A Formal System for Pattern-Constellations

Florin Cojocariu

09.30.2025

Contents

1. Introduction and Motivation

This document presents a unified formal system for describing pattern-constellations (PCs), their linguistic expression, and the operations that connect them. The formalism progresses from pre-linguistic pattern-recognition through primitive language use to abstract conceptual structures, ultimately addressing the problem of reference.

Pattern Constellations are introduced and explained elsewhere; this document is focused on scaffolding a formalism around them. For the purposes here, we'll recall that a Pattern Constellation is a stable attractor[1] configuration in a neural network that integrates sensory, motor, emotional, social, and (once learned) linguistic patterns. Formed through Hebbian learning where co-occurring patterns wire together. Activated as a unified whole when any partial pattern triggers it. Exists as a weighted complex network structure that can be measured in NN models, not an abstract entity. PCs are formalized in §2 below.

Core philosophical commitment: Pattern-recognition events are atomic and pre-linguistic. Language emerges as a coordination mechanism for these events, not as a bridge to mind-independent reality.

2. Primitive Pattern-Recognition Events

2.1. Experiential Actualization

Notation: $E(\{A\})$ or $E(\{A, a^o\})$[2]

Reads as: *"Pattern-constellation A is experientially activated" or "Pattern-constellation A labelled with the object-word a^o is experientially activated"*

Definition: Marks when a pattern-recognition event occurs in direct experience. These are atomic events with no internal structure - pattern and recognition are unified inside the PC, not separate.

Two forms:

Pre-linguistic: $E(\{A\})$

- Pure sensory-motor-emotional pattern recognition

[1] This is best modelled by Hopfield Networks, a type of NN that learns rapidly and remembers patterns. Its theoretical basis is built around minimum energy configuration and in this respect a learned pattern is literally an attractor.

[2] Both E and $\exists$ are used in the sense of "existence", but according to the two different uses we have for "existence", un-mediated, and mediated. To recognize something in the real world is, in terms of mental states, to become aware of its existence, thus E. However, for conceptual things, it is rather their possibility for existence in the abstract space that is recognized, hence the $\exists$ from logic.

- No linguistic component
- Present in animals and pre-linguistic infants

With integrated label: $E(\{A, a^o\})$

- Experience includes linguistic label
- a^o fires together with sensory patterns
- Post-language-learning stage

Examples:

- $E(\{CAT\})$: Cat seeing mouse (no linguistic mediation)
- $E(\{DOG, dog^o\})$: Child seeing dog (with label integrated)
- $E(\{DANGER\})$: Sensing threat (pre-linguistic) (in order to make the differences clear, Pattern Cosntellation names will be written in CAPS. CAT is not the word “ ”cat”—neither its concrete label use, cat^o, nor its metaphorical or abstract one, cat^c - it is just a coherent way to label a complex entity built in our minds as we learn what a cat is)

Properties:

- Atomic (cannot be decomposed into pattern + recognition)
- Subjective (occurs in individual experience)
- Can occur with or without linguistic component
- Grounds all subsequent linguistic operations

Neural substrate: Hopfield-like attractor states, learned through a Hebbian process and activated by any partial sensory input

2.2. Linguistic Mediation

Notation: $\exists(\{A, a^c\})$

Reads as: “Pattern-constellation A is linguistically mediated through concept-word a^c.

Definition: Marks when a pattern-constellation is expressed, coordinated, or referenced through language using the abstract/conceptual mode of the word.

