Meaning in Context Symposium
Welcome

Hinrich Schütze

Center for Information and Language Processing, LMU Munich

2015-09-14

Participants

Alessandro Lenci, Pisa
Ann Copestake, Cambridge
Annie Zaenen, Stanford
Bonnie Webber, Edinburgh
Georgiana Dinu, IBM
Hans Kamp, Stuttgart
Ido Dagan, Bar-Ilan
Jerry Hobbs, ISI
Katrin Erk, UT Austin
Lauri Karttunen, Stanford
Louise McNally, UPF
Marco Baroni, Trento
Mark Steedman, Edinburgh
Mirella Lapata, Edinburgh
Phil Blunsom, Oxford
Sebastian Löbner, Düsseldorf
Sebastian Pado, Stuttgart
Sebastian Riedel, UCL
Stephen Clark, Cambridge
Tomas Mikolov, Facebook

Program

Monday, 14 September 2015

Morning Session	Chair: Hinrich Schuetze
10:00 - 10:30	P Welcome Address
10:30 - 11:30	P Marco Baroni, University of Trento
11:30 - 12:00	P Discussants: Georgiana Dinu, Hans Kamp
12:00 - 12:30	P Plenary Discussion
12:30 - 13:30	Lunch Break
Afternoon Session	
13:30 - 15:15	W Dynamic & Formal
15:15 - 15:45	Coffee Break
15:45 - 17:00	W Dynamic & Formal

Program

Tuesday, 15 September 2015

Morning Session	Chair: NN
09:00 - 10:30	W Dynamic & Formal
10:30 - 11:00	Coffee Break
11:00 - 12:00	P Tomas Mikolov, Facebook
12:00 - 12:30	P Plenary Discussion
12:30 - 13:30	Lunch Break
Afternoon Session	
13:30 - 15:15	W Dynamic & Formal
15:15 - 15:45	Coffee Break
15:45 - 17:00	W Dynamic & Formal

Program

Wednesday, 16 September 2015

Morning Session	Chair: Alessandro Lenci
09:00 - 10:00	P Katrin Erk, UT Austin
10:00 - 10:30	P Plenary Discussion
10:30 - 11:00	Coffee Break
11:00 - 11:45	P WG Dynamic Reports to Plenum
11:45 - 12:30	P WG Formal Reports to Plenum
12:30 - 13:30	Lunch Break
Afternoon Session	
13:30 - 15:15	W KR & Sentence
15:15 - 15:45	Coffee Break
15:45 - 17:00	W KR & Sentence
19:00	Conference Dinner (Osterwaldgarten)

Program

Thursday, 17 September 2015

Morning Session	Chair: NN
09:00 - 10:00	P Phil Blunsom, Oxford
10:00 - 10:30	P Plenary Discussion
10:30 - 11:00	Coffee Break
11:00 - 12:30	W KR & Sentence
12:30 - 13:30	Lunch Break
Afternoon Session	
13:30 - 15:15	W KR & Sentence
15:15 - 15:45	Coffee Break
15:45 - 17:00	W KR & Sentence

Program

Friday, 18 September 2015

Morning Session	Chair: Mirella Lapata
09:00 - 09:45	P WG KR Reports to Plenum
09:45 - 10:30	P WG Sentence Report to Plenum
10:30 - 11:00	Coffee Break
11:00 - 12:15	P Plenary Discussion
12:30 - 14:15	Lunch Break (Café Reitschule)
Afternoon Session	
14:15 - 15:15	P Wrap Up: Follow-Up Activities
15:15 - 15:45	Coffee Break
15:45 - 21:00	P CIS Meetings, Dinner

Four working groups

Four working groups

- WG Dynamic, Mo/Tu: Dynamic Distributional Semantics

Four working groups

- WG Dynamic, Mo/Tu: Dynamic Distributional Semantics
- WG Formal, Mo/Tu: Formal and Distributional Semantics

