Shadows of meaning

Chung-chieh Shan
Cornell University

21 December 2011


衆瞽
探象之圖

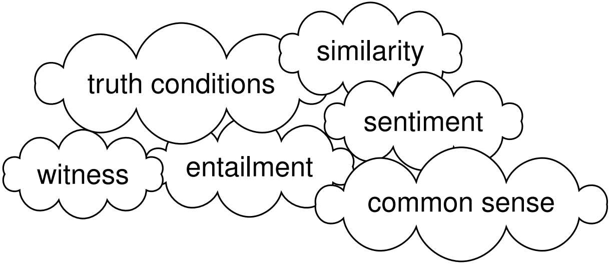
What is meaning?

Semantics with no treatment of truth conditions is not semantics.

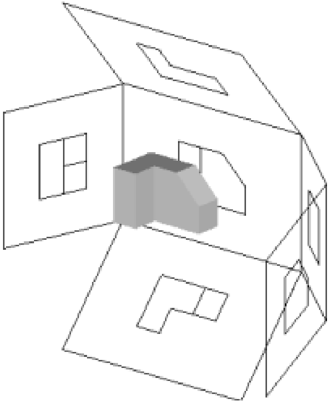
. . .

In order to say what a meaning is, we may first ask what a meaning does, and then find something that does that.

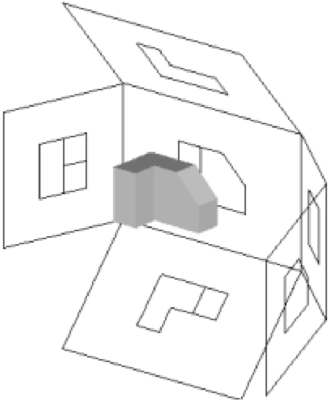
—Lewis "General semantics"



Informative shadows from random projections



Informative shadows from random projections

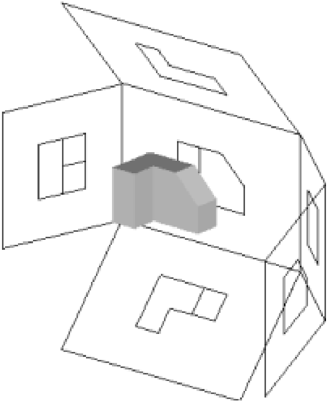


Hope in progress that “meaning” is not so polysemous:

1. Generics about kinds as topological spaces
2. Distributional semantics from language models

Reconstruct non-just-so stories from projections.

Informative shadows from random projections



Hope in progress that “meaning” is not so polysemous:

1. Generics about kinds as topological spaces
2. Distributional semantics from language models

Reconstruct non-just-so stories from projections.

Generic statements refer to kinds (Carlson)

Lions are carnivores.
Lions have four legs.
Lions have manes.
Lions give birth to live young.
Lions roar.
Lions are female. (not)

Lions are widespread.
Lions are extinct. (not)

Not universal, not existential, not proportional, not quantificational.

Generic statements refer to kinds (Carlson)

Lions are carnivores.
Lions have four legs.
Lions have manes.
Lions give birth to live young.
Lions roar.
Lions are female. (not)

Lions are widespread.
Lions are extinct. (not)

Not universal, not existential, not proportional, not quantificational.

What does it take for a kind to satisfy a property that applies to members of the kind?

Generic statements are default generalizations (Leslie)

See a few lions eat flesh, then encounter a new lion.
See a few lions give birth to live young, then encounter a new lion.
See a few lions that are male, then encounter a new lion.
See a few lions at the zoo, then encounter a new lion.

Generic statements are default generalizations (Leslie)

See a few lions eat flesh, then encounter a new lion.
See a few lions give birth to live young, then encounter a new lion.
See a few lions that are male, then encounter a new lion.
See a few lions at the zoo, then encounter a new lion.

- Tolerate negative counterexamples
  Lions are female. (not)
  Peacocks have big blue tails. (not if females had big pink tails)
- Generalize along characteristic dimensions
  Birds lay eggs.
  Bees reproduce. Bees are not sterile.
- Consider the function and purpose of artifacts
  OrangeCrusher 2000s crush oranges. (even if never used)
  Firefighters extinguish fires. (even if no fires)
- Sense disposition for disease, disaster, danger, disgust
  Mosquitoes carry the West Nile virus.
  Sharks attack bathers.

From cognition to logic

Logic is good for formalizing statements and their idealized inference patterns.

Predicate logic is good for formalizing quantified statements.

What is good for formalizing generic statements?

From cognition to logic

Logic is good for formalizing statements and their idealized inference patterns.

Predicate logic is good for formalizing quantified statements.

What is good for formalizing generic statements?

Adam Bjorndahl, Will Starr and I propose to use topology:

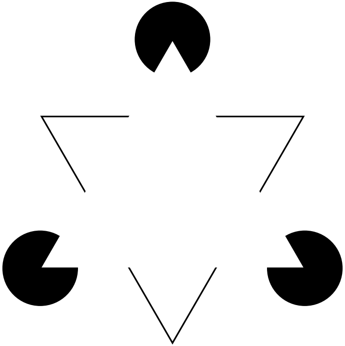
- Kinds are topological spaces
- Generic properties are large sets

One notion of large sets is what mathematicians call *generic*—

General-viewpoint assumption (Huffman)

General-viewpoint assumption (Huffman)

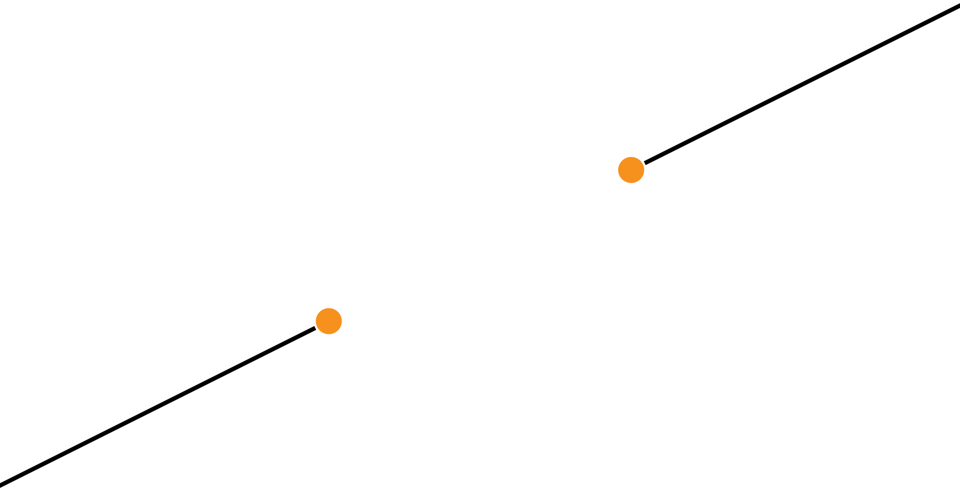
Kanizsa's subjective contours



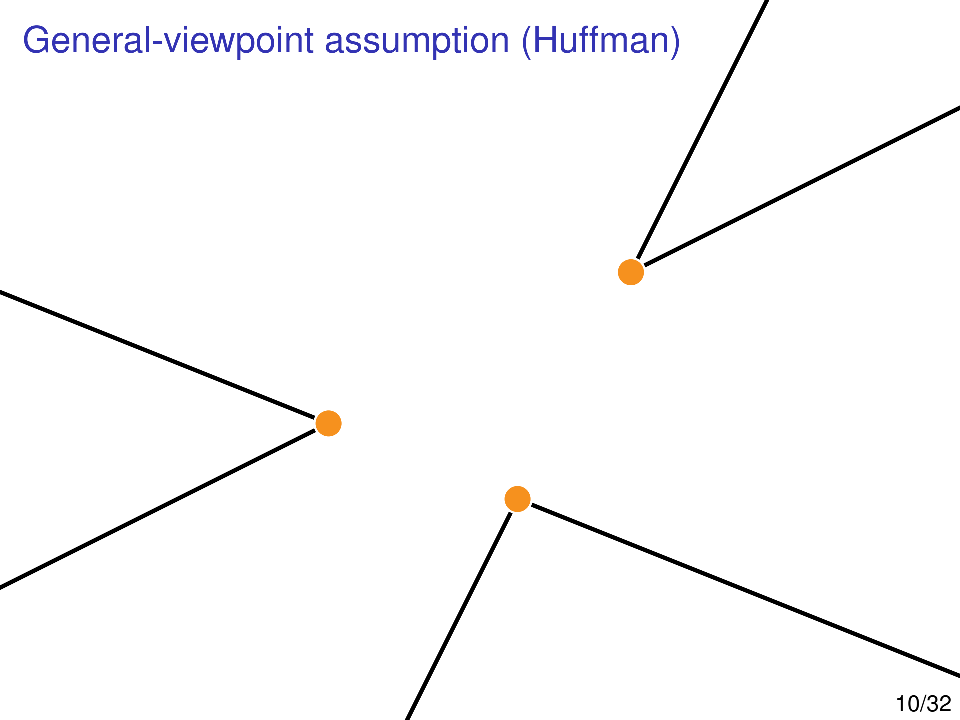
Lining up is robust, not an accident: it holds after jiggling

