Relation learning in a neurocomputational architecture supports cross-domain transfer

Leonidas A. A. Doumas (alex.doumas@ed.ac.uk), Guillermo Puebla (guillermo.puebla@ed.ac.uk)
Department of Psychology University of Edinburgh

Andrea E. Martin (andrea.martin@mpi.nl)
Max Planck Institute for Psycholinguistics

John E. Hummel (jehummel@illinois.edu)
Department of Psychology University of Illinois

Abstract

Humans readily generalize, applying prior knowledge to novel situations and stimuli. Advances in machine learning have begun to approximate and even surpass human performance, but these systems struggle to generalize what they have learned to untrained situations. We present a model based on well-established neurocomputational principles that demonstrates human-level generalisation. This model is trained to play one video game (Breakout) and performs one-shot generalisation to a new game (Pong) with different characteristics. The model generalizes because it learns structured representations that are functionally symbolic (viz., a role-filler binding calculus) from unstructured training data. It does so without feedback, and without requiring that structured representations are specified a priori. Specifically, the model uses neural co-activation to discover which characteristics of the input are invariant and to learn relational predicates, and oscillatory regularities in network firing to bind predicates to arguments. To our knowledge, this is the first demonstration of human-like generalisation in a machine system that does not assume structured representations to begin with.

Keywords: predicate learning; generalisation; neural networks; symbolic-connectionism; neural oscillations

Introduction

Recently deep neural network (DNN) systems have reached and even exceeded human levels of performance on a range of cognitive tasks (for a review see, Hassabis, Kumaran, Summerfield, & Botvinick, 2017). For example, DNNs have learned to master an impressive number of games (Mnih et al., 2015; Silver et al., 2017). DNNs are general, in that they can learn to perform a variety of tasks without a priori background knowledge. Nevertheless, while DNNs readily perform interpolation (i.e., generalisation to untrained items from within the bounds of the training set), they struggle to perform extrapolation (i.e., generalisation to items from outside the bounds of the training set). For example, a network trained to play Breakout must be completely retrained to play Pong (Mnih et al., 2015).

In contrast, a person is able to quickly catch on to playing a game like Pong after learning to play a game like Breakout. After all, Breakout and Pong are very similar: In both games the objective is to use a paddle to keep a ball in play, and to hit the ball toward some goal. While in Breakout the ball is played vertically towards blocks at the top of the screen, and in Pong the ball is played horizontally towards an opponent paddle.

Accounts of how humans generalize are frequently based on powerful symbolic languages that include structured relations (or predicates), which can be promiscuously applied to new arguments (Doumas & Hummel, 2012; Lake, Ullman, Tenenbaum, & Gershman, 2017). In this view, we have abstract representations like right-of and above. These representations allow us to characterize different domains with the same representations, and generalize what we have learned about these representations across domains. Structured models, however, face a challenge that is complementary to that which DNNs face: They characteristically require the modeler to specify a collection of necessary representational structures in advance of any actual learning (e.g, Lake, Salakhutdinov, & Tenenbaum, 2015).

We have previously proposed a neural network model of how structured representations are instantiated in a biologically plausible neural system, and how such representations are learned in the first place (Doumas, Hummel, & Sandhofer, 2008). The model, called DORA, uses unsupervised comparison to discover which characteristics of the input are invariant, and to learn functional predicates; it then applies these predicates to arguments in a symbolic fashion, using oscillatory regularities to dynamically bind predicates and arguments. DORA learns representations that are functionally and formally symbolic from flat vector data, without feedback, and without requiring that structured representations be specified a priori.

In the following we show that after learning to play one video game, Breakout, the representations that DORA learns support generalisation to a completely new game, Pong, in one shot. Importantly, DORA's learning and reasoning rely intimately on the phase dynamics that carry binding information in the model.

