Probabilistic Representations of Linguistic Meaning ("PReLiM")

Jason Eisner

JOHNS HOPKINS UNIVERSITY

Jelinek Memorial Workshop, July 2014

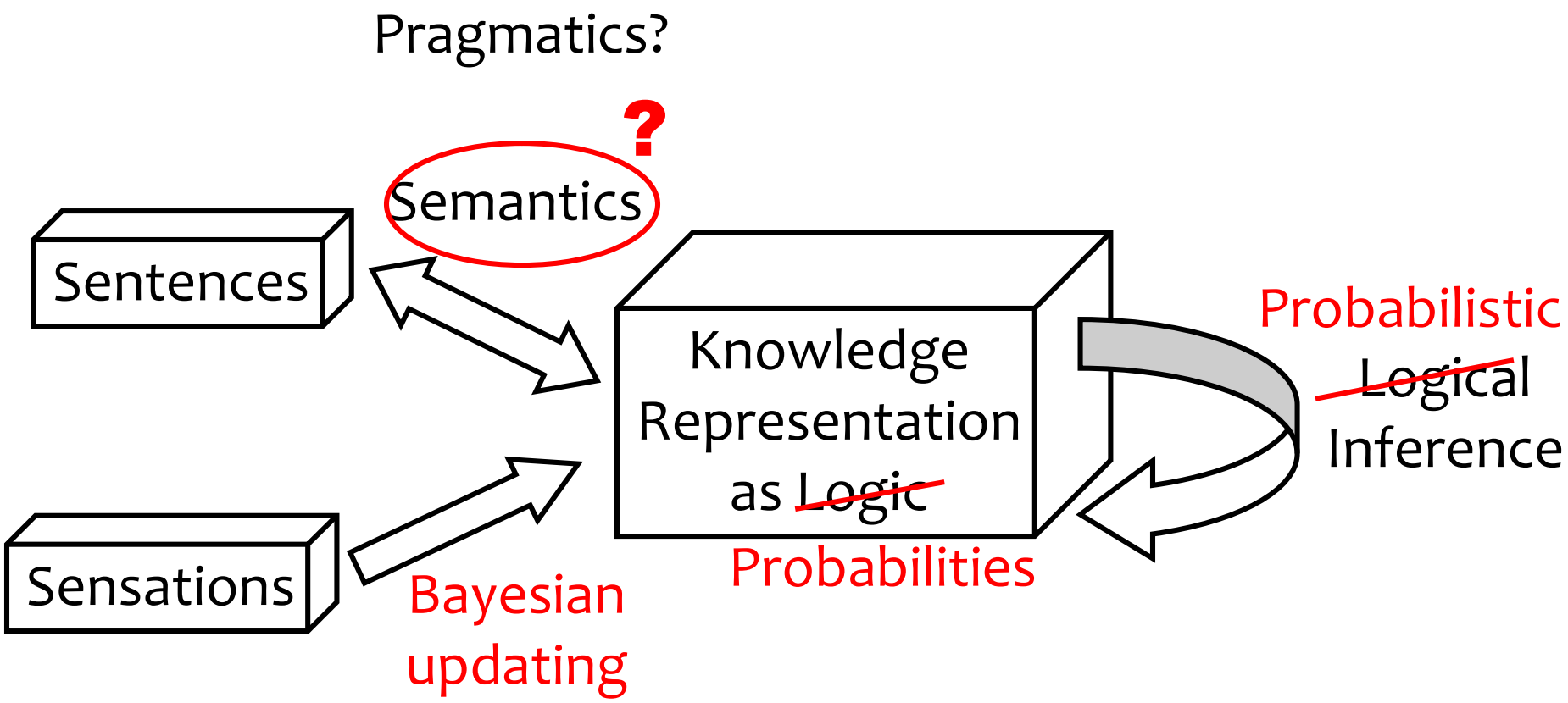
Week 1: A blue-sky workshop

- What should computational semantics look like in 10 years?

- More of what semanticists think about
- Integrated with reasoning and background knowledge ("language of thought")
- Integrated with pragmatics
- Fuller probabilistic treatment (generative)

Why a "probabilistic treatment"?

