UNIVERSITÄT OSNABRÜCK

Methods of Artificial Intelligence

Knowledge Representation & Conceptual Spaces

Lucas Bechberger

Overview

- Different types of knowledge
- Different ways of representing knowledge
  - Frames
  - Semantic Networks & Ontologies
  - Prototypes
  - Machine Learning
- Conceptual Spaces

Why Knowledge Representation?

- Imagine the following scenario:

  An old woman sits in her arm chair, close to the fire place. "Robby", she calls, and her household robot appears in the living room. "How can I help you?" - "Please make me a cup of tea. The same one as yesterday."

  What knowledge does Robby need in order to fulfill this request?

- Discuss this question with your neighbors! (2 minutes)

Different types of knowledge

- Knowledge about Natural Language (e.g., grammar)
- Knowledge about objects (e.g., tea)
- Knowledge about events (e.g., “I made tea yesterday”)
- Knowledge about relations (e.g., “tea is a beverage”)
- Knowledge about procedures (e.g., how to make tea)
- Knowledge about knowledge (e.g., “I know that I know X”)
- ...

Different ways of representing a table

	Hamburg	Berlin
Bielefeld	253.6	384.4
Gießen	442.0	468.2

```
distance(Bielefeld, Hamburg, 253.6)
distance(Bielefeld, Berlin,    384.4)
distance(Giessen, Hamburg, 442.0)
distance(Giessen, Berlin,      468.2)
```

```
,Hamburg,Berlin
Bielefeld,253.6,384.4
Gießen,442.0,468.2
```

```
<table>
  <tr>
    <td></td><td>Hamburg</td><td>Berlin</td>
  </tr>
  <tr>
    <td>Bielefeld</td><td>253.6</td><td>384.4</td>
  </tr>
  <tr>
    <td>Giessen</td><td>442.0</td><td>468.2</td>
  </tr>
</table>
```

```
\begin{table}\centering
\begin{tabular}{|l|l|l|}
\cline{1-3}
\vbox to1,88ex{\vspace{1pt}\vfil\hbox
to12,80ex{\hfil \hfil}} &
\vbox to1,88ex{\vspace{1pt}\vfil\hbox
to12,80ex{\hfil Hamburg\hfil}} &
\vbox to1,88ex{\vspace{1pt}\vfil\hbox
to12,80ex{\hfil Berlin\hfil}} \\
\cline{1-3}
[...]
\cline{1-3}
\end{tabular}
\end{table}
```

How can we represent knowledge?

- Logics (predicate, first order, fuzzy, modal, …)
- Frames
- Semantic Networks & Ontologies
- Prototypes
- Machine Learning
- Conceptual Spaces
- ...

Frames

- **Name**: how is this concept called?
- **Slots**: what are attributes/properties of this concept?
- **Fillers**: values that can fill a given slot
  - Slots can be filled by other frames (e.g., “mammal” or “hooves”)

Horse
ISA: mammal
Has-part: hooves, mane, tail, nose, legs, back
Motion: walk, trot, gallop
Eats: oats, hay, grass, carrots
Speech: neigh, whinny, snort

[Minsky1974]

Frames: Alignment

Horse
ISA: mammal
Has-part: hooves, mane, tail, nose, legs, back
Motion: walk, trot, gallop
Eats: oats, hay, grass, carrots
Speech: neigh, whinny, snort

✓

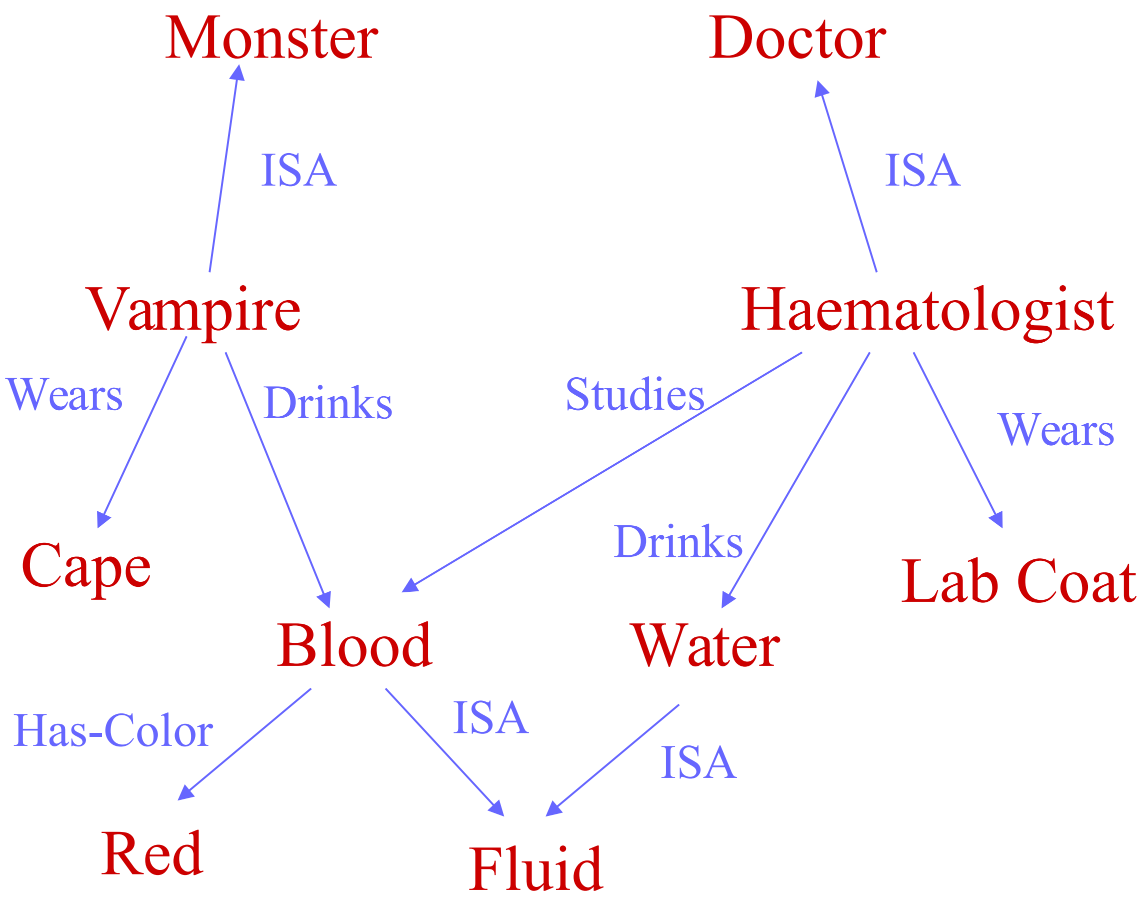
Cow
ISA: mammal
Has-part: udder, tail, nose, legs, back
Motion: walk, trot
Eats: hay, grass
Speech: moo

