# A Solution to the Binding Problem for Compositional Connectionism


**John E. Hummel, Keith J. Holyoak, Collin Green, Leonidas A. A. Doumas, Derek Devnich, Aniket Kittur, Donald J. Kalar**

UCLA Dept of Psychology



### Abstract

Achieving compositional connectionism means finding a way to represent role-filler bindings in a connectionist system without sacrificing role-filler independence. Role-filler binding schemes based on varieties of conjunctive coding (the most common approach in the connectionist literature) fail to preserve role-filler independence. At the same time, dynamic binding of roles to fillers (e.g., by synchrony of firing) represents bindings without sacrificing independence, but is inadequate for storing bindings in long-term memory. An appropriate combination of dynamic binding (for representation in working memory) and conjunctive coding (for long-term storage and token formation) provides a platform for compositional connectionism, and has proven successful in simulating numerous aspects of human perception and cognition.


## Compositional Systems, Role-Filler Independence and Binding

Compositional systems combine and recombine a finite number of representational elements (symbols or sub-symbols) into a much larger (infinite, if recursion is permitted) number of specific structures. Examples include natural languages and formal symbol systems (e.g., propositional notation, mathematical notation, and computer programming languages). An important property of a compositional system is that the meaning and representation of individual elements in a structure do not vary as a function of their position in the structure as a whole. For example, "*loves*," "John" and "Mary" can be combined to form two different propositions, *loves* (John, Mary) and *loves* (Mary, John). The meaning and representation of these elements (which can be objects, relations or relational roles) is fundamentally independent of each element's place in the compositional structure as a whole (e.g.,, Mary is represented in the same way whether she is a lover or a beloved). At the same time, the meaning of the compositional structure does depend on the elements' configuration—i.e., the *bindings* of arguments (here, John and Mary) to relational roles (*lover* and *beloved*).


Copyright © 2004, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.


### Role-Filler Independence

This property of formal compositional systems is so fundamental that it is easy to overlook. Yet it is essential: If the representation of an element (role or object) varied as a function the argument or role to which it was bound, then the resulting system would not be properly described as "compositional": Different expressions consisting of the "same" symbols would not consist of the same symbols; they would be different expressions with different symbols. They would simply be *different*, with nothing systematic in common.

The importance of role-filler independence is apparent in human relational reasoning. Relational generalization—inferences and generalizations that depend on relations between objects rather than just the features of those objects—is only possible if relational roles are represented independently of their fillers (Hummel & Holyoak, 1997, 2003). For example, suppose someone knows that John loves Mary, Mary loves Zack, and John is jealous of Zack. This person then observes that Sally loves Tom, and Tom loves Cindy. A plausible analogical inference is that Sally will be jealous of Cindy. This inference is based on the relational correspondences (i.e., participation in corresponding roles) of John to Sally, Mary to Tom, and Zack to Cindy. Making the inference requires the reasoner to discover these correspondences and to use them to guide inferences about the second (novel) situation (e.g., it is not a plausible inference that Tom will be jealous of Sally). Importantly, it is only possible to discover and use the correspondences if the *loves* relation is represented in the same way regardless of who loves whom (i.e., independently of its fillers), and the people involved are represented independently of their role bindings (e.g., the same Mary is both a lover and beloved; see Hummel & Holyoak, 2003).

### The Binding Problem

Representing roles independently of their fillers, while necessary, is not sufficient for achieving compositionality. It is also necessary to specify how fillers are bound to relational roles. The only difference between *loves* (John, Mary) and *loves* (Mary, John) is in the bindings of fillers (people) to the roles of the *loves* relation. Yet the two statements mean very different things.

Moreover, as illustrated by the jealousy example, in order to reason using compositional representations, it is not even sufficient to represent roles independently of their fillers and represent their bindings. It is also necessary to integrate multiple role-filler bindings: It is not just Sally's role as *lover*, or Tom's role as *beloved* that causes Sally to be jealous; it is Sally's role as *lover with* Tom as the *beloved and* Tom's role as a *lover without* Sally as the corresponding *beloved* that together cause Sally to be jealous. Compositional representations are neither simply collections of roles and fillers, nor simply collections of role-filler bindings. They are integrated systems of role-filler bindings.

## Achieving Role-Filler Independence and Binding in a Connectionist System

Most of the debate between connectionists and proponents of symbolic models concerns whether human mental representations are genuinely compositional, and so whether it is even desirable to achieve compositionality in a cognitive model. We have argued (Holyoak & Hummel, 2000; Hummel & Holyoak, 1997, 2003) that mental representations are indeed compositional. For the present, we simply claim that it is at least possible they might be, and that it is therefore interesting to consider how a connectionist system might achieve true compositionality. We therefore assume, for the purposes of this paper, that the goal is to achieve true compositionality in a connectionist system, and emphasize the representational and computational constraints on achieving that goal.

