Inteligenţă Artificială
Cursurile 3-4

Putem modela activități cognitive?

Curs: Dan Cristea

Laboratoare: Ionuţ Pistol, Mădălina Răschip, Marius Zbancioc

AlterEgo

- Un agent înzestrat cu funcții cognitive, vitale, comunicative, procesuale asemănătoare omului
- Cuplat la aceeași realitate fizică la care sunt eu conectat
- Care copiază (pațial sau total) personalitatea mea
- Care crește odată cu mine și reprezintă un fel de oglindă a mea (frate geamăn sau alter ego)

Jocul

- Dezvoltă acest agent ca un joc aflat la dispoziția mea:
  - pot controla o unică instanță
  - trebuie să modelez personalitatea acestui agent astfel încât el să reproducă cât mai fidel comportamentul meu
    - => să previzioneze hotărârile pe care le voi lua eu însumi

Igor Aleksander

- Professor of Artificial Intelligence at Imperial College
  - Impossible Minds. My Neurons, My Consciousness, Imperial College Press, 1996, 2015
  - Artificial Neuroconsciousness: An Update, http://web.archive.org/web/19970302014628/http://www.ee.ic.ac.uk/research/neural/publications/iwann.html

Postulatul fundamental: conștiința și activitatea neuronală

- *The personal sensations that lead to the consciousness of an organism are due to the firing patterns of some neurons, such neurons being part of a larger number which form the state variables of a neural state machine, the firing patterns having been learned through a transfer of activity between sensory input neurons and the state neurons.*

Corollary 1: The brain is a state machine

- *The brain of a conscious organism is a state machine whose state variables are the outputs of neurons. This implies that a definition of consciousness be developed in terms of the elements of a state machine theory.*
  - Such machines can be probabilistic, where links between states are defined as probabilities, they can have a finite or an infinite number of states.
  - Define the characteristics of state structure that are necessary for and specific to organisms that are said to be conscious.

Corollary 2: Inner Neuron Partitioning

- *The inner neurons of a conscious organism are partitioned into at least three sets:*
  - Perceptual Inner Neurons: *responsible for perception and perceptual memory;*
  - Auxiliary Inner Neurons: *responsible for inner 'labelling' perceptual events;*
  - Functional Inner Neurons: *responsible for 'life-support' functions - not involved in consciousness.*

Corollary 3: Conscious and Unconscious States

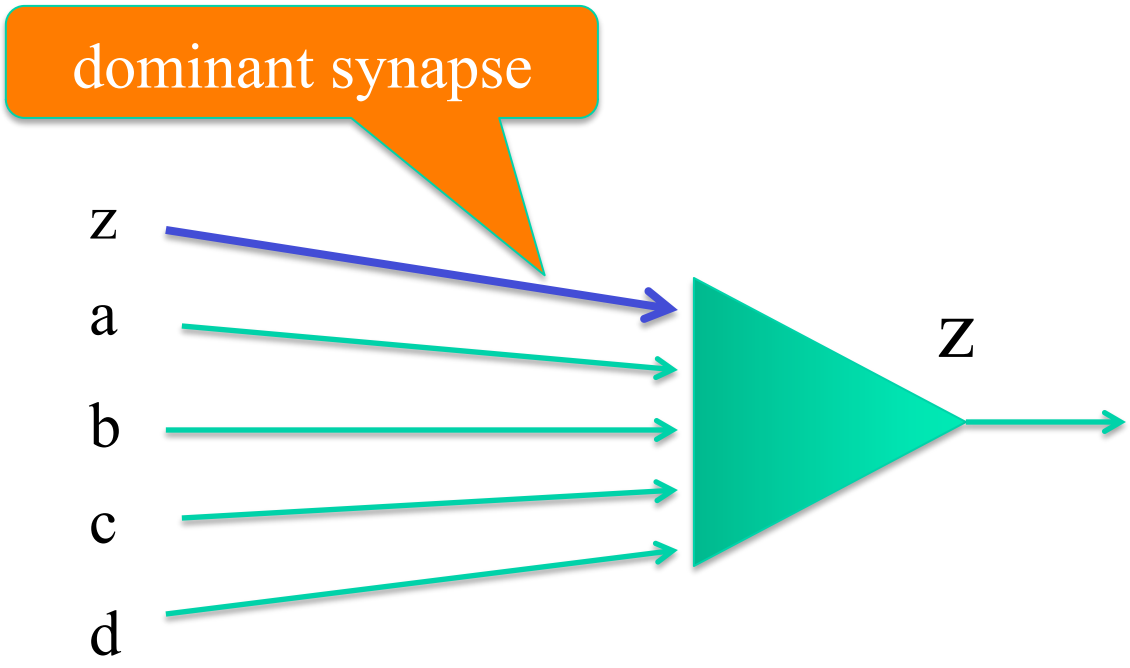
- *Consciousness in a conscious organism resides directly in the perceptual inner neurons in two fundamental modes:*
  - Perceptual: *which is active during perception - when sensory neurons are active;*
  - Mental: *which is active even when sensory neurons are inactive. The activity of the inner perceptual neurons ranges over the same states in both these modes.*

=>The same perceptual neurons can enter semi-conscious or unconscious states that are not related to perception.

Corollary 4: Perceptual Learning and Memory

- *Perception is a process of the input* sensory *neurons causing selected perceptual* inner *neurons to fire and others not. This firing pattern on inner neurons is the inner representation of the percept - that which is felt by the conscious organism.*
  - ***Learning*** *is a process of adapting not only to the firing of the input neurons, but also to the firing patterns of the other perceptual inner neurons.*
  - ***Generalisation*** *in the neurons (i.e. responding to patterns similar to the learnt ones) leads to representations of world states being self-sustained in the inner neurons and capable of being triggered by inputs similar to those learned originally.*

Learning



z=1 => learning (the “configuration” of <a, b, c, d>)
z=0 => reproducing (generalisation)

Comentarii până aici

- Conștiința este un rezultat al activității neuronale
- Aceleași patternuri de aprindere ale neuronilor pot fi obținute pe mai multe căi
- Putem vorbi de conștiință doar dacă senzații similare celor produse de percepții pot fi susținute în rețea în absența intrărilor
- Nu toate stările neuronilor perceptuali interni au corelații senzoriale => visul, starea de anestezie, activitatea subconștientă

Surse de inspirație

- Proiecte care au legătură cu ideea noastră
  - ALEAR: comunități de roboți care dezvoltă spontan limbajul
  - POETICON++: robot umanoid inzestrat cu capacități perceptuale (vizuale, tactile, auditive) ale cărui mișcări sunt controlate prin limbaj
- Teorii ajutătoare:
  - rețele neuronale, automate probabiliste

ALEAR (**Artificial Language Evolution on Autonomous Robots**) – an FP7 project

Achievement of open-ended cognitive development and open-ended verbal dialogues among fully embodied situated agents (humanoid robots, mechanisms which include sensori-motor intelligence, scripts for establishing the turn-taking interaction among them, perceptual processes, processes that perform the conceptualisation of what to say, the expression of these conceptualisations in language and processes that perform the parsing of sentences and their interpretation in sensori-motor experience). ALEAR proved that humanoid robots may evolve their own artificial languages adapted to the environment and task settings in which they are placed.

ALEAR - partners

- Humboldt University, Berlin - coordinator
- SONY CSL – Paris (scientific coordinator: prof. Luc Steels)
- Osnabruck University
- Autonomous University of Barcelona
- Vrije Universiteit Brussel – Brussels
- "Alexandru Ioan Cuza" University of Iasi

An Evolutionist View Over Conceptualization in Language

Dan Cristea
"Alexandru Ioan Cuza" University of Iaşi, Dept. of Computer Science
Romanian Academy, Institute for Computer Science
dcristea@info.uaic.ro

What is this talk about?

- The magic words in this workshop:
  - classification, hierarchies, conceptualisation …

- How concepts aroused in humans and how have they been expressed in language?
  - Why do we speak the way we do?
  - Why are there so many languages although most of us operate with the same concepts?
  - Is there a way to prove scientifically hypotheses about the evolution of languages?

The "Talking Heads" experiment

- Goal: how has the language evolved?
  - In a community, over time: language emerges through self organisation
  - In individuals, meaning is build in a cumulative growth process

ALEAR – Main Objective

"Carefully controlled experiments in which autonomous humanoid robots self-organise rich conceptual frameworks and communication systems with similar features as those found in human languages."

Approach

- *Whole systems approach*: experiment with robots => physical embodiment, sensori-motor, perception, conceptualisation, language
- *Self-generated*: as opposed to designed or acquired through inductive machine learning
- *Multi-agent*: as opposed to stand-alone
- *Evolutionary*: start from scratch and see how a communication system forms and further develops

Setting

- Cognitive agents:
  - Physical aspects:
    - body
    - sensors
    - articulators
    - physical location
    - objects and agents located in the environment
  - Mental properties
    - behaviour
    - memory
    - lexicon
    - grammar
    - etc.
- The two aspects are separated: a real agent exists only when a virtual agent is loaded in a physical robot body

The robots

- Physically embodied autonomous agents
  - Motor-sensory processing
    - perception,
    - movements,
    - actions
  - Conceptual processing
    - recognise objects,
    - learn a lexicon,
    - build representations of concepts
    - towards the development of grammar

