Multi-agent language coordination and cognition

Nina Gierasimczuk
(results based on joint work with Dariusz Kalociński, Franek Rakowski, and Jakub Uszyński)

DTU Compute
Technical University of Denmark

IDA: Driving AI, March 18, 2022

Semantic Universals

Linguistic universals:

- properties shared among different natural languages
- observed in phonology, syntax, and semantics;
- facilitate (logical) reasoning.

Linguistic universals:

- properties shared among different natural languages
- observed in phonology, syntax, and semantics;
- facilitate (logical) reasoning.

The source:

- language is in a way a product of human cognition, so
- (possible) languages that do not accommodate cognitive constraints that underlie linguistic universals are evolutionarily extinct;
- explanation may require more than one influencing factor, e.g., balance between processing costs and expressiveness.

Convexity and Monotonicity

Convexity of Colour Terms (e.g., *red* or *blue*)

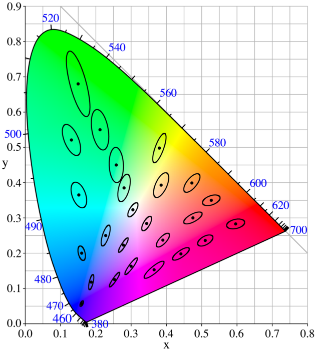
- Colour terms are associated with regions of colour space.
- In that space the distance between any two points is well-defined.
- Convexity: if two points belong to a region,
  then any point on the shortest route between those points is in that region.

Convexity universal: natural simple colour terms are convex.

Convexity of Colour Terms (e.g., *red* or *blue*)

- Colour terms are associated with regions of colour space.
- In that space the distance between any two points is well-defined.
- Convexity: if two points belong to a region,
  then any point on the shortest route between those points is in that region.

Convexity universal: natural simple colour terms are convex.



Convexity and Monotonicity

Monotonicity of Gradable Adjectives (e.g., *tall* or *cold*)

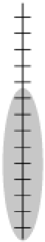
- Meanings are identified with subsets of linearly ordered sets of degrees.
- Monotonicity: if truthful application of the adjective to a given degree extends to any greater (or lesser) degree.

Monotonicity universal: natural simple gradable adjectives are monotone.

Monotonicity of Gradable Adjectives (e.g., *tall* or *cold*)

- Meanings are identified with subsets of linearly ordered sets of degrees.
- Monotonicity: if truthful application of the adjective to a given degree extends to any greater (or lesser) degree.

Monotonicity universal: natural simple gradable adjectives are monotone.

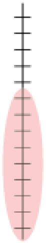


Convexity and Monotonicity

Monotonicity of Gradable Adjectives (e.g., *tall* or *cold*)

- Meanings are identified with subsets of linearly ordered sets of degrees.
- Monotonicity: if truthful application of the adjective to a given degree extends to any greater (or lesser) degree.

Monotonicity universal: natural simple gradable adjectives are monotone.

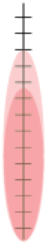


Convexity and Monotonicity

Monotonicity of Gradable Adjectives (e.g., *tall* or *cold*)

- Meanings are identified with subsets of linearly ordered sets of degrees.
- Monotonicity: if truthful application of the adjective to a given degree extends to any greater (or lesser) degree.

Monotonicity universal: natural simple gradable adjectives are monotone.

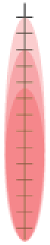


Convexity and Monotonicity

Monotonicity of Gradable Adjectives (e.g., *tall* or *cold*)

- Meanings are identified with subsets of linearly ordered sets of degrees.
- Monotonicity: if truthful application of the adjective to a given degree extends to any greater (or lesser) degree.

Monotonicity universal: natural simple gradable adjectives are monotone.

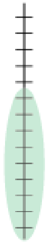


Convexity and Monotonicity

Monotonicity of Gradable Adjectives (e.g., *tall* or *cold*)

- Meanings are identified with subsets of linearly ordered sets of degrees.
- Monotonicity: if truthful application of the adjective to a given degree extends to any greater (or lesser) degree.

