

Vector Symbolic Architectures: A New Building Material for Artificial General Intelligence¹

Simon D. LEVY^{a,2}, and Ross GAYLER^b

^a *Washington and Lee University, USA*

^b *Veda Advantage Solutions, Australia*

Abstract. We provide an overview of Vector Symbolic Architectures (VSA), a class of structured associative memory models that offers a number of desirable features for artificial general intelligence. By directly encoding structure using familiar, computationally efficient algorithms, VSA bypasses many of the problems that have consumed unnecessary effort and attention in previous connectionist work. Example applications from opposite ends of the AI spectrum – visual map-seeking circuits and structured analogy processing – attest to the generality and power of the VSA approach in building new solutions for AI.

Keywords. Vector Symbolic Architectures, associative memory, distributed representations, Holographic Reduced Representation, Binary Spatter Codes, connectionism

Introduction

Perhaps more so than any other sub-field of computer science, artificial intelligence has relied on the use of specialized data structures and algorithms to solve the broad variety of problems that fall under the heading of “intelligence”. Although initial enthusiasm about general problem-solving algorithms [1] was eventually supplanted by a “Society of Mind” view of specialized agents acting in concert to accomplish different goals [2], the dominant paradigm has always been one of discrete atomic symbols manipulated by explicit rules.

The strongest challenge to this view came in the 1980’s, with the emergence of connectionism [3], popularly (and somewhat misleadingly) referred to as neural networks. In contrast to the rigid, pre-specified solutions offered by “Good Old-Fashioned AI” (GOFAI), connectionism offered novel learning algorithms as solutions to a broad variety of problems. These solutions used mathematical tools like learning internal feature representations through supervised error correction, or back-propagation [4], self-organization of features without supervision [5], and the construction of content-addressable memories through energy minimization [6]. In its most radical form, the connectionist ap-

¹The authors thank Chris Eliasmith, Pentti Kanerva, and Tony Plate for useful discussion on the topics presented here, and three anonymous reviewers for helpful suggestions.

²Corresponding Author: Simon D. Levy, Computer Science Department, Washington and Lee University, Lexington, Virginia 24450 USA; E-mail: levys@wlu.edu.

proach suggested that the structure assumed by GOFAI and traditional cognitive science might be dispensed with entirely, to be supplanted by general mechanisms like sequence learning [7]. In support of the representations derived by such models, proponents cited the advantages of distributing informational content across a large number of simple processing elements [8]: distributed representations are robust to noise, provide realistically “soft” limits on the number of items that can be represented at a given time, and support distance metrics. These properties enable fast associative memory and efficient comparison of entire structures without breaking down the structures into their component parts.

The most serious criticism of connectionism held that neural networks could not arrive at or exploit systematic, compositional representations of the sort used in traditional cognitive science and AI [9]. Criticism that neural networks are in principle unable to meet this requirement was met in part by compositional models like RAAM [10]; however, RAAM’s reliance on the back-propagation algorithm left it open to criticism from those who pointed out that this algorithm did not guarantee a solution, and could be computationally intractable [11].³

In the remainder of this paper, we describe a new class of connectionist distributed representations, called Vector Symbolic Architectures (VSA), that addresses all of these concerns. VSA representations offer all of the desirable features of distributed (vector) representations (fast associative lookup, robustness to noise) while supporting systematic compositionality and rule-like behavior, and they do not rely on an inefficient or biologically implausible algorithm like back-propagation. The combination of these features makes VSA useful as a general-purpose tool or building material in a wide variety of AI domains, from vision to language. We conclude with a brief description of two such applications, and some prospects for future work.

1. Binding and Bundling in Tensor Products

Systematicity and compositionality can be thought of as the outcome of two essential operations: binding and bundling. Binding associates fillers (*John, Mary*) with roles (LOVER, BELOVED). Bundling combines role/filler bindings to produce larger structures. Crucially, representations produced by binding and bundling must support an operation to recover the fillers of roles: it must be possible to ask “Who did what to whom?” questions and get the right answer.

Vector Symbolic Architectures is a term coined by one of us [13] for a general class of distributed representation models that implement binding and bundling directly. These models can trace their origin to the Tensor Product model of Smolensky [14]. Tensor-product models represent both fillers and roles as vectors of binary or real-valued numbers. Binding is implemented by taking the tensor (outer) product of a role vector and a filler vector, resulting in a mathematical object (matrix) having one more dimension than the filler. Given vectors of sufficient length, each tensor product will be unique. Bundling can then be implemented as element-wise addition (Figure 1), and bundled structures can be used as roles, opening the door to recursion. To recover a filler (role) from a bundled tensor product representation, the product is simply divided by the role (filler) vector.

