Available online at www.sciencedirect.com
ScienceDirect
Cognitive Systems Research 11 (2010) 204–207
Cognitive Systems Research
www.elsevier.com/locate/cogsys

Book review

Rogers, T. T., & McClelland, J. L. Review of semantic cognition: A parallel distributed processing approach. MIT Press

Action editor: Stefan Wermter

J.F. Glazebrook

Department of Mathematics and Computer Science, Eastern Illinois University, 600 Lincoln Avenue, Charleston, IL 61920–3099, USA

Available online 9 October 2009

This book has its origins in some of the earlier studies concerning the organization of knowledge in semantic memory, in particular, that of Quillian (1968) who showed that the taxonomic hierarchy in such organizations can provide an efficient means for storage and retrieval of (semantic) information. One key feature was that category membership at each level of the hierarchy can involve several properties shared by all members of the more specific subcategories. Applications of Quillian's findings were taken up by a number of researchers, such as Warrington (1975) who suggested that deficits in semantic cognition, as seen in cases of fluent aphasia (more recently referred to as *semantic dementia*), may arise when only selection at the higher level of taxonomy seemed possible, at the cost of foresaking finer level categorical specifications. In other words, a somewhat reverse procedure to infant learning and recognition patterns was observed whereby in the latter case, global category structures were the ones that were first grasped, followed by a progressive descent through the taxonomic hierarchy to achieve an awareness of various specifications (see e.g. Keil, 1979; Mandler, 1988).

However, in the Quillian model there arose 'distance issues' of general properties which were seen as more strongly bound compared to specifics, although a similar model was adopted by Hinton (1981, 1986) for storing propositional knowledge in a distributed connectionist network. Besides seeking useful representations, Hinton's modifications included back propagation trained networks in order to answer ancestral questions for a given number of families, and to some extent created the genesis of *Parallel Distributed Processing* (PDP) for the authors' work on a theory of semantic cognition. Following Hinton (1981) and Rummelhart et al. (1986), the origin of ideas is summarized in chap. 2 of the book (p. 55) as:

- Performance in semantic tasks occurs through the propagation of activation among simple processing units, via weighted connections.
- Connection weights encode the knowledge that determines which distributed representations arise internally, and that governs the joint use of these representations to determine the outcome of the process.
- The connection weights are learned, so that semantic cognitive abilities arise from experience.

The framework as developed from this scheme follows from Rummelhart (1990) (see also Rummelhart & Todd, 1993) who devised a simplified version of Hinton's family-tree model using a single direction 'feedforward' architecture, in part implementing a network of nonlinear processing, trained on the 'experiences' class of the Quillian hierarchy. One principle is that the network gains distributed internal representations of 'items' by assimilating properties in a variety of contexts. The task involves showing that the propositional context embedded within a Quillian-type hierarchy, could be recovered within the distributed representations as obtained in a PDP trained with back-propagation supporting those same kinds of inferences.

Briefly stated, the design of the network as adapted by the authors from the Rummelhart model, is that of a sequence of directed 'layers':

Item *Representation* *Relation*/(*Hidden*) *Attribute* trained by similar means to the 'experiences' of the Quillian model, via a series of 'epochs' or 'sweeps' that foster the learning process. The *Hidden* units are coupled to *Relations* (ISA, "is", "can", "has"). For instance, units

E-mail address: jfglazebrook@eiu.edu

doi:10.1016/j.cogsys.2009.09.001

corresponding to "canary" and "can", are activated in the input; subsequently, the network is trained to activate the outputs "move", "fly", "sing", "is yellow", etc. Cumulative error detection from properties assists inducing the eventual distinction of representational classes (such as plants in distinction to animals). The internal representations obtained from a newly introduced item, can in turn be used to make further inferences by training the network to establish other properties (once established the canary " is a bird', then "canary is, ..., has", etc.). The authors implement further 'cognitive' modifications to allow for greater category coherence and for patterns of progressive differentiation due to the back-propagation feature trained with a similarity structure. In this way, items with similar representations and having shared properties, will be learned faster than properties differentiating such items. For instance, the property "is yellow" can be true of the canary, daisy and sunfish, whereas "has wings" is true for a canary and sparrow, say. The property "has wings" is one that varies coherently along with other properties, compared to "is yellow", and the network thus picks it up faster. In other words, the in-built feature of 'coherent covariation' trains the network to represent two given birds as possessing a number of essential similarities, that make them noticeably different from e.g. two fish. Questions such as these and others comprise the first two chapters. Further insight into the framework of the methodology and methods of simulation is taken up in chap. 3.[1]