Achieving genuinely compositional connectionism means achieving both role-filler independence and explicit role-filler binding in a connectionist system. That is, it is necessary to identify a means of binding roles to fillers in a way that does not affect the representation of either.

### Binding in Connectionist Systems

The role of binding in compositionality is more salient than that of role-filler independence, so it is no surprise that modelers interested in compositional connectionism of one form or another (e.g., Halford et al., 1998; Hinton, 1990; Hummel & Biederman, 1992; Hummel & Holyoak, 1997, 2003; Kanerva, 1998; Plate, 1991; Shastri & Ajjanagadde, 1993; Smolensky, 1990) have typically paid more attention to the binding problem than to the independence problem.

**Conjunctive Coding**The most straightforward way to achieve binding in a connectionist representation is to use *conjunctive coding*: to designate units that represent *conjunctions* of roles and fillers. (More generally, "conjunctive coding" refers to designating units to represent conjunctions of any type.) For example, Hinton's (1990) model represents family relations by designating three separate pools of units: One to represent the "agent" of a relation, another to represent the "patient" of the relation, and a third to represent the relation itself. Each unit thus represents a conjunction of a person (or person feature) and a particular role. To represent "Mary is the wife of George," this model activates the pattern for "Mary" on the agent units, the pattern for "wife" on the relation units, and the pattern for "George" on the patient units. To represent "George is the husband of Mary," it activates "George" on the agent units, "Mary" on the patient units, and "husband" on the relation units. The limitation of this approach is that it violates role-filler independence. Mary in the agent role of the "wife" relation is represented on one set of units, whereas Mary as the patient of "husband" is represented on a completely separate set of units. As a result, nothing learned about Mary in the former role need transfer to Mary in the latter (formally equivalent) role. For example, if the model knows that Mary, the agent of *wife-of*, is jealous of George's affair with Sally, then it could simultaneously know that Mary, the patient of the *husband-of* relation, is completely ignorant of it, or even happy about it. The two representations of Mary are simply unrelated to one another.

The reason for this separation is that, to a connectionist system, all that matters is which units represent a concept (e.g., Mary). Learning in a connectionist system is specific to particular units: If unit A learns a strengthened connection to unit B, then that learning has no direct effect on the connection between C and D. Thus, when Mary, the agent of *wife-of* (represented, say, by units A – G) learns about her husband's affair, Mary the patient of *husband-of* (represented by units T – X) does not.

This failure to transfer between the two occurrences of Mary reflects the violation role-filler independence intrinsic to conjunctive coding. If Mary were represented independently of her roles in the *wife-of* and *husband-of* relations (i.e., by the same units regardless of the role to she happened to be bound at the time), then Mary the patient of *husband-of* would learn exactly the same things as Mary the agent of *wife-of*. Conjunctive coding does not permit learning (or inference) to extend beyond specific role-filler *conjunctions*, so it neither allows a filler to be treated as the same entity in different roles, nor allows a role to be treated as the same role with different fillers. In other words, conjunctive coding cannot support relational generalization (Hummel & Holyoak, 2003). This problem renders conjunctive coding inadequate (or, as we will suggest later, fundamentally incomplete) as a solution to role-filler binding in compositional connectionist systems.

**Tensor Products** A more sophisticated approach to connectionist role-filler binding is based on vector multiplication in the form of the outer (i.e., tensor) product of two or more vectors (e.g., Halford et al, 1998; Smolensky, 1990), and related schemes (e.g., Kanerva, 1998; Plate, 1991). When two vectors, $\mathbf{u}$ and $\mathbf{v}$, are multiplied to form a tensor (a rank two tensor), the value of the *ij*th element of the tensor is simply the product of the *i*th element of $\mathbf{u}$ and the *j*th element of $\mathbf{v}$. Similarly, the *ijk*th element of a rank three tensor is the product of $\mathbf{u}_i$, $\mathbf{v}_j$ and $\mathbf{w}_k$. A tensor is thus

Copyright © 2004, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

a matrix that is treated like an activation vector. Tensors of any rank can be formed in this way. "Mary is the wife of George" could be represented by the rank three tensor, **mwg**, formed by multiplying **m (**representing Mary) by **w** (representing *wife-of*) and **g** (representing George); "George is the husband of Mary" could likewise be **ghm**.