³Recent results with a linear version of RAAM using principal component analysis instead of back-prop [12] show promise for overcoming this problem.

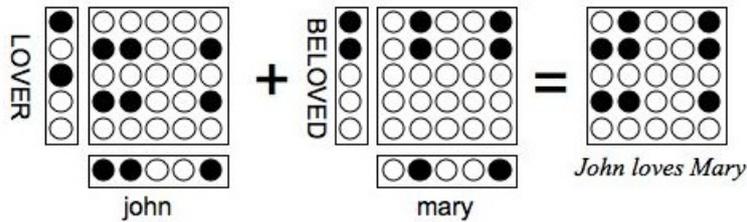


Figure 1. Tensor product representation of *John loves Mary*.

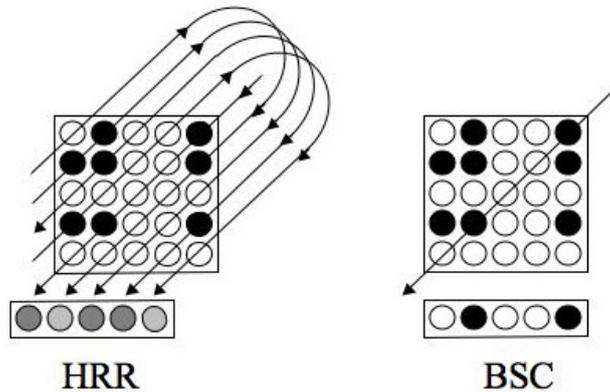


Figure 2. Methods for keeping fixed dimensionality in tensor-product representations.

2. Holographic Reduced Representations and Binary Spatter Codes

Because the dimension of the tensor product increases with each binding operation, the size of the representation grows exponentially as more recursive embedding is performed. The solution is to collapse the $N \times N$ role/filler matrix back into a length- N vector. As shown in Figure 2, there are two ways of doing this. In Binary Spatter Coding, or BSC [15], only the elements along the main diagonal are kept, and the rest are discarded. If bit vectors are used, this operation is the same as taking the exclusive or (XOR) of the two vectors. In Holographic Reduced Representations, or HRR [16], the sum of each diagonal is taken, with wraparound (circular convolution) keeping the length of all diagonals equal. Both approaches use very large ($N > 1000$ elements) vectors of random values drawn from a fixed set or interval.

Despite the size of the vectors, VSA approaches are computationally efficient, requiring no costly backpropagation or other iterative algorithm, and can be done in parallel. Even in a serial implementation, the BSC approach is $O(N)$ for a vector of length N , and the HRR approach can be implemented using the Fast Fourier Transform, which is $O(N \log N)$. The price paid is that most of the crucial operations (circular convolution, vector addition) are a form of lossy compression that introduces noise into the representations. The introduction of noise requires that the unbinding process employ a “cleanup memory” to restore the fillers to their original form. The cleanup memory can be imple-

mented using Hebbian auto-association, like a Hopfield Network [6] or Brain-State-in-a-Box model [17]. In such models the original fillers are attractor basins in the network's dynamical state space. These methods can be simulated by using a table that stores the original vectors and returns the one closest to the noisy version.

3. Applications

As a relatively new technology, VSA is just beginning to be used in the AI and cognitive science communities. Its support for compositional structure, associative memory, and efficient learning make it an appealing "raw material" for a number of applications. In this concluding section we review some of these applications, and outline possibilities for future work.

3.1. Representing Word Order in a Holographic Lexicon

Jones and Mewhort [18] report using a holographic / convolution approach similar to HRR, for incorporating both word meaning and sequence information into a model lexicon. Their holographic BEAGLE model performed better than (300-dimensional) Latent Semantic Analysis [19] on a semantic-distance test, and, unlike LSA, BEAGLE can predict results from human experiments on word priming.

3.2. Modeling Surface and Structural Properties in Analogy Processing

Experiments on children's ability to process verbal analogies show the importance of both surface information (who or what the sentence is about) and structural information (who did what to whom) [20]. Eliasmith and Thagard [21] have successfully modeled these properties in DRAMA, an HRR-based model of analogy processing. Because HRR and other VSA approaches support the combination of surface and structural information through simple vector addition, DRAMA is able to model both components of analogy in a single representation.

