Research at Google

Statistical Models for Frame-Semantic Parsing

Dipanjan Das*
Google

Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore

June 27, 2014

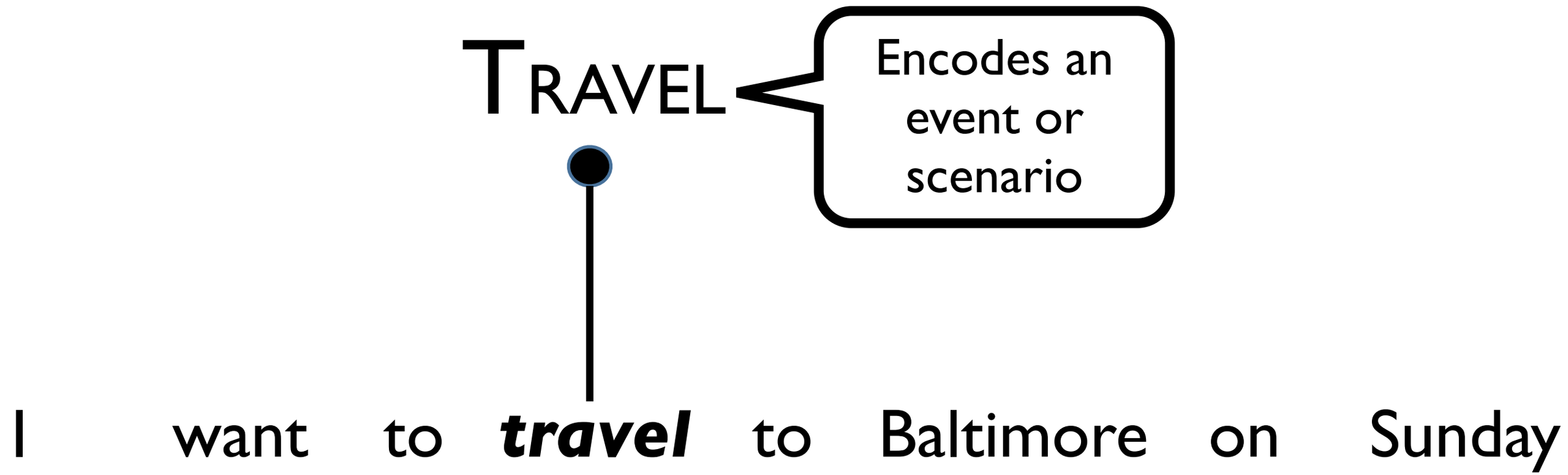
*Thanks to Desai Chen, Kuzman Ganchev, Karl Moritz Hermann, André Martins, Nathan Schneider, Noah Smith and Jason Weston

