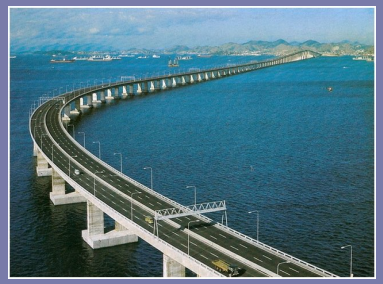


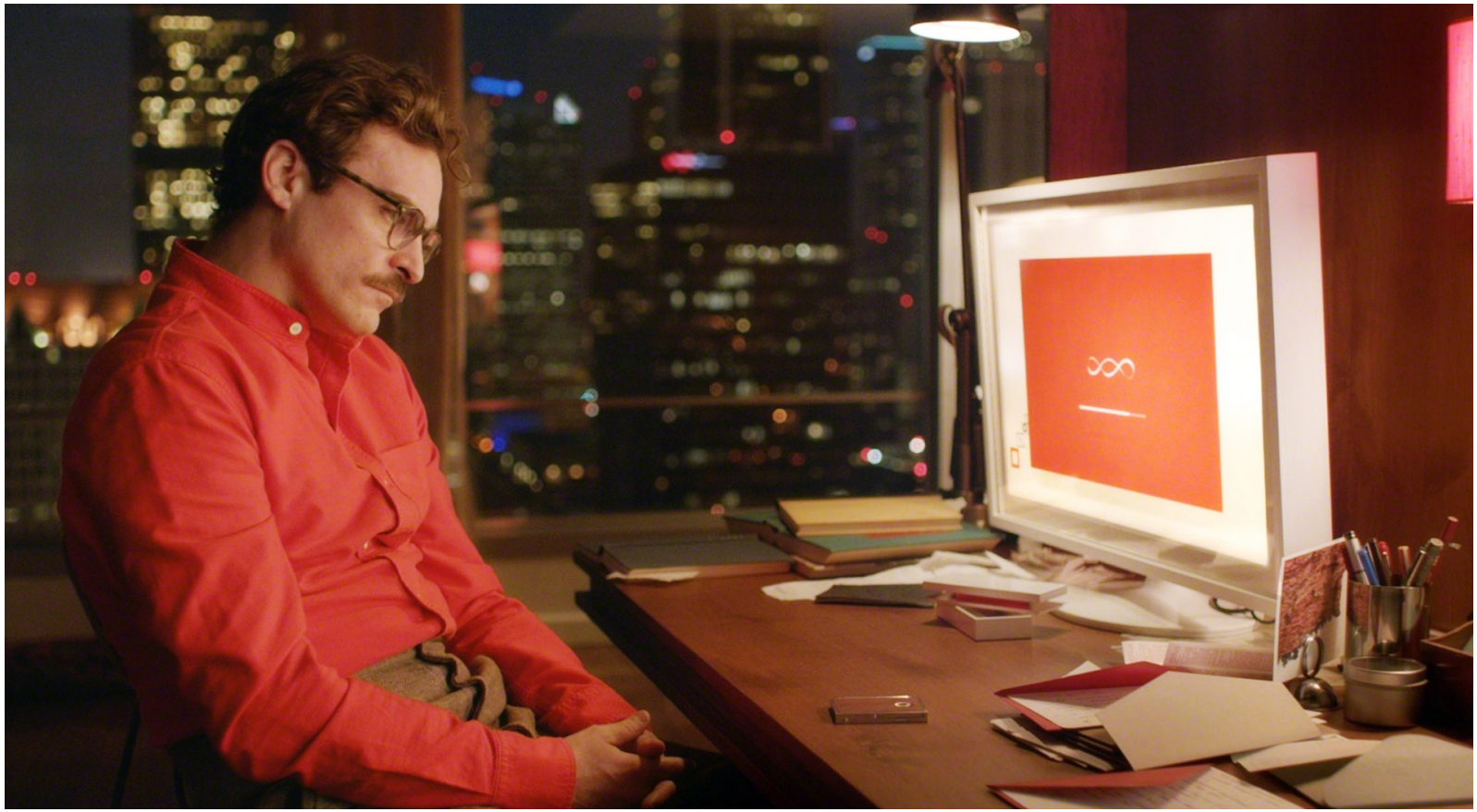
Lean Logic for Lean Times: Entailment and Contradiction Revisited

Valeria de Paiva
Nuance Communications,
NL and AI Lab, Sunnyvale, CA

Nuance Comms,
AI and NL Lab, Sunnyvale, CA

Ron Kaplan "**Beyond the GUI: It's Time for a Conversational User Interface**", Wired 2013

