Common Sense and Language

Benjamin Van Durme

Hobbs (1987)

We use words to talk about the world. Therefore, to understand what words mean, we must have a prior explication of how we view the world.

Credit: Benjamin Van Durme

JOHNS HOPKINS
UNIVERSITY

Common Sense and Language

Common sense **for** language

Common sense **from** language

Common Sense and Language

Common sense **for** language

Common sense from language

Credit: Benjamin Van Durme

Definitions

Shared Knowledge

If …

Shared Knowledge

If …

I know p

Shared Knowledge

If …

    I know p

    You know p

Shared Knowledge

If …

   I know p

   You know p

   I know that you know p

Shared Knowledge

If …

   I know p

   You know p

   I know that you know p

   You know that I know p

Shared Knowledge

If …

    I know **p**

    You know **p**

    I know that you know **p**

    You know that I know **p**

Then **p** is *shared knowledge*

JOHNS HOPKINS UNIVERSITY

Background Knowledge

If p is shared across some group,

then we say that p is *background knowledge*

Common Sense

When that group is really big then p is called:
commonly known background knowledge

or just simply: *common sense*

JOHNS HOPKINS
UNIVERSITY

Nunberg (1987)

*… this extensive body of knowledge …
underlies the use of natural-language … when
they interpret utterances –* *this is what
constitutes their "common-sense beliefs"* …

Credit: Benjamin Van Durme

Mother Jones. June 5, 2009.

JOHNS HOPKINS
UNIVERSITY

Inference

Credit: Benjamin Van Durme

Clark (1975)

[*bridging inferences*] *conveyed by language ... draw on one's knowledge of natural objects and events that goes beyond one's knowledge of language itself.*

Credit: Benjamin Van Durme

JOHNS HOPKINS
U N I V E R S I T Y
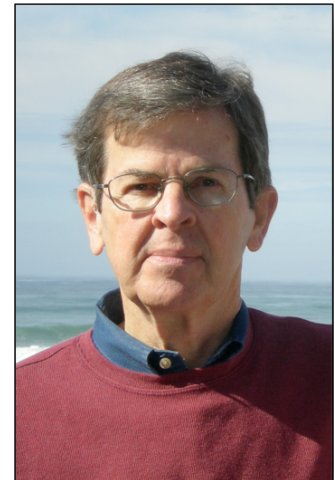
McCarthy (1959)

a program has common sense if it automatically deduces for itself a sufficiently wide class of immediate consequences of anything it is told and what

*it already knows*
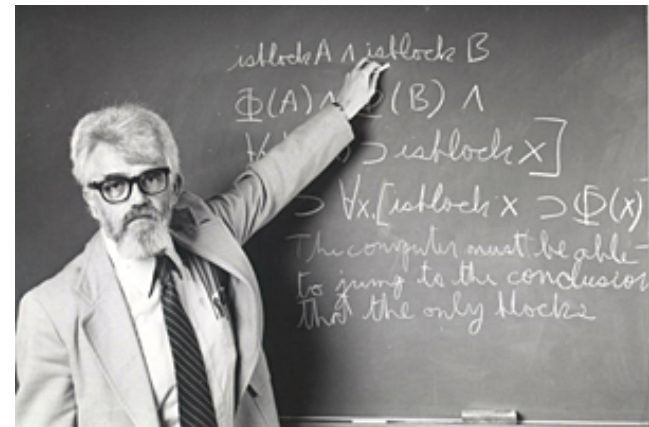
Stanford Report. June 18, 2003.

JOHNS HOPKINS
U N I V E R S I T Y

Credit: Benjamin Van Durme

Hobbs *et al.* (1993)

Interpretation as Abduction:

The interpretation of a text is the minimal explanation of why the text would be true

Credit: Benjamin Van Durme

JOHNS HOPKINS
UNIVERSITY

Schank (1975)

It has been apparent … within the domain of *natural language understanding* *… that the eventual limit to our solution … would be our ability to characterize world knowledge.*

JOHNS HOPKINS
U N I V E R S I T Y

Credit: Benjamin Van Durme
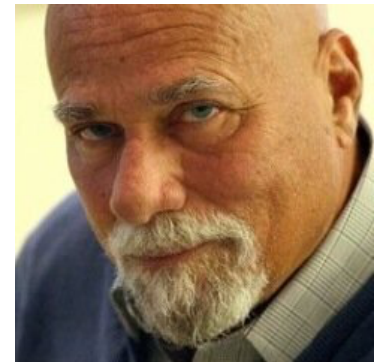
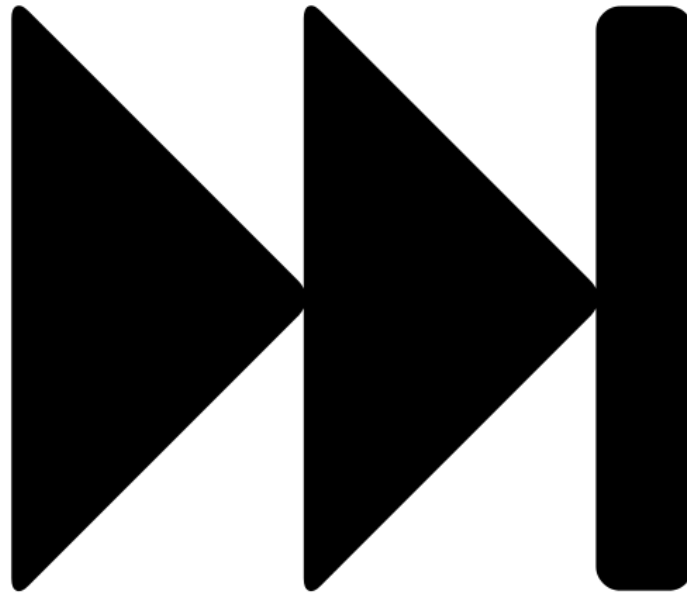# Schank (1975)

In order to build a real understanding system it will be necessary to organize the knowledge that facilitates understanding.

JOHNS HOPKINS UNIVERSITY

Credit: Benjamin Van Durme
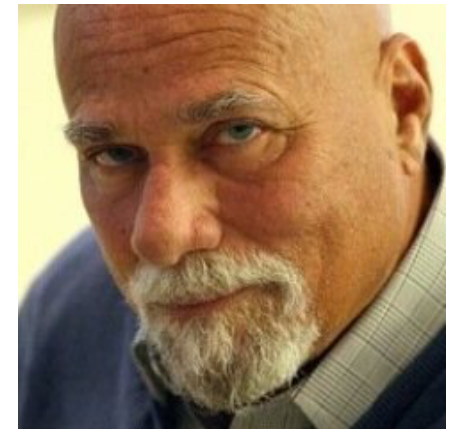
[fast forward forty years]

Credit: Benjamin Van Durme

JOHNS HOPKINS
UNIVERSITY

Bos (2013)

The bottleneck in achieving high recall is the lack of a systematic way to produce relevant background knowledge …

the knowledge acquisition problem.