Horse
ISA: mammal
Has-part: hooves, mane, tail, nose, legs, back
Motion: walk, trot, gallop
Eats: oats, hay, grass, carrots
Speech: neigh, whinny, snort

✗

Sofa
ISA: furniture
Has-part: seat, back, arm rest
Made-of: leather
Color: black
Seats: three

Semantic Networks & Ontologies



- A semantic network consists of labelled nodes and directed labeled edges.
- The nodes in a semantic network correspond to **concepts.**

Semantic Networks & Ontologies

- A semantic network can be expressed by triples
  `<source_concept, relation_label, target_concept>`
  e.g., `<Vampire, Drinks, Blood>`

- Reasoning
  - Description Logics: Classes, Relations, Instances

- Conceptual distance
  - number of edges between two nodes?

- Making things more complex
  - Relation labels can also be concepts
    - e.g., `<Drinks, ISA, Consumes>`
  - Edges can be weighted (e.g., reflecting uncertainty)

Semantic Networks & Ontologies

- Used in the Semantic Web:
  - Web Ontology Language (**OWL**)
    - Define classes & their relations
  - Resource Description Format (**RDF**)
    - Standard for describing triples
  - SPARQL Protocol and RDF Query Language (**SPARQL**):
    - SQL-like language to write queries about triples

- Example: https://knowledgestore.fbk.eu/

- How many RDF triples does DBpedia contain?

  3 million   30 million   300 million   3 billion

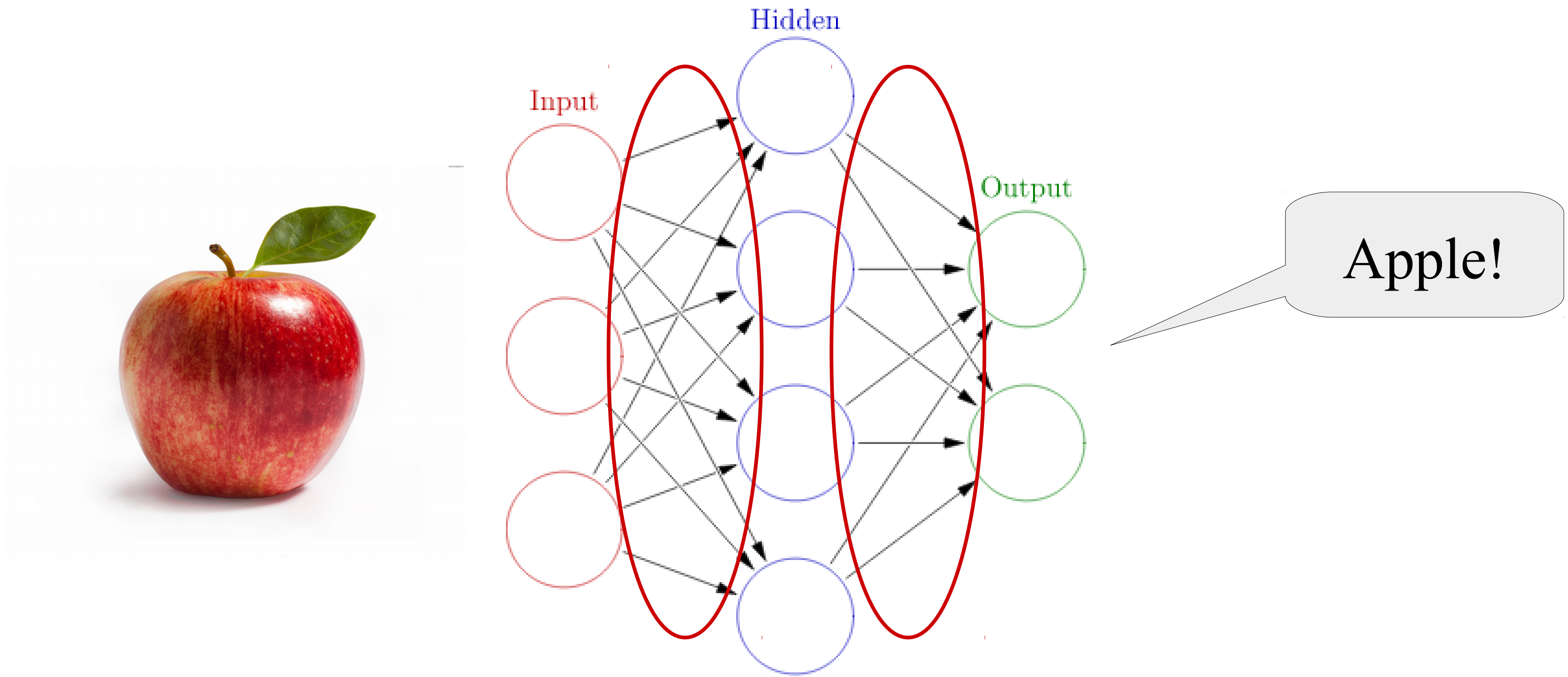
Prototype theory

- "Imagine an apple!"

- Some instances are more typical than others
- Represent via feature lists:
  - Apple: round, red, ~7cm diameter, has stem, …
  - How many of the features are present in a given example?

[Rosch1975]

Machine Learning



- Artificial neural network learns to correctly classify images
- ANN “knows” what apples look like
- Where is this knowledge encoded?
- Problem: not really interpretable...

