Lund University
Cognitive Science

Peter Gärdenfors

Conceptual spaces, cognitive semantics and robotics

Early AI

Two main flavours:

(1) KR Symbolic representations

- … came up against the frame problem etc
- Formal ontologies

(2) ML Machine learning

- Learning is biased, data greedy, opaque, and brittle

Current AI

- Focus on categorization
- Deep learning ≈ back propagation 2.0
- Contents of concepts captured opaquely in neural networks
- Generates data for nouns and adjectives Verbs are generally missing
- Limited models of reasoning
- Here I propose *event structure* as an appropriate semantic foundation for AI and robotic systems
- Modelled in conceptual spaces

Three levels of modelling in KR and ML

Symbolic models

Based on a given set of predicates with known denotation.

Representations based on logical and syntactic operations.

Problem: Where do the predicates come from?

Conceptual spaces

Based on a set of quality dimensions.

Representations based on topological and geometrical notions.

Problems: Where do the dimensions come from?

Connectionist models

Based on a (uninterpreted) inputs from receptors.

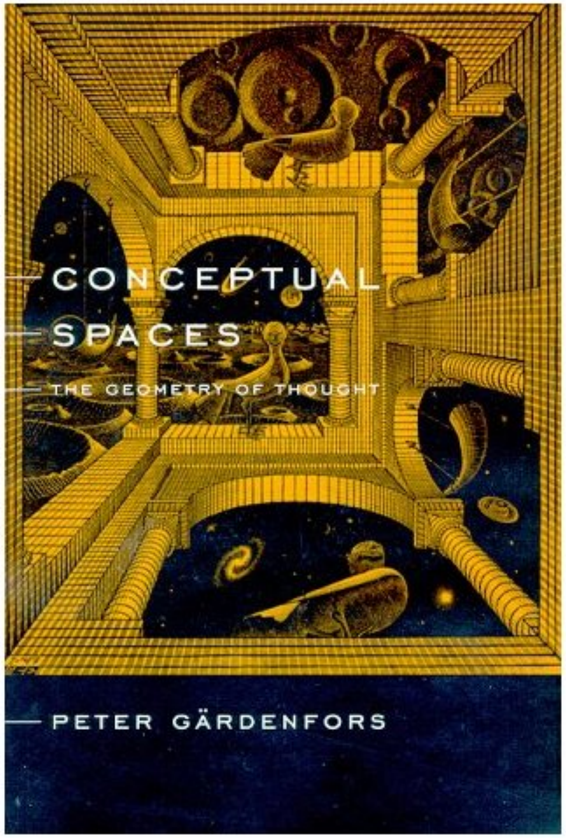
Distributed representations by dynamic connection weights.

Problems: What is represented in the network?

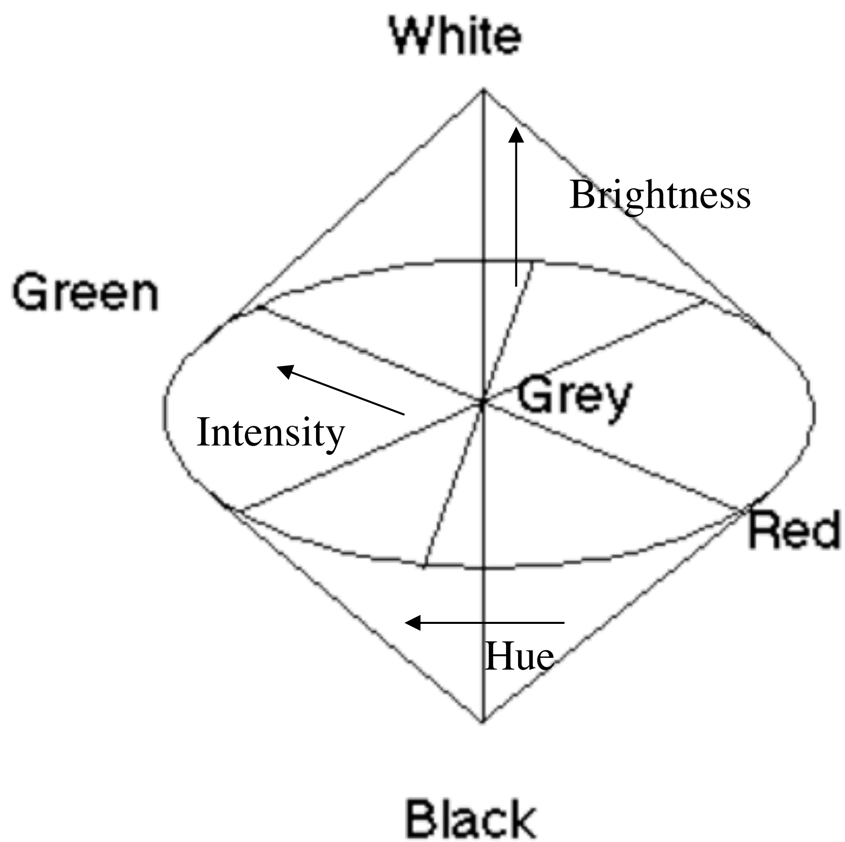
Learning is in general very slow.