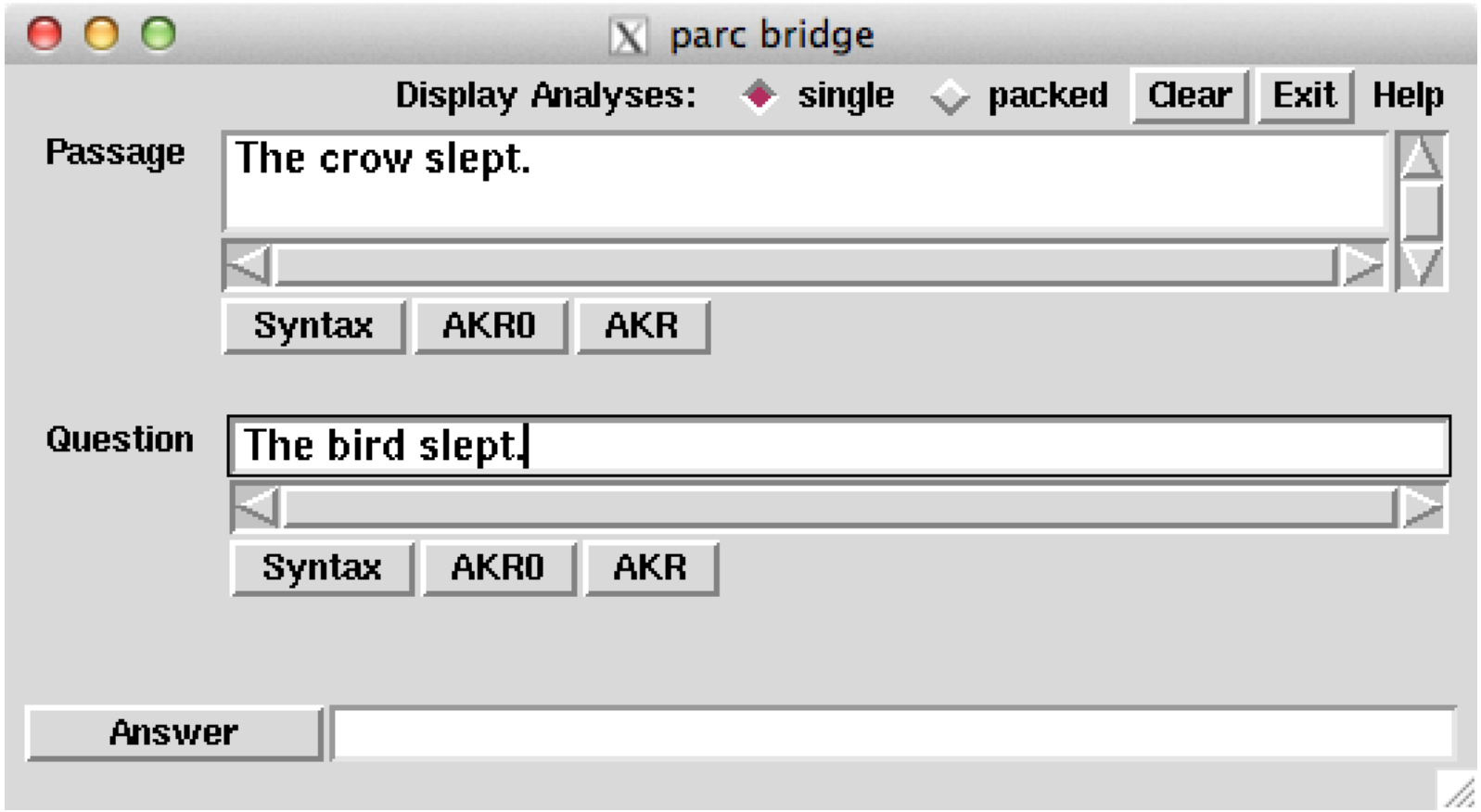
The Future is Meaning…



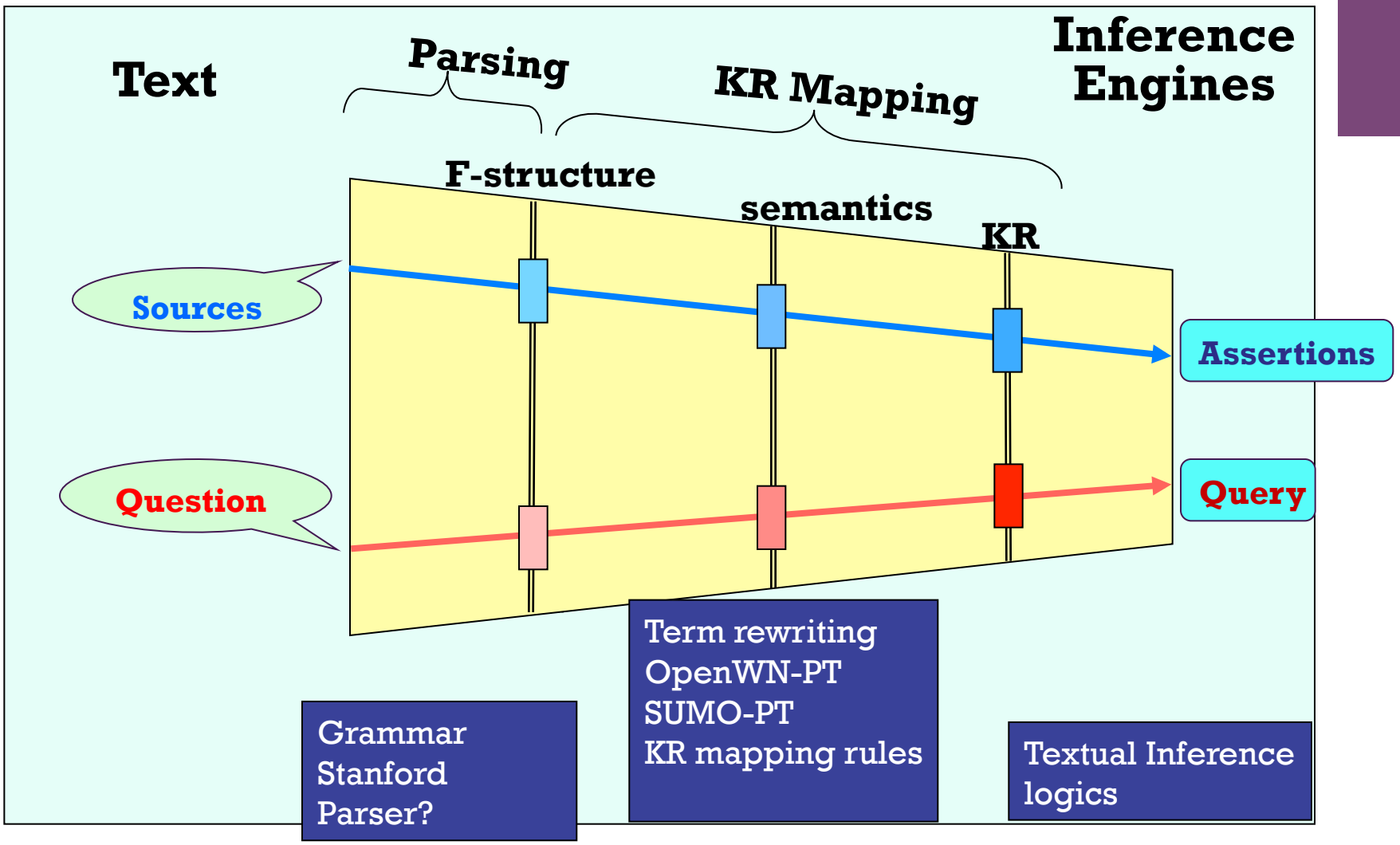
The difficulty is to decide how to get there
CAVEAT EMPTOR

Not so recent past…

PARC's Bridge System (2000-2008 for me)



PARC's Bridge System (2000-2008)



Powerset

Acquired by Microsoft, 2008

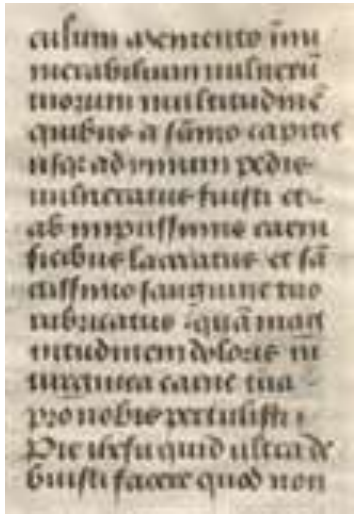
and Cuil…


cuil
Search
Search 121,617,892,992 web pages
About Cuil | Your Privacy
© 2008 Cuil, Inc

Another story

https://www.parc.com/event/934/adventures-in-searchland.html

Goals in 2010: a bridge



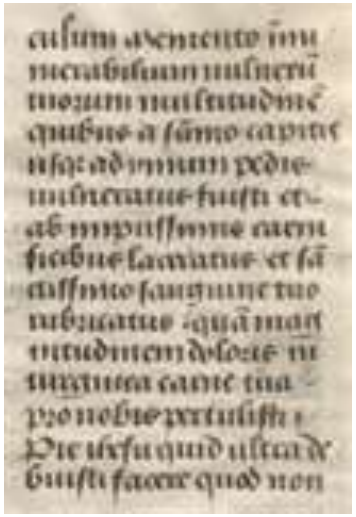
Improve Lexical Resources and Inferential Systems to work with Logic coming from free form text.

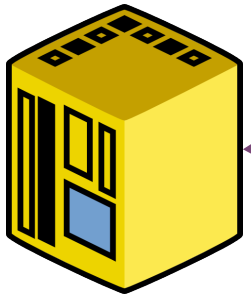
+ Goals in 2010: a bridge

Preprocessing

Grammatical features

Linguistic semantics





Automatic theorem Prover/system

logic

Reality Check…

- Pre-processing is MOST of the processing!
- XLE research license, but hard to use, hard to modify, hard to understand decades of code. Besides, need several lexicons that DO NOT exist openly or for Portuguese, notably WordNet.
  - Spent the next four years working on OpenWordNet-PT.
  - Google Translate, Open MultiLingual Wordnet, BabelNet, FreeLing use our Portuguese wordnet.
- There are several open toolkits that could be used for the processing needed. More usable, more community, less expertise required:

  StanfordNLP

  OpenNLP

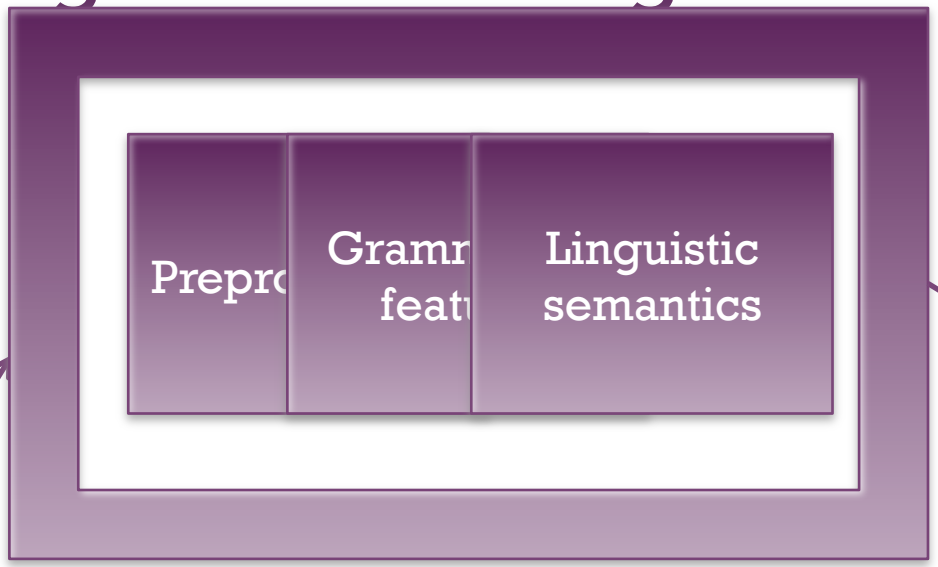
  NLTK

  FreeLing (English and Spanish…)

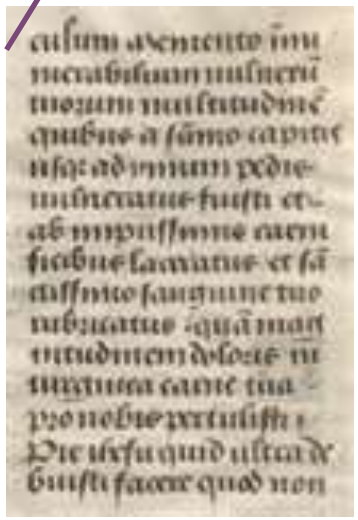
Reasoning/Inference

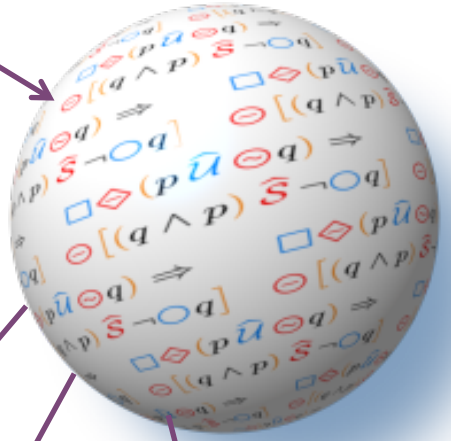
- **Which kind?**
- **Textual entailment** methods recognize, generate, and extract pairs ⟨T,H⟩ of natural language expressions, such that a human who reads (and trusts) T would infer that H is most likely also true (Dagan, Glickman & Magnini, 2006)
- Example:
  (T) The drugs that slow down Alzheimer's disease work best the earlier you administer them.
  (H) Alzheimer's disease can be slowed down using drugs.
  $T \Rightarrow H$
- A series of competitions since 2004, ACL "Textual Entailment Portal", many different systems…

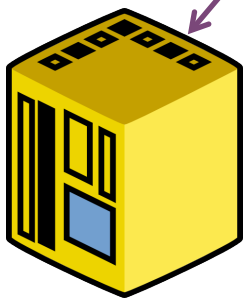
Bridge: Text to Logic

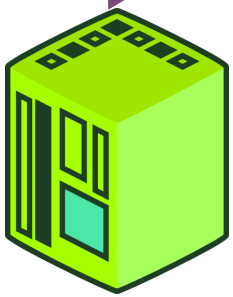


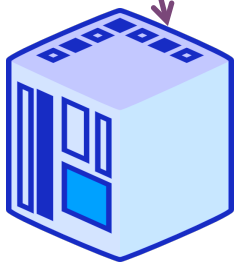
Linguistics BlackBox











...

