What is a Stimulus?
A Computational Perspective on an Associative Learning Model

Rui P. Cardoso[1], Niall Donnelly[2], Edward Keedwell[2], Lucy Cheke[3] and Murray Shanahan[1]

[1]Imperial College London
[2]University of Exeter
[3]University of Cambridge
rui.cardoso@imperial.ac.uk

Abstract

Comparative and animal cognition literature describes many models of associative learning and a multitude of experimental protocols for exploring learning phenomena. These methodologies can serve as inspiration for reinforcement learning (RL) algorithms and tasks (Shanahan et al., 2020). However, there is a considerable gap between animal cognition and RL research, both conceptually and in the assumptions made about the learning process. Associative learning models assume the presence of a "stimulus" guiding a behavioural response, which in the field of RL usually translates loosely into a state-action pair. Our research attempts to investigate and bridge this gap by implementing the A-Learning model (Ghirlanda et al., 2020) into an embodied AI system, using the purpose-built Animal-AI environment (Beyret et al., 2019). Here we present early findings of our research.

Introduction

Since the beginning of psychology as a discipline, researchers have attempted to formally explain the processes by which human and non-human animals learn to interact with their environment. From Hebbian (Hebb, 1949) and Pavlovian (Pavlov, 1927; Rescorla and Wagner, 1972) models of learning to contemporary associative learning theories (Ghirlanda et al., 2020; Lind, 2018; Honey et al., 2020), a recurrent theme has been the ability of human and non-human species to recognise a "stimulus" in their environment, from which various behavioural phenomena can be observed, e.g. blocking (Moore and Schmajuk, 2008) or second-order conditioning (Rashotte et al., 1977), among others.

Associative learning theories have been demonstrably powerful and reliable in predicting the learned behaviour of animals, and they provide experimental protocols which can serve as inspiration for RL algorithms and tasks (Shanahan et al., 2020). However, theorists disagree on the level at which a "stimulus" should be defined, and thus the translation from formal mathematical models (such as the learning rule from Rescorla and Wagner (1972)) into behaving artificial agents is nontrivial. The transformation of a so-called 'neutral' stimulus into a 'conditioned' stimulus via changes to associative strength or stimulus-response weightings or the assumed ability to distinguish unconditioned stimuli from neutral stimuli are aspects which remain vague, especially when attempting to formalise them in an artificial learning model. Consequently, a divide of sorts can be discerned between animal learning researchers and AI researchers. The former usually leave a "stimulus" object formally undefined, whereas the latter typically define it as a "state", which can be thought of as the sum of all currently present stimuli (Sutton and Barto, 2018).

Here we attempt to investigate and bridge this gap by implementing the A-Learning model (Ghirlanda et al., 2020) into an embodied AI system, using the purpose-built Animal-AI environment (Beyret et al., 2019). We define a "stimulus" as a *representation* of the agent's visual input, which is the result of combining an autoencoder (Dong et al., 2018) with a clustering algorithm (Jain et al., 1999). Here we present early findings of our research.

Related Work

Ghirlanda et al. (2020) present the A-Learning model, which is an attempt to bridge the gap between animal cognition theory and machine learning with a clearer mathematical formulation. This model reproduces the main features of other associative learning models, such as instrumental and Pavlovian conditioning. Lind and Vinken (2021) conclude that the A-Learning formulation suffices as a mechanism for general animal intelligence.