Key distinction from E:

- E uses a^o mode (sensory-integrated, grounded in experience) but can be completely pre-verbal
- $\exists$ uses a^c mode (abstract, conceptual, can operate detached from immediate experience) and cannot exist outside language

Examples:

- Child says "Dogs are animals" $\rightarrow \exists(\{DOG, dog^c\})$ (conceptual statement, no dog present)
- Talking about justice $\rightarrow \exists(\{JUSTICE, justice^c\})$ (purely conceptual, no direct E possible)
- Planning: "We'll see dogs at the park" $\rightarrow \exists(\{DOG, dog^c\})$ (anticipatory, no current E)

Relationship to E:

- $E(\{A\})$ can occur without $\exists(\{A, a^c\})$ (pre-linguistic experience)
- $\exists(\{A, a^c\})$ can occur without current $E(\{A, a^o\})$ (talking about absent things)
- But $\exists(\{A, a^c\})$ typically develops from repeated $E(\{A, a^o\})$ through learning

Key distinction: E and $\exists$ mark **different functional modes**:

- E: grounded in immediate experience, uses x^o(label mode)
- $\exists$: operating in conceptual/linguistic space, uses x^c (concept mode)

2.3. The Fundamental Transition

Notation: $E(\{A, a^o\}) \rightarrow \exists(\{A, a^c\})$

Reads as: "From experiencing-with-label to linguistically-mediating-abstractly"

What this captures:

- Move from sensory-grounded mode (xˆo) to conceptual mode (xˆc)
- Language becoming detached from immediate experience
- Ability to talk about A when not experiencing A
- Foundation for abstract thought, planning, communication about absent things

Not a temporal sequence (both can co-occur) but a **functional transition** showing:

- How same word operates in different modes
- Progression from grounding to abstraction
- Basis for all higher linguistic operations

2.4. Combined Cases

Simultaneous experience and linguistic mediation: $E(\{A, a^o\}) \wedge \exists(\{A, a^c\})$

Experiencing something while also talking about it abstractly/conceptually.

Examples:

$E(\{MOUNTAIN, mountain^o\}) \wedge \exists(\{MOUNTAIN, mountain^c\})$ - Seeing the mountain

AND discussing mountain geography

$E(\{PAIN, pain^o\}) \wedge \not\exists(\{PAIN, pain^c\})$ - Feeling pain but unable to articulate/conceptualize it

$\not E(\{JUSTICE\}) \wedge \exists(\{JUSTICE, justice^c\})$ - Justice exists conceptually but not as direct experience

3. Pattern-Constellations

3.1. Definition

Notation: $\{A\}$ where A is capitalized

Reads as: "The pattern-constellation A"

Definition: A unified attractor state in neural space integrating multiple pattern types:

- Sensory patterns (visual, auditory, tactile, olfactory, gustatory)
- Motor patterns (action affordances, manipulation schemes)
- Emotional patterns (valence, arousal, specific feelings)
- Social patterns (conventional uses, shared practices)
- Linguistic patterns (words in $x^o$and x^c modes, once integrated)

Key insight: The capital letter (A, DOG, SNOW) names the pre-linguistic pattern structure itself, not any word.

Neural implementation: Hopfield network with Hebbian learning

- Patterns that co-occur during learning wire together
- Form unified low-energy configuration (attractor basin)
- Activated as whole when partial pattern encountered

3.2. Stages of Constellation Development

Stage 1: $\{A\}$ - Pre-linguistic

Pure sensory-motor-emotional integration $E(\{A\})$ possible No linguistic component Example: {MOUSE} in cat's brain

Stage 2: $\{A, a^o\}$ - Label integrated

Word "a" integrated as label through learning $E(\{A, a^o\})$ now occurs (experiencing with label) a^o

embedded in constellation Example: {DOG, dog^o} in child who has learned "dog"

Stage 3: $\{A, a^o, a^c\}$ - Dual function

Same word operates in two modes: - a^o: label mode (grounded, sensory-integrated) - a^c: concept mode (abstract, LPC operations) Both $E(\{A, a^o\})$ and $\exists(\{A, a^c\})$ possible Example: $\{DOG, dog^o, dog^c\}$ in language-competent speaker

Stage 4: $\mathcal{R}(a^c, a^o)$ - Reference proper

Explicit coordination between modes Answers "What does a refer to?" Meta-linguistic capability Presupposes Stage 3 development

