

Conceptual Spaces: The Geometry of Thought

Gärdenfors, P.

Igor de Camargo e Souza Câmara

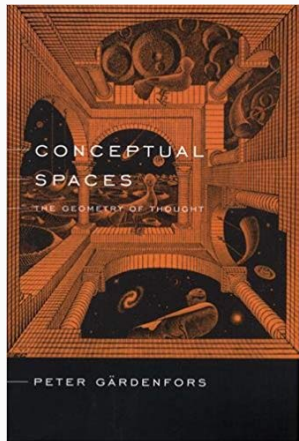
3 de abril de 2019

IME - USP

Introdução

Introdução

- Publicado em 2000.
- Projeto de pesquisa amplo, abarcando uma série de disciplinas (computação, linguística, ciências cognitivas, filosofia).
- Continuado em uma série de artigos e, mais recentemente, no livro *The Geometry of Meaning: Semantics Based on Conceptual Spaces* (2014).
- Proposta simultaneamente **explicativa** e **construtiva**.



Espaços Conceituais

Gardenfors identificou dois níveis na discussão sobre representação de conceitos.

Clássico

- Paradigma aristotélico.
- Conceitos são determinados por condições necessárias e suficientes, descritos em uma linguagem simbólica.
- Raciocínio e inferência conceitual são vistos como manipulação, segundo regras, desses símbolos.
 - Por exemplo, manipulação sintática na teoria linguística de Noam Chomsky.
- Cognição := Entradas sensoriais \rightarrow Mente \rightarrow Gera sentenças \rightarrow Computa essas sentenças de acordo com regras.

Conexionista

- Ligada à tradição filosófica do empirismo, cuja base é *associação de ideias*.
- Retomada com força pelo trabalho de McClelland & Rumelhart com redes neurais artificiais e verbos irregulares.
- Atividades nervosas dependem de sua interação com o entorno - i.e. com a estrutura da rede neural.
- Sem semântica explícita: é difícil descrever o que está sendo representado.

Gardenfors faz duas propostas:

1. Os paradigmas de representação conceitual não são explicações concorrentes de um mesmo fenômeno, mas visões com diferente nível de resolução de um mesmo fenômeno. Assim, são explicações compatíveis que devem coexistir.
2. Há um terceiro nível, intermediário entre os dois. É o nível dos **espaços conceituais**.

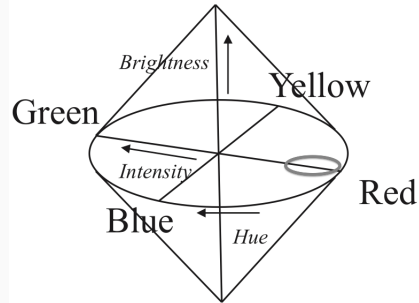
1. Espaços conceituais são geometrias/topologias sub-simbólicas que permitem classificar conceitos e propriedades.
2. Objetos são pontos nesses espaços e conceitos e propriedades são regiões.
3. Seu foco está em elucidar aspectos que resistem ao escrutínio dos outros níveis de resolução, como o **aprendizado de conceitos**, julgamento de **similaridade** (entre objetos e entre conceitos) e alguns efeitos como **protótipos**.
4. Não há um único conjunto de dimensões possível para representar um conceito, mas há indicativos de algumas características gerais do que deve ser seguido.

Gardenfors compara os três níveis de resolução para representação conceitual a três maneiras de se orientar em uma selva:

1. **Conexionismo**: as decisões tomadas são sempre locais, fruto da interação do sujeito com o ambiente que o circunda mais diretamente. Por exemplo: virar à direita, evitando um cipó mais à esquerda. Contornar um barranco etc.
2. **Espaços conceituais**: há a possibilidade de articular *instruções*. Por exemplo: “ande 30 metros para oeste, vire à direita e desça a colina”.
3. **Simbólico**: agora, também é possível dar nomes às coisas. Esses nomes têm uma carga referencial baseada em um conhecimento comum. Exemplo: “passe à direita da montanha X, então cruze o rio Y.”

Espaço das Cores

- Há mais de uma maneira de representar o espaço das cores, mas a predominante é baseada em três dimensões: (1.) **hue**, (2.) **chromaticness**, e (3.) **brightness**.
- Essas dimensões não são inteiramente independentes, por isso não formam um espaço tridimensional euclidiano, mas um **spindle**.
- Essa representação para visão humana tem fundamento fisiológico: humanos possuem três tipos de receptores cromáticos (cones).



- O essencial em um espaço conceitual não é o espaço em si, mas como ele representa o julgamento de similaridade entre objetos (exemplo: vermelho é mais similar a laranja do que a azul).

Domínio Propriedades

- Os espaços conceituais são constituídos por dimensões. Essas dimensões podem ser **integrais** - isto é, que são dependentes - ou **separáveis**.
- Uma dimensão, ou um pequeno número de dimensões integrais, forma aquilo que Gardenfors chama de **domínio**.
- **Espaço conceitual** é uma coleção de um ou mais domínios.
- **Propriedades** são regiões em um domínio.
 - Exemplo: “vermelho” é uma propriedade.