Credit: Benjamin Van Durme

JOHNS HOPKINS
U N I V E R S I T Y

Build a KB by hand?

Early musings

Hobbs and Navarretta (1993)

Professionals

Cyc (Lenat 1995), Project Halo (Friedland and Allen 2003)

Crowdsourced

OpenMind (Singh 2002, Havasi *et al* 2007), Learner (Chklovski 2003)

Games with a purpose

Verbosity (von Ahn *et al* 2004)

Credit: Benjamin Van Durme

Common Sense and Language

Common sense **for** language

Common sense from language

Common Sense and Language

Common sense for language

Common sense from language

Credit: Benjamin Van Durme

MindNet: acquiring and structuring semantic information from text

Stephen D. Richardson, William B. Dolan, Lucy Vanderwende
Microsoft Research
One Microsoft Way
Redmond, WA 98052
U.S.A.

automatic procedures … provide the only credible prospect for acquiring world knowledge on the scale needed to support common-sense reasoning

(effort spanning: 1992, 1993, 1996, 1998, 2005, …)

Credit: Benjamin Van Durme

JOHNS HOPKINS
UNIVERSITY

Many Examples

"Knowledge" Patterns

Montemagni and Vanderwende (1992)
KNEXT (Schubert 2002)
Liakata and Pulman (2002)
Almuhareb and Poesio (2004)
ALICE (Banko and Etzioni 2007)
Paşca and Van Durme (2007)
DART (Clark and Harrison 2009)
Berant et al (2011)
…

"Language" Patterns

Church and Hanks (1989)
Hearst (1992)
Zernik (1992)
Resnik (1993)
Rosenfeld (1994)
Abney (1996)
Lin (1998)
VerbOcean (Chklovski and Pantel 2004)
Snow et al. (2005)
Chambers and Jurafsky (2008)
PPDB (Ganitkevitch et al 2013)

JOHNS HOPKINS
UNIVERSITY

Credit: Benjamin Van Durme

Basic properties

Combinations, correlations, scripts

Basic properties

Combinations, correlations, scripts

Credit: Benjamin Van Durme

Instances $a_1, a_2, ..., b_1, ...$

Classes $A, B, ...$

$$a_1, a_2, ... \in A$$

$$b_1, b_2, ... \in B$$

$$P(a_1), P(a_2), ... \Rightarrow P(A)$$

$$R(a_1, b_7), R(a_2, b_4), ... \Rightarrow R(A, B)$$

Basic properties

KNEXT (what people describe)

Query Logs (what people search for)

Social Media (what people say)

Basic properties

KNEXT (what people describe)

Query Logs (what people search for)

Social Media (what people say)

Credit: Benjamin Van Durme

KNowledge EXTraction (KNEXT)

...
Gordon *et al.* (2010)
Gordon *et al.* (2010)
Gordon *et al.* (2009)
Van Durme and Gildea (2009)
Van Durme *et al.* (2009)
Van Durme and Schubert (2008)
Van Durme *et al.* (2008)
Schubert and Tong (2003)
Schubert (2002)

Credit: Benjamin Van Durme

Schubert (2002)

there is a largely untapped source of general knowledge in texts, lying … beneath the explicit assertional content

JOHNS HOPKINS
U N I V E R S I T Y

Credit: Benjamin Van Durme

Rep. Ronnie Flippo (D. AL.), one of the members of the delegation, says he was particularly impressed by Mr. Krenz's ready admission that East Germany needs to change.

Rep. Ronnie Flippo (D. AL.), one of the members of the delegation, says he was particularly impressed by Mr. Krenz's ready admission that East Germany needs to change.

```
(((:I (:Q DET ELECTED-REPRESENTATIVE) SAY[V] (:Q DET PROPOS))
  (:I (:Q DET DELEGATION[N]) HAVE[V]
      (:Q DET (:F PLUR MEMBER[N])))
  (:I (:Q DET MALE-INDIVIDUAL) (:F BE[PASV] IMPRESS[V])
      (:P BY[P] (:Q DET ADMISSION[N])))
  (:I (:Q DET ADMISSION[N]) READY[A])
  (:I (:Q DET COUNTRY) NEED[V] (:F KA CHANGE[V]))))
```

Credit: Benjamin Van Durme

Rep. Ronnie Flippo (D. AL.), one of the members of the delegation, says he was particularly impressed by Mr. Krenz's ready admission that East Germany needs to change.

AN ELECTED-REPRESENTATIVE MAY SAY A PROPOSITION.

A DELEGATION MAY HAVE MEMBERS.

A MALE-INDIVIDUAL MAY BE IMPRESSED BY AN ADMISSION.

AN ADMISSION CAN BE READY.

A COUNTRY MAY NEED TO CHANGE.

JOHNS HOPKINS
UNIVERSITY

Credit: Benjamin Van Durme

Rep. Ronnie Flippo (D. AL.), one of the members of the delegation, says he was particularly impressed by Mr. Krenz's ready admission that East Germany needs to change.

AN ELECTED-REPRESENTATIVE MAY SAY A PROPOSITION.
A DELEGATION MAY HAVE MEMBERS.
A MALE-INDIVIDUAL MAY BE IMPRESSED BY AN ADMISSION.
AN ADMISSION CAN BE READY.
A COUNTRY MAY NEED TO CHANGE.

JOHNS HOPKINS
UNIVERSITY

Credit: Benjamin Van Durme

Rep. Ronnie Flippo (D. AL.), one of the members of the delegation, says he was particularly impressed by Mr. Krenz's ready admission that East Germany needs to change.

AN ELECTED-REPRESENTATIVE MAY SAY A PROPOSITION.
A DELEGATION MAY HAVE MEMBERS.
A MALE-INDIVIDUAL MAY BE IMPRESSED BY AN ADMISSION.
AN ADMISSION CAN BE READY.
A COUNTRY MAY NEED TO CHANGE.