Conceptual Spaces: Motivation

- If a computer has the following logical knowledge:

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

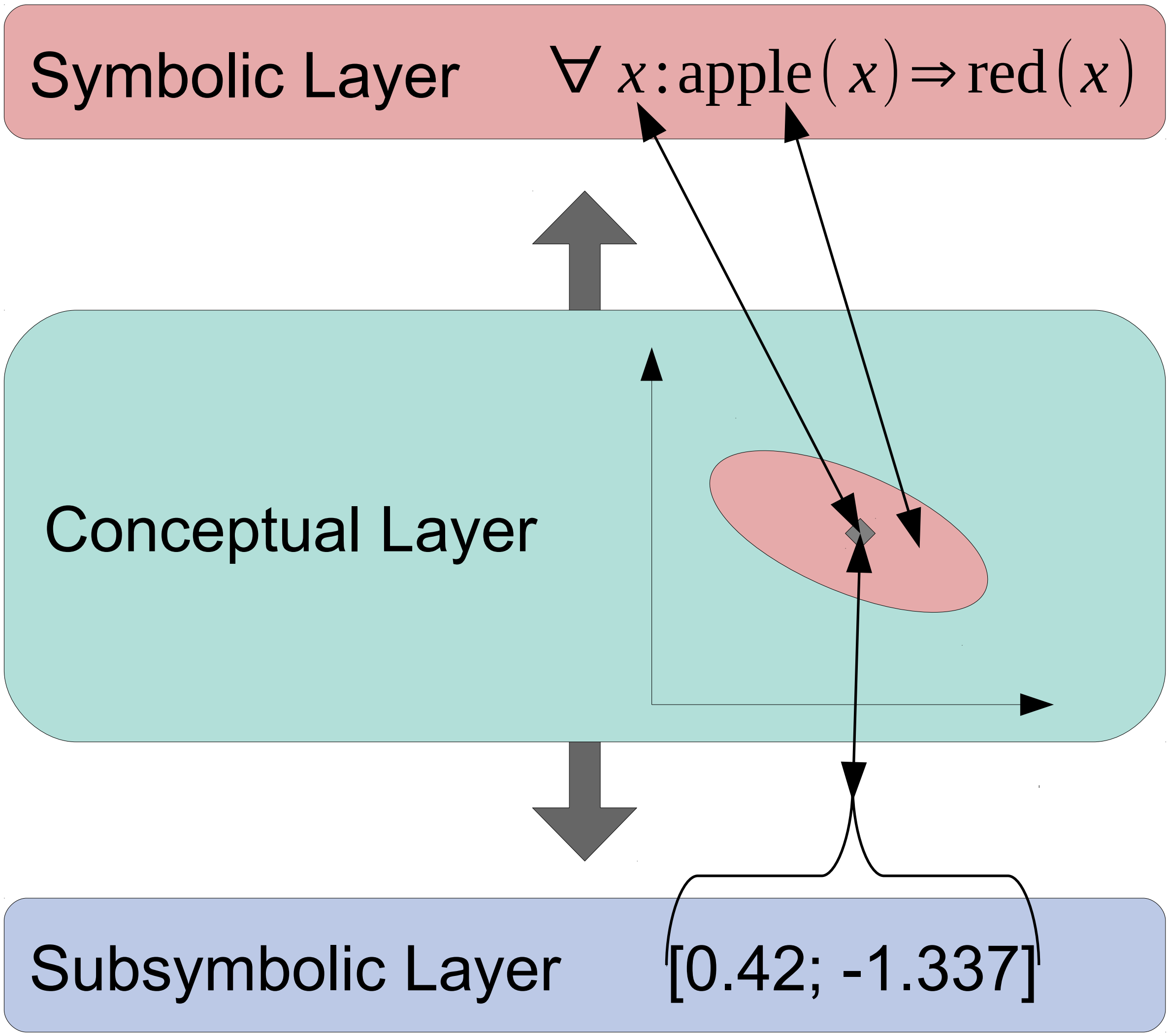
- Does it know what „red“ means?
- Does it know what „apple“ means?

- 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

Conceptual Spaces: Motivation



Formal Logics, Ontologies, ...