Conceptual spaces

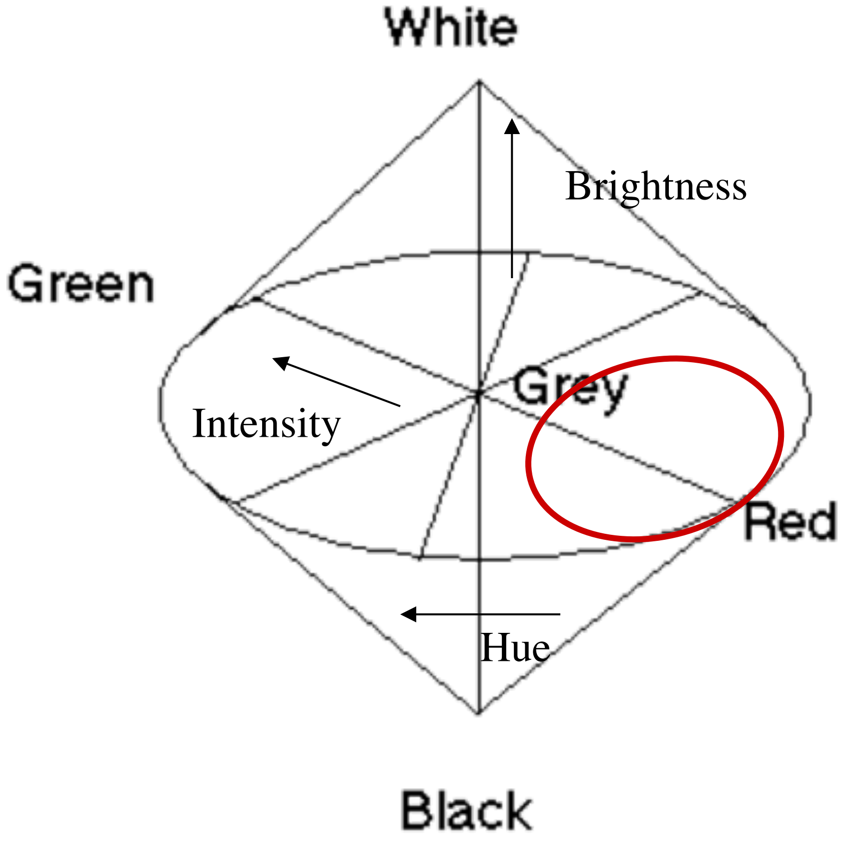
- Consists of a number of dimensions (colour, size, shape, weight, location, force …)
- Dimensions have topological or geometric structures
- Dimensions are sorted into *domains*
- Concepts are represented as *convex* regions of conceptual spaces