JOHNS HOPKINS
UNIVERSITY

Credit: Benjamin Van Durme

cs.rochester.edu/research/knext

13 years of extracting millions of statements of general world knowledge

JOHNS HOPKINS UNIVERSITY

Credit: Benjamin Van Durme

Basic properties

KNEXT (what people describe)

Query Logs (what people search for)

Social Media (what people say)

JOHNS HOPKINS
U N I V E R S I T Y

Credit: Benjamin Van Durme

Basic properties

KNEXT (what people describe)

Query Logs (what people search for)

Social Media (what people say)

Credit: Benjamin Van Durme

Query Logs (what people search for)

…

Van Durme and Paşca (2008)

Paşca and Van Durme (2008)

Paşca, Van Durme and Garera (2007)

Paşca and Van Durme (2007)

Credit: Benjamin Van Durme

JOHNS HOPKINS
UNIVERSITY

china capital city

capital france

capital of spain
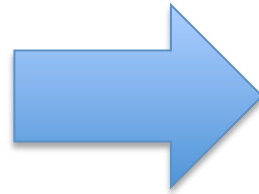
usa capital

what is the capital of denmark
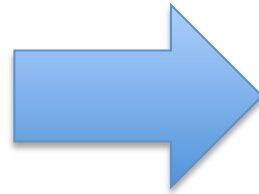
wikimedia.org

JOHNS HOPKINS
UNIVERSITY

Credit: Benjamin Van Durme

Country:	capital, population, president, map, capital city, currency, climate, flag, culture, leader
Drug:	side effects, cost, structure, benefits, mechanism of action, overdose, long term use, price, synthesis, pharmacology
Company:	ceo, future, president, competitors, mission statement, owner, website, organizational structure, logo, market share
City:	population, map, mayor, climate, location, geography, best, culture, capital, latitude
Painter:	paintings, works, portrait, death, style, artwork, bibliography, bio, autobiography, childhood

Paşca and Van Durme (2007)

Country: capital, population, president, map, capital city, currency, climate, flag, culture, leader

Drug: side effects, cost, structure, benefits, mechanism of action, overdose, long term use, price, synthesis, pharmacology

Company: ceo, future, president, competitors, mission statement, owner, website, organizational structure, logo, market share

City: population, map, mayor, climate, location, geography, best, culture, capital, latitude

Painter: paintings, works, portrait, death, style, artwork, bibliography, bio, autobiography, childhood

Paşca and Van Durme (2007)

Prince (1978)
Stereotypical Tacit Assumptions

PEOPLE have:

parents, siblings, a spouse, a home,
a television, a clock, neighbors, …

COUNTRIES have:

a leader, a president, a queen, a duke,
citizens, land, a language, a history, …

Credit: Benjamin Van Durme

Country: capital, population, president, map, capital city, currency, climate, flag, culture, leader

Paşca and Van Durme (2007)

we know that
    COUNTRIES have:
        a leader, a president, a queen, a duke,
        citizens, land, a language, a history, …

Prince (1978)

JOHNS HOPKINS
U N I V E R S I T Y

Credit: Benjamin Van Durme

Basic properties

KNEXT (what people describe)

Query Logs (what people search for)

Social Media (what people say)

JOHNS HOPKINS
UNIVERSITY

Basic properties

KNEXT (what people describe)

Query Logs (what people search for)

Social Media (what people say)

Social Media

...

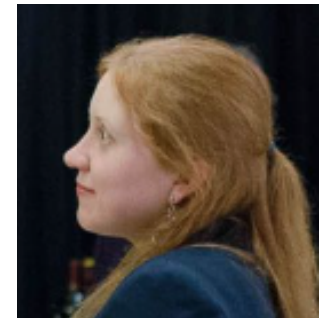Volkova and Van Durme (2015)

Volkova *et al.* (2014)

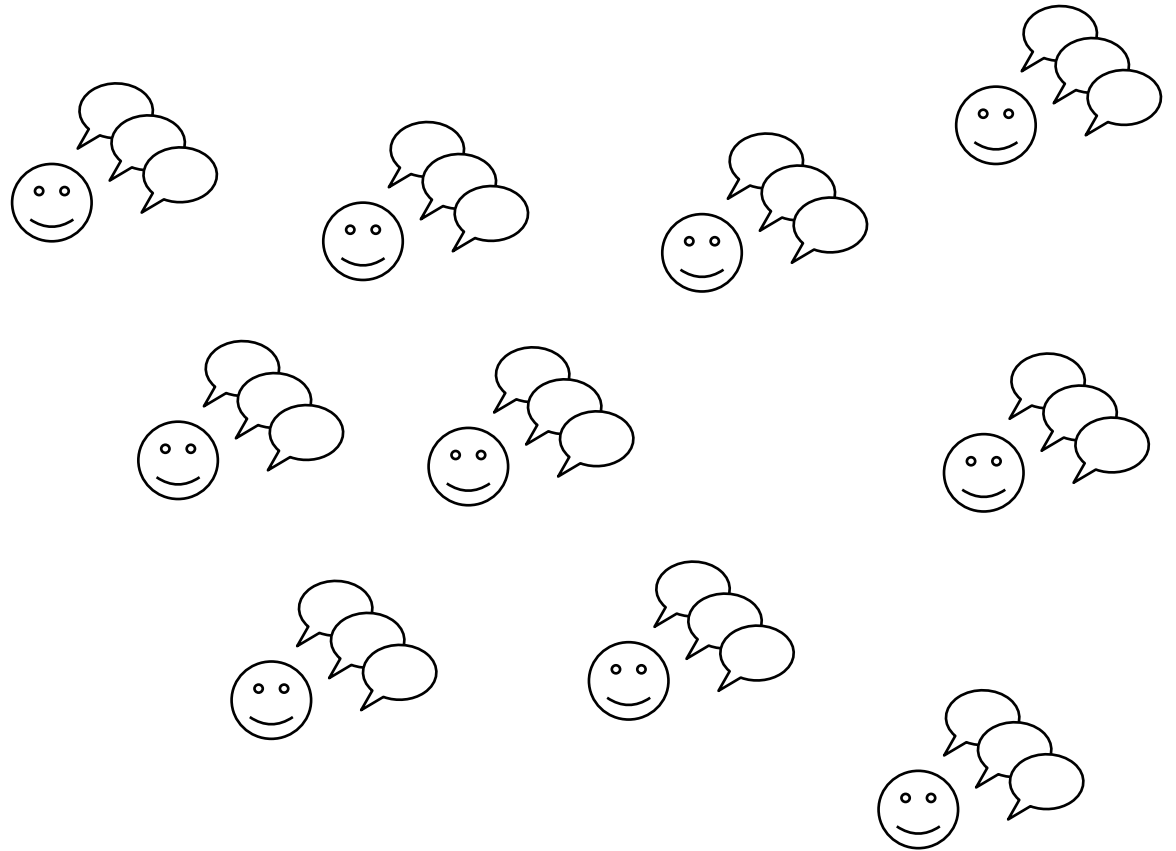Beller *et al.* (2014)

Beller *et al.* (2014)

