

Conceptual Spaces

A geometric framework for representing concepts

Lucas Bechberger

- If a computer has the following entry in its knowledge base:

$$\forall x: \text{apple}(x) \Rightarrow \text{red}(x)$$

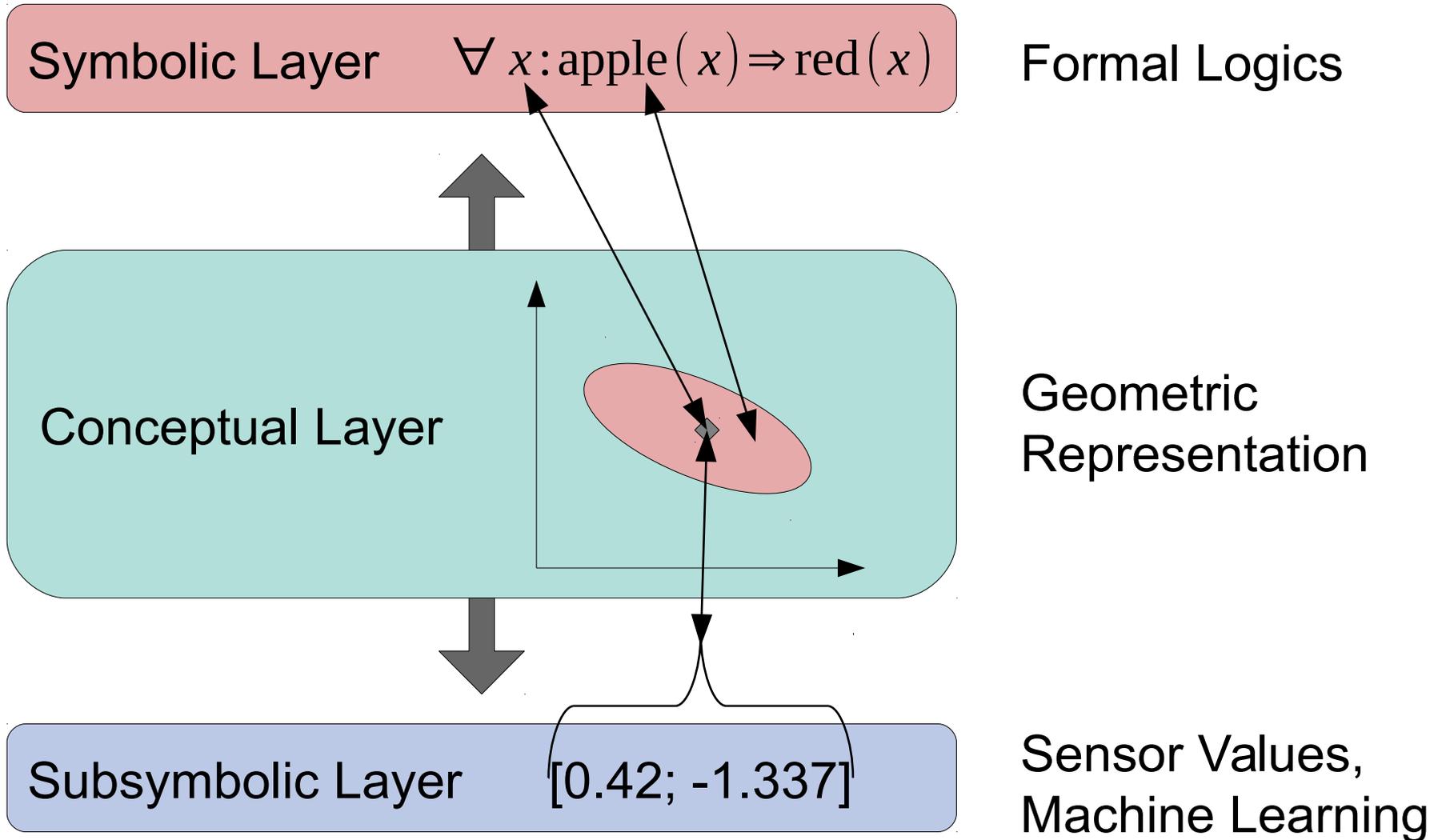
- Does it know what „red“ means?
- Does it know what „apple“ means?

- Most likely not!

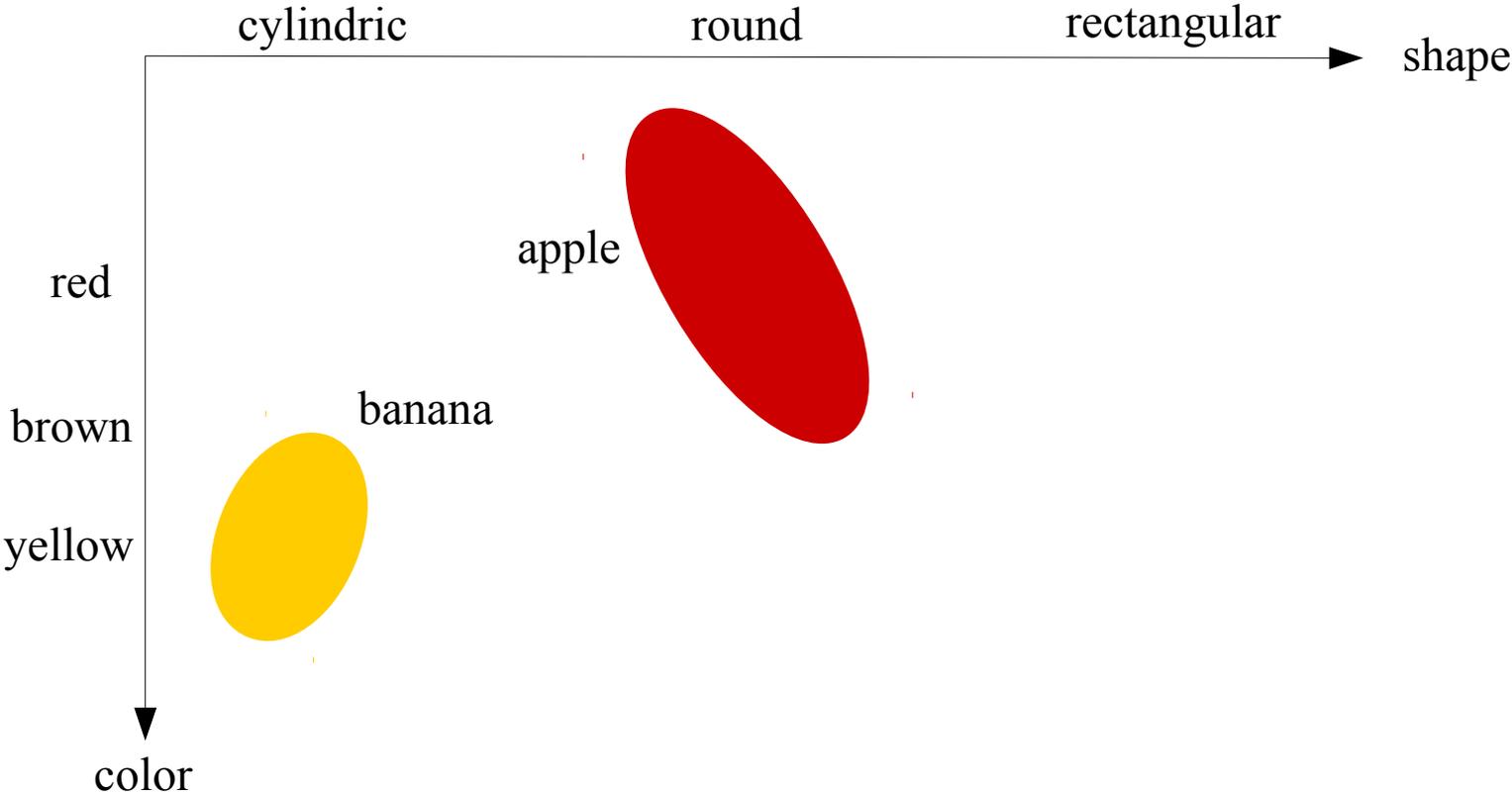
- For a computer „red“ is just an arbitrary symbol

$$\forall x: \text{klj8}(x) \Rightarrow \text{42x8e45}(x)$$

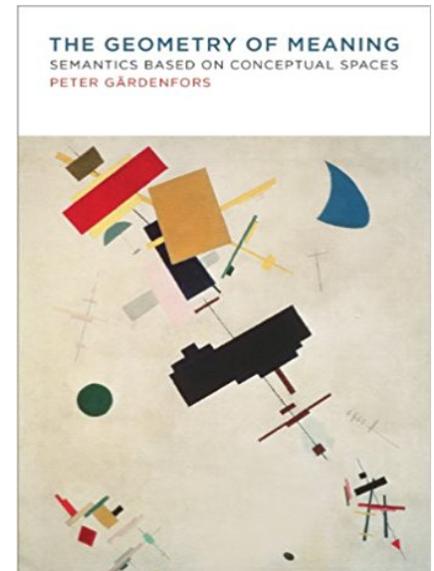
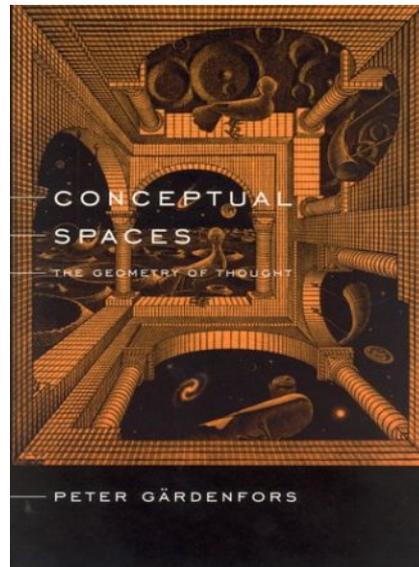
- „Symbol grounding problem“ [Harnad1990]:
 - How can abstract symbols contain any meaning?
 - They need to be grounded in reality



- High-dimensional space [Gärdenfors2000]
 - Dimensions represent measurable qualities
 - Concepts are convex regions in this space