Research at Google

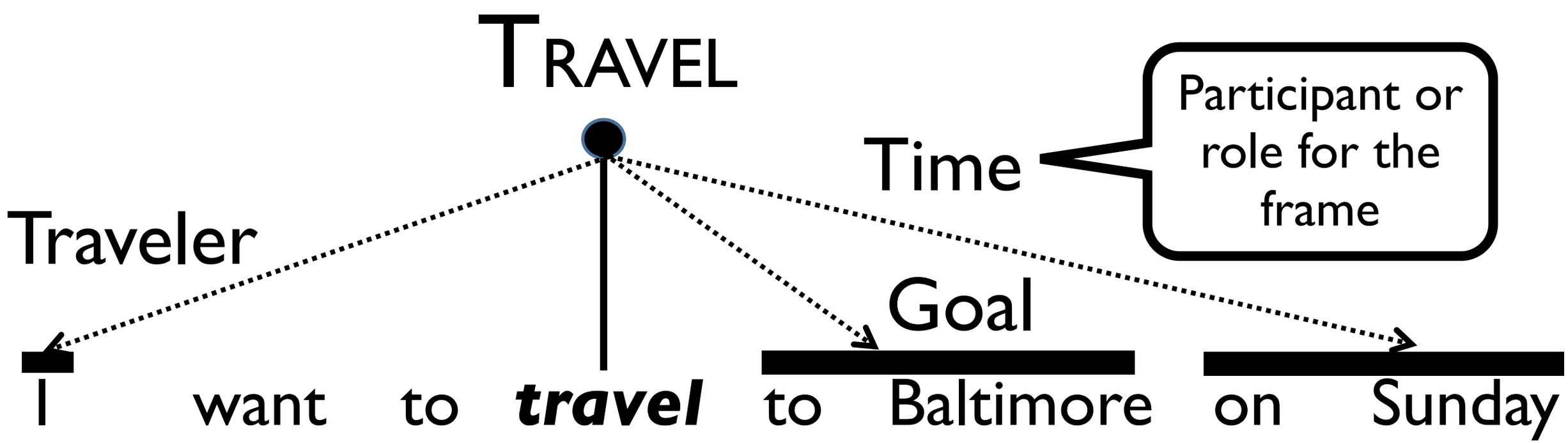
Frame-Semantic Parsing

I want to travel to Baltimore on Sunday

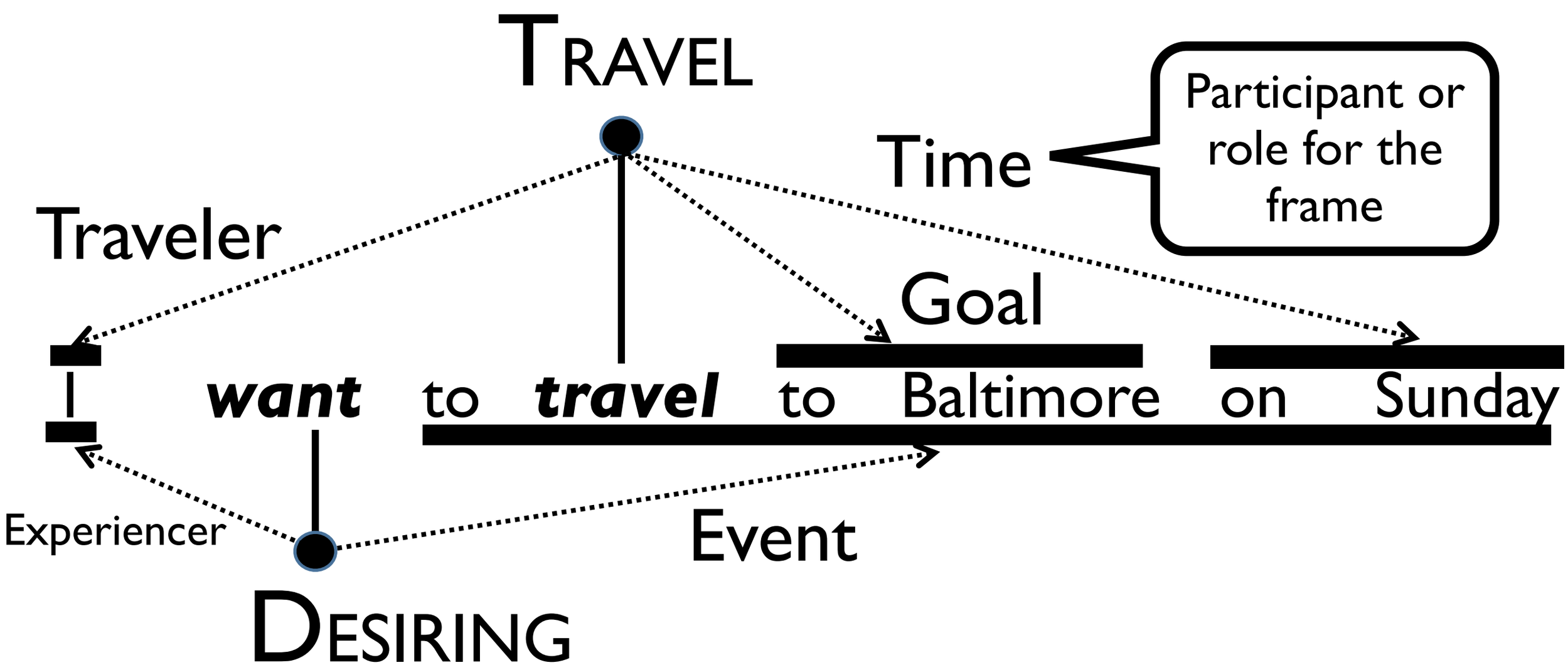
Frame-Semantic Parsing



Frame-Semantic Parsing



Frame-Semantic Parsing



Research at Google

Outline

- Why should we build statistical models for frame semantics?
- Overview of statistical methods
- Future directions