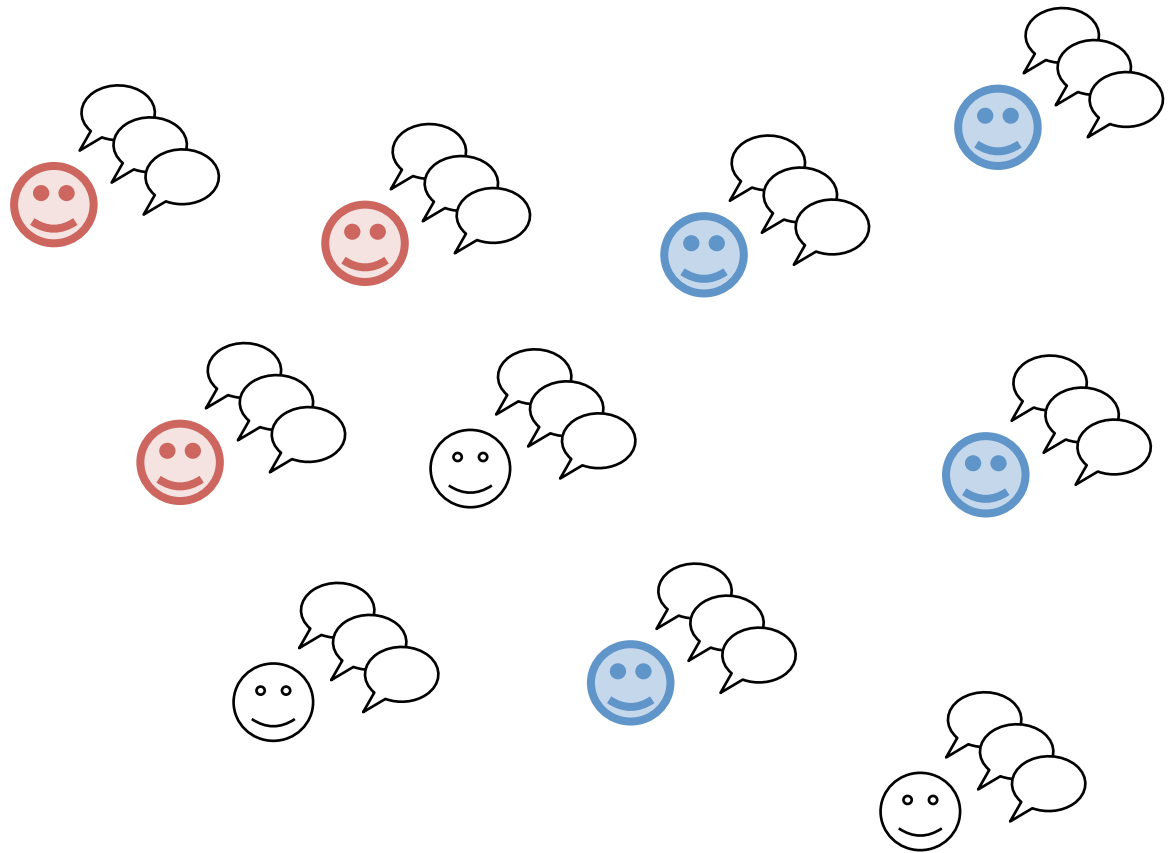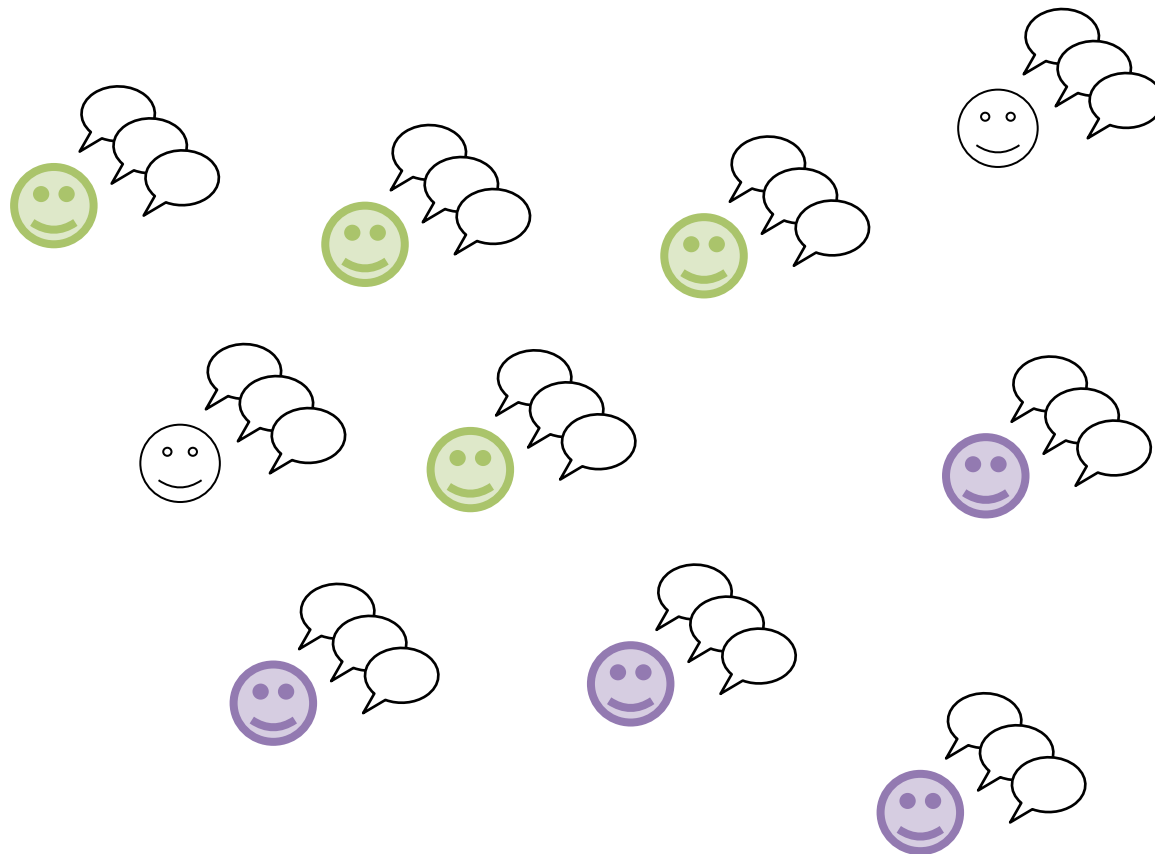Bergsma and Van Durme (2013)

Bergsma *et al.* (2013)

Van Durme (2012)