Reason 1: Integration

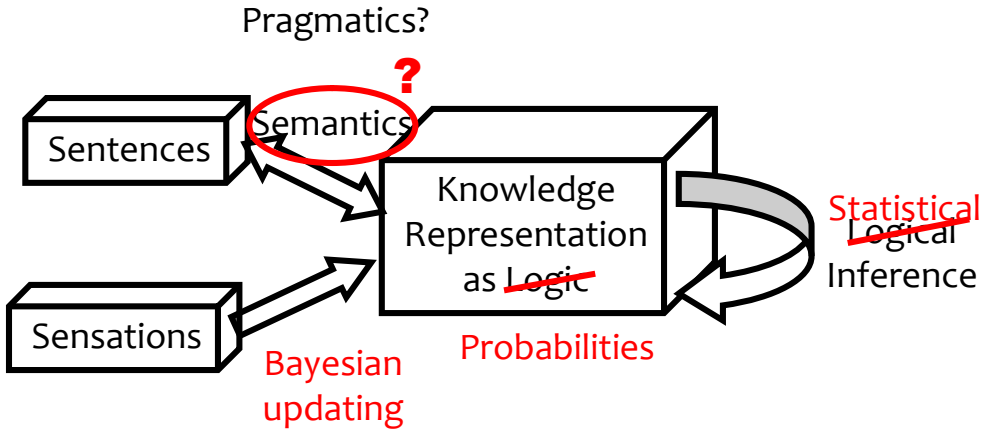


Why a "probabilistic treatment"?

Reason 1: Integration

- This is going to be harder than just Bayes nets.
  - What is the probabilistic version of *modal* logic?
  - How do we deal with richly *structured* beliefs?
  - How about *incomplete*, *vague*, or *contradictory* ones?

- We may need some new ideas about probabilistic modeling.



Why a "probabilistic treatment"?

Reason 1: Integration

- Natural language is related to the "language of thought."
- Fillmore 1982:

  "This sentence did not give you this information directly; you had to 'compute' some of it by constructing, in your imagination, a complex context within which each of the lexically signaled framings was motivated."

  (*"The decedent while on land and in mufti last weekend ate a typical breakfast and read a novel high in flip strength."*)

- What happens in a car crash?
- Think of a widow. How old is she?
  - How do you know?
  - Compositional semantics: How about "the widow of the firefighter"?
  - "The widow of the village chieftain"?

Why a "probabilistic treatment"?

Reason 2: Stochasticization has paid off for computational linguistics

- For some decades, computational linguistics has been trying to reinvent linguistics "so it works"
- From grammars to probability distributions (from "what's possible?" to "what's probable?")
- Starting with phonetics and collocations

Thanks, Fred

JOHNS HOPKINS UNIVERSITY

Why a "probabilistic treatment"?

Reason 2: Stochasticization has paid off for computational linguistics

- **Phonetics:** Gaussian mixtures, n-grams, …
- **Phonology, morphology:** probabilistic FSTs, stochastic OT, graphical models over many strings, …
- **History of words:** probabilistic evolutionary models
- **Word collocations:** n-gram models, topic models, embeddings, …
- **Syntax:** PCFGs, selectional preference models, …
- **Translation:** synchronous grammars, …

Model the joint distribution of elements of form and meaning → Algorithms for robust comprehension, production, and learning

Why a "probabilistic treatment"?

Reason 2: Stochasticization has paid off for computational linguistics

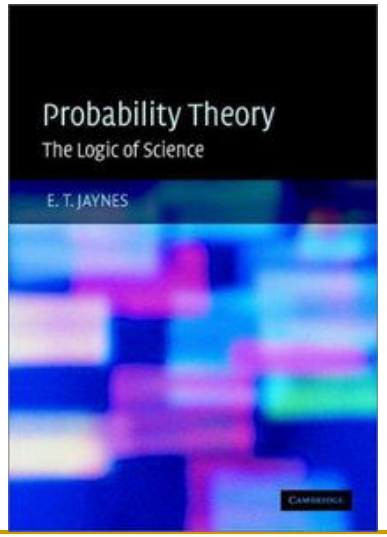
- This is because **knowledge of language** includes **knowledge of probabilities.**

- It's amazing that we can communicate so effectively at such high bandwidth!
- Seems to require prior probabilities to help reconstruct what is left out.
  - Certainly for syntax … and also for semantics.

Why a "probabilistic treatment"?

Reason 2: Stochasticization has paid off for computational linguistics

- Models get richer in linguistic insight over time.
  - The goal of science is to find the underlying probability distributions that can explain and predict our observations.