Teleporting

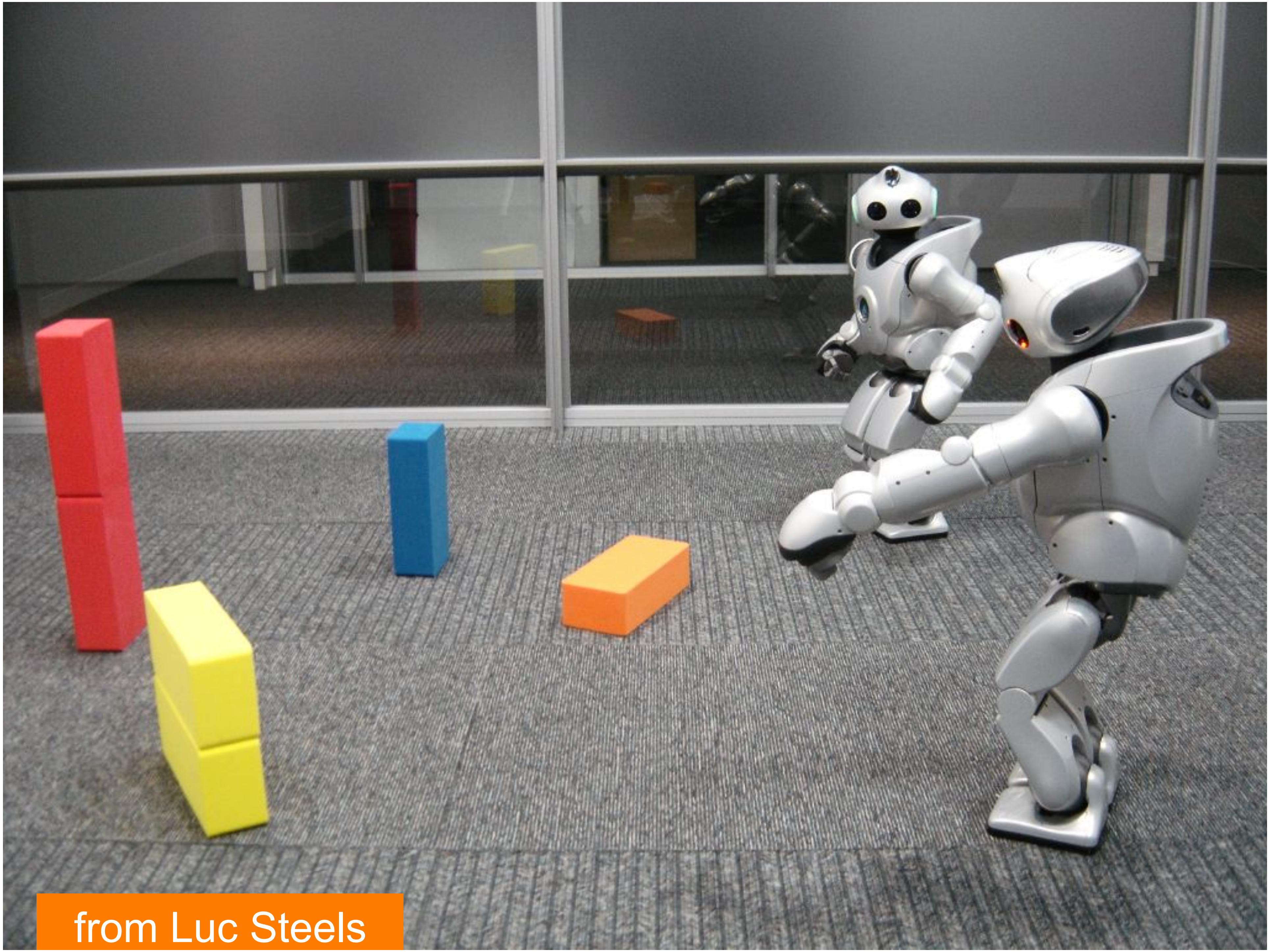
- Develop categories in one location and enrich his learning experience by moving to another location
- The transmission of language from one generation to the next
- Intercultural exchange and language contacts by migrating mental bodies in different parts of the world

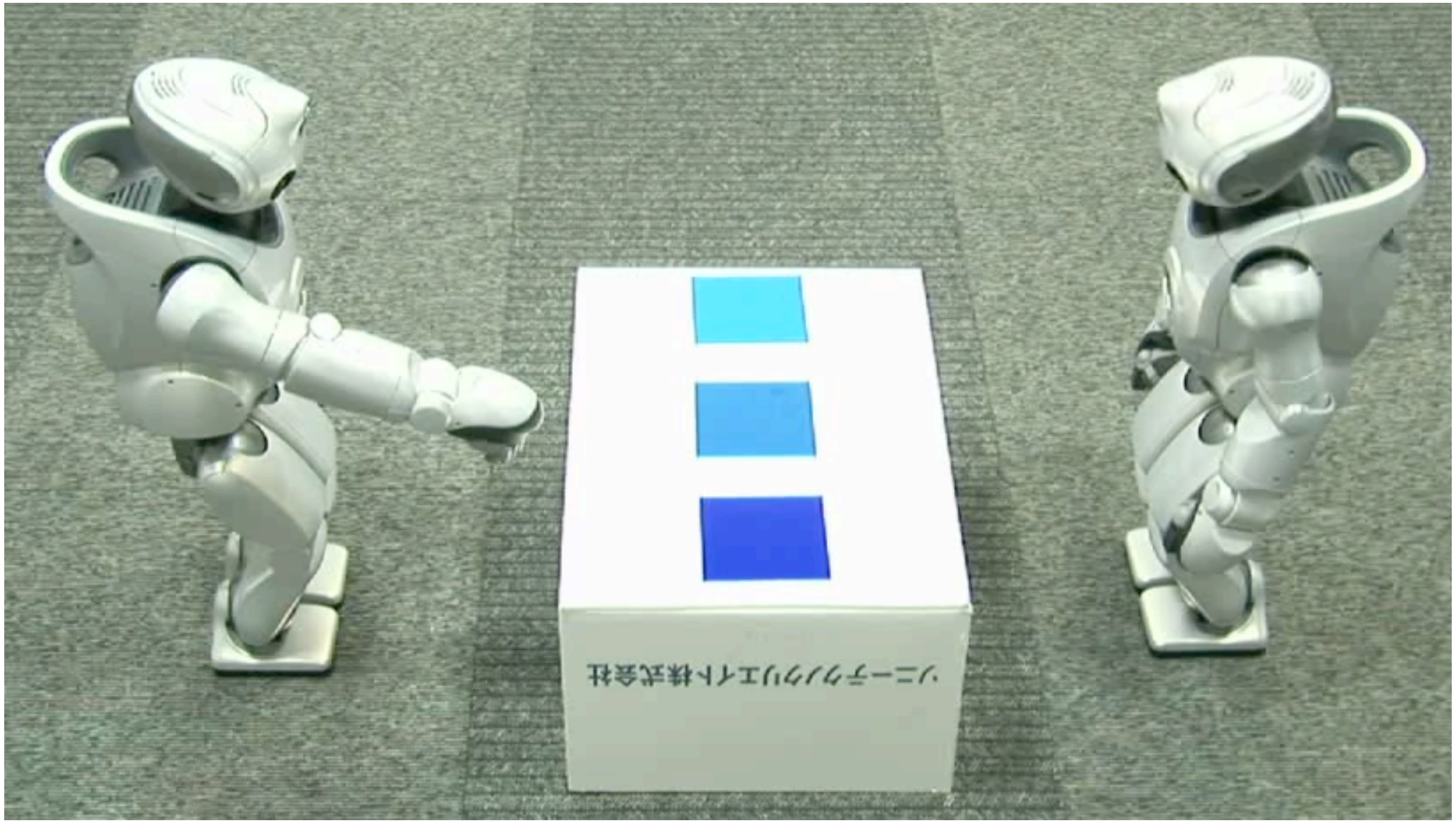
The “guessing game”

- Two physically instantiated agents: *speaker* and *hearer*
- Why *game*?
  - Because neither agent can look into the mind of the other. They only interact through the external environment
- What triggers the game behaviour?
  - There is an innate motivation programmed: agents try to maximise their communicative success

Rules of the game

- A table on which objects of different forms and colors are projected
  - agents in different virtual states queue up for places in physical bodies
  - a game: a *speaker* and a *hearer*
    - both speaker and hearer visualise the area: *the context*
    - the speaker chooses one object, *the topic*, and gives a *verbal hint* to the hearer
    - the hearer then has to guess the object
    - the game succeeds if the identified object by the hearer is the one chosen by the speaker
    - otherwise, it fails: the speaker points to the topic, the hearer tries to repair his internal state to be more successful in the future
      - the speaker weakens his hypothesis that the words he has used are right
      - the hearer tries to guess what meaning the speaker might have used


from Luc Steels



The Mondriaan Experiment

from Luc Steels

What happens when agents play a language game

- Speaker perceives the scene
- Speaker conceptualises what to say
- Speaker applies inventory to produce sentence
- Sentence transmitted to hearer

- Hearer decodes sentence
- Hearer parses sentence using own inventory
- Hearer applies meaning to his own perception of the scene

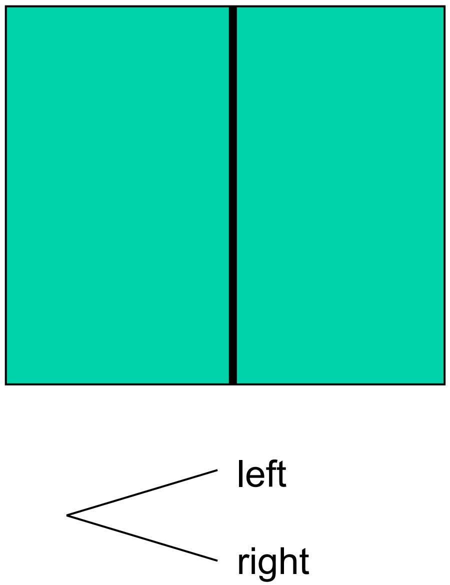
Perception and categorisation

- Sensory channels
  - software processes interpreting specific real world information:
    - HPOS (horizontal position)
    - VPOS (vertical position)
    - GRAY (gray level)
    - others
  - domain: 0-10 (continuous) discretised to a number of discrete values ➔ categorisation
  - categorisation could be specific to individuals