Model description

DORA is a symbolic-connectionist model descended from LISA (Hummel & Holyoak, 2003). Its operation is summarized as follows. (1) DORA starts with representations of differentiated objects encoded as flat feature vectors. (2) Through a process of analogical mapping, objects are compared (and co-activated) and their feature vectors are superimposed. (3) DORA learns a representation of the overlaid pattern through Hebbian learning. The resulting representation is an encoding of what the compared objects have in common. (4) The learned representations are bound to objects by systematic asynchrony of firing, resulting in functional single-place predicates. (5) Co-occurring sets of single-place predicates are linked to form functional multi-place relations. Below we provide a conceptual overview of DORA's operation at a high level of abstraction (for computational details see, Doumas et al., 2008).

Computational macrostructure

DORA has a long-term-memory (LTM; see Fig. 1) composed of bidirectionally connected layers of units. Units in LTM are referred to as token units. Token units in the lowest layer of LTM are bidirectionally connected to a common pool of feature units. Token units are yoked to inhibitors that integrate input from their yoked unit and token units in higher layers, and fire after reaching a threshold. Yoked inhibitors serve the purpose of implementing phasic firing and refractory periods in the token units, which are important for implementing dynamic binding in the network.

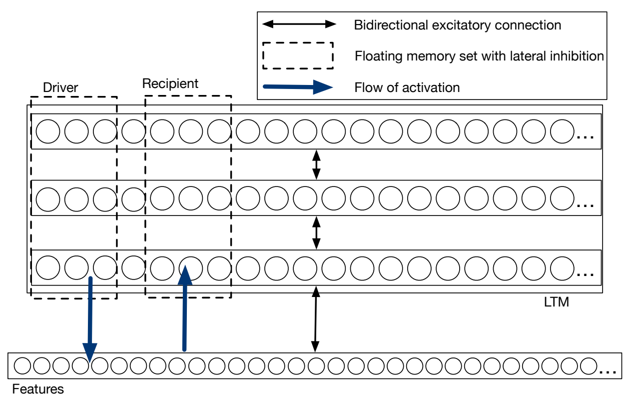

Figure 1: Macrostructure of the DORA network.

Potentiated sets of token units, or memory sets (dashed boxes in Fig. 1), correspond to DORA's working memory. Memory sets include, the driver, DORA's current focus of attention, and the recipient, DORA's current active memory. Token units in the same layer inhibit one another within, but not across, memory sets. Activation in the model flows from the token units in the driver to token units in the recipient and LTM via the shared pool of feature units.

Representation learning

At a very high level, DORA's learning algorithm has three important features (for details see (Doumas et al., 2008)). DORA starts with representations of single objects encoded as a flat vectors of features. A localist token unit connects to the features defining the specific object (Fig. 2A). As the first step in learning, DORA compares multiple objects. Compared objects become co-active, and pass activation to their constituent features. Any features shared by the compared objects will receive roughly twice as much input and become roughly twice as active as unshared features (Fig. 2Bi). The process of comparison serves to highlight features shared (and unshared) by compared objects.

Second, DORA learns an explicit representation of the shared properties using Hebbian learning. During comparison, units are recruited in the token layer i (TL i), connected to the features, and token layer ii (TL ii), and learn connections to active units in proportion to their activation (Fig. 2Bii). As a result, DORA learns a token unit that conjunctively codes for the rough featural overlap of the compared objects (unit *r* in Fig. 2Bii), and another token unit linking the newly learned unit to one of the compared objects units (unit *r*+b Fig. 2Bii).

Crucially, these new representations are bound to arguments via time-based binding (wherein binding information is carried by when units fire), and therefore function as single-place predicates (Doumas et al., 2008). In DORA, bound predicates and arguments fire in direct sequence. For example, to bind a predicate *r* to an object b, the units representing *r* fire followed by the units representing b, and to bind a predicate *l* to an object p, the units representing *l* fire followed by the units representing p (Fig. 2Ci). The binding signal is explicit and dynamic (i.e., binding information can be created and destroyed on the fly): Binding *r* to p and *l* to b only involves the units representing *r*, *l*, b, and p firing in a different order (Fig. 2Cii). Time-based binding emerges naturally in a system with lateral inhibition, unit refraction, and conjunctive encodings of linked units (Doumas & Hummel, 2005; Doumas et al., 2008).

