

Compositionality Continuum as Neural Basis of Intelligence

Paper Reading

Srinitish Srinivasan 19th August 2026

What is Composition?

- Compositionality is the core of cognition
- Intelligent systems combine finite set parts into many new expressions.
- Definition: a function is represented by its function of its parts and by rules that specify how these parts combine.
- Language does this all the time. Words combine through grammar.
- How does RNN, large language models do this?

What kinda questions are we asking?

- Is a system compositional? Are LLMs compositional?
- This is the wrong question because composition isnt binary!!!!
- It lies on a continuum
- Questions: Are the outputs compositional in nature but do the internal operations resemble composition?
- A tiny model may have clean reusable models but solve very simple tasks. A huge LLM may converse but its internal mechanisms are hard to classify . A rat may lack language and yet bind abstract sequence roles to new locations.

Expressivity and complexity

- Each reusable component, how powerful it is? A single feature detector has low expressivity. Motifs that accumulates evidence or stores a memory is richer.
- How sophisticated are the recombination rules? Modules simply may add or factorize
- Flexible recombination needs enough modules and diverse contexts to reveal which combinations are valid.

Single Variable Encoding

- Single variable encoding: A single variable is aligned to a single neuron. One factor is reusable but less expressive.
- Its a base case. Modular expressivity can be built from here.

Abstract Variable Encoding

- Functions don't just encode raw stimulus features. It organizes activity so that a task relevant variable can be decoded across contexts. A rule learned for one stimulus type can be transferred to others.
- If an animal learns respond when tone is above the boundary, the same abstract relation can support respond when visual contrast is above the boundary. Relational computation is the same despite change in sensory input.

RNNs

- In RNNs and biological circuits, tasks unfold over time. Networks here don't encode one static embedding, it evolves over a sequence of states.
- One dynamical sequence feeds to another, then the outgoing state of the first sequence, must be incompatible with the input region of the next motif.

Pro+Memory->Memory Anti

- Anti means respond opposite the stimulus
- Memory means hold the stimulus through a delay
- New Task: memory anti means the RNN can combine the memory motif with the anti response motif instead of learning the entire task from scratch.
- Compositionality is visible inside the model as shared sequences. Not just as task accuracy.

Binding

- Binding means a particular object, location, word is assigned to an abstract role inside a larger structure.
- Once the system learns the role structure, it can generalize on new objects.
- The cat eats the fish follows the structure subject->verb->object
- The cat eats the florp(unseen). The structure is still subject->verb->object.
- Florp binds to object.

- Rat navigation. A rat learns a cyclic structure $A \rightarrow B \rightarrow C \rightarrow D \rightarrow A$. New physical locations can then bind to the above structure.