Geometric Representation

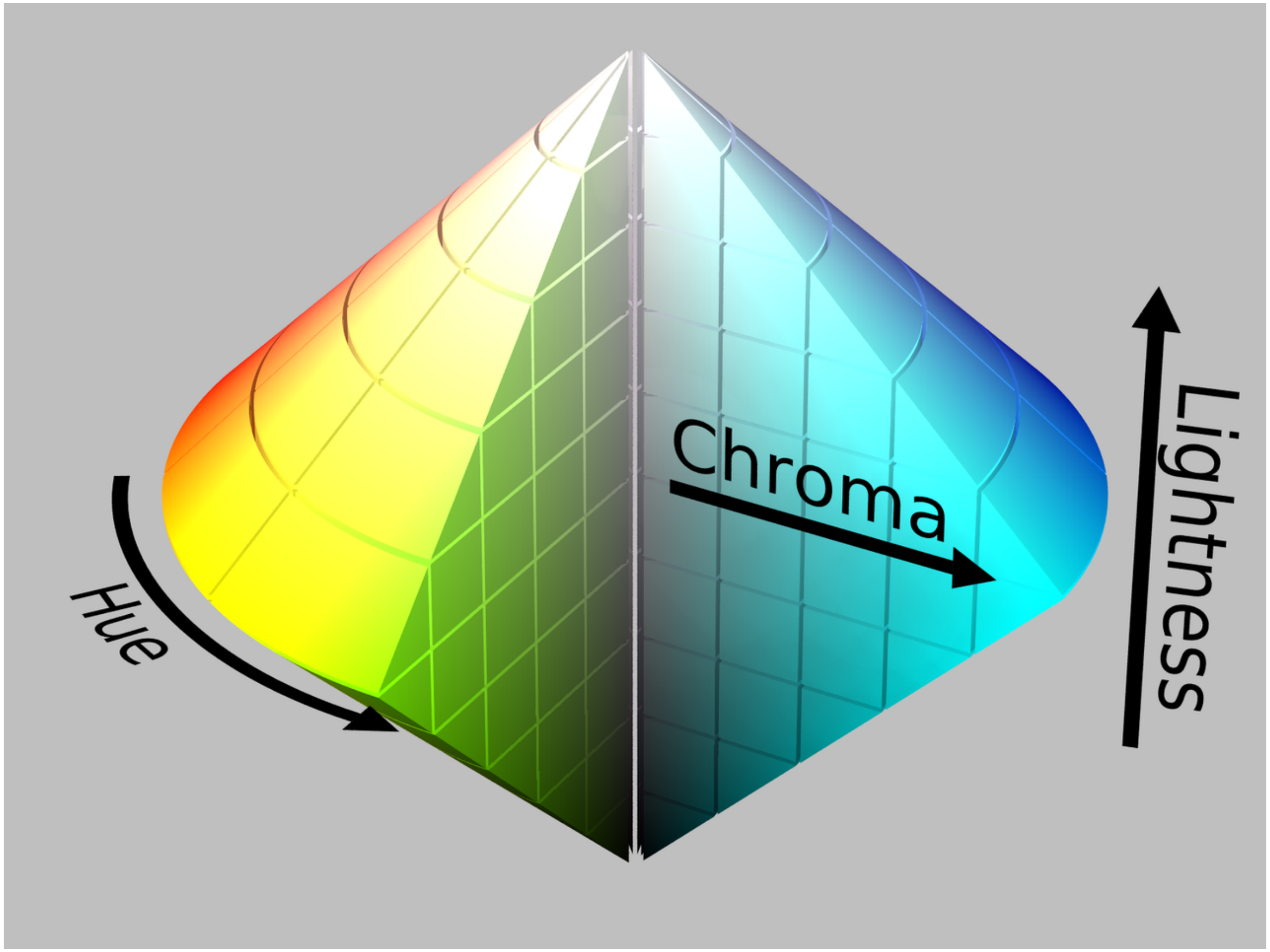
Sensor Values, Machine Learning

Dimensions & Domains

- Quality dimensions
  - Different ways stimuli are judged to be similar or different
  - Interpretable by humans → meaningful features
  - E.g., temperature, weight, brightness, pitch, height
  - Difference to word embeddings!
- Domain
  - Set of integral dimensions that inherently belong together
  - Each domain represents a different perceptual modality
  - Color: hue, saturation, and brightness
  - Sound: pitch and volume

[Gärdenfors2000]

Example: the color domain



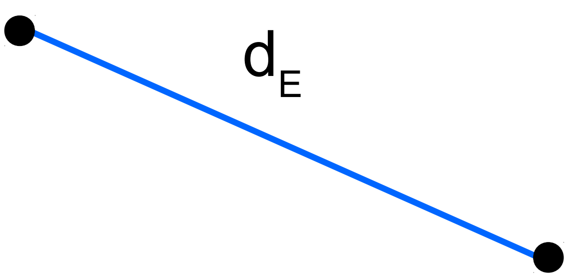
https://en.wikipedia.org/wiki/HSL_and_HSV#/media/File:HSL_color_solid_dblcone_chroma_gray.png

UNIVERSITÄT OSNABRÜCK

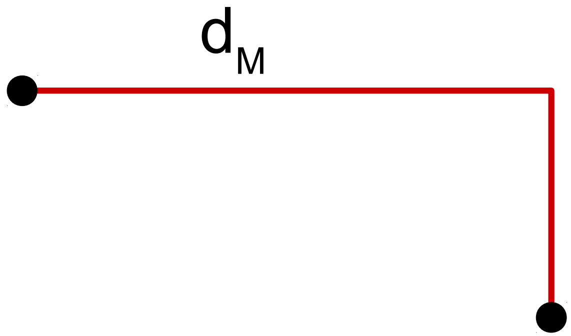
Distance & Similarity

- Distance in a conceptual space
  - Distances on individual dimensions need to be aggregated
  - Within a domain → Euclidean distance
  - Between domains → Manhattan distance
  - Context can influence distance by putting weights on dimensions

Euclidean Distance



Manhattan Distance



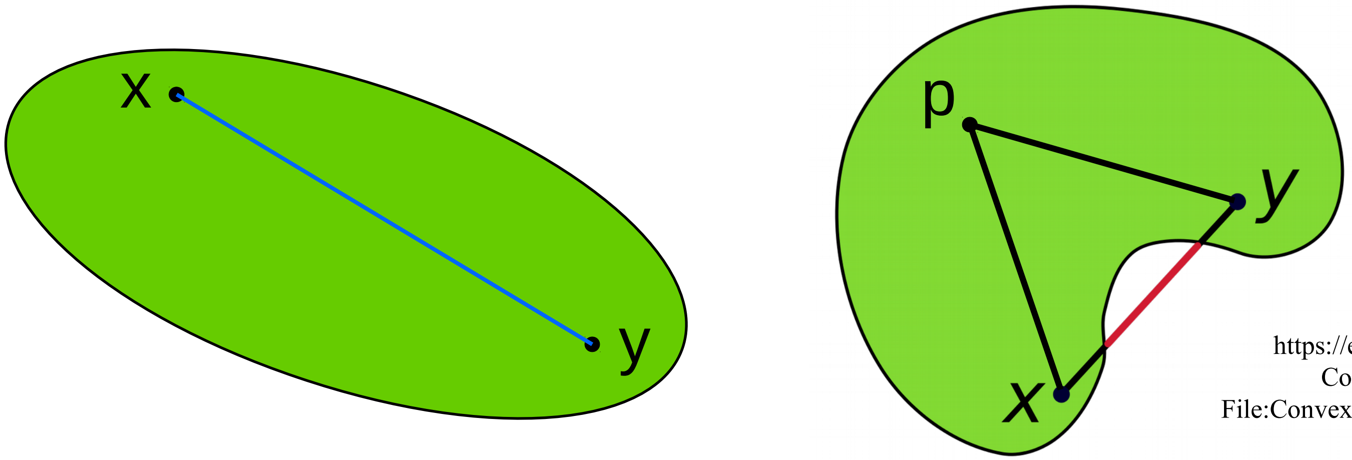
- Distance in a CS is inversely related to similarity
  - The closer two things are in a CS, the more similar they are
  - Exponential decay: $e^{-c \cdot d(x,y)}$

Betweenness & Convexity

- Betweenness is a predicate *B(x,y,z)*
  - "*y* is between *x* and *z*"
  - Given a metric *d*, we can define:

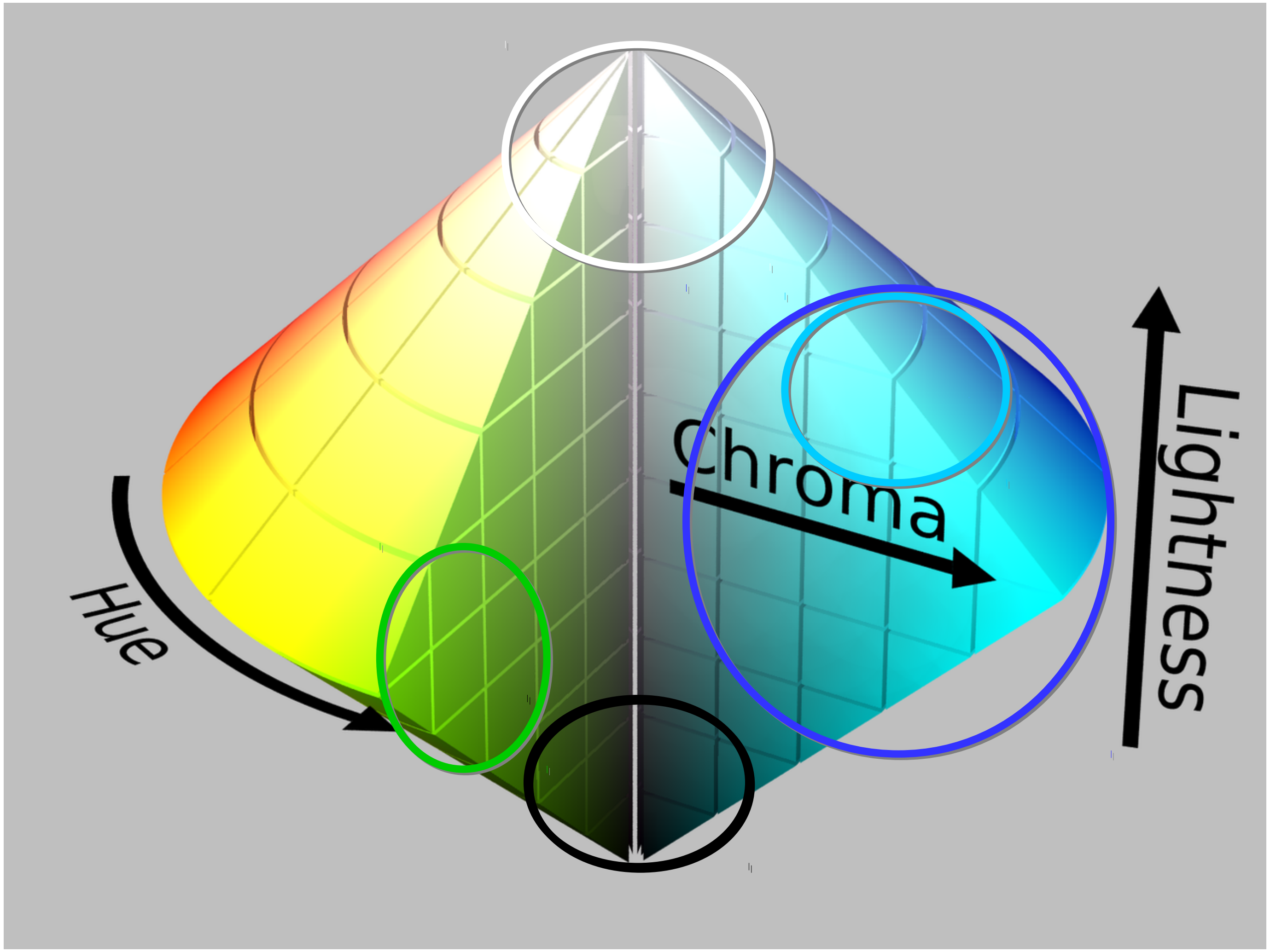
$$B(x,y,z) :\leftrightarrow d(x,y) + d(y,z) = d(x,z)$$

- *Convex region C:*
  - $\forall x,z \in C : \forall y : B(x,y,z) \Rightarrow y \in C$
  - If *x* and *z* are in *C*, then everything between them also is



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

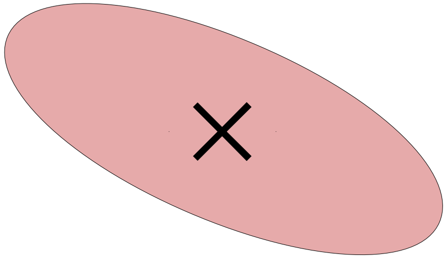
Example: the color domain



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

Connection to prototype theory

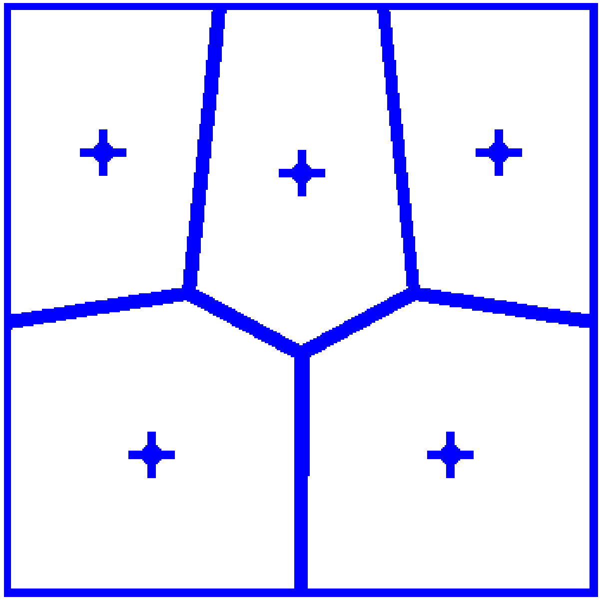
- Prototype theory of concepts
  - Each concept is mentally represented by a prototype
  - Prototype = abstract summary representation
    - E.g., average instance
  - Categorization: compare stimulus to all prototypes
    - → best match wins



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

Connection to prototype theory

- Voronoi tessellation
  - Given a set of central points $\{p_1, \ldots, p_n\}$
  - 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

Concepts

- Example: „apple“
  - Color: red
  - Shape: round
  - Texture: smooth
  - Taste: sweet

- Defined across multiple domains: combination of properties
  - Different „importance“ to the concept (influenced by context)
  - Potentially correlated (e.g., taste and color)

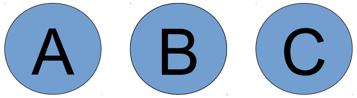
- Criterion C:
  - A natural concept is represented as a set of convex regions in a number of domains together with an assignment of salience weights to the domains and information about how the regions in different domains are correlated.

