‘Going Dynamic’ in Distributional Semantics

Alessandro Lenci

Università di Pisa
Dipartimento di Filologia Letteratura e Linguistica
and
Scuola Normale Superiore

IN SUPREMÆ DIGNITATIS · 1343 ·

CoLing Lab
Computational Linguistics Laboratory

Context and Meaning in Distributional Semantics

The Distributional Hypothesis

The semantic similarity between two expressions E and E' is a function of their distribution in linguistic contexts

Distributional Representations

The distributional representation of E is a mathematical object (e.g., a vector, matrix, etc.) representing the statistical distribution E in contexts

Semantic Similarity / Relatedness

Semantic similarity (relatedness) between E and E' is measured with the similarity between their distributional representations

Context and Meaning in Distributional Semantics

The Distributional Hypothesis

The semantic similarity between two expressions E and E' is a function of their distribution in linguistic contexts

Distributional Representations

The distributional representation of E is a mathematical object (e.g., a vector, matrix, etc.) representing the statistical distribution E in contexts

Semantic Similarity / Relatedness

Semantic similarity (relatedness) between E and E' is measured with the similarity between their distributional representations

Context and Meaning in Distributional Semantics

The Distributional Hypothesis

The semantic similarity between two expressions E and E' is a function of their distribution in linguistic contexts

Distributional Representations

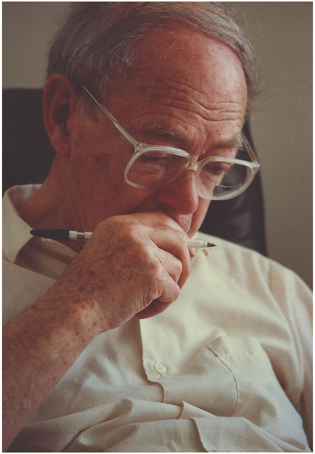
The distributional representation of E is a mathematical object (e.g., a vector, matrix, etc.) representing the statistical distribution E in contexts

Semantic Similarity / Relatedness

Semantic similarity (relatedness) between E and E' is measured with the similarity between their distributional representations

Context and Meaning in Distributional Semantics

Zellig S. Harris

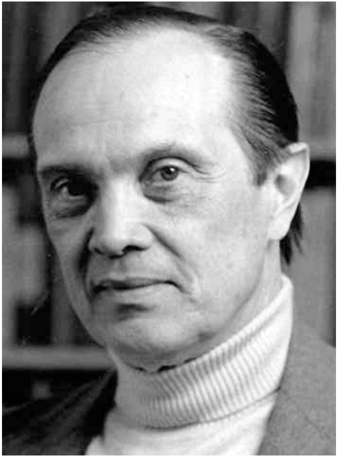


If we consider words or morphemes A and B to be more different in meaning than A and C, then we will often find that the distributions of A and B are more different than the distributions of A and C. In other words, difference in meaning correlates with difference of distribution.

(Harris 1954: 156)

Contextual Representations

George A. Miller



The contextual representation of a word is knowledge of how that word is used. [. . .] That is to say, a word's contextual representation [. . .] is an abstract cognitive structure that accumulates from encounters with the word in various (linguistic) contexts. [. . .] Two words are semantically similar to the extent that their contextual representations are similar.

(Miller and Charles 1991: 5)

Context and Meaning in Dynamic Semantics

- Dynamic Semantics (= DRT, in this talk) assumes a two-way interaction between linguistic expressions and contexts (Kamp 1981, Kamp and Reyle 1993, Van Eijck and Kamp 2011, Kamp 2013):
  - i.) the content of an expression E used in a context C depends on C
  - ii.) once this content has been determined, it leads to an update of C to a new context C' and this updated context C' helps determine the content of the next expression.

Context and Meaning in Dynamic Semantics

- Dynamic Semantics (= DRT, in this talk) assumes a two-way interaction between linguistic expressions and contexts (Kamp 1981, Kamp and Reyle 1993, Van Eijck and Kamp 2011, Kamp 2013):
    - i.) the content of an expression E used in a context C depends on C
    - ii.) once this content has been determined, it leads to an update of C to a new context C' and this updated context C' helps determine the content of the next expression.

Context and Meaning in Dynamic Semantics

- Dynamic Semantics (= DRT, in this talk) assumes a two-way interaction between linguistic expressions and contexts (Kamp 1981, Kamp and Reyle 1993, Van Eijck and Kamp 2011, Kamp 2013):
  - i.) the content of an expression E used in a context C depends on C
  - ii.) once this content has been determined, it leads to an update of C to a new context C' and this updated context C' helps determine the content of the next expression.

Context and Meaning in Dynamic Semantics

- Semantic content is a context-change potential, affecting the interpretation of following expressions (cf. also Heim 1983)

  “The slogan ‘You know the meaning of a sentence if you know the conditions under which it is true’ is replaced by this one: ‘You know the meaning of a sentence if you know the change it brings about in the information state of anyone who accepts the news conveyed by it’. Thus, meaning becomes a dynamic notion: the meaning of a sentence is an operation on information states.” (Veltman 1996)

Two Different Notions of Context Dependence

- Type (kind) level context dependence
  - the content of the type *E* depends on the linguistic contexts in which the tokens of *E* occur
- Token level context dependence
  - the content of the token *E* depends on the linguistic context in which *E* occurs

Two Different Notions of Context Dependence

- Type (kind) level context dependence
  - the content of the type *E* depends on the linguistic contexts in which the tokens of *E* occur
- Token level context dependence
  - the content of the token *E* depends on the linguistic context in which *E* occurs

Two Different Notions of Context Dependence

- Context-dependence in Distributional Semantics is at type (kind) level
  - the content of the type *dog* depends on the tokens of linguistic contexts in which *dog* occurs
    - *The dog barks*, *The dog is running fast*, *I own a brown dog*, ... $\Rightarrow \overrightarrow{dog}$
- Distributional vectors represent (part of) the conceptual content expressed by linguistic types and Distributional Semantics may be regarded as a model of semantic memory (Jones et al. 2015)
- Some attempts at modelling token-level context dependence in Distributional Semantics (Erk and Padó 2008, Mitchell and Lapata 2010, Baroni and Zamparelli 2010, Baroni et al. 2014, among many others)

