A theory of the detection and learning of structured representations of similarity and relative magnitude

Leonidas A.A. Doumas[1] (alex.doumas@ed.ac.uk); Aaron Hamer[1] (aaron.hamer@ed.ac.uk); Guillermo Puebla[1] (pueblaramirezg@gmail.com); Andrea E. Martin[1, 2] (andrea.martin@mpi.nl)

[1]Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ, United Kingdom

[2]Max Planck Institute for Psycholinguistics; Nijmegen, The Netherlands

Abstract

Responding to similarity, difference, and relative magnitude is ubiquitous in the animal kingdom. However, humans seem unique in the ability to represent relative magnitude and similarity as abstract relations that take arguments (e.g., *greater-than* (*x*,*y*)). While many models use structured relational representations of magnitude and similarity, little progress has been made on how these representations arise. Models that use these representations assume access to computations of similarity and magnitude a priori. We detail a mechanism for producing invariant responses to "same", "different", "more", and "less" which can be exploited to compute similarity and magnitude as an evaluation operator. Using DORA (Doumas, Hummel, & Sandhofer, 2008), these invariant responses can serve to learn structured relational representations of relative magnitude and similarity from pixel images of simple shapes.

Introduction

Reacting to similarity, and magnitude ("same"/ "different", "more"/"less"; SDML) are hallmarks of complex organisms. For example, gerbils use the retinal size of a stimulus to estimate its distance (Goodale, Ellard, & Booth, 1990), rats choose the larger of two food rewards (Kim et al., 2015), and pigeons learn to group pictures of 16 identical items in one set, and pictures of 16 different items in a different set (Young, Wasserman, & Garner, 1997).

Humans, however, go beyond simple detection of relative magnitude and similarity. We make an analogies between a nucleus and the sun because they are both *larger* than their orbiting bodies (electrons and planets). We infer this relationship because we represent relative magnitude and similarity as abstract relations that take arguments (i.e., as predicates; see Holyoak, 2012).

Our ability to reason about abstract SDML manifests in a variety of domains such as analogy (e.g., Holyoak & Thagard, 1995), categorisation (e.g., Medin, Goldstone, & Gentner, 1993), and concept learning (e.g., Doumas & Hummel, 2013). While models that use structured representations have had success in accounting for how humans use abstract SDML, these models say little about where the representations they use come from in the first place. For example, SME (Falkenhainer, Forbus, & Gentner, 1989), STAR (Halford et al., 1998), and LISA (Hummel & Holyoak, 1997, 2003) account for many phenomena from the analogy literature, but require the relations they use to make these analogies be hand-coded by the modeler. Similarly, Bayesian models of concept development and learning (e.g., Kemp, 2012; Kemp & Tenenbaum, 2007, 2009; Lake et al., 2016) assume relational structures a priori, starting with a vocabulary of objects and relations and learning new concepts by building new combinations of these innate elements.

Some models attempt to account for the origins of abstract concepts without assuming innate representations of relational concepts. For example, BART (Lu, Chen, & Holyoak, 2012) uses feature lists generated by human subjects or corpora analysis to find properties associated with items in the world which instantiate particular relations. BART has difficulty with some edge cases of relational cognition (e.g., reasoning about something like an atom being *bigger* than something else when it has not experienced instances where an atom was bigger than anything), but the model makes a serious effort to account for development of analogy-making with minimal assumptions about the starting representations of the learning system.

In a similar vein, DORA (Doumas, Hummel, & Sandhofer, 2008) explains how structured representations (i.e., predicates) can be acquired from unstructured representations (i.e., feature vectors). While DORA learns relational representations that can take any arguments (including edge cases and completely novel arguments ; Doumas et al., 2008), DORA assumes a system to detect the invariant features that underlie the abstract concepts that it learns.

A complete account of how people acquire structured representations of abstract SDML relations must solve three problems. First, there must be some *invariant* features which remain constant across instances of the relation which the perceptual/cognitive system can learn to detect. Second, the system must isolate these invariants from other properties of the objects engaged in the relation to be learned. Third, the system must learn a *predicate* representation of the relational properties (i.e., an explicit entity that can be bound to arbitrary, novel arguments).