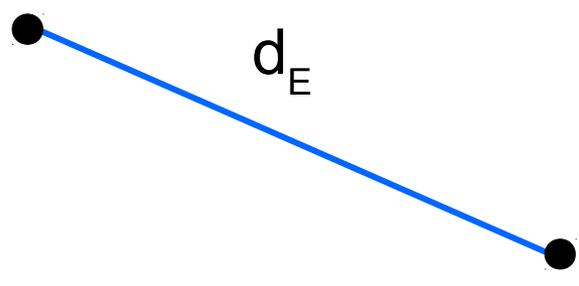


Conceptual Spaces

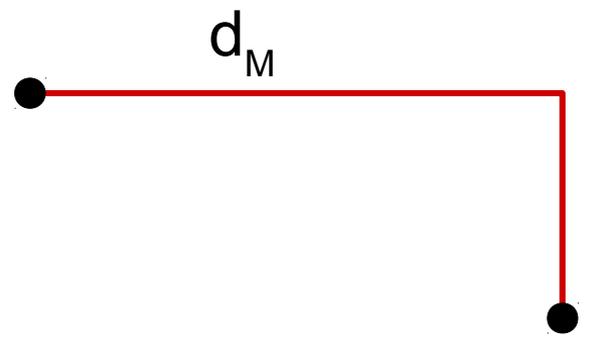


- Quality dimensions
 - Different ways stimuli are judged to be similar or different
 - E.g., temperature, weight, brightness, pitch
- Domain
 - Set of integral dimensions that are separable from all other dimensions
 - Color: hue, saturation, and brightness
- Distance in this space is inversely related to similarity
 - Integral dimensions → Euclidean distance (length of line segment)
 - Separable dimensions → Manhattan distance (sum of distances)

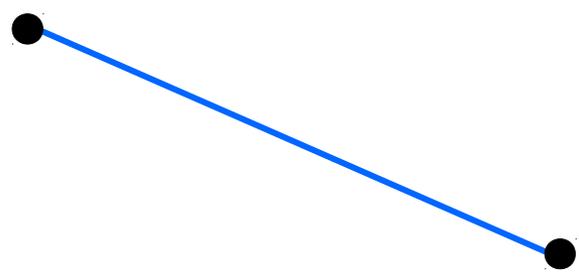
Euclidean Distance



Manhattan Distance



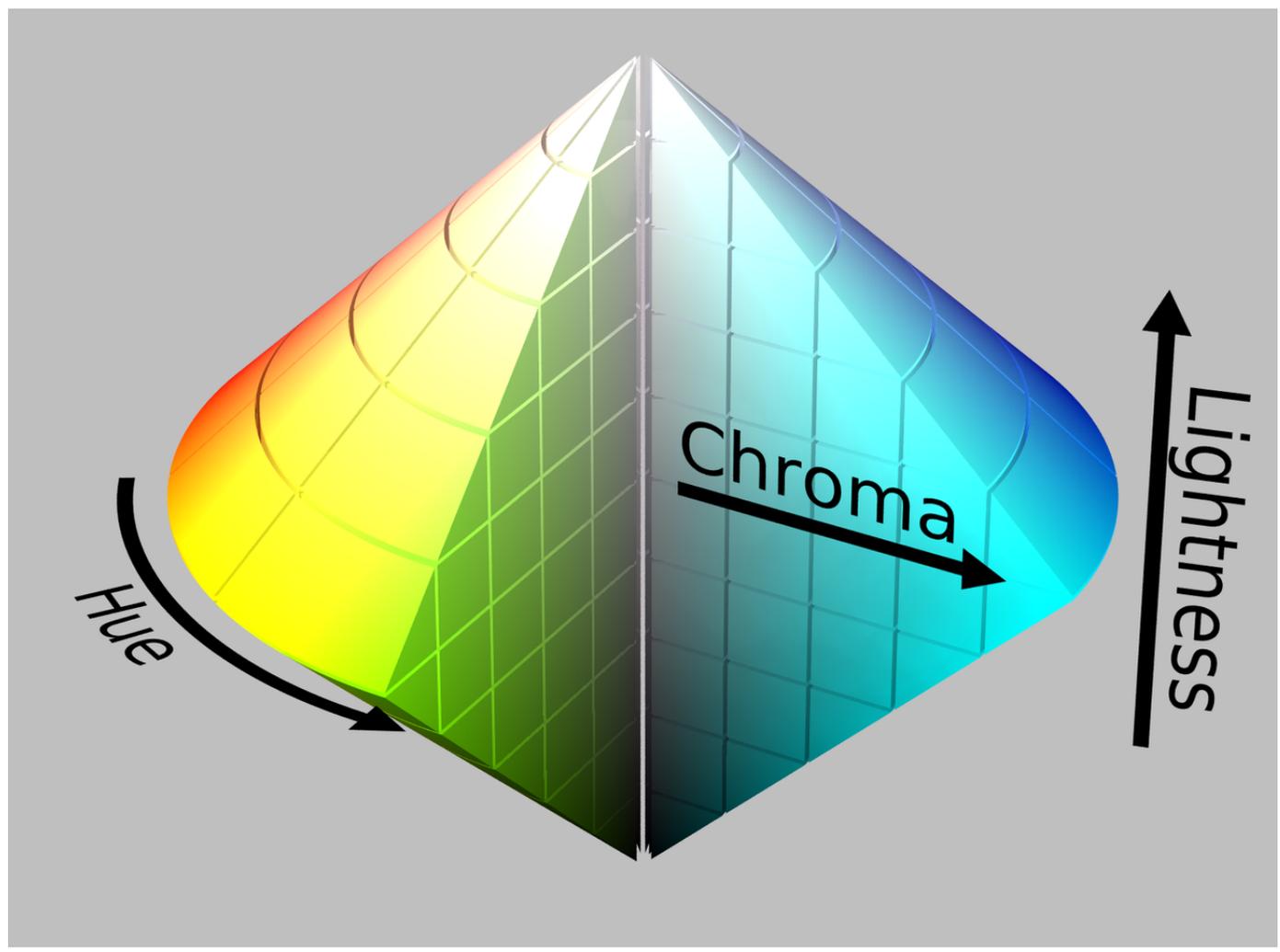
Euclidean Betweenness



Manhattan Betweenness



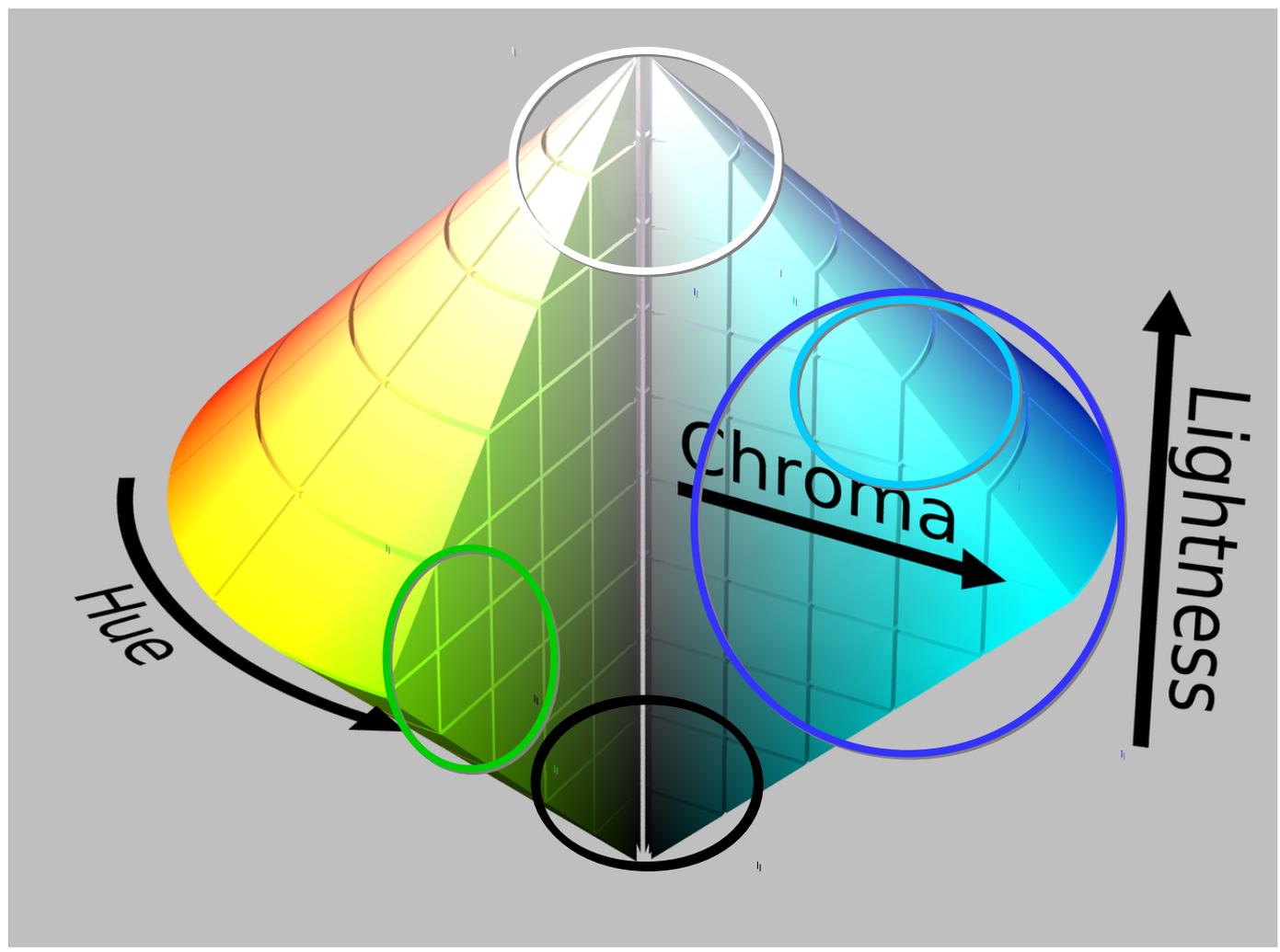
Example: the color domain



https://en.wikipedia.org/wiki/HSL_and_HSV#/media/File:HSL_color_solid_dblcone_chroma_gray.png

- Look at one domain (e.g., color)
 - Regions correspond to properties (e.g., red)
 - Properties are mostly expressed by adjectives
- Criterion P:
 - A natural property is a convex region of a domain in a conceptual space.