The color spindle

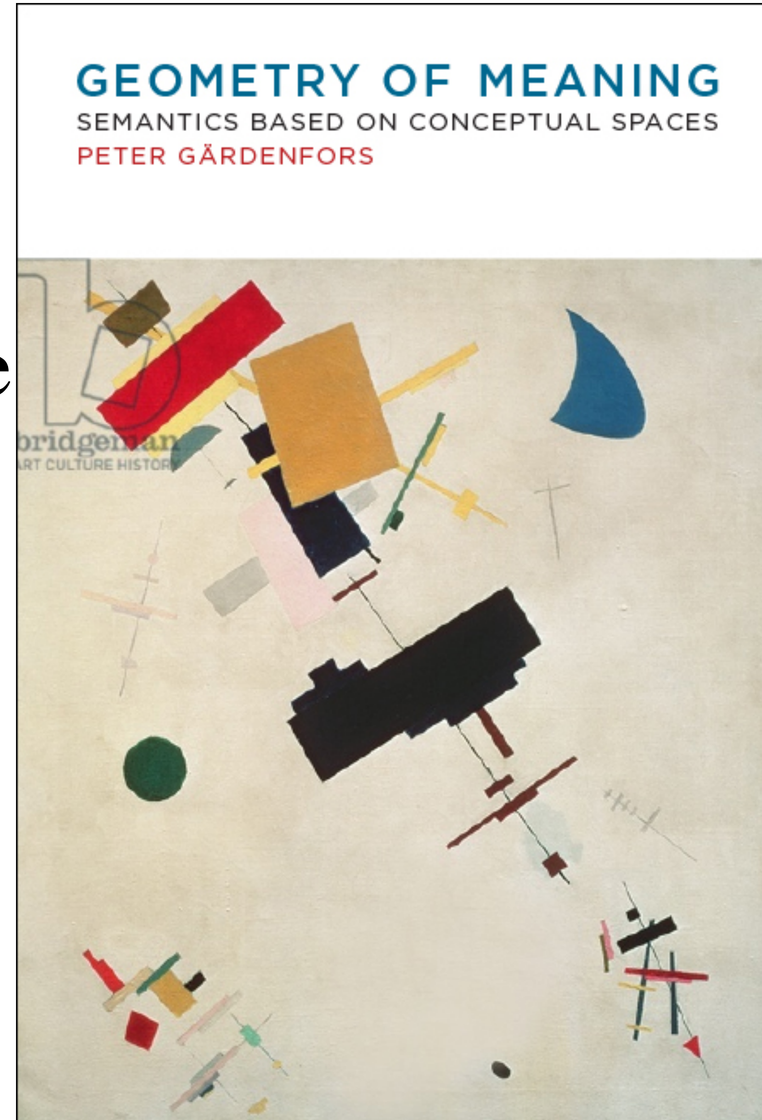


The color spindle



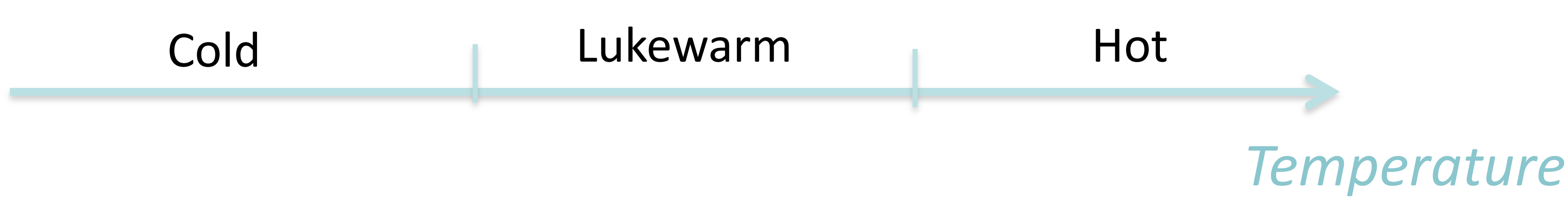
Why are there word classes?

- Word classes are defined syntactically in linguistics
- Each use of a word is supposed to belong to a single class
- ”Round” can be noun, verb, adjective, adverb, preposition
- My aim: A cognitively grounded theory of the *semantics* of word classes



Properties and adjectives

- *Thesis:* (single domain constraint for adjectives) The meaning of an adjective can be described as a convex region in a single domain
- Examples: *heavy* (weight), *hot* (temperature), *bitter* (taste), *round* (shape)
- No adjective means ”green or orange”
- No adjective means ”long and hot”



An example of an object category: ”Apple”

Domain	*Region*
Color	**Red-green-yellow**
Taste	**Values for sweetness, sourness etc**
Shape	**"Round" region of shape space**
Nutrition	**Values for sugar, vitamin C, fibres etc**

Properties vs. Object categories

Properties: A convex region in a single domain

Object categories: A number of convex regions in different domains; together with (1) prominence values of the domains and (2) information about how the regions in different domains are correlated

Subclasses of nouns characterised by domains

- *Places* only require the space domain
- *Mass nouns* do not contain the shape domain
- *Concrete* nouns require object category domains
- *Abstract* nouns do not contain the space domain
- *Agents* require the force domain
- *Intentional agents* require the goal domain

Humans are excellent at identifying actions

Representational hypothesis for actions

- The fundamental cognitive representation of an action is the *pattern of forces* that generates it
- Actions are more or less *similar* and show prototype effects
- An *action concept* is a convex region in the space of force patterns

Representing verb meanings

- *Thesis:* (single domain constraint for verbs) The meaning of a verb root is a *convex* region of vectors that depends only on *a single domain*.
- Examples: *push* refers to the force vector of an event (and thus the force domain), *move* refers to changes in the spatial domain of the result vector and *heat* refers to changes in temperature.
- There are no verbs that mean ‘walk and burn’ (multiple domains) and there are no verbs that mean ‘crawl or run’ (not convex)

Predictions from the theory

- Explains *similarities* of verb meanings
- Explains *subcategories* of verbs
- *March*, *stride*, *strut*, *saunter*, *tread* etc are subregions of *walk*
- Explains metaphorical uses of verbs: ”move the meeting to Friday”, ”vacuuming with his mouth”, ”slice a tennis ball”, ”eat up a profit”
- Predicts the division into manner and result verbs
- Manner verbs describe force vector (cause) and result verbs the result vector (effect)

