

Accumulating over a set of such experiences (Figure 2 left) with objects of different properties, the robot grows its causal knowledge (certainty) of properties relevant to the task. Furthermore, in cases when the system is presented a cylinder of a weight never experienced before, all the past experiences (due to the same shape) will be recalled and a weighted averaging of the rewards expected due to these past experiences is used as an estimate of net anticipated reward (see Bhat et al. (2016) for more details). As the number of experiences with objects of different weights increases, the accuracy in the prediction of reward increases systematically. Figure 2 (right) show robot’s predictions in four different random orders of 8 objects. Results show that the causal knowledge is same at the end of explorations in all cases and error in prediction (i.e., difference between robots prediction and Archimedes principle) rapidly decreases in all cases.

4 CONCLUSION

The work takes the topic of affordances from the level of *object-action* to the level of *property-action*, in line with emerging studies from neurosciences, and suggests a possible mechanism for causal learning in animals and robots. The work emphasizes that reasoning and learning always have to go hand-in-hand and grow cumulatively and continuously in lifetime of a learner, be it a natural or an artificial cognitive agent.

REFERENCES

- Timothy A. Allen and Norbert J. Fortin. The evolution of episodic memory. *Proceedings of the National Academy of Sciences of the United States of America*, 110(SUPPL2):10379–10386, 2013. ISSN 00278424. doi: 10.1073/pnas.1301199110. URL http://www.pnas.org/content/110/Supplement_{_}2/10379.full.
- Fabien Baradel, Natalia Neverova, Julien Mille, Greg Mori, and Christian Wolf. COPHY: Counterfactual Learning of Physical Dynamics. In *ICLR*, apr 2020. URL <http://arxiv.org/abs/1909.12000>.
- Peter Battaglia, Razvan Pascanu, Matthew Lai, Danilo Rezende, and Koray Kavukcuoglu. Interaction networks for learning about objects, relations and physics. In *Advances in Neural Information Processing Systems*, pp. 4509–4517, 2016.
- Ajaz A. Bhat, Vishwanathan Mohan, Giulio Sandini, and Pietro Morasso. Humanoid infers Archimedes’ principle: Understanding physical relations and object affordances through cumulative learning experiences. *Journal of the Royal Society Interface*, 13(120):20160310, 2016. ISSN 17425662. doi: 10.1098/rsif.2016.0310. URL <http://rsif.royalsocietypublishing.org/lookup/doi/10.1098/rsif.2016.0310>.
- Ajaz A. Bhat, Sharath C. Akkaladevi, Vishwanathan Mohan, Christian Eitzinger, and Pietro Morasso. Towards a learnt neural body schema for dexterous coordination of action in humanoid and industrial robots. *Autonomous Robots*, 41(4):945–966, 2017. ISSN 15737527. doi: 10.1007/s10514-016-9563-3. URL <http://link.springer.com/10.1007/s10514-016-9563-3>.
- Christopher David Bird and Nathan John Emery. Rooks Use Stones to Raise the Water Level to Reach a Floating Worm. *Current Biology*, 19(16):1410–1414, 2009. ISSN 09609822. doi: 10.1016/j.cub.2009.07.033. URL <http://dx.doi.org/10.1016/j.cub.2009.07.033>.
- Lucy G. Cheke, Elsa Loissel, and Nicola S. Clayton. How do children solve Aesop’s fable? *PLoS ONE*, 7(7):e40574, 2012. ISSN 19326203. doi: 10.1371/journal.pone.0040574. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3405099{&}tool=pmcentrez{&}rendertype=abstract>.
- Alison Gopnik, David M. Sobel, David Danks, Clark Glymour, Laura E. Schulz, and Tamar Kushnir. A Theory of Causal Learning in Children: Causal Maps and Bayes Nets. *Psychological Review*, 111(1):3–32, 2004. ISSN 0033295X. doi: 10.1037/0033-295X.111.1.3.
- Thomas L. Griffiths and Joshua B. Tenenbaum. Structure and strength in causal induction. *Cognitive Psychology*, 51(4):334–384, dec 2005. ISSN 00100285. doi: 10.1016/j.cogpsych.2005.05.004.
- Daniel Hanus, Natacha Mendes, Claudio Tennie, and Josep Call. Comparing the performances of apes (gorilla gorilla, pan troglodytes, pongo pygmaeus) and human children (homo sapiens) in the floating peanut task. *PLoS ONE*, 6(6), 2011. ISSN 19326203. doi: 10.1371/journal.pone.0019555.
- J. J. Hopfield. Searching for memories, Sudoku, implicit check bits, and the iterative use of not-always-correct rapid neural computation. *Neural Computation*, 20(5):1119–1164, may 2008. ISSN 08997667. doi: 10.1162/neco.2008.09-06-345. URL <http://www.mitpressjournals.org/doi/abs/10.1162/neco.2007.09-06-345{#}.VquamvkrLIU>.
- Raban Iten, Tony Metger, Henrik Wilming, Lídia del Rio, and Renato Renner. Discovering Physical Concepts with Neural Networks. *Physical Review Letters*, 124(1), jan 2020. ISSN 0031-9007. doi: 10.1103/physrevlett.124.010508.
- Sarah A. Jelbert, Alex H. Taylor, Lucy G. Cheke, Nicola S. Clayton, and Russell D. Gray. Using the aesop’s fable paradigm to investigate causal understanding of water displacement by new caledonian crows. *PLoS ONE*, 9(3):e92895, jan 2014. ISSN 19326203. doi: 10.1371/journal.pone.0092895. URL <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0092895>.