Four working groups

- WG Dynamic, Mo/Tu: Dynamic Distributional Semantics
- WG Formal, Mo/Tu: Formal and Distributional Semantics
- WG KR, We/Th: Natural Language Based Knowledge Representation

Four working groups

- WG Dynamic, Mo/Tu: Dynamic Distributional Semantics
- WG Formal, Mo/Tu: Formal and Distributional Semantics
- WG KR, We/Th: Natural Language Based Knowledge Representation
- WG Sentence, We/Th: Vector Representations of Sentences and Other

WG Formal (Erk/Pado) (1)

WG Formal (Erk/Pado) (1)

- Philosophy of language aspects

WG Formal (Erk/Pado) (1)

- Philosophy of language aspects
  - Use-based vs. truth-based meaning: the one, the other, both?

WG Formal (Erk/Pado) (1)

- Philosophy of language aspects
  - Use-based vs. truth-based meaning: the one, the other, both?
  - "Transgressions" of distributional semantics: mixture of linguistic & world knowledge, inclusion of multi-modal information, "pragmatic" footprint: problematic or not?

WG Formal (Erk/Pado) (2)

WG Formal (Erk/Pado) (2)

- The role of distributional and formal representation in semantics

WG Formal (Erk/Pado) (2)

- The role of distributional and formal representation in semantics
  - Many of us seem to subscribe to models of semantics where both formal and distributional components have a role to play

WG Formal (Erk/Pado) (2)

- The role of distributional and formal representation in semantics
  - Many of us seem to subscribe to models of semantics where both formal and distributional components have a role to play
  - What does the division of labor look like - where are the interfaces? (with regard to representation, with regard to inference, etc.)

WG Formal (Erk/Pado) (2)

- The role of distributional and formal representation in semantics
  - Many of us seem to subscribe to models of semantics where both formal and distributional components have a role to play
  - What does the division of labor look like - where are the interfaces? (with regard to representation, with regard to inference, etc.)
  - Current models seem to couple the formal and distributional aspects loosely rather than opting for a unified single model: is that reasonable?

WG Formal (Erk/Pado) (2)

- The role of distributional and formal representation in semantics
  - Many of us seem to subscribe to models of semantics where both formal and distributional components have a role to play
  - What does the division of labor look like - where are the interfaces? (with regard to representation, with regard to inference, etc.)
  - Current models seem to couple the formal and distributional aspects loosely rather than opting for a unified single model: is that reasonable?
  - At first glance, current models look rather scattered with no common aim or method. What do they have in common?

WG Formal (Erk/Pado) (3)

WG Formal (Erk/Pado) (3)

- Next steps for combined formal and distributional semantic models

WG Formal (Erk/Pado) (3)

- Next steps for combined formal and distributional semantic models
  - What phenomena should such models aim to cover?

WG Formal (Erk/Pado) (3)

- Next steps for combined formal and distributional semantic models
  - What phenomena should such models aim to cover?
  - What general architecture should they use?

WG Formal (Erk/Pado) (3)

- Next steps for combined formal and distributional semantic models
  - What phenomena should such models aim to cover?
  - What general architecture should they use?
  - What are the most pressing open questions?

WG Formal (Erk/Pado) (3)

- Next steps for combined formal and distributional semantic models
  - What phenomena should such models aim to cover?
  - What general architecture should they use?
  - What are the most pressing open questions?
  - What should/will happen with the current "zoo" of models in a couple of years?

WG Dynamic (Baroni/Schütze)

WG Dynamic (Baroni/Schütze)

- Will focus on reference: anaphora, definite descriptions etc.

WG Dynamic (Baroni/Schütze)

- Will focus on reference: anaphora, definite descriptions etc.
- Is there a distributed-representation way of modeling reference or should this be left to formal semantics?

WG Dynamic (Baroni/Schütze)

- Will focus on reference: anaphora, definite descriptions etc.
- Is there a distributed-representation way of modeling reference or should this be left to formal semantics?
- Not a topic: how distributed semantics can be a component in, say, coreference resolution systems.

