Inducing semantic relations from conceptual spaces

A data-driven approach to common sense reasoning

Joaquín Derrac
(joint work with Steven Schockaert)

School of Computer Science & Informatics
Cardiff University, Cardiff, UK
j.derrac@cs.cardiff.ac.uk
s.schockaert@cs.cardiff.ac.uk
Commonsense reasoning: interpolation

Undergraduate students are exempt from council tax in the UK.

Phd students are exempt from council tax in the UK.

Master’s students are exempt from council tax in the UK.

**Interpolation**: if B is conceptually between A and C, then B is likely to possess any property that is shared by A and C.
Commonsense reasoning: a fortiori inference

It is illegal for children younger than 18 to drink beer in the UK.

It is illegal for children younger than 18 to drink whiskey in the UK.

The Shawshank Redemption received a 15 certificate from the BBFC.

American Psycho received a 15 or 18 certificate from the BBFC.

**A fortiori inference**: if concept A has a property X, and concept B is more severe than A, then B has property X, or a property which is more severe than X.
Commonsense reasoning: analogical inference

Poached salmon pairs well with Chardonnay
Salmon tartare pairs well with Pinot Gris
Grilled steak pairs well with Chianti
Beef carpaccio pairs well with Dolcetto

**Analogical inference:** concepts that differ in analogical ways have properties that differ in analogical ways
Semantic relatedness

Automating these commonsense reasoning patterns requires detailed knowledge about how concepts/properties are semantically related.

**Conceptual betweenness**
master’s students are conceptually between undergraduate students and PhD students

**Relative attributes**
whiskey is stronger than beer
American psycho is more violent than the Shawshank redemption

**Analogue proportions**
Salmon tartare relates to poached salmon like beef carpaccio relates to grilled steak
Pinot gris relates to chardonnay like dolcetto relates to chianti
How to induce the required semantic relations from the web?
Semantic relatedness

**Observation 1**: all the required relations are of a qualitative nature (vs. similarity based approaches)
Semantic relatedness

**Observation 1:** all the required relations are of a qualitative nature (vs. similarity based approaches)

**Observation 2:** all the required relations can be interpreted as spatial relations in a semantic space / conceptual space / vector-space model

**Euclidean spaces** can be used to represent the **meaning** of natural language terms:

- “Semantic spaces” in computational linguistics
- “Vector-space models” in information retrieval
- “Conceptual spaces” in knowledge representation (Gärdenfors 2000)
To obtain the required semantic relations, we will instead rely on inducing conceptual spaces from text corpora. Conceptual spaces [5] are metric spaces in which the meaning of natural language concepts and properties can be represented. In most applications, conceptual spaces are assumed to be Euclidean. They are typically high-dimensional, with each dimension corresponding to a primitive cognitive feature. Specific entities then correspond to points in the conceptual space, while natural concepts and properties are posited to correspond to convex regions. Figure 1 shows a simple example of a two-dimensional conceptual space of vehicles, although it should be noted that most conceptual spaces will have a considerably higher number of dimensions. An important observation is that many types of semantic relations between vehicles correspond to qualitative spatial relations in this conceptual space. For example, the semantic is-a relationship corresponds to a spatial part-of relationship (e.g. the region for bicycle is included in the region for vehicle, because every bicycle is also a vehicle). Furthermore, we can identify conceptual betweenness with geometric betweenness (e.g. the region for motorbike is geometrically between the regions for bicycle and car, and accordingly the properties of a motorbike can be thought of as being intermediate between those of a bicycle and those of a car). Vagueness can be modelled by modelling concepts as fuzzy sets, or more simply, as nested sets of convex regions. For example, in Figure 1, elevators are considered to be borderline cases of vehicles. Finally, relative properties such as “more environmentally friendly” and “more technologically advanced” can be represented in such a space.
Semantic relatedness

Observation 1: all the required relations are of a qualitative nature (vs. similarity based approaches)

Observation 2: all the required relations can be interpreted as spatial relations in a semantic space / conceptual space / vector-space model

Spatial relations can be used to represent the semantic relatedness between natural language terms:

- conceptual betweenness
- relative attributes
- analogical proportions
- geometric betweenness
- direction relations
- parallelograms
Semantic relatedness

**Observation 1**: all the required relations are of a qualitative nature (vs. similarity based approaches)

**Observation 2**: all the required relations can be interpreted as spatial relations in a semantic space / conceptual space / vector-space model

**Main idea**: learn a semantic space from a relevant text corpus and evaluate which spatial relationships hold in that semantic space to discover fine-grained, symbolic representations of semantic relatedness in an entirely unsupervised way
Inducing a semantic space of place types from Flickr

- photos tagged with “restaurant”
- photos tagged with “tapas bar”
- photos tagged with “gothic church”

PPMI weighted tag co-occurrence vectors

Classical multi-dimensional scaling

50-dimensional Euclidean space
Direction towards less “formal” places
post-rock is conceptually between indie rock and dance
indie rock is to pop music as electro is to dance
Automating interpolative reasoning
(case study: enriching place type taxonomies)
Place type taxonomies are to some extent arbitrary

**Cyc:**
- ice cream shop is-a retail store
- retail store disjoint-from restaurant-organisation

**Tripadvisor:**
- ice cream shop is-a restaurant
- restaurant disjoint-from shopping
Place type taxonomies: categorisation

**top-level foursquare categories**

1: Arts & Entertainment  
2: College & University  
3: Food  
4: Professional & Other Places  
5: Nightlife Spots  
6: Residences  
7: Great Outdoors  
8: Shops & Services  
9: Travel & Transport

**Assign these place types to the appropriate category**

- Piano bar
- Fish Market
- Hotel
Place type taxonomies: categorisation

**Arts & Entertainment**
- Aquarium
- Arcade
- Art Gallery
- Bowling Alley
- Music Venue
- Jazz Club
- Piano Bar
- Rock Club

**Shop & Service**
- Antique Shop
- Arts & Crafts Store
- Automotive Shop
- Food & Drink Shop
- Butcher
- Cheese Shop
- Farmers Market
- Fish Market
- Food Court

**Travel & Transport**
- Airport
- Airport Food Court
- Airport Gate
- Airport Lounge
- General Travel
- Hotel
- Bed & Breakfast
Categorising places using interpolation

Can we use betweenness to find natural categories of place types?

natural category = convex region

Main idea:
- For each category C and all place types a and b known to belong to category C, determine to what degree c is between a and b in the semantic space
- Use k-NN like voting scheme to make decision

"if restaurant and pub both belong to category C, then tapas bar also belongs to category C"
Categorising places using interpolation

**Table 3.** Classification accuracy for the region-based encoding.

<table>
<thead>
<tr>
<th></th>
<th>GeoNames</th>
<th>Foursquare</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Humans (average): 0.5716</td>
<td>Humans (average): 0.7667</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 4.** Identification results (MAP) for the region-based encoding.

<table>
<thead>
<tr>
<th></th>
<th>GeoNames</th>
<th>Foursquare</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Humans (average): 0.5716</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Moreover, it implicitly encodes many alternative ways in which place types could be grouped, which is useful in applications where taxonomies are used for their predictive value. Our experimental results show that using betweenness outperforms similarity-based methods such as k-NN on a number of place type clas-
Automating a fortiori reasoning
(case study: enriching film databases)
Inducing a semantic space of films

POS tagger, chunker
(Open NLP project)

PPMI weighted term co-occurrence vectors

Classical multi-dimensional scaling

100-dimensional Euclidean space

Train SVM for each term, filter by kappa-score

interpretable dimensions
Inducing a semantic space of films

Direction towards more “violent” films
<table>
<thead>
<tr>
<th>Film</th>
<th>Adjectives</th>
<th>Nouns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Blair Witch Project</strong></td>
<td>spooky, scary, scarier, not scary, creepy, pretty creepy</td>
<td>witch, scary movies, spooky, a horror flick, a horror movie, scares</td>
</tr>
<tr>
<td><strong>The Godfather</strong></td>
<td>italian, corrupt, immoral, unsurpassed, absolutely wonderful</td>
<td>organised crime, the gangsters, the mob, gangsters, the assassination, loyalty</td>
</tr>
<tr>
<td><strong>Fight Club</strong></td>
<td>insightful, provocative, disturbing, depressed, depressing</td>
<td>conformity, society, voyeurism, our society, a dark comedy</td>
</tr>
<tr>
<td><strong>Gladiator</strong></td>
<td>epic, historically accurate, historical, lavish, magnificent</td>
<td>epics, the battle scenes, battle scenes, an epic, the epic</td>
</tr>
</tbody>
</table>
Categorising films using a fortiori inference

