Logic and Natural Language Semantics: NLP applications

RAFFAELLA BERNARDI

DISI, UNIVERSITY OF TRENTO E-MAIL: BERNARDI@DISI.UNITN.IT

Contents

1	Back to the general picture						
	1.1	Logic as "science of reasoning"	5				
2	Formal Semantics & DB access						
	2.1	Research questions and Examples	7				
3	Questio	stion Answering (QA) 8					
4	Textual Entailment						
	4.1	Kind of inferences	10				
	4.2	Approaches: pro and contra	12				
	4.3	RTE PASCAL: observation	13				
	4.4	Logical inference & Shallow features	14				
	4.5	Challenges for shallow approaches	15				
	4.6	Natural Logic & Shallow features	16				
	4.7	NatLog against FraCaS data set	17				
	4.8	From FraCaS to RTE data set	18				
5	Back to philosophy of language						
	5.1	Back to words	20				
6	Logical	words: Quantifiers	21				

6.1	Quantifiers from the FS angles					
	6.1.1	Determiners: which relation	23			
	6.1.2	Conservativity and Extension	25			
	6.1.3	Symmetry	26			
	6.1.4	Monotonicity	27			
	6.1.5	Effects of Monotonicity	28			
6.2	Quantif	iers from the "language as use" angle	30			
	6.2.1	Quantifier Phrases and scalar implicature	31			
	6.2.2	Positive and Negative Quantifiers	32			
	6.2.3	Performative utterances with Quantifiers Phrases	33			
	6.2.4	QP and anaphora	34			
Concl	usion		35			

7

1. Back to the general picture



1.1. Logic as "science of reasoning"

Stoics put focus on propositional logic reasoning.

- Aristotele studies the relations holding between the structure of the premises and of the conclusion.
- Frege introduces quantifiers symbols and a way to represent sentences with more than one quantifiers.
- Tarski provides the model theoretical interpretation of Frege quantifiers and hence a way to deal with entailment involving more complex sentences that those studied by the Greek.
- Montague by studying the syntax-semantics relation of linguistic structure provides the framework for building FOL representation of natural language sentences and hence natural language reasoning.
- Lambek calculus captures the algebraic principles behind the syntax-semantic interface of linguistic structure and has been implemented.

How can such results be used in real life applications?

Have we captured natural language reasoning?

2. Formal Semantics & DB access

Query answering over ontology Take a reasoning system used to query a DB by exploiting an ontology:

$$\mathcal{O}, \mathcal{DB} \models q$$

the DB provides the references and their properties (A-Box of description logic), the ontologies the general knowledge (T-Box).

Formal Semantics allows the development of natural language interface to such systems:

- It allows domain experts to enter their knowledge in natural language sentences to build the ontology;
- It allows users to query the DB with natural language questions.

Good application: the DB provides entities and FS the meaning representation based on such entities.

2.1. Research questions and Examples

Research Can we get the users use natural language in a free way or do we need to control their use? Which controlled language? Develop a real case scenario.

Example of systems designed for writing unambiguous and precise specifications text:

- PENG: http://web.science.mq.edu.au/~rolfs/peng/
- ACE: http://attempto.ifi.uzh.ch/site/
- See Ian Pratt papers on fragments of English.

More at: http://sites.google.com/site/controllednaturallanguage/

3. Question Answering (QA)

Mixture of deep and shallow approaches. The deep approaches exploit NPL parsing and relation extraction methods.

AquaLog it's a QA system thattakes queries expressed in natural language and an ontology as input and returns answers drawn from one or more knowledge bases (KBs), which instantiate the input ontology with domain-specific information.

More: http://technologies.kmi.open.ac.uk/aqualog/

IBM Watson is QA system developed in IBM's DeepQA project; it won the American TV quiz show, Jeopardy. "The DeepQA hypothesis is that by complementing classic knowledge-based approaches with recent advances in NLP, Information Retrieval, and Machine Learning to interpret and reason over huge volumes of widely accessible naturally encoded knowledge (or "unstructured knowledge") we can build effective and adaptable open-domain QA systems."

More: http://www.research.ibm.com/deepqa/deepqa.shtml

4. Textual Entailment

Textual entailment recognition is the task of deciding, given two text fragments, whether the meaning of one text is entailed from another text. Useful for QA, IR, IE, MT, RC (reading comprehension), CD (comparable documents). Given a text T and an hypothesis H

 $T \models H$

if, typically, a human reading T would infer that H is most probably true.