Desafios da teoria tradicional das propriedades

- Há duas teorias tradicionais (simbólicas) das propriedades: **teoria extensional** (modelos de primeira ordem) e **teoria intensional** (semântica de mundos possíveis).
- Ambas teorias enfrentam dificuldades. Citaremos duas:
 1. Propriedades naturais (com relação, entre outras coisas, ao raciocínio indutivo).
 2. Representação de protótipos.

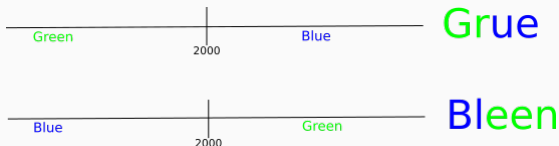
- A ideia de classes naturais está presente há muito tempo na filosofia.

Scientific disciplines frequently divide the particulars they study into kinds and theorize about those kinds. To say that a kind is natural is to say that it corresponds to a grouping that reflects the structure of the natural world rather than the interests and actions of human beings.

- Não há consenso sobre o que caracteriza uma classe natural.

Propriedades naturais

- Goodman propôs o que chamou de segunda charada da indução.



- Se vemos uma série de instâncias de esmeraldas verdes ($V(x_i) \wedge E(x_i)$), indutivamente concluímos que todas as esmeraldas são verdes ($\forall x(E(x) \rightarrow V(x))$). Mas **não** concluímos que todas as esmeraldas são *verdezul*, ainda que tenhamos observado os mesmos fenômenos $VZ(x_i) \wedge E(x_i)$. Por que?
- Uma solução: verde é uma propriedade natural; verdezul, não. E só induzimos em propriedades que julgamos ser naturais.

- Gardenfors propõe caracterizar propriedades naturais como certas regiões do espaço conceitual.
- Essas regiões devem ter os objetos mais importantes em posição central e devem ser pouco arbitrárias.
- Essas imposições se justificam, evolutivamente, por economia cognitiva.
- O critério proposto é **CP**: propriedades naturais são *regiões convexas* de um espaço natural.

- A teoria clássica dos conceitos representa pertencimento a um conceito como uma questão de “sim” ou “não”. Não faz sentido perguntar quanto um objeto pertence a um determinado conceito; nem qual, entre dois objetos, pertence mais.
- No raciocínio humano, essa distinção é feita. Isso é constatado sobretudo na teoria de protótipos, formulada originalmente por [2].
- Um tico-tico é mais tipicamente um pássaro do que uma ema.
- Protótipos dependem de conhecimento local, bem como do contexto. A cor do protótipo de limão pode ser verde em uma região do mundo e amarela em outra.

Protótipos em espaços conceituais

- A ideia de Gardenfors é que os protótipos não devem ser, necessariamente, objetos existentes, mas apenas pontos arbitrários no espaço conceitual.
- Esses pontos são pontos focais de aglomerados (*clusters*) de objetos¹.



¹Esses aglomerados podem ser determinado tanto com aprendizado supervisionado (um adulto mostrando para uma criança: “este é um cachorro”) quanto não supervisionado (algoritmos de clustering etc.).

- Os protótipos serviriam para determinar os conceitos em si através do método das **Voronoi Tessellations** - partição do espaço em regiões convexas com protótipos por centro.

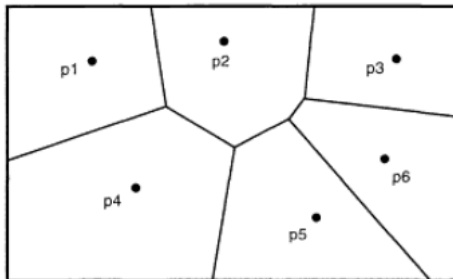


Figure 3.13

Voronoi tessellation of the plane into convex sets.

Protótipos em espaços conceituais

- Uma vantagem dessa abordagem é a economia cognitiva: não é preciso se lembrar da pertinência de cada ponto do espaço a um determinado conceito, basta ter uma coleção de protótipos e determinar a partição a partir daí.
- Também ajuda a explicar o aprendizado: é possível generalizar (formar propriedades) a partir de um número pequeno de casos concretos. As bordas (e o número de) das propriedades mudam conforme novos casos vão sendo vistos.
- Um exemplo prático das ciências cognitivas: Petito usa esse método para determinar partição de sons em consoantes com sucesso.

Conceitos

- As teorias tradicionais da representação de conceitos não os distinguem das propriedades - na lógica, ambos são predicados.
- A teoria dos espaços conceituais os diferencia da seguinte maneira: **propriedades** estão localizadas em um domínio (por exemplo, domínio das cores). Já os **conceitos** são regiões em uma série de domínios (por exemplo, uma fruta é caracterizada pela composição química, pela cor, pelo formato, pelo gosto etc.).
- Essa distinção tem base no tratamento dado pela linguagem natural: propriedades são, em geral, adjetivos (*vermelho, azedo etc.*; conceitos, substantivos (*maçã, cachorro etc.*).