Cyc

UL+ECD

KIML+Maude

Model Theoretic and Proof Theoretic semantics

- Introductory courses on semantics for natural languages talk about creating **representations** for the sentences in a logical formalism (the **logical forms**) and uncovering **truth conditions** for these translated sentences.
- This is Model Theoretic semantics
- An alternative view, the ***proof-theoretic paradigm*** of semantics claims that the most basic criterion is to establish when sentences follow from others, when they are consistent with each other, when they contradict each other. In short their entailment behavior. (cf. Nissim Francez)
- Relations of entailment and contradiction are the key data of semantics, as traditionally viewed as a branch of linguistics. The ability to recognize such semantic relations is clearly **not a sufficient criterion** for language understanding: there is more than just being able to tell that one sentence follows from another. But we would argue that it is a **minimal, necessary criterion**.
- Hence **Lean Logic**

Inference Rules

Inheritance rules

$$\frac{\text{Nina has a canary, canary} \sqsubseteq \text{bird}}{\text{Nina has a bird}} \qquad \frac{\text{Ed kissed Nina, kiss} \sqsubseteq \text{touch}}{\text{Ed touched Nina}}$$

$$\frac{\text{Every carp is a fish, carp} \sqsupseteq \text{koi}}{\text{Every koi is a fish}}$$

$$\frac{\text{She didn't give him a bird, bird} \sqsupseteq \text{canary}}{\text{She didn't give him a canary}}$$

Dowty/Hoeksema examples

Logical languages?

- Assuming you believe that text can be transformed coherently and in a principled fashion into logical formulas…
- There are still many design choices to make:
  - Which logical language? FOL vs. HOL, modal or not, DRT, lambda-calculus or not?
  - Which logical system?
- Logical language and logical system tend to get confused, but there are many logical systems that use the same language, FOL, for instance.
- The same logical system, say IPL, intuitionistic propositional logic can be represented using many languages e.g. axioms, sequents, Natural Deduction, etc.

Knowledge Inference Management Language (KIML)

- A representation language based on events, concepts, roles and contexts, McCarthy-style
- Using events, concepts and roles is traditional in NL semantics (Lasersohn)
- Usually equivalent to FOL (first-order logic), ours a small extension, contexts are like modalities.
  Language based on PARC linguists' intuitions
- Exact formulation of system still being decided: e.g. not considering temporal assertions, yet…
- (used to call both language and logic AKR, abstract knowledge representation)

Example: a crow slept

Conceptual Structure:
role(cardinality restriction,crow-1,sg)
role(sb,sleep-4,crow-1)

subconcept(crow-1,
[crow#n#1,crow#n#2,brag#n#1])
subconcept(sleep-4,[sleep#v#1,sleep#v#2])
Contextual Structure:
instantiable(crow-1,t)
instantiable(sleep-4,t)
top context(t)
Temporal Structure:
trole(when,sleep-4,interval(before,Now))

KIML versus FOL

- In FOL write ∃ Crow ∃ Sleep.Sleep(crow)
  Instead we will use basic concepts from a parameter ontology *O*
- *O* (can be Cyc, SUMO, UL, WN, DBPedia, Freebase, etc…)
- Instead of FOL predicates have Skolem constants crow-1 a subconcept of an ambiguous list of concepts:
  subconcept(crow-1,[crow#n#1,crow#n#2,brag#n#1])
- Same for sleep-2 and have roles relating concepts
  role(sb,sleep-4,crow-1)
  meaning that the sb=subject of the sleeping event is a crow concept

What is Different?

- Corresponding to formulas in FOL, KIML has a collection of assertions that, read conjunctively, correspond to the semantics of (fragments of) sentences in English.
- Concepts in KIML – similar to Description Logic concepts primitive concepts from an idealized version of the chosen
- Ontology on-the-fly concepts, always sub-concepts of some primitive concept. concepts are as fine or as coarse as needed by the application
- Roles connect concepts: deciding which roles with which concepts a big problem… for linguists
- Roles assigned in a consistent, coherent and maximally informative way by the NLP module

Contexts for Quantification

- Contexts for modelling **negation**, **implication**, as well as **propositional attitudes** and other intensional phenomena. (Similar to modal DRT)
- There is a first initial context (written as t), roughly what the author of the sentence takes the world to be.
- Contexts used for making existential statements about the existence and non-existence in specified possible worlds of entities that satisfy the intensional descriptions specified by our concepts.
- Propositional attitudes predicates (knowing, believing, saying,...) relate contexts and concepts in our logic.
- Concepts like knowing, believing, saying introduce context that represents the proposition that is known, believed or said.

Ed knows that the crow slept

alias(Ed-0,[Ed])
role(prop,know-1,ctx(sleep-8))
role(sb,know-1,Ed-0)
role(sb,sleep-8,crow-6)
subconcept(Ed-0,[male#n#2]) subconcept(crow-6,
[crow#n#1,crow#n#2,brag#n#1])
subconcept(know-1,[know#v#1,...,sleep-
together#v#1]) subconcept(sleep-8,
[sleep#v#1,sleep#v#2]) context(ctx(sleep-8)),
context(t) context-lifting-
relation(veridical,t,ctx(sleep-8)) context-
relation(t,ctx(sleep-8),crel(prop,know-1))
instantiable(Ed-0,t)
instantiable(crow-6,ctx(sleep-8))
instantiable(sleep-8,ctx(sleep-8))

Inference to build reps and to reason with them

- In previous example can conclude:
  instantiable(sleep-8,t)
  if knowing X implies X is true.
  (Can conclude instantiable(crow-6,t) too, for definitiveness reasons..)
- Happening or not of events is dealt with by the instantiability/ uninstantiability predicate that relates concepts and contexts e.g. Negotiations prevented a strike
- Contexts can be:
  veridical, antiveridical or averidical with respect to other contexts.
- Have 'context lifting rules' to move instantiability assertions between contexts. (MacCarthy style)
- Want to describe the logic system making these rules work. HOW?