Two Different Notions of Context Dependence

- Context-dependence in Distributional Semantics is at type (kind) level
    - the content of the type *dog* depends on the tokens of linguistic contexts in which *dog* occurs
        - *The dog barks*, *The dog is running fast*, *I own a brown dog*, $\ldots \Rightarrow \overrightarrow{dog}$
- Distributional vectors represent (part of) the conceptual content expressed by linguistic types and Distributional Semantics may be regarded as a model of semantic memory (Jones et al. 2015)
- Some attempts at modelling token-level context dependence in Distributional Semantics (Erk and Padó 2008, Mitchell and Lapata 2010, Baroni and Zamparelli 2010, Baroni et al. 2014, among many others)

Two Different Notions of Context Dependence

- Context-dependence in Distributional Semantics is at type (kind) level
  - the content of the type *dog* depends on the tokens of linguistic contexts in which *dog* occurs
    - *The dog barks*, *The dog is running fast*, *I own a brown dog*, … $\Rightarrow \overrightarrow{dog}$
- Distributional vectors represent (part of) the conceptual content expressed by linguistic types and Distributional Semantics may be regarded as a model of semantic memory (Jones et al. 2015)
- Some attempts at modelling token-level context dependence in Distributional Semantics (Erk and Padó 2008, Mitchell and Lapata 2010, Baroni and Zamparelli 2010, Baroni et al. 2014, among many others)

Two Different Notions of Context Dependence

- Context-dependence in Dynamic Semantics is at token level

  A man chased a dog. The dog chased another dog.

x y q z
man(x)
dog(y)
dog(q)
q=y
dog(z)
chase(q,z)

- Some attempts at adding type-level context dependence by representing constants in DRT as distributional vectors (McNally 2015, McNally and Boleda 2016)

x y
$\overrightarrow{man}$(x)
$\overrightarrow{dog}$(y)

Two Different Notions of Context Dependence

- Context-dependence in Dynamic Semantics is at token level

  A man chased a dog. The dog chased another dog.

x y q z
man(x) dog(y) dog(q) q=y dog(z) chase(q,z)

- Some attempts at adding type-level context dependence by representing constants in DRT as distributional vectors (McNally 2015, McNally and Boleda 2016)

x y
$\overrightarrow{man}$(x) $\overrightarrow{dog}$(y)

Dynamic Distributional Semantics

- ‘Going Dynamic’ in Distributional Semantics or ‘Going Distributional’ in Dynamic Semantics requires a strong integration of type-level and token-level context dependence
- This in turns requires a richer notion of context

Dynamic Distributional Semantics

- ‘Going Dynamic’ in Distributional Semantics or ‘Going Distributional’ in Dynamic Semantics requires a strong integration of type-level and token-level context dependence
- This in turns requires a richer notion of context

One Context, Many Contexts

Kamp (2016) “Entity Representations and Articulated Contexts: An Exploration of the Semantics and Pragmatics of Definite Noun Phrases”, manuscript

Articulated contexts

An Articulated Context is a 4-tuple $\langle K_{dis}, K_{enc}, K_{gen}, K_{env} \rangle$, where

i.) K_{dis} is the representation of the discourse context (with possible occurrences of indexical discourse referents to capture the contributions of the utterance context)

ii.) K_{enc} is a set of representations of “known entities”

iii.) K_{gen} is a set of representations of items of “(generic) world knowledge”

iv.) K_{env} is a set of representations of elements from the immediate (perceptual) environment

One Context, Many Contexts

Kamp (2016) "Entity Representations and Articulated Contexts: An Exploration of the Semantics and Pragmatics of Definite Noun Phrases", manuscript

Articulated contexts

An Articulated Context is a 4-tuple $\langle K_{dis}, K_{enc}, K_{gen}, K_{env} \rangle$, where

i.) K_{dis} is the representation of the discourse context (with possible occurrences of indexical discourse referents to capture the contributions of the utterance context)

ii.) K_{enc} is a set of representations of "known entities"

iii.) K_{gen} is a set of representations of items of "(generic) world knowledge"

iv.) K_{env} is a set of representations of elements from the immediate (perceptual) environment

One Context, Many Contexts

Kamp (2016) “Entity Representations and Articulated Contexts: An Exploration of the Semantics and Pragmatics of Definite Noun Phrases”, manuscript

Articulated contexts

An Articulated Context is a 4-tuple $\langle K_{dis}, K_{enc}, K_{gen}, K_{env} \rangle$, where

i.) K_{dis} is the representation of the discourse context (with possible occurrences of indexical discourse referents to capture the contributions of the utterance context)

ii.) K_{enc} is a set of representations of “known entities”

iii.) K_{gen} is a set of representations of items of “(generic) world knowledge”

iv.) K_{env} is a set of representations of elements from the immediate (perceptual) environment

One Context, Many Contexts

Kamp (2016) “Entity Representations and Articulated Contexts: An Exploration of the Semantics and Pragmatics of Definite Noun Phrases”, manuscript

Articulated contexts

An Articulated Context is a 4-tuple $\langle K_{dis}, K_{enc}, K_{gen}, K_{env} \rangle$, where

i.) K_{dis} is the representation of the discourse context (with possible occurrences of indexical discourse referents to capture the contributions of the utterance context)

ii.) K_{enc} is a set of representations of “known entities”

iii.) K_{gen} is a set of representations of items of “(generic) world knowledge”

iv.) K_{env} is a set of representations of elements from the immediate (perceptual) environment

The Structure and Content of K_{gen}

- "I will have next to nothing to say about content and form of K_{gen}. The problems presented by K_{gen} are very different from those connected with K_{enc} and K_{env}." (Kamp 2016)
- K_{gen} has a crucial role for the resolution of bridging definite descriptions (Clark 1997)

  I walked into the room. The chandelier sparkled brightly.
- "he has to know that entities like the one retrieved are always, or regularly or at least sometimes, coming in the company of entities of the kind described by α, so that he can infer with reasonable plausibility that the two of them are related in this way." (ibid.)
- "K_{gen} has been described as a collection of propositions that express general connections between things, states and events within our world" (ibid., fn. 73)

The Structure and Content of K_{gen}

- "I will have next to nothing to say about content and form of K_{gen}. The problems presented by K_{gen} are very different from those connected with K_{enc} and K_{env}." (Kamp 2016)
- K_{gen} has a crucial role for the resolution of bridging definite descriptions (Clark 1997)

  I walked into the room. The chandelier sparkled brightly.