Third, DORA learns representations of multi-place relational structures, by linking systematically co-occurring predicate-argument sets. When structurally similar sets of predicate-argument pairs are in the driver and recipient and are compared, a systematic pattern of firing necessarily emerges: Specifically, similar predicate-argument pairs will be co-active in the driver and recipient, and out will fire of phase with any other predicate-argument pairs (Fig. 2Di). DORA uses the same Hebbian recruitment and learning described in the second step above to link sets of predicate-argument pairs. Explicitly, DORA recruits a unit at token layer iii (TL iii), and learns connections to active units in TL ii via Hebbian learning as they become active (Fig. 2Di). Thus, DORA learns to link a set of predicate-argument pairs into a single structure (Fig. 2Dii). The resulting structure effectively encodes and behaves like a multi-place predicate (for details see Doumas et al., 2008; Hummel & Holyoak, 2003).

Processing

DORA is a settling network. It starts in some state, such as a set of units in driver (e.g., chosen at random from LTM, or based on DORA's current perceptual state such as a videogame screen shot). Token units in the driver compete (via lateral inhibition) to become active, and activation flows to token units in the recipient and LTM via shared feature units. DORA eventually settles into some state (e.g., with some units active in driver and recipient). Due to the refraction of

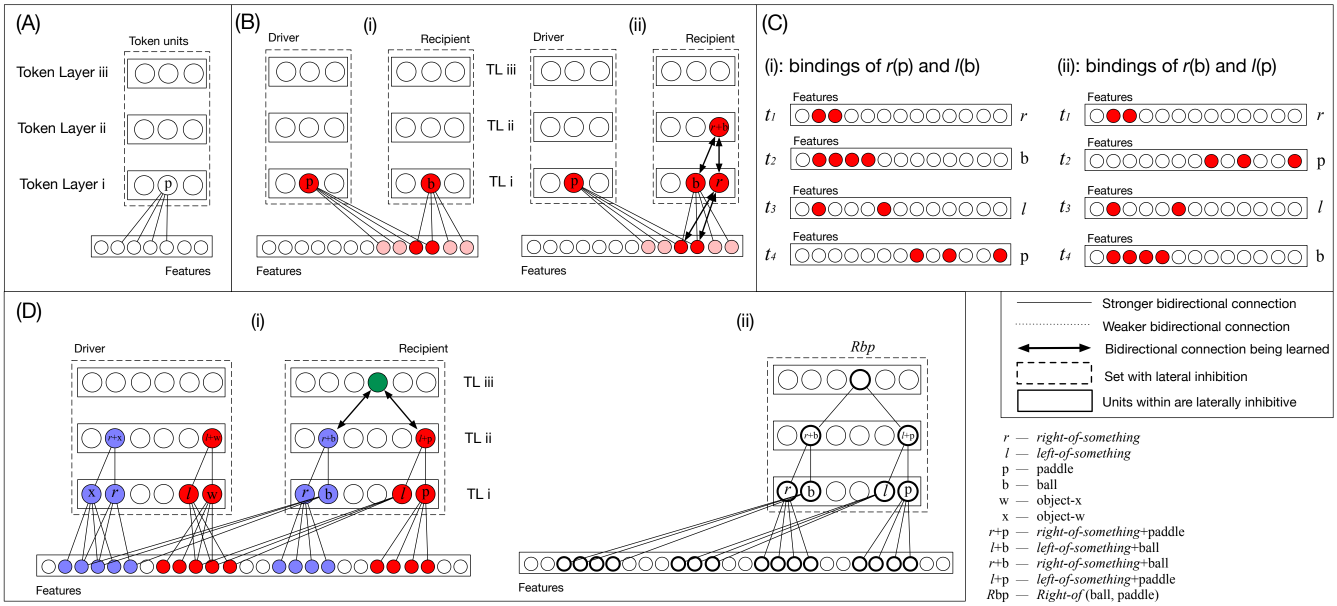

Figure 2: Representation learning in DORA. (A) DORA starts with representations of objects encoded as flat feature vectors. (B) (i) By comparing objects b and p, shared features receive more input and become more active. (ii) Using Hebbian learning DORA learns an explicit representation of the featural overlap of b and p—unit labeled *r*—and links *r* to b (unit labeled *r*+b). (C) Illustration of binding in DORA. (i) To bind *r* to b and *l* to p units coding *r* fire at t1, followed by units for b at t2, *l* at t3 and p at t4. (ii) Complementary binding information is carried by a different sequence of firing. (D) Learning multi-place relations. (i) Similar predicate-argument sets are compared in driver and recipient. Driver units activate featurally similar units in recipient: Violet units in the driver will activate violet units in recipient, and red units in the driver will cause red recipient units to become active. Using Hebbian learning DORA learns a conjunctive encoding of the predicate-argument pairs in recipient as they become active (green unit in Tiii). (ii) The resulting representation encodes a multi-place relational proposition.

nodes and yoked inhibitors, this state will eventually become upset and the process will start again.

