Vector symbolic architectures are a viable alternative for Jackendoff's challenges

Ross W. Gayler

r.gayler@mbox.com.au

Abstract: The authors, on the basis of brief arguments, have dismissed tensor networks as a viable response to Jackendoff's challenges. However, there are reasons to believe that connectionist approaches descended from tensor networks are actually very well suited to answering Jackendoff's challenges. I rebut their arguments for dismissing tensor networks and briefly compare the approaches.

Van der Velde and de Kamps (V&K) have proposed neural blackboard architectures (NBAs) in response to Jackendoff's (2002) challenges. Their note 1 dismisses tensor networks (Smolensky 1990) as a viable alternative. However, Gayler (2003) argues that vector symbolic architectures (VSAs)—connectionist approaches descended from tensor networks—are very well suited to answering Jackendoff's challenges. There is not space here to repeat those arguments. Rather, I will rebut V&K's arguments for dismissing tensor networks and briefly compare the approaches. Regardless of the ultimate relative success of NBAs and VSAs, the field of cognitive neuroscience will benefit from having plausible alternatives that can be compared and contrasted.

V&K's note 1 claims that "tensor networks fail to instantiate combinatorial structures." Tensor product binding was developed specifically to address the issue of combinatorial representation. Although Smolensky's (1990) presentation was somewhat abstract and mathematical, he chose his primitive representations and operations to be easily implemented in connectionist systems. Items are represented by numerical vectors (distributed representation) and the operators are simple vector sums and products. Smolensky demonstrated tensor network implementation of variables and binding and also, as an example of the computational power of tensor networks, demonstrated a connectionist implementation of the CONS, CAR, and CDR operators of the LISP programming language. Halford and colleagues (1998) discuss cognitive symbolic operations that can be implemented in tensor networks.

V&K's note 1 cites Fodor and McLaughlin (1990) in support of the claim that "tensor networks fail to instantiate combinatorial structures," but this is not the focus of their paper. Their focus is Fodor's and Pylyshyn's (1988) claim that connectionism either cannot account for cognition or, if it can, is a mere implementation of a "classical" theory. Fodor and McLaughlin (1990) is a response to attempts by Smolensky to rebut Fodor and Pylyshyn (1988). In that context, Fodor and McLaughlin concede that "It's not . . . in doubt that tensor products *can represent* constituent structure" (p. 200).

Fodor and McLaughlin argue that although tensor networks represent constituent structure, the constituents are not causally effective in processing of the composite. They claim (correctly) that "When a tensor product vector . . . is tokened, its components are not" (1990, p. 198) and (incorrectly) that "Merely counterfactual representations have no causal consequences; only actually tokened representations do" (p. 199). Binding and unbinding by a tensor network (Smolensky 1990) suffices as a counterexample to demonstrate the falsity of the latter claim.

V&K claim that "tensor networks fail to instantiate combinatorial structures . . . [because] a tensor is just a list of constituents organized in a particular fashion (i.e., as an n-dimensional list for a rank-n tensor)" (note 1). However, exactly the same claim could be made of the storage tape of a Turing machine or the CONS cells of a LISP program, yet no one would dispute the ability of a Turing machine to represent combinatorial structures.

The final argument in note 1 is the most significant; V&K point out that "adding constituents to the tensor increases the dimensions of the tensor, which requires adjustments to all components in the cognitive system that can interact with the tensor." This is the major problem with tensor networks as an implementation rather than an abstract formalism. However, there have been 15 years of further development since Smolensky (1990), and this problem was soon solved (Plate 1991).

In a tensor network, two items are bound by forming the outer product of the vectors representing the items. That is, if each primitive item is represented by a vector of n elements, their combination contains n^2 elements. If this composite item were bound with another primitive item, the result would contain n^3 elements, and so on. This has major practical consequences for resource requirements and the connections between processing components of a tensor network. The number of elements increases dramatically with the binding order, which means that the resource requirements may be excessive and that bindings of arbitrarily high order cannot be represented on a fixed set of resources. Also, the connections between processing components must be dimensioned to accommodate the highest order representation chosen by the designer.

An abstract solution to this problem was proposed by Hinton (1988; 1990) and a specific, practical implementation demonstrated by Plate (1991; 1994; 2003). Hinton introduces the idea of “reduced descriptions” as a means to represent compositional structures in fixed size connectionist architectures. A reduced description is a representation of a composite item that is the same size as any of the component items and from which the component items can be generated. Plate demonstrates holographic reduced representations (HRRs) as a specific implementation of reduced descriptions. HRRs can be conceptualised as a compression of the tensor product to a vector the same size as each of the components being bound.

HRRs use a specific, highly ordered compression of the tensor product. However, Wilson and Halford (1994) show that the majority of tensor product elements can be destroyed without compromising performance, and Plate (2000; 2003) shows that many alternative compressions of the tensor product, including randomly disordered compressions, would suffice. Other compressions were developed independently (Gayler 1998; Gayler & Wales 1998; Kanerva 1994; 1997; Rachkovskij & Kussul 2001). Collectively, these are the vector symbolic architectures, and various members of this family are compared in Gayler (1998) and Plate (1997).