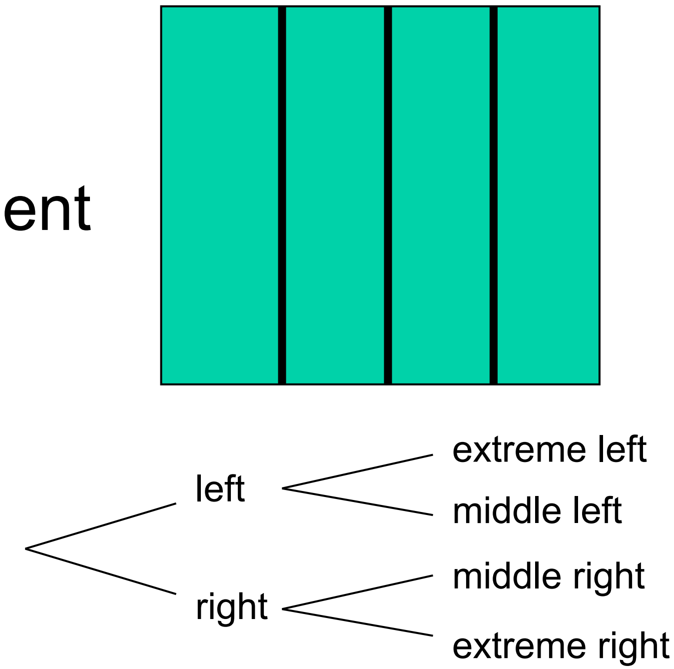
The categories trees

- Each individual can develop his own tree of categories for each sensory channel
- Example for HPOS:

agent A

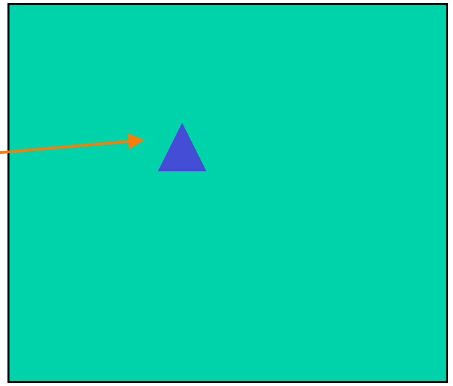


agent B



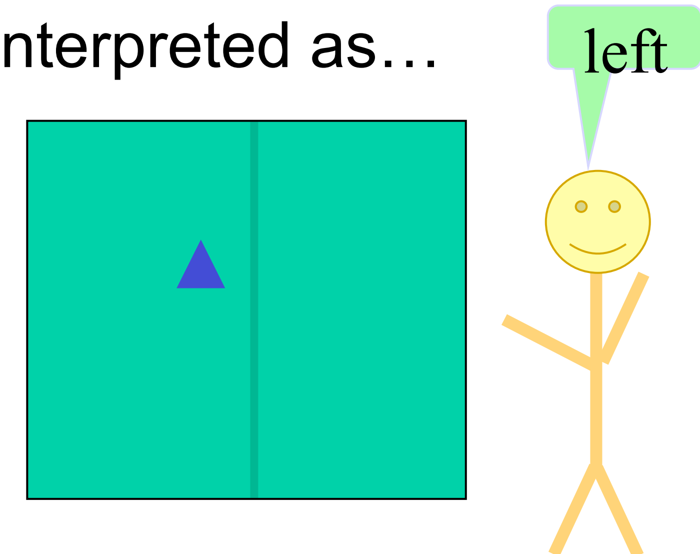
Categorisation in individuals

This object

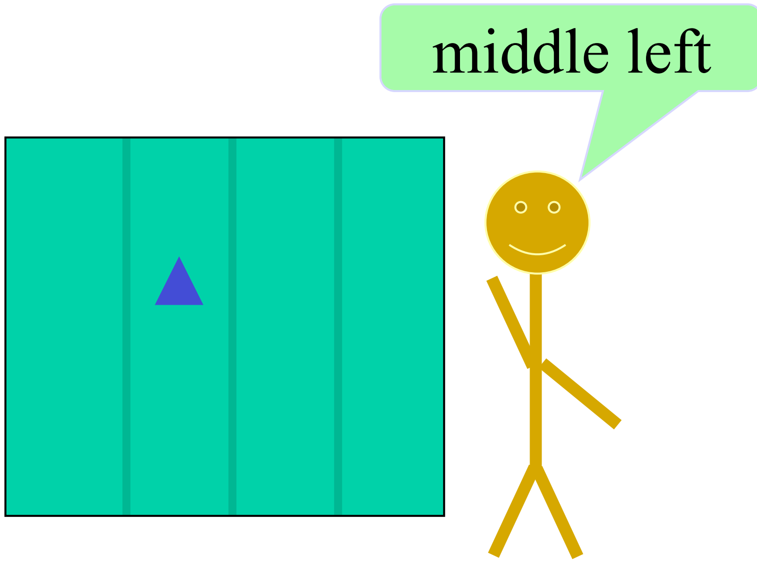


Important notice: the symbols **left** and **middle left** are our conventions to notate the HPOS property values – they do not belong to the acquired lexicon!