Monotonicity universal: natural simple gradable adjectives are monotone.

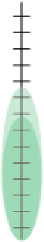


Convexity and Monotonicity

Monotonicity of Gradable Adjectives (e.g., *tall* or *cold*)

- Meanings are identified with subsets of linearly ordered sets of degrees.
- Monotonicity: if truthful application of the adjective to a given degree extends to any greater (or lesser) degree.

Monotonicity universal: natural simple gradable adjectives are monotone.



Convexity and Monotonicity

Monotonicity of Gradable Adjectives (e.g., *tall* or *cold*)

- Meanings are identified with subsets of linearly ordered sets of degrees.
- Monotonicity: if truthful application of the adjective to a given degree extends to any greater (or lesser) degree.

Monotonicity universal: natural simple gradable adjectives are monotone.

Quantity terms

Natural languages include a variety of quantity terms, among them:

- numerals (e.g., *one, two, three*)
- quantifiers (e.g., *at least two*, *a few*, *half of*).

Quantity terms

Natural languages include a variety of quantity terms, among them:

- numerals (e.g., *one, two, three*)
- quantifiers (e.g., *at least two*, *a few*, *half of*).

Even the societies lacking developed number systems
use words denoting initial natural numbers.

Quantity terms

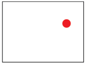
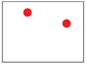
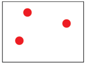
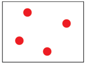
Natural languages include a variety of quantity terms, among them:

- numerals (e.g., *one, two, three*)
- quantifiers (e.g., *at least two*, *a few*, *half of*).

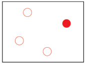
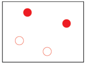
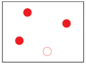
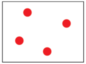
Even the societies lacking developed number systems
use words denoting initial natural numbers.

Theoretical studies of quantity terms in mathematical logic (GQ Theory) give ways to isolate and rigorously define candidates for such universals: among them convexity and monotonicity.

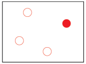
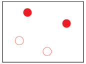
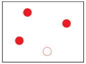
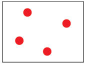
Extensions of quantity terms

model	numeral
	*one*
	*two*
	*three*
	*four*

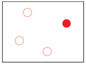
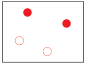
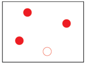
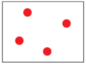
Extensions of quantity terms

model	numeral	proportion
	*one*	1/4
	*two*	1/2
	*three*	3/4
	*four*	1

Extensions of quantity terms

model	numeral	proportion	quantifiers
	*one*	1/4	*some*
	*two*	1/2	*some, half*
	*three*	3/4	*some, most*
	*four*	1	*some, all, most*

Extensions of quantity terms

model	numeral	proportion	quantifiers
	*one*	1/4	*some*
	*two*	1/2	*some, half*
	*three*	3/4	*some, most*
	*four*	1	*some, all, most*

Straightforward link to (generalized) quantifiers of type $\langle 1 \rangle$ and $\langle 1, 1 \rangle$.

Universals in the number triangle space

(0,0)

(1,0) (0,1)

(2,0) (1,1) (0,2)

(3,0) (2,1) (1,2) (0,3)

(4,0) (3,1) (2,2) (1,3) (0,4)

. . . .

Number Cognition: Approximate Number Sense

How to ground such abstract semantics in cognition?

Human cognition is hypothesised to be equipped with an evolutionarily old mechanism of number cognition: the **approximate number sense** (ANS, for short).

Number Cognition: Approximate Number Sense

How to ground such abstract semantics in cognition?

Human cognition is hypothesised to be equipped with
an evolutionarily old mechanism of number cognition:
the **approximate number sense** (ANS, for short).

Check it out, it's fascinating and explains a lot about a lot!

Number Cognition: Approximate Number Sense

How to ground such abstract semantics in cognition?