General-viewpoint assumption (Huffman)

General-viewpoint assumption (Huffman)

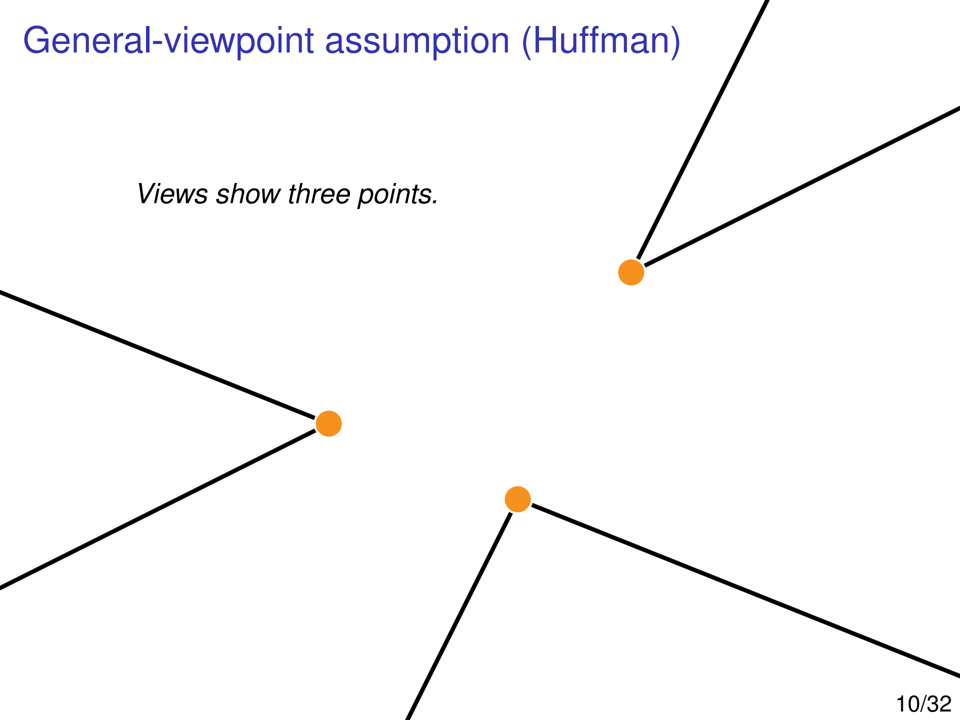


General-viewpoint assumption (Huffman)



General-viewpoint assumption (Huffman)

Views show three points.



Spaces

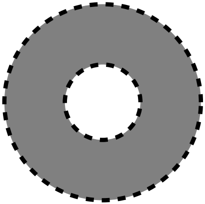
A topological space is

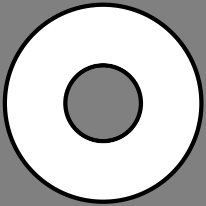
- a set X of points, and
- certain sets of points designated as open.

(Unions and finite intersections of open sets must be open.)

Example: metric spaces, for instance

- $X = \mathbb{R}^2$
- $U \subseteq X$ is open iff
  $\forall p \in U,\ \exists \epsilon > 0,\ \forall q \in X,\ d(p, q) < \epsilon \ \rightarrow\ q \in U$





Spaces

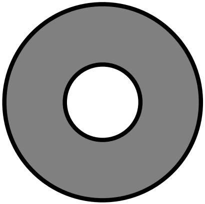
A topological space is

- a set X of points, and
- certain sets of points designated as open.

(Unions and finite intersections of open sets must be open.)

Example: metric spaces, for instance

- $X = \mathbb{R}^2$
- $U \subseteq X$ is open iff
  $\forall p \in U, \;\; \exists \epsilon > 0, \;\; \forall q \in X, \;\; d(p, q) < \epsilon \;\rightarrow\; q \in U$



Spaces

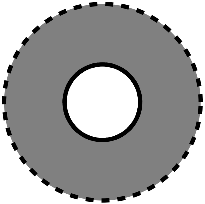
A topological space is

- a set X of points, and
- certain sets of points designated as open.

(Unions and finite intersections of open sets must be open.)

Example: metric spaces, for instance

- $X = \mathbb{R}^2$
- $U \subseteq X$ is open iff
  $\forall p \in U,\ \exists \epsilon > 0,\ \forall q \in X,\ d(p,q) < \epsilon \ \rightarrow\ q \in U$



Spaces

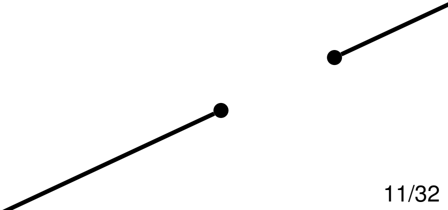
A topological space is

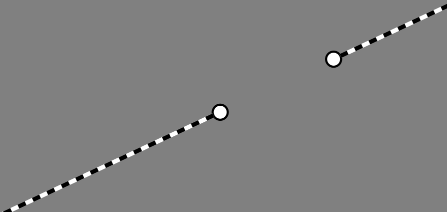
- a set X of points, and
- certain sets of points designated as open.

(Unions and finite intersections of open sets must be open.)

Example: metric spaces, for instance

- $X = \mathbb{R}^2$
- $U \subseteq X$ is open iff
  $\forall p \in U,\ \exists \epsilon > 0,\ \forall q \in X,\ d(p, q) < \epsilon \ \rightarrow\ q \in U$





Spaces

A topological space is

- a set X of points, and
- certain sets of points designated as open.

(Unions and finite intersections of open sets must be open.)

Example: metric spaces, for instance

- $X = \mathbb{R}^2$
- $U \subseteq X$ is open iff
  $\forall p \in U,\ \exists \epsilon > 0,\ \forall q \in X,\ d(p,q) < \epsilon\ \rightarrow\ q \in U$

Example: partial orders, for instance

- $X =$ strings
- $U \subseteq X$ is open iff
  $\forall p \in U,\ \forall q \in X,\ p \preceq q\ \rightarrow\ q \in U$

(Kripke models)

Spaces

A topological space is

- a set X of points, and
- certain sets of points designated as open.

(Unions and finite intersections of open sets must be open.)

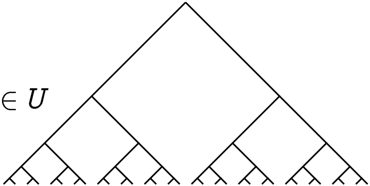
Example: metric spaces, for instance

- $X = \mathbb{R}^2$
- $U \subseteq X$ is open iff
  $\forall p \in U,\ \exists \epsilon > 0,\ \forall q \in X,\ d(p,q) < \epsilon \ \rightarrow\ q \in U$

Example: partial orders, for instance

- $X =$ strings
- $U \subseteq X$ is open iff
  $\forall p \in U,\ \forall q \in X,\ p \preceq q \ \rightarrow\ q \in U$

(Kripke models)



Spaces

A topological space is

- a set X of points, and
- certain sets of points designated as open.

(Unions and finite intersections of open sets must be open.)