Concept combination

- "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***
- "purple apple"
  - ***purple*** is incompatible with ***apple***'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

Plausible/Common-sense reasoning

- 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

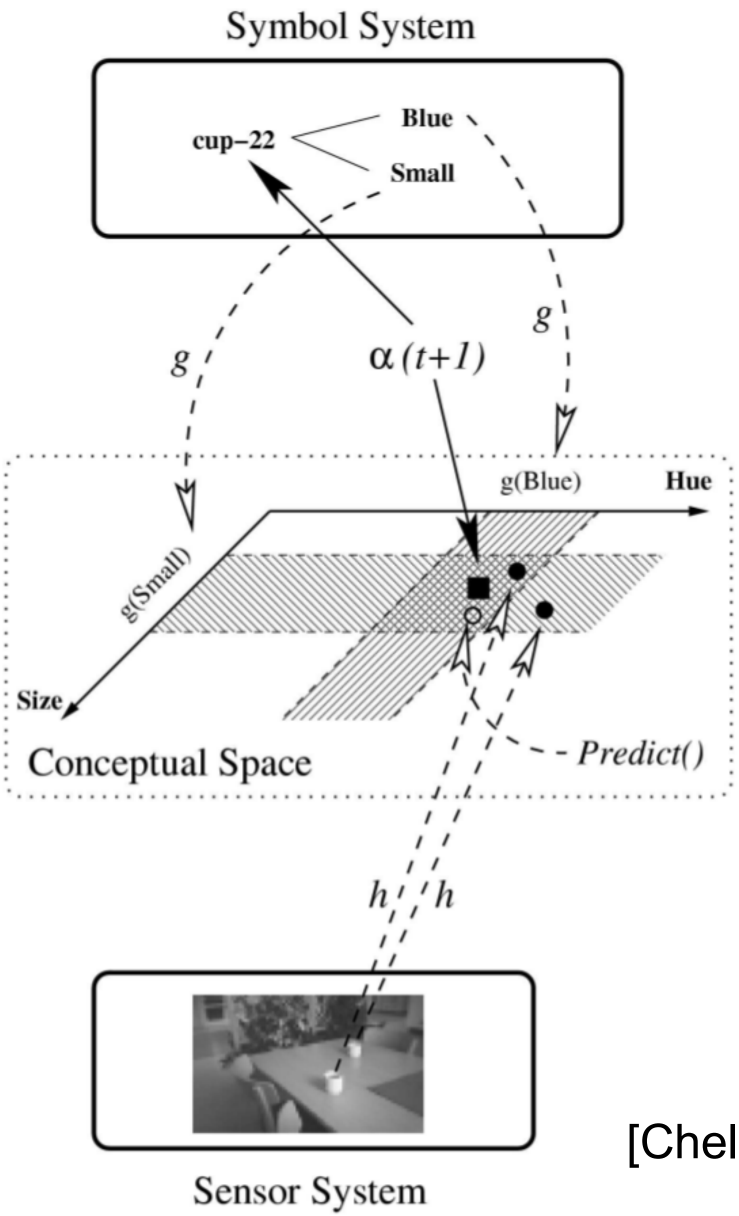


- Derive conceptual space from textual data (word vectors)
- Find interpretable dimensions in this space
- Some example results:
  - “wine shop” is between “gourmet shop” and “liquor store”
  - Difference between “Jurassic Park” and “Kill Bill: Vol. 1”:
    - “dinosaurs”, “the expedition”, “the scientist”

[Derrac2015]

Anchoring in robotics

- Symbol system
  - Symbols (“cup-22”)
  - Predicates (“blue”)
  - g: predicates → areas
- Sensor system
  - Takes measurements at each time step
  - h: measurements → points
- Anchor
  - α: time → symbols x points
  - Ties symbols to observations
  - Needs to be updated