Use Cohen’s Kappa measure, we identify the 200 terms which most strongly correspond to a direction in the semantic space.
## Identifying salient directions: films

<table>
<thead>
<tr>
<th>Category</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horror movies</td>
<td>zombie, much gore, slashers, vampires, scary monsters, ...</td>
</tr>
<tr>
<td>Killer</td>
<td>stabbings, a psychopath, serial killer, ...</td>
</tr>
<tr>
<td>Supernatural</td>
<td>a witch, ghost stories, mysticism, a demon, the afterlife, ...</td>
</tr>
<tr>
<td>Scientist</td>
<td>experiment, the virus, radiation, the mad scientist, ...</td>
</tr>
<tr>
<td>Criminal</td>
<td>the mafia, robbers, parole, the thieves, the mastermind, ...</td>
</tr>
<tr>
<td>The animation</td>
<td>the voices, drawings, the artwork, the cartoons, anime, ...</td>
</tr>
<tr>
<td>Touching</td>
<td>inspirational, warmth, dignity, sadness, heartwarming, ...</td>
</tr>
<tr>
<td>Budget</td>
<td>a low budget film, b movies, independent films, ...</td>
</tr>
<tr>
<td>Political</td>
<td>socialism, idealism, terrorism, policy, leaders, protests, ...</td>
</tr>
<tr>
<td>Clever</td>
<td>schemes, satire, smart, witty dialogue, ingenious, ...</td>
</tr>
<tr>
<td>Bizarre</td>
<td>odd, twisted, peculiar, lunacy, surrealism, obscure, ...</td>
</tr>
<tr>
<td>Predictable</td>
<td>forgettable, unoriginal, formulaic, implausible, contrived, ...</td>
</tr>
<tr>
<td>Twists</td>
<td>unpredictable, betrayals, many twists and turns, deceit, ...</td>
</tr>
<tr>
<td>Tragic</td>
<td>anguish, sorrow, fatal, misery, bitter, heartbreaking, ...</td>
</tr>
<tr>
<td>Romantic</td>
<td>lovers, romance, the chemistry, kisses, true love, ...</td>
</tr>
<tr>
<td>Eerie</td>
<td>paranoid, spooky, impending doom, dread, ominous, ...</td>
</tr>
<tr>
<td>Scary</td>
<td>shivers, chills, creeps, frightening, the dark, goosebumps, ...</td>
</tr>
<tr>
<td>Cheesy</td>
<td>camp, corny, tacky, laughable, a guilty pleasure, ...</td>
</tr>
</tbody>
</table>
Identifying salient directions: place types

<table>
<thead>
<tr>
<th>Place Type</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>nature</td>
<td>glacial, geology, wilderness, forests, animal, erosion, naturalworld, plants, sun, peaceful ...</td>
</tr>
<tr>
<td>chicago</td>
<td>denver, queens, texas, newyork, boston, vancouver, uptown, sf, ...</td>
</tr>
<tr>
<td>cathedrale</td>
<td>romanempire, cappella, frescoes, vatican, colosseo, worldheritagelist, tourisme, loirevalley, piazza, toureiffel,</td>
</tr>
<tr>
<td>aircraft</td>
<td>flughafen, 737, controltower, boeing, landing, ...</td>
</tr>
<tr>
<td>sauna</td>
<td>accommodation, massage, chalet, wellness, piscine, hottub, jaccuzzi, luxuryhotel, ...</td>
</tr>
<tr>
<td>coal</td>
<td>electricity, railways, steel, furnace, industry, pollution, ...</td>
</tr>
<tr>
<td>bike</td>
<td>motorcycle, cyclist, scooter, ducati, lane, busstop, ...</td>
</tr>
<tr>
<td>lettuce</td>
<td>sandwich, tuna, noodles, poultry, bananas, tomatoes, lamb, sausage, steak, desserts, carrots, spinach, ...</td>
</tr>
<tr>
<td>sailingboat</td>
<td>lifeboat, fjord, motorboat, lakefront, windsurfing, kayaks, sunshinecoast, mooring, shipwreck, waterscapes, sail, houseboat, ...</td>
</tr>
<tr>
<td>finance</td>
<td>jobs, investment, canarywharf, officebuilding, skyscraper, ...</td>
</tr>
<tr>
<td>science</td>
<td>learning, career, biology, classroom, laboratory, physics, nasa, planetarium, technology, ...</td>
</tr>
</tbody>
</table>
Categorising films using a fortiori inference