- "he has to know that entities like the one retrieved are always, or regularly or at least sometimes, coming in the company of entities of the kind described by α, so that he can infer with reasonable plausibility that the two of them are related in this way." (ibid.)
- "K_{gen} has been described as a collection of propositions that express general connections between things, states and events within our world" (ibid., fn. 73)

The Structure and Content of K_{gen}

- "I will have next to nothing to say about content and form of K_{gen}. The problems presented by K_{gen} are very different from those connected with K_{enc} and K_{env}." (Kamp 2016)
- K_{gen} has a crucial role for the resolution of bridging definite descriptions (Clark 1997)

  I walked into the room. The chandelier sparkled brightly.
- "he has to know that entities like the one retrieved are always, or regularly or at least sometimes, coming in the company of entities of the kind described by α, so that he can infer with reasonable plausibility that the two of them are related in this way." (ibid.)
- "K_{gen} has been described as a collection of propositions that express general connections between things, states and events within our world" (ibid., fn. 73)

The Structure and Content of K_{gen}

- "I will have next to nothing to say about content and form of K_{gen}. The problems presented by K_{gen} are very different from those connected with K_{enc} and K_{env}." (Kamp 2016)
- K_{gen} has a crucial role for the resolution of bridging definite descriptions (Clark 1997)

  I walked into the room. The chandelier sparkled brightly.
- "he has to know that entities like the one retrieved are always, or regularly or at least sometimes, coming in the company of entities of the kind described by α, so that he can infer with reasonable plausibility that the two of them are related in this way." (ibid.)
- "K_{gen} has been described as a collection of propositions that express general connections between things, states and events within our world" (ibid., fn. 73)

Distributional Semantics Meet DRT

General assumptions

- Language comprehension consists in incrementally building a discourse semantic representation (DRS) from the linguistic input
- DRSs are models for mental representations (Hamm et al. 2006, Kamp 2016)
- Language is a set of instructions used to create a mental representation of an event or situation that is described by linguistic forms (Zwaan and Radvansky 1998)
- The goal of the comprehender is to identify the event or situation the speakers wants to convey, and this is the event that best explains the linguistic cues used in the sentence (Kuperberg 2016)
- Language comprehension always occurs in an Articulated Context
- K_{gen} contains (distributional) information about events and their participants that is activated by linguistic cues
- The distributional content of a linguistic expression can be viewed as a context change potential that updates K_{disc} with information activated from K_{gen}, and from the other components of the Articulated Context

Distributional Semantics Meet DRT

General assumptions

- Language comprehension consists in incrementally building a discourse semantic representation (DRS) from the linguistic input
- DRSs are models for mental representations (Hamm et al. 2006, Kamp 2016)
- Language is a set of instructions used to create a mental representation of an event or situation that is described by linguistic forms (Zwaan and Radvansky 1998)
- The goal of the comprehender is to identify the event or situation the speakers wants to convey, and this is the event that best explains the linguistic cues used in the sentence (Kuperberg 2016)
- Language comprehension always occurs in an Articulated Context
- K_{gen} contains (distributional) information about events and their participants that is activated by linguistic cues
- The distributional content of a linguistic expression can be viewed as a context change potential that updates K_{disc} with information activated from K_{gen}, and from the other components of the Articulated Context

Distributional Semantics Meet DRT

General assumptions

- Language comprehension consists in incrementally building a discourse semantic representation (DRS) from the linguistic input
- DRSs are models for mental representations (Hamm et al. 2006, Kamp 2016)
- Language is a set of instructions used to create a mental representation of an event or situation that is described by linguistic forms (Zwaan and Radvansky 1998)
- The goal of the comprehender is to identify the event or situation the speakers wants to convey, and this is the event that best explains the linguistic cues used in the sentence (Kuperberg 2016)
- Language comprehension always occurs in an Articulated Context
- K_{gen} contains (distributional) information about events and their participants that is activated by linguistic cues
- The distributional content of a linguistic expression can be viewed as a context change potential that updates K_{disc} with information activated from K_{gen}, and from the other components of the Articulated Context

Distributional Semantics Meet DRT

General assumptions

- Language comprehension consists in incrementally building a discourse semantic representation (DRS) from the linguistic input
- DRSs are models for mental representations (Hamm et al. 2006, Kamp 2016)
- Language is a set of instructions used to create a mental representation of an event or situation that is described by linguistic forms (Zwaan and Radvansky 1998)
- The goal of the comprehender is to identify the event or situation the speakers wants to convey, and this is the event that best explains the linguistic cues used in the sentence (Kuperberg 2016)
- Language comprehension always occurs in an Articulated Context
- K_{gen} contains (distributional) information about events and their participants that is activated by linguistic cues
- The distributional content of a linguistic expression can be viewed as a context change potential that updates K_{disc} with information activated from K_{gen}, and from the other components of the Articulated Context

Distributional Semantics Meet DRT

General assumptions

- Language comprehension consists in incrementally building a discourse semantic representation (DRS) from the linguistic input
- DRSs are models for mental representations (Hamm et al. 2006, Kamp 2016)
- Language is a set of instructions used to create a mental representation of an event or situation that is described by linguistic forms (Zwaan and Radvansky 1998)
- The goal of the comprehender is to identify the event or situation the speakers wants to convey, and this is the event that best explains the linguistic cues used in the sentence (Kuperberg 2016)
- Language comprehension always occurs in an Articulated Context
- K_{gen} contains (distributional) information about events and their participants that is activated by linguistic cues
- The distributional content of a linguistic expression can be viewed as a context change potential that updates K_{disc} with information activated from K_{gen}, and from the other components of the Articulated Context