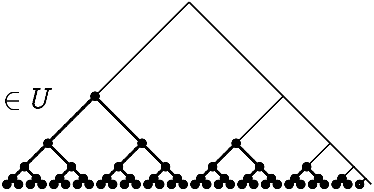
Example: metric spaces, for instance

- $X = \mathbb{R}^2$
- $U \subseteq X$ is open iff
  $\forall p \in U,\ \exists \epsilon > 0,\ \forall q \in X,\ d(p,q) < \epsilon \ \rightarrow\ q \in U$

Example: partial orders, for instance

- $X =$ strings
- $U \subseteq X$ is open iff
  $\forall p \in U,\ \forall q \in X,\ p \preceq q \ \rightarrow\ q \in U$

(Kripke models)



Operations on spaces

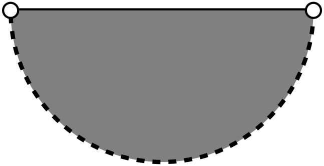
Product, sum, . . . , restriction to a subspace

Operations on spaces

Product, sum, . . . , restriction to a subspace

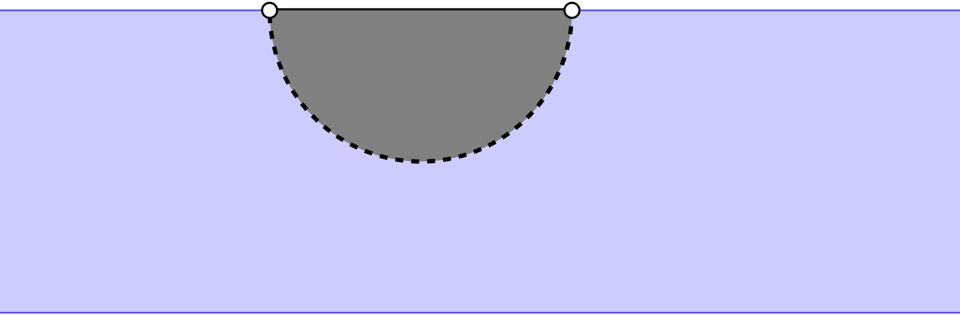
Operations on spaces

Product, sum, . . . , restriction to a subspace



Operations on spaces

Product, sum, . . . , restriction to a subspace



Kinds are topological spaces

What are the points? At least the actual lions.
Maybe also metaphysically possible lions. But not just.

Points are like situations or discourse contexts:

- Points may be situations or discourse contexts.
- Points reflect how human cognition carves up the world.
- The metaphysical status of points is unclear.
- One can worry too much about points.
  (as with viewpoints and individuals)

Kinds are topological spaces

What are the points? At least the actual lions.
Maybe also metaphysically possible lions. But not just.

Points are like situations or discourse contexts:

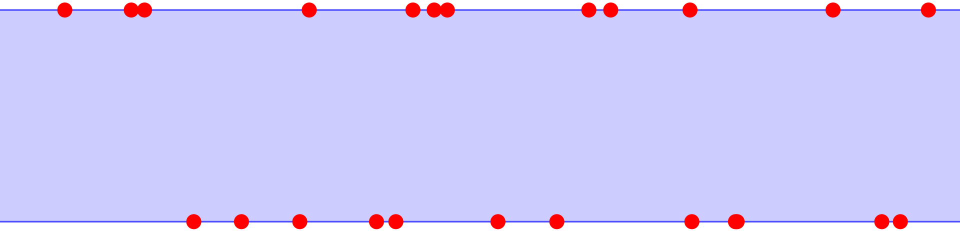
- Points may be situations or discourse contexts.
- Points reflect how human cognition carves up the world.
- The metaphysical status of points is unclear.
- One can worry too much about points.
  (as with viewpoints and individuals)

Kinds are topological spaces

What are the points? At least the actual lions.
Maybe also metaphysically possible lions. But not just.

Points are like situations or discourse contexts:

- Points may be situations or discourse contexts.
- Points reflect how human cognition carves up the world.
- The metaphysical status of points is unclear.
- One can worry too much about points.
  (as with viewpoints and individuals)



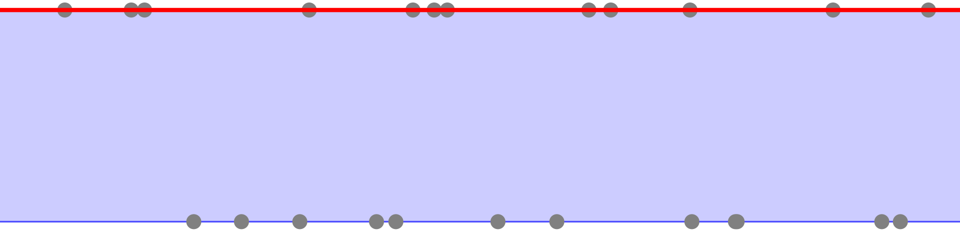
Actual lions are *discrete*.

Kinds are topological spaces

What are the points? At least the actual lions.
Maybe also metaphysically possible lions. But not just.

Points are like situations or discourse contexts:

- Points may be situations or discourse contexts.
- Points reflect how human cognition carves up the world.
- The metaphysical status of points is unclear.
- One can worry too much about points.
  (as with viewpoints and individuals)



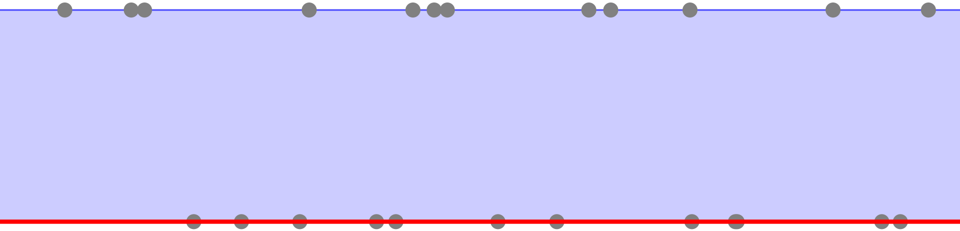
A property is a set of points.

Kinds are topological spaces

What are the points? At least the actual lions.
Maybe also metaphysically possible lions. But not just.

Points are like situations or discourse contexts:

- Points may be situations or discourse contexts.
- Points reflect how human cognition carves up the world.
- The metaphysical status of points is unclear.
- One can worry too much about points.
  (as with viewpoints and individuals)



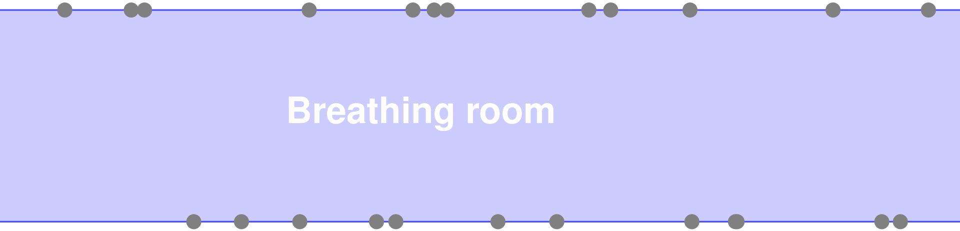
Neither-male-nor-female points are *dense*.

Kinds are topological spaces

What are the points? At least the actual lions.
Maybe also metaphysically possible lions. But not just.

Points are like situations or discourse contexts:

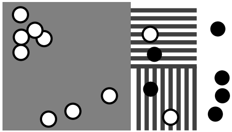
- Points may be situations or discourse contexts.
- Points reflect how human cognition carves up the world.
- The metaphysical status of points is unclear.
- One can worry too much about points.
  (as with viewpoints and individuals)



Intensionality arises from the extra points *and their topology*—

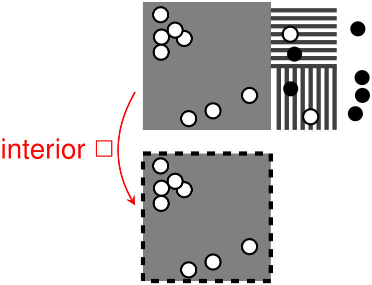
Open sets

A qualitative yet geometric notion of nearness.