The geometry of prepositions

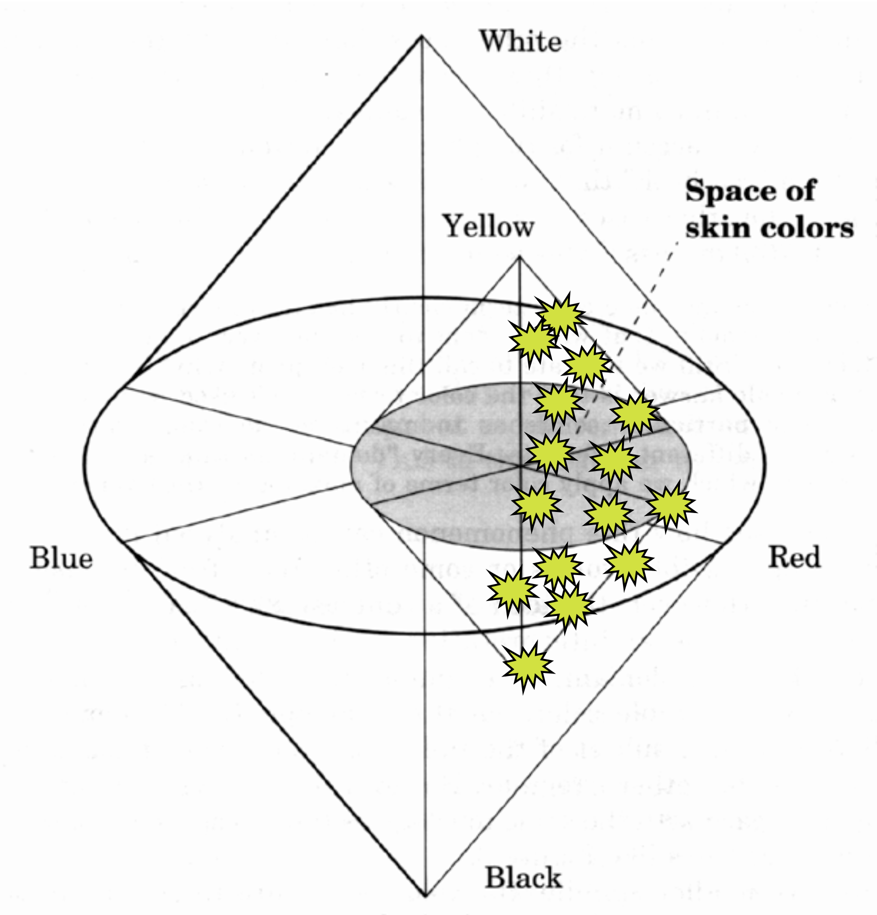
- *Thesis*: (single domain constraint for prepositions) Prepositions denote *convex* regions in a *single domain*
- Spatial domain (*above*), force domain (*on*), time domain (*during*), epistemic domain (*despite*)
- Complicating factor: Prepositions have many metaphorical uses

Geometric *structures* of word classes

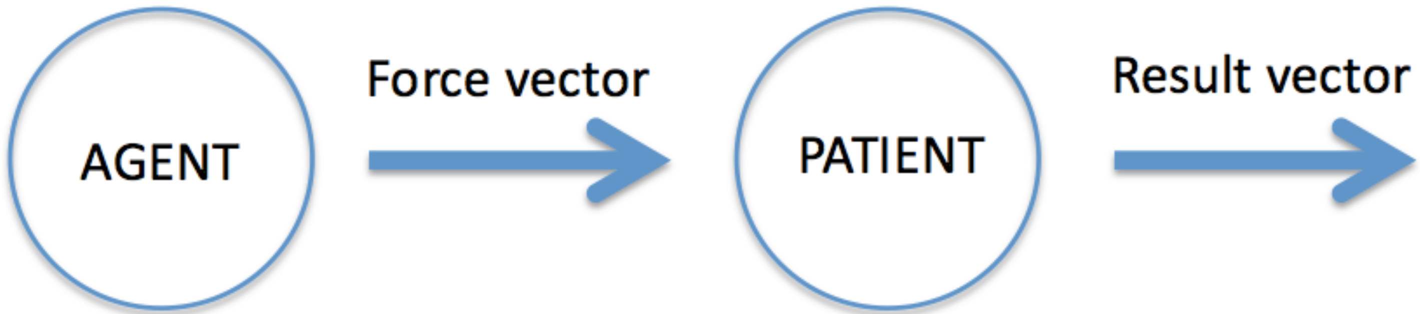
• *Adjectives*	Regions of domains
• *Nouns*	Regions of products of domains
• *Verbs*	Vector regions of domains
• *Prepositions*	Regions and vectors with landmarks
• *Adverbs*	Multipliers of vectors
• *Demonstratives*	Regions of perceptual domains
• *Quantifiers*	Properties of sets

Combinations of word meanings

- Context effects
- Contrast classes: red book, red wine, red skin, red hair, …
- Why do we speak in sentences?
- General answer: Sentences are about *events*