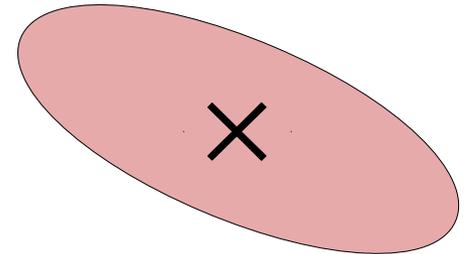
Example: the color domain



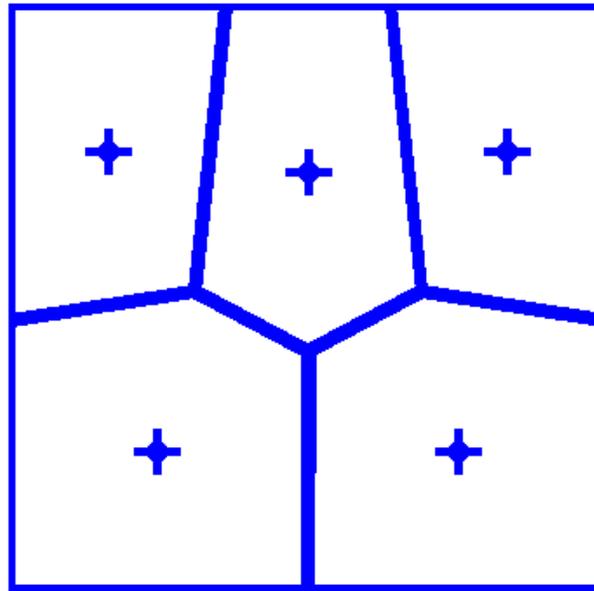
https://en.wikipedia.org/wiki/HSL_and_HSV#/media/File:HSL_color_solid_dblcone_chroma_gray.png

- Prototype theory of concepts
 - Each concept is mentally represented by a prototype
 - Prototype = abstract summary representation
 - E.g., average instance
 - Categorization: compare stimulus to all prototypes
 - → best match wins

- Conceptual spaces
 - Each concept is represented by a convex region
 - Central point of this region can be interpreted as prototype



- Voronoi tessellation
 - Given a set of central points $\{p_1, \dots, p_n\}$
 - Assign each point in the space to its closest p_i



<https://commons.wikimedia.org/wiki/File:CentroidalVoronoiTessellation2.png>

- Set of prototype points generates convex sets

- Example: „apple“
 - Color: red
 - Shape: spherical
 - Texture: smooth
 - Taste: sweet
- Defined across multiple domains: combination of properties
 - Different „importance“ to the concept (influenced by context)
 - Potentially correlated
- Criterion C:
 - A natural concept is represented as a set of convex regions in a number of domains together with an assignment of salience weights to the domains and information about how the regions in different domains are correlated.

- Start with a set of dimensions
- Group them into domains
- Properties = convex regions within a single domain
- Concepts span multiple domains
 - Salience values and correlation information

Applications of Conceptual Spaces

The Framework in Action

- Adjectives refer to properties
 - “red”, “tall”, “round”, “sweet” all refer to single domains
- Nouns refer to concepts
 - “apple”, “dog”, “tree” are based on a combination of domains
- Verbs refer to actions
 - “push”, “walk”, “bend” refer to the force domain
- Prepositions refer to the spatial domain
 - “above”, “into”, “across” refer to positions and paths with respect to a landmark

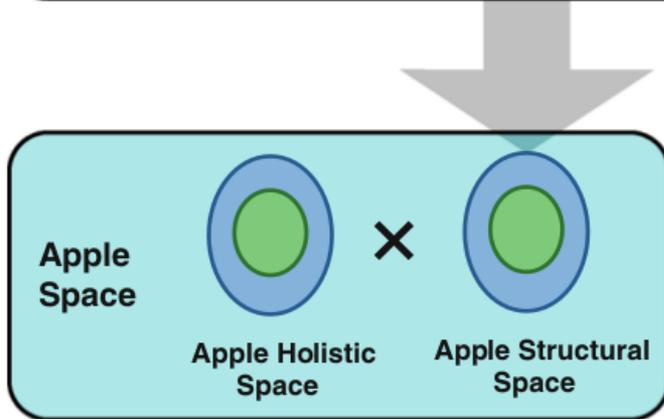
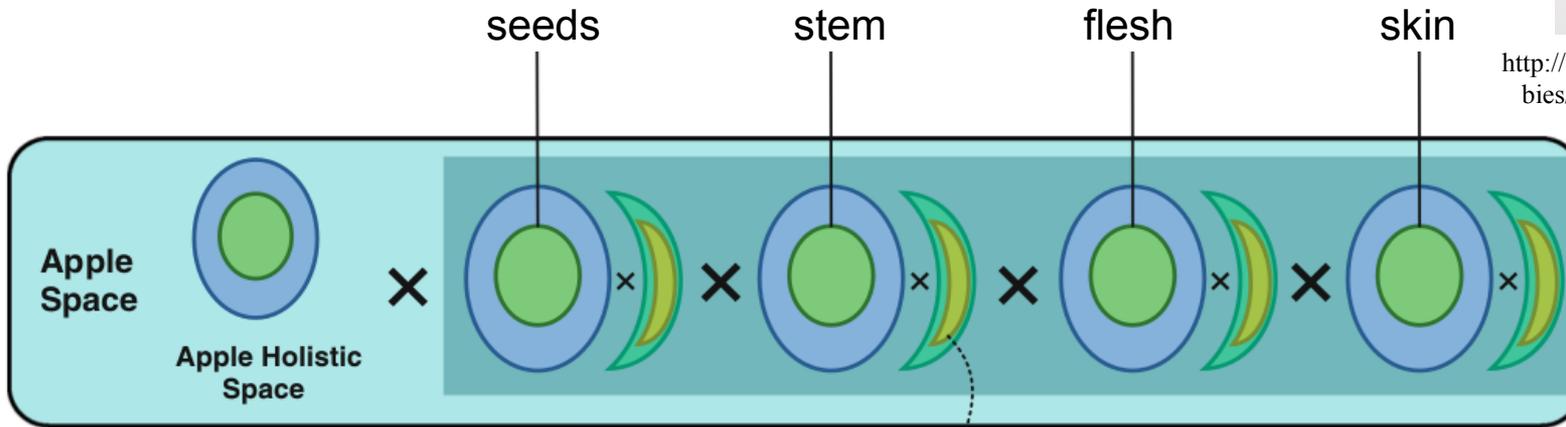
[Gärdenfors2014]

- “green banana”
 - **green** is compatible with with **banana**’s color information
 - Narrow down the color region
 - Correlations between domains yield further updates:
 - Consistency is **solid**
 - Taste is **bitter**
 - “pink elephant”
 - **pink** is incompatible with **elephant**’s color information
 - Replace the color information
 - “stone lion”
 - **stone** is incompatible with most domains of **lion** (e.g. life span, habitat)
 - Compatible domains: shape, size, and color
 - Remove incompatible domains, keep compatible ones
- [Gärdenfors2000]

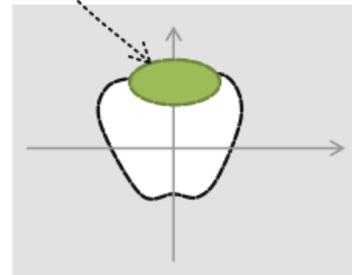
- Typically, objects consist of multiple parts
 - Apple: seeds, stem, flesh, skin
- Idea: represent parts & their relations



<http://www.blogoftheworld.com/freebies/high-resolution-fruits-stock-photos/>



Apple Stem Structure Domain

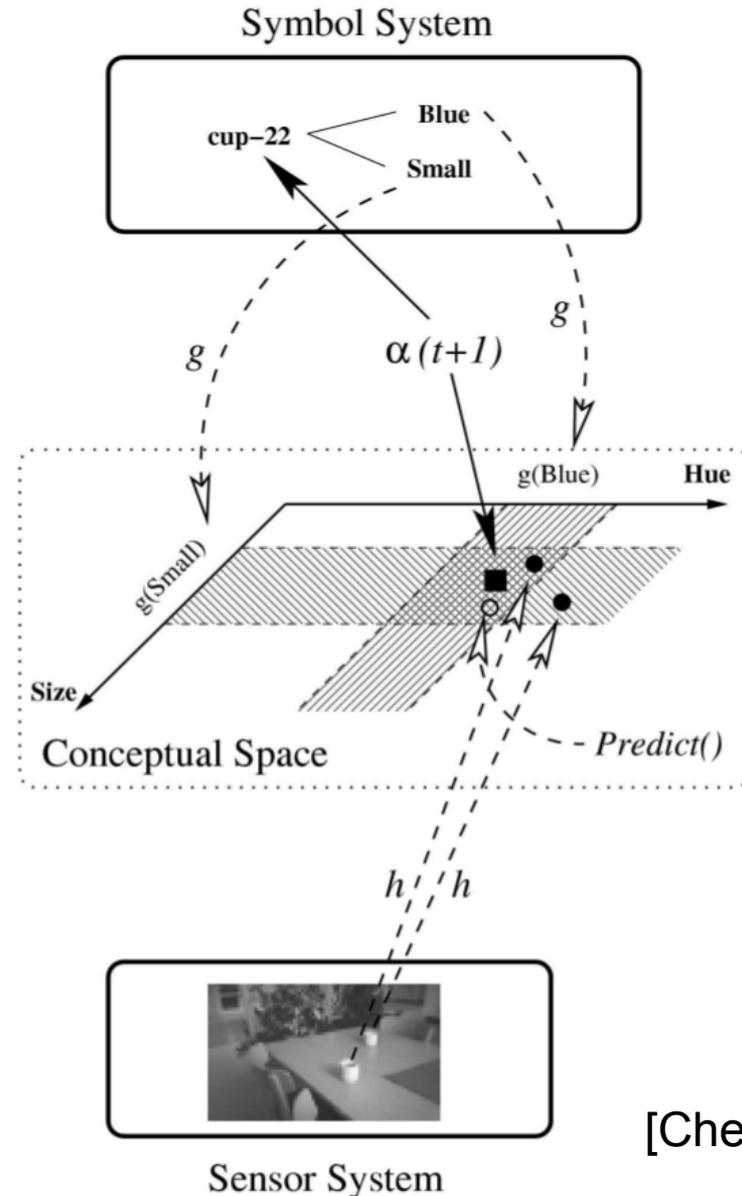


[Fiorini2013]