• QA application:

T: "Norways most famous painting, The Scream by Edvard Munch, was recovered Saturday, almost three months after it was stolen from an Oslo museum." **H**: "Edvard Munch painted The Scream"

• IR application:

T: "Google files for its long awaited IPO". H: "Google goes public."

4.1. Kind of inferences

Benchmarks and an evaluation forum for entailment systems have been developed. Main kind of inferences:

syntactic inference: e.g. nominalization:

T: Sunday's election results demonstrated just how far the pendulum of public opinion has swung away from faith in *Koizumi's promise to bolster the Japanese economy* and make the political system more transparent and responsive to the peoples'needs.

H: Koizumi promised to bolster the Japanese economy.

other syntactic phenomena: apposition (), predicate-complement construction, coordination, embedded clauses, etc.

lexically based inferences: simply based on the presence of synonyms or the alike.

phrasal-level synonymy: T: "The three-day G8 summit ..." and **H**: "The G8 summit last three days"

4.2. Approaches: pro and contra

(slide by MacCartney)



4.3. RTE PASCAL: observation

Approaches used are combination of the following techniques:

- Machine Learning Classification systems
- Transformation-based techniques over syntactic representations
- Deep analysis and logical inferences
- Natural Logic

Overview of the task and the approaches: Ido Dagan, Bill Dolan, Bernardo Magnini, and Dan Roth. "Recognizing textual entailment: Rational, evaluation and approaches" Lucy Vanderwende and William B. Dolan: "What Syntax can Contribute in the Entailment Task." found that in the RTE dataset:

- 34% of the test items can be handled by syntax,
- 48% of the test items can be handled by syntax plus a general purpose thesaurus

http://pascallin.ecs.soton.ac.uk/Challenges/RTE/

4.4. Logical inference & Shallow features

Johan Bos and Katja Markert "When logical inference helps determining textual entailment (and when it doesn't)"

Task comparison: word overlap vs. FOL inferences (FOL theorem prover and finite model building techniques.) combined with ML techniques. Background knowledge built using WordNet (is-a relation) and manually coded inference rules.

- word overlap: 61.6 % accuracy
- both methods: 60.6 % accuracy.

4.5. Challenges for shallow approaches

The following valid entailment

T Every firm polled saw costs grow more than expected, even after adjusting for inflation. **H**: Every big company in the pool reported cost increases.

becomes invalid if we replace "every" with "some".

The sentences are difficult to be translated in FOL, but the entailment is also challenging for systems relying on lexical and syntactic similarity.

4.6. Natural Logic & Shallow features

B. MacCartney and C. D. Manning "An extended model of natural logic." presents a Natural Logic system (NatLog).

The system implements what semanticists have discovered about the some semantics properties of linguistic expressions that effects inference, like monotonicity of quantifier phrases (e.g. nobody) or other expressions (e.g. without):



4.7. NatLog against FraCaS data set

FraCas (project carried out by semanticists in the mid-1990s) dataset (http://www-nlp.
stanford.edu/~wcmac/downloads/fracas.xml)

topic	start	count	8	single- premise
Quantifiers	1	80	23 %	50
Plurals	81	33	10 %	24
Anaphora	114	28	8 %	6
Ellipsis	142	55	16 %	25
Adjectives	197	23	7 %	15
Comparatives	220	31	9 %	16
Temporal	251	75	22 %	39
Verbs	326	8	2 %	8
Attitudes	334	13	4 %	9
	topic Quantifiers Plurals Anaphora Ellipsis Adjectives Comparatives Temporal Verbs Attitudes	topicstartQuantifiers1Plurals81Anaphora114Ellipsis142Adjectives197Comparatives220Temporal251Verbs326Attitudes334	topic start count Quantifiers 1 80 Plurals 81 33 Anaphora 114 28 Ellipsis 142 55 Adjectives 197 23 Comparatives 220 31 Temporal 251 75 Verbs 326 8 Attitudes 334 13	topic start count % Quantifiers 1 80 23 % Plurals 81 33 10 % Anaphora 114 28 8 % Ellipsis 142 55 16 % Adjectives 197 23 7 % Comparatives 220 31 9 % Temporal 251 75 22 % Verbs 326 8 2 % Attitudes 334 13 4 %

NatLog achieve 70% accuracy over FraCaS data set entailment – only worked on entailment involving one sentence as premise.

4.8. From FraCaS to RTE data set

RTE differs from FraCaS in several ways: longer premises, more natural, diversity of types of inferences (including paraphrases, temporal reasoning, etc.)

NatLog combined with Stanford RTE systems led to accuracy gain of 4% against RTE data set.