Open sets

A qualitative yet geometric notion of nearness. Gives operations:

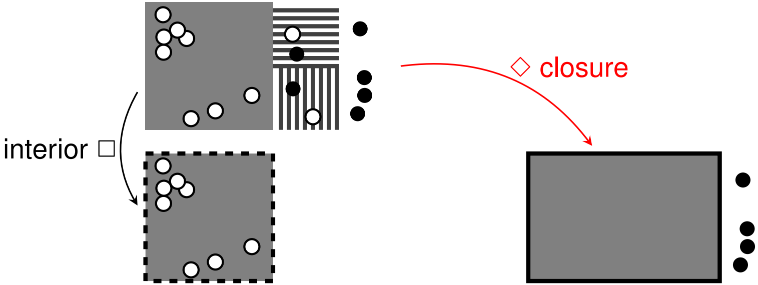


$\Box A = A \leftrightarrow A$ is open

Where A holds robustly, locally.

Open sets

A qualitative yet geometric notion of nearness. Gives operations:

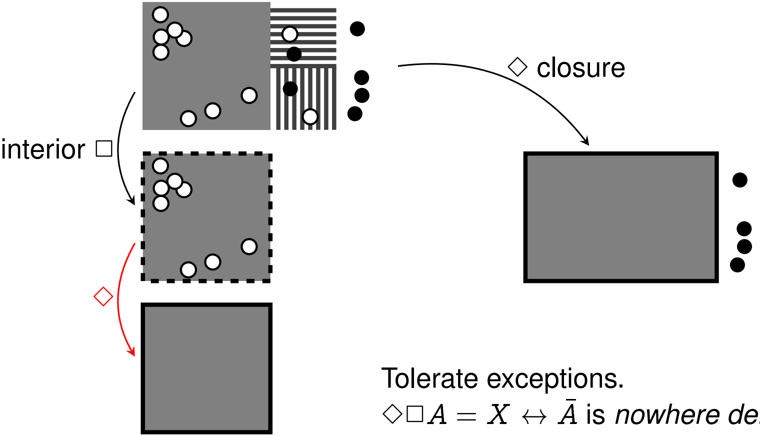


$\Diamond A = A \leftrightarrow A$ is *closed*

$\Diamond A = X \leftrightarrow A$ is *dense*

Open sets

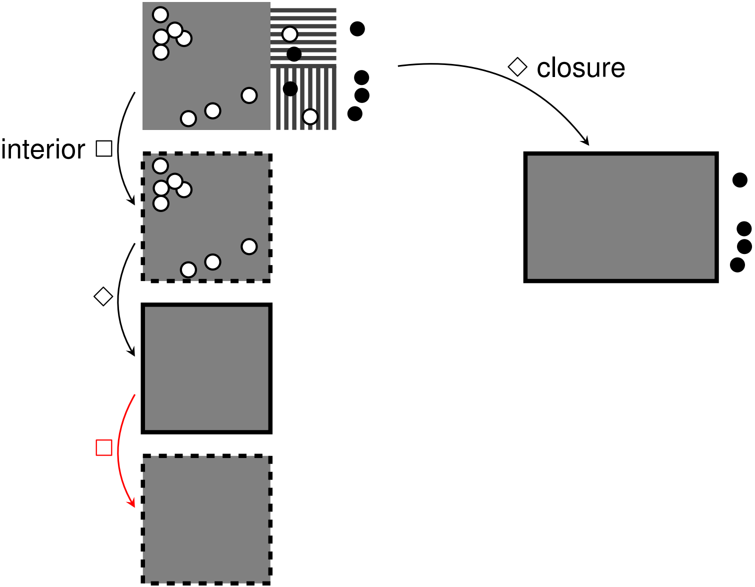
A qualitative yet geometric notion of nearness. Gives operations:



Tolerate exceptions.
$\Diamond\Box A = X \leftrightarrow \bar{A}$ is *nowhere dense*

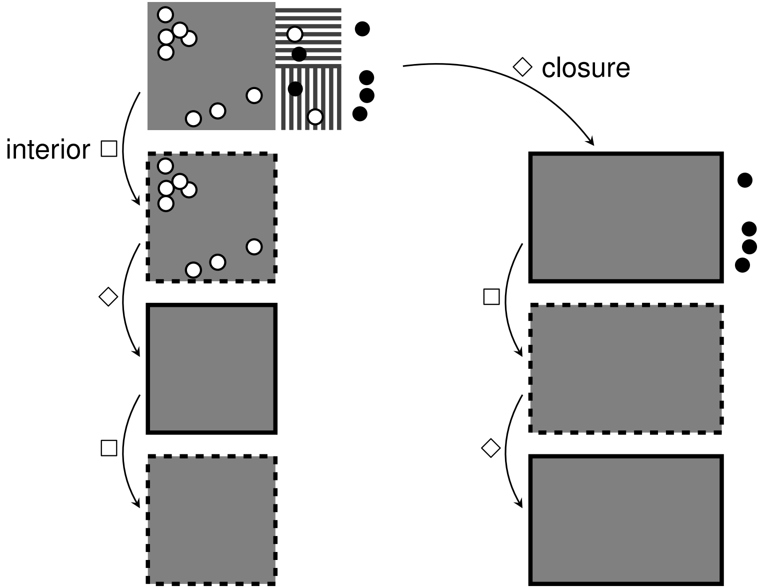
Open sets

A qualitative yet geometric notion of nearness. Gives operations:



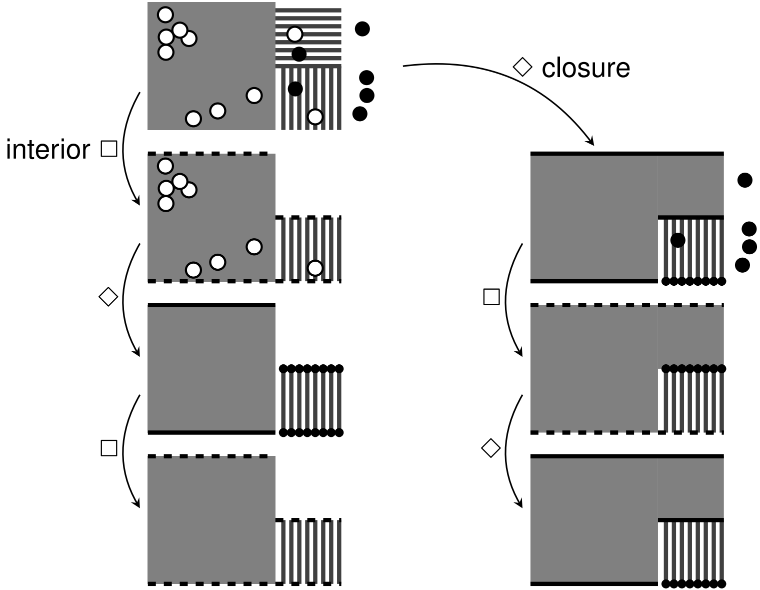
Open sets

A qualitative yet geometric notion of nearness. Gives operations:



Open sets

A qualitative yet geometric notion of nearness. Gives operations:



Open sets are sensible properties

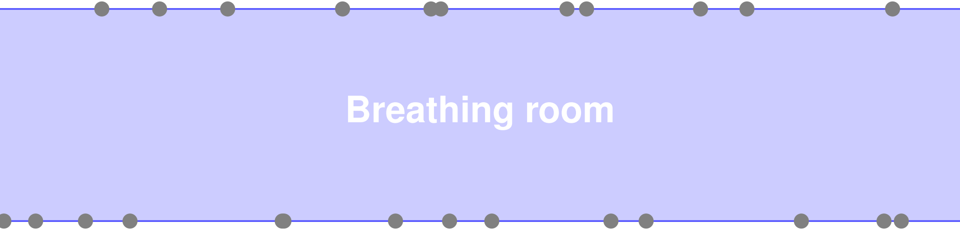
A set of points is a property. An open set is a “positive” (Leslie), “human-graspable” (Kratzer) property.