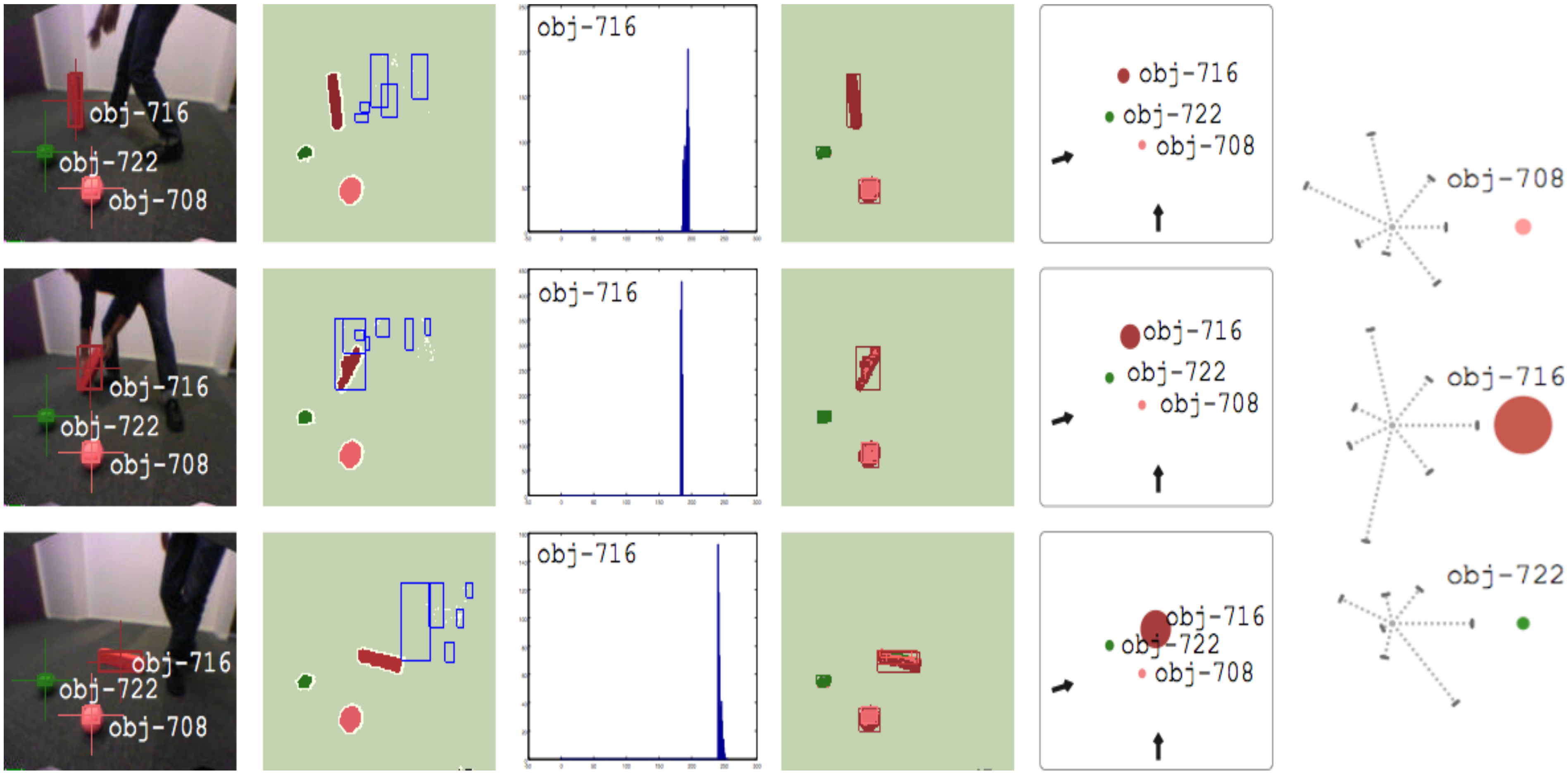
will be interpreted as…



by agent A

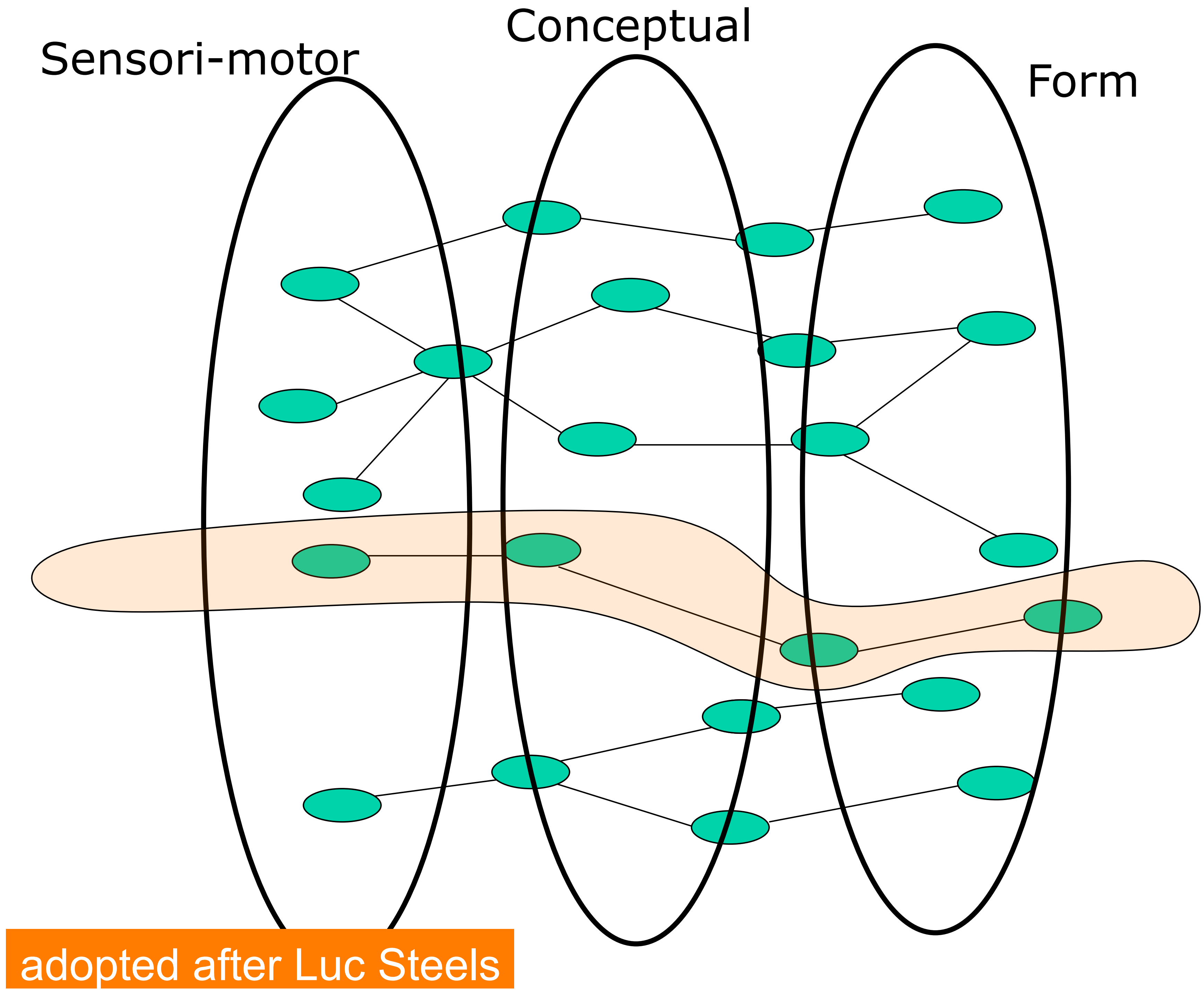


by agent B

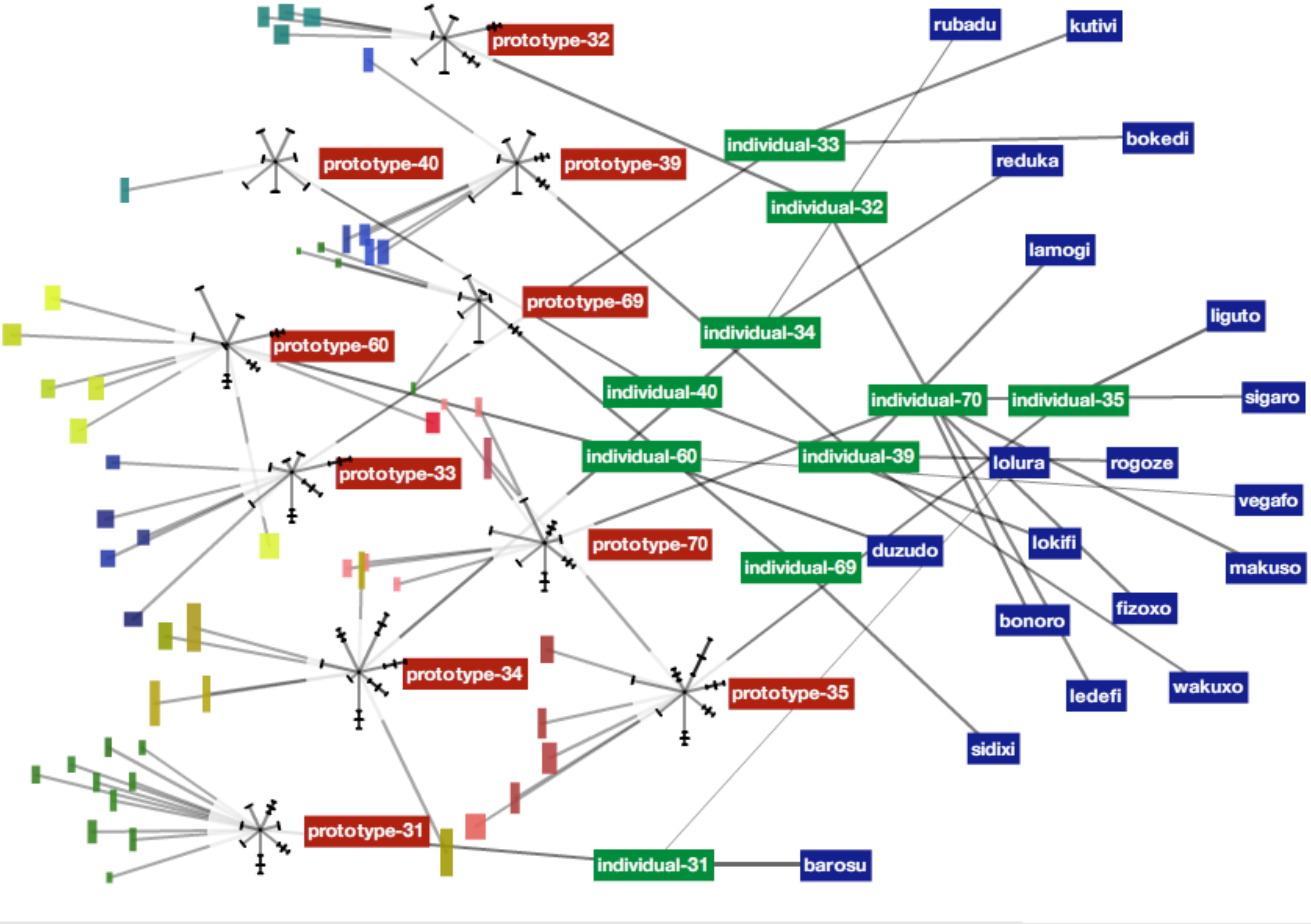


Sensory experience: feature values of segment along various dimensions

from Luc Steels

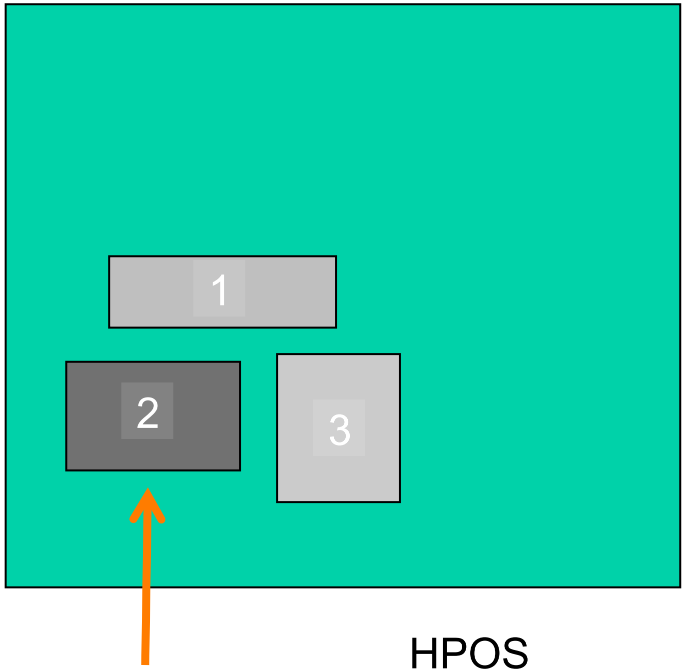


Sensori-motor
Conceptual
Form
adopted after Luc Steels



from Luc Steels

About perception again: why do we use some features and not others?



HPOS

VPOS

HIGHT

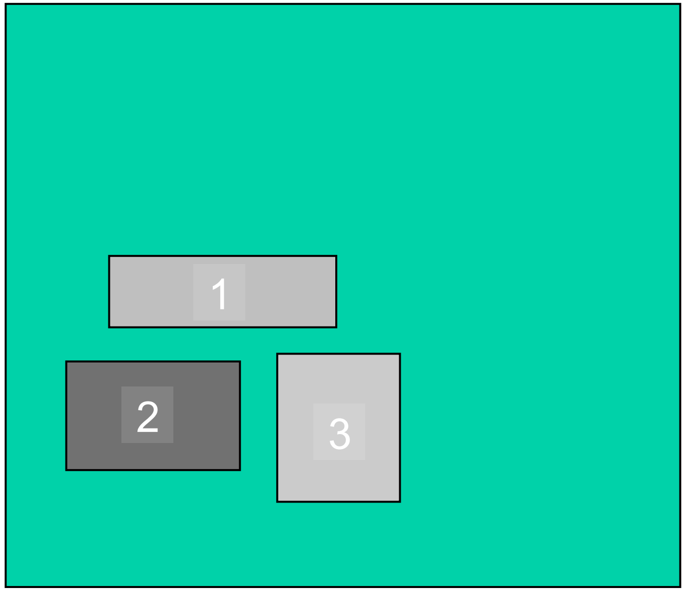
WIDTH

GRAY

AREA

- Salience: the property of one feature to distinguish the topic in the context:
  - the minimum distance between the topic's value for that feature and all the other objects' values for that feature

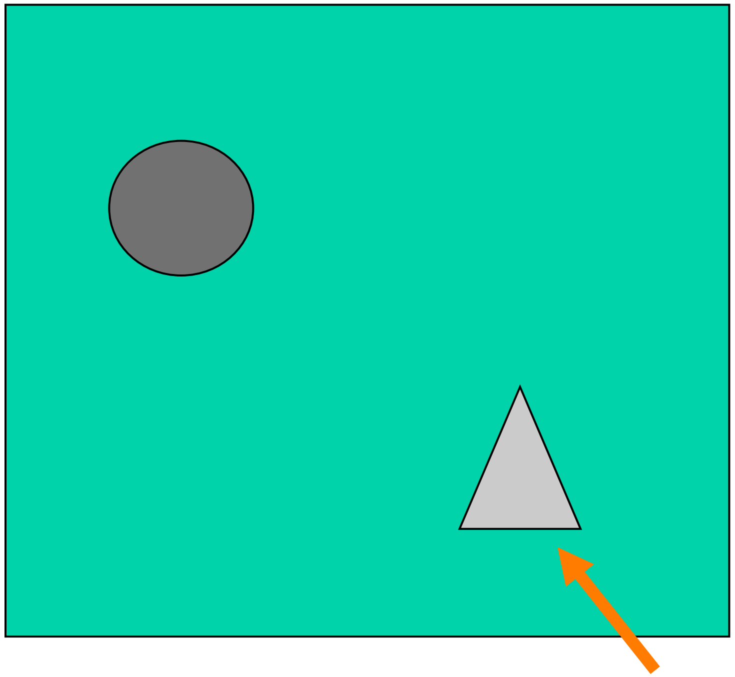
About perception again: why do we use some features and not others?