Distributional Semantics Meet DRT

General assumptions

- Language comprehension consists in incrementally building a discourse semantic representation (DRS) from the linguistic input
- DRSs are models for mental representations (Hamm et al. 2006, Kamp 2016)
- Language is a set of instructions used to create a mental representation of an event or situation that is described by linguistic forms (Zwaan and Radvansky 1998)
- The goal of the comprehender is to identify the event or situation the speakers wants to convey, and this is the event that best explains the linguistic cues used in the sentence (Kuperberg 2016)
- Language comprehension always occurs in an Articulated Context
- K_{gen} contains (distributional) information about events and their participants that is activated by linguistic cues
- The distributional content of a linguistic expression can be viewed as a context change potential that updates K_{disc} with information activated from K_{gen}, and from the other components of the Articulated Context

Distributional Semantics Meet DRT

General assumptions

- Language comprehension consists in incrementally building a discourse semantic representation (DRS) from the linguistic input
- DRSs are models for mental representations (Hamm et al. 2006, Kamp 2016)
- Language is a set of instructions used to create a mental representation of an event or situation that is described by linguistic forms (Zwaan and Radvansky 1998)
- The goal of the comprehender is to identify the event or situation the speakers wants to convey, and this is the event that best explains the linguistic cues used in the sentence (Kuperberg 2016)
- Language comprehension always occurs in an Articulated Context
- K_{gen} contains (distributional) information about events and their participants that is activated by linguistic cues
- The distributional content of a linguistic expression can be viewed as a context change potential that updates K_{disc} with information activated from K_{gen}, and from the other components of the Articulated Context

K_{gen} as Generalized Event Knowledge (GEK)

- K_{gen} stores generalized knowledge about events and their participants
- GEK derives from first-hand experience and from linguistic experience (e.g., from linguistic descriptions of events)
- Language provides multiple cues that can be used to focus and activate various aspects of events, participants, locations, etc. (McRae and Matsuki 2009: 1419)

  "the specific choice of verb can be used to bring to mind somewhat different scenarios, such as *eating* versus *dining*. In terms of the possible entities that participate in such events, knowing that a *waitress* is involved, for example, invokes a certain type of eating event. The phrase *hamburgers and hot dogs* produces a different type of scenario than does *turkey and stuffing*, including perhaps information about location and time of year. Instrument nouns can cue certain types of eating, as in *eating with a fork* versus *eating with a stick*. Finally, event nouns like *breakfast* or location nouns like *cafeteria* cue specific types of eating scenarios."

K_{gen} as Generalized Event Knowledge (GEK)

- K_{gen} stores generalized knowledge about events and their participants
- GEK derives from first-hand experience and from linguistic experience (e.g., from linguistic descriptions of events)
- Language provides multiple cues that can be used to focus and activate various aspects of events, participants, locations, etc. (McRae and Matsuki 2009: 1419)

  "the specific choice of verb can be used to bring to mind somewhat different scenarios, such as *eating* versus *dining*. In terms of the possible entities that participate in such events, knowing that a *waitress* is involved, for example, invokes a certain type of eating event. The phrase *hamburgers and hot dogs* produces a different type of scenario than does *turkey and stuffing*, including perhaps information about location and time of year. Instrument nouns can cue certain types of eating, as in *eating with a fork* versus *eating with a stick*. Finally, event nouns like *breakfast* or location nouns like *cafeteria* cue specific types of eating scenarios."

K_{gen} as Generalized Event Knowledge (GEK)

- K_{gen} stores generalized knowledge about events and their participants
- GEK derives from first-hand experience and from linguistic experience (e.g., from linguistic descriptions of events)
- Language provides multiple cues that can be used to focus and activate various aspects of events, participants, locations, etc. (McRae and Matsuki 2009: 1419)

  "the specific choice of verb can be used to bring to mind somewhat different scenarios, such as *eating* versus *dining*. In terms of the possible entities that participate in such events, knowing that a *waitress* is involved, for example, invokes a certain type of eating event. The phrase *hamburgers and hot dogs* produces a different type of scenario than does *turkey and stuffing*, including perhaps information about location and time of year. Instrument nouns can cue certain types of eating, as in *eating with a fork* versus *eating with a stick*. Finally, event nouns like *breakfast* or location nouns like *cafeteria* cue specific types of eating scenarios."

The Elements of GEK

- Both dynamic and static situations or eventualities are included in GEK (Vendler 1967, Dowty 1979, Rothstein 2004)
  - e.g., the information that student read books and that books have pages are both parts of GEK
- GEK is highly structured, and organized under various levels of complexity, granularity, and schematicity
  - fully-specified micro-events
    - e.g., students read books, surfers surf in the sea, etc.
  - schematic events with entities that co-occur in the same situation, abstracting away from the specific events linking them
    - e.g, surfers, boards, waves, and wax tend to co-occur in the situations
  - complex scenarios, much like scripts, frames or narrative schemas, which include various sub-events and complex temporal and causal relations about them
    - the surfing scenario includes events such as bringing a surf board, diving in the sea, swimming, etc.