On closer examination of the behaviour of the network, it is seen that varying sensitivity to patterns of coherent covariance may actually reflect upon processes of conceptual formation, particularly observed in that of infants, a subject that is discussed in chap. 4. Moreover, the PDP mechanism suggests that conceptual attainment is influenced by 'domain-general' learning induced by perceptual experience such as when twelve-month old infants tend to make an initial grasping of external, more prominent characteristics, such as 'legs' or 'wheels' for grouping objects, and so eventually learn how to differentiate between categories. Related is the contrast between items shared in a same "basic" category and the rate at which the network can give specific labels to domains of expertise. "Basic" means here those more obvious or meaningful categories, such as when a given cat, say, may belong to categories such as 'Siamese', 'tabby', 'cat', 'animal', 'living things', in which case, 'cat' can be considered as "basic". In this respect, the authors make the reasonable claims (chap. 5, p. 176):

1. Children first learn to label objects with their basic-level name, instead of with more general or more specific names.
2. In free-naming tasks, adults prefer to name at the basic level even when they know more general or specific names.
3. Adults are faster in verifying category membership at the basic level.
4. Adults are quicker to verify properties of objects that are shared by exemplars of a "basic" category.

Accordingly, infant learning processes are often influenced by parental linguistic interactions which can coax wide semantic distinctions from more general to basic levels, aided by the fact that 'cats', for instance, are more environmentally common as pets (whereas 'elephants' are not). In the opposite direction, foresaking the basic names for more general names may be a possible symptom of (semantic) dementia. Likewise, the network detects 'expertise' as arising from increased frequency of exposure to information occurring in the expert domain, and in a related sense, the authors make some modifications to the hypothesis of Rosch (1975) concerning basic-level lexical acquisition with respect to word frequencies. In support of their claims, the authors in chap. 5 commence summarizing the simulations used in PDP, the training trials and various computational artifacts, with an ample degree of technical details.

The PDP mechanism also demonstrates overall sensitivity towards, and facility in acquiring, the natural and intuitive categorical properties ('category coherence'). Several comparisons can be made with respect to the *modus operandi* of theory–theorists, where it is debatable if theories themselves, sometimes lacking in representations, can have influence upon semantic task performance. The authors advocate the necessity of "causal theories", to an extent gauged by the weighting of attributes, in turn, reflecting upon how certain properties "hang together", cognitively speaking. For instance, under the concept of "flight", representations such as "has wings", "has feathers" and "has bones", are significant because they are causally related. In this context of 'category coherence', there is the following list of questions following Murphy & Medin (1985), to an extent summarizing the task at hand (chap. 6, p.239):

1. Why do some sets of objects seem to "hang together" in a psychologically natural way while others do not?
2. Why are some constellations of properties easier to learn and remember than others?
3. Why do children and adults sometimes draw incorrect inferences about the properties of familiar and unfamiliar objects?
4. Why are some attributes important for categorizing an object, while others are irrelevant?
5. Why might a given property be important for categorizing some objects, but irrelevant for others?

Through appropriate gearing of PDP, semantic judgements can be influenced by causal knowledge thus extend-

[1] It is worth pointing out that the PDP++ software for creating the corresponding simulations has been made freely available: O'Reilly et al., 1995. The PDP++ Software. (Carnegie Mellon University, Pittsburgh, PA, 15213. Updated versions are maintained by Randall C. O'Reilly available from http://psych.colorado.edu/oreilly/PDP++/PDP++.html.