The representations that DORA learns support a variety of operations from the LISA and DORA including analogical mapping, schema refinement, and relational generalisation. During mapping, DORA discovers structural correspondences between token units in the driver and recipient. During refinement, DORA learns a schema from the featural intersection of mapped items in driver and recipient. During generalisation, DORA implements a version of copy-with-substitution-and-generalisation (CWSG) (Holyoak, Novick, & Melz, 1994), wherein information from one situation is carried over into a mapped situation (described in more detail in section following).

DORA's algorithm is capable of composing features into structured representations of relations and arguments (i.e., propositions). However, in order to learn relational representations, there need to be invariants that characterise the underlying relations—e.g., to learn a representation of above that captures every instance of aboveness, there must be some detectable property(ies) that remain constant over all instances of aboveness (Biederman, 2013). We have developed a novel algorithm to discover invariants for relative magnitude (e.g., "same", "more", "less") based on the known properties of neural encodings of absolute magnitude and eye movements (Doumas, Hamer, Puebla, & Martin, 2017). The algorithm exploits the invariants that emerge when neural encodings of absolute magnitude are superimposed.

Model in context

DORA is a model of representation learning. It assumes that objects are differentiated and makes no strong claims about how choices between available options (i.e., moves in a video game) are made. As such, we situated DORA's predicate learning algorithm between a visual pre-processor, and tabular Q-learning (Watkins, 1989) (see Fig. 3). The visual pre-processor served to differentiate objects, and the tabular Q-learning allowed DORA to learn associations between representational states and move options in a game.

The visual pre-processor used edge detection (via local contrast) with an inbuilt bias such that any enclosed edges were treated as a single object. The pre-processor delivered representations of individual objects characterised by their colour and their location, (represented in raw pixels).

Simulations

We ran three sets of simulations. Simulation 1 compared DORA to several other networks for their capacity to generalize to Pong after training on Breakout. Simulation 2 served

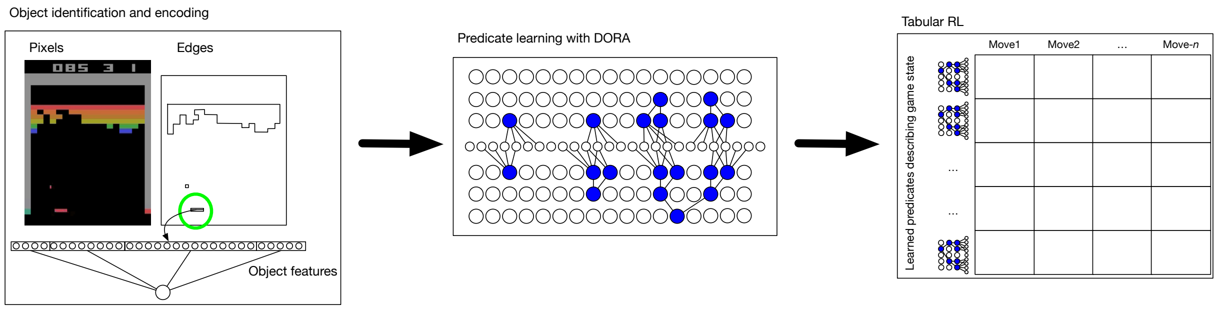

Figure 3: Predicate learning in context.

to evaluate the capacity of the relational representations that DORA learned from video game screens to support human-level analogical reasoning. Simulation 3 extended the test of cross-domain transfer: DORA learned representations from instances unrelated to games, and then used these representations to learn to play one game and generalize to another