WG Dynamic (Baroni/Schütze)

- Will focus on reference: anaphora, definite descriptions etc.
- Is there a distributed-representation way of modeling reference or should this be left to formal semantics?
- Not a topic: how distributed semantics can be a component in, say, coreference resolution systems.
- Memory?

WG Dynamic (Baroni/Schütze)

- Will focus on reference: anaphora, definite descriptions etc.
- Is there a distributed-representation way of modeling reference or should this be left to formal semantics?
- Not a topic: how distributed semantics can be a component in, say, coreference resolution systems.
- Memory?
- Presentation of computational models: Baroni, Blunsom, Mikolov, ...

WG Dynamic (Baroni/Schütze)

- Will focus on reference: anaphora, definite descriptions etc.
- Is there a distributed-representation way of modeling reference or should this be left to formal semantics?
- Not a topic: how distributed semantics can be a component in, say, coreference resolution systems.
- Memory?
- Presentation of computational models: Baroni, Blunsom, Mikolov, ...
- Are anaphora variables and does that mean we have a neural/distributed binding problem?

WG Dynamic (Baroni/Schütze): Questions

WG Dynamic (Baroni/Schütze): Questions

- How do we represent the broader discourse context and the entities it defines in DS?

WG Dynamic (Baroni/Schütze): Questions

- How do we represent the broader discourse context and the entities it defines in DS?
- How do we build such discourse representations compositionally and dynamically?

WG Dynamic (Baroni/Schütze): Questions

- How do we represent the broader discourse context and the entities it defines in DS?
- How do we build such discourse representations compositionally and dynamically?
- How do we learn a reference-aware DS model from the data? From which data?

WG Dynamic (Baroni/Schütze): Questions

- How do we represent the broader discourse context and the entities it defines in DS?
- How do we build such discourse representations compositionally and dynamically?
- How do we learn a reference-aware DS model from the data? From which data?
- How do we represent sentences that are "referring" in DS?

WG Dynamic (Baroni/Schütze): Questions

- How do we represent the broader discourse context and the entities it defines in DS?
- How do we build such discourse representations compositionally and dynamically?
- How do we learn a reference-aware DS model from the data? From which data?
- How do we represent sentences that are "referring" in DS?
- Actually, is representing sentences the right approach or do alternatives exist in which our DS model represents the subpart of discourse context that is necessary to interpret subsequent discourse, but does not represent sentences as vectors?

WG KR (Dagan/Riedel)

WG KR (Dagan/Riedel)

- Natural language knowledge representation (KR) / natural language knowledge base (KB)

WG KR (Dagan/Riedel)

- Natural language knowledge representation (KR) / natural language knowledge base (KB)
- Vision: a kind of knowledge graph, but the nodes are propositions, not entities

WG KR (Dagan/Riedel)

- Natural language knowledge representation (KR) / natural language knowledge base (KB)
- Vision: a kind of knowledge graph, but the nodes are propositions, not entities
- Which propositions to represent and how to represent them, including their internal structure?

WG KR (Dagan/Riedel)

- Natural language knowledge representation (KR) / natural language knowledge base (KB)
- Vision: a kind of knowledge graph, but the nodes are propositions, not entities
- Which propositions to represent and how to represent them, including their internal structure?
- Merging information across NL propositions

WG KR (Dagan/Riedel)

- Natural language knowledge representation (KR) / natural language knowledge base (KB)
- Vision: a kind of knowledge graph, but the nodes are propositions, not entities
- Which propositions to represent and how to represent them, including their internal structure?
- Merging information across NL propositions
- Relations between different propositions (e.g., causality)

WG KR (Dagan/Riedel)

- Natural language knowledge representation (KR) / natural language knowledge base (KB)
- Vision: a kind of knowledge graph, but the nodes are propositions, not entities
- Which propositions to represent and how to represent them, including their internal structure?
- Merging information across NL propositions
- Relations between different propositions (e.g., causality)
- Inference mechanisms over knowledge representations