- Symbol system
 - Symbols (“cup-22”)
 - Predicates (“blue”)
 - g : predicates \rightarrow areas

- Sensor system
 - Takes measurements at each time step
 - h : measurements \rightarrow points

- Anchor
 - α : time \rightarrow symbols x points
 - Ties symbols to observations
 - Needs to be updated



[Chella2004]

- Interpolative reasoning
 - Bachelor students are exempt from paying council tax in the UK
 - PhD students are exempt from paying council tax in the UK
 - What about Master students?
- Similarity and betweenness are important 
- Derive conceptual space from textual data (word vectors)
- Find interpretable dimensions in this space
- Some example results:
 - “wine shop” is between “gourmet shop” and “liquor store”
 - Difference between “Jurassic Park” and “Kill Bill: Vol. 1”:
 - “dinosaurs”, “the expedition”, “the scientist”

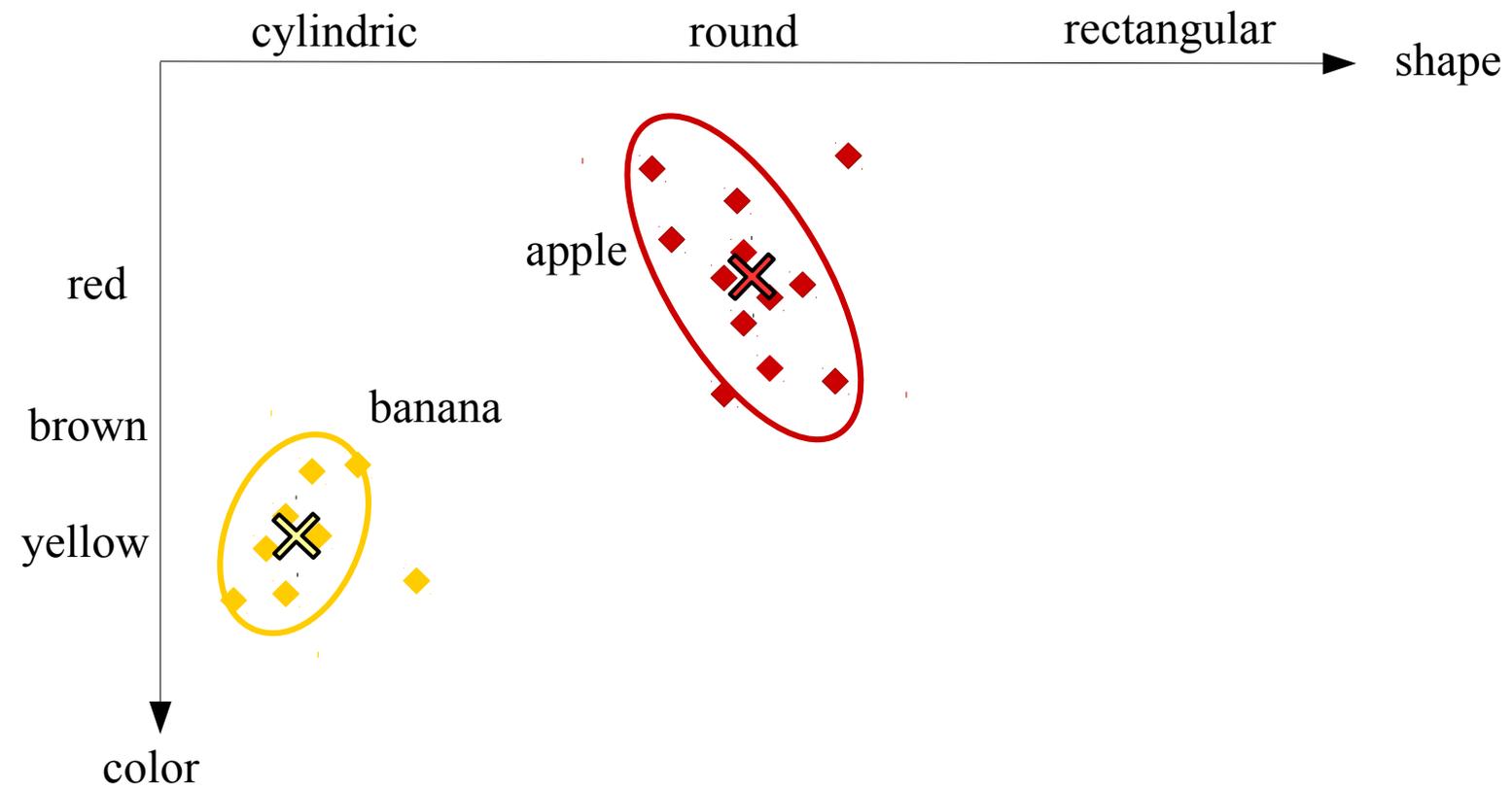
[Derrac2015]

Concept Formation in Conceptual Spaces

My PhD research

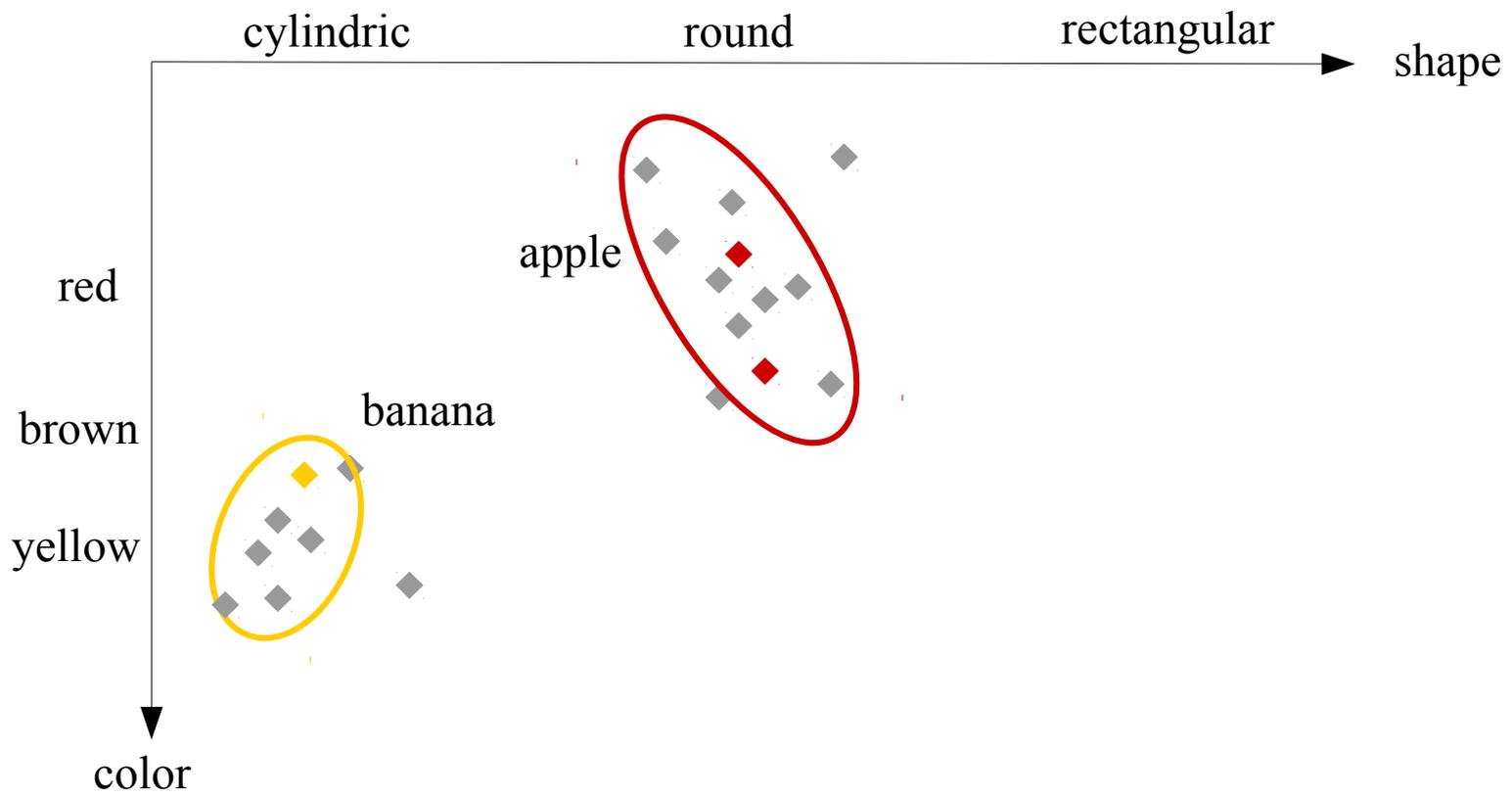
What's missing?

- How can concepts be „discovered“?
 - Theory tells us how to deal with a set of labeled data points...

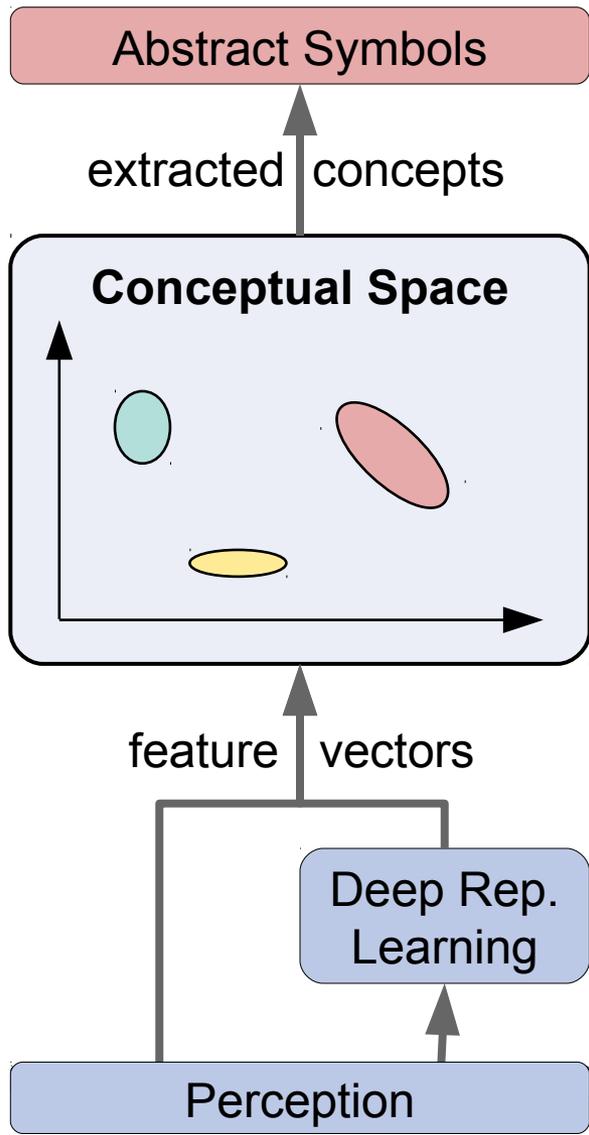


What's missing?