Research at Google

Outline

- Why should we build statistical models for frame semantics?
- Overview of statistical methods
- Future directions

Motivation

Research at Google

Motivation

Deeper understanding beyond syntax

Research at Google

Motivation

Deeper understanding beyond syntax

Grouping of predicates and arguments into clusters

Research at Google

Motivation

Deeper understanding beyond syntax

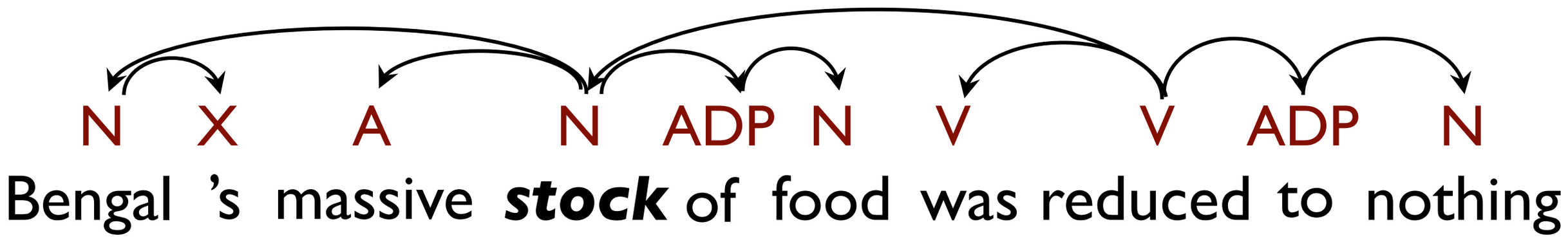
Grouping of predicates and arguments into clusters