- Markus Kiefer and Friedemann Pulvermüller. Conceptual representations in mind and brain: Theoretical developments, current evidence and future directions. *Cortex*, 48(7):805–825, 2012. ISSN 00109452. doi: 10.1016/j.cortex.2011.04.006.
- Sang Wan Lee, John P. O’Doherty, and Shinsuke Shimojo. Neural Computations Mediating One-Shot Learning in the Human Brain. *PLoS Biology*, 13(4), 2015. ISSN 15457885. doi: 10.1371/journal.pbio.1002137.
- Alex Martin. GRAPES Grounding representations in action, perception, and emotion systems: How object properties and categories are represented in the human brain. *Psychonomic Bulletin and Review*, 23(4):979–990, aug 2016. ISSN 15315320. doi: 10.3758/s13423-015-0842-3.
- Gema Martin-Ordas, Josep Call, and Fernando Colmenares. Tubes, tables and traps: Great apes solve two functionally equivalent trap tasks but show no evidence of transfer across tasks. *Animal Cognition*, 11(3):423–430, 2008. ISSN 14359448. doi: 10.1007/s10071-007-0132-1.
- Vishwanathan Mohan, Giulio Sandini, and Pietro Morasso. A neural framework for organization and flexible utilization of episodic memory in cumulatively learning baby humanoids. *Neural computation*, 26(12):2692–734, dec 2014. ISSN 1530-888X. doi: 10.1162/NECO_a.00664. URL http://www.mitpressjournals.org/doi/abs/10.1162/NECO_a.00664. URL <http://www.ncbi.nlm.nih.gov/pubmed/24179229>.
- Hae Jeong Park and Karl Friston. Structural and functional brain networks: From connections to cognition. *Science*, 342(6158):1238411, nov 2013. ISSN 10959203. doi: 10.1126/science.1238411. URL <http://science.sciencemag.org/content/342/6158/1238411.abstract><http://www.ncbi.nlm.nih.gov/pubmed/24179229>.
- Karalyn Patterson, Peter J Nestor, and Timothy T Rogers. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature reviews. Neuroscience*, 8(12): 976–87, dec 2007. ISSN 1471-0048. doi: 10.1038/nrn2277. URL <http://dx.doi.org/10.1038/nrn2277>.
- Judea Pearl. *Causality*. Cambridge University Press, 2009.
- Matthew A.Lambon Ralph, Elizabeth Jefferies, Karalyn Patterson, and Timothy T. Rogers. The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, 18(1):42–55, dec 2017. ISSN 14710048. doi: 10.1038/nrn.2016.150. URL <http://dx.doi.org/10.1038/nrn.2016.150>.
- Giacomo Rizzolatti and Laila Craighero. The Mirror-Neuron System. *Annual Review of Neuroscience*, 27(1):169–192, jul 2004. ISSN 0147-006X. doi: 10.1146/annurev.neuro.27.070203.144230.
- Shohei Shimizu, Patrik O. Hoyer, Aapo Hyvärinen, and Antti Kerminen. A linear non-gaussian acyclic model for causal discovery. *Journal of Machine Learning Research*, 7:2003–2030, 2006. ISSN 15337928.
- Olaf Sporns. The human connectome: a complex network. *Annals of the New York Academy of Sciences*, 1224:109–25, apr 2011. ISSN 1749-6632. doi: 10.1111/j.1749-6632.2010.05888.x. URL <http://www.ncbi.nlm.nih.gov/pubmed/21251014>.
- Caiming Xiong, Nishant Shukla, Wenlong Xiong, and Song Chun Zhu. Robot learning with a spatial, temporal, and causal and-or graph. In *Proceedings - IEEE International Conference on Robotics and Automation*, volume 2016-June, pp. 2144–2151. Institute of Electrical and Electronics Engineers Inc., jun 2016. ISBN 9781467380263. doi: 10.1109/ICRA.2016.7487364.