We solve the first problem with an extension to DORA which produces invariant responses to similarity and relative magnitude. We have previously shown how DORA can solve the second and third of these problems (Doumas et al., 2008). We begin with a brief overview of DORA, describe the process which produces invariant features for SDML, and provide simulations demonstrating how DORA solves all three problems to learn structured relational representations of SDML.

Model

DORA

DORA (Doumas, et al., 2008) is a symbolic-connectionist model, based on the LISA (Hummel &

Holyoak, 1997, 2003) model of analogy. DORA learns structured relational representations from unstructured representations of objects (e.g. feature vectors).

LISAese Representations We begin by describing the end state of DORA's representations (i.e., its representations *after* it has gone through learning). Relational propositions are represented by a hierarchy of distributed and localist codes (see Figure 1). At the bottom, semantic units code the features of objects and roles in a distributed fashion. In the next layer, localist predicate-object (PO) units representing individual predicates (or roles) and objects, are connected to these distributed semantic representations. In the next layer, localist role-binding (RB) units link predicates and objects into specific role-filler pairs. At the top of the hierarchy, localist proposition (P) units link RB units into complete relational propositions. Importantly, while we use different names for the units in different layers, and different shapes to distinguish these units in diagrams, we do so only for the purposes of expositional brevity. These are just nodes in different layers of a network. RB units are just like PO units, except for the fact that they are in a different layer, and, therefore, take input from and pass input to different layers of units.

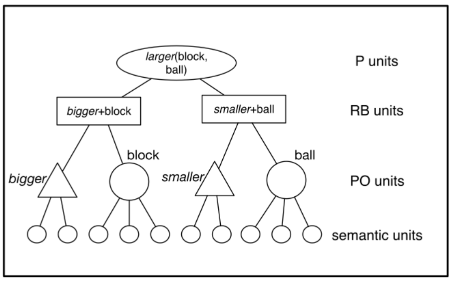

Figure 1. Complete relational proposition in DORA. Units in different layers are coded using different shapes for the purposes of exposition.

Propositions in DORA are divided into four mutually-exclusive sets of layered networks: a *driver*, one or more *recipients*, *long-term memory* (LTM), and the *emerging recipient* (EM). Each set consists of a layered network of PO, RBs, and P units (i.e., there are specific layers coding for PO, RB, and P units in the driver, and another set of layers coding for PO, RB, and P units in the recipient). Semantic units are shared across all networks (i.e., driver and recipient units are connected to the same pool of semantic units). The driver corresponds to the current focus of attention and controls the flow of activation. Units in the driver pass activation to the semantic units. Because the semantic units are shared by all sets, activation flows from the driver to the other three sets. DORA operations (e.g., *mapping* and *relation learning*, detailed below) proceed as a product of units in the driver activating their semantic units, which in turn activates units in the various other sets.

When a relational representation enters the driver the binding of roles to their fillers must be represented dynamically without violating their independence (i.e., it is not sufficient to represent bindings using only conjunctive units; see, e.g., Doumas & Hummel, 2005; von der Malsburg, 1999). DORA uses systematic asynchrony of firing to dynamically bind roles to their fillers (see Doumas et al., 2008). As a relational representation in the driver becomes active, bound objects and roles fire in direct sequence. Information about role-filler bindings is carried by proximity of firing (e.g., with roles firing directly before their fillers). This sequence-based binding keeps roles and their fillers distinct and thus independent. Using the example in Figure 1, in order to bind *bigger* to block and *smaller* to ball (and so represent *larger* (block, ball)), the units corresponding to *bigger* fire directly followed by the units corresponding to block, followed by the units for coding *smaller* followed by the units for ball.

Mapping DORA uses LISA's mapping algorithm (see Hummel & Holyoak, 1997; Doumas et al., 2008). DORA learns *mapping connections* between units of the same type in the driver and recipient (e.g., between PO units in the driver and PO units in the recipient). These connections grow whenever corresponding units in the driver and recipient are active simultaneously. The connections act as mappings between corresponding structures in separate analogs. They also permit correspondences learned in mapping to influence correspondences learned later.