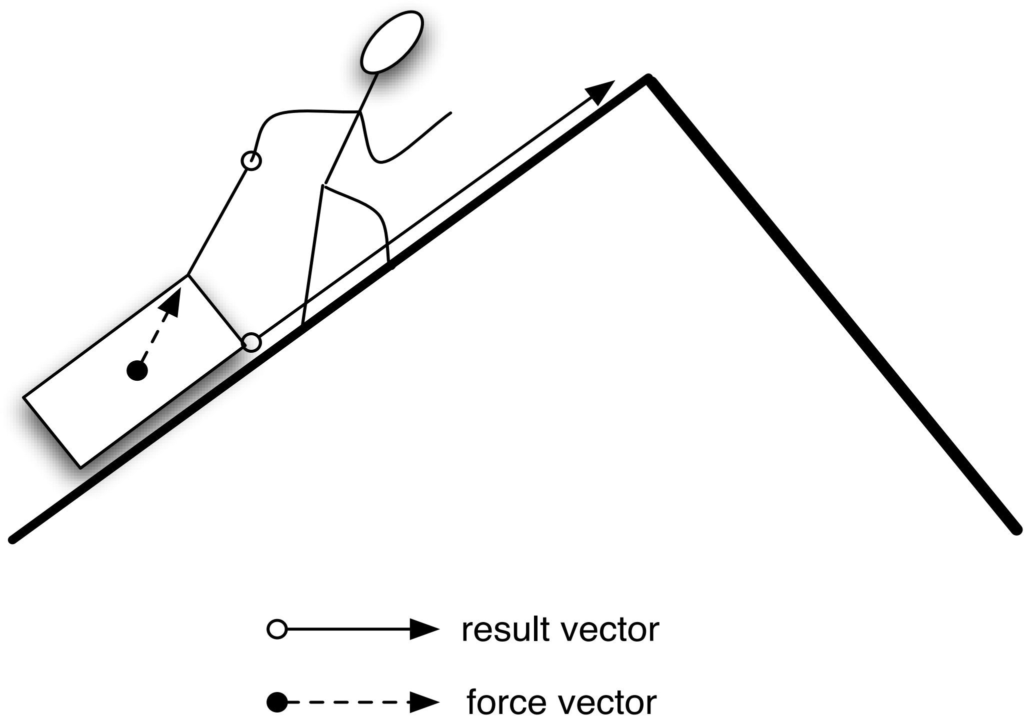


A two-vector model of an event



- The *force vector* (pattern) acts on an patient
- From force space (generated by an action)
- The *result vector* describes the changes of the properties of the patient
- Changes in location or in category space
- Agent is not always necessary (fall, grow …)
- Sometimes agent = patient

Example: pulling a sled



Event semantics

- Entities have *roles* in events:
- Agent, patient, action, result, recipient, instrument, …
- Event representations generate new expectations: actions about results, agents about patients ...
- Expectations partly controls attention

Causal relations determined by the two vectors

- Force – cause
- Result – effect
- Unlike most philosophical theories cause and effect are of different kinds
- Judea Pearl: ”I postulate that the major impediment to achieving accelerated learning speeds as well as human level performance should be overcome by … equipping learning machines with causal reasoning tools.”

Sentences refer to (aspects of) events

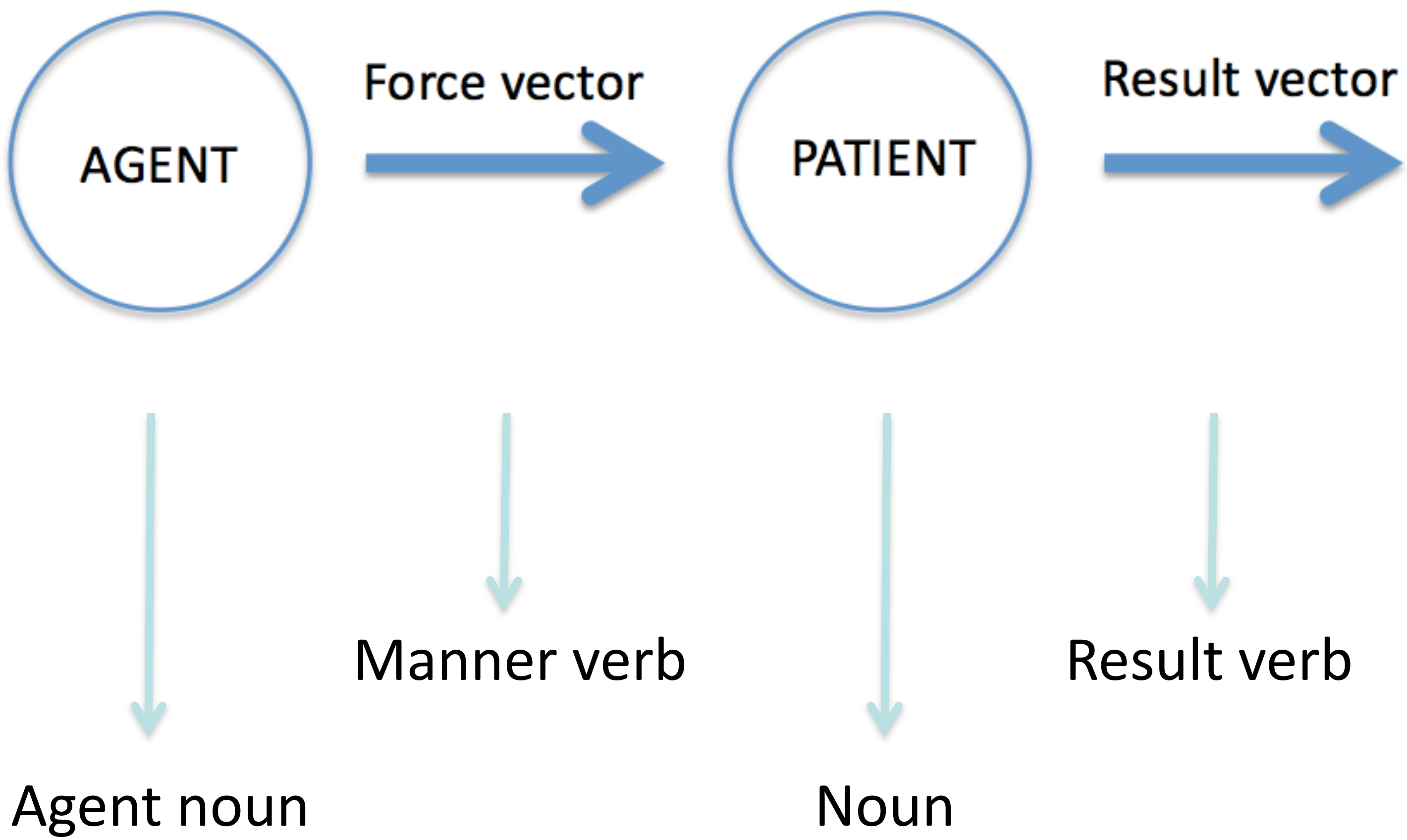
Thesis: A sentence expresses a part of an event involving at least a force or a result vector (and one entity)