Having responded to the arguments of note 1, I feel that a brief comparison of NBAs and VSAs is worthwhile. NBAs use localist representation, so effort is required to connect the representations that need to interact. VSAs use distributed representations, so many representations are simultaneously present over the same connectionist units. This replaces the problem of connecting the interacting representations with the problem of keeping most superposed representations logically separated to avoid unwanted interactions. Another difference is the treatment of novel items. Localist representation requires the recruitment of neural units to new assemblies, whereas distributed representation requires only a novel pattern which is implemented over the same units as familiar patterns. VSAs do not require a pool of unallocated units or a recruitment process. These two comparisons suggest the breadth of issues that can be explored in the range of possible connectionist cognitive architectures spanned by NBAs and VSAs.

ACKNOWLEDGEMENTS

My thanks go to Graeme Halford, Pentti Kanerva, Steven Phillips, Tony Plate, and Roger Wales for their helpful comments.

REFERENCES

Fodor, J. & McLaughlin, B. P. (1990) Connectionism and the problem of systematicity: Why Smolensky's solution doesn't work. *Cognition* 35: 183–204.

Fodor, J. A. & Pylyshyn, Z. W. (1988) Connectionism and cognitive architecture: A critical analysis. *Cognition* 28: 3–71. Available at http://citeseer.ist.psu.edu/446398.html.

Gayler, R. W. (1998) Multiplicative binding, representation operators, and analogy [Abstract of poster]. In: *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences*, ed. K. Holyoak, D. Gentner & B. Kokinov. New Bulgarian University. Available at http://cogprints.org/502/.

Gayler, R. W. (2003) Vector symbolic architectures answer Jackendoff's challenges for cognitive neuroscience. In: *Proceedings of the Joint International Conference on Cognitive Science*, ed. P. P. Slezak. University of New South Wales. Available at http://cogprints.org/3983/.

Gayler, R. W. & Wales, R. (1998) Connections, binding, unification, and analogical promiscuity. In: *Advances in analogy research: Integration of theory and data from the cognitive, computational, and neural sciences,* ed. K. Holyoak, D. Gentner & B. Kokinov New Bulgarian University. Available at http://cogprints.org/500/.

Halford, G., Wilson, W. H. & Phillips S. (1998) Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences* 21: 803–31. Available at http://www.bbsonline.org/documents/a/00/00/05/50/index.html.

Hinton, G. E. (1988) Representing part-whole hierarchies in connectionist networks. In: *Proceedings of the Tenth Annual Conference of the Cognitive Science Society,* pp. 48–54, Erlbaum.

Hinton, G. E. (1990) Mapping part-whole hierarchies into connectionist networks. *Artificial Intelligence* 46: 47–75. Available at http://dx.doi.org/10.1016/0004-3702(90)90004-J

Jackendoff, R. (2002) *Foundations of language: Brain, meaning, grammar, evolution.* Oxford University Press.

Kanerva, P. (1994) The spatter code for encoding concepts at many levels. In: *ICANN '94, Proceedings of International Conference on Artificial Neural Networks,* pp. 226–29, ed. M. Marinaro and P.G. Morasso. Springer-Verlag. Available at http://citeseer.ist.psu.edu/666731.html.

Kanerva, P. (1997) Fully distributed representation. In: *Proceedings of the 1997 Real World Computing Symposium (RWC'97, Tokyo).* Real World Computing Partnership. Available at http://citeseer.ist.psu.edu/247558.html.

Plate, T. A. (1991) Holographic Reduced Representations: Convolution algebra for compositional distributed representations. In: *Proceedings of the 12th International Joint*

Conference on Artificial Intelligence, ed. J. Mylopoulos & R. Reiter, Morgan Kaufmann. Available at http://www.d-reps.org/tplate/papers/plate.ijcai91.pdf.

Plate, T. A. (1994) *Distributed representations and nested compositional structure.* Doctoral dissertation. Department of Computer Science, University of Toronto, Toronto, Canada. Available at http://citeseer.ist.psu.edu/90353.html.

Plate, T. A. (1997) A common framework for distributed representation schemes for compositional structure. In: *Connectionist systems for knowledge representation and deduction,* ed. F. Maire, R. Hayward & J. Diederich. Queensland University of Technology Press. Available at http://citeseer.ist.psu.edu/32374.html.

Plate, T. A. (2000). Randomly connected sigma-pi neurons can form associator networks. *Network: Computation in Neural Systems* 11: 321–32. Available at http://www.d-reps.org/tplate/papers/network00.pdf.

Plate, T. A. (2003) *Holographic reduced representation: Distributed representation for cognitive structures.* CSLI.

Rachkovskij, D. A. & Kussul, E. M. (2001) Binding and normalization of binary sparse distributed representations by context-dependent thinning. *Neural Computation* 13: 411–52. Available at http://cogprints.org/1240/.

Smolensky, P. (1990) Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence* 46: 159–216. Available at http://dx.doi.org/10.1016/0004-3702(90)90007-M.

Wilson, W. H. & Halford, G. S. (1994) Robustness of tensor product networks using distributed representations. In: *Proceedings of the Fifth Australian Conference on Neural Networks (ACNN'94), Brisbane, Australia*, pp. 258–61, ed. A. C. Tsoi & T. Downs, University of Queensland. Available at http://www.cse.unsw.edu.au/~billw/reprints/wilson.lobotomy.pdf.