Connections Binding Unification & Analogical Promiscuity

Ross W. Gayler
102 Through Road
Burwood VIC 3125, AUSTRALIA
r.gayler@psych.unimelb.edu.au

Roger Wales
Department of Psychology, The University of Melbourne
Parkville VIC 3052, AUSTRALIA
r.wales@psych.unimelb.edu.au

Context of Paper

- This is a working hypothesis about the nature of cognition and connectionist computation
- It is not specifically psychological (more an engineering approach)
- It is intentionally extreme to generate ideas (a vision statement)
- Connectionist implementation work is in progress
- Results of the implementation work are encouraging

Thesis

- Higher cognition implemented by a connectionist system is <u>fundamentally</u> analogical. That is:
  - Connectionist representation is effectively symbolic
  - Analogical mapping by continuous systematic substitution is the core cognitive process.
- This conclusion is drawn from a design argument
  - The properties are necessary consequences of meeting the task demands of higher cognition by using connectionist processing capabilities

Design Argument

- The problem is: How to (broadly) design a connectionist system to
  - support higher cognition
  - remain faithful to connectionist constraints
- This drives a series of design decisions in response to technical problems in adapting a connectionist system to higher cognition
  - The decisions lead to effectively symbolic representation
  - This leads, in turn, to the centrality of analogical mapping
  - The design decisions are plausible extrapolations of research on binding

Design Boundary Conditions

- Functional requirements - "Higher Cognition" is defined by a task:
  - The ability to follow novel, cross-modal instructions
- The design must respect connectionist constraints:
  - Many interconnected processing units
  - Local processing (each unit is connected to a small fraction of the units)
  - Scalar outputs (the set of outputs over units is interpretable as a vector)
  - Simple combination function (sum or product)
  - Fixed architecture (no construction of networks on the fly, as in ACME)

Problem 1: Representing Novel Concepts

- Arbitrary, novel concepts are needed to be represent the instructions
- This rules out local representations
  - The fixed, finite stock of units implies a fixed, finite stock of concepts
- Not all distributed representations are adequate
  - Feature-based (subsymbolic) representations have the same problem. The fixed stock of features gives a fixed space of representable concepts
- **Decision 1: Don't require fixed meanings for individual units**
  - Context dependent meanings for individual units are acceptable

Problem 2: Immediate Learning

- The task definition requires immediate learning (i.e. construction and storage of the representation of the novel instruction)
- Learning by iterative weight adjustment (e.g. backprop) is too slow
- We need single step association of arbitrary patterns
- We need the automatic capability to decompose associated patterns
- **Decision 2: Implement learning by binding of patterns**
  - Patterns are composable and decomposable with binding operators
    - unbind(bind(x,y),x) = y
  - We are using multiplicative binding. Other binding operators are available

Problem 3: Representational Compatibility

- Processing requires the interaction of short & long term knowledge
- Standard NN representations of short & long term knowledge differ
  - Short term knowledge is represented as an activation <u>vector</u>,
  - Long term knowledge is represented as a weight <u>matrix</u>
  - Short and long term knowledge interact via the processing function
  - Typically, only Short/Long interaction is supported and not Short/Short
- The time scale of our task requires Short/Short interaction
- **Decision 3: Use identical short & long term representations**
  - Identical representation does not require identical implementation
  - In multiplicative binding, activation <u>vectors</u> (short term implementation) and weight <u>vectors</u> (long term implementation) are treated equivalently
  - Short/Short and Short/Long interactions are identical. Long/Long interactions can only occur indirectly via the short term representations.

Interpretations of Bindings as Representations

- Bindings may be interpreted as virtual neurons
  - Create arbitrary pattern associations (grandmother cells on demand)
- Bindings guarantee systematicity (clumps of capability)
  - Any bound component may be used to retrieve the other components
- Bindings may be interpreted as rules over constants (a & b $\Rightarrow$ c)
  - With graded similarity of constants (but most similarities are zero)
  - And with additivity of rules: bind(x,y) + bind(x,z) = bind(x,y+z)

Problem 4: Avoiding Spurious Similarity

- Which pattern should be used to represent a novel concept?
- Connectionist processing is driven by similarity
- But the new concept representation is constructed by immediate learning before we know whether it should be similar to any others
- **Decision 4: Novel concept patterns should be dissimilar**
- This avoids spurious associations between concepts
- Dissimilarity of novel concept representations is a natural consequence of representation based on binding
  - Binding any novel component causes dissimilarity
  - If **not(d $\approx$ e)** then **not(bind(a,b,c,d) $\approx$ bind(a,b,c,e))**

Claim 1: Patterns as Symbols

- Classical symbols are atomic and discrete
  - Most patterns in a high-dimensional space are dissimilar (effectively atomic and discrete). They are similar if composed from related patterns.
- Classical symbols can be arbitrarily composed
  - Patterns may be arbitrarily composed with binding operators
- Classical symbols have arbitrary referents
  - Patterns do not necessarily have inherent meanings because units have no fixed meanings and they may be arbitrarily associated by binding

Problem 5: Using Unique Patterns

- How can dissimilar, arbitrary tokens be causal in processing? (Standard connectionist processing is driven by pattern similarity)
- What carries the representational information? Not the atomic token patterns because they are essentially arbitrary.
- Each token participates in multiple bindings. The structural relationships between bindings carry the information
- Detection of structural equivalence is graph matching. (Are both structures identical after systematic substitution of tokens?)

- **Decision 5: Use systematic substitution as the core process**
  - The identity substitution is legitimate and might be common. (There may be some representations that <u>do</u> have fixed meanings)

Rules with Variables

- What are the consequences of systematic substitution?
  - Every constant becomes a variable
  - Every episodic representation becomes a rule with variables

---

Claim 2: Analogy is Core

- Analogical mapping is implemented by systematic substitution
- Systematic substitution is needed to use arbitrary tokens
- The basic cognitive process <u>is</u> analogical mapping
- Analogy is necessary because:
  - Each episode is novel (especially if goals and context are included)
  - Each episodic representation is novel, therefore tokens are dissimilar
  - Only structural similarity is available to drive connectionist processing

---

Basis of Decomposition

- Composite patterns need to be decomposed.
- There are no unique reference axes (no fixed unit meanings)
- There is no unique architectural decomposition into constituents
- Composite patterns are decomposed with respect to other representations (structural consistency between representations)

Regenerative Structures

- Short term patterns are continuously decomposed with respect to each other and long term patterns
- The decomposition is inexact and includes new constituents
- Short term composite patterns are regenerated from the constituents
- Regenerating patterns are continually evolving and competing
- The evolution of patterns engages short and long term knowledge

---

Summary

- Use distributed representations with no fixed unit meanings
- Pattern association is implemented by binding and unbinding
- Short and long term knowledge have the same representation
- New concepts get represented by novel, unique patterns
- Processing structures of unique patterns uses systematic substitution
- Patterns are effectively symbols
- Analogy is the core cognitive process