3.3. Example Constellations

{DOG} $\rightarrow$ **{DOG,** dog^o**}** $\rightarrow$ **{DOG,** dog^o**,** dog^c**}**:

Stage 1 - Pre-linguistic {DOG}: Visual: four-legged, furry, tail Auditory: barking, panting Tactile: soft fur, warm Motor: petting, playing Emotional: affection, excitement E({DOG})$ occurs in pre-linguistic child/animal

Stage 2 - Label integrated {DOG, dog^o}: [All DOG patterns] + Linguistic: "dog"^o integrated $E(\{DOG, dog^o\})$ occurs when seeing dog

Stage 3 - Conceptual function {DOG, dog^o, dog^c}: [All above] + Linguistic: dog^c for abstract use $\exists(\{DOG, dog^c\})$ when talking about dogs abstractly "Dogs are animals" uses dog^c

{SNOW} $\rightarrow$ **{SNOW,** $snow^o$**}** $\rightarrow$ **{SNOW,** $snow^o$**,** $snow^c$**}**:

- Stage 1: Visual (white, crystalline), Tactile (cold, wet), Motor (scooping)
- Stage 2: + $snow^o$ as label
- Stage 3: + $snow^c$ for abstract/conceptual use

3.4. Properties

Integration: All components activate together (Hebbian co-activation)

Partial activation:

- See dog $\rightarrow E(\{DOG, dog^o\})$ - entire constellation activates
- Hear "dog" $\rightarrow$ can trigger $E(\{DOG, dog^o\})$ or $\exists(\{DOG, dog^c\})$ depending on context

Learning: Constellations built through repeated co-occurrence

- Bulk phase in childhood

- Continuous refinement throughout life
- Environmental interaction shapes constellation structure

Individual variation: $\{\text{A}\}$ varies between individuals based on:

- Personal learning history
- Cultural context
- Frequency and type of encounters

Overlap: $\{\text{A}\} \approx \{\text{B}\}$ when constellations share substantial pattern structure

- Enables translation
- Grounds communication
- Explains why reference coordination succeeds

4. The Developmental Path: From Experience to Reference

4.1. Complete Sequence

Stage 1: $\{A\}$ - Pre-linguistic pattern-constellation - $E(\{A\})$ events - Pure sensory-motor-emotional integration - Present in animals, pre-linguistic infants

Stage 2: $\{A, a^o\}$ - Word integrated as label through Hebbian learning - a^o becomes part of $\{A\}$ - $E(\{A, a^o\})$ - experiencing with label - Primitive referential function - Child can name but not yet conceptualize abstractly - Animals can make spefic sounds as labels

Stage 3: $\{A, a^o, a^c\}$ - Abstract conceptual function emerges - Same word, dual modes - $E(\{A, a^o\})$ - experiencing with label - $\exists(\{A, a^c\})$ - conceptual/abstract linguistic use - Participates in Linguistic PCs - Can talk about A when absent

Stage 4: $\mathcal{R}(a^c, a^o)$ - Reference proper - Conceptual mode coordinates with referential mode - Internal to language but grounded via $\{\text{A}\}$ - Meta-linguistic capability - "What does 'a' mean/refer to?"

This is the genetic/developmental priority:

- $\{A\}$ precedes $\{A, a^o\}$ precedes $\{A, a^o, a^c\}$ precedes $\mathcal{R}(a^c, a^o)$
- x^o (labeling) precedes x^c (conceptualizing) precedes $\mathcal{R}(x^c, x^o)$ (reference)

4.2. Key Transitions

$\{A\} \rightarrow \{A, a^o\}$: Language learning

Repeated co-occurrence: $E(\{A\}) \wedge$ hearing "a" $\rightarrow$ Hebbian: "a" wires with A patterns $\rightarrow$ {A, aˆo} formed $\rightarrow E(\{A, a^o\})$ now possible