Credit: Benjamin Van Durme

JOHNS HOPKINS
UNIVERSITY

Credit: Benjamin Van Durme

FEMALE

MALE

Credit: Benjamin Van Durme

JOHNS HOPKINS
UNIVERSITY

Above 25

Under 25

Credit: Benjamin Van Durme

College Student

Credit: Benjamin Van Durme
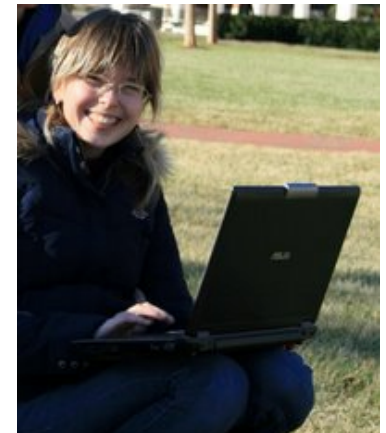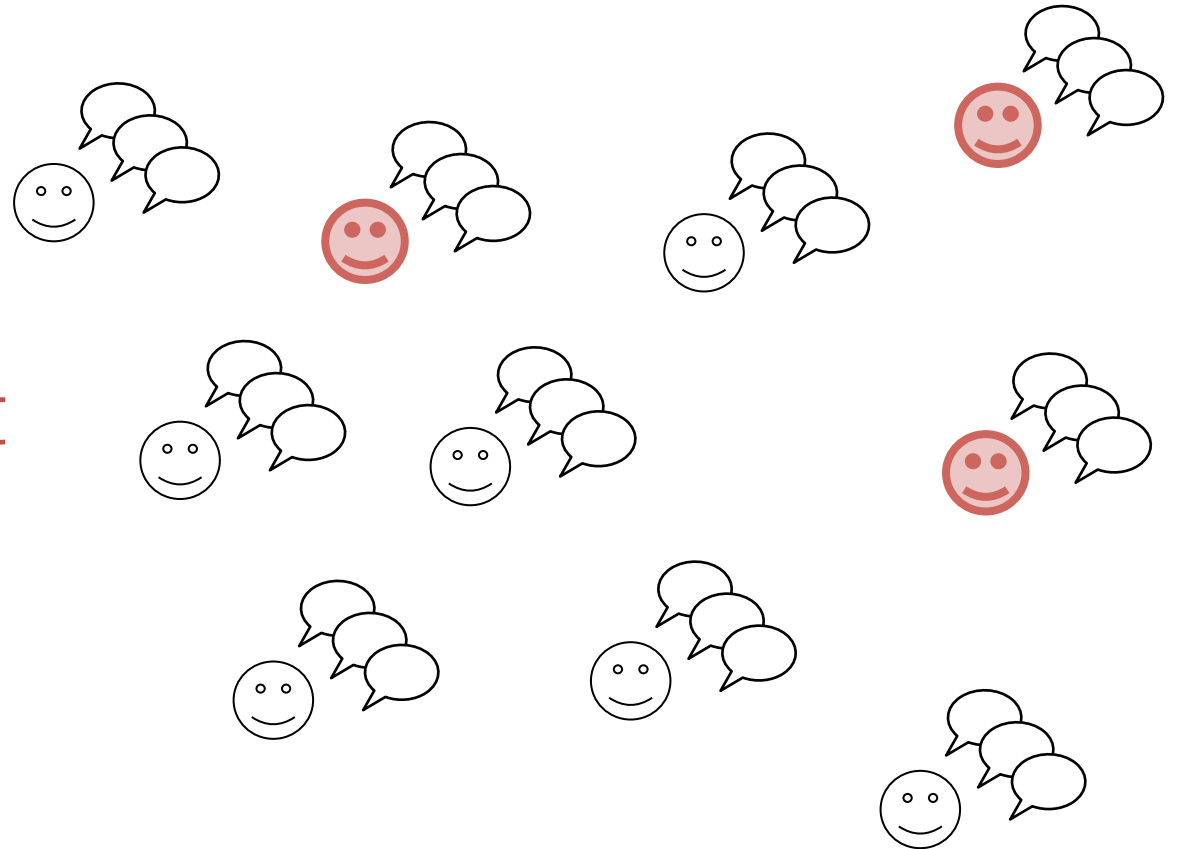
Characteristic Attributes

Reporter

Check out my article featured in this …

College Student

Still ain't bought my cap and gown

Actor

Brett just took one of my lines but …

Credit: Benjamin Van Durme

Characteristic Activities

Student

I finish school *in March …*

DJ

Quick cut session before *I spin tonight*

Christian

I thank God *4 His blessings in my life …*

Credit: Benjamin Van Durme

Basic properties

KNEXT (what people describe)

Query Logs (what people search for)

Social Media (what people say)

Credit: Benjamin Van Durme

Basic properties

KNEXT (what people describe)

Query Logs (what people search for)

Social Media (what people say)

Credit: Benjamin Van Durme

Basic properties

Combinations, correlations, scripts

Basic properties

Combinations, correlations, scripts

Minsky (1974)

A frame is … for representing a stereotyped situation, like being in a certain kind of living room, or going to a child's birthday party.

Credit: Benjamin Van Durme

Omni. June, 1980.

Schank (1975)

Two types of frames are scripts and plans ... used to understand and generate stories and actions, and there can be little understanding without them.

JOHNS HOPKINS
U N I V E R S I T Y
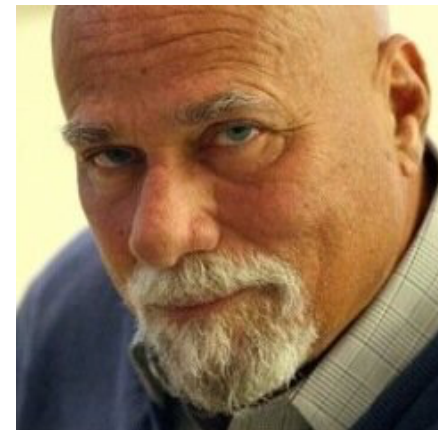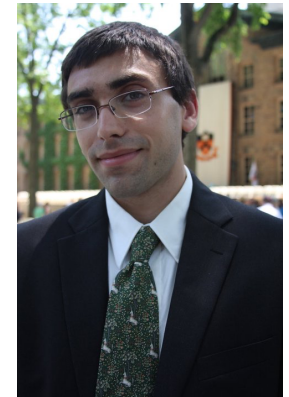
Scripts, Discourse Models

Ferraro and Van Durme (submitted)

Rudinger *et al.* (submitted)

Rudinger *et al.* (2015)

Kjersten and Van Durme (2012)

Van Durme and Lall (2009)

Van Durme (2008)