- How can concepts be „discovered“?
 - Theory tells us how to deal with a set of labeled data points...
 - ... but not how to deal with a stream of unlabeled data points!



My overall idea

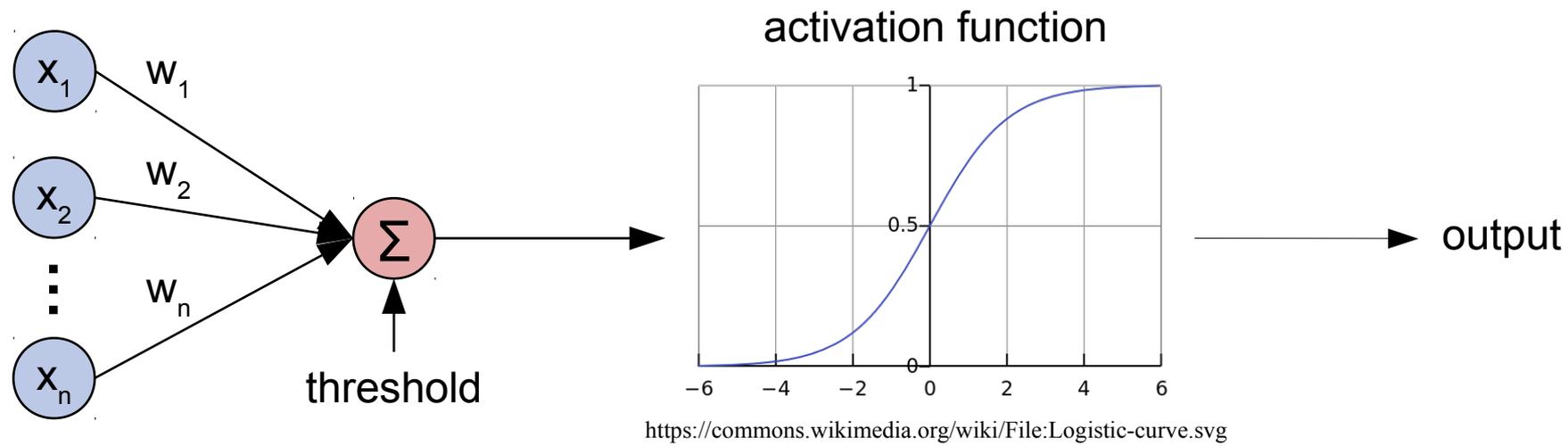


Deep (Representation) Learning

A neurally inspired machine learning technique

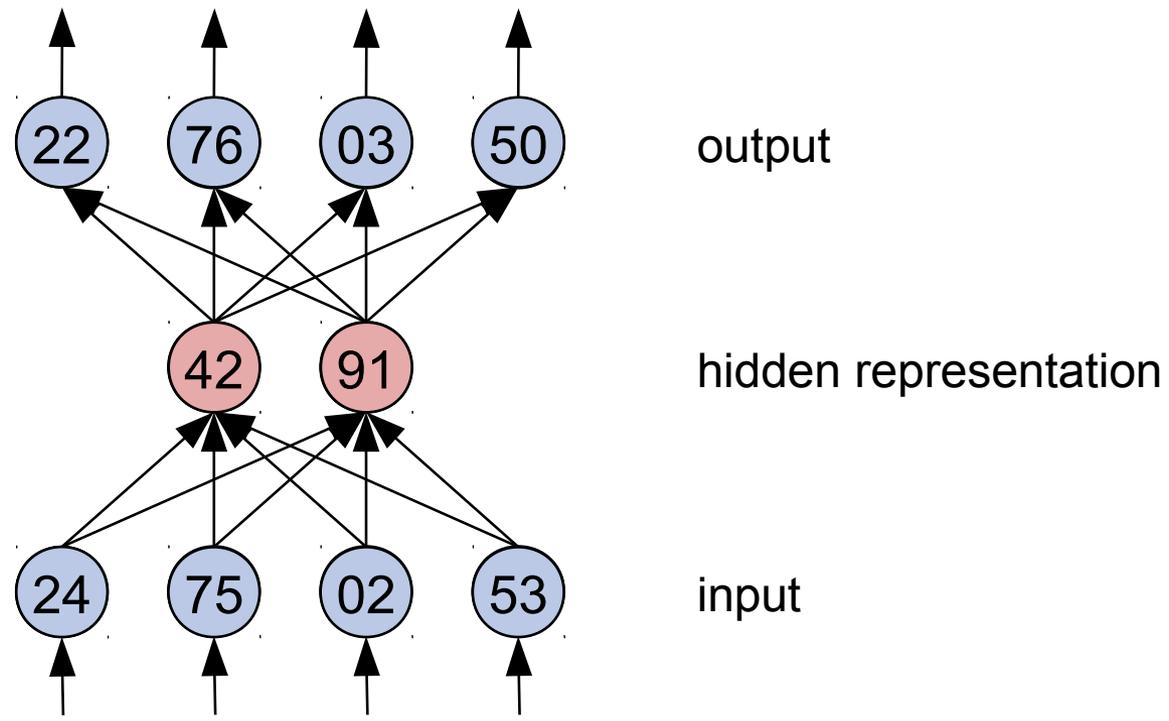
[Bengio2014]

- Inspired by biological neurons



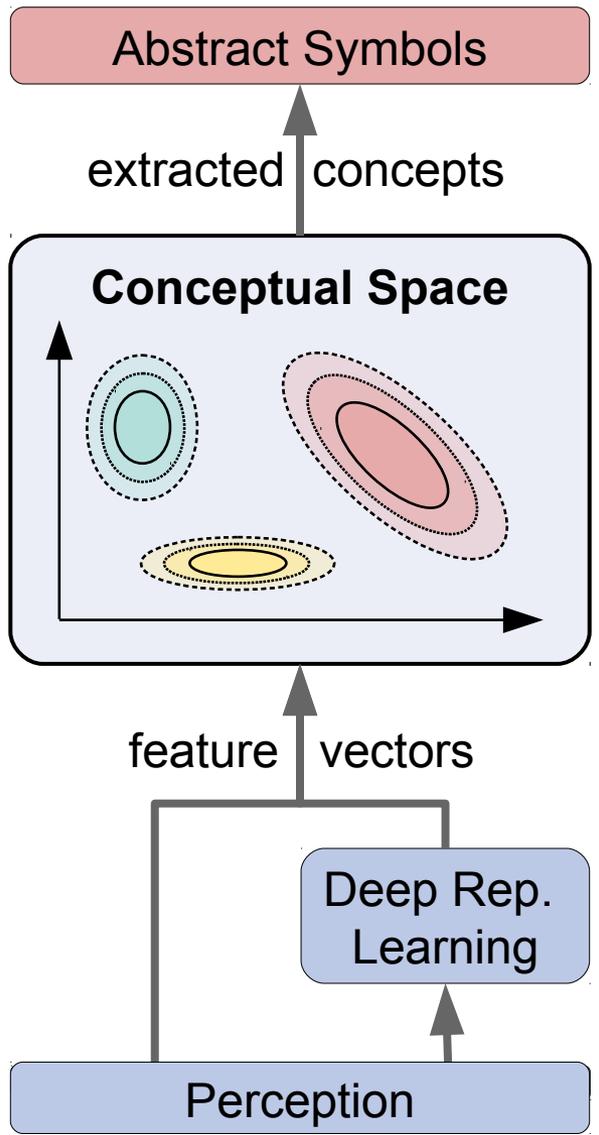
- Error = computed output - expected output
- Backpropagation: Adjust weights based on error gradient

- Artificial Neural Network: combine many artificial neurons
- Autoencoder: compress and reconstruct input