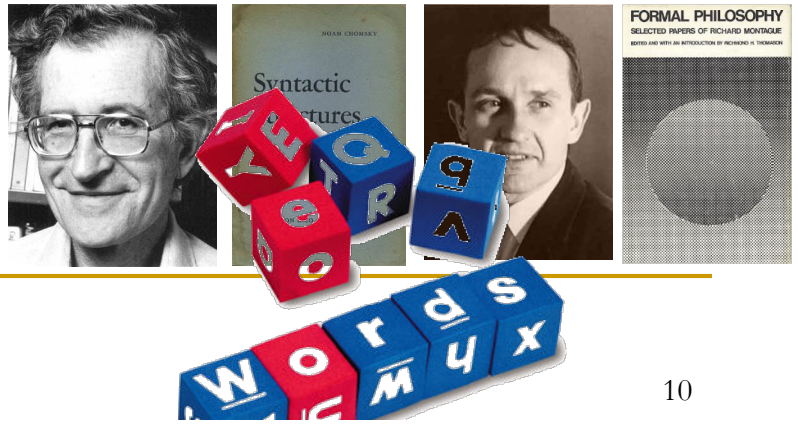
E. T. Jaynes (2003), Probability Theory: The Logic of Science

JOHNS HOPKINS UNIVERSITY

Why a "probabilistic treatment"?

Reason 2: Stochasticization has paid off for computational linguistics

- Models get richer in linguistic insight over time.
  - The goal of science is to find the underlying probability distributions that can explain and predict our observations.
  - Linguists shouldn't be alarmed. "Do not think that I have come to abolish the Law or the Prophets; I have not come to abolish them but to fulfill them."



Why a "probabilistic treatment"?

Reason 2: Stochasticization has paid off for computational linguistics

- Models get richer in linguistic insight over time.
  - The goal of science is to find the underlying probability distributions that can explain and predict our observations (Jaynes 2003).
  - Linguists shouldn't be alarmed. "Do not think that I have come to abolish the Law or the Prophets; I have not come to abolish them but to fulfill them."
  - But we're only human so it takes time. ☺
  - And in fact, it helps to start simple.

Why a "probabilistic treatment"?

Reason 2: Stochasticization has paid off for computational linguistics

- Models get richer in linguistic insight over time.
  - Helps to start simple.

- Case study: PCFGs
  - Formalism just tries to capture the most important things about syntax: phrase types, hierarchical structure
  - **Flexible:** no commitment to particular nonterminals or rules
  - Explanatory adequacy? *(sufficiently strong inductive bias?)*
    - Missing a lot of things that probably belong in UG
    - But you can get more specific theories via priors over grammars
  - Descriptive adequacy? *(sufficiently weak inductive bias?)*
    - Maybe not, but the formal treatment points the way to more complicated variants (TAG, HPSG, continuous nonterminals …)

Why a "probabilistic treatment"?

Reason 2: Stochasticization has paid off for computational linguistics

- Models get richer in linguistic insight over time.
  - Helps to start simple.
- Case study: PCFGs
  - Case study: (stochastic) Optimality Theory
    - Formalism just tries to capture the most important things about phonology: interaction among constraints
    - **Flexible:** no commitment to particular constraints or representations

Why a "probabilistic treatment"?

Reason 2: Stochasticization has paid off for computational linguistics

- Models get richer in linguistic insight over time.
  - Helps to start simple.
- Case study: PCFGs
  - Case study: (stochastic) Optimality Theory
    - Case study: probabilistic semantics?
      - What should we capture initially?
      - What are the basic kinds of objects we're dealing with?
      - How do they fit together?
      - We might not handle verb aspect right away, or physical predicates; but there should be somewhere to fit them in

Why a "probabilistic treatment"?

Reason 3: Lots of specific semantic/pragmatic phenomena …

- You're probably going to lose now, since that rook is pinned.
- She may have left. Her car is gone.
- I should cancel the milk.
- Children must be carried. *[sign on the escalator]*
- Most parents are likelier to give in on weekends.
- If you can afford a cheap car, you can afford expensive coffee.
- If you had poured the coffee, it wouldn't have spilled.