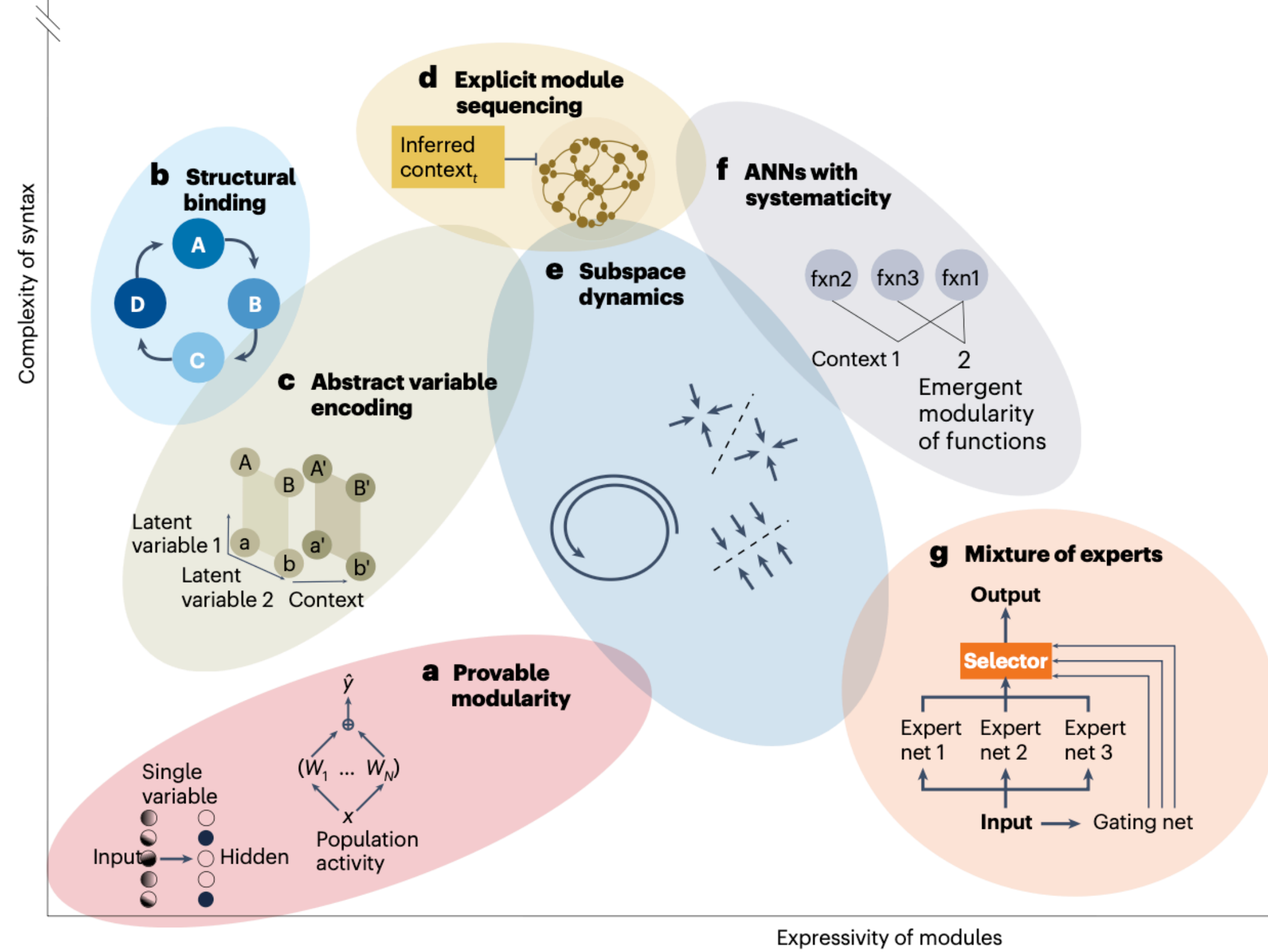
Discontinuity of Language

- Natural language is uniquely human. Syntax and semantics are deeply entangled.
- Classical ideas say that syntax should be abstract reusable independent of component. The human language does not behave so neatly.
- Purely grammatical rules can become hard to understand without semantic support
- But if the sentence has some semantic mean, even if grammatically it doesnt really make sense, people can still understand what it means.

Pattern matching and Compositional Implementation

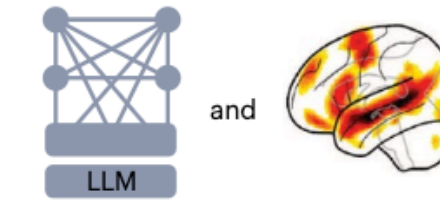
- In pattern matching, system learns correlations between input and output. If test cases, resemble training data, it may perform well.
- Fails when tasks need genuine recombination outside the training distribution
- LLMs show compositional behavior but a lot of it could be part of the training distribution itself.

Compositional neural implementations



Unknown neural implementations

h Human-level natural language



Formal logic and mathematical reasoning
 $p \mid \sim p$

So what next?

- Scaling alone does not encourage compositional mechanisms.
- Encourages curriculum learning, tasks designed to test whether internal modules can recombine out of distribution