Relation Learning DORA uses comparison to isolate shared properties of objects and to represent them as explicit structures. DORA begins with simple feature-vector representations of objects (i.e., a node connected to a set of semantic features describing that object). When DORA compares two objects, the two representations are activated simultaneously. For instance, if DORA compares a block that is larger than some object to a plate that is larger than some other object (e.g., when the block is larger than a ball and the plate is larger than a fork), then the nodes representing the block and plate fire together (Figure 2a). Semantic features shared by the compared objects (i.e., features common to the block and the plate) receive twice as much input and thus become roughly twice as active as features connected to one but not the other (Figure 2b). DORA then learns connections between a newly recruited PO unit and active semantic units via Hebbian learning (Figure 2c). In Hebbian learning the strength of a learned connection is a function of unit activation (i.e., stronger connections are learned to more active units). Consequently, the new PO unit becomes most strongly connected to the highly active semantic units. The new PO becomes an explicit representation of the feature overlap between the block and plate. In this example, DORA forms an explicit representation of the semantics of *bigger* things (i.e., the features common to both the block and plate). The new PO functions as a predicate representation of *bigger* because it can be dynamically bound to fillers via an RB unit (Figure 2d).

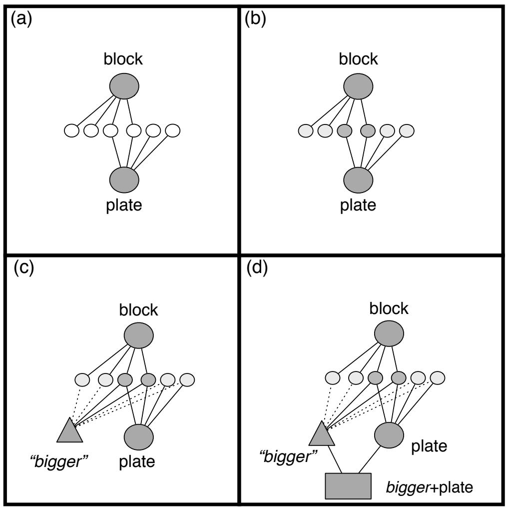

Figure 2. Comparison-based predication in DORA. DORA learns a representation of bigger by comparing a block that is bigger than some object to a plate that is bigger than some other object. (a) DORA compares a block and a plate. Units representing both become active. (b) Feature units shared by the block and the plate become more active than unshared features (darker grey). (c) A new PO unit learns connections to features in proportion to their activation (solid lines indicate stronger connection weights). The new unit codes the featural overlap of the block and plate (i.e., the role "*bigger*"). (d) This new PO unit functions as a predicate when dynamically bound to fillers.

DORA learns representations of multi-place relations by linking sets of co-occurring role-filler pairs into hierarchical relational structures. Continuing the example, when DORA compares a plate that is larger than a fork to a block that is larger than a ball, it will map *larger* (plate) to *larger* (block) and *smaller* (fork) to *smaller* (ball) (Figure 3a). When constituent sets of role-filler pairs are mapped, a distinct pattern of firing emerges—namely, mapped RB units fire together and out of synchrony with any other RB units; Figure 3b-d). This pattern is a reliable signal that DORA exploits to combine sets of role-filler pairs into multi-place relations. In response to the pattern, DORA recruits a P unit that learns connections to any active RB units in the recipient (Figure 3e-g) via Hebbian learning. The result is a P unit linking the RB units in the recipient into a complete relational structure (*larger* (block, ball); Figure 3i).

Producing invariant responses for basic SDML

A comparison-based solution to the problem of learning an invariant feature coding for "more", "less", and "same" requires the assumption that initially available magnitude information is coded by a direct neural proxy: All else being equal, higher magnitude items are coded (at least early in processing) by more neurons than comparatively lower magnitude items. For example, a larger item will be coded by more neurons than a smaller item. There is a preponderance of evidence for this assumption. In visual processing, larger items take up more space on the retina (e.g., Wandell, 1995) and are coded by larger swaths of the visual cortex (e.g., Engel et al., 1994).