Simulation 1

We compared (1) an implementation of DORA with Q-learning against (2) DQN; (3) DQN with the same pre-processed inputs used by DORA; (4) a supervised deep neural network (DNN) with the same pre-processed inputs used by DORA with fixed frame skipping; (5) a supervised DNN with the same pre-processed inputs used by DORA with random frame skipping; (6) Humans (two Breakout and Pong novices). We trained all these systems to play one videogame (Breakout), and then tested their ability to generalize to a different videogame (Pong) without any explicit training. Finally, we evaluated these systems' ability to switch back to playing the original game, after time spent learning to play the second.

For the first 250 games of Breakout, DORA made random moves, generating game states from which it learned structured representations in an unsupervised manner as described above. DORA successfully learned predicate representations encoding to instances such as *more-y*(object1, object2) and *more-x*(object1, object2). DORA then attempted to learn to play Breakout using the representations that it had learned during the first 250 games to represent the current game screen and then made a response. Associations between these learned representations and successful moves were learned via tabular Q-learning. Fig. 4a shows the performance of all networks on Breakout as an average score of the last 100 games played, and a high score. All systems performed quite well, reaching levels of performance that matched or exceeded human participants. As would be expected, DORA took far fewer games to learn to play Breakout than any of the other networks (1,000 vs. 10,000,000 games for DORA and DQN, respectively).

We then tested the capacity of the networks to play a new videogame, Pong. DORA had learned to play Breakout by learning associations between relational configurations and

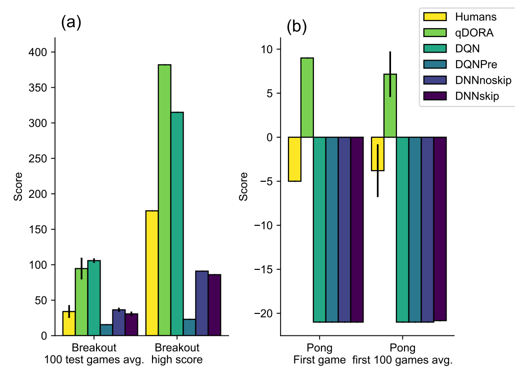

Figure 4: Game play performance for DORA and DNNs.

actions. During its first game of Pong, DORA represented the game state using the relations it had learned playing Breakout. DORA discovered a correspondence between the action sets in the two games: particularly, more-y/less-y of the paddle (the paddle moves up and down) in Pong and more-x/less-x of the paddle (the paddle moves horizontally) in Breakout. This correspondence allowed DORA to infer via relational generalisation the configurations that reward specific moves in Pong. For example, just as more-x(ball, paddle) tends to reward a more-x move of the paddle in Breakout, more-y(ball, paddle) rewards a more-y move in Pong (see Fig. 4).

Fig. 4b shows the performance of the human players and the networks on the first game of Pong after training on Breakout and the average performance over the first 100 games playing Pong. Like a human player, DORA performed at a high level on Pong on a single exposure to the game and continued to play Pong at a high level. By contrast, all other networks showed poor performance –which is unsurprising given previous results using DNNs and transferring to different contexts.

Importantly, the generalisation failure of the networks using the same visual processing as DORA demonstrates that the visual processing does not produce representations that support the generalisation performance demonstrated by DORA. Using both Q-learning and supervised learning, the unstructured (i.e., non-symbolic) representations produced

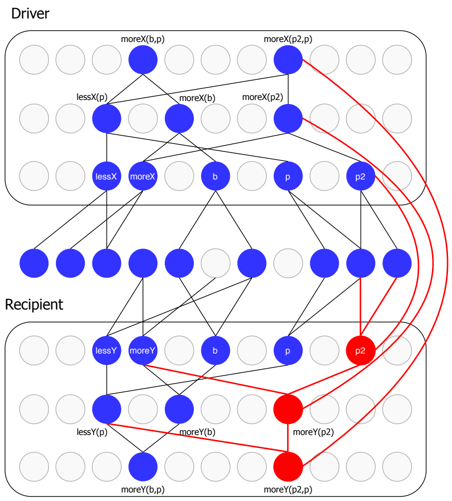

Figure 5: Relational generalisation in DORA.

by the visual processor and the DQN did not support extrapolatory generalisation, while the predicate representations learned by DORA did.