After scaling:

obj	HPOS	VPOS	HEIGHT	WIDTH	GRAY	AREA
1	0.25	0.45	0.30	0.66	0.45	0.70
2	**0.20**	**0.32**	**0.40**	**0.50**	**0.90**	**0.74**
3	0.42	0.31	0.50	0.30	0.42	0.76
sal	**0.05**	**0.01**	**0.10**	**0.16**	**0.45**	**0.02**

Lexicalisation – associating meanings to words



Game 125

A segments the scene in 2 objects

A categorizes the topic as HPOS[right]

A says: “mo”

B does not know “mo”

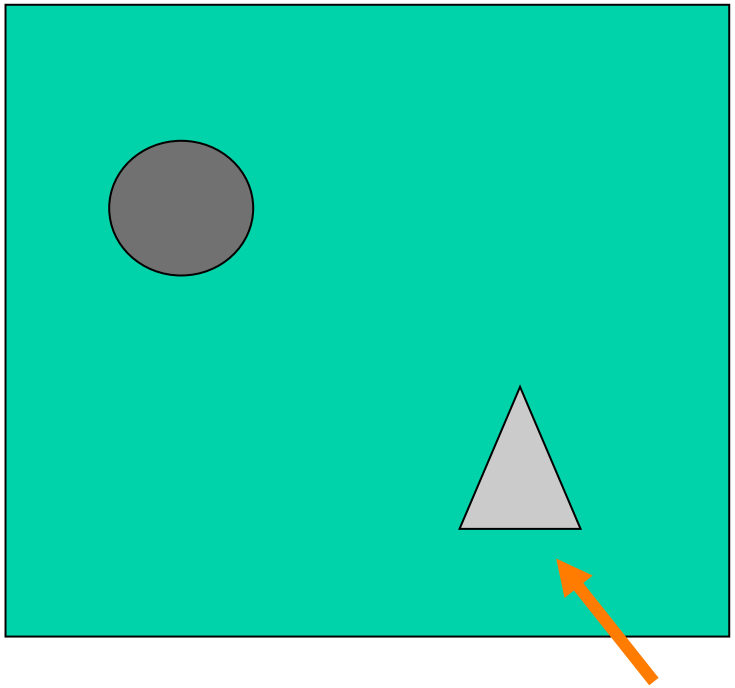
B says: “mo?”

A points to the topic

B categorizes the topic as HPOS[right]

B stores “mo” as HPOS[right]

Lexicalisation – interpretation



Game 205

A segments the scene in 2 objects

A categorizes the topic as HPOS[right]

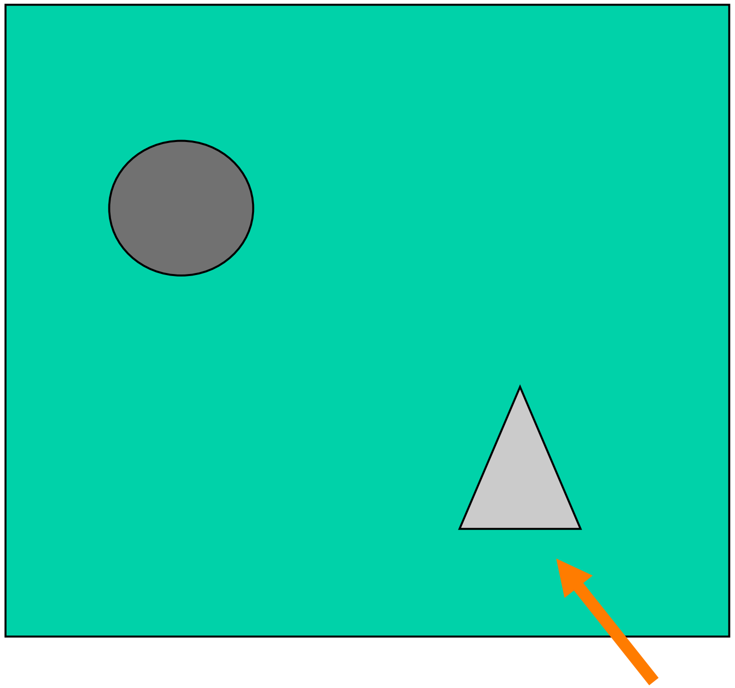
A says: “mo”

B interprets “mo” as HPOS[right]

B points to the topic

A says “OK”

Lexicalisation – synonymy



Game 245

A segments the scene in 2 objects

A categorizes the topic as HPOS[right]

A says: “mo”

B does not know “mo”

B says: “mo?”

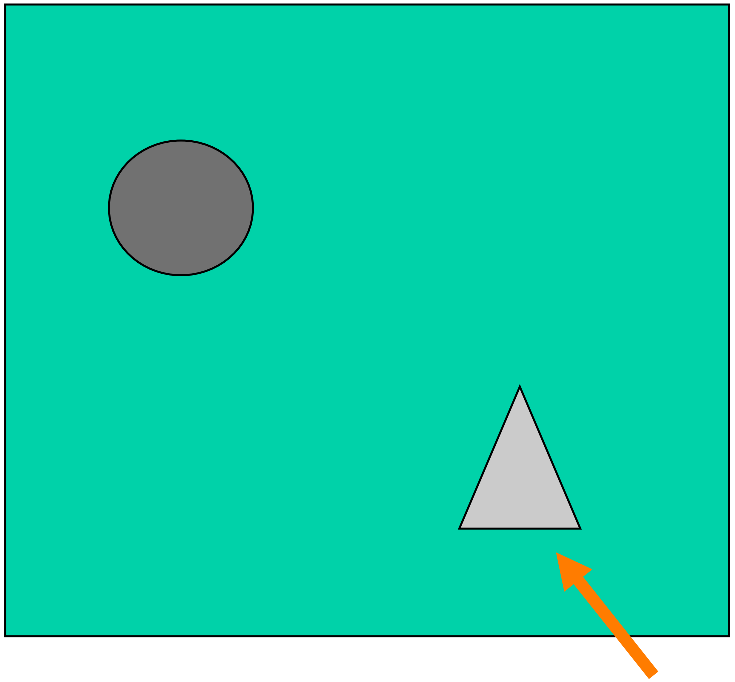
A points to the topic

B categorizes the topic as HPOS[right]

B has a word for HPOS[right]: “mogash”

B stores “mo” as a synonym for “mogash”

Differences in conceptualization produce "subtle" social polysemy



- A has a two-values conceptualization of HPOS
- B has a four-values conceptualization of HPOS

Game 280

A segments the scene in 2 objects

A categorizes the topic as HPOS[right]

A says: "mo"

B does not know "mo"

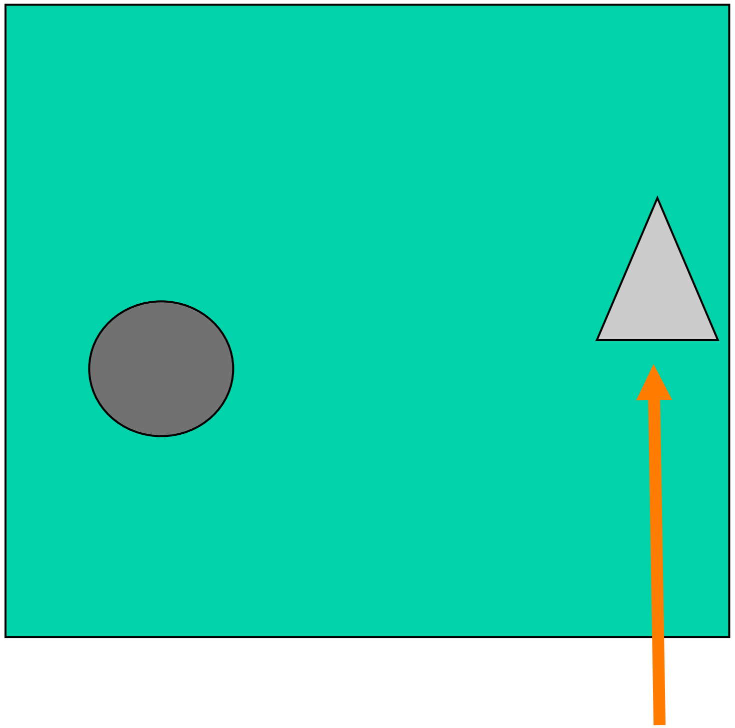
B says: "mo?"

A points to the topic

B categorizes the topic as HPOS[middle right]

B stores "mo" as HPOS[middle right]

Subtle social differences in meaning can give rise to generalisations



Game 302

A segments the scene in 2 objects

A categorizes the topic as HPOS[right]

A says: “mo”

B knows “mo” as HPOS[middle right]

B does not recognize an object in the scene having this value

B says: “mo?”

A points to the topic

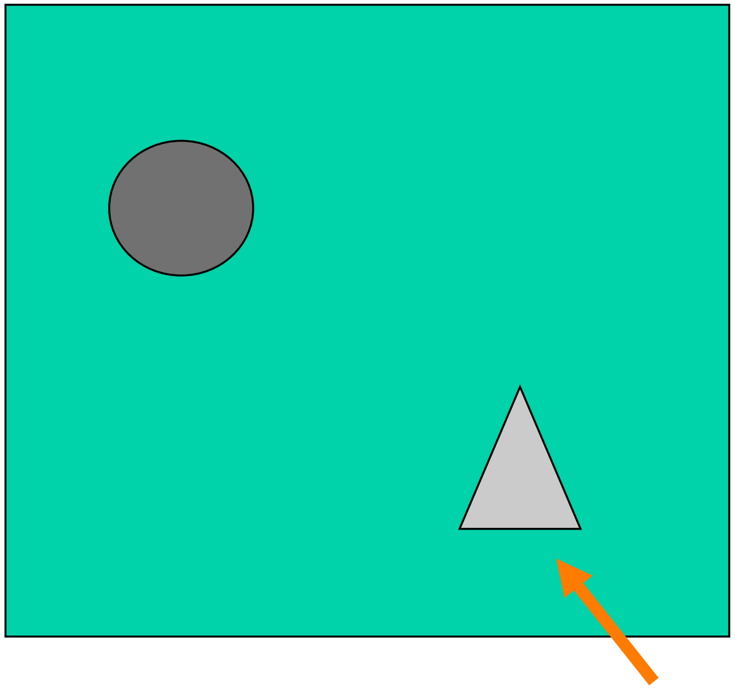
B categorizes the topic as HPOS[extreme right]

B stores “mo” as HPOS[extreme right]

Now B knows “mo” as both HPOS[middle right] and HPOS[extreme right]

By repetition he can infer a new category which subsumes both HPOS[middle right] and HPOS[extreme right], which should be HPOS[right], and this will be called “mo”

Ambiguity-1



Game 325

A segments the scene in 2 objects

A categorizes the topic as HPOS[right]

A says: “mo”

B does not know “mo”

B says: “mo?”