Patients and agents ➔NPs / Vectors ➔VPs

In analogy with the visual process, a sentence *focuses* on some parts of an event

“Victoria hits Oscar” vs “Oscar was hit by Victoria”

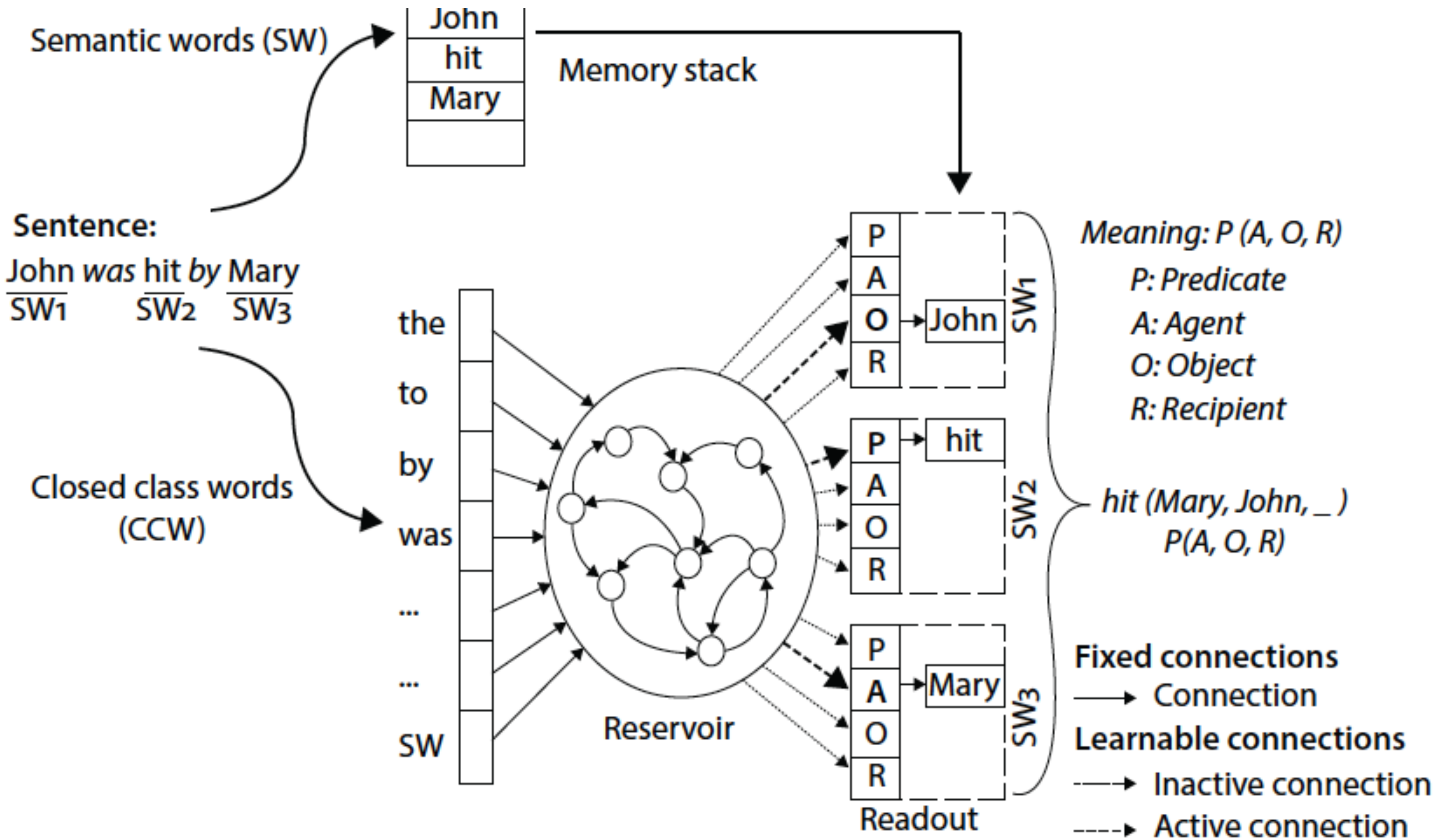
From events to word classes



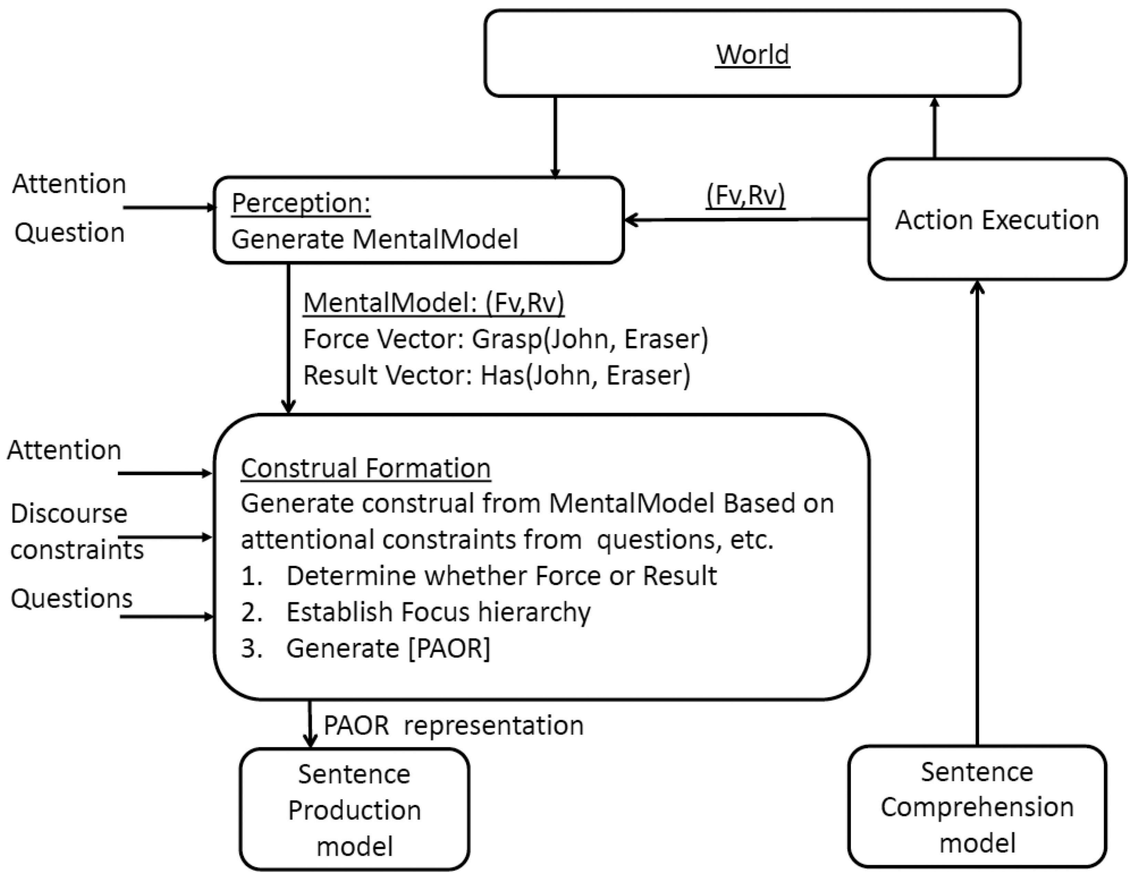
Many possible construals of an event

- **A**gent **F**orce **P**atient **R**esult
- AF ”She scrubs” (intransitive)
- AR ”She cleans”
- FP ”The table was scrubbed” (passive)
- PR ”The table was cleaned” (passive)
- AFP ”She scrubs the table” (transitive)
- APR ”She cleans the table” (transitive)
- AFPR ”She scrubs the table clean”