These properties are sensible grounds for similarity claims:

Lions are like tigers because they both have paws.
\# *Tigers are like snakes because they both do not have manes.*
Tweetie is like Lulu because they both are ravens/black.
\# *Fido is like Quincy because they both are not ravens/black.*

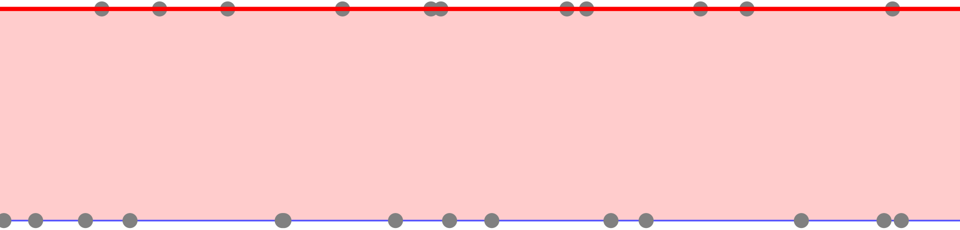
Open sets are sensible properties

A set of points is a property. An open set is a "positive" (Leslie), "human-graspable" (Kratzer) property.



Open sets are sensible properties

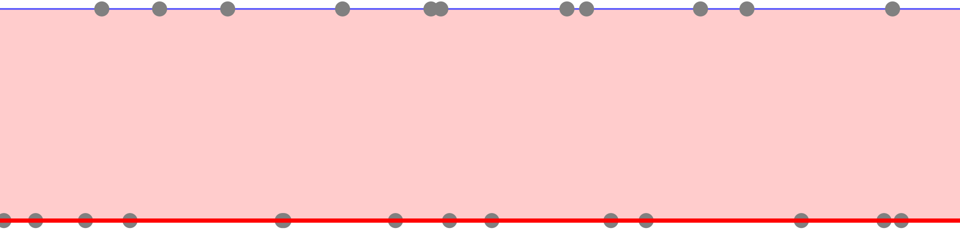
A set of points is a property. An open set is a “positive” (Leslie), “human-graspable” (Kratzer) property.



Lions have manes.

Open sets are sensible properties

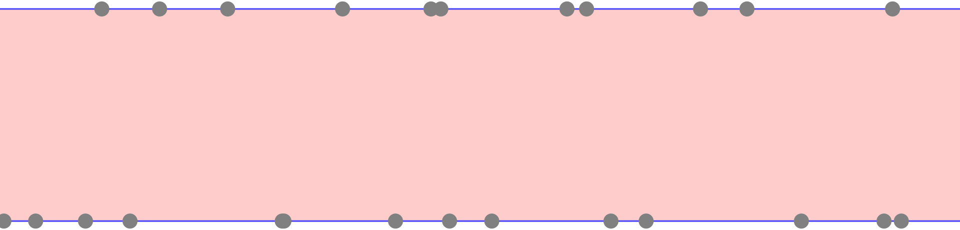
A set of points is a property. An open set is a “positive” (Leslie), “human-graspable” (Kratzer) property.



Lions give birth to live young.

Open sets are sensible properties

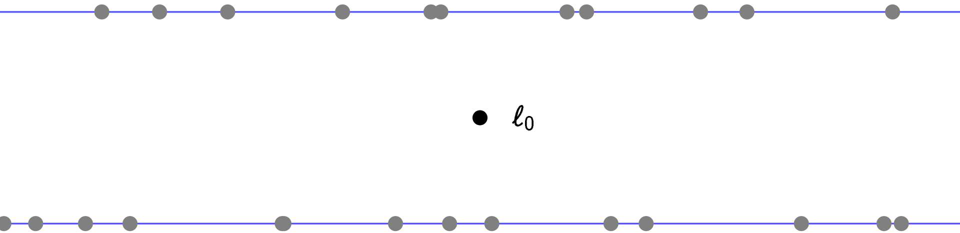
A set of points is a property. An open set is a “positive” (Leslie), “human-graspable” (Kratzer) property.



Lions have manes and give birth to live young.

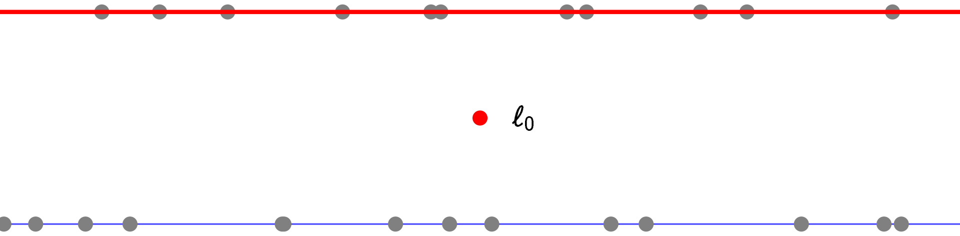
Open sets are sensible properties

A set of points is a property. An open set is a “positive” (Leslie), “human-graspable” (Kratzer) property.



Open sets are sensible properties

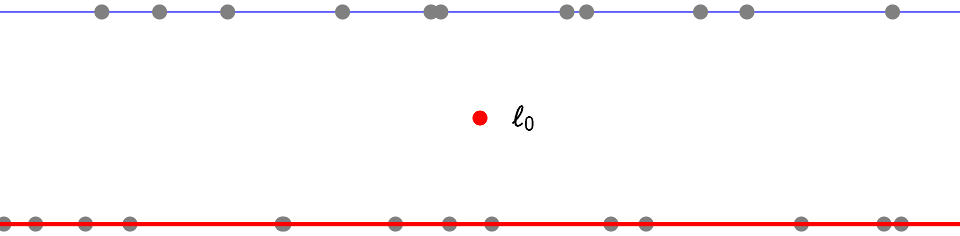
A set of points is a property. An open set is a "positive" (Leslie), "human-graspable" (Kratzer) property.



Lions have manes.

Open sets are sensible properties

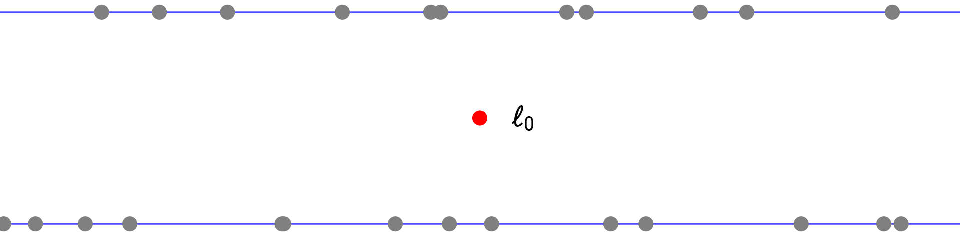
A set of points is a property. An open set is a "positive" (Leslie), "human-graspable" (Kratzer) property.



Lions give birth to live young.

Open sets are sensible properties

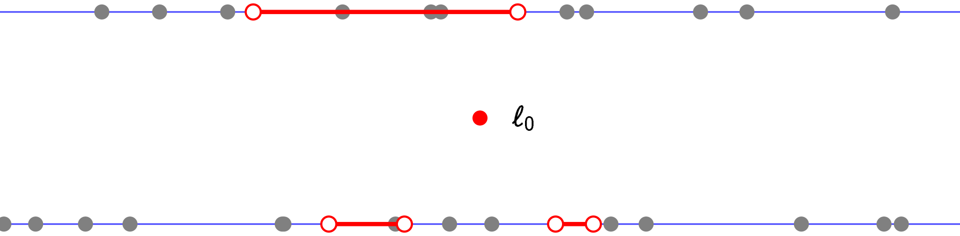
A set of points is a property. An open set is a "positive" (Leslie), "human-graspable" (Kratzer) property.



Lions have manes and give birth to live young.

Open sets are sensible properties

A set of points is a property. An open set is a "positive" (Leslie), "human-graspable" (Kratzer) property.



Put ℓ_0 in every nonempty open set.
In other words, let $\ell_0 \succeq$ every element of partial order.
This generic lion ℓ_0 may have properties that no actual lion has!

Add more points for generic statements about subkinds:
$\ell_0 \succ \ell_f \succ$ every actual female lion. Just so?