$\{A, a^o\} \rightarrow \{A, a^o, a^c\}$: Conceptual development

Using "a" in varied contexts: - Categorization: "Dogs are animals" - Absence: "Where is the dog?" - Comparison: "Like a dog" - Abstract: "Dogness"

$\rightarrow a^c$ function emerges $\rightarrow \exists(\{A, a^c\})$ now possible

$\{A, a^o, a^c\} \rightarrow \mathcal{R}(a^c, a^o)$: Meta-linguistic awareness

Explicit coordination between modes: - "What does 'dog' mean?" - "A dog is a dog" - Philosophical reflection on reference

$\rightarrow \mathcal{R}(a^c, a^o)$ possible

4.3. Empirical Evidence

Rod/Cap structure (Empirical observed manifolds in sentence embedding space, see essay on subject) validates this sequence.

Rods: x^o like uses (concrete, literal, label-like) (Stage 2)

- Tight clustering around word
- Stable, sensory-grounded
- "The dog ran" - dog^o in $E(\{DOG, dog^o\})$ - is a sentence embedded in the rod manifold.

Caps: x^c uses (metaphorical, idiomatic, abstract) (Stage 3)

- Diffuse, context-dependent
- Abstract, conceptual
- "Dogs are animals" - dog^c in $\exists(\{DOG, dog^c\})$ - sentence embedded in the cap manifold

LLMs: develop the rod manifold, tight grouped embeddings in the sentence embedding space for literal uses, and the cap manifold - difuse, large, loosely grouped embeddings for metaphorical or idiomatic uses of the word.

- Only $\exists$ realm (no E)
- Only x^c mode; no sensory-grounded, $x^o$is emulated -developed word "rods" to account for literal, concrete use in sentences (x^o substitution)

- Can compute $\mathcal{R}$ formally but without grounding (see below)

5. Reference vs. Predication

5.1. Pure Reference

Form: $\mathcal{R}(x^c, x^o)$ (same token, different functions)

Definition: The word in conceptual mode picking out the word in referential mode.

Function: Answers "what does word x refer to?"

Grounding:

$\mathcal{R}(x^c, x^o)$ where $x^o \in \{X, x^o, x^c\}$

When $\mathcal{R}$ operates: - x^c (concept mode in $\exists$ realm) coordinates with x^o (label mode in E realm) - x^o activates full constellation $\{X\}$ including sensory patterns - Reference achieved through internal linguistic coordination - Grounded through $\{X\}$ sensory integration learned via $E(\{X, x^o\})$

Examples:

$\mathcal{R}(dog^c, dog^o)$ - "What does 'dog' refer to?" Answer: $\{DOG, dog^o, dog^c\}$ activated via dog^o Includes all sensory-motor-emotional patterns of DOG

$\mathcal{R}(snow^c, snow^o)$ - "What does 'snow' refer to?" Answer: $\{SNOW, snow^o, snow^c\}$ with full sensory integration

$\mathcal{R}(rose^c, rose^o)$ - "A rose is a rose" Shows fundamental reference structure

Why identity statements matter:

"A rose is a rose" = $\mathcal{R}(rose^c, rose^o)$

Not trivial because: - Shows dual functionality (x^c and x^o) - Reveals reference structure - Demonstrates internal linguistic coordination - Points to $\{ROSE\}$ grounding

5.2. Predication

Form: $\mathcal{P}(x^c, y^c)$ where x $\neq$ y

Definition: Using concept x to characterize/describe referent y.

Function: Says something ABOUT y using concept x.