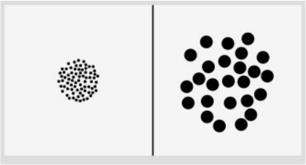
Human cognition is hypothesised to be equipped with
an evolutionarily old mechanism of number cognition:
the **approximate number sense** (ANS, for short).

Check it out, it's fascinating and explains a lot about a lot!

From a purely functional perspective:

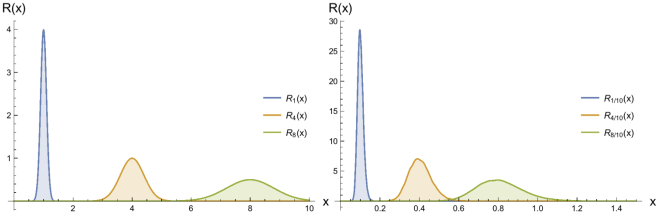
- ANS allows for instant perception of quantities at the cost of accuracy,
- with an error proportional to the intensity of the perceived input.

Number Cognition: Approximate Number Sense



"Which side has more dots?"

The Model: Reactive units

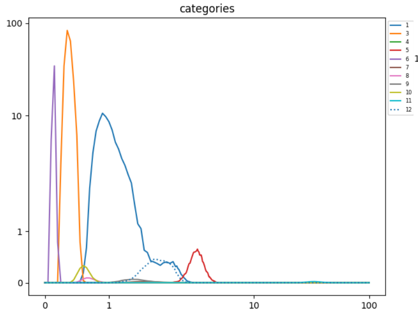


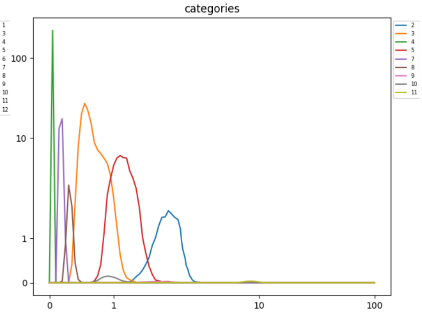
The Model: Single-agent level

On the level of a single agent:

- perceives stimuli according to the ANS activation pattern (reactive unit);
- groups reactive patterns into **categories (concepts)**;
- uses a category to discriminate between stimuli (discrimination game);
- each agent has her own language (binding of words to categories).

The Model: Examples of emerged categories





Language

Let F be a dictionary and C a set of categories.

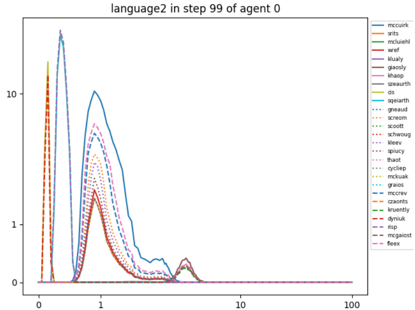
Language is a function $L : F \times C \to [0, \infty)$.

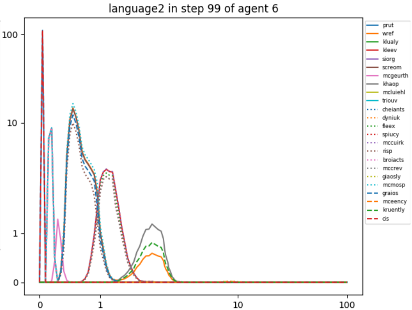
$L(f, c)$ is the strength of the connection between word f and category c.

Language is a dynamical object that changes during the lifetime of an agent:

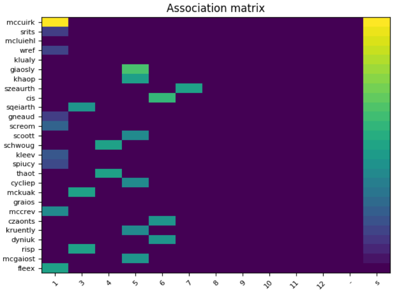
Additions and deletions in dictionary and categories,
and varying the strengths of connections.