Applied logician's job as a puzzle

- When modeling a system as a logic you can start from the implemented system…
- Or you can start from a logic that looks similar to the system, that could be a good fit for it.
- Hopefully the two lines of attack converge, eventually

From implemented system to Logic

- *A Basic Logic for Textual inference* (with D. Bobrow, C. Condoravdi, R. Crouch, R. Kaplan, L. Karttunen, T. King and A. Zaenen), AAAI Workshop on Inference for Textual Question Answering, Pittsburgh PA, July 2005.
- Textual Inference Logic: Take Two, (with D. G. Bobrow, C. Condoravdi, R. Crouch, L. Karttunen, T. H. King, R. Nairn and A. Zaenen) *Workshop on Contexts and Ontologies, Representation and Reasoning*, CONTEXT 2007.
- Bridges from Language to Logic: Concepts, Contexts and Ontologies (V. de Paiva) in 5th LSFA'10, Natal, Brazil, 2010.
- Contexts for Quantification, Valeria de Paiva, in Proceedings of CommonSense2013
- Similar to a Natural Logic extension? Dowty199:4 `studying deductive systems which approximate the class of common linguistic inferences to some interesting degree or in some interesting way'
- Moss, Icard, Djalali and Pratt-Hartmann: maybe we can augment/bend "Syllogystic inference"

Problems

- System changes under your feet
- Writing documentation is boring
- Specs don't need to be complete, logical systems do
- Big system, big representations, hard to know what it should do
- Layers of tangled code, personal incompetence
- Mostly: Proprietary code goes away, logic is for all

From Logics to Systems

- Easier and more rewarding. Other people can find **other** applications for the systems you've developed.
- *Natural Deduction and Context as (Constructive) Modality*. Proceedings of the 4th CONTEXT 2003, Stanford
  - *Constructive CK for Contexts* (with Michael Mendler), Worskhop on Context Representation and Reasoning, Paris, France, July 2005.
- *(Towards) Constructive Hybrid Logic* (with T. Brauner), Methods for Modalities 3, LORIA, Nancy, France, September, 2003.(2006)
- Constructive Description Logics: what, why and how. Context Rep and Reasoning, Riva del Garda, 2006.
  - Constructive Description Logics Hybrid-Style.(with Rademaker, Hauesler) Electr. Notes Theor. Comput. Sci. 273: 21-31 (2011)
  - Contextual Constructive Description Logics (with Natasha Alechina), preprint

TODAY: Textual Entailment and Connexive Logic

- Connexive logic is motivated by ideas of coherence or 'connection' between premises and conclusions of valid inferences.
- No, I had never heard of it till very recently. Wansing has an updated entry (2014) in SEP about it
- I've known about Relevant Logic for a while. Even had a phd student trying to produce categorical models for it.
- But I used to be a purist. Wanted beautiful proof theory and clean, principled logic systems. More pragmatic now.
- Can do a relevant lambda-calculus, λ I-calculus easily

Implicational Relevant Logic

$$\Gamma, x\colon A, \Gamma' \vdash x\colon A$$

$$\frac{\Gamma, x\colon A, \Gamma' \vdash t\colon B}{\Gamma, \Gamma' \vdash \lambda x\colon A.t\colon A \to B}\ (x \in FV(t)) \qquad \frac{\Gamma \vdash t\colon A \to B \quad \Gamma \vdash u\colon A}{\Gamma \vdash tu\colon B}$$

Figure 1: Typing Judgements for IRL (implicational relevant logic)

Can we use this to prove "theorem" like the ones next?

Experimental Results: a few `theorems'

- 1. a crow was thirsty ⊢ a thirsty crow
- 2. a thirsty crow ⊢ a crow
- 3. ed arrived and the crow flew away ⊢ the crow flew away
- 4. ed knew that the crow slept ⊢ the crow slept
- 5. ed did not forget to force the crow to fly ⊢ the crow flew
- 6 the crow came out in search of water ⊢ the crow came out
- 7. a crow was thirsty ⊢ a bird was thirsty

Maude implementation of Entailment and Contradiction, with Vivek Nigam LSFA 2014

Which Inference Rules?