Examples:

$\mathcal{P}(flower^c, rose^c)$ - "A rose is a flower" - Using flower-concept to characterize roses - $flower^c$ operates in $\exists(\{FLOWER, flower^c\})$ - $rose^o$ embedded in $\{ROSE, rose^o, rose^c\}$

$\mathcal{P}(white^c, snow^c)$ - "Snow is white" - Using white-concept to characterize snow - Predication, not pure reference

$\mathcal{P}(animal^c, dog^c)$ - "A dog is an animal" - Relating different constellations

Structure:

$\mathcal{P}(x^c, y^c)$ where x $\neq$ y: - x^c: predicative concept ($\exists$ realm) - y^c: referential target (can have E realm grounding through y^o) - Different words/constellations - Saying: "y has property x" or "y belongs to category x"

Key distinction:

Reference: $\mathcal{R}(x^c, x^o)$ - same token, internal coordination Predication: $\mathcal{P}(x^c, y^c)$ - different tokens, conceptual relation

5.3. Recursive Structures

Left-nesting (toward presence):

$\mathcal{R}(\mathcal{R}(x^c, x^o), x^o)$ $\mathcal{R}(\mathcal{R}(\mathcal{R}(x^c, x^o), x^o), x^o)$

Each application reinforces x^o (referential grounding) Contracts toward $\{X\}$ (sensory-grounded constellation) Movement toward immediate presence Pulling toward $E(\{X, x^o\})$ realm Asymptotic approach (Kant's "thing in itself")

Right-nesting (deferring):

$\mathcal{R}(x^c, \mathcal{R}(x^c, x^o))$ $\mathcal{R}(x^c, \mathcal{R}(x^c, \mathcal{R}(x^c, x^o)))$

Concept-word keeps re-entering Never grounds in final referent Semantic expansion without closure Staying in $\exists(\{X, x^c\})$ realm

Language mediates between:

- E **pole**: x^o, sensory immediacy, $\{X\}$ activation, grounded experience
- $\exists$ **pole**: x^c, conceptual elaboration, LPC operations, abstract thought

6. Linguistic Pattern Constellations (LPCs)

6.1. Motivation

Not all patterns are directly sensory-grounded. Some patterns emerge at the level of linguistic patterns themselves.

Key insight: Mathematical and logical concepts are patterns visible in how we use language, not patterns of direct sensory experience.

6.2. Hierarchy of Patterns

Level 1: Primitive PCs

$\{A\}$ or $\{A, a^o, a^c\}$ = Constellation integrating: - Sensory patterns - Motor patterns
- Emotional patterns - Linguistic patterns (a^o and a^c)

Built through: $E(\{A, a^o\})$ events during learning Mediated through: $\exists(\{A, a^c\})$ for abstract use

Level 2: Linguistic Pattern Constellations (LPCs)

$\langle\langle$a$\rangle\rangle$ = Pattern of how word "a" behaves in linguistic space; for example, it can be shown that all sentences with a specific word (like "cat") are distributed on non-arbitrary manifolds in the sentence embedding space; these are patterns of use for that word. These manifolds have at least two distinct regions for literal, concrete (x^o use) and metaphorical, idiomatic, or abstract (x^c use) sentences.

Can be seen as meta-constellation over primitive PCs Patterns IN linguistic space ($\exists$ realm) Built through observing regularities in $\exists(\{A, a^c\})$ usage Can be written $\langle\langle a^o, a^c \rangle\rangle$ when emphasizing dual modes

Level 3: Mathematical/Logical Symbols

Symbols functioning as x^o for LPCs Patterns of patterns of language

6.3. Notation

Primitive PC: $\{A\}$ and **Simple PCs**: $\{A, a^o\}$, $\{A, a^o, a^c\}$ - braces with capitals. **LPC**: $\langle\langle a^o, a^c \rangle\rangle$ or, simplified,$\langle\langle$a$\rangle\rangle$ - double angle brackets

Parallel structure:

- Primitive: $\{A, a^o\}$ - word embedded in sensory-grounded constellation $\{A\}$
- Linguistic: $\langle\langle$a$\rangle\rangle$ - word's pattern in linguistic space