WG KR (Dagan/Riedel)

- Natural language knowledge representation (KR) / natural language knowledge base (KB)
- Vision: a kind of knowledge graph, but the nodes are propositions, not entities
- Which propositions to represent and how to represent them, including their internal structure?
- Merging information across NL propositions
- Relations between different propositions (e.g., causality)
- Inference mechanisms over knowledge representations
- Use cases and applications (e.g., question answering)

WG KR (Dagan/Riedel): Initial discussion

WG KR (Dagan/Riedel): Initial discussion

- Is this a semantic net and if not, what is different?

WG KR (Dagan/Riedel): Initial discussion

- Is this a semantic net and if not, what is different?
- How can propositions be recognized in text?

WG KR (Dagan/Riedel): Initial discussion

- Is this a semantic net and if not, what is different?
- How can propositions be recognized in text?
- The role of distributional semantics in merging information across propositions

WG KR (Dagan/Riedel): Initial discussion

- Is this a semantic net and if not, what is different?
- How can propositions be recognized in text?
- The role of distributional semantics in merging information across propositions
- Does it scale? $O(n^2)$ comparisons for n nodes

WG KR (Dagan/Riedel): Initial discussion

- Is this a semantic net and if not, what is different?
- How can propositions be recognized in text?
- The role of distributional semantics in merging information across propositions
- Does it scale? $O(n^2)$ comparisons for n nodes
- Event coreference may be crucial:
  relationship to Martha Palmer's ERE?

WG KR (Dagan/Riedel): Initial discussion

- Is this a semantic net and if not, what is different?
- How can propositions be recognized in text?
- The role of distributional semantics in merging information across propositions
- Does it scale? $O(n^2)$ comparisons for n nodes
- Event coreference may be crucial:
  relationship to Martha Palmer's ERE?
- What is the best way of taking into account context for recognizing propositions in text, for merging them and for event coreference?

WG KR (Dagan/Riedel): Targeted deliverables

WG KR (Dagan/Riedel): Targeted deliverables

- Initial specification of a largely new research area (position-paper style)

WG KR (Dagan/Riedel): Targeted deliverables

- Initial specification of a largely new research area (position-paper style)
- Concrete component task specifications

WG KR (Dagan/Riedel): Targeted deliverables

- Initial specification of a largely new research area (position-paper style)
- Concrete component task specifications
- A set of analyzed/annotated examples (serving the other deliverables)

WG Sentence (Clark/Dinu)

Starting point

Neural network-based machine translation models are providing competitive results compared to the state-of-the-art in phrase- and syntax-based SMT, and at the heart of the neural MT model is a vector-based representation of sentences which effectively acts as an interlingua. The relative success of neural MT models is somewhat puzzling, and raises a number of questions.

WG Sentence (Clark/Dinu)

WG Sentence (Clark/Dinu)

Famous tweet by Ray Mooney

You can't cram the meaning of a whole %&$# sentence into a single $&# vector

WG Sentence (Clark/Dinu)

Famous tweet by Ray Mooney

You can't cram the meaning of a whole %&$# sentence into a single $&# vector

WG Sentence (Clark/Dinu)

Famous tweet by Ray Mooney

You can't cram the meaning of a whole %&$# sentence into a single $&# vector

VS.

WG Sentence (Clark/Dinu): Questions

WG Sentence (Clark/Dinu): Questions

- Is there a limit to how much information can be represented in a fixed-size vector? If so, is this at odds with a view of language in which syntax/morphology is highly productive?

WG Sentence (Clark/Dinu): Questions

- Is there a limit to how much information can be represented in a fixed-size vector? If so, is this at odds with a view of language in which syntax/morphology is highly productive?
- If these methods accurately encode sentence meaning in fixed-length representations, what can we learn about how they perform meaning composition?