GEK and Distributional Semantics

- We refer with GEK_{DS} to the distributional subset of GEK that can be derived from co-occurrences in the linguistic input
- Events in GEK_{DS} contain information directly extracted from parsed sentences in corpora
- We represent events in GEK_{DS} with attribute-value matrices (AVM) specifying their participants and roles
  - attributes are syntactic dependencies (e.g. SUBJ, COMP–IN, etc.), as a surface approximation of deeper semantic roles
  - values are distributional vectors of dependent lexemes
    - "out-of-context" distributional vector encodings of lexical items

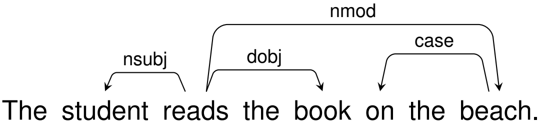
GEK and Distributional Semantics

- We refer with GEK_{DS} to the distributional subset of GEK that can be derived from co-occurrences in the linguistic input
- Events in GEK_{DS} contain information directly extracted from parsed sentences in corpora
- We represent events in GEK_{DS} with attribute-value matrices (AVM) specifying their participants and roles
  - attributes are syntactic dependencies (e.g. SUBJ, COMP−IN, etc.), as a surface approximation of deeper semantic roles
  - values are distributional vectors of dependent lexemes
    - “out-of-context” distributional vector encodings of lexical items

GEK and Distributional Semantics

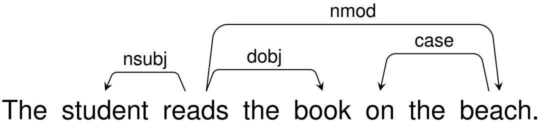
- We refer with *GEK_{DS}* to the distributional subset of GEK that can be derived from co-occurrences in the linguistic input
- Events in *GEK_{DS}* contain information directly extracted from parsed sentences in corpora
- We represent events in *GEK_{DS}* with attribute-value matrices (AVM) specifying their participants and roles
    - attributes are syntactic dependencies (e.g. SUBJ, COMP–IN, etc.), as a surface approximation of deeper semantic roles
    - values are distributional vectors of dependent lexemes
        - “out-of-context” distributional vector encodings of lexical items

GEK and Distributional Semantics



EVENT	
NSUBJ	$\overrightarrow{student}$
HEAD	$\overrightarrow{read}$
DOBJ	$\overrightarrow{book}$
NMOD-ON	$\overrightarrow{beach}$

GEK and Distributional Semantics



EVENT	
NSUBJ	$\overrightarrow{student}$
HEAD	$\overrightarrow{read}$
DOBJ	$\overrightarrow{book}$
NMOD-ON	$\overrightarrow{beach}$

Events in GEK are Hierarchically Structured

EVENT	
NSUBJ	$\overrightarrow{student}$
DOBJ	$\overrightarrow{book}$

EVENT	
NSUBJ	$\overrightarrow{student}$
HEAD	$\overrightarrow{read}$
DOBJ	$\overrightarrow{book}$

EVENT	
NSUBJ	$\overrightarrow{student}$
HEAD	$\overrightarrow{buy}$
DOBJ	$\overrightarrow{book}$

EVENT	
NSUBJ	$\overrightarrow{student}$
HEAD	$\overrightarrow{write}$
DOBJ	$\overrightarrow{book}$

Lexical Items as Cues to GEK in K_{gen}

- The lexicon is a repository of constructions (i.e., words and other more schematic elements) stored in long-term memory
- Constructions cue (i.e. activate) portions of GEK in K_{gen}
- Each construction *Cxn* is defined by a FORM and a content (SEM), represented with AVMs as in Sign-Based Construction Grammar (Sag 2012, Michaelis 2013)
- SEM is formed by two types of information:
  - a set of events stored in the GEK in K_{gen} and activated by the construction
  - a set of semantic neighbors (NEI) of the construction

$$\begin{bmatrix} \text{FORM} & \textit{student} \\ \text{SEM} & \begin{bmatrix} \text{GEK} & \langle e_1, \sigma_1 \rangle, \ldots, \langle e_n, \sigma_n \rangle \\ \text{NEI} & \langle n_1, s_1 \rangle, \ldots, \langle n_n, s_n \rangle \end{bmatrix} \end{bmatrix}$$

Lexical Items as Cues to GEK in K_{gen}

- The lexicon is a repository of constructions (i.e., words and other more schematic elements) stored in long-term memory
- Constructions cue (i.e. activate) portions of GEK in K_{gen}
- Each construction *Cxn* is defined by a FORM and a content (SEM), represented with AVMs as in Sign-Based Construction Grammar (Sag 2012, Michaelis 2013)
- SEM is formed by two types of information:
  - a set of events stored in the GEK in K_{gen} and activated by the construction
  - a set of semantic neighbors (NEI) of the construction

$$\begin{bmatrix} \text{FORM} & \textit{student} \\ \text{SEM} & \begin{bmatrix} \text{GEK} & \langle e_1, \sigma_1 \rangle, \ldots, \langle e_n, \sigma_n \rangle \\ \text{NEI} & \langle n_1, s_1 \rangle, \ldots, \langle n_n, s_n \rangle \end{bmatrix} \end{bmatrix}$$

Lexical Items as Cues to GEK in K_{gen}

- The lexicon is a repository of constructions (i.e., words and other more schematic elements) stored in long-term memory
- Constructions cue (i.e. activate) portions of GEK in K_{gen}
- Each construction *Cxn* is defined by a FORM and a content (SEM), represented with AVMs as in Sign-Based Construction Grammar (Sag 2012, Michaelis 2013)
- SEM is formed by two types of information:
  - a set of events stored in the GEK in K_{gen} and activated by the construction
  - a set of semantic neighbors (NEI) of the construction

$$\begin{bmatrix} \text{FORM} & \textit{student} \\ \text{SEM} & \begin{bmatrix} \text{GEK} & \langle e_1, \sigma_1 \rangle, \ldots, \langle e_n, \sigma_n \rangle \\ \text{NEI} & \langle n_1, s_1 \rangle, \ldots, \langle n_n, s_n \rangle \end{bmatrix} \end{bmatrix}$$