ing that obtained via correlating the various attributes. For instance, how ‘‘living by the ocean’’ affords more specific representations for certain classes of birds, compared to ‘‘having the color white’’. In a similar spirit, the model may reveal results of ‘illusory correlations’ such as when a correct attribute might be noted serendipitously, though this attribute may in fact be disguised in an actual observation. Likewise, apparent alterations to the Rummelhart model along with the latter’s coherent covariance of characteristics, allow the PDP to unfold a certain degree of concept ‘coalescence’. This is best explained with reference to the studies of Carey (1985) concerning how children can grasp sociobiological concepts. She hypothesized that patterns of induced projection differ significantly between the ages of 3 and 10 years. Typically, whereas young children usually make distinctions between plants and animals at an early stage, older children eventually are able to conceive and appreciate the common conceptual features of ‘life’ as involved. This is the essence of ‘concept coalescence’. In order to recover this property, the authors outline (in chap. 7) several techniques as applied to the model, such as the addition at the *Attribute* layer of new output units (referred to as ‘queems’), and making various adjustments to weights in the network. In many cases, weights emanating from the *representation* and *relation* compartments into *Hidden*, adjust slowly while carefully searching for different qualities of similarity between objects housed within different categories. The model can thus be trained to distribute the various properties (cf the findings of Gelman & Markman, 1986). However, there are certain limitations here, since regulators such as ‘‘is’’, ‘‘has’’’, and ‘‘can’’, do not by themselves account for the contextual depth required in order for the eventual attainment of concepts. Here is where the overlying causal experience of events appears to be significantly influential for the purpose of semantic judgements. Still, the operative PDP remains as a valid blueprint for developing further, more refined conceptual models.

The overall significance of the role of causal properties is discussed further in chap. 8. Concerning ‘‘causal powers’’, a quotation from Wilson & Keil (2000) in a way sets the stage (chap. 8, p. 299):

‘‘ It may seem as though ‘‘causal powers’’ is another name for a property or object, but the real sense seems more one of an interaction between a kind of thing and the world in which it is situated. Thus we can understand and explain something in terms of its causal powers, which means not just listing its properties as sets of things attached to it, but rather listing its *disposition to behave in certain ways in certain situations.*’’

It is in this chapter that the authors address certain comparisons and differences with the approach of theory-theorists. Whereas the authors agree about the general sensitivity of the causal nature of events and objects, and the use of verbal explanations for the purpose of judging semantic tasks, they do not fully accept that causal knowledge is in any way privileged or special. Instead, the claim is that causal properties have a good chance of arising from those same learning mechanisms that bring awareness to other properties. With this in mind, the authors propose with details how the network architecture of Rummelhart’s model can be extended beyond input–output pairs to those of situation-outcome, thus capturing to an extent the mode of domain-general learning and specifically events involving verbal communication, listening, temporal context, etc. (for instance, the network’s familiarity with the behaviour of a conscientious waiter in a restaurant, leads to concluding that this same person should later present the bill, and not the individual that has eaten the meal and then pays for it). One factor is that at its present stage of development, the PDP model is not in itself a mechanism for unfolding one theory after another. Overall, the authors seek for a balance for instrumentation of causal properties; some are easier to learn than others. For instance, how ‘‘white’’ is causally central to polar bears as a property linked to survival characteristics, whereas it is not considered casually central to the concept of ‘‘refrigerator’’.

Many of the above considerations lead the authors to account for the core principles of the model and to create a basis for obtaining insight into further redesigning of the Rummelhart model. These can be summarized by the authors’ own dictum as follows (chap. 9, pp. 348–367):

1. (*Projective-Error Driving Learning*) Adjust each parameter of the mind in proportion to the extent that its adjustment will reduce the discrepancy between predicted and observed events.
2. (*Sensitivity to Coherent Covariation*) Coherent covariation of properties across items and contexts drives conceptual differentiation and determines what properties are central and what properties are incidental to different concepts.
3. (*Similarity-Based Distributed Representation*) Use similarity-based distributed representations to promote generalization and sensitivity to patterns of coherent covariation.
4. (*Convergence Principle*) Organize the processing and representation of the objects of thought so that all different sorts of information about all objects, in all contexts, converge on the same units and connections.
5. (*Gradual, Structure-Sensitive Learning*) Adjust the weights on the connections slowly, over the course of many experiences sampled from a cognitive domain, so that they will be sensitive to the overall structure of experience.
6. (*Activation-Based Representation of Novel Objects*) Use a pattern of activation constrained by prior semantic learning to represent a novel object, to avoid disturbing the knowledge that has built up gradually in connections.