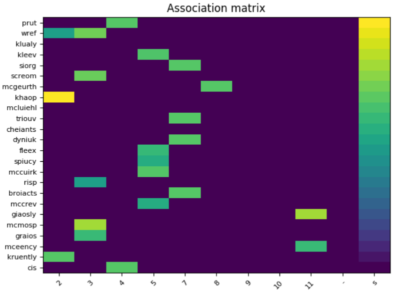
The Model: Examples of Emerged Languages





THE MODEL: CATEGORY-WORD BINDING





The Model: The Coordination Game

On the level of interaction between agents:

- two agents (speaker and hearer) meet on a blind date;
- they are shown two stimuli (one of them is topic-known to speaker);
- **speaker:**
  - perceives the stimuli;
  - finds a best category that distinguishes the topic from the other;
  - finds a word that best corresponds to it in her language;
  - utters that word.
- **hearer:**
  - looks up the uttered word in his language;
  - looks up the category that has the strongest binding;
  - points to the stimulus with the highest response in the category.

Success: The correct associations are increased, other are decreased.
Failure: The guessed associations are decreased.

Meaning and pragmatic meaning

Meaning: between-word interaction

$$[f]^L =_{df} \{q \in Q : \Sigma_{c \in C(L)} L(f, c)\langle c|R_q\rangle > 0\}.$$

In words, a stimulus q contributes to the meaning of f if some category c associated with f gives a positive response to q.

Pragmatic Meaning: context-dependence

$$[f]^L_p = \{q : \exists\, c\, [f \in \underset{\substack{f' \in F(L) \\ L(f',c)>0}}{\text{argmax}} L(f', c), c \in \underset{\substack{c \in C(L) \\ \langle c|R_q\rangle > 0}}{\text{argmax}} \langle c|R_q\rangle]\}.$$

In words, q is an element of the pragmatic meaning of f if f is maximally (and non-negatively) associated with c that gives a maximal (and non-negative) response to q.

Convexity

A strictly ordered set of stimuli $(Q, <)$ with the *less than* relation.

The **pragmatic meaning** of f in L is convex if it is a convex set in $(Q, <)$.

Convexity

A strictly ordered set of stimuli $(Q, <)$ with the *less than* relation.

The **pragmatic meaning** of f in L is convex if it is a convex set in $(Q, <)$.

Monotonicity

$$\text{Many good chess players know advanced tactics.} \tag{1}$$

$$\text{Many good chess players know tactics.} \tag{2}$$

Convexity

A strictly ordered set of stimuli $(Q, <)$ with the *less than* relation.

The **pragmatic meaning** of f in L is convex if it is a convex set in $(Q, <)$.

Monotonicity

Few beginning chess players know tactics. (1)

Few beginning chess players know advanced tactics. (2)

Convexity

A strictly ordered set of stimuli $(Q, <)$ with the *less than* relation.

The **pragmatic meaning** of f in L is convex if it is a convex set in $(Q, <)$.

Monotonicity

We attribute monotonicity to **meanings** rather than pragmatic meanings:

pragmatic meaning represents only the fragment of the overall meaning—the one that is most contextually and linguistically salient. For example, even though *most* can be truly used in situations where all objects posses a given property, it then might make more sense to use *all* instead.

Convexity and monotonicity

Convexity

A strictly ordered set of stimuli $(Q, <)$ with the *less than* relation.

The **pragmatic meaning** of f in L is convex if it is a convex set in $(Q, <)$.

Monotonicity

Let f be a word in language L. We say that $[f]^L$ is **monotone** if it is upward closed with respect to $\leq$ (i.e., if $q \in [f]^L$ and $q \leq q'$ then $q' \in [f]^L$) or downward closed with respect to $\leq$ (analogously).