Grounding of natural language to an ontology

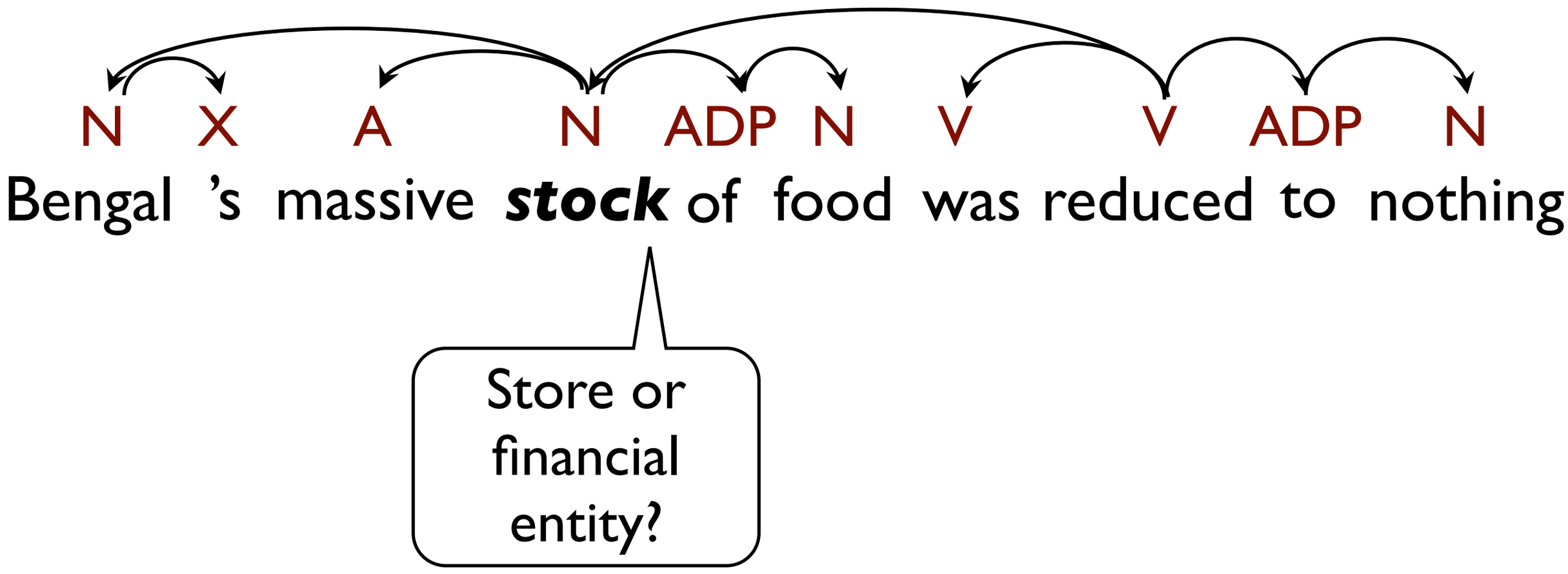
Motivation

Bengal 's massive stock of food was reduced to nothing

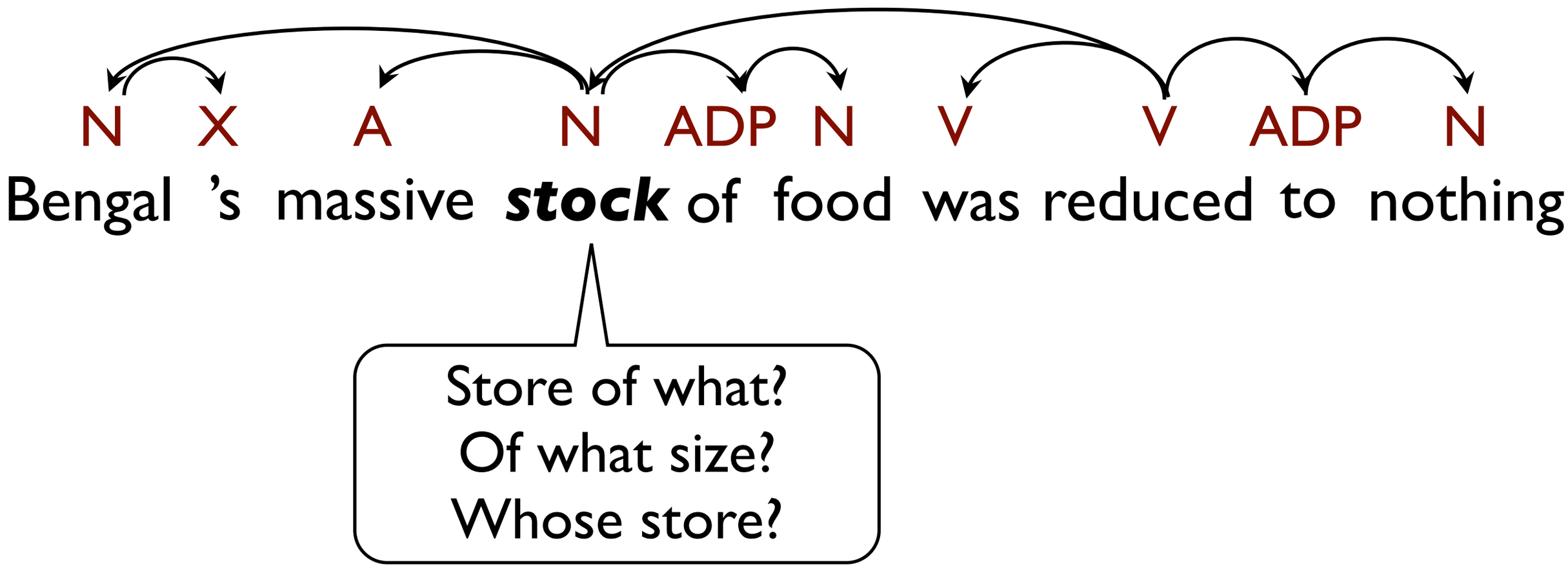
Motivation


N X A N ADP N V V ADP N
Bengal 's massive ***stock*** of food was reduced to nothing