There are other issues to be addressed with the intention of making improvements to the model. One way is to look at why certain connectionist models fall short of being optimal and do not generalize in a suitable way. The roots of the matter may be in that of 'architecture dependence' (cf Fodor, 2000) and a tradition for some models to incorporate Bayesian inference techniques. One possible way is to compare with those cognitive systems that rely upon module-like structures geared to managing evolutionary factors and projective representations in tune with the task of deriving semantic structure from experience, while underscoring the importance of the convergence principle. Also, one of the most interesting tasks is to further explore how the model can be adapted to study certain neurophysiological processes by virtue of the fact that two complementary learning systems are basic to the acquisition of knowledge. The hypothesis is that at the core of the model, the slow-learning semantic/conceptual system may be associated to the neural processes of the brain's neocortex, and the complementary fast-learning system for registering new information, may be associated with the medial temporal regions. Thus the authors suggest that semantic processing of information spans several regions of the brain while listing an ample number of selected references supporting this claim. In particular, those relating to the various senses and how, for instance, fMRI images typically reveal certain common patterns found in forms of brain damage. Needless to say, there is a huge amount of literature on this latter topic that time and space would not permit including (I would imagine that would amount to including at least two additional chapters). In any event, this particular direction for application of the techniques of PDP remains an important and ambitious task.

Altogether, the book represents a solid, first-rate account of the authors' research and development of the Rummelhart model over a period of years and provides a concise background to this particular field of research while at the same time making apt comparisons with the significant work that has been achieved by other authors. It is also a well written book with the methodology clearly outlined; there are final sections that provide details of simulation and training patterns. Those researchers working in the connectionist and related fields of research ought to find the book to be a worthwhile companion volume, and one which may eventually prove to be a valuable contribution towards the ontology of cognition.

References

Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Fodor, J. (2000). *The mind doesn't work that way: The scope and limits of computational psychology*. Boston, MA: MIT Press/Bradford Books.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23*, 183–209.

Hinton, G. E. (1981). Implementing semantic networks in parallel hardware. In G. E. Hinton & J. A. Anderson (Eds.), *Parallel models of associative memory* (pp. 161–187). Hillsdale, NJ: Erlbaum.

Hinton, G. E. (1986). Learning distributed representations of concepts. In *Proceedings of the eighth annual conference of the cognitive science society* (pp. 1–12). Hillsdale, NJ: Erlbaum.

Keil, F. (1979). *Semantic and conceptual development: An ontological perspective*. Cambridge, MA: Harvard University Press.

Mandler, J. M. (1988). How to build a baby: On the development of an accessible representational system. *Cognitive Development, 3*, 113–136.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review, 92*, 289–316.

Quillian, M. R. (1968). Semantic memory. In M. Minsky (Ed.), *Semantic information processing* (pp. 227–270). Cambridge, MA: MIT Press.

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*, 192–233.

Rummelhart, D. E. (1990). Brain style computation: Learning and generalization. In S. F. Zornetzer, J. L. Davis, & C. Lau (Eds.), *An Introduction to Neural and Electronic Networks* (pp. 405–420). San Diego, CA: Academic Press.

Rummelhart, D. E., & Todd, P. M. (1993). Learning and connectionist representations. In D. E. Meyer & S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence and cognitive neuroscience* (pp. 3–30). Cambridge, MA: MIT Press.

Rummelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. E. (1986). Schemata and sequential thought processes in PDP models. In J. L. McClelland, D. E. Rummelhart, & the PDP research group (Eds.). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 2, pp. 7–57). Cambridge, MA: MIT Press.

Warrington, E. K. (1975). Selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology, 27*, 635–657.

Wilson, R. A., & Keil, F. C. (2000). The shadows and shallows of explanation. In F. C. Keil & R. Wilson (Eds.), *Explanation and cognition* (pp. 87–114). Cambridge, MA: MIT Press.