See N. Chambers, D. Cer, T. Grenager, D. Hall, C. Kiddon, B. MacCartney, M. de Marneffe, D. Ramage, E. Yeh, C. Manning "Learning Alignments and Leveraging Natural Logic"

More on Natural Logic:

http://www.stanford.edu/~icard/logic&language/programme.html

5. Back to philosophy of language

Frege:

- 1. Linguistic signs have a reference and a sense:
 - (i) "Mark Twin is Mark Twin" vs. (ii) "Mark Twin is Samuel Clemens".
 - (i) same sense and same reference vs. (ii) different sense and same reference.
- 2. Both the sense and reference of a sentence are built compositionaly.
- 3. "sense" what is common to the sentences that yield the same consequent.
- 4. knowing a concept means to know in which network of concepts it lives in.

Lead to the Formal Semantics studies of natural language who focused on "meaning" as "reference".

Wittgenstein's claims brought philosophers of language to focus on "meaning" as "sense" leading to the "language as use" view.

5.1. Back to words

But,

the "language as use" school has focused on content words meaning.

vs.

Formal semantics school has focused mostly on the grammatical words and in particular on the behaviour of the "logical words".

- **content words** or open class: are words that carry the content or the meaning of a sentence and are open-class words, e.g. noun, verbs, adjectives and most adverbs.
- **grammatical words** or closed class: are words that serve to express grammatical relationships with other words within a sentence; they can be found in almost any utterance, no matter what it is about, e.g. such as articles, prepositions, conjunctions, auxiliary verbs, and pronouns.

Among the latter, one can distinguish the **logical words**, viz. those words that corresponds to logical operators: negation, conjunction, disjunction, quantifiers.

6. Logical words: Quantifiers

Logical words are not used always as their corresponding logic operators. E.g: "Another step and I shoot you" (= if you take another step, I shoot you.)

Quantifiers are typical logical words.

- In the previous lectures we have seen them at work on the syntax-semantic interface, putting attention on the scope ambiguity.
- In the following, we will see what has been understood about quantifiers from a Formal Semantics perspective and from an empirical view.

6.1. Quantifiers from the FS angles

Formal semantics have studied the properties that characterize all natural language determiners so to demarcate their class with the class of all logically possible determiners, and explain facts as the following: not all QP can be negated:

- not every man
- *not few man

Further info in B. Partee, A. ter Meulen and R. Wall "Mathematical methods in linguistics"

6.1.1. Determiners: which relation



Consider the four subsets of the universe of discourse as above. A determiner that refer:

- to only (ii) is an intersective determiner;
- to only (i) is a co-intersective determiner;
- to both (i) and (ii) are proportional determiners.

[(ii) viz. $NP' \cap Pred'$]

[(i) viz. NP' - Pred']

• Intersective determiners only (ii), $NP' \cap Pred'$

At least six robots = $\{\langle NP', Pred' \rangle : |NP' \cap Pred'| \ge 6\}$ no robots = $\{\langle NP', Pred' \rangle : |NP' \cap Pred'| = 0\}$

- Co-intersective determiners only (i), NP' Pred'every robot' = $\{\langle NP', Pred' \rangle : NP' \subseteq Pred'\}$
- Proportional determiners both (i) and (ii): more than 50% of the' = $\{\langle NP', Pred' \rangle : |NP' \cap Pred'| > |NP'|/2\}$ most of the' = $\{\langle NP', Pred' \rangle : |NP' \cap Pred'| > |NP' \cap -Pred'|\}$

Natural Language Determiners do not make references to area (iii) Pred' - NP' and (iv) $U - (NP' \cup Pred')$.

Indifference to area (iii) is known as "conservativity"; indifference to area (iv) is known as "extension".

6.1.2. Conservativity and Extension

Conservativity DET is conservative if

DET(NP')(Pred') is true iff $DET(NP')(NP' \cap Pred')$ is true.

Extension DET has extension if

DET(NP')(Pred') remains true if the size of the universe outside changes.

For further information: A. Szabolsci "Quantification", 2010

6.1.3. Symmetry "Q A are B" iff "Q B are A". ????

symmetric quantifiers are: some, no, at least five, exactly three, an even number of, infinitely many; non-symmetric: all, most, at most one-third of the.

Symmetry is a feature of (most of) the quantifiers allowed in so-called existential there sentences

"There are at least five men in the garden is fine" vs. "*There are most men in the garden is not."