Modelando conceitos

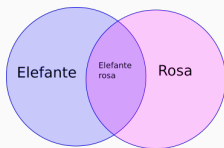
- Uma questão: quais domínios envolvidos na caracterização de um conceito?
- Varia contextualmente e dependendo do objetivo da pesquisa.
- Por exemplo: a forma pode ser determinante para um leigo determinar a espécie de um animal. Para o biólogo, a presença de determinadas estruturas é mais importante.
- Além do conjunto de dimensões que é levada em conta, também a **saliência** dessas dimensões varia com o contexto.
- Por exemplo, o peso é uma dimensão mais saliente na caracterização de um piano no contexto de uma mudança, mas não é tão saliente no contexto da classificação dos instrumentos de uma orquestra.

Critério C:

*A **natural concept** is represented as a set of 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.*

Combinando conceitos

- Na abordagem clássica, combinação de conceitos é representada pela operação de intersecção entre conjuntos.



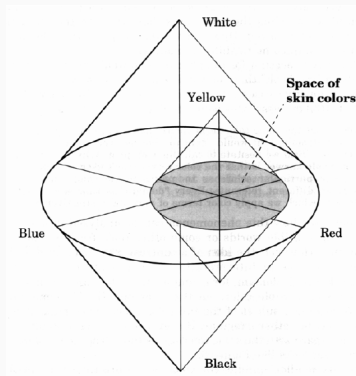
- Há diversos problemas nesse entendimento.
 1. Muitos casos em que não funciona: *stone lion* não é o conjunto de objetos que são leões e são de pedra, que é \emptyset .
 2. Nas linguagens naturais, a ordem dos conceitos na combinação é relevante: *honey bee* \neq *bee honey*.

Combinando conceitos

- Não há uma teoria que explique todas as combinações de conceitos, mas há apontamentos gerais de como fazê-lo no paradigma dos espaços conceituais.
- Há casos em que há a intersecção, e então a combinação é a restrição das dimensões em um ou mais domínios. Exemplo: maçã verde restringe o espaço cromático de maçã - que incluía vermelho e verde - para verde apenas.
- Há casos em que o novo conceito herda pedaços de cada um dos conceitos. Exemplo: *pet bird* herda o habitat doméstico de pássaro, mas não o aspecto *peludo*, que é típico para esses animais. Herda ter penas do conceito de pássaro.
- Casos em que os conceitos são incompatíveis (*stone lion*) há a sobreposição da região de um contexto sobre outro. Isso faz com que uma abordagem possível seja a da **revisão**. [4][3]
- De modo geral, essa transmissão depende da saliência das dimensões para os conceitos considerados. Regiões mais salientes têm precedência.

Classes de contraste

- O sentido de determinadas caracterizações de conceitos depende do contexto em que elas são feitas.
- Exemplo: *tall chihuahua* \nrightarrow *tall dog*.
- Esse efeito é solucionado pelas **classes de contraste**, classes que delimitam um domínio para objetos de um determinado conceito.



Concept Combination Rule: the combination CD of two concepts C and D is determined by letting the regions for the domains of C , confined to the contrast class defined by D , replace the values of the corresponding regions for D [1, p. 122]

Aspectos não monotônicos de conceitos

1. Mudança de uma categoria geral para uma subordinada.
 - Ao passar de uma categoria geral (pássaro), para uma subordinada (ema), abandonamos algumas propriedades (pequeno).
2. Efeitos contextuais.
 - O contexto torna algumas instâncias mais típicas do que outras.
3. Efeitos de classe de contraste.
 - *Red wine* \nrightarrow *Red objects*.
4. Efeitos de metáfora.
 - Imprensa marrom não é marrom.

Conclusão

- A teoria de Gardenfors estabelece ligações imediatas com a representação conceitual conforme vista sob a ótica das ciências cognitivas.
- Apesar disso, ela é computacionalmente tratável e implementável.
- Algumas abordagens “rivais” conseguem reproduzir efeitos que seriam vantagem de sua teoria. Por exemplo: há lógicas não monotônicas que reproduzem esses aspectos inferenciais.
- No quesito proposito, o livro ainda é um pouco “cru” - o que se pode fazer com essa teoria, em termos de implementação?

Três perguntas gerais:

1. Por que?
2. Para quê?
3. Como?



P. Gärdenfors.

***Conceptual spaces: The Geometry of Thought*, volume 4.**

MIT Press, Cambridge, Massachusetts, 2000.



E. Rosch and Others.

Principles of categorization.

1978.



R. Wassermann.

Revising Concepts.

(February), 2013.



R. Wassermann and E. Fermé.

A Note on Prototype Revision.

Spinning Ideas - Electronic Essays Dedicated to Peter Gardenfors on His Fiftieth Birthday, pages 1–16, 1999.

Obrigado