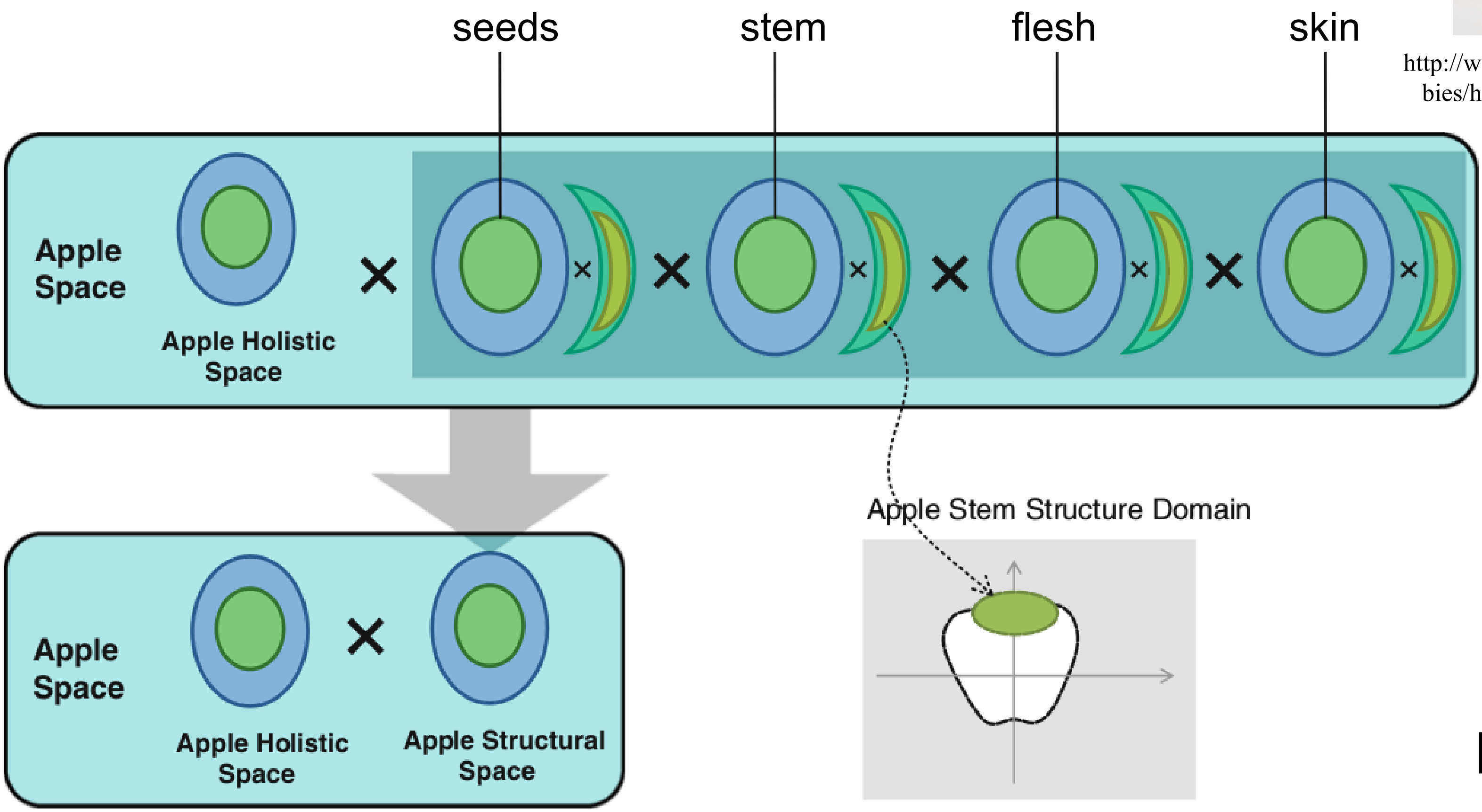


[Chella2004]

Representing composite concepts

- 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/



[Fiorini2013]

Semantic grounding of word classes

- 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]

Conceptual Spaces

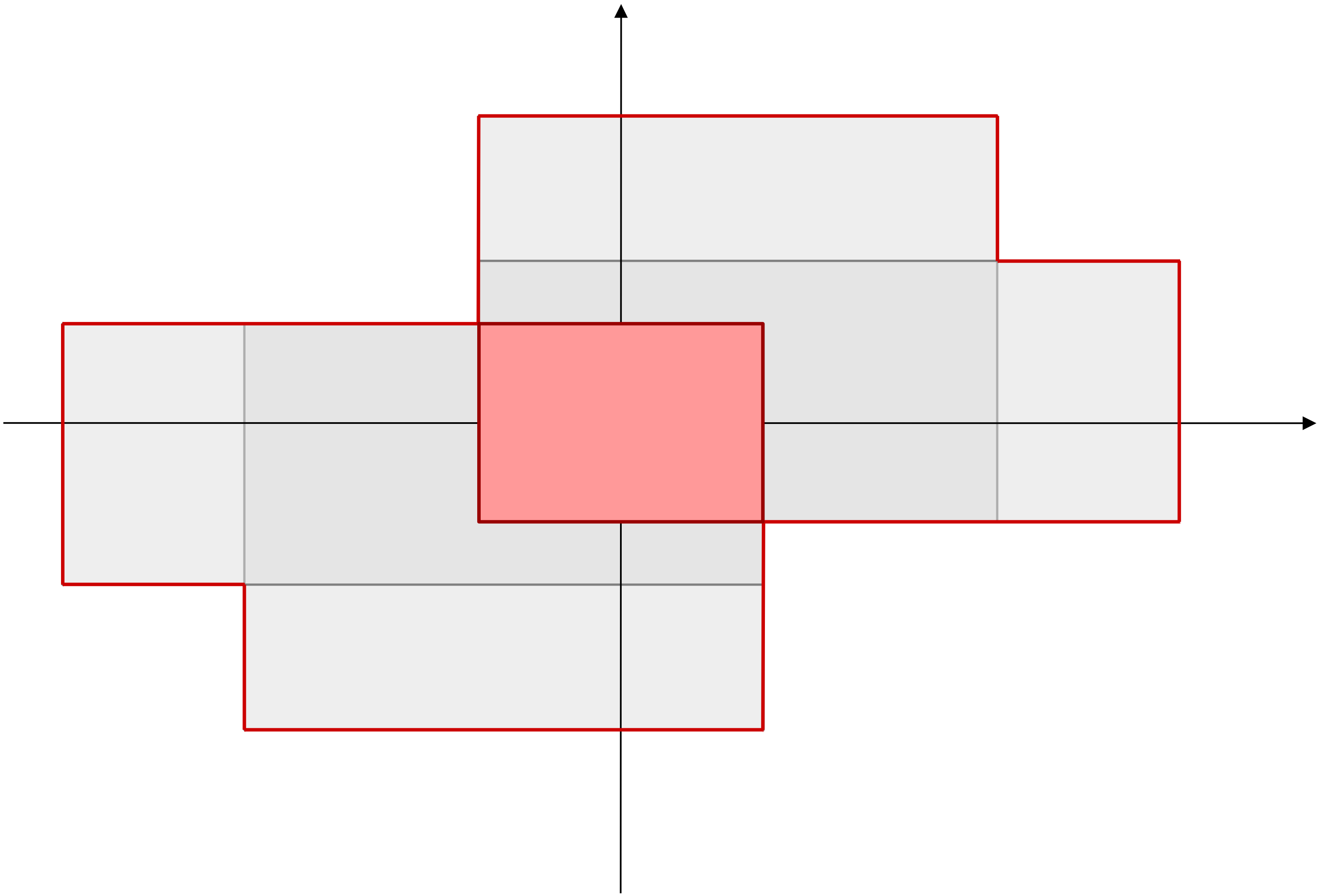
- Start with a set of interpretable dimensions
- Group them into domains
- Properties = convex regions within a single domain
- Concepts span multiple domains
  - Salience values and correlation information

- Connections to:
  - Machine learning (feature space)
  - Prototypes (centroid of region)
  - Frames (domain structure)
  - Ontologies (ISA-relation is simply subsethood)
  - Logics (logical AND = intersection of regions, etc.)
  - Linguistics (adjectives = properties, nouns = concepts, ...)