6.1.4. Monotonicity A DET is a two argument function. It can be upward (\uparrow) or downward monotonic in any of the two argument. E.g. for the predicate argument: If Pred1 \leq Pred2, then

- $DET(NP)(Pred1) \le DET(NP)(Pred2)$
- $DET(NP)(Pred2) \le DET(NP)(Pred1)$
- when neither of the two holds, the DET is

For instance:

- $\operatorname{all}(\downarrow)(\uparrow)$, $\operatorname{some}(\uparrow)(\uparrow)$, $\operatorname{no}(\downarrow)(\downarrow)$,
- at least five, no more than ten, infinitely many, are monotone in both arguments
- $most(|)(\uparrow)$, at least two-thirds of the(|)(\uparrow)
- Exactly three(|)(|), between two(|,|) and seven(|,|)

when $DET(-)(\uparrow)$ when $DET(-)(\downarrow)$ non-monotone (-)(|)

6.1.5. Effects of Monotonicity

Licensors of NPI Monotonicity is crucial for explaining the distribution of polarity items: "No one will ever succeed" vs. *"Someone will ever succeed"

Constrain on coordination NPs can be coordinated by conjunction and disjunction iff they have the same direction of monotonicity, where but requires NPs of different monotonicity (but it does not allow iteration freely)

- *John or no student saw Jane.
- *All the woman and few men walk.
- John but no student saw Jane.
- All the woman, few men but several student walk

vd. phd thesis con JvB

 $\mathbf{1}$

6.2. Quantifiers from the "language as use" angle

Quantifiers have been studied in details from the FS angle but have been mostly ignored by the empirical based community which has focused on content words.

They have been studied in Pragmatics and Psycholinguistics.

- QP and scalar implicature
- QP and performative action
- QP and anaphora

Conjecture: The results found in the pragmatics and psycholinguistics research suggest that DS can bring highlight on QP uses.

6.2.1. Quantifier Phrases and scalar implicature Determiners like "no", "few", "some", "all", are scalar expressions: they can be order on a scale with respect to the strengths of the information that they convey. Their use involves pragmatic inferences called scalar implicature (Grice 1975):

the participants in a conversation expect that each will tailor their contribution to be as informative as required but no more informative than is required.

One meaning: "Some N": $\{X | N \land X \neq 0\}$, $\lambda X . \exists x . N(x) \land X(x)$ [some \exists and possibly \forall] But different uses:

- R: If you ate some of the cookies, then I won't have enough for the party.
 M: I ate some of the cookies. In fact, I ate all of them. [some ∃ and ∀]
- R: Where are the apples that I bought?M: I ate some of them.

[some \exists but $\neg \forall$]

6.2.2. Positive and Negative Quantifiers Quantifiers can have positive or negative polarity, even when denoting the same vague quantity.

- Positive Polarity: "A few", "quite a few", "many"
- Negative Polarity: polarity: "Few", "very few", "not many".

They have different pragmatic functions: encouraging (occur in positive context) vs. discouraging (occur in negative context) a putative course of action. **6.2.3. Performative utterances with Quantifiers Phrases** Hilton et al (2005): "The logical vocabulary of natural language has its origins in the necessity of achieving successful co-ordination of social interactions through communication, and is very well adapted to communicating perception of risk, danger and opportunity."

Using Austin (1976) distinction, sentences using logic expressions can be conceived as utterance aiming to get the hearer to do things, called "performative utterances" rather than as statements merely describing states of affairs, called "constative" statements. E.g.

- 1. If there are a few wolves in the West valley, then don't hunt there. [not good for the wolves]
- 2. If there are few wolves in the West valley, then don't hunt there. [not good for you]

(1) and (2) as constative utterances have the same truth-values; but as performative utterances create different expectations: (1) "wolves" are seen as meat (something positive) vs. in (2) "wolves" are seen as predator (something negative).

6.2.4. QP and anaphora Some of the fans went to the game.

Reference set those fans who went to the game.

Complement set those fans who did not go to the game.

Positive polarity QP put the focus on the Reference set vs. Negative polarity QP put the focus on the Complement set, but also the Reference set is available for anaphora.

Given the sentences below

- (a) A few of the students attended the lecture. They ...
- (b) Few of the students attended the lecture. They \dots

people continue (a) speaking of properties of the reference set (e.g. "They listen carefully and took notes."), and (b) speaking on the complement set (e.g. "They decided to stay at home instead.")

7. Conclusion

We have seen that:

- Natural language syntactic structures convey information to grasp their semantic.
- Many tasks involving natural language need some level of semantic understanding. (e.g. NLDB, QA, TE).
- FOL based systems can be used for domain specific and controlled natural language applications.
- For open domain and free text application, needs to combine relational representation and ML methods
- Logical words have been mostly ignored both by the shallow system developers and by the empirically driven theoreticians.

Tomorrow we go back to Montague framework but from an empirical view.