The totally obvious

$$\frac{}{s \to s} \qquad \frac{s \to t \qquad s \to r}{r \to t}$$

Inference Rules

Inheritance rules

$$\frac{\text{Nina has a canary, canary} \sqsubseteq \text{bird}}{\text{Nina has a bird}} \qquad \frac{\text{Ed kissed Nina, kiss} \sqsubseteq \text{touch}}{\text{Ed touched Nina}}$$

$$\frac{\text{Every carp is a fish, carp} \sqsupseteq \text{koi}}{\text{Every koi is a fish}}$$

$$\frac{\text{She didn't give him a bird, bird} \sqsupseteq \text{canary}}{\text{She didn't give him a canary}}$$

Dowty/Hoeksema examples

Modifiers Inference Rules

$$\frac{\text{Ed arrived in the city by bus}}{\text{Ed arrived in the city}} \qquad \frac{\text{Ed did not arrive in the city}}{\text{Ed did not arrive in the city by bus}}$$

$$\frac{\text{Ed arrived in the city, Ed} \sqsubseteq \text{person}}{\text{A person arrived in the city}} \qquad \frac{\text{Ed arrived in Rome, Rome} \sqsubseteq \text{city}}{\text{Ed arrived in a city}}$$

Note that *Ed did not arrive in the city by bus* does **not** entail that *Ed did not arrive in the city*.

Within Contexts

- MacCartney and Manning's Relations

symbol	name	example	definition
a ≡ b	equiv	sofa ≡ couch	a ⊆ b and b⊆a
a ⊏ b	forw ent	bulldog ⊏ dog	a ⊆ b
a ⊐ b	back ent	fish ⊐ carp	a ⊇ b
a N b	negation	bat N non-bat	a∩ b= ∅ and $a \cup b = U$
a‖ b	alternance	cat ‖dog	a∩ b= ∅ and $a \cup b \neq U$
a⌣ b	cover		a∩ b≠ ∅ and $a \cup b = U$
a♯ b	indep		all other cases

Within Contexts

- A constructive version of Djalali's Synthetic Logic

R, S	$\sqsubset$	$\sqsupset$	N	$\parallel$
$\sqsubset$	$\sqsubset$	·	$\parallel$	$\parallel$
$\sqsupset$	·	$\sqsupset$	·	·
N	·	$\parallel$	$\equiv$	$\sqsupset$
$\parallel$	·	$\parallel$	$\sqsubset$	·

Within contexts

Let Γ be a set of $\mathcal{L}$ formulas. Then the rules of the calculus are of the form below, where R, S and T are taken from the table above.

$$\frac{\Gamma \vdash aRb \qquad \Gamma \vdash bSc}{\Gamma \vdash aTc}$$

The following proof rules are also included in the calculus:

$$\frac{}{\Gamma \vdash a \equiv a} \qquad \frac{\Gamma \vdash a \equiv b}{\Gamma \vdash b \equiv a} \qquad \frac{\Gamma \vdash a \equiv b \qquad \Gamma \vdash b \equiv c}{\Gamma \vdash a \equiv c}$$

$$\frac{\Gamma \vdash a \sqsubset b}{\Gamma \vdash b \sqsupset a} \qquad \frac{\Gamma \vdash a \sqsupset b}{\Gamma \vdash b \sqsubset a} \qquad \frac{\Gamma \vdash a \| b}{\Gamma \vdash b \| a}$$

$$\frac{\Gamma \vdash a\mathbf{N}b}{\Gamma \vdash b\mathbf{N}a} \qquad \frac{a \in \Gamma}{\Gamma \vdash a}$$

Unprincipled hack?

- Perhaps. But I'm not the first one to do it.

Definition 2.4 (Inference rules) Inference rules of GS are defined as follows:

$$\frac{}{s \sqsubset s}\ ax \qquad \frac{s \sqsubset t \quad t \sqsubset u}{s \sqsubset u}\ (\sqsubset) \qquad \frac{s \sqsubset t \quad t \vdash u}{s \vdash u}\ (\mathsf{H}) \qquad \frac{\mathcal{P} \quad \mathcal{Q}}{\mathcal{P} \cup \mathcal{Q}}\ (+) \qquad \frac{\mathcal{P}}{\mathcal{P}'}\ (-) \text{ for } \mathcal{P}' \subsetneq \mathcal{P}$$

For $(+)$ rule, we assume that $\mathcal{P}$ and $\mathcal{Q}$ are different sets, i.e., $\mathcal{P} \neq \mathcal{Q}$.

Since a formulas $\mathcal{P} = \{P_1, \ldots, P_n\}$ means the conjunction $P_1 \wedge \cdots \wedge P_n$, $(+)$ and $(-)$ rules can be considered as generalizations of usual $\wedge$-introduction and $\wedge$-elimination rules of Gentzen's natural deduction.

Generalized Syllogistic Inference System
based on Inclusion and Exclusion Relations
Koji Mineshima, Mitsuhiro Okada, and Ryo Takemura
Department of Philosophy, Keio University. 2010

Implicative Commitment Rules

- Preserving polarity:
  "Ed managed to close the door" → "Ed closed the door"
  "Ed didn't manage to close the door" → "Ed didn't close the door".

- The verb "forget (to)" inverts polarities:
  "Ed forgot to close the door" → "Ed didn't close the door"
  "Ed didn't forget to close the door" → "Ed closed the door".