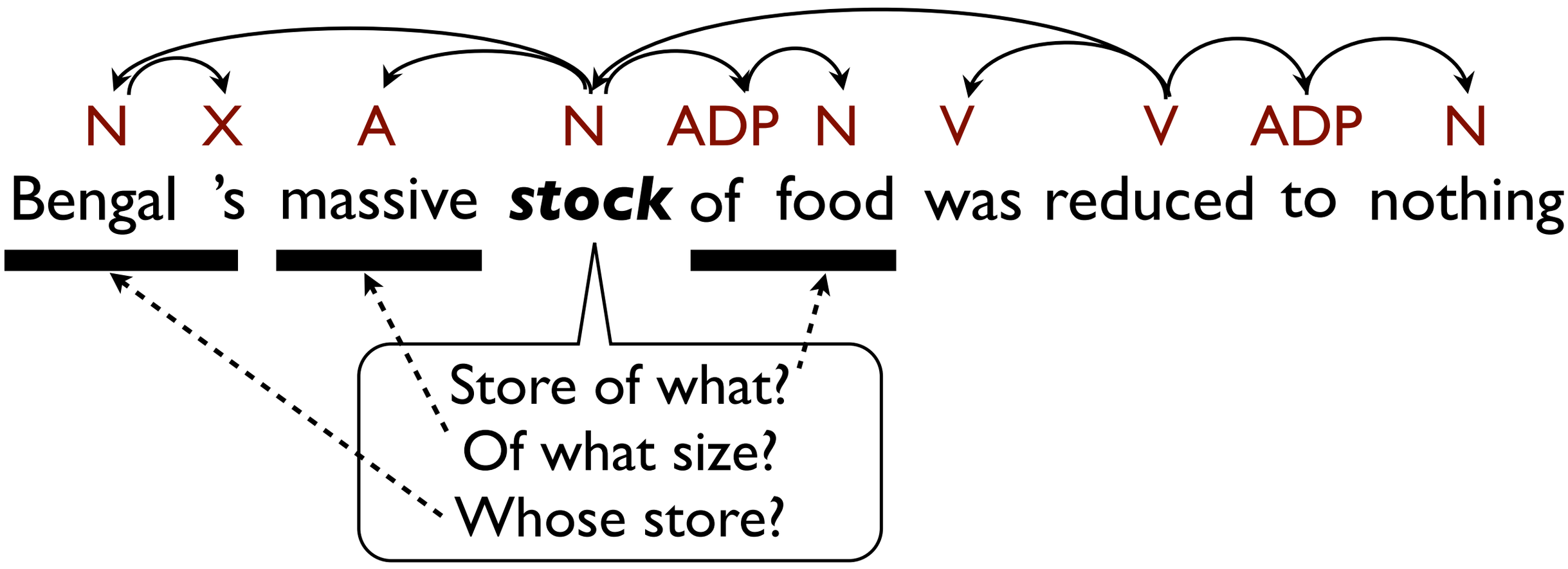
Motivation


N X A N ADP N V V ADP N
Bengal 's massive **stock** of food was reduced to nothing
Store or financial entity?