Jonsson et al. (2021) implement a learning simulator which acts as an interface to implement associative learning and RL algorithms. However, these explorations leave aside the challenge of incorporating perceptual and motor elements into behaviour, leaving both stimulus and response as symbolically represented. Such an approach provides a significantly impoverished account of animal behaviour as it limits both the inputs and outputs to a small and predefined selection, rather than the "blooming, buzzing confusion" (James, 2004) within which biological agents must learn about the world. Beyret et al. (2019) have designed a purpose-built platform called Animal-AI, which enables the

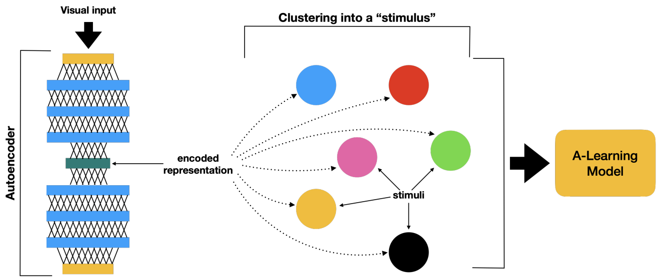

Figure 1: Conceptual model of our implementation

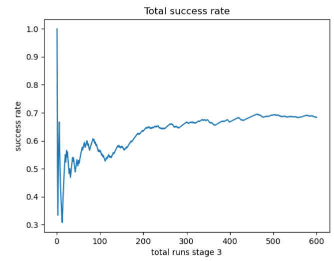

Figure 2: Success rate at finding the correct pit over time

testing of particular learned behaviours by an embodied AI agent in a 3D noise-laden stochastic virtual environment.

Our research implements the A-Learning model, as formally defined by Ghirlanda et al. (2020), within the Animal-AI environment. Our objective is to test whether an agent designed to recognise and process stimuli in accordance with Learning Theory (and the A-Learning model specifically) demonstrates the kind of behavioural phenomena that these theories were created to explain, and — beyond this — whether such an agent may come to behave in an “animal-like” manner across cognitive tasks in general.

Conceptual Model

Figure 1 shows a conceptual model of the proposed solution to the question of the stimulus. We have used the Animal-AI environment (Beyret et al., 2019) to implement the A-Learning model. Animal-AI has enabled us to build an agent which perceives the environment through a visual input and reward signals, and interacts with it via a number of actions. We use an autoencoder to extract a “stimulus” from the visual input of the agent in two steps: firstly, the autoencoder produces an encoded representation of the visual input of the agent; secondly, the encoded representation is clustered together with other representations into a category of inputs, each of which we consider as a “stimulus”. The autoencoder is pretrained to reconstruct samples of the visual input of the agent. The clustering of visual encodings into stimulus categories is based on the *cosine distance* between the encodings and the *centroid* of each cluster (a constantly updated mean of all encodings belonging to the cluster). Initially, a maximum cut-off distance is defined in order to initialise the clusters. When a maximum number of clusters is reached, new encodings are added to the nearest cluster, in a similar fashion to K-means clustering (Lloyd, 1982).

We have created a test suite on Animal-AI based on experiments carried out on pigeons by Mondragón and Hall (2015). In these, the agent is given the choice of falling down two pits, one of which has a reward and the other results in a punishment. A posterboard indicates, according to its colour or shape, which pit the agent should choose in order to get the reward. The agent is first pretrained on simpler tasks which only involve falling down either pit in order to get a reward. It is then trained on the actual task over a number of episodes, each lasting a certain number of time steps. The agent receives a penalty if it “times out”. At the end of an episode, a sequence $s_1 \to r_1 \to s_2 \to r_2 \to \ldots \to s_n \to r_n$ of stimuli followed by responses (actions) is produced, which is used to update both stimulus values and stimulus-response values in accordance to the A-Learning formulation (Ghirlanda et al., 2020). The training procedure balances exploration with exploitation.

Preliminary Results

Our preliminary results show that the A-Learning agent is capable of solving the task inspired by Mondragón and Hall (2015) with a success rate of approximately 70% over time (see Figure 2). This means that the agent is able to achieve a performance similar to that which pigeons achieve on a similar task.

Conclusions and Future Work

Our initial findings highlight some of the challenges in translating associative learning theories, designed to *explain* and *predict* animal learning and behaviour, into models which can be used by an artificial agent in order to learn and behave in a virtual environment in a similar way to how an animal might. It has proved particularly important to clarify what is meant by a “stimulus” in the interaction between such an agent and virtual environment, as well as to design an appropriate training curriculum which can shape the behaviour of the agent in the desired way.