Notions of largeness

	$\bar{A}$ is nowhere dense ($X = \Diamond\Box A = \Box\Diamond\Box A$)	$\bar{A}$ is *meager* (countable union of nowhere dense)	A is not meager
$\boxdot A,\ \boxdot B \vDash \boxdot(A \cap B)$	✓	✓	
$\boxdot A \vDash \boxdot(A \cup B)$	✓	✓	
$\boxdot A_i \vDash \boxdot \bigcap_{i=1}^{\infty} A_i$	×	✓	
$\boxdot\emptyset \vDash \bot$	✓	×	
Degenerates to $\forall$ in discrete space	✓	×	
Preserved by restriction to open subspace	✓	✓	

We want to axiomatize $\boxdot$ separately from $\Box$, $\Diamond$.

Notions of largeness

	$\bar{A}$ is nowhere dense ($X = \Diamond\Box A$ $= \Box\Diamond\Box A$)	$\bar{A}$ is *meager* (countable union of nowhere dense)	A is not meager
$\boxdot A,\ \boxdot B \vDash \boxdot(A \cap B)$	✓	✓	
$\boxdot A \vDash \boxdot(A \cup B)$	✓	✓	
$\boxdot A_i \vDash \boxdot \bigcap_{i=1}^{\infty} A_i$	×	✓	
$\boxdot\emptyset \vDash \bot$	✓	×	
Degenerates to $\forall$ in discrete space	✓	×	
Preserved by restriction to open subspace	✓	✓	

We want to axiomatize $\boxdot$ separately from $\Box$, $\Diamond$.

Notions of largeness

	$\bar{A}$ is nowhere dense ($X = \Diamond\Box A$ $= \Box\Diamond\Box A$)	$\bar{A}$ is *meager* (countable union of nowhere dense)	A is not meager
$\boxdot A,\ \boxdot B \models \boxdot(A \cap B)$	✓	✓	
$\boxdot A \models \boxdot(A \cup B)$	✓	✓	
$\boxdot A_i \models \boxdot \bigcap_{i=1}^{\infty} A_i$	×	✓	
$\boxdot\emptyset \models \bot$	✓	×	
Degenerates to $\forall$ in discrete space	✓	×	
Preserved by restriction to open subspace	✓	✓	

We want to axiomatize $\boxdot$ separately from $\Box$, $\Diamond$.

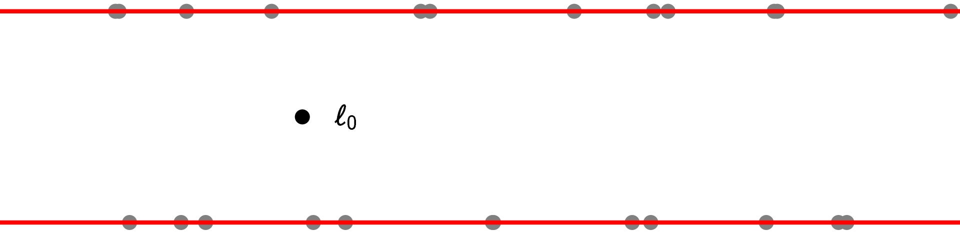
Positive alternatives and public announcement

Lions give birth to live young.
Lions are female. (not)

Positive alternatives and public announcement

Lions give birth to live young.
Lions are female. (not)

Maybe bringing up *male* or *female* restricts to the set $\mathrm{Male} \cup \mathrm{Female}$, not an open set in the original lion space. (In fact, its interior is empty.)



Generics and kinds: summary

Kinds are topological spaces.

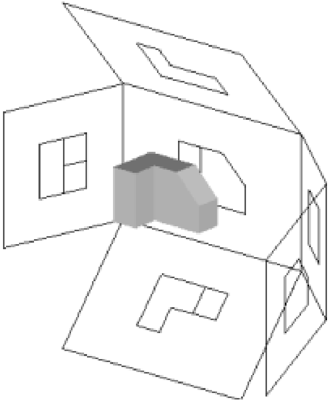
- Points are not just actual individuals.
- Open sets are sensible properties.

Generic properties are large sets.

- Inference patterns arise from $\boxbox = \Box\Diamond\Box$
  and public announcement.

> *'But it has always happened that the more I hate men individually the more I love humanity.'*
>
> *—Dostoyevski "The Brothers Karamazov"*

Informative shadows from random projections



Hope in progress that “meaning” is not so polysemous:

1. Generics about kinds as topological spaces
2. Distributional semantics from language models

Reconstruct non-just-so stories from projections.

Thanks

Bolzano

- European Masters Program in Language and Communication Technologies

Trento

- Marco Baroni
- Raffaella Bernardi
- Roberto Zamparelli

Rutgers

- Jason Perry
- Matthew Stone

Cornell

- John Hale
- Mats Rooth

Distributional semantics

For information retrieval, bag of words in each document.

$$
\begin{array}{c}
\\ \vdots \\ \\ \\ \\ \vdots
\end{array}
\begin{array}{cccc}
\dots\text{fact} & \text{factor} & \text{factory}\dots & \\
\end{array}
$$

$$
\begin{pmatrix}
 & & & \\
\dots 136 & 68 & 3 & \dots \\
\dots\ 31 & 7 & 0 & \dots \\
\dots\ 11 & 6 & 1 & \dots \\
 & & &
\end{pmatrix}
$$

Stopwords, stemming, tagging

Normalize by document and by word

Inner products for keyword/similarity retrieval (why does it work?)

Distributional semantics

For information retrieval, bag of words in each document.

$$\begin{array}{c} \\ \vdots \\ \\ \\ \\ \vdots \end{array} \begin{pmatrix} & \text{...fact} & \text{factor} & \text{factory...} & \\ \\ \ldots & 136 & 68 & 3 & \ldots \\ \ldots & 31 & 7 & 0 & \ldots \\ \ldots & 11 & 6 & 1 & \ldots \\ \\ \end{pmatrix}$$

$$= \begin{array}{c} \vdots \\ \\ \\ \\ \vdots \end{array} \begin{pmatrix} \quad \\ \\ \\ \\ \\ \end{pmatrix} \times \begin{array}{c} \text{...fact} \quad \text{factor} \quad \text{factory...} \\ \begin{pmatrix} \qquad\qquad\qquad\qquad \\ \\ \\ \end{pmatrix} \end{array}$$

Reduce dimensionality—topic models

Distributional semantics

Rows are signs (phrases). Columns are contexts (nearby words).

$$\begin{array}{rc} & \ldots\text{fact} \quad \text{factor} \quad \text{factory}\ldots \\ \begin{array}{l}\vdots\\ \text{pamphlet}\\ \text{pan}\\ \text{pancake}\\ \text{pancreas}\\ \vdots\end{array} & \left(\begin{array}{ccc} \quad & \quad & \quad \\ & & \\ & & \\ & & \\ & & \\ & & \end{array}\right)\end{array}$$

All mixed up with world knowledge and pragmatic debris.
But this shadow of meaning does do

- similarity, relevance, analogy (inner product)
- entailment: lexical (feature inclusion), quantifier (?)
- sentiment

Two views: geometric (e.g., cosine distance) and probabilistic (e.g., KL divergence). See *information geometry*.