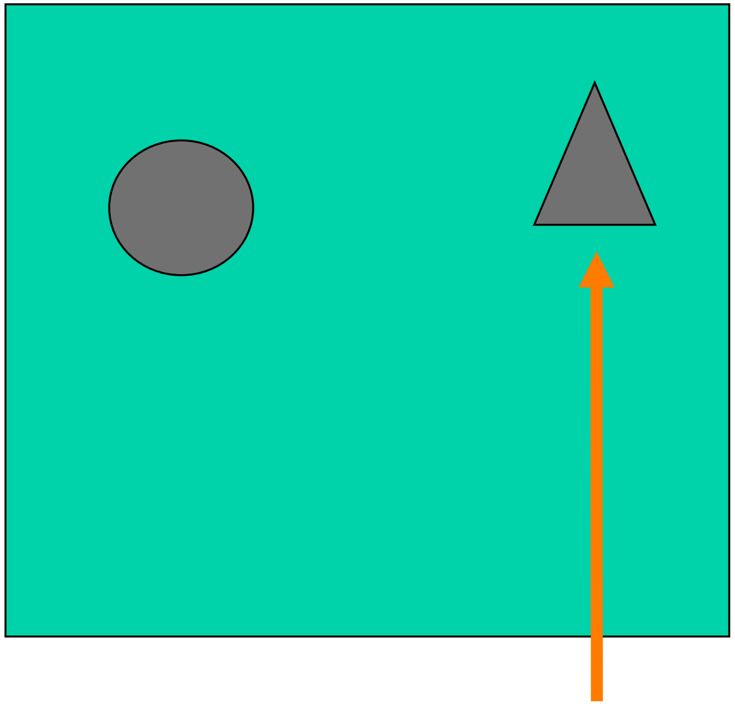
A points to the topic

B categorizes the topic as HPOS[right] and VPOS[low] and GRAY[light]

B stores “mo” as HPOS[right] OR VPOS[low] OR GRAY[light]

However, by positive feedback the lexicon will converge towards an efficient usage

Recovering from ambiguity



It is not known a priory whether "mo" will be stabilized by B as only HPOS[right] or only VPOS[low] or the union of the two.

An example:

Game 340

A segments the scene in 2 objects

A categorizes the topic as HPOS[right]

A says: "mo"

B knows "mo" as HPOS[right] OR VPOS[low] OR GRAY[light]

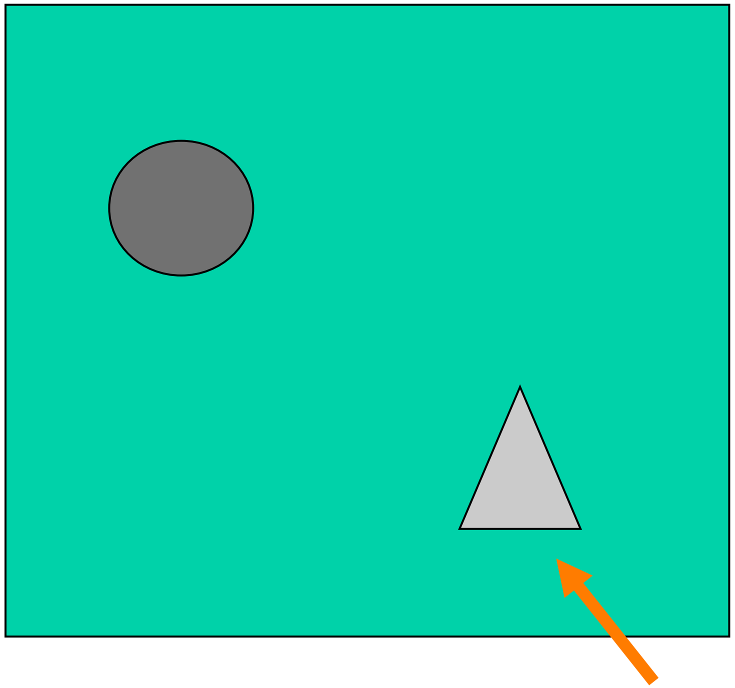
B recognizes an object in the scene with the value HPOS[right] and no object with the value VPOS[low] or GRAY[light]

B points to the topic

A says "OK"

B diminishes the meaning of "mo" as VPOS[low] and GRAY[light] and augments its meaning as HPOS[right]

Ambiguity-2



Suppose Game 325 takes place as following, instead:

Game 325'

A segments the scene in 2 objects

A categorizes the topic as HPOS[right]

A says: “mo”

B does not know “mo”

B says: “mo?”

A points to the topic

B categorizes the topic as VPOS[low]

B stores “mo” as VPOS[low]]

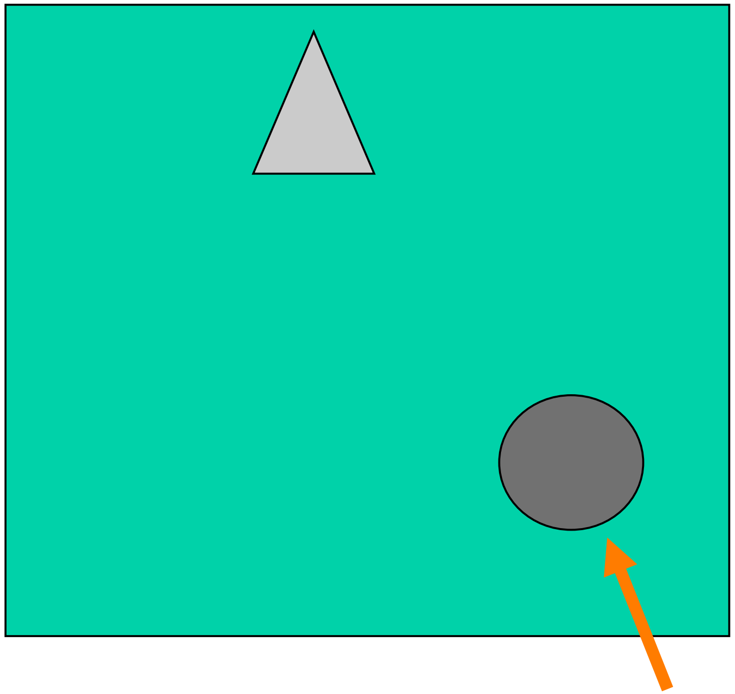
At this moment A and B understand different concepts by “mo”.

Ambiguity maintained

After Game 325' A knows "mo" as meaning HPOS[right] and B acquired it as VPOS[low]

Then we have this:

Game 390



A segments the scene in 2 objects

A categorizes the topic as HPOS[right]

A says: "mo"

B knows "mo" as VPOS[low]

B recognizes an object in the scene with the value VPOS[low]

B points to the topic

A says "OK"

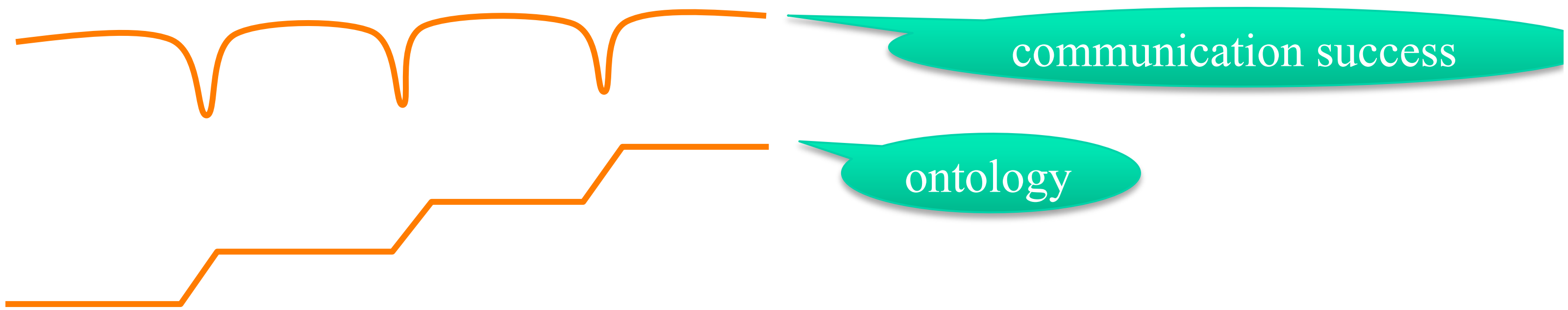
Now and again, the two agents do not realize that they give different meanings to "mo"

What is proved?

- A lexicon in a single agent
  - new words are invented or adopted
  - scores of association between forms (words) and meanings (concepts) go up and down
  - a “virgin”, “newly born” agent will catch up with a lexicon already existent in a population
- A common lexicon that stabilizes in a group of agents
  - however, differences could still coexist
  - the lexicon is sensible to the grows or reduction of population
  - the lexicon can absorb some shocks of contacts with other groups or can be destabilized

What the plots show?

- Ontology size as compared to communication success
  - Communication success dips each time the ontology is enlarged with new concepts, since new words have to be invented to deal with them.
  - However, the agents clearly manage to become again successful in guessing.



What the plots show?

- Increasing the population size
  - The agents create a word without knowing that one word already exist somewhere in the population (as it takes time to propagate) ➔ the risk of synonymy increases.
  - However, a steady progress towards an effective communication system is noticed.

What the plots show?

- What happens when two populations interact?
  - There is an initial destabilisation period when the coherence is low as the ambiguity increases.
  - However, the new community catches up and a new common lexicon emerges, abundant in synonyms.

How to prove the origins of language?

- In a simplified form:

  language = lexicon + grammar

  - The lexicon gives names for concepts and objects: the guessing games
  - The grammar expresses relations between concepts and objects: how to put it in terms of interactions between agents?