Use Cohen’s Kappa measure, we identify the 200 terms which most strongly correspond to a direction in the semantic space.

Each of these terms induces a ranking of all terms, which gives us a purely symbolic representation of a semantic space. We use a variation of FOIL to learn categorisation rules from these rankings.

**if** $x$ is more violent and less funny than most horror films

**then** $x$ is a horror film
Categorising films using a fortiori inference

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Genres</th>
<th>Ratings</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Acc.</td>
<td>F1</td>
<td>Acc.</td>
</tr>
<tr>
<td>FOIL₀</td>
<td>0.922</td>
<td>0.558</td>
<td>0.836</td>
</tr>
<tr>
<td>FOIL₁₀₀</td>
<td>0.918</td>
<td>0.575</td>
<td>0.860</td>
</tr>
<tr>
<td>FOIL₅₀₀</td>
<td>0.925</td>
<td>0.581</td>
<td>0.865</td>
</tr>
<tr>
<td>FOIL₂₅₀₀</td>
<td>0.928</td>
<td>0.57</td>
<td>0.861</td>
</tr>
<tr>
<td>NN</td>
<td>0.903</td>
<td>0.507</td>
<td>0.831</td>
</tr>
<tr>
<td>C₄.₅_MDS</td>
<td>0.903</td>
<td>0.480</td>
<td>0.807</td>
</tr>
<tr>
<td>C₄.₅_dir</td>
<td>0.912</td>
<td>0.515</td>
<td>0.824</td>
</tr>
<tr>
<td>SVM_MDS</td>
<td>0.910</td>
<td>0.516</td>
<td>0.852</td>
</tr>
<tr>
<td>SVM_BoW</td>
<td>0.894</td>
<td>0.375</td>
<td>0.788</td>
</tr>
</tbody>
</table>
Conclusions

Motivation: reasoning about messy knowledge
- Requires combination of deductive and inductive reasoning
- Conclusions need to be intuitive to explain/justify

Main idea: identify semantic relatedness with spatial relations in semantic spaces
- Geometric betweenness supports interpolative reasoning
- Interpretable directions support a fortiori and analogical reasoning

Case study: semantic space of place types from Flickr
- Betweenness based classifier competitive with state of the art
- Semantic space representations sufficient to outperform humans
- Qualitative abstraction of semantic space more suitable than taxonomies for organising place types?

Case study: semantic space of films from a corpus of reviews
- Induce purely qualitative representation based on directions/rankings
- Unsupervised method for developing critique based search
- Used in variant of FOIL to outperform standard classifiers
- Easy to publish a & combine with linked data
Acknowledgment + references

Steven Schockaert
Cardiff University

Henri Prade
IRIT, Toulouse

• S. Schockaert and H. Prade. Interpolative and extrapolative reasoning in propositional theories using qualitative knowledge about conceptual spaces, Artificial Intelligence, 2013
• S. Schockaert and H. Prade. Interpolative reasoning with default rules, IJCAI, 2013
• J. Derrac, S. Schockaert. Enriching taxonomies of place types using Flickr, FOIKS, 2014
• J. Derrac, S. Schockaert. Characterising semantic relatedness using interpretable directions in conceptual spaces, ECAI, 2014