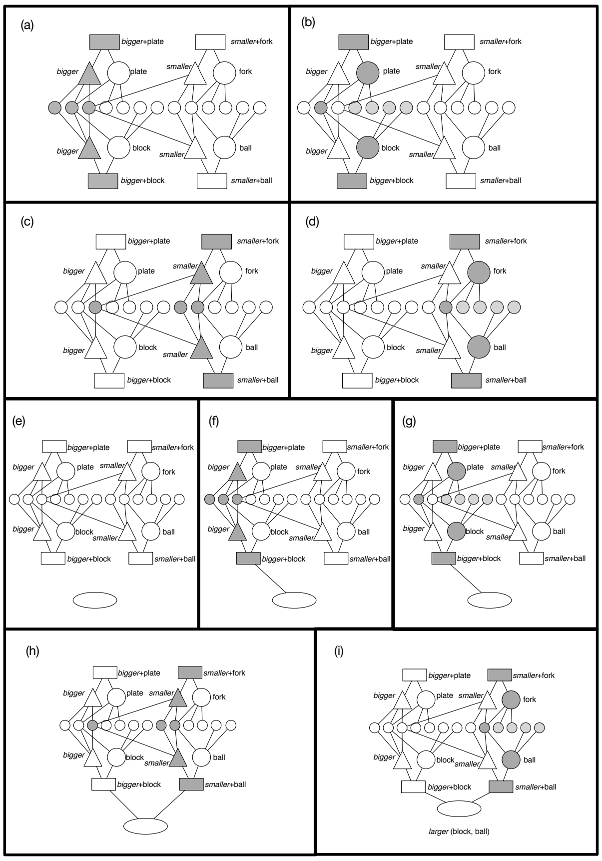

Figure 3. DORA learns a representation of the whole relation *larger* (block, ball) by mapping *bigger*(plate) to *bigger*(block) and *smaller*(fork) to *smaller*(ball). (a) The units coding *bigger* fire; (b) the units for plate and block fire; (c) the units for *smaller* fire; (d) the units for fork and ball fire. (e) DORA recruits a P unit in the recipient. (f-g) DORA learns a connection between the new P unit and the active RB unit (the unit coding for *bigger*(block)). (h-i) The P unit learns connections to the active RB unit (coding for *smaller*(ball)). The result is a structure coding for *larger*(block, ball).

Basic magnitude calculation is accomplished by comparison. When the model attends to two representations with specific magnitude values (e.g., two POs attached to absolute size are present in the driver together; Figure 4a), the representations of the absolute magnitude semantics are co-activated and the PO units attached to these semantic units compete via lateral inhibition (Figure 4b). The POs will eventually settle, with either one PO becoming more active and inhibiting the other to inactivity, or, when both POs code for the same absolute magnitude, with both POs in a steady state of co-activation. More semantic units can then respond to the particular pattern of firing in the driver POs. Some units are excited by two active POs in the driver, others are excited by a single highly active PO early in firing, or by a single highly active PO late in firing (these regions of excitement are easily learnable via simple neural threshold tuning). The active POs learn connections to the active semantic unit by Hebbian learning. If a single PO is active, that unit will learn connections to the semantics that are activated by a single highly active driver PO early in firing (which becomes the invariant signal for "more"; Figure 4c). When the active PO becomes inhibited (because of asynchronous binding), the second PO (the one inhibited by the winning PO) will become active (Figure 4d). That unit learns connections to the semantics that are activated by a single highly active driver PO late in firing (which becomes the invariant signal for "less"; Figure 4d). Otherwise, if two POs are co-active (i.e., they code the same magnitude), then they will learn connections to the semantics which are activated by two active driver POs (which becomes the invariant signal of "sameness".