- Hidden neurons = dimensions in our conceptual space

My overall idea

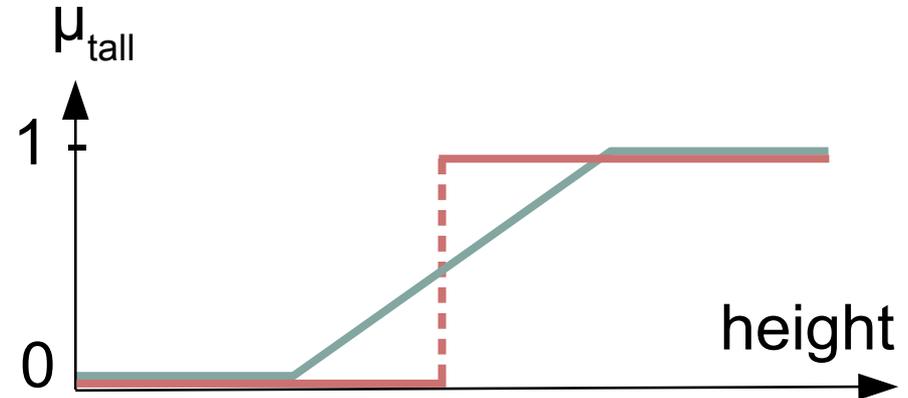


Fuzziness

The world is not black and white

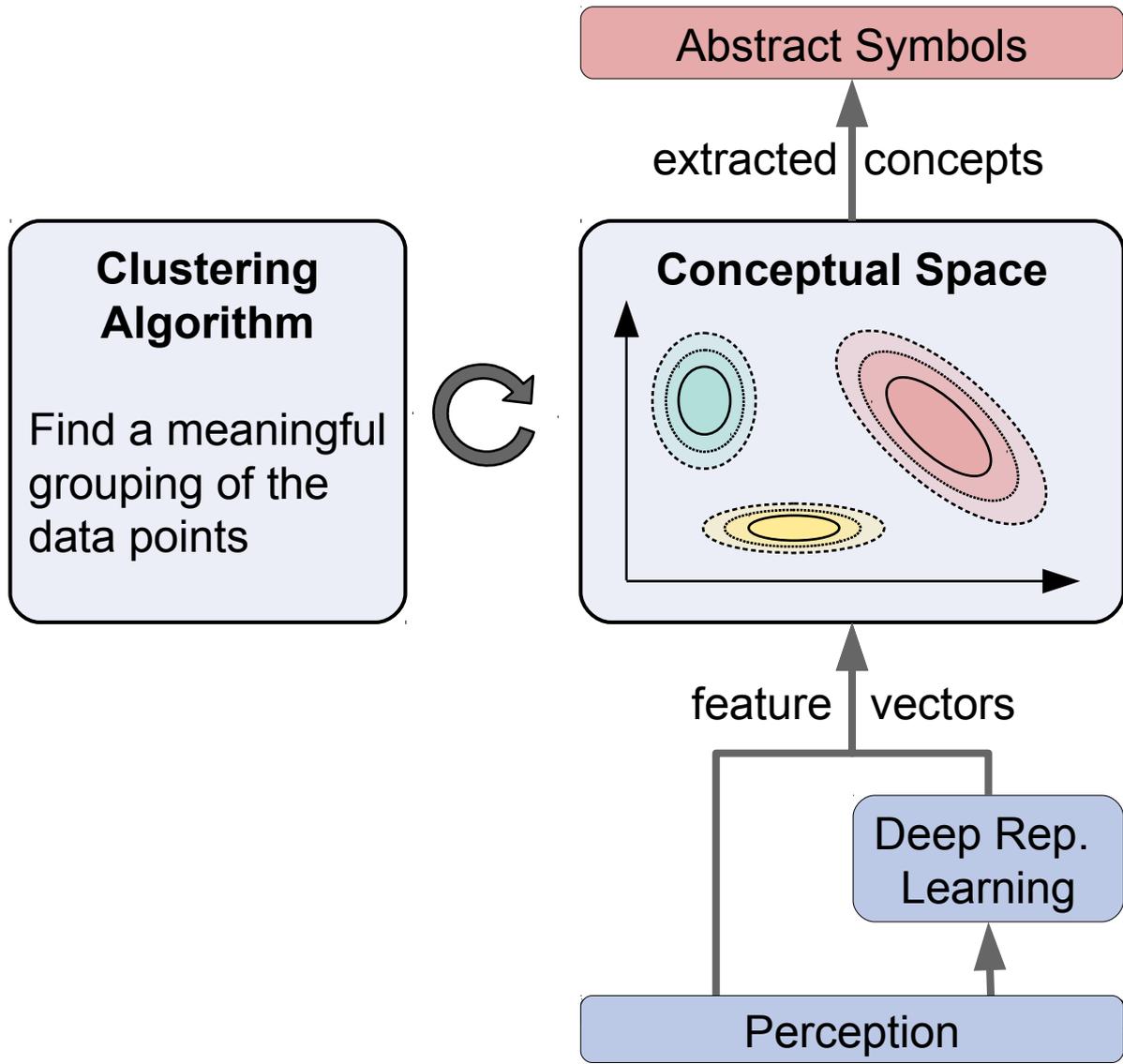
[Zadeh 1965]

- When is a person “tall”?
 - 1,80 m?
 - 1,79 m?
 - 1,78 m?
 - ...
 - 0,00 m?!



- Every threshold is somewhat arbitrary
- Also applies to other concepts (e.g., “red”)
- One solution: degrees of membership → fuzzy sets
 - Imprecise boundaries

My overall idea



Clustering

An unsupervised machine learning technique

- Using only data points without class information
- Goal: find meaningful groupings
- Assumption: closeness in feature space indicates similarity

- Example: k-means

- 1) Pick k random points as initial cluster centers
- 2) Assign each data point to the closest cluster center
- 3) Recalculate cluster centers: midpoint of assigned data points
- 4) Repeat until nothing changes

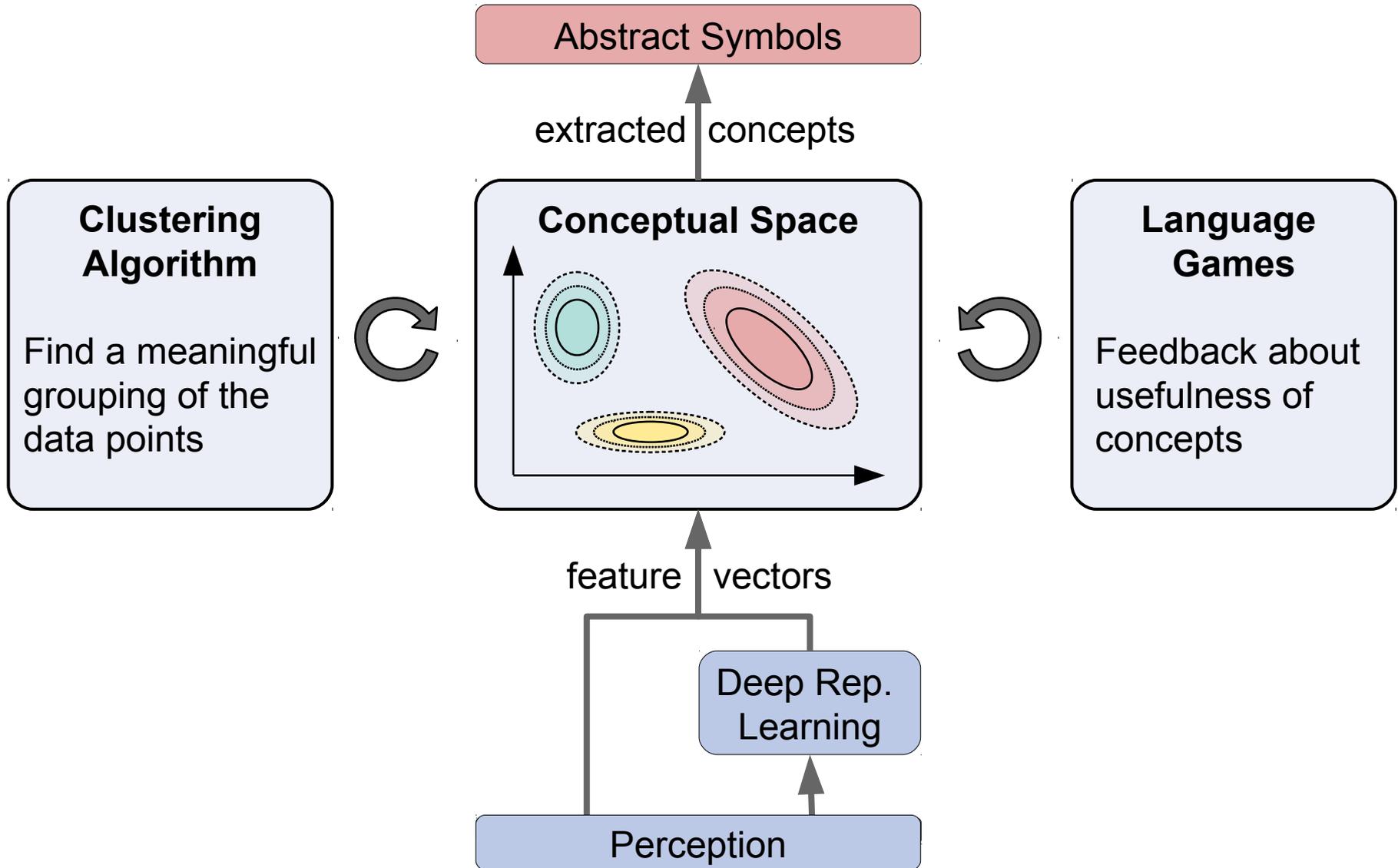


<http://shabal.in/visuals/kmeans/6.html>

- Problems: not incremental, fixed number of clusters

- Wish list:
 - Incremental (stream of observations)
 - Semi-supervised (take into account scarce feedback)
 - Unknown number of clusters
 - Fuzzy
 - Hierarchical
- Good news: some approaches seem (partially) fitting!
- Bad news: none of them fits perfectly!
- → I will probably need to combine different approaches

My overall idea



Language Games

Grounding concepts in communication

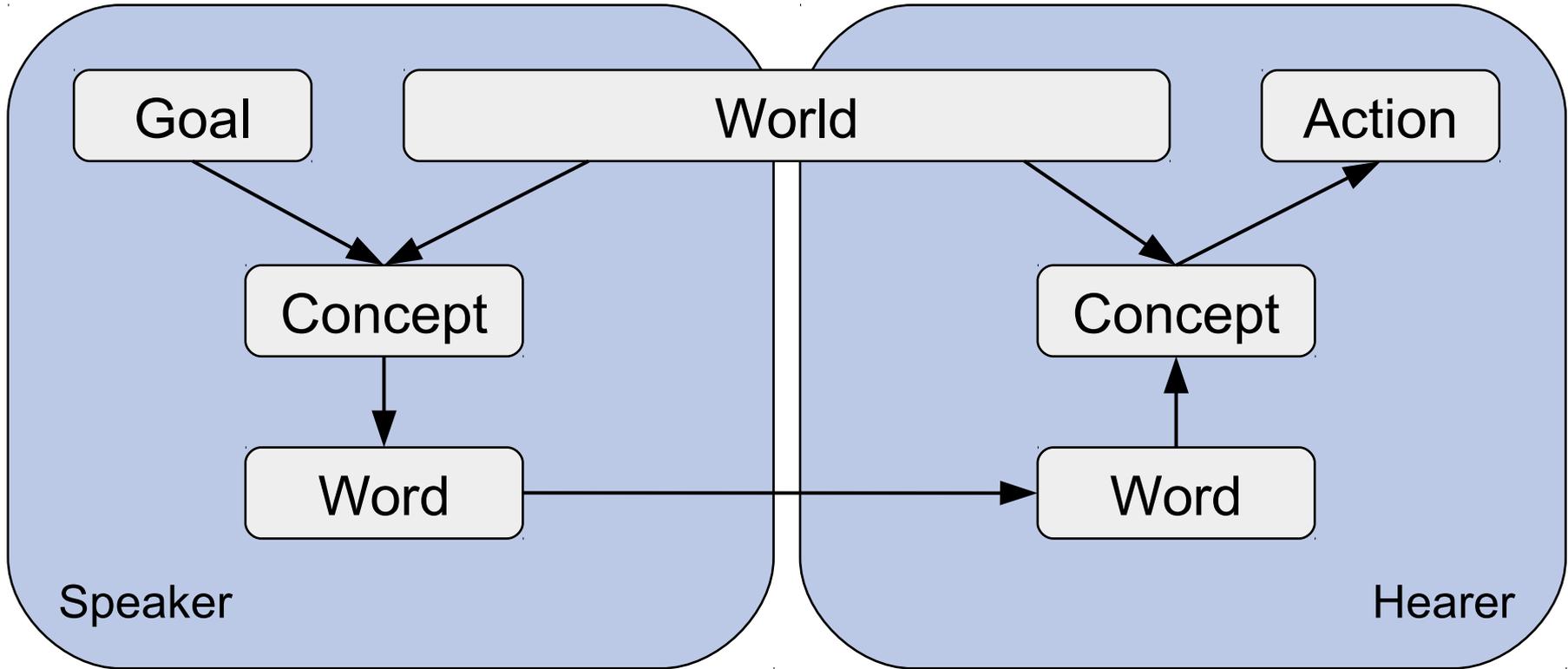
[Steels2015]