WG Sentence (Clark/Dinu): Questions

- Is there a limit to how much information can be represented in a fixed-size vector? If so, is this at odds with a view of language in which syntax/morphology is highly productive?
- If these methods accurately encode sentence meaning in fixed-length representations, what can we learn about how they perform meaning composition?
- Sentence vectors in neural MT models not only represent the source-side sentence; they are also used to generate the target side. Limitations in NLG when the input is a vector?

WG Sentence (Clark/Dinu): Questions

- Is there a limit to how much information can be represented in a fixed-size vector? If so, is this at odds with a view of language in which syntax/morphology is highly productive?
- If these methods accurately encode sentence meaning in fixed-length representations, what can we learn about how they perform meaning composition?
- Sentence vectors in neural MT models not only represent the source-side sentence; they are also used to generate the target side. Limitations in NLG when the input is a vector?
- Are there problems in natural language semantics, e.g. quantification, which cannot be solved using vector representations? Should we be trying to solve all of semantics using neural networks?

WG Sentence (Clark/Dinu): Questions

- Is there a limit to how much information can be represented in a fixed-size vector? If so, is this at odds with a view of language in which syntax/morphology is highly productive?
- If these methods accurately encode sentence meaning in fixed-length representations, what can we learn about how they perform meaning composition?
- Sentence vectors in neural MT models not only represent the source-side sentence; they are also used to generate the target side. Limitations in NLG when the input is a vector?
- Are there problems in natural language semantics, e.g. quantification, which cannot be solved using vector representations? Should we be trying to solve all of semantics using neural networks?
- Can vector-based representations be usefully extended beyond sentences, e.g. to represent elements of discourse?

Four working groups

- WG Dynamic, Mo/Tu: Dynamic Distributional Semantics
- WG Formal, Mo/Tu: Formal and Distributional Semantics
- WG KR, We/Th: Natural Language Based Knowledge Representation
- WG Sentence, We/Th: Vector Representations of Sentences and Other

Organization

Organization

- Two rooms

Organization

- Two rooms
- Please do not take cups and plates up to the main lecture hall.

Instructions to Rapporteurs (1)

- The reporting session for your working group has a length of 45 minutes.
- Two reporting sessions are Wednesday, 11:00-12:30, two Friday, 09:00-10:30
- You probably want to leave some time for discussion, but it's up to you.
- One person or several people can report.
- Each WG should answer the MIC-specific questions (first block below).
- The other two blocks ("general questions", "detailed questions") are intended to help you with the report, but you don't have to address them.

Instructions to Rapporteurs (2)

- Key results of your working group
- Fundamental research questions / research problems that you identified
- Points of agreement / points of disagreement
- New ideas / new approaches / new insights you want to share with the plenum
- What has changed in the two years since Dagstuhl?
- In which areas do you see most interesting work happening in the next two years?

Instructions to Rapporteurs (3)

- What is the relation between formal semantics and distributional semantics? Joint or separate? One a component of the other? Division of labor?
- (Updated) Dagstuhl questions from 2013
  - Hardest problems for formal semantics that distributional semantics can handle well
  - Hardest problems for distributional semantics that formal semantics can handle well
  - Compositionality
  - Dynamics in the sense of dynamic semantics like DRT
- Shared tasks?
- Any planned publications?
- Other follow-up activities

Instructions to Rapporteurs (4)

Instructions to Rapporteurs (4)

∗ Representations for linguistic entities (formal, distributed etc.) are context-independent static representations in some formalisms and operators on context in other formalisms.

Instructions to Rapporteurs (4)

∗ Representations for linguistic entities (formal, distributed etc.) are context-independent static representations in some formalisms and operators on context in other formalisms.
∗ Word representations are often context-independent static representations. However, CCG and RNN-type NNs treat words as operators on context. What is the right approach?