- There are six such classes, depending on whether positive environments are taken to positive or negative ones. (Nairn, Condoravdi and Kartunnen 2006)

- Accommodating this fine-grained analysis into traditional logic description is further work. (Nairn et al 2006 presents an implemented recursive algorithm for composing these rules)

Contexts as Modalities again

$$\Gamma, x\colon A, \Gamma' | \Delta \vdash x_M\colon A \qquad \Gamma | \Delta, x\colon A, \Delta' \vdash x_I\colon A$$

$$\frac{\Gamma | \Delta, x\colon A \vdash t\colon B}{\Gamma | \Delta \vdash \lambda x\colon A.t\colon A \to B}\ (x \in FV(t)) \qquad \frac{\Gamma | \Delta \vdash t\colon A \to B \quad \Gamma | \Delta \vdash u\colon A}{\Gamma | \Delta \vdash tu\colon B}$$

$$\frac{\Gamma | \_ \vdash t\colon A}{\Gamma | \Delta \vdash \Box t\colon \Box A} \qquad \frac{\Gamma | \Delta \vdash t\colon \Box A \quad \Gamma, x\colon A | \Delta \vdash u\colon B}{\Gamma | \Delta \vdash \texttt{let}\ t\ \texttt{be}\ \Box x\ \texttt{in}\ u\colon B}$$

Figure 2: Typing Judgements for 'dual relevant and modal logic'

Need more typing here.
Criteria for success?

Conclusions

- Proof-of-concept framework
- KIML assertions and TIL inference system for textual entailment
- Relevant intuitionistic modal logic for ECD
- A framework can be implemented in Maude and used to prove in an semi-automated fashion whether a sentence follows from another (LSFA2014)
- 'shallow theorem proving' for common sense applications?
- Many problems**: black box**, ambiguity, temporal information, paraphrasing, etc..
- Applied logic needs empirical justification
- After all these years, work is still only starting to check this route
- (Newer, more sophisticated work in Stanford and other places carries on Angeli, Bowman, MacCartney, Bos, etc..)

Thanks!

References

OpenWordNet-PT: An Open Brazilian WordNet For Reasoning. de Paiva, Valeria, Alexandre Rademaker, and Gerard de Melo. In *Proceedings of the 24th International Conference On Computational Linguistics*. http://hdl.handle.net/10438/10274.

Embedding NomLex-BR Nominalizations Into OpenWordnet-PT. Coelho, Livy Maria Real, Alexandre Rademaker, Valeria De Paiva, and Gerard de Melo. 2014. In *Proceedings of the 7th Global WordNet Conference*. Tartu, Estonia.

Bridges from Language to Logic: Concepts, Contexts and Ontologies Valeria de Paiva (2010) Logical and Semantic Frameworks with Applications, LSFA'10, Natal, Brazil, 2010.

`A Basic Logic for Textual inference", AAAI Workshop on Inference for Textual Question Answering, 2005.

``Textual Inference Logic: Take Two", CONTEXT 2007.

``Precision-focused Textual Inference", Workshop on Textual Entailment and Paraphrasing, 2007.

PARC's Bridge and Question Answering System Proceedings of Grammar Engineering Across Frameworks, 2007.

Towards a Maude Rewriting Framework for Textual Entailment

- Maude is an implementation of **rewriting logic** developed at SRI and Illinois.
- Maude modules (rewrite theories) consist of a term-language plus sets of equations and rewrite-rules. Terms in rewrite theory are constructed using operators (functions taking 0 or more arguments of some sort, which return a term of a specific sort).
- Hand-correct the representations given by the NLP module: goal is not to obtain correct representations, but to work logically with correct representations
- Thus get an implementation of TIL, using the traditional rewriting system Maude to reason about the logical representations produced by the BlackBox module we are considering.

A Rewriting Framework

- A rewrite theory is a triple (Σ ,E,R), with (Σ ,E) an equational theory with Σ a signature of operations and sorts, and E a set of (possibly conditional) equations, and with R a set of (possibly conditional) rewrite rules.
- A few logical predicates for our natural languages representations: (sub)concepts, roles, contexts and a few relations between these.
- But the concepts that the representations would use in a minimally working system in the order of 150 thousand, concepts in WordNet.
- Scaling issues?

Maude Rewriting

- Basic rewriting sorts: Relations, SBasic and UnifSet
- TIL basic assertions such as canary $\sqsubseteq$ bird belong to Relations.
- Concept and contextual assertions, such as instantiable(drink-0,t) belong to the SBasic basic statements sort.
- The third basic sort, UnifSet, contains unification of skolem constants, such as crow-6 := bird-1. This last sort is necessary for for unifying skolem constants.

BlackBox Inference?

- Can use Xerox's PARC Bridge system as a black box to produce NL representations of sentences in KIML (Knowledge Inference Management Language).
- KIML + inference rules = TIL (Textual Inference Logic)
- Translate TIL formulas to a theory in Maude, the SRI rewriting system.
- Use Maude rewriting to prove Textual Entailment "theorems".