Lexical Items as Cues to GEK in K_{gen}

- The lexicon is a repository of constructions (i.e., words and other more schematic elements) stored in long-term memory
- Constructions cue (i.e. activate) portions of GEK in K_{gen}
- Each construction *Cxn* is defined by a FORM and a content (SEM), represented with AVMs as in Sign-Based Construction Grammar (Sag 2012, Michaelis 2013)
- SEM is formed by two types of information:
  - a set of events stored in the GEK in K_{gen} and activated by the construction
  - a set of semantic neighbors (NEI) of the construction

$$\begin{bmatrix} \text{FORM} & \textit{student} \\ \text{SEM} & \begin{bmatrix} \text{GEK} & \langle e_1, \sigma_1 \rangle, \ldots, \langle e_n, \sigma_n \rangle \\ \text{NEI} & \langle n_1, s_1 \rangle, \ldots, \langle n_n, s_n \rangle \end{bmatrix} \end{bmatrix}$$

Lexical Items as Cues to GEK in K_{gen}

- We model K_{gen} as a set of pairs $\langle e, \sigma \rangle$, such that:
  - e is an event stored in GEK_{DS}
  - σ is a score expressing the salience of the event with respect to the construction that cues it (e.g., $P(e|Cxn)$)
- Each event in GEK_{DS} may be cued by several lexical items, as part of their semantic content

$$\left[read\ \boxed{1} \left\langle \begin{bmatrix} \textbf{EVENT} & \\ \text{NSUBJ} & \overrightarrow{student} \\ \text{HEAD} & \boxed{1} \\ \text{NOBJ} & \overrightarrow{book} \\ \text{NMOD-ON} & \overrightarrow{beach} \end{bmatrix}, \sigma_i \right\rangle \right] \left[book\ \boxed{1} \left\langle \begin{bmatrix} \textbf{EVENT} & \\ \text{NSUBJ} & \overrightarrow{student} \\ \text{HEAD} & \overrightarrow{read} \\ \text{NOBJ} & \boxed{1} \\ \text{NMOD-ON} & \overrightarrow{beach} \end{bmatrix}, \sigma_k \right\rangle \right]$$

Lexical Items as Cues to GEK in K_{gen}

- We model K_{gen} as a set of pairs $\langle e, \sigma \rangle$, such that:
  - e is an event stored in GEK_{DS}
  - σ is a score expressing the salience of the event with respect to the construction that cues it (e.g., $P(e|Cxn)$)
- Each event in GEK_{DS} may be cued by several lexical items, as part of their semantic content

$$\begin{bmatrix} \textit{read}\ \boxed{1} \\ \left\langle \begin{bmatrix} \textbf{EVENT} & \\ \text{NSUBJ} & \overrightarrow{\textit{student}} \\ \text{HEAD} & \boxed{1} \\ \text{NOBJ} & \overrightarrow{\textit{book}} \\ \text{NMOD-ON} & \overrightarrow{\textit{beach}} \end{bmatrix}, \sigma_i \right\rangle \end{bmatrix} \begin{bmatrix} \textit{book}\ \boxed{1} \\ \left\langle \begin{bmatrix} \textbf{EVENT} & \\ \text{NSUBJ} & \overrightarrow{\textit{student}} \\ \text{HEAD} & \overrightarrow{\textit{read}} \\ \text{NOBJ} & \boxed{1} \\ \text{NMOD-ON} & \overrightarrow{\textit{beach}} \end{bmatrix}, \sigma_k \right\rangle \end{bmatrix}$$

Lexical Items as Cues to GEK

$$
\begin{bmatrix}
\text{FORM} & \textit{student}\ \boxed{1} \\
\text{SEM} & \begin{bmatrix}
\text{GEK} & \left[\left\langle \begin{bmatrix} \textbf{EVENT} & \\ \text{NSUBJ} & \boxed{1} \\ \text{HEAD} & \overrightarrow{\textit{read}} \\ \text{NOBJ} & \overrightarrow{\textit{book}} \\ \text{NMOD-ON} & \overrightarrow{\textit{beach}} \end{bmatrix}, \sigma_1 \right\rangle, \left\langle \begin{bmatrix} \textbf{EVENT} & \\ \text{NSUBJ} & \boxed{1} \\ \text{HEAD} & \overrightarrow{\textit{study}} \\ \text{NMOD-IN} & \overrightarrow{\textit{library}} \end{bmatrix}, \sigma_2 \right\rangle, \ldots \right] \\
\text{NEI} & \left[\langle \overrightarrow{\textit{pupil}}, s_1 \rangle, \langle \overrightarrow{\textit{learner}}, s_1 \rangle, \ldots \right]
\end{bmatrix}
\end{bmatrix}
$$

A Dynamic Interpretation of the Distributional Hypothesis

- The "standard" Distributional Hypothesis
  - You know the content of an expression *E* if you know the contexts in which *E* occurs
- The Articulated Context of *E* contains information about likely events activated by *E*

The Dynamic Distributional Hypothesis

You know the content of an expression *E* if you know the changes it causes in the expectations about the likely events represented in the Articulated Context

A Dynamic Interpretation of the Distributional Hypothesis

- The “standard” Distributional Hypothesis
  - You know the content of an expression *E* if you know the contexts in which *E* occurs
- The Articulated Context of *E* contains information about likely events activated by *E*

The Dynamic Distributional Hypothesis

You know the content of an expression *E* if you know the changes it causes in the expectations about the likely events represented in the Articulated Context

Expectations Change Potentials

The surfer

- rides the waves in the ocean
- has a board
- is in the water
- is on the beach
- is naked
- has a wetsuit
- put the wax onto the board
- drinks a beer
- ...

The surfer reads

- is on the beach
- is naked
- reads a book
- reads a comic
- reads a newspaper
- has a board
- ...

Expectations Change Potentials

The surfer

- rides the waves in the ocean
- has a board
- is in the water
- is on the beach
- is naked
- has a wetsuit
- put the wax onto the board
- drinks a beer
- . . .

The surfer reads

- is on the beach
- is naked
- reads a book
- reads a comic
- reads a newspaper
- has a board
- . . .

Expectations Change Potentials

The surfer

- rides the waves in the ocean
- has a board
- is in the water
- is on the beach
- is naked
- has a wetsuit
- put the wax onto the board
- drinks a beer
- ...