Dominey's reservoir computing

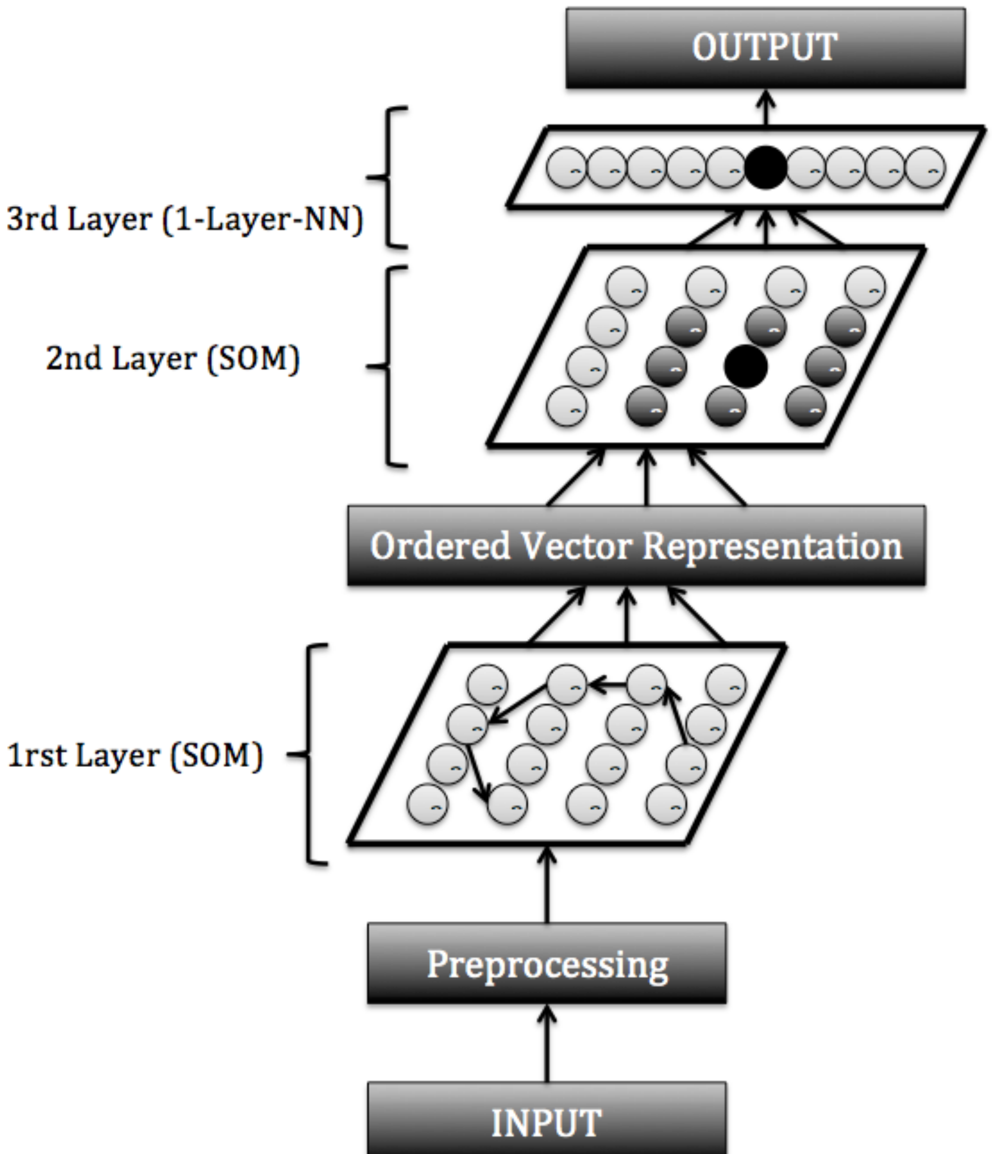


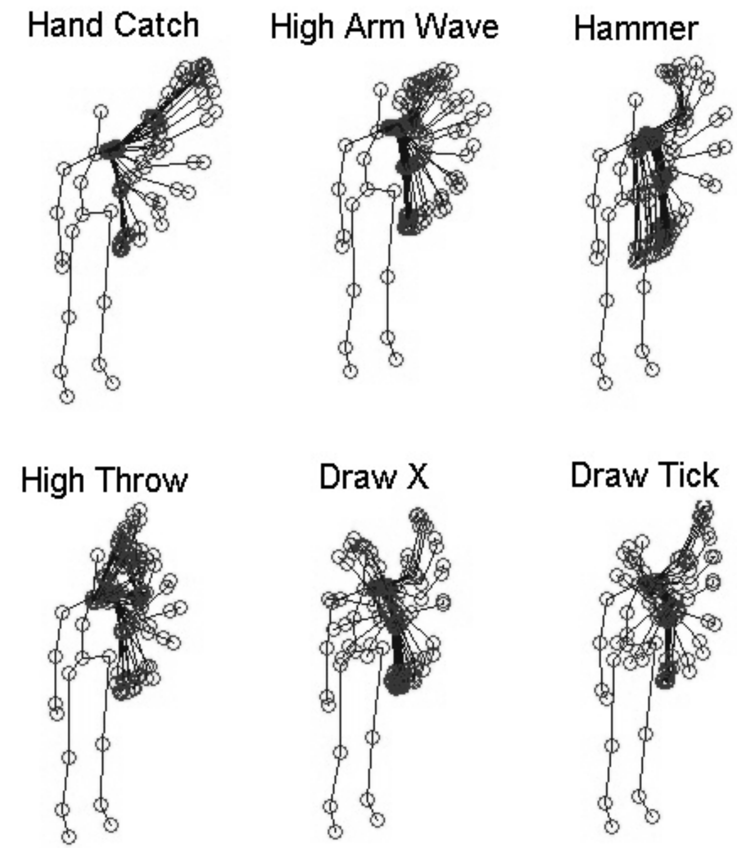
Extension to manner and result verbs



Sentence comprehension

LUND action recognition system





Real time action identification

References

- P. Gärdenfors (2019): “Using event representations to generate robot semantics”, *ACM Transactions on Human-Robot Interaction*, Volume 8(4), Article No. 21, 1–21.
- A.-L. Mealier, G. Pointeau, P. Gärdenfors and P. F. Dominey (2016): “Construals of meaning: The role of attention in robotic language production”, *Interaction Studies 17(1)*, 48-76.
- Z. Gharaee, P. Gärdenfors and M. Johnsson (2017): “Online recognition of actions involving objects”, *Biologically Inspired Cognitive Architectures*, DOI 10.1016/j.bica.2017.09.007