Credit: Benjamin Van Durme

JOHNS HOPKINS
U N I V E R S I T Y

$$R(a_1, b_2)$$

$$R(a_1, b_2)$$

$$R'(a_1, c_7)$$

Credit: Benjamin Van Durme

$$R(a_1, b_2) \ \& \ R'(a_1, c_7)$$
$$R(a_3, b_4) \ \& \ R'(a_3, c_9)$$
$$R(a_9, b_8) \ \& \ R'(a_9, c_2)$$
$$\ldots$$

$$R(\boxed{a_1}, b_2) \; \& \; R'(\boxed{a_1}, c_7)$$

$$R(a_3, b_4) \; \& \; R'(a_3, c_9)$$

$$R(a_9, b_8) \; \& \; R'(a_9, c_2)$$

$$\ldots$$

$$R(a_1, b_2) \ \& \ R'(a_1, c_7)$$

$$R(\boxed{a_3}, b_4) \ \& \ R'(\boxed{a_3}, c_9)$$

$$R(a_9, b_8) \ \& \ R'(a_9, c_2)$$

$$\dots$$

$$R(a_1, b_2) \ \& \ R'(a_1, c_7)$$

$$R(a_3, b_4) \ \& \ R'(a_3, c_9)$$

$$R(\boxed{a_9}, b_8) \ \& \ R'(\boxed{a_9}, c_2)$$

$$\ldots$$

$$R(a_1, b_2) \ \& \ R'(a_1, c_7)$$

$$R(a_3, b_4) \ \& \ R'(a_3, c_9)$$

$$R(a_9, b_8) \ \& \ R'(a_9, c_2)$$

$$\ldots$$

$$R(a_i \in A, B) \ \Rightarrow \ R'(a_i, C)$$

Credit: Benjamin Van Durme

John studied for the exam and aced it.

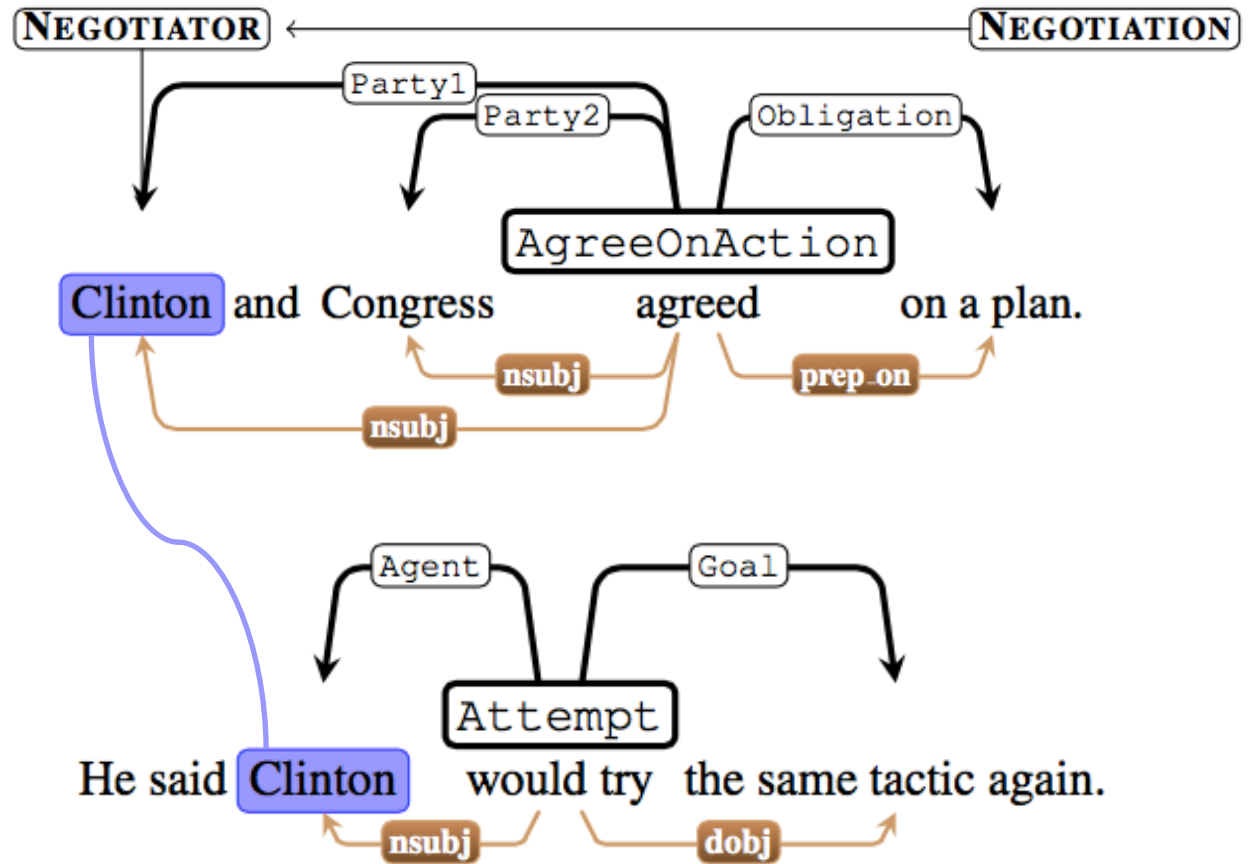His teacher congratulated him.

if:

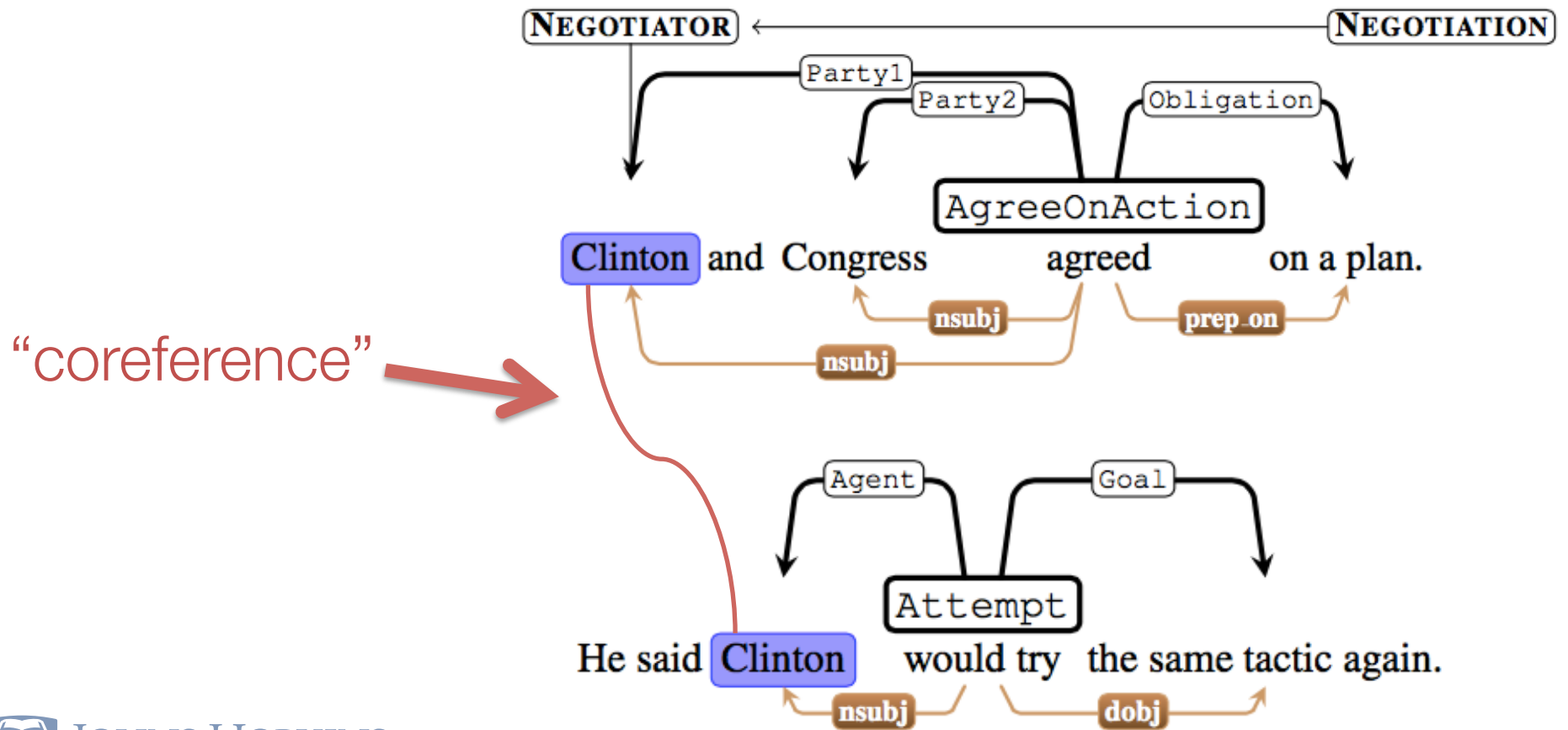a person studies for an exam,
and that person aces that exam

then:

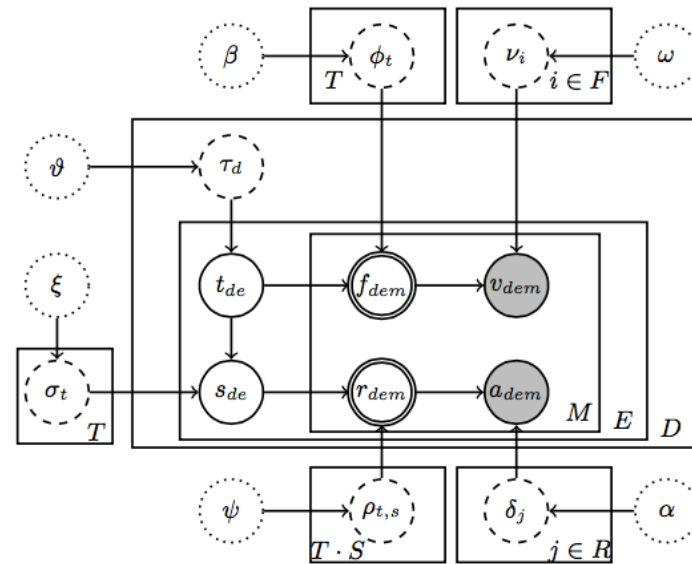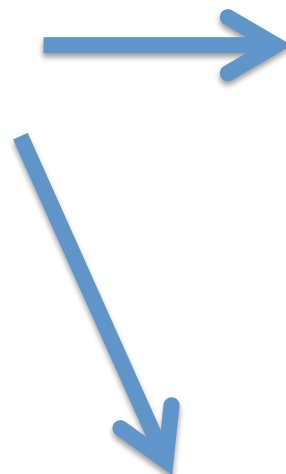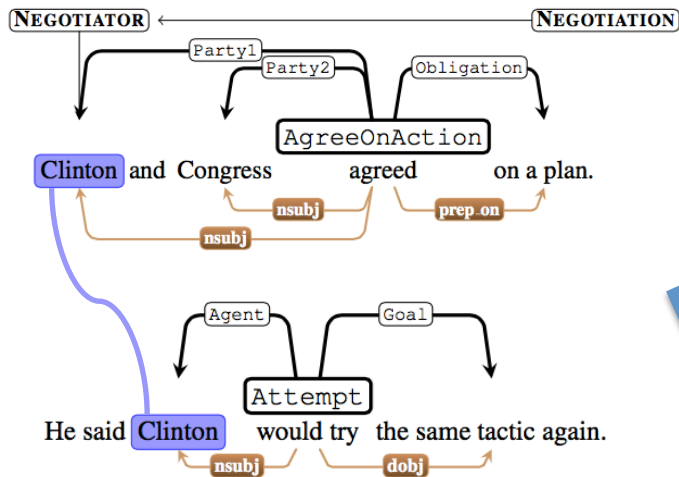a teacher might congratulate that person

JOHNS HOPKINS
UNIVERSITY

Clinton and Congress agreed on a plan.

He said Clinton would try the same tactic again.

Credit: Benjamin Van Durme

"coreference"

Credit: Benjamin Van Durme

Negotiator ← Negotiation

Party1
Party2
Obligation

AgreeOnAction

Clinton and Congress agreed on a plan.