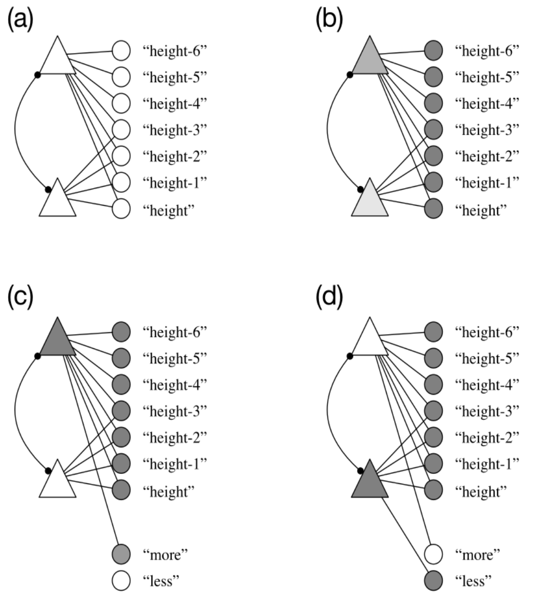

Figure 4. The SDML detector working on POs coding different values on a dimension. For the purposes of clarity, only the predicate POs and their semantics are depicted in this figure. (a) Two POs coding for different heights are in the driver. (b) The semantics coding for absolute dimensional information become active and the two POs compete to become active. (c) The unit coding for the greater value on the dimension (here height-6) becomes active first, thus marking it as "more". The PO learns a connection to the semantic that responds to winning the SDML competition (i.e., the invariant of "more"). (d) The unit coding for the lesser value on the dimension (here height-3) will become active last, thus marking it as "less". The predicate is connected to the semantic unit coding for losing the SDML competition, or the invariant of "less".

In short, comparing different magnitudes in a network in which magnitude information is coded by an absolute proxy (as in the human neural system) produces one of three patterns. (1) Both units settle into a state of similar co-activation—which occurs when two representations of the same magnitude are compared. (2) One unit becomes more active and forces the second unit to inactivity—which occurs when a unit codes for a greater magnitude. (3) One unit becomes active after it has been inhibited by a winning unit—which occurs when a unit codes for a lesser magnitude. Whatever units respond to these patterns naturally or through tuning become implicit invariant codes for the presence of "sameness", "moreness", and "lessness", respectively. Vitally, the same patterns will emerge and the same codes will become active when specific relative magnitudes are present even cross dimensionally. That is, the same patterns emerge and units become active during an instance of different absolute height, or width, or colour. What is left for the system is to learn explicit representations of these invariant semantics that are not tied to any specific magnitudes (e.g, a PO connected to semantics encoding 'more' & 'height', without strong connections to any specific height) and can take other POs as arguments. In other words, exactly the learning that DORA does.

Simulations

Simulation 1

We tested whether DORA could learn structured representations of relative SDML relations starting with information about sets of shapes with features representing absolute values on dimensions. This simulation mirrored what happens during development when a child learns from experience without a teacher or guide.

The model began with pixel images of basic shapes (differing in shape, colour, size, width, and height). These images were pre-processed with a feedforward neural network that learned via back-propagation to deliver absolute shape, colour, size, width, and height information (akin to that information delivered by early visual processing). Each processed image was represented by a PO attached to the delivered features. In addition, each shape was also attached to a set of 10 extraneous features selected randomly from a set of 100 features, included as noise (as objects in the world contain several features extraneous to any particular learning goal). Each shape was then randomly paired with another to create pairs of shapes over which relations were learned. We created 100 pairs of objects in this manner and placed them in DORA's LTM.

We then allowed DORA to attempt to learn from these basic representations. On each learning trial, DORA selected one pair of objects from LTM at random and ran (or attempted to run) retrieval, mapping, SDML comparison, predication, and multi-place relation learning, and stored any representations that it learned in LTM. In short, we are testing whether unguided learning from simple shape objects is sufficient for DORA to learn structured representations of relative SDML relations.

We defined a relational quality metric as the mean of connection weights to relevant features (i.e., those defining a relative magnitude on some specific dimension (e.g., 'more'+'height', or 'less'+'width')) divided by the mean of all other connection weights + 1 (1 was added to the mean of all other connection weights to normalize the quality measure to between 0 and 1). A higher quality denoted stronger connections to the semantics defining a specific SDML relation relative to all other connections. We measured the relational quality of the last 100 items DORA had learned after each 100 learning trials for 1000 total learning trials. Importantly, we tested all representations that the model learned (not just those that instantiated the relevant relations) and included these in the relational selectivity calculation.