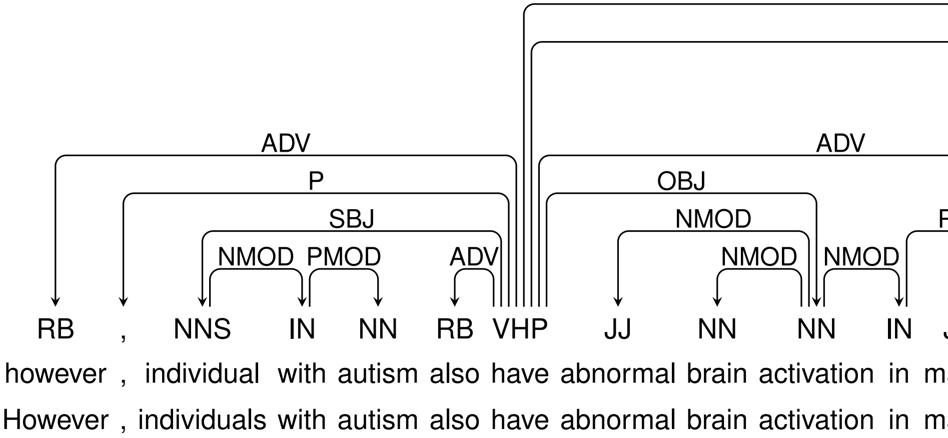
Beyond counting words

Perform different tasks without going back to the corpus?
Phrase meanings? (sparse data; compositionality)
Use syntactic structure? (word dependencies easier; sparse data)

RB	,	NNS	IN	NN	RB	VHP	JJ	NN	NN	IN	J
however	,	individual	with	autism	also	have	abnormal	brain	activation	in	m
However	,	individuals	with	autism	also	have	abnormal	brain	activation	in	m

Beyond counting words

Perform different tasks without going back to the corpus?
Phrase meanings? (sparse data; compositionality)
Use syntactic structure? (word dependencies easier; sparse data)



Beyond counting words

Rows are utterances. Columns are worlds and utterance contexts?

$$
\begin{array}{c c}
 & \begin{array}{cccc} w_1 & w_2 & w_3 & \dots \end{array} \\
\begin{array}{c} S_1 \\ S_2 \\ S_3 \\ \vdots \end{array} &
\begin{pmatrix} 1 & 0 & 1 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}
\end{array}
$$

Beyond counting words

Rows are utterances. Columns are worlds and utterance contexts? Decomposition reveals entities?

$$
\begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ \vdots \end{array}
\begin{array}{c}
\begin{array}{cccc} w_1 & w_2 & w_3 & \dots \end{array} \\
\begin{pmatrix} 1 & 0 & 1 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}
\end{array}
$$

$$
= \begin{array}{c} S_1 \\ S_2 \\ S_3 \\ \vdots \end{array}
\begin{pmatrix} \quad & \quad \\ & \\ & \\ & \end{pmatrix}
\times
\begin{array}{c}
\begin{array}{cccc} w_1 & w_2 & w_3 & \dots \end{array} \\
\begin{pmatrix} \quad & \quad & \quad & \quad \\ & & & \\ & & & \end{pmatrix}
\end{array}
$$

Beyond counting words

Rows are utterances. Columns are worlds and utterance contexts? Decomposition reveals entities?

$$\begin{array}{c} \\ S_1 \\ S_2 \\ S_3 \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccc} w_1 & w_2 & w_3 & \dots \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \end{array}$$

$$= \begin{array}{c} S_1 \\ S_2 \\ S_3 \\ \vdots \end{array} \begin{pmatrix} \quad \\ \quad \\ \quad \\ \quad \end{pmatrix} \times \begin{array}{c} \begin{array}{cccc} w_1 & w_2 & w_3 & \dots \end{array} \\ \begin{pmatrix} \quad & \quad & \quad & \quad \\ \quad & \quad & \quad & \quad \end{pmatrix} \end{array}$$

Rows are phrases. Columns include linguistic contexts?

From language model to distributional semantics

Sparse data motivates language modeling to produce virtual infinite corpus: not frequencies observed but probabilities estimated (smoothed, factored).

Let the distributional meaning of a phrase S be the probability distribution over its contexts C.

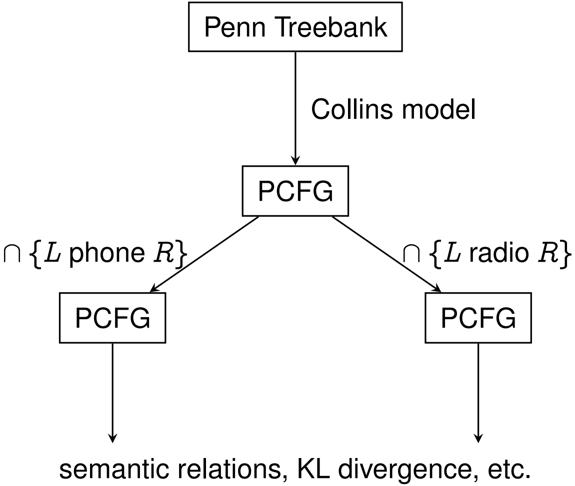
$$[\![S]\!] = \lambda C.\ \frac{\Pr(C[S])}{\sum_{C'} \Pr(C'[S])}$$

$$[\![\text{red army}]\!] = \lambda(L, R).\ \frac{\Pr(L \text{ red army } R)}{\sum_{(L',R')} \Pr(L' \text{ red army } R)}$$

$$[\![\text{red S}]\!] = \lambda(L, R).\ \frac{[\![S]\!](L \text{ red}, R)}{\sum_{(L',R')} [\![S]\!](L' \text{ red}, R')}$$

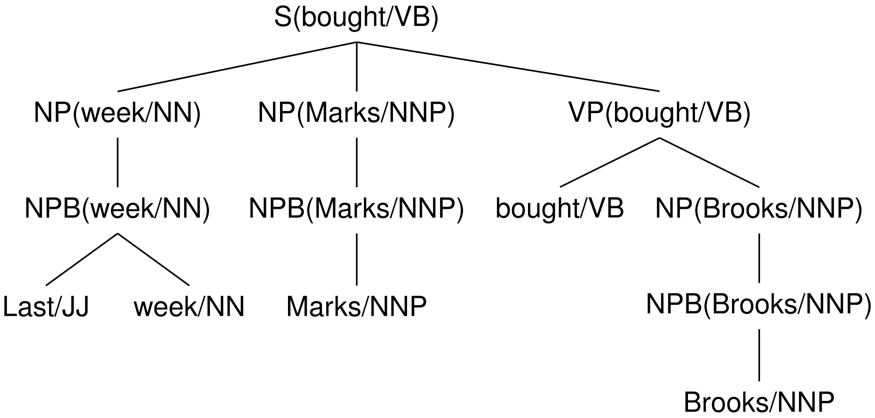
Probabilities from any model: bag of words, Markov, PCFG. . .
Pass the buck.

From Penn Treebank to distributional semantics



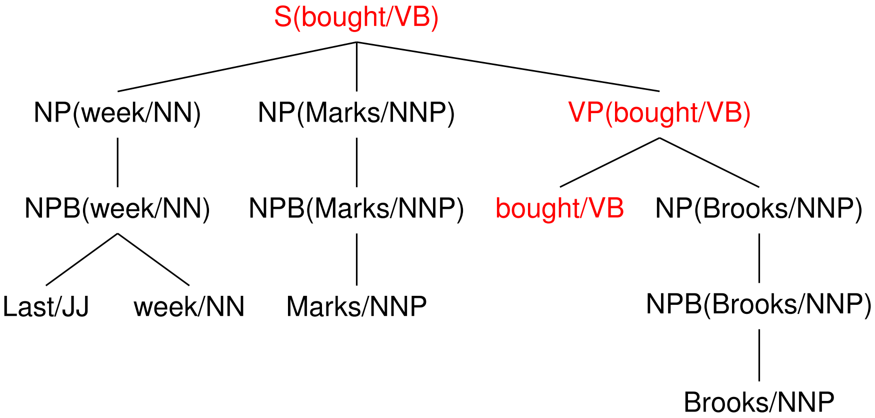
Collins model

Lexicalized PCFG for parsing (1997)
Not for generation (Post & Gildea 2008)
Bikel (2004) exegesis



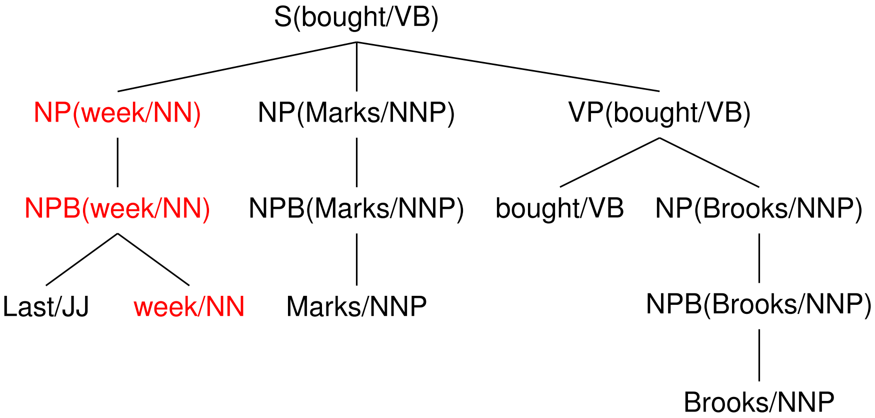
Collins model

Lexicalized PCFG for parsing (1997)
Not for generation (Post & Gildea 2008)
Bikel (2004) exegesis



Collins model

Lexicalized PCFG for parsing (1997)
Not for generation (Post & Gildea 2008)
Bikel (2004) exegesis



Summary statistics

Standard English training set: Wall Street Journal §§02–21

- 39 832 sentences
- 950 028 word tokens
  44 113 unique words
  10 437 unique words that occur 6+ times
- 28 basic nonterminal labels
  42 parts of speech

Tiny for a corpus today.