- Concepts are already grounded in perception
- ... but there are many ways in which the conceptual space can be divided up into concepts
- Still, humans seem to share their concepts (otherwise we could not communicate)
- Idea: use of concepts in communication gives further constraints

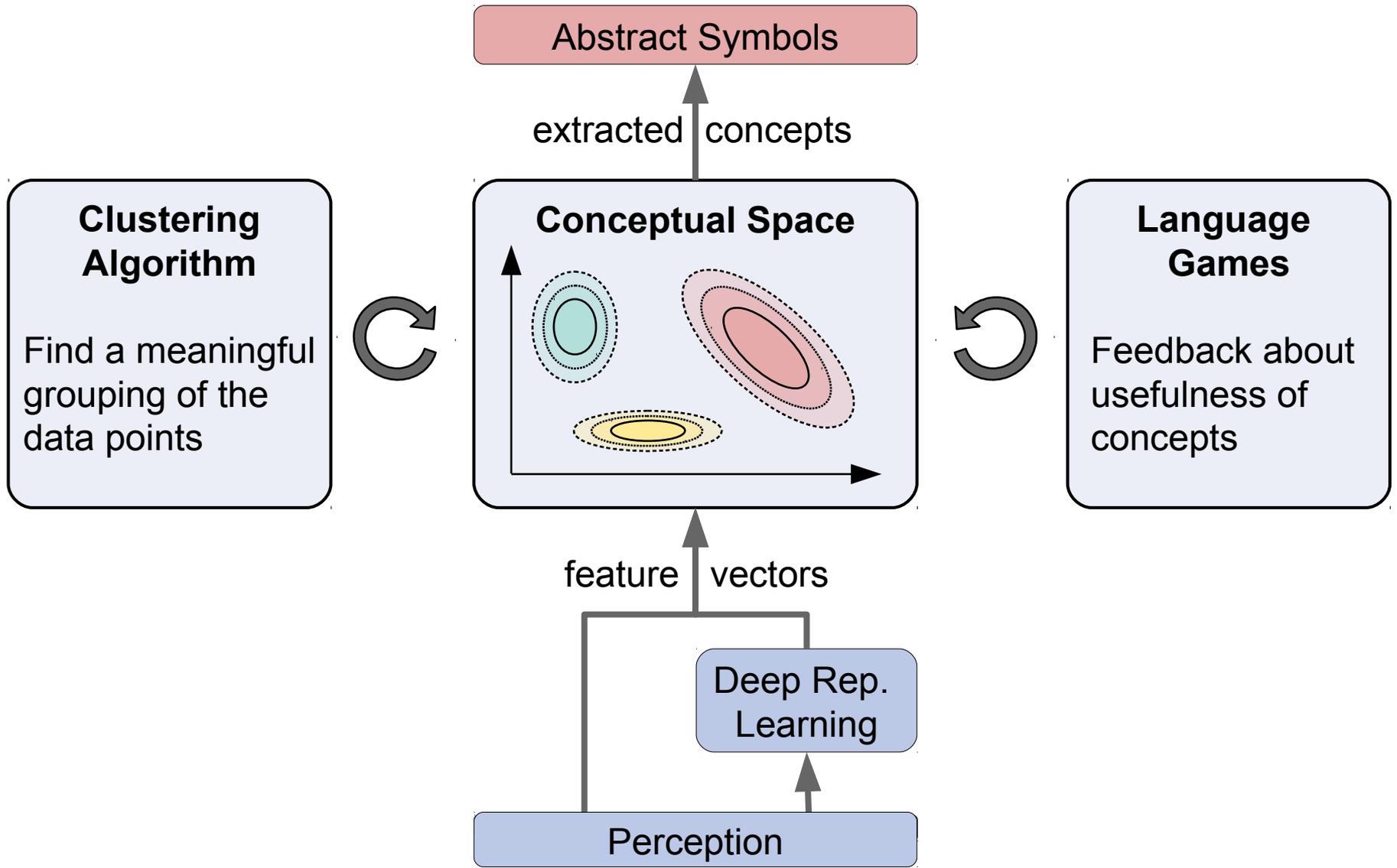
Language games



<http://shop.labbe.de/bastel-basics/stifte-kreiden/prime-jumbo.html>



My overall idea

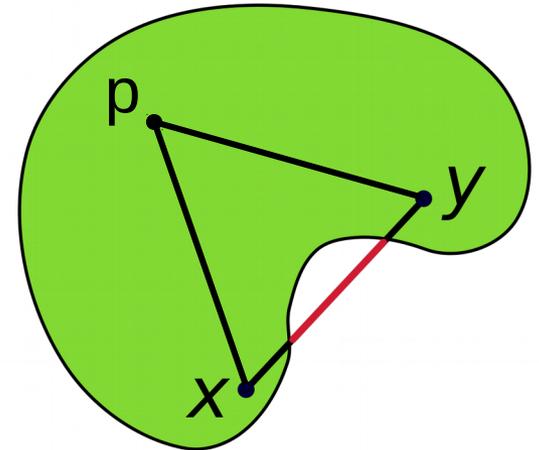
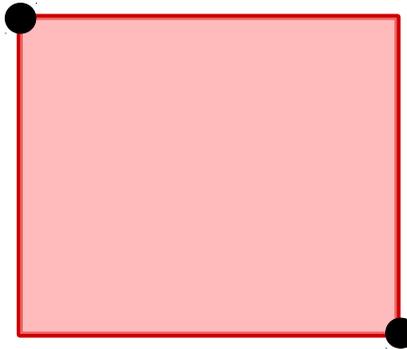
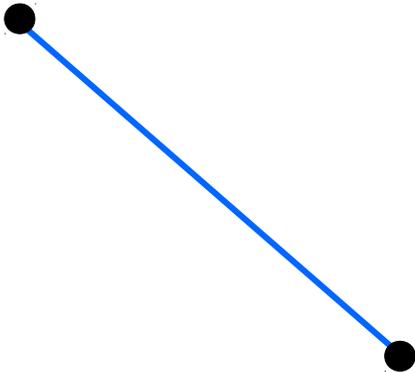


Why is this interesting?

- Principled way of creating symbols from perception
- Thorough mathematical formalization of the conceptual spaces framework (submitted to IJCAI-17)
- Combine language games with cognitive framework
- New type of clustering algorithm
- Learning in conceptual spaces

What have I done so far?

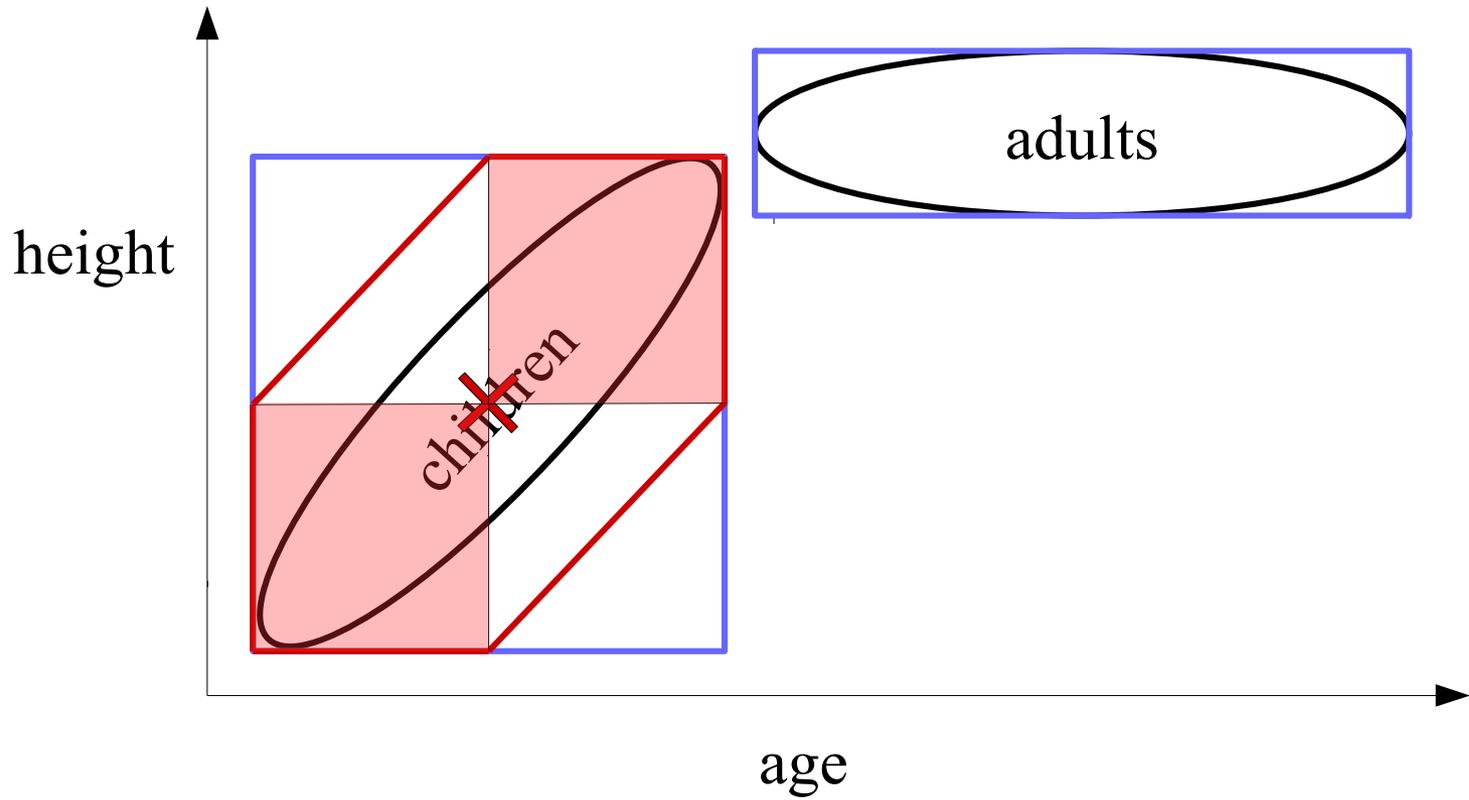
- Remember: Euclidean and Manhattan betweenness



https://en.wikipedia.org/wiki/Convex_set#/media/File:Convex_polygon_illustration2.svg