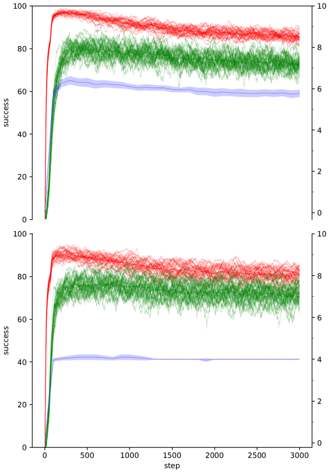
Experiment

- The code written in Python and Mathematica.
- A simulation consists of 30 trials.
- Within a trial, 10 agents evolve across 3000 steps.
- At each step agents are paired randomly for a guessing game.
- Separate simulations for **numeric-based** and **quotient-based** stimuli.
- Varied conditions of agents **with** and **without the ANS**.
- This gives us 4 conditions.

Results: Model Validity

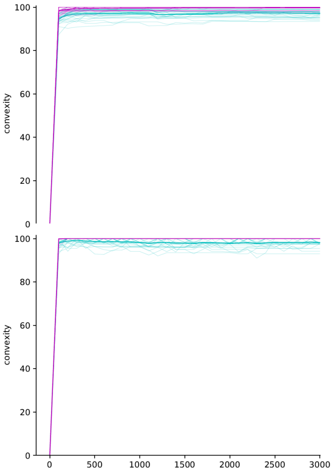
- numeric (top), quotient (bottom);
- cumulative discriminative success,
- communicative success,
- mean size of active lexicon (right)

Each red/green line corresponds to a single trial.



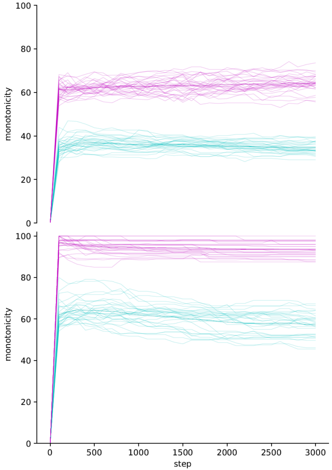
Results: Convexity

- population-level convexity for each trial
- numeric (top), quotient (bottom)
- with ANS and without ANS



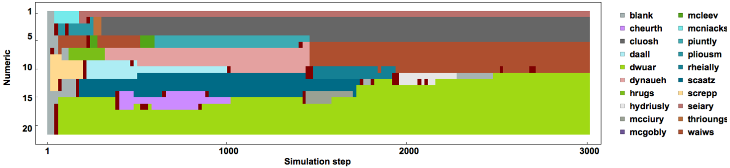
Results: Monotonicity

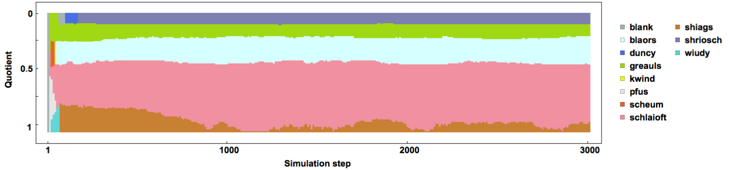
- population-level monotonicity for each trial
- numeric (top), quotient (bottom)
- with ANS and without ANS



Results: Language over Time

Active lexicon for numeric (top) and quotient stimuli (bottom) in the ANS condition for agent 4.





Results: Language across Agents

Active lexicons of all agents from a selected trial at the last simulation step for quotient stimuli.

Conclusions

- Model based on the existing approach to colour expressions (Steels 2005).
- Perceptual layer inspired by ANS.
- Agent-based simulations result in communicatively usable lexicons.

Monotonocity

- ANS facilitates monotonicity.
- Why? ANS extends the scope of a category, especially for larger inputs, because the long-tailed categories can provide positive responses even for distant quantities. Moreover, the more relaxed, vague semantics of quantity expressions under ANS likely allows an easier upward or downward merge with other categories.

Convexity

- ANS does not significantly facilitate convexity.
- But it does to a larger extent than precise number perception.
- In general convexity must stem from other layers of the model.

Thank you!