Simulation 2

Simulation 2 was designed to investigate whether the representations DORA learned in Simulation 1 have the characteristics of structured relational representations. One of the key properties of structured representations is that they should be applicable (i.e., generalise) across contexts: The representation of *larger* should apply just as easily (and mean the same thing) in a domain like the game Breakout and an analogical reasoning task. To this end, we tested whether the representations that DORA learned in Simulation 1 would immediately generalise to support human level analogical reasoning by testing whether—with no additional learning—they: (i) support solving cross mappings; (ii) support mapping similar, but non-identical predicates; (iii) support mapping objects with no featural overlap, including completely novel objects, if they play similar roles; and (iv) provide a basis for mapping the arguments of a n-place relation onto those of an m-place relation even when n and m are unequal (i.e., whether they, like people, can violate the *n-ary restriction*, according to which an n-place predicate can only map to another n-place predicate). As the representations that DORA learned from video games should be immediately generalisable to these other tasks, this simulation used only those representations learned during simulation 1 and included no new representation learning.

During a cross-mapping, an object (object1) is mapped to a featurally less similar object (object2) rather than a featurally more similar object (object3) because it (object1) plays the same role as the less similar object (object2). For example, if cat1 chases mouse1 and mouse2 chases cat2, then the structural cross-mapping places cat1 into correspondence with mouse2 because both are bound to the *chaser* role. The ability to find such a mapping is a key property of genuinely relational (i.e., as opposed to feature-based) processing. Cross-mappings serve as a stringent test of the structure sensitivity of a representation as they require violating featural or statistical similarity.

To test the representations DORA learned in Simulation 1 for their ability to support cross-mappings we randomly selected two of the representations DORA had learned for a given relation (e.g., both coded for *above*). DORA bound these representations to new objects, creating two new propositions, P1 and P2, such that the agent of P1 was featurally identical to the patient of P2 and vice versa. DORA then mapped P1 onto P2. We repeated this procedure 10 times (each time with a different randomly chosen relational representaitons). In each simulation, DORA successfully mapped the agent of P1 to the agent of P2 (the correct relational map) and vice-versa. DORA's success indicates that the relations it learned in the first part of this simulation apply immediately across tasks and supporting cross-mapping.

We then tested whether DORA's representations support mapping similar but nonidentical relations (such as mapping above to greater-than) and support mapping objects with no featural overlap that play similar roles. We selected two of the refined relations that DORA had learned during Simulation 1, P1 and P2 (e.g, *above*(x,y) or *wider*(x,y)), such that each role in P1 shared roughly 50% of its features with a corresponding role in P2 (e.g., the role *more-height* has 50% of its features in common with the role *more-width*). To assure that no mappings would be based on object similarity and that the mapping would work with completely novel object, none of the objects that served as arguments of the relations had any featural overlap and the object features were units that we added to DORA solely for these simulations (i.e., these were feature units DORA had not "experienced" previously). We repeated this process 10 times, each time with a different pair of relations. Each time, DORA mapped the agent role of P1 to the agent role of P2 and the patient role of P1 to the patient role of P2, and, despite their lack of featural overlap, corresponding objects always mapped to one another (because of their bindings to mapped roles).

Finally, we tested whether the representations DORA learned can violate the n-ary restriction, mapping the arguments of an n-place predicate onto those of an m-place predicate when $n \neq m$. Models of relational thinking based on propositional notation or labelled graphs are unable to map predicates with different numbers of arguments, but people have little difficulty doing so, as evidenced by our ability to map the arguments of, say, bigger (Sam, Larry) on onto those of small (Joyce) and big (Susan) (see Hummel & Holyoak, 1997).

To test DORA's ability to solve such mappings, we randomly selected a relation, P1, that DORA had learned in the previous part of this simulation. We then created a single place predicate (p2) that shared 50% of its features with the agent role of P1 and none of its features with the patient role.