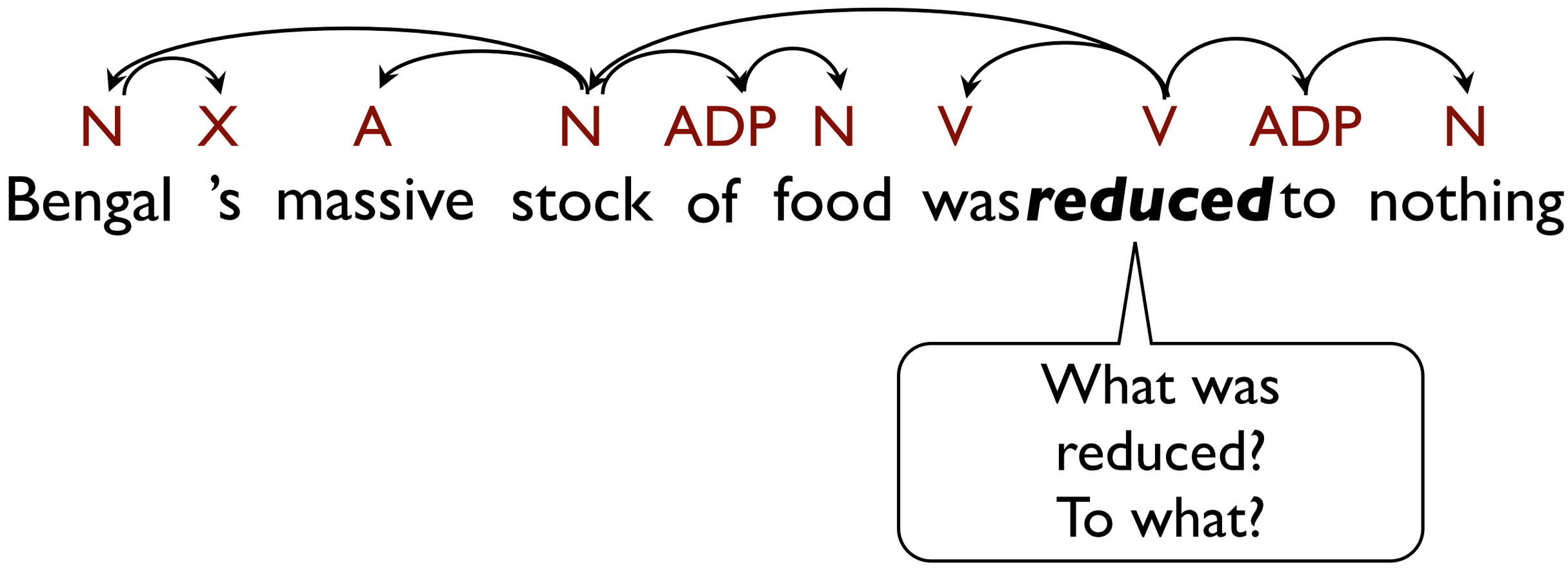
Motivation


N X A N ADP N V V ADP N
Bengal 's massive ***stock*** of food was reduced to nothing
Store of what?
Of what size?
Whose store?