Although perhaps not apparent at first blush, tensor products (and their relatives, such as spatter codes [Kenevera, 1998] and holographic reduced representations (HRRs; Plate, 1992]) are a form of conjunctive coding: the *ijk*th element of **mwg** depends on the *i*th element of **m**, the *j*th value of **w** *and* the *k*th element of **g**. No element of a tensor represents an object independently of its role bindings or vice-versa. As a result, a tensor-based representation of, for example, "$x = 8$" (e.g., **xe8**) may not overlap at all with a tensor representation of "$y = 8$" (e.g., **ye8**). The overlap between **xe8** and **ye8**—for example, the dot product **xe8•ye8**—is equal to the *product* **(x•y)(e•e)(8•8)** (Holyoak & Hummel, 2000). Thus if **x•y** = 0 (i.e., if the representation of the variable $x$ shares no units with the representation of $y$), then the entire product **xe8•ye8** = 0. That is, the tensors **xe8** and **ye8** fail to specify what "$x = 8$" has in common with "$y = 8$". Tensor products thus suffer the same shortcomings as other forms of conjunctive coding, making them incapable of supporting true composition in a connectionist system.

**Dynamic Binding** The deep limitation of tensor products and other forms of conjunctive coding is that they represent role-binding information in the *identities* of units. Changing the role to which an object is bound (or vice versa) necessarily changes which units represent that role or object: it changes the very representation of the object or role. An analogous convention in a traditional symbol system would be to represent the number eight with the symbol "8" when it is bound to the variable $x$, and with "9" when bound to $y$. This "solution" to the binding problem is no solution at all: Binding is something a compositional system *does to* the elements of that system; it is not a property the elements themselves. Binding is something that is applied *dynamically* to elements in order to indicate that, in this particular structure, these elements are bound together in this way.

Representing bindings dynamically requires a degree of freedom external to the identities of the bound symbols/entities. Activation is a plausible binding tag for connectionist systems, but it is already "in use" representing other information (Hummel & Biederman, 1990, 1992). For this reason, SRNs (e.g., Ellman, 1990) and other connectionist approaches that do not represent bindings dynamically (i.e., independently of units and their activations) must, as a logical consequence, be non-compositional: Every unit in an SRN must necessarily represent some conjunction of role and filler information (in which case it fails to represent them independently), or else represent some roles/fillers independently of other fillers/roles (in which case it does not specify their bindings), or some combination of the two (in which case some are not independent and others not bound).

Many kinds of binding tags can be imagined for connectionist systems: Units could fire in different "flavors" or "colors" as a function of how they are bound together (e.g., with units that are bound together firing in the same color and different bound groups firing in different colors). At present, however, the most neurally plausible proposed binding tag is based on the use of time: Units fire in synchrony (or in systematic asynchrony; Love, 1999) when they are bound together, and out of synchrony when they are not. Binding by synchrony is used in numerous connectionist/neural models (see, Hummel & Holyoak, 1997, for a review) and there is also support for it, albeit controversial, in the neurophysiological literature (see Singer, 2000, for a review).

The virtue of dynamic binding is that it allows the same units to represent a given object or relational role regardless of the role or object to which it happens to be bound. That is, it permits a connectionist system to achieve role-filler binding without sacrificing role-filler independence. For example, to represent "John loves Mary," the units representing John fire in synchrony with the units representing *lover*, while the units for Mary fire in synchrony with those for *beloved* (and out of synchrony with John and *lover*). "Mary loves John" would be represented by exactly the same units, only their synchrony relations would be reversed (Hummel & Biederman, 1992; Hummel & Holyoak, 1997; Shastri & Ajjanagadde, 1993; von der Malsburg, 1981).

Dynamic binding (e.g., through synchrony) is thus a step toward compositional connectionism. However, it is capacity-limited, making it impractical for storage in long-term memory (LTM). Although the capacity limits of dynamic binding are consistent with the capacity limits of working (WM), dynamic binding must be augmented with conjunctive coding for the purposes of storage in LTM and token formation in WM (Hummel & Holyoak, 1997, 2003). Combined appropriately, conjunctive coding and dynamic binding permit a connectionist system to achieve role-filler binding, role-filler independence and integration of multiple role-filler bindings simultaneously. That is, they give rise to truly compositional connectionism.