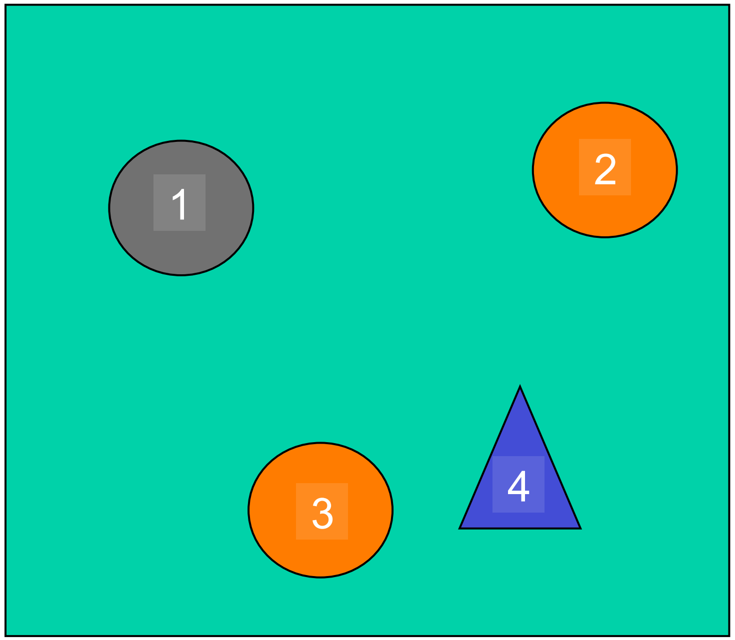
Guessing games implicit assumptions

- Experiments are made in a controlled setting that simplifies many aspects:
  - The world is simplified to the content of the table (scene)
  - The agents have the attention focused towards the scene
  - Their “aim” is to identify objects (they are motivated)
  - The agents dispose of a set of channels which are sufficient to put in evidence identifying properties of the objects populating the scene
  - The maximum vocabulary – sufficient to cover the concepts, as values produced by channels, is strictly limited
  - The words used by the agents apply to property values and not to objects
- This setting is assumed (given, programmed)

Establishing settings for exercising the birth of a grammar

- Setting 1:
  - The world: the content of the table (scene)
  - Attention: focused towards the scene
  - Motivation: identifying objects
  - Channels: identify properties of objects (enriched)
  - Maximum vocabulary: strictly limited
  - The agents already share a background vocabulary for naming properties of objects (as in phase 1), not directly objects
  - Only one property is not enough for disambiguation ➔ necessity to use combinations of words to express conjunctions
  - Implicit supposition: combinations express conjunction and not disjunction

Putting words together



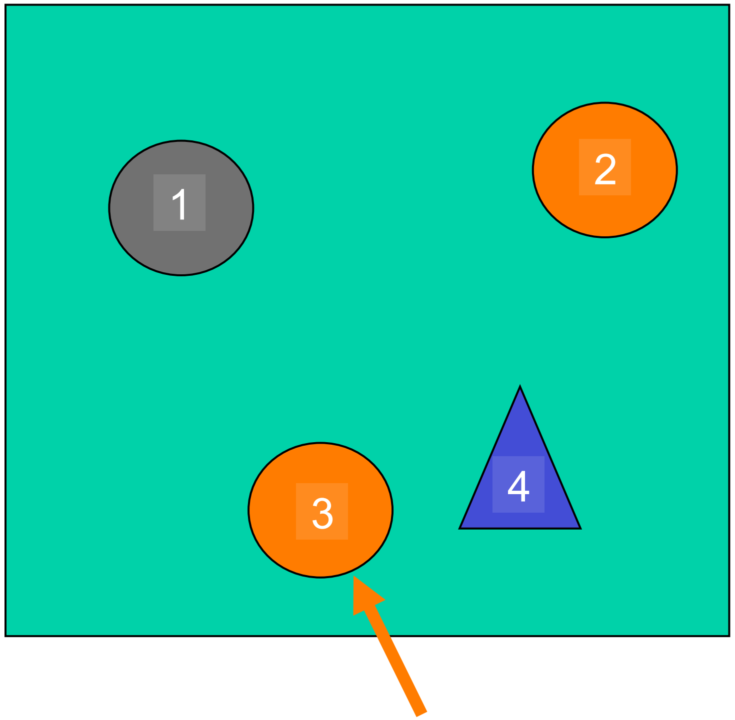
New channels:

COLOR: black, red, blue

SHAPE: circle, triangle

Important: the above symbols are our conventions to notate the mentioned property values – they do not belong to the acquired lexicon!

Putting words together



Game 1014

A segments the scene in 4 objects

A categorizes the topic as {VPOS[low], SHAPE[circle], COLOR[red]}

A has the lexicon:

- "bagadiru" for VPOS[low]
- "gugeawa" for SHAPE[circle]
- "camende" COLOR[red]

A correctly identifies on the decision tree that VPOS[low] AND SHAPE[circle] are sufficient to identify the topic

A says: "bagadiru gugeawa"

B has the lexicon:

- "bagadiru" for VPOS[low]
- "camende" for COLOR[red]

B says: "gugeawa?"

A points to the topic

B categorizes the topic as {VPOS[low], SHAPE[circle], COLOR[red]}

B correctly discovers on the decision tree that either SHAPE[circle] or COLOR[red], in combination with VPOS[low] are sufficient to identify the topic

B stores "gugeawa" as both SHAPE[circle], and COLOR[red], with a confidence = 0.5

Conclusions of a set of experiments of this kind

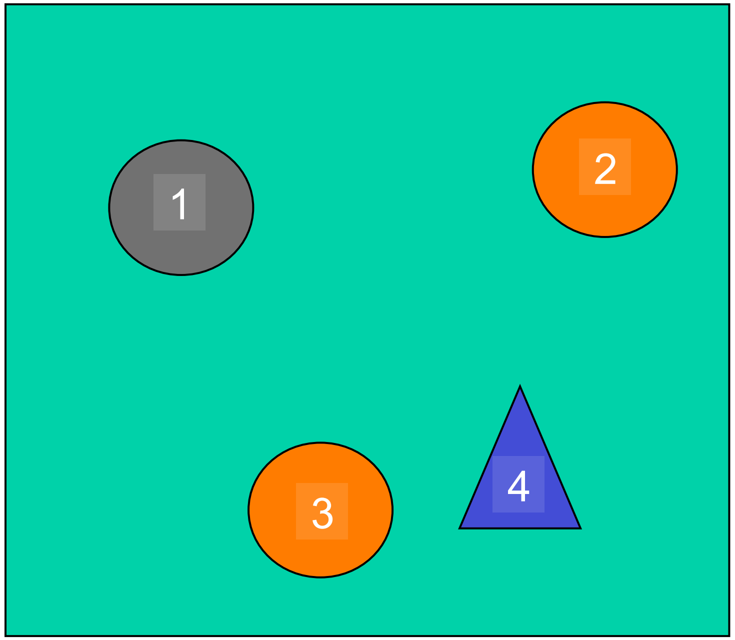
- Will not give rise to a grammar of expressing conjunctions: the implicit assumption was that putting words together restricts the selection (in conformity with most modern languages)
- The order of words is not important

  red circular or *circular red*

One step further in building a grammar

- Setting 2:
  - The world: the content of the table (scene)
  - Attention: focused towards the scene
  - Motivation for: identifying spatial relations among objects
  - Channels: identify properties of objects BUT ALSO spatial relations among objects
  - Maximum vocabulary: strictly limited
  - The agents have as background a common vocabulary for naming properties of objects (as in phase 1), not directly objects
  - Implicit supposition: in a linear expression Obj_1 R Obj_2, the focus is Obj_1

New channels expressing spatial relations



- HREL: left-of, right-of
- VREL: above, below

obj_1 left-of obj_2 (0.9)
obj_1 left-of obj_3 (0.4)
obj_1 left-of obj_4 (0.6)
obj_3 left-of obj_2 (0.1)
obj_3 left-of obj_4 (0.9)
…

But grammar is all about form

“HREL[right-of(obj_2)]” expresses the concept “whatever is to the right of obj_2“

One way to say that in this relation is obj_1: “$obj_{1\ \text{lexical-item-for}}$(right-of) obj_2”

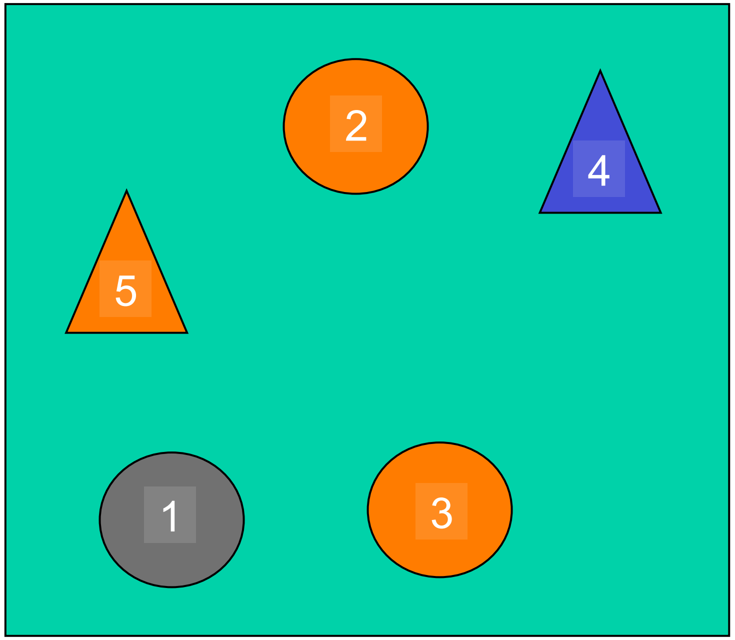
But, the agent knows a term for “right”: “mo”

Then, he might combine this term with a new word expressing the concept of “relation”, for instance “ga”: “mo-ga”

Expressing relations between objects

Initial lexicon:

- “mo” = HPOS[right]
- “bagadiru” = VPOS[low]
- “gugeawa” = SHAPE[circle]
- “zamira” = SHAPE[triangle]
- “camende” = COLOR[red]
- “gamaru” = COLOR[gray]