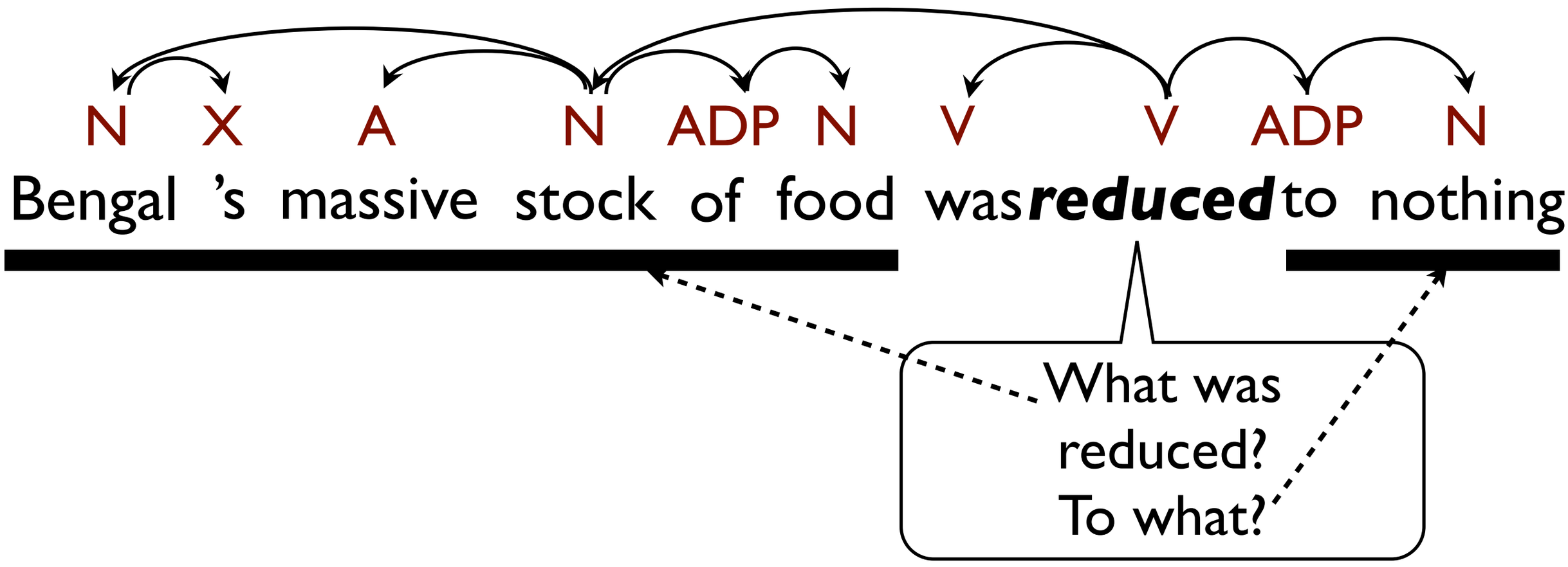
Motivation


N X A N ADP N V V ADP N
Bengal 's massive **stock** of food was reduced to nothing
Store of what?
Of what size?
Whose store?