Participants

"Big, architectural thinkers"

- Jason Eisner, JHU (CS/cogsci)
- Oren Etzioni, UW/Allen Institute (CS)
- Shalom Lappin, KCL (philosophy)
- Dan Lassiter, Stanford (psycholinguistics)
- Percy Liang, Stanford (CS/stats)
- Staffan Larsson , Gothenburg (philosophy/linguistics)
- David McAllester, TTI-Chicago (CS)
- James Pustejovsky, Brandeis (linguistics)
- Benjamin Van Durme, JHU (CS/cogsci)

- JHU students:
  - Nick Andrews (CS)
  - Drew Reisinger (cogsci)
  - Darcey Riley (CS)
  - Rachel Rudinger (CS)

Virtual Participants

Couldn't be here this week but asked to be involved

- Joshua Tenenbaum, MIT (cogsci)
- Noah Goodman, Stanford (psych)

- Barbara Partee, UMass (ling)
- Gerhard Jaeger, Tübingen (ling)
- Michael Franke, Amsterdam (ling)
- Christopher Potts, Stanford (ling)
- Kyle Rawlins, JHU (ling)

- Robin Cooper, Gothenburg (phil)
- Igor Douven, Groningen (phil)

- Luke Zettlemoyer, UW (CS)
- Ido Dagan, Bar-Ilan (CS)
- Len Schubert, Rochester (CS)
- Dan Klein, Berkeley (CS)

- JHU students:
  - Frank Ferraro (CS)
  - Pushpendre Rastogi (CS)

(Kyle and Frank are joining next week)

Setting the stage

- **Mon am: Introductions**
  - Our goals and interests
  - Our desiderata and warnings of pitfalls
- **Mon pm:** James Pustejovsky (plenary talk)
  - Why it is important to distinguish "possible" from "probable" meaning shifts: How distributions impact linguistic theory
- **Tue am: Taking stock**
  - Hard examples
  - What's already understood
  - Planning our time

Towards a probabilistic language of thought

- **Tue am:** Shalom Lappin (plenary talk)
  - A Rich Probabilistic Type Theory for the Semantics of Natural Language
- **Tue pm:** Knowledge representation
  - Belief, theory of mind
  - Metaphor and meaning shift
  - Chalktalk by Darcey/Jason on locally renormalized PCFG?
- **Tue pm:** Oren Etzioni (plenary talk)
  - Semantics, Science, and 10-Year-Olds
- **Wed am:** Worlds and situations
  - Generics, quantifiers
  - Modals, conditionals, counterfactuals
  - Chalktalk by Drew on dialogue scenario?

Pragmatics

- **Wed am:** Dan Lassiter (plenary talk)
  - Bayesian Pragmatics
- **Wed pm:** Pragmatics
  - Meta-reasoning (chalktalk by Dan?)
  - Presuppositions and implicatures
  - Game theory

Linguisticization

- **Wed pm:** David McAllester (plenary talk)
  - The Problem of Reference
- **Thu am:** Linguisticization
  - Lexical semantics, event semantics
  - Framing
  - Linguistic marking
    - definiteness, information structure, modality, evidentials, classifiers, conventional implicatures

Grounding

- **Thu am:** Staffan Larsson (plenary talk)
  - Perceptual Semantics and Coordination in Dialogue
- **Thu pm:** Grounding
  - Perception
  - Vagueness
  - Temporal and spatial reasoning

Making it happen

- **Thu pm:** Percy Liang (plenary talk)
  - The State of the Art in Semantic Parsing
- **Fri am:** Remaining difficult issues
  - E.g., imprecise language, contradictory beliefs, linguistic ambiguity about contrast sets
- **Fri am:** Martha Palmer (plenary talk)
  - Designing Abstract Meaning Representations for Machine Translation
- **Fri pm:** Practical next steps toward "semantic AI"
  - Chalktalks by Rachel, Nick?

- **Fri pm:** Ben Van Durme (plenary talk)
  - Reexamining Semantic Roles

Goals

- What are the constraints on a full theory?
  - Listen closely to each other …
- Can we agree on a baseline theory (like PCFG)?
  - Or at least rule out directions that won't work?
- What conceptual / mathematical work still needs to be done?
- What first steps can we take toward actually building semantic AI?
  - Starting with the remaining 3 weeks (student work on PReLiM; Martha's CLAMR group)

Week 1: A blue-sky workshop

- What should computational semantics look like in 10 years? Let's get the foundations right.

- More of what semanticists think about
- Integrated with reasoning and background knowledge ("language of thought")
- Integrated with pragmatics
- Fuller probabilistic treatment (generative)