3.3. Variables and Quantification

GOFAI has excelled at representing and reasoning with universally and existentially quantified variables; e.g., $\forall x \text{ Computer}(x) \rightarrow \text{HasBuggyProgram}(x) \vee \text{Broken}(x)$. It has been known for some time, however, that human performance on such reasoning tasks differs in interesting ways from simple deductive logic [22]. Recent work by Eliasmith [23] shows that HRR encodings of logical rules yield results similar to those seen in the experimental literature.

3.4. Future Work: VSA in Visual Map-Seeking Circuits

Arathorn's Map Seeking Circuits (MSCs) [24] are recurrent neural networks for recognizing transformed images, using localist representations. We propose treating the localist MSC as an input recognition device, with the localist output values subsequently encoded into VSA representations indicating items in the agent's environment, and their spatial relationships to the agent. This would allow the representation and manipulation of multiple simultaneous items in the agent's environment.

References

- [1] Newell, A., Simon, H.A.: Gps, a program that simulates human thought. *Lernende Automaten* (1961) 109–124
- [2] Minsky, M.L.: *The Society of Mind*. Simon and Schuster (1988)
- [3] Rumelhart, D., McClelland, J., eds.: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. MIT Press (1986)
- [4] Rumelhart, D., Hinton, G., Williams, R.: Learning internal representation by error propagation. In Rumelhart, D., McClelland, J., eds.: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Volume 1. MIT Press (1986)
- [5] Kohonen, T.: *Self-Organizing Maps*. 3 edn. Springer-Verlag, Secaucus, NJ (2001)
- [6] Hopfield, J.: Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences* **79** (1982) 2554–2558
- [7] Elman, J.: Finding structure in time. *Cognitive Science* **14** (1990) 179–211
- [8] McClelland, J., Rumelhart, D., Hinton, G.: The appeal of parallel distributed processing. In Rumelhart, D., McClelland, J., eds.: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Volume 1. MIT Press (1986)
- [9] Fodor, J., Pylyshyn, Z.: Connectionism and cognitive architecture: A critical analysis. *Cognition* **28** (1988) 3–71
- [10] Pollack, J.: Recursive distributed representations. *Artificial Intelligence* **36** (1990) 77–105
- [11] Minsky, M.L., Papert, S.A.: *Perceptrons: Expanded Edition*. MIT Press (1988)
- [12] Voegtlin, T., Dominey, P.F.: Linear recursive distributed representations. *Neural Netw.* **18**(7) (2005) 878–895
- [13] Gayler, R.: Vector symbolic architectures answer jackendoff’s challenges for cognitive neuroscience. In Slezak, P., ed.: *ICCS/ASCS International Conference on Cognitive Science*. CogPrints, Sydney, Australia, University of New South Wales (2003) 133–138
- [14] Smolensky, P.: Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence* **46** (1990) 159–216
- [15] Kanerva, P.: The binary spatter code for encoding concepts at many levels. In Marinaro, M., Morasso, P., eds.: *ICANN ’94: Proceedings of International Conference on Artificial Neural Networks*. Volume 1., London, Springer-Verlag (1994) 226–229
- [16] Plate, T.: *Holographic reduced representations*. Technical Report CRG-TR-91-1, Department of Computer Science, University of Toronto (1991)
- [17] Anderson, J., Silverstein, J., Ritz, S., Jones, R.: Distinctive features categorical perception and probability learning: some applications of a neural model. *Psychological Review* **84**(5) (1977) 413–451
- [18] Jones, M.N., Mewhort, D.J.K.: Representing word meaning and order information in a composite holographic lexicon. *Psychological Review* **114** (2007)
- [19] Foltz, T.L.T., Laham, D.: Introduction to latent semantic analysis. *Discourse Processes* **25** (1998) 259–284
- [20] Gentner, D., Toupin, C.: Systematicity and surface similarity in the development of analogy. *Cognitive Science* **10** (1986) 277–300
- [21] Eliasmith, C., Thagard, P.: Integrating structure and meaning: a distributed model of analogical mapping. *Cognitive Science* **25**(2) (2001) 245–286
- [22] Wason, P.: Reasoning. In Foss, B., ed.: *New Horizons in Psychology*. Penguin, Harmondsworth (1966)
- [23] Eliasmith, C.: Cognition with neurons: A large-scale, biologically realistic model of the wason task. In Bara, G., Barsalou, L., Bucciarelli, M., eds.: *Proceedings of the 27 th Annual Meeting of the Cognitive Science Society*. (2005)
- [24] Arathorn, D.W.: *Map-Seeking Circuits in Visual Cognition*. Stanford University Press (2002)