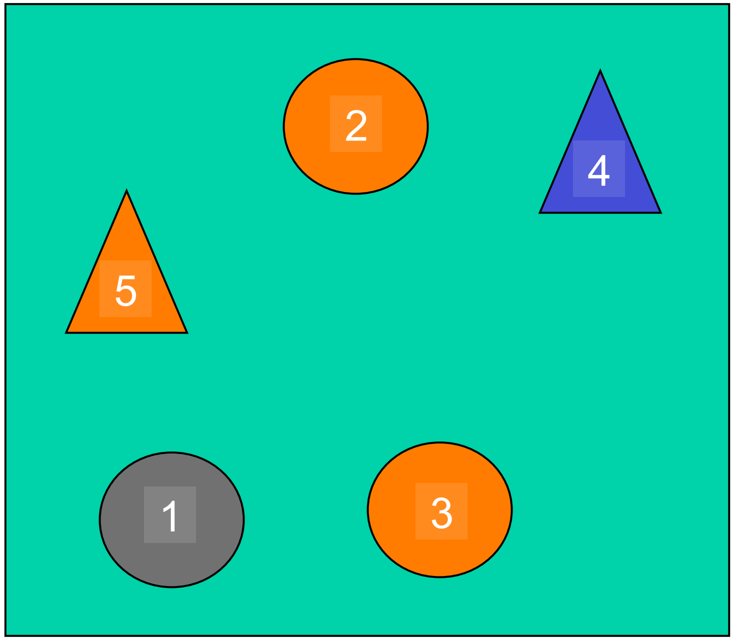
Derived expressions:

“zamira mo-ga gugeawa”

Expressing relations between objects

Initial lexicon:

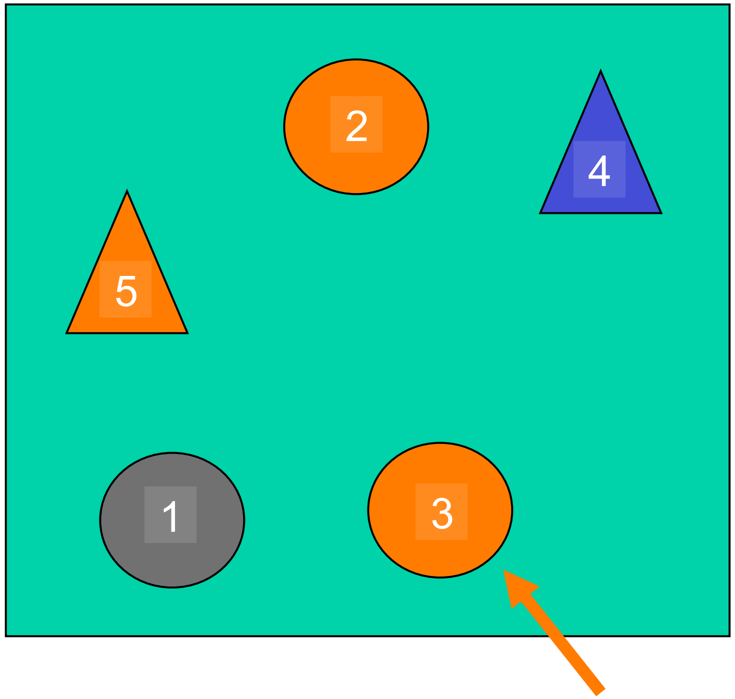
- "mo" = HPOS[right]
- "bagadiru" = VPOS[low]
- "gugeawa" = SHAPE[circle]
- "zamira" = SHAPE[triangle]
- "camende" = COLOR[red]
- "gamaru" = COLOR[gray]



Derived expressions:

- "camende gugeawa mo-ga bagadiru gamaru gugeawa"

Guessing relations



Game 2020

A segments the scene in 4 objects

A categorizes the topic as {VPOS[low], HPOS[right], SHAPE[circle], COLOR[red],HREL[right-of(SHAPE[circle])], HREL[right-of(COLOR[gray])], …}

Both A and B have the lexicon:

- "mo" = HPOS[right]
- "bagadiru" = VPOS[low]
- "gugeawa" = SHAPE[circle]
- "zamira" = SHAPE[triangle]
- "camende" = COLOR[red]
- "gamaru" = COLOR[gray]

A correctly identifies on the decision tree that HREL[right-of(obj_1)] is sufficient to identify the topic

A identifies obj_3 as {SHAPE[circle], VPOS[low], COLOR[red]}

A says: "camende gugeawa mo-ga bagadiru gamaru gugeawa"

B says: "mo-ga?"

A points to the topic

B identifies "camende gugeawa" as either obj_2 or obj_3 but, based on A's pointing, eliminates obj_2

B identifies "bagadiru gamaru gugeawa" as obj_1

B stores "mo-ga" as HREL[right-of()]

How to diminish the amount of initial assumptions?

- Remember our *Setting 2*:
  - …
  - **Implicit supposition: in a linear expression $\mathbf{Obj_1}$ R $\mathbf{Obj_2}$, the focus is $\mathbf{Obj_1}$**

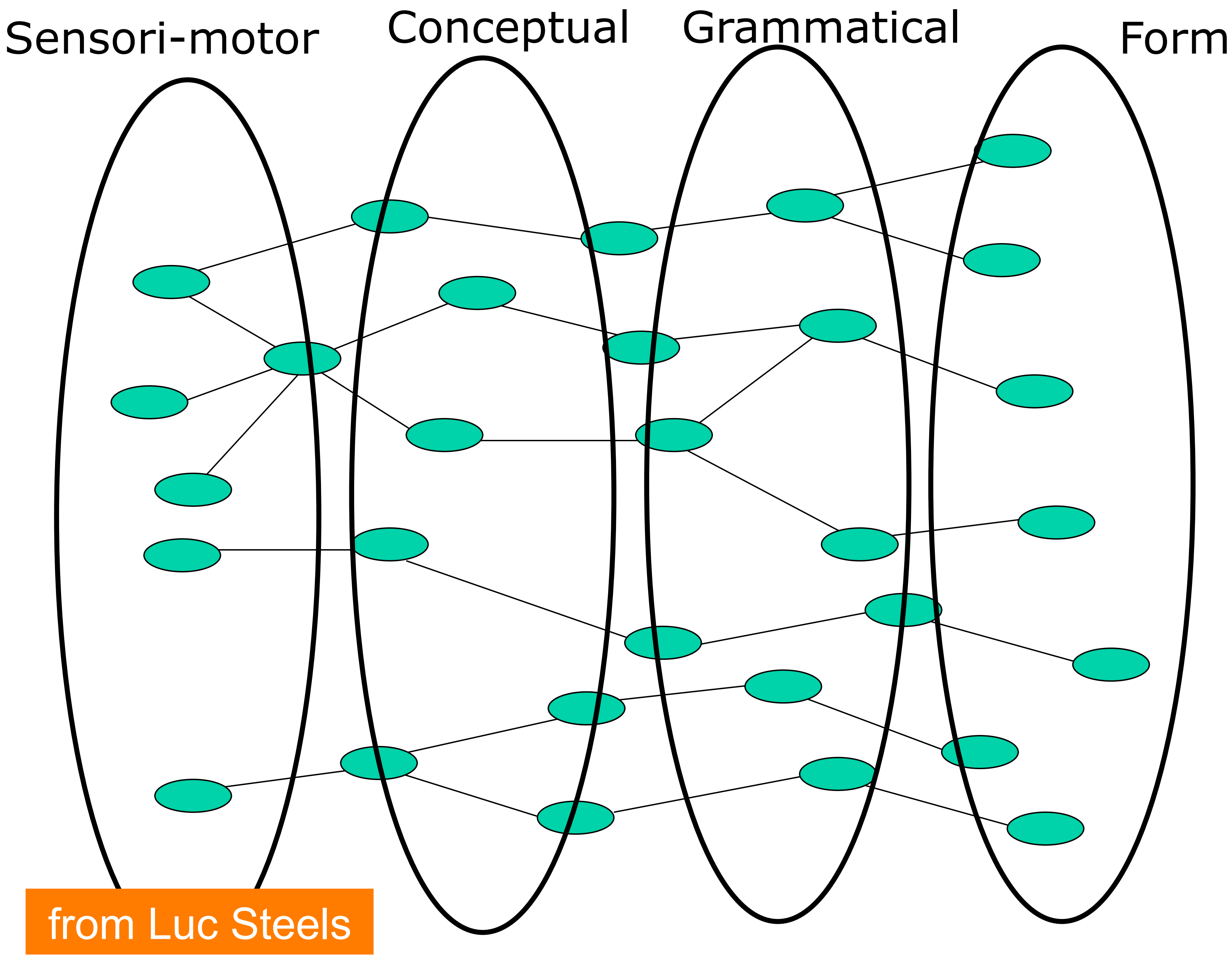
- This is a direct and artificial immixture in the very heart of the birth process of the language!
- Solution: parameterize all grammatical features of the language and let them evolve naturally

Fluid Construction Grammar

- Structures to represent the information needed in language processing about a specific sentence (feature structures)
- Structures to represent the lexical and grammatical constructions (rules)
- Operations of Unify and Merge
- Structures specifying how new rules are built (templates)

a chemical metaphor

from Luc Steels


Sensori-motor
Conceptual
Grammatical
Form
from Luc Steels

Constructions

- Semantic and syntactic categories do not operate in isolation
- They are part of frames (schemas, patterns)
- Constructions are mappings from semantic to syntactic schema
- Constructions can also add additional meaning to meaning of the parts and add additional form

Fillmore, Kay, Michaelis, Croft, Goldberg, ...

from Luc Steels

Conclusion

- The language is considered from an evolutionistic point of view
- Lexicon formation: proved experimentally
- Grammar and dialogue: in the study in ALEAR
  - Fluid Construction Grammar
    - integrates syntax and semantics (Fillmore's roles)
    - unification and merge mechanisms

Going further

- Discourse phenomena
  - Modelling cohesion: anaphora
  - Modelling coherence: why do we make discourses the way we do?

And further…

- Can we build a humanoid robot able to acquire the ability to talk the way we do?
- What type of minimal cognitive mechanism should we place in an artificial agent for letting “him” evolve towards the human performance in discourse?
- Is it that “human type of discourse” could be achieved spontaneously in agents, during a long series of interactions with similar agents?