The objects bound to the agent and patient role of P1 each shared 50% of their features with the object bound to p2. DORA attempted to map P1 to p2. We repeated this process 10 times, each time with a different relation from DORA's LTM, and each time DORA successfully mapped the agent role of P1 to p2, along with their arguments. We then repeated the simulation such that p2 shared half its featural content with the patient (rather than agent) role of P1. In 10 additional simulations, DORA successfully mapped the patient role of P1 to p2 (along with their arguments).

General Discussion

We have shown that a machine system can perform extrapolatory generalisation. Specifically, DORA used predicate learning to discover symbolic representations from video game screen shots without feedback, and without assuming any structured representations a priori. Crucially, the predicate representations that DORA learned allowed it to extrapolate its knowledge to untrained situations. Specifically, the model was able to use the representations that it learned playing Breakout to successfully play a new game, Pong, and to perform a number of analogical reasoning tasks. Just like human players, generalizing to a new game, like Pong, or a new task, like analogy making, was fast (zero-shot) and did not affect the system's ability to play the previously learned game. In contrast, four different DNNs failed to transfer knowledge from Breakout to Pong.

These results demonstrate that extrapolatory and cross-domain generalisation can be greatly facilitated by learning and explicitly representing the relations—rather than just the literal features—characterizing the domain in question. By learning the relations characterizing the abstract structure of Breakout (e.g., relations between the locations of the paddle and the ball), DORA was prepared to discover analogical correspondences between Breakout and Pong. On this approach to domain learning, cross-domain transfer is not a matter of learning a wholly new domain but is instead a matter of learning how old knowledge applies to new problems. We argue that this approach is precisely the approach the human mind takes to learning, both within and across domains.

Our relation-based approach represents a fundamental departure from DNN or other statistical machine learning approaches, which learn only the statistical relations between input states (e.g., features of a game screen), and output states (e.g., moves of the game paddle left/right or up/down). Statistical relations between input and output features are, in a trivial sense, relations, so DNNs can in this sense be described as "learning relations". But the simulations described here demonstrate that there is a fundamental difference between the capacity to learn statistical relations between features (the approach in traditional machine learning) and the ability to learn an open-ended set of abstract structured relational representations, both between and within the domains to be learned. The capacity to learn an open-ended set of relations and represent them as explicit predicates confers the capacity to profit from prior learning for reasoning a new domain, rather than suffering from it. Moreover, as demonstrated here and elsewhere (e.g., Doumas et al., 2008), structured relational representations can be learned without having to assume that such structures exist a priori (Lake et al., 2015).

References

Biederman, I. (2013). Human object recognition: Appearance vs. shape. In *Shape perception in human and computer vision* (pp. 387–397). Springer.

Doumas, L. A., Hamer, A., Puebla, G., & Martin, A. E. (2017). A theory of the detection and learning of structured representations of similarity and relative magnitude. In *The 39th annual conference of the cognitive science society (cogsci 2017)* (pp. 1955–1960).

Doumas, L. A., & Hummel, J. E. (2005). Approaches to modeling human mental representations: What works, what doesn't and why. *The Cambridge handbook of thinking and reasoning, ed. KJ Holyoak & RG Morrison*, 73–94.

Doumas, L. A., & Hummel, J. E. (2012). Computational models of higher cognition. *Oxford handbook of thinking and reasoning*, 52–66.

Doumas, L. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological review*, *115*(1), 1.

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-inspired artificial intelligence. *Neuron*, *95*(2), 245–258.

Holyoak, K. J., Novick, L. R., & Melz, E. R. (1994). *Component processes in analogical transfer: Mapping, pattern completion, and adaptation.* Ablex Publishing.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological review*, *104*(3), 427.

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological review*, *110*(2), 220.

Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, *350*(6266), 1332–1338.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and brain sciences*, *40*.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... others (2015). Human-level control through deep reinforcement learning. *Nature*, *518*(7540), 529.

Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., ... others (2017). Mastering the game of Go without human knowledge. *Nature*, *550*(7676), 354.

Watkins, C. J. C. H. (1989). *Learning from delayed rewards.* (Unpublished doctoral dissertation). University of Cambridge.