The surfer reads

- is on the beach
- is naked
- reads a book
- reads a comic
- reads a newspaper
- has a board
- ...

Reading surfer

Expectations Change Potentials

The surfer reads

- is on the beach
- is naked
- reads a book
- reads a comic
- reads a newspaper
- has a board
- ...

The surfer reads in the library

- is at a table
- reads a book
- is dressed
- is sitting on a chair
- there are bookshelves
- ...

Improbable Libraries

Alex Johnson, *Improbable Libraries. A Visual Journey to the World's Most Unusual Libraries*, 2015

Expectations Update

Bicknell K. *et al.* (2010), "Effects of event knowledge in processing verbal arguments", *Journal of Memory and Language*, 63: 489-505

- Self-paced reading and ERP studies show that the choice of agent noun alters the event the verb describes, by modifying the verb expectations about its patient argument
  (1) The **journalist**$_{AG}$ checked the **spelling**$_{PA}$ of his latest report (congruent)
  (2) The **mechanic**$_{AG}$ checked the **spelling**$_{PA}$ of his latest report (incongruent)

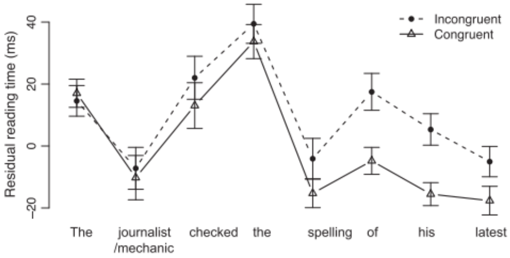
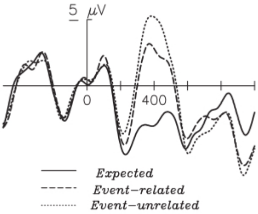


Fig. 1. Mean residual reading times. Error bars show one standard error above and below the mean, calculated by participants.

Expectations Update

Metusalem *et al.* (2012), "Generalized event knowledge activation during online sentence comprehension", *Journal of Memory and Language*, 66: 545-567

- Subjects activate elements in GEK related to a discourse representation even if these violate local linguistic restrictions
  - *Michelle had a toothache for several months. She knew she should do something about it, but held off. She finally got checked out when she was told she could get some anesthetic to reduce the PAIN/ DENTIST/ DRIVER and ease her discomfort*
    - *PAIN* **linguistically expected**
    - *DENTIST* **linguistically unexpected, but event-related**
    - *DRIVER* **linguistically unexpected, but event-unrelated**



Events and Bridging Definite Descriptions

1. The surfer rides the waves in the ocean. The board is white
2. The surfer is in the ocean. The board is white.
3. The surfer is on the beach. ?The board is white
4. The surfer reads the book in the library. *The board is white

Events and Bridging Definite Descriptions

1. The surfer rides the waves in the ocean. The board is white
2. The surfer is in the ocean. The board is white.
3. The surfer is on the beach. ?The board is white
4. The surfer reads the book in the library. *The board is white

Events and Bridging Definite Descriptions

1. The surfer rides the waves in the ocean. The board is white
2. The surfer is in the ocean. The board is white.
3. The surfer is on the beach. ?The board is white
4. The surfer reads the book in the library. *The board is white

Events and Bridging Definite Descriptions

1. The surfer rides the waves in the ocean. The board is white
2. The surfer is in the ocean. The board is white.
3. The surfer is on the beach. ?The board is white
4. The surfer reads the book in the library. *The board is white

Distributional Dynamics: First Hypothesis

- K_{gen} acts as long-term memory containing GEK activated by lexical items
- K_{disc} acts as a working memory which is updated with information coming from K_{gen} language processing
- Given an Articulated Context $\langle K_{dis}, K_{enc}, K_{gen}, K_{env} \rangle$, discourse comprehension is carried out online by the following steps:
  - activation in K_{disc} by a given lexical item w_i of the GEK associated with it in K_{gen}, GEK_{w_i}
  - integration and update of the existing GEK in K_{disc} with GEK_{w_i}

Distributional Dynamics: First Hypothesis

- K_{gen} acts as long-term memory containing GEK activated by lexical items
- K_{disc} acts as a working memory which is updated with information coming from K_{gen} language processing

- Given an Articulated Context $\langle K_{dis}, K_{enc}, K_{gen}, K_{env} \rangle$, discourse comprehension is carried out online by the following steps:
  - activation in K_{disc} by a given lexical item w_i of the GEK associated with it in K_{gen}, GEK_{w_i}
  - integration and update of the existing GEK in K_{disc} with GEK_{w_i}

Distributional Dynamics: First Hypothesis

- K_{gen} acts as long-term memory containing GEK activated by lexical items
- K_{disc} acts as a working memory which is updated with information coming from K_{gen} language processing

- Given an Articulated Context $\langle K_{dis}, K_{enc}, K_{gen}, K_{env} \rangle$, discourse comprehension is carried out online by the following steps:
  - activation in K_{disc} by a given lexical item w_i of the GEK associated with it in K_{gen}, GEK_{w_i}
  - integration and update of the existing GEK in K_{disc} with GEK_{w_i}

Distributional Dynamics: First Hypothesis

nsubj

The surfer reads

$$\left\langle K_{disc} \left[\text{GEK} \quad F(surfer_{K_{gen}}, read_{K_{gen}}) \right] \right\rangle$$

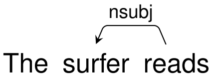
nmod

nsubj case

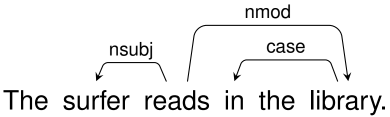
The surfer reads in the library.

$$\left\langle K_{disc} \left[\text{GEK} \quad F((surfer\ read)_{K_{gen}}, library_{K_{gen}}) \right] \right\rangle$$

Distributional Dynamics: First Hypothesis



$$\left\langle K_{disc} \left[\text{GEK} \quad F(surfer_{K_{gen}}, read_{K_{gen}}) \right] \right\rangle$$



$$\left\langle K_{disc} \left[\text{GEK} \quad F((surfer\ read)_{K_{gen}}, library_{K_{gen}}) \right] \right\rangle$$