Simplified Collins Model 1

- 575 936 nonterminals
  15 564 terminals
  12 611 676 rules

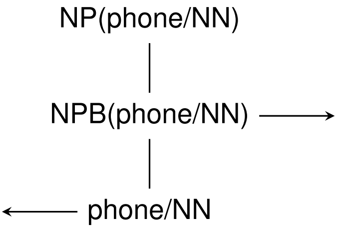
Big for a grammar today.

Pilot evaluation using BLESS data set

Concept	Relation	Relatum
phone	coord	computer
phone	coord	radio
phone	coord	stereo
phone	coord	television
phone	hyper	commodity
phone	hyper	device
phone	hyper	equipment
phone	hyper	good
phone	hyper	object
phone	hyper	system
phone	mero	cable
phone	mero	dial
phone	mero	number
phone	mero	plastic
phone	mero	wire
phone	random-n	choice
phone	random-n	clearance
phone	random-n	closing

Baroni and Lenci Evaluation of Semantic Spaces (2011)

Only head nouns observed in corpus:



Compute KL divergences among distributions over *modifier-nonterminal sequences*

Pilot evaluation using BLESS data set

Concept	Relation	Relatum
phone	coord	computer
phone	coord	radio
phone	coord	stereo
phone	coord	television
phone	hyper	commodity
phone	hyper	device
phone	hyper	equipment
phone	hyper	good
phone	hyper	object
phone	hyper	system
phone	mero	cable
phone	mero	dial
phone	mero	number
phone	mero	plastic
phone	mero	wire
phone	random-n	choice
phone	random-n	clearance
phone	random-n	closing

Baroni and Lenci Evaluation of Semantic Spaces (2011)

Only head nouns observed in corpus:

NP(phone/NN)

NPB(phone/NN) →

← phone/NN

Compute KL divergences among distributions over *modifier-nonterminal sequences*

Pilot evaluation using BLESS data set

Concept	Relation	Relatum
phone	coord	computer
phone	coord	radio
phone	coord	stereo
phone	coord	television
phone	hyper	commodity
phone	hyper	device
phone	hyper	equipment
phone	hyper	good
phone	hyper	object
phone	hyper	system
phone	mero	cable
phone	mero	dial
phone	mero	number
phone	mero	plastic
phone	mero	wire
phone	random-n	choice
phone	random-n	clearance
phone	random-n	closing

Baroni and Lenci Evaluation of Semantic Spaces (2011)

Only head nouns observed in corpus:

NP(phone/NN)

NPB(phone/NN) ⟶

⟵ phone/NN

Compute KL divergences among distributions over *modifier-nonterminal sequences*

Pilot evaluation using BLESS data set

38 Concept	Relation	687 Relatum
phone	173 coord	computer
phone	coord	radio
phone	coord	stereo
phone	coord	television
phone	125 hyper	commodity
phone	hyper	device
phone	hyper	equipment
phone	hyper	good
phone	hyper	object
phone	hyper	system
phone	490 mero	cable
phone	mero	dial
phone	mero	number
phone	mero	plastic
phone	mero	wire
phone	561 random-n	choice
phone	random-n	clearance
phone	random-n	closing

Baroni and Lenci Evaluation of Semantic Spaces (2011)

Only head nouns observed in corpus:

NP(phone/NN)

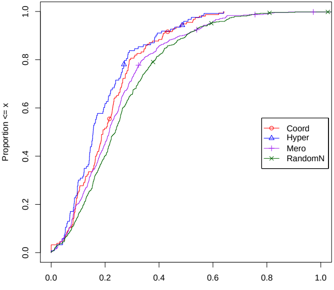
NPB(phone/NN) ⟶

⟵ phone/NN

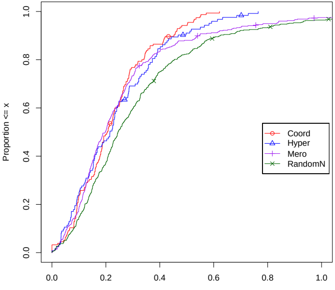
Compute KL divergences among distributions over *modifier-nonterminal sequences*

$D_{\mathrm{KL}}(\text{Concept} \,||\, \text{Relatum})$ $D_{\mathrm{KL}}(\text{Relatum} \,||\, \text{Concept})$

NP
NPB →

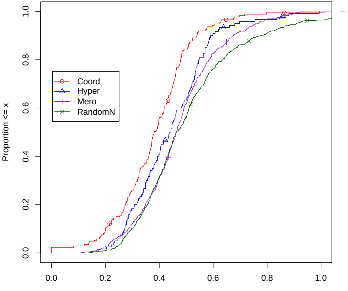


r_np_npb_KL_ by Relation (Kruskal−Wallis rank sum test p=1.73002e−05)
n:1288 m:61

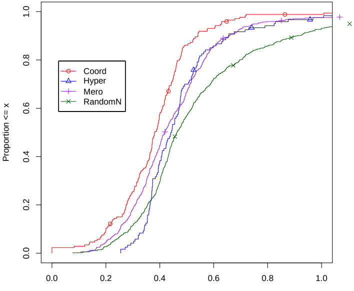


r_np_npb_LK_ by Relation (Kruskal−Wallis rank sum test p=5.88196e−06)
n:1288 m:61

NPB
← NNs



l_npb_KL by Relation (Kruskal−Wallis rank sum test p=2.54141e−13)
n:1340 m:9

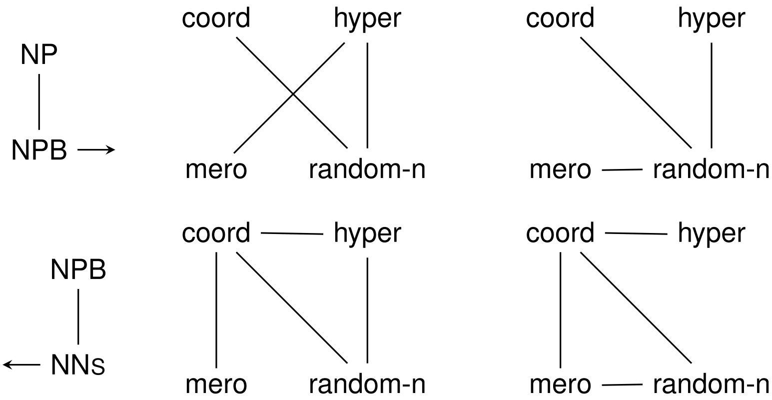


l_npb_LK by Relation (Kruskal−Wallis rank sum test p=1.0453e−15)
n:1340 m:9

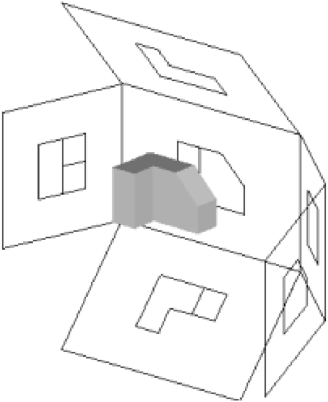
Mann-Whitney-Wilcoxon rank sum test

Edges indicate $p < .01$

	D_{KL}(Concept ‖ Relatum)	D_{KL}(Relatum ‖ Concept)



Summary



Generics about kinds as topological spaces

- System 1 reasoning *in discrete space* is System 2 reasoning

Distributional semantics from language models

- Estimate felicity *in context* from observed use