Figure 5 shows the quality of the representations that DORA learned. DORA learned representations of whole relational structures encoding relative magnitudes and similarity on all the encoded dimensions. DORA learned representations of *bigger* (one predicate PO connected most strongly to the semantics 'more' & 'size', the other connected to 'less' & 'size'), *wider* (predicate POs connected to 'more' & 'width, and 'less' & 'width'), *taller* (predicate POs connected to 'more' & 'height, and 'less' & 'height), *same-size* (predicate POs both connected most strongly to 'same' & 'size;), *same-width* (predicate POs both connected most strongly to 'same' & 'width'), *same-height* (predicate POs both connected most strongly to 'same' & 'height'), *same-colour* (predicate POs both connected most strongly to 'same' & 'colour'), and *same-shape* (predicate POs both connected most strongly to 'same' & 'shape'). The results indicate that DORA can learn structured representations of relative SDML relations from objects that include only absolute values on dimensions even with the addition of extraneous noise.

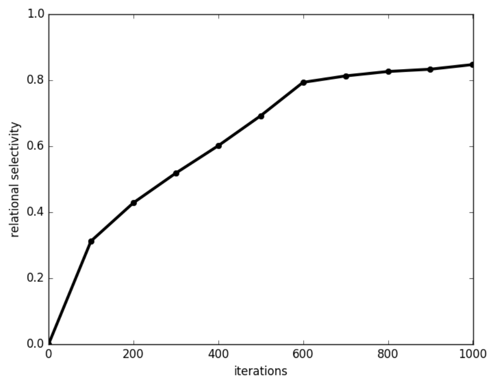

Figure 5. Results of DORA's learning.

Simulation 2

A crucial question remains: do the representations DORA learns meet the requirements of relational representations? Some hallmark of relational representations (see Holyoak, 2012) are that they, (i) form the basis of solving cross mappings; (ii) support mapping similar, but non-identical predicates; and (iii) form the basis of overcoming the n-ary restriction.

During cross-mapping, an object (object1) is mapped to a featurally less similar object rather than a featurally more similar object because it (object1) plays the same role as the less similar object. Cross-mappings serve as a stringent test of the structure sensitivity of a representation as they require violating featural or statistical similarity.

We tested the relations that DORA had learned in the previous simulations for their ability to support finding cross-mappings. We selected two of the refined relations that DORA had learned during the previous simulation at random. We bound the relations to new objects, creating two new propositions, P1 and P2 such that the agent of P1 was semantically identical to the patient of P2 and patient of P1 was semantically identical to the agent of P2, and allowed DORA to attempt to map P1 and P2. We repeated this procedure 10 times, each time with a different randomly-chosen pair of relations. All 10 times DORA successfully mapped the agents and patients of P1 and P2. The relations DORA learned in the first simulation satisfy the requirement of cross-mapping.

We also tested whether the relations that DORA has learned would support mapping to similar but non-identical relations (such as mapping *higher* to *greater-than*). Humans successfully map such relations (e.g., Bassok, Wu, & Olseth, 1995; Gick & Holyoak, 1983), an ability that Hummel and Holyoak (1997, 2003) have argued depends on the semantic-richness of relational representations. We selected one of the refined relations that DORA had learned during the previous simulation, R1, and constructed a new relation, R2, that shared 50% of its semantics (in each role) with the selected relation. So that mappings could not be based on object similarity, none of the objects that served as arguments of the relations had any semantic overlap. We repeated this process 10 times. Each time, DORA mapped the agent role of R1 to the agent role of R2 and the patient role of R1 to the patient role of R2, and, despite their lack of semantic overlap, corresponding objects always mapped to one another (because of their bindings to mapped roles).

Finally, we tested the model's ability to find mappings that violate the *n-ary restriction*: the restriction that an *n*-place predicate may not map to an *m*-place predicate when $n \neq m$. Almost all models of structured cognition follow the *n*-ary restriction (namely, those that represent propositions using traditional propositional notation and its isomorphs; see Doumas & Hummel, 2005). However, the restriction does not appear to apply to human reasoning, as evidenced by our ability to easily find correspondences between *bigger* (Sam, Larry) on one hand, and *small* (Joyce), *big* (Susan), on the other (Hummel & Holyoak, 1997).