nsubj   prep_on
nsubj

Agent   Goal

Attempt

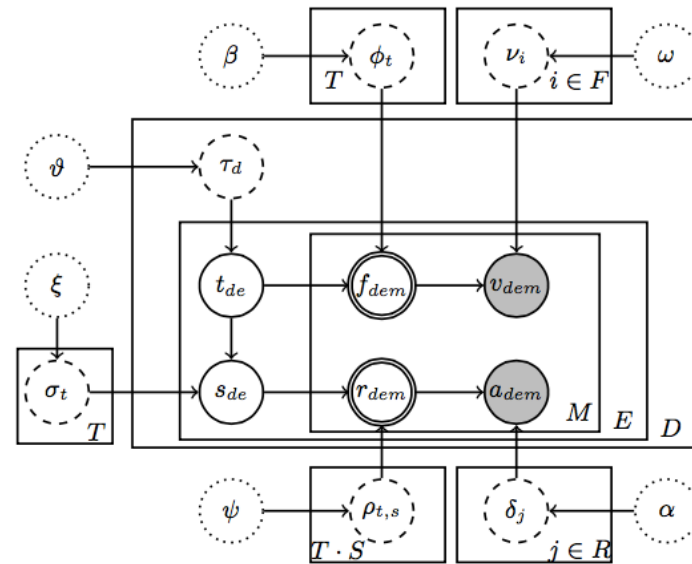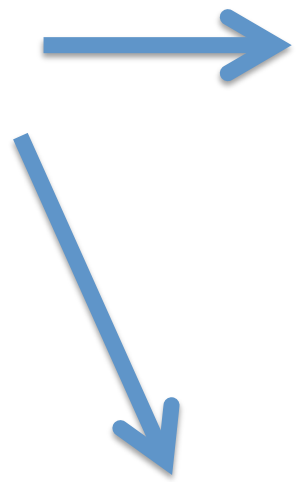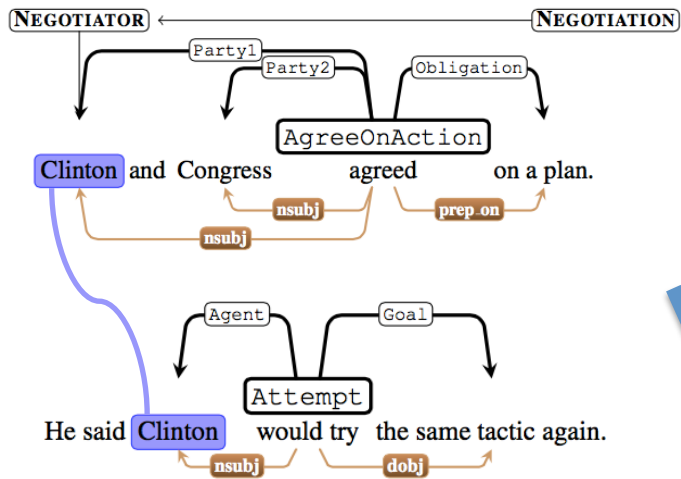He said Clinton would try the same tactic again.

nsubj   dobj

generative

$$p(w|[w_1, \ldots, w_C]) = \frac{\exp(\mathbf{t}_w^\intercal \hat{\mathbf{t}}_s + b_w)}{\sum_{w' \in \mathcal{V}} \exp(\mathbf{t}_{w'}^\intercal \hat{\mathbf{t}}_s + b_s)}$$

$$\text{where } \hat{\mathbf{t}}_s = \sum_{j=1}^{C} \mathbf{m}_j \odot \mathbf{c}_{w_j}$$

Credit: Benjamin Van Durme

discriminative

JOHNS HOPKINS
UNIVERSITY

NEGOTIATOR ← NEGOTIATION

Party1
Party2
Obligation

AgreeOnAction

Clinton and Congress agreed on a plan.

nsubj   prep_on
nsubj

Agent   Goal

Attempt

He said Clinton would try the same tactic again.

nsubj   dobj

$$p(w|[w_1,\ldots,w_C]) = \frac{\exp(\mathbf{t}_w^\intercal \hat{\mathbf{t}}_s + b_w)}{\sum_{w'\in\mathcal{V}}\exp(\mathbf{t}_{w'}^\intercal \hat{\mathbf{t}}_s + b_s)}$$

$$\text{where } \hat{\mathbf{t}}_s = \sum_{j=1}^{C} \mathbf{m}_j \odot \mathbf{c}_{w_j}$$

Credit: Benjamin Van Durme

JOHNS HOPKINS
UNIVERSITY

Language Models?

State of the art

Credit: Benjamin Van Durme

JOHNS HOPKINS
U N I V E R S I T Y

Basic properties

Combinations, correlations, scripts

Basic properties

Combinations, correlations, scripts

Bias??

Credit: Benjamin Van Durme
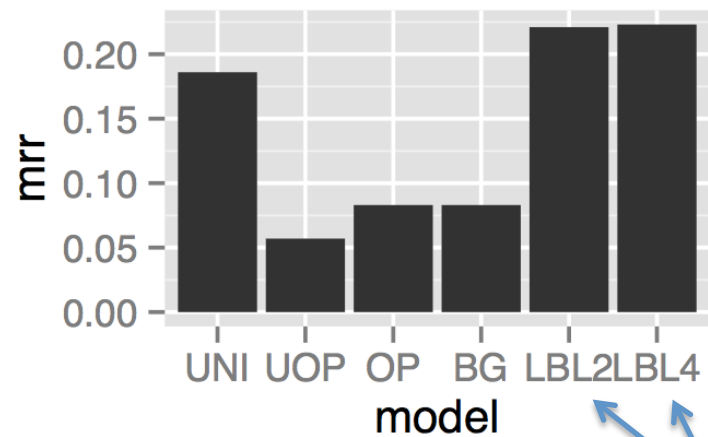
Reporting Bias

Gordon & Van Durme (2013)

Van Durme (2009)

JOHNS HOPKINS
UNIVERSITY

Credit: Benjamin Van Durme

Reporting Bias
aka "Blinking and Breathing (B&B)"

Word	Teraword	Word	Teraword
spoke	11,577,917	breathed	725,034
laughed	3,904,519	hugged	610,040
murdered	2,843,529	blinked	390,692
inhaled	984,613	exhaled	168,985

JOHNS HOPKINS
U N I V E R S I T Y

Credit: Benjamin Van Durme

Reporting Bias
aka "Blinking and Breathing (B&B)"

Word	Teraword	Word	Teraword
spoke	11,577,917	breathed	725,034
laughed	3,904,519	hugged	610,040
murdered	2,843,529	blinked	390,692
inhaled	984,613	exhaled	168,985

Reporting Bias
aka "Blinking and Breathing (B&B)"

Word	Teraword	Word	Teraword
spoke	11,577,917	breathed	725,034
laughed	3,904,519	hugged	610,040
murdered	2,843,529	blinked	390,692
inhaled	984,613	exhaled	168,985

Credit: Benjamin Van Durme

Reporting Bias
aka "Blinking and Breathing (B&B)"

Word	Teraword	Word	Teraword
spoke	11,577,917	breathed	725,034
laughed	3,904,519	hugged	610,040
murdered	2,843,529	blinked	390,692
inhaled	984,613	exhaled	168,985

JOHNS HOPKINS UNIVERSITY

Reporting Bias
aka "Blinking and Breathing (B&B)"

Word	Teraword	Word	Teraword
spoke	11,577,917	breathed	725,034
laughed	3,904,519	hugged	610,040
murdered	2,843,529	blinked	390,692
inhaled	984,613	exhaled	168,985

JOHNS HOPKINS
UNIVERSITY

Credit: Benjamin Van Durme

Basic properties

Combinations, correlations, scripts

Bias??

Credit: Benjamin Van Durme

Basic properties

Combinations, correlations, scripts

Summary

Common sense for language
Lots of interest in having it

Common sense from language
… lots of interest in getting it

(and lots of open problems remaining)

Credit: Benjamin Van Durme