Instructions to Rapporteurs (4)

∗ Representations for linguistic entities (formal, distributed etc.) are context-independent static representations in some formalisms and operators on context in other formalisms.
∗ Word representations are often context-independent static representations. However, CCG and RNN-type NNs treat words as operators on context. What is the right approach?
∗ The most common way of representing sentences is context-independent. However, DRT and its ilk treat sentences as operators on context. What is the right approach?

Instructions to Rapporteurs (4)

∗ Representations for linguistic entities (formal, distributed etc.) are context-independent static representations in some formalisms and operators on context in other formalisms.
∗ Word representations are often context-independent static representations. However, CCG and RNN-type NNs treat words as operators on context. What is the right approach?
∗ The most common way of representing sentences is context-independent. However, DRT and its ilk treat sentences as operators on context. What is the right approach?
∗ Composition: context-dependent or context-independent?

Instructions to Rapporteurs (4)

* Representations for linguistic entities (formal, distributed etc.) are context-independent static representations in some formalisms and operators on context in other formalisms.
* Word representations are often context-independent static representations. However, CCG and RNN-type NNs treat words as operators on context. What is the right approach?
* The most common way of representing sentences is context-independent. However, DRT and its ilk treat sentences as operators on context. What is the right approach?
* Composition: context-dependent or context-independent?
* Include a graph in your presentation that shows architectural design and data flow: what is the role of context in natural language understanding?

Instructions to Rapporteurs (4)

∗ Representations for linguistic entities (formal, distributed etc.) are context-independent static representations in some formalisms and operators on context in other formalisms.
∗ Word representations are often context-independent static representations. However, CCG and RNN-type NNs treat words as operators on context. What is the right approach?
∗ The most common way of representing sentences is context-independent. However, DRT and its ilk treat sentences as operators on context. What is the right approach?
∗ Composition: context-dependent or context-independent?
∗ Include a graph in your presentation that shows architectural design and data flow: what is the role of context in natural language understanding?
∗ There is a (at this point completely hypothetical) possibility of a third meeting in Stanford, 2016 or 2017. Which topics would you suggest this third meeting address?

Coherence

Coherence

- Please make sure to keep the MIC questions in mind

Coherence

- Please make sure to keep the MIC questions in mind
- Devote some time to discussing them

Coherence

- Please make sure to keep the MIC questions in mind
- Devote some time to discussing them
- Should be easy for WGs Dynamic and Sentence

Coherence

- Please make sure to keep the MIC questions in mind
- Devote some time to discussing them
- Should be easy for WGs Dynamic and Sentence
- Also for Formal? "Next steps for combined formal and distributional semantic models"

Coherence

- Please make sure to keep the MIC questions in mind
- Devote some time to discussing them
- Should be easy for WGs Dynamic and Sentence
- Also for Formal? "Next steps for combined formal and distributional semantic models"
- More tricky for KR? "Best way of taking into account context for recognizing propositions in text, for merging them and for event coreference?"

Coherence

- Please make sure to keep the MIC questions in mind
- Devote some time to discussing them
- Should be easy for WGs Dynamic and Sentence
- Also for Formal? "Next steps for combined formal and distributional semantic models"
- More tricky for KR? "Best way of taking into account context for recognizing propositions in text, for merging them and for event coreference?"
- There should be some synergy between working groups – don't organize your working group as an isolated event that is not related to the rest of MIC.

Instructions to Rapporteurs (4)

∗ Representations for linguistic entities (formal, distributed etc.) are context-independent static representations in some formalisms and operators on context in other formalisms.
∗ Word representations are often context-independent static representations. However, CCG and RNN-type NNs treat words as operators on context. What is the right approach?
∗ The most common way of representing sentences is context-independent. However, DRT and its ilk treat sentences as operators on context. What is the right approach?
∗ Composition: context-dependent or context-independent?
∗ Include a graph in your presentation that shows architectural design and data flow: what is the role of context in natural language understanding?
∗ There is a (at this point completely hypothetical) possibility of a third meeting in Stanford, 2016 or 2017. Which topics would you suggest this third meeting address?

Questions? Comments?