The Update Function

- The update function F is a compositional function that unifies the events AVMs of two lexical items and updates their scores:
  - $F(GEK_{w_1}, GEK_{w_2}) = GEK_{w_1,w_2}$
- F is actually formed by two functions F_e and F_σ:
  1. F_e unifies two event AVMs e_i and e_j, producing a new event AVM e_k:

$$F_e(e_i, e_j) = e_k = e_i \sqcup e_j \quad (1)$$

  2. F_σ updates the event weights of the successfully unified events, by combining the weights of e_i and e_j into a new weight assigned to e_k, e.g., by summation:

$$F_\sigma(\sigma_i, \sigma_j) = \sigma_k = \sigma_i + \sigma_j \quad (2)$$

The Update Function

- The update function F is a compositional function that unifies the events AVMs of two lexical items and updates their scores:
  - $F(GEK_{w_1}, GEK_{w_2}) = GEK_{w_1, w_2}$
- F is actually formed by two functions F_e and F_σ:
  1. F_e unifies two event AVMs e_i and e_j, producing a new event AVM e_k:

$$F_e(e_i, e_j) = e_k = e_i \sqcup e_j \qquad (1)$$

  2. F_σ updates the event weights of the successfully unified events, by combining the weights of e_i and e_j into a new weight assigned to e_k, e.g., by summation:

$$F_\sigma(\sigma_i, \sigma_j) = \sigma_k = \sigma_i + \sigma_j \qquad (2)$$

Some Conclusions

- Distributional Semantics is usually viewed as a method to “squeze” semantic similarity from linguistic contexts
- Actually, there is much more semantically relevant information that can be extracted from distributional data
- Distributional Semantics can be used to model portions of event knowledge stored in K_{gen}
- The notion of Articulated Context offers promising synergies between Distributional and Dynamic Semantics
- K_{gen} and K_{disc} are likely to strongly interact during sentence processing and their interaction need to be explored in depth
- The update of K_{disc} during language comprehension can include an update function of distributional data about GEK, activated in K_{gen}
- Dynamic Distributional Semantics (or Distributional Dynamic Semantics) may offer new opportunities to model cognitive data

Some Conclusions

- Distributional Semantics is usually viewed as a method to “squeze” semantic similarity from linguistic contexts
- Actually, there is much more semantically relevant information that can be extracted from distributional data
- Distributional Semantics can be used to model portions of event knowledge stored in K_{gen}
- The notion of Articulated Context offers promising synergies between Distributional and Dynamic Semantics
- K_{gen} and K_{disc} are likely to strongly interact during sentence processing and their interaction need to be explored in depth
- The update of K_{disc} during language comprehension can include an update function of distributional data about GEK, activated in K_{gen}
- Dynamic Distributional Semantics (or Distributional Dynamic Semantics) may offer new opportunities to model cognitive data

Some Conclusions

- Distributional Semantics is usually viewed as a method to “squeze” semantic similarity from linguistic contexts
- Actually, there is much more semantically relevant information that can be extracted from distributional data
- Distributional Semantics can be used to model portions of event knowledge stored in K_{gen}
- The notion of Articulated Context offers promising synergies between Distributional and Dynamic Semantics
- K_{gen} and K_{disc} are likely to strongly interact during sentence processing and their interaction need to be explored in depth
- The update of K_{disc} during language comprehension can include an update function of distributional data about GEK, activated in K_{gen}
- Dynamic Distributional Semantics (or Distributional Dynamic Semantics) may offer new opportunities to model cognitive data

Some Conclusions

- Distributional Semantics is usually viewed as a method to “squeze” semantic similarity from linguistic contexts
- Actually, there is much more semantically relevant information that can be extracted from distributional data
- Distributional Semantics can be used to model portions of event knowledge stored in K_{gen}
- The notion of Articulated Context offers promising synergies between Distributional and Dynamic Semantics
- K_{gen} and K_{disc} are likely to strongly interact during sentence processing and their interaction need to be explored in depth
- The update of K_{disc} during language comprehension can include an update function of distributional data about GEK, activated in K_{gen}
- Dynamic Distributional Semantics (or Distributional Dynamic Semantics) may offer new opportunities to model cognitive data

Some Conclusions

- Distributional Semantics is usually viewed as a method to “squeze” semantic similarity from linguistic contexts
- Actually, there is much more semantically relevant information that can be extracted from distributional data
- Distributional Semantics can be used to model portions of event knowledge stored in K_{gen}
- The notion of Articulated Context offers promising synergies between Distributional and Dynamic Semantics
- K_{gen} and K_{disc} are likely to strongly interact during sentence processing and their interaction need to be explored in depth
- The update of K_{disc} during language comprehension can include an update function of distributional data about GEK, activated in K_{gen}
- Dynamic Distributional Semantics (or Distributional Dynamic Semantics) may offer new opportunities to model cognitive data

Some Conclusions

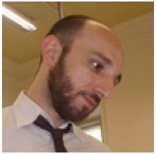
- Distributional Semantics is usually viewed as a method to “squeze” semantic similarity from linguistic contexts
- Actually, there is much more semantically relevant information that can be extracted from distributional data
- Distributional Semantics can be used to model portions of event knowledge stored in K_{gen}
- The notion of Articulated Context offers promising synergies between Distributional and Dynamic Semantics
- K_{gen} and K_{disc} are likely to strongly interact during sentence processing and their interaction need to be explored in depth
- The update of K_{disc} during language comprehension can include an update function of distributional data about GEK, activated in K_{gen}
- Dynamic Distributional Semantics (or Distributional Dynamic Semantics) may offer new opportunities to model cognitive data

CoLing Lab

http://colinglab.humnet.unipi.it/

- This research is conducted in collaboration with:
  - Emmanuele Chersoni

  - Gianluca Lebani



Grazie!!!

Thank You!!!

Danke!!!