Future research will design training curricula and learning strategies which improve on the current ability of the agent to solve more challenging tasks, including binary choice maze environments, such as T-Maze (Deacon and Rawlins, 2006) and Y-Maze (Kraeuter et al., 2019) configurations. We will also investigate more complex phenomena, such as blocking (Moore and Schmajuk, 2008) and second-order conditioning (Rashotte et al., 1977). We will likewise assess whether the A-Learning agent can solve tasks testing for object permanence with a test suite which has been designed for the Animal-AI environment (Voudouris et al., 2022), containing over 5,000 tests.

References

Beyret, B., Hernández-Orallo, J., Cheke, L., Halina, M., Shanahan, M., and Crosby, M. (2019). The animal-ai environment: Training and testing animal-like artificial cognition.

Deacon, R. M. and Rawlins, J. N. P. (2006). T-maze alternation in the rodent. *Nature protocols*, 1(1):7–12.

Dong, G., Liao, G., Liu, H., and Kuang, G. (2018). A review of the autoencoder and its variants: A comparative perspective from target recognition in synthetic-aperture radar images. *IEEE Geoscience and Remote Sensing Magazine*, 6(3):44–68.

Ghirlanda, S., Lind, J., and Enquist, M. (2020). A-learning: A new formulation of associative learning theory.

Hebb, D. O. (1949). Organization of behavior. new york: Wiley. *J. Clin. Psychol*, 6(3):335–307.

Honey, R. C., Dwyer, D. M., and Iliescu, A. F. (2020). Elaboration of a model of pavlovian learning and performance: Heidi. *Journal of Experimental Psychology: Animal Learning and Cognition*, 46.

Jain, A. K., Murty, M. N., and Flynn, P. J. (1999). Data clustering: A review. volume 31.

James, W. (2004). *The principles of psychology, Vol I.*

Jonsson, M., Ghirlanda, S., Lind, J., Vinken, V., and Enquist, M. (2021). Learning simulator: A simulation software for animal and human learning. *Journal of Open Source Software*, 6(58):2891.

Kraeuter, A.-K., Guest, P. C., and Sarnyai, Z. (2019). The y-maze for assessment of spatial working and reference memory in mice. *Pre-clinical models: Techniques and protocols*, pages 105–111.

Lind, J. (2018). What can associative learning do for planning? *Royal Society open science*, 5(11):180778.

Lind, J. and Vinken, V. (2021). Can associative learning be the general process for intelligent behavior in non-human animals? *bioRxiv*.

Lloyd, S. P. (1982). Least squares quantization in pcm. *IEEE Transactions on Information Theory*, 28.

Mondragón, E. and Hall, G. (2015). Analysis of the role of stimulus comparison in discrimination learning in pigeons. *Learning and Motivation*, 49.

Moore, J. and Schmajuk, N. (2008). Kamin blocking. *Scholarpedia*, 3.

Pavlov, I. (1927). Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. *Oxford Univ. Press*.

Rashotte, M. E., Griffin, R. W., and Sisk, C. L. (1977). Second-order conditioning of the pigeon's keypeck. *Animal Learning & Behavior*, 5(1):25–38.

Rescorla, R. A. and Wagner, A. R. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Clasical conditioning II: current research and theory*.

Shanahan, M., Crosby, M., Beyret, B., and Cheke, L. (2020). Artificial intelligence and the common sense of animals.

Sutton, R. S. and Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. The MIT Press, second edition.

Voudouris, K., Donnelly, N., Rutar, D., Burnell, R., Burden, J., Hernández-Orallo, J., and Cheke, L. G. (2022). Evaluating object permanence in embodied agents using the animal-ai environment. *https://ceur-ws. org/Vol-3169/paper2. pdf*.