Both are built around the word "a" but at different levels of abstraction

6.4. Key Difference: E vs $\exists$ for LPCs

Primitive PCs:

- Can have $E(\{A, a^o\})$ - direct experience
- Can have $\exists(\{A, a^c\})$ - linguistic mediation
- Word "a" grounded in sensory experience of A

LPCs:

- Cannot have E($\langle\langle$a$\rangle\rangle$) - no direct sensory experience of linguistic patterns
- Only $\exists$ ($\langle\langle$a$\rangle\rangle$) - exist purely in linguistic realm
- But $\langle\langle$a$\rangle\rangle$ ultimately traced to E events through grounding chain in $E(\{a^o\})$
- The pattern emerges from how "a" is used across many $\exists$ language operations; in turn, these operations are determined by the sum of E experience across all people.

7. Complete Formal System Summary

7.1. Core Ontology

Experiential Realm (E):

- $E(\{A\})$: Pre-linguistic pattern-recognition
- $E(\{A, a^o\})$: Experience with integrated label
- Sensory-motor-emotional grounding
- x^omode operates here
- Hopfield attractor states

Linguistic Realm ($\exists$):

- $\exists(\{A, a^c\})$: Linguistic/conceptual mediation
- Abstract, can operate detached from immediate E
- x^c mode operates here
- Linguistic PCs (LPCs) emerge here; x^c functions as label for patterns of use for word x, $\langle\langle$x$\rangle\rangle$

Fundamental Transition:

- $E(\{A, a^o\}) \rightarrow \exists(\{A, a^c\})$: From grounded to abstract
- From x^o to x^c
- From immediate to mediated

Pattern-Constellations:

- $\{A\}$: Pre-linguistic (Stage 1) - capital letters for clarity
- $\{A, a^o\}$: Label integrated (Stage 2) -$\{A, a^o, a^c\}$: Dual function (Stage 3)
- $\langle\langle \text{a} \rangle\rangle$: Linguistic PC (meta-level)

Operations:

- $\mathcal{R}(x^c, x^o)$: Pure reference (answers "what does x refer to?")
- $\mathcal{R}(x^c, y^c)$: Predication (says something about y)

7.2. Developmental Hierarchy

Level 0: Sensory-motor experience Pre-linguistic interaction with environment

Level 1: $E(\{A\})$ Primitive pattern-recognition Animals, pre-linguistic infants

Level 2: $E(\{A, a^o\})$ Label integrated through learning Word fires with sensory patterns Hebbian: co-activation $\rightarrow$ co-wiring

Level 3: $\exists(\{A, a^c\})$ Abstract/conceptual linguistic use Can talk about A when absent Participates in LPCs

Level 4: $\mathcal{R}(a^c, a^o)$ Explicit reference operation Meta-linguistic awareness "What does 'a' refer to?"

Level 5: Predication $\mathcal{R}(x^c, y^o)$ Relating different PCs "y is x"

Level 6: $\langle\langle \text{a} \rangle\rangle$ formation LPCs emerge from $\exists$ patterns Patterns of language use

Level 7: Abstract symbols Act for LPCs as x^o acts for PCs Mathematics, logic

7.3. Key Relationships

Embedding:

$x^o \in \{X, x^o, x^c\}$

x^obuilt through: $E(\{X, x^o\})$ events x^oactivates: full $\{X\}$ constellation

Coordination:

$\mathcal{R}(x^c, x^o)$ coordinates:_ $\exists$ realm (x^c) with E realm grounding (x^o)

Reference succeeds through: Constellation activation built from E events

Translation:

$\{A\}$speaker $\approx \{B\}$ hearer when: overlap of (E events) sufficient

$\exists(\{A, a^c\})$speaker coordinates with $\exists(\{B, b^c\})$hearer via: Similar $E(\{A, a^o\})$ and $E(\{B, b^o\})$ histories

florin.cojocariu@s.unibuc.ro, October 1, 2025