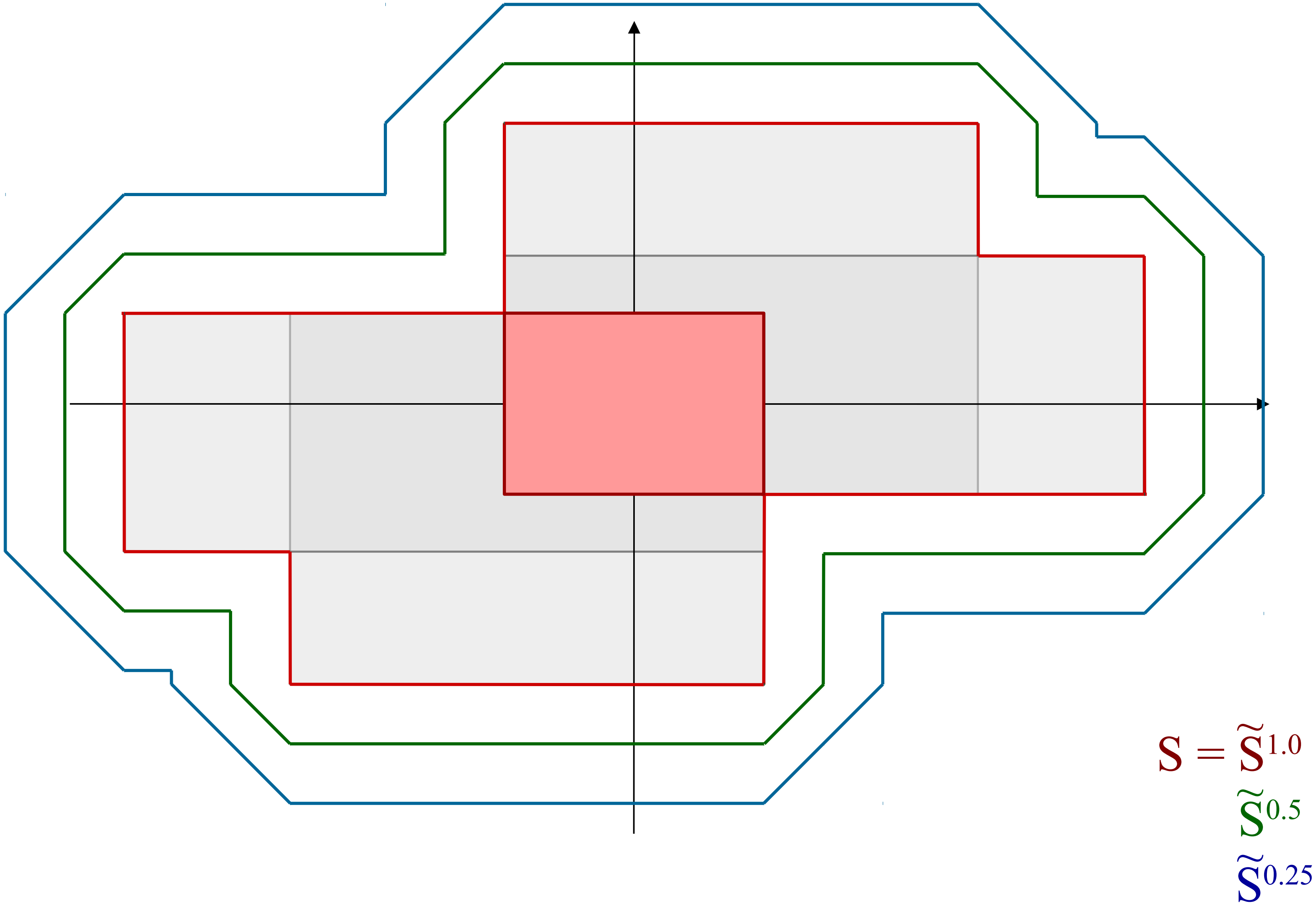
My research on Conceptual Spaces

- There are still open questions from an AI perspective...
- How can we mathematically formalize and implement this framework?
- Where do the dimensions of a conceptual space come from?
- How can an autonomous agent learn about regions in a conceptual space in a cognitively plausible setting?

Formalizing Conceptual Spaces



Formalizing Conceptual Spaces

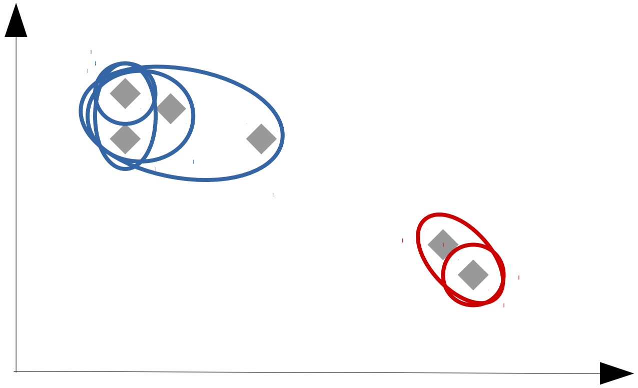


Where do dimensions come from?

- Isn't it quite obvious which dimensions we need?
  - Color: hue, saturation, brightness
  - Temperature: temperature
  - Emotions: valence, arousal
  - …
- … but what about shape?
  - it's surprisingly hard to define this domain with a handful of dimensions
  - Roundness, convexity, number of corners?
  - But how to extract those from images?
- Idea: **learn** the dimensions of a given domain with ANNs

Learning new concepts

- We look for meaningful regions in the conceptual space
  - Concepts = clusters of data points
- Observed objects usually come without class information
  - Unsupervised learning
- Observing one object at a time, limited memory
  - Stream of data points, incremental processing



Final Slide

- Take home messages
  - There are different types of knowledge
  - There are different ways of representing knowledge
    - Logics
    - Frames
    - Semantic Networks & Ontologies
    - Prototypes
    - Parameters learned in Machine Learning
    - Conceptual Spaces

- Camouflaged advertising
  - There will be a seminar on conceptual spaces in the summer term!

References

- [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

References

- [Derrac2015]
  - J. Derrac and S. Schockaert, “Inducing Semantic Relations from Conceptual Spaces: A Data-Driven Approach to Plausible Reasoning”, Artificial Intelligence, 2015
- [Minsky1974]
  - Minsky, M. (1974). A framework for representing knowledge.
- [Rosch1975]
  - Rosch, E. (1975). Cognitive representations of semantic categories. Journal of experimental psychology: General, 104(3), 192.