To test DORA's ability to violate the n-ary restriction, we randomly selected a refined relation (R1) that DORA had learned in the previous simulation. We then created a single place predicate (r2) that shared 50% of its semantics with the agent role of R1 and none of its semantics with the patient role. The objects bound to the agent and patient role of R1 each shared 50% of their semantics with the object bound to r2. DORA attempted to map R1 to r2. We repeated this process 10 times, and each time DORA successfully mapped the agent role of R1 to r2, along with their arguments. We repeated the simulation such that r2 shared half its semantic content with the patient (rather than agent) role of R1. In 10 additional simulations, DORA successfully mapped the patient role of R1 to r2 (along with their arguments). In short, in all our simulations DORA overcame the *n*-ary restriction, mapping the single-place predicate r2 onto the most similar relational role of R1.

Conclusion

We have shown how structured relational representations of magnitude and similarity can be learned from objects with only absolute magnitude values. Our model exploits regularities that emerge in a connectionist network when distributed representations are compared or co-activated. These regularities serve as invariant signals that the model can learn to exploit to bootstrap the detection of relative magnitude differences and similarities. When linked with the DORA predicate learning algorithm, the system learns structured predicate representations of these relative magnitudes and similarities, and then can exploit the resulting representations to solve problems.

Our account provides a trajectory for similarity cognition that maps to cognitive complexity across species and maturational trajectories in humans. This trajectory reveals three distinct levels of abstraction in SDML computation; (i) implicit detection of SDML (responding based on the regular firing that occurs when absolute magnitudes are compared), (ii) implicit generalization of SDML (or learning based on the presence or absence of a particular feature; e.g., learning to respond based on the presence or absence of the 'more' feature), and (iii) predicate representations of SDML (or full-fledged relational representations that support complex cognitive capacities like analogy and reasoning).

This distinction may explain why humans solve some tasks involving similarity judgments without the extensive training that other animals require (e.g., Young, Wasserman, & Garner, 1997). Humans may solve the task relationally rather than relying on generalized implicit similarity judgments.

Many cognitive architectures and task models rely on stimulus recognition. This theory explains how stimulus recognition might be computed. We believe that providing a computational account for a function existing models depend on represents a significant architectural contribution.

References

Bassok, M., & Olseth, K. L. (1995). Object-based representations: Transfer between cases of continuous and discrete models of change. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(6), 1522.

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. (2013). Comparison and mapping facilitate relation discovery and predication. *PloS one*, *8*(6), e63889.

Doumas, L. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological review*, *115*(1), 1-43

Engel, S.A., Rumelhart, D.E., Wandell, B.A., Lee, A.T., Glover, G.H., Chichilnisky, E.J., & Shadlen, M.N. (1994). fMRI of human visual cortex, *Nature, 369*, 525.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial intelligence*, *41*(1), 1-63.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive psychology*, *15*(1), 1-38.

Goodale, M. A., Ellard, C. G., & Booth, L. (1990). The role of image size and retinal motion in the computation of absolute distance by the Mongolian gerbil (Meriones unguiculatus). *Vision Research*, *30*(3), 399-413.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, *21*(06), 803-831.

Holyoak, K. J. (2012). Analogy and relational reasoning. *The Oxford handbook of thinking and reasoning*, 234-259.

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps*.

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

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

Kemp, C. (2012). Exploring the conceptual universe. *Psychological review*, *119*(4), 685.

Kemp, C., & Tenenbaum, J. B. (2009). Structured statistical models of inductive reasoning. *Psychological review*, *116*(1), 20.

Kim, K. U., Huh, N., Jang, Y., Lee, D., & Jung, M. W. (2015). Effects of fictive reward on rat's choice behavior. *Scientific reports*, *5*, 8040.

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

Lu, H., Chen, D., & Holyoak, K. J. (2012). Bayesian analogy with relational transformations. *Psychological review*, *119*(3), 617.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological review*, *100*(2), 254.

Von der Malsburg, C. (1999). The what and why of binding: the modeler's perspective. *Neuron*, *24*(1), 95-104.

Wandell, B. (1995). *Foundations of Vision*. Sinaur Associates Inc.: Sunderland, MA.

Young, M. E., Wasserman, E. A., & Garner, K. L. (1997). Effects of number of items on the pigeon's discrimination of same from different visual displays. *Journal of Experimental Psychology: Animal Behavior Processes*, *23*(4), 491.