## Using Compositional Connectionism to Model Cognition

Our research group has been exploiting these principles for over a decade to use compositional connectionism to simulate aspects of object recognition (Hummel & Biederman, 1990, 2992; Hummel & Stankiewicz, 1996, 1998; Hummel, 2001), memory retrieval, analogical mapping, and the complex relations between them (Hummel & Holyoak, 1997), as well as analogical inference, schema induction and relational generalization generally (Hummel & Holyoak, 2003). We have used the same principles to simulate relations between effortless "reflexive" forms of inference and more effortful "reflective" inferences

Copyright © 2004, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

(Hummel & Choplin, 2000), the role of specialized (e.g., perceptual) computing "modules" in reasoning (Hummel & Holyoak, 2001), the origins, nature and role of capacity limits in human thinking (Kubose, Holyoak & Hummel, 2002), and the impact of aging (Viskontas et al., 2004), fronto-temporal degeneration (Morrison et al., in press) and dual-tasking on reasoning performance.

Taken as a whole, our research program suggests that seeking to achieve role-filler binding, independence and integration simultaneously in a connectionist system is a powerful formula for simulating numerous aspects of human perception and cognition, and for realizing the promise of compositional connectionism.

## References


Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Brain and Behavioral Sciences*, *21*, 803-864.

Hinton, G. E. (Ed.) (1990) . *Connectionist symbol processing*. Cambridge, MA: MIT Press.

Holyoak, K. J., & Hummel, J. E. (2000). The proper treatment of symbols in a connectionist architecture. In E. Dietrich and A. Markman (Eds.). *Cognitive dynamics: Conceptual change in humans and machines* (pp. 229 - 264). Hillsdale, NJ: Erlbaum.

Hummel, J. E. (2001). Complementary solutions to the binding problem in vision: Implications for shape perception and object recognition. *Visual Cognition*, *8*, 489 - 517.

Hummel, J. E. & Biederman, I. (1990). Dynamic binding: A basis for the representation of shape by neural networks. *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society* (pp. 614-621). Hillsdale, NJ: Erlbaum.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, *99*, 480-517.

Hummel, J. E., & Choplin, J. M. (2000). Toward an integrated account of reflexive and reflective reasoning. In L. R. Gleitman and A. K. Joshi (Eds.), *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society* (pp. 232-237). Mahwah, NJ: Erlbaum.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review*, *104*, 427-466.

Hummel, J. E., & Holyoak, K. J. (2001). A process model of human transitive inference. In M. Gattis (Ed.). *Spatial schemas in abstract thought* (pp. 279-305). Cambridge, MA: MIT Press.

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review*, *110*, 220-264.

Hummel, J. E., & Stankiewicz, B. J. (1996). An architecture for rapid, hierarchical structural description. In T. Inui and J. McClelland (Eds.). *Attention and performance XVI: Information integration in perception and communication* (pp. 93-121). Cambridge, MA: MIT Press.

Hummel, J. E., & Stankiewicz, B. J. (1998). Two roles for attention in shape perception: A structural description model of visual scrutiny. *Visual Cognition*, *5*, 49-79.

Kanerva, P. (1998). *Sparse distributed memory*. Cambridge, MA: MIT Press.

Kubose, T. T., Holyoak, K. J., & Hummel, J. E. (2002). The role of textual coherence in incremental analogical mapping. *Journal of Memory and Language, 47*, 407-435.

Love, B. C. (1999). Utilizing time: Asynchronous binding. *Advances in Neural Information Processing Systems*, *11*, 38-44.

Morrison, R.G., Krawczyk, D.C., Holyoak, K.J., Hummel, J.E., Chow, T.W., Miller, B.L., & Knowlton. (2004). A neurocomputational model of analogical reasoning and its breakdown in Frontotemporal Lobar Degeneration. *Journal of Cognitive Neuroscience*, *16*, 260-271.

Plate, T. (1991). Holographic reduced representations: Convolution algebra for compositional distributed representations. In J. Mylopoulos & R. Reiter (Eds.), *Proceedings of the 12th International Joint Conference on Artificial Intelligence* (pp. 30-35). San Mateo: Morgan Kaufmann.

Shastri, L., & Ajjanagadde, V. (1993). From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings using temporal synchrony. *Behavioral and Brain Sciences, 16*, 417-494.

Singer, W. (2000). Response synchronization, a universal coding strategy for the definition of relations. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd edition). Cambridge, MA: MIT Press.

Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence*, *46*, 159-216.

Viskontas, I., Morrison, R. G., Holyoak, K., J., Hummel, J. E., & Knowlton, B. J. (2004). Relational integration, attention and reasoning in older adults. Manuscript in preparation, Dept. of Psychology, UCLA.

von der Malsburg, C. (1981). The correlation theory of brain function. Internal Report 81-2, Department of Neurobiology, Max-Planck-Institute for Biophysical Chemistry.



Copyright © 2004, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.