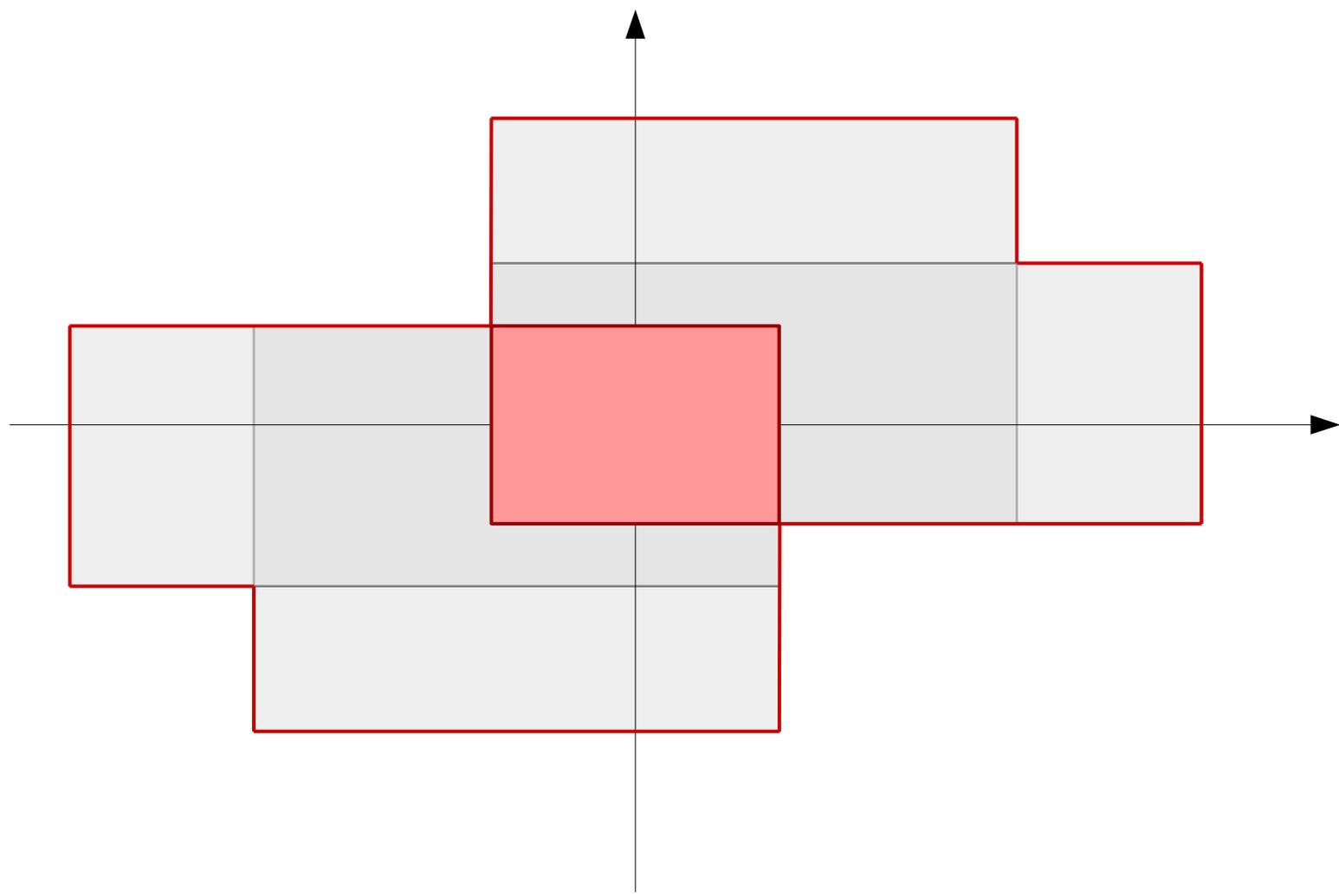
- Convex region C :
 - For all points $x, y \in C$ and for all z between x and y : $z \in C$
- Star-shaped region S :
 - Central point $p \in S$, for all z between p and $y \in S$: $z \in S$

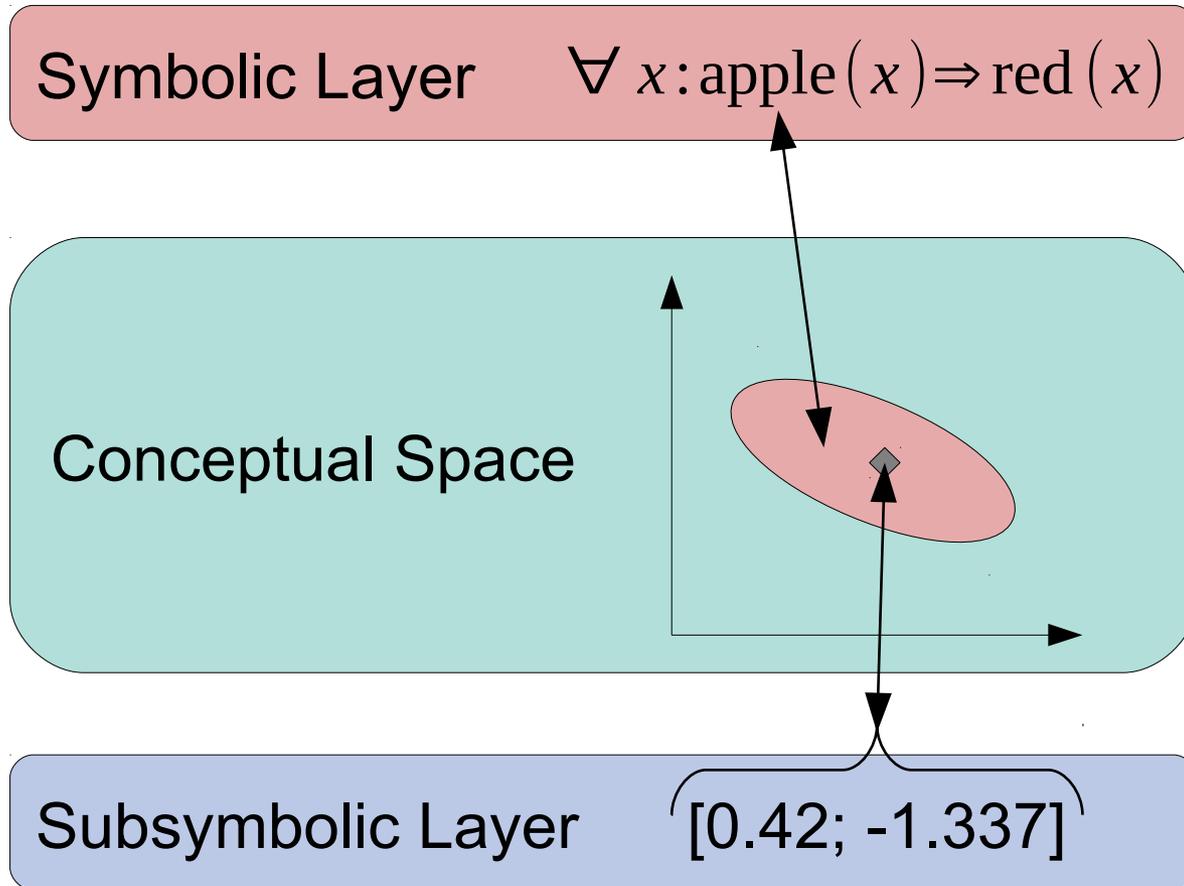
What have I done so far?



→ use star-shaped sets as basis for a formalization

What have I done so far?





- There are many possible applications of this theory

Thank you for your attention!

Questions? Comments? Discussions?



<https://www.lucas-bechberger.de>

- [Harnad1990]
 - S. Harnad, “The Symbol Grounding Problem”, Physica D: Nonlinear Phenomena, 1990
- [Gärdenfors2000]
 - P. Gärdenfors, “Conceptual Spaces: The Geometry of Thought”, MIT press, 2000
- [Gärdenfors2014]
 - P. Gärdenfors, „The Geometry of Meaning: Semantics Based on Conceptual Spaces“, MIT Press, 2014
- [Fiorini2013]
 - S. Fiorini, P. Gärdenfors, and M. Abel, “Representing Part-Whole Relations in Conceptual Spaces”, Cognitive Processing, 2013
- [Chella2004]
 - A. Chella, S. Coradeschi, M. Frixione, and A. Saffotti, “Perceptual Anchoring via Conceptual Spaces”, 2004

- [Derrac2015]
 - J. Derrac and S. Schockaert, “Inducing Semantic Relations from Conceptual Spaces: A Data-Driven Approach to Plausible Reasoning”, Artificial Intelligence, 2015
- [Bengio2014]
 - Y. Bengio, A. Courville, and P. Vincent: “Representation Learning: A Review and New Perspectives”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013
- [Zadeh1965]
 - L. Zadeh, "Fuzzy sets" Information and control, 1965
- [Baraldi1997]
 - A. Baraldi and F. Parmiggiani, “Novel Neural Network Model Combining Radial Basis Function, Competitive Hebbian Learning Rule, and Fuzzy Simplified Adaptive Resonance Theory”, Applications of Soft Computing, 1997
- [Steels2015]
 - Luc Steels, „The Talking Heads experiment: Origins of words and meanings“, Language Science Press, 2015