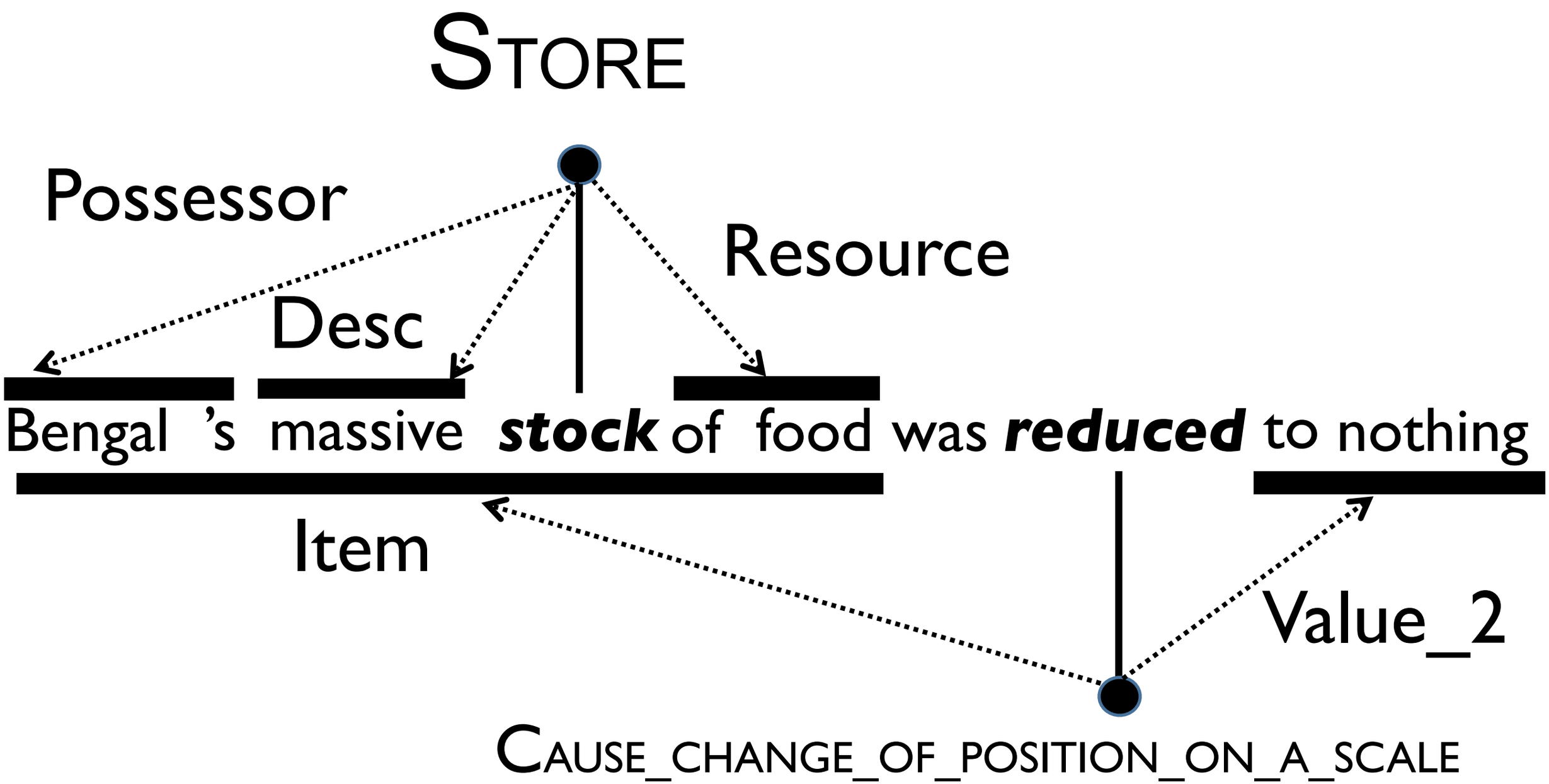
Motivation


N X A N ADP N V V ADP N
Bengal 's massive stock of food was reduced to nothing
What was reduced? To what?

Motivation


N X A N ADP N V V ADP N
Bengal 's massive stock of food was reduced to nothing
What was reduced? To what?

Motivation

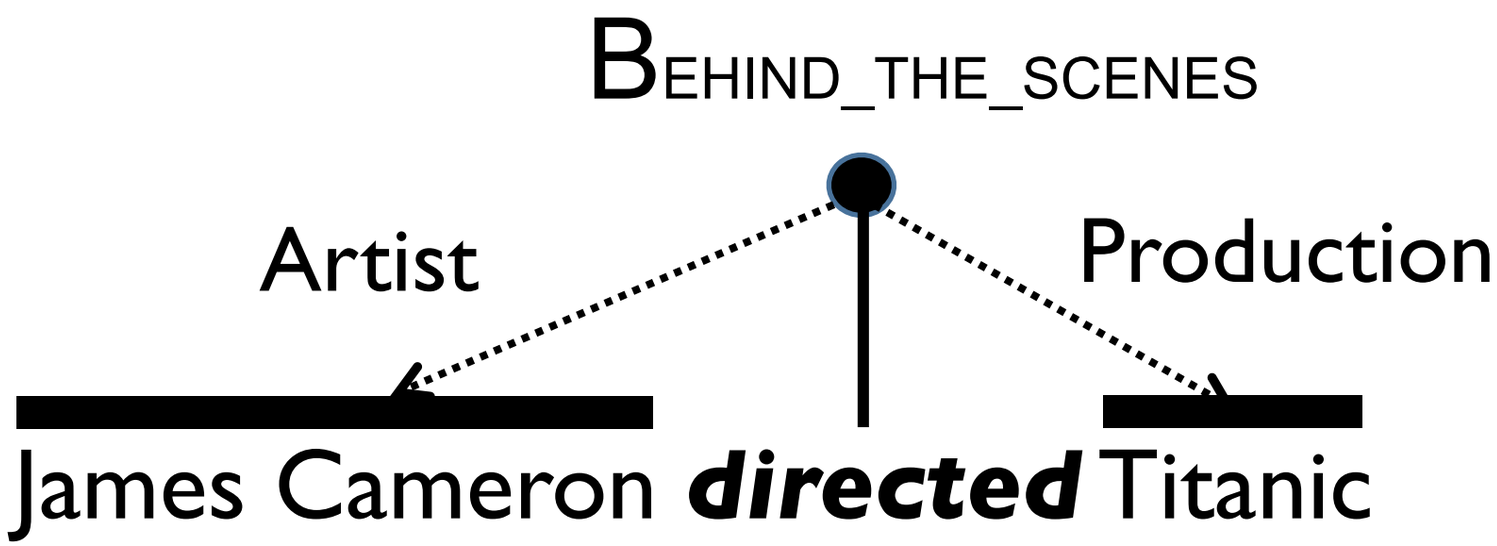

STORE
Possessor
Desc
Resource
Bengal 's massive ***stock*** of food was ***reduced*** to nothing
Item
Value_2
CAUSE_CHANGE_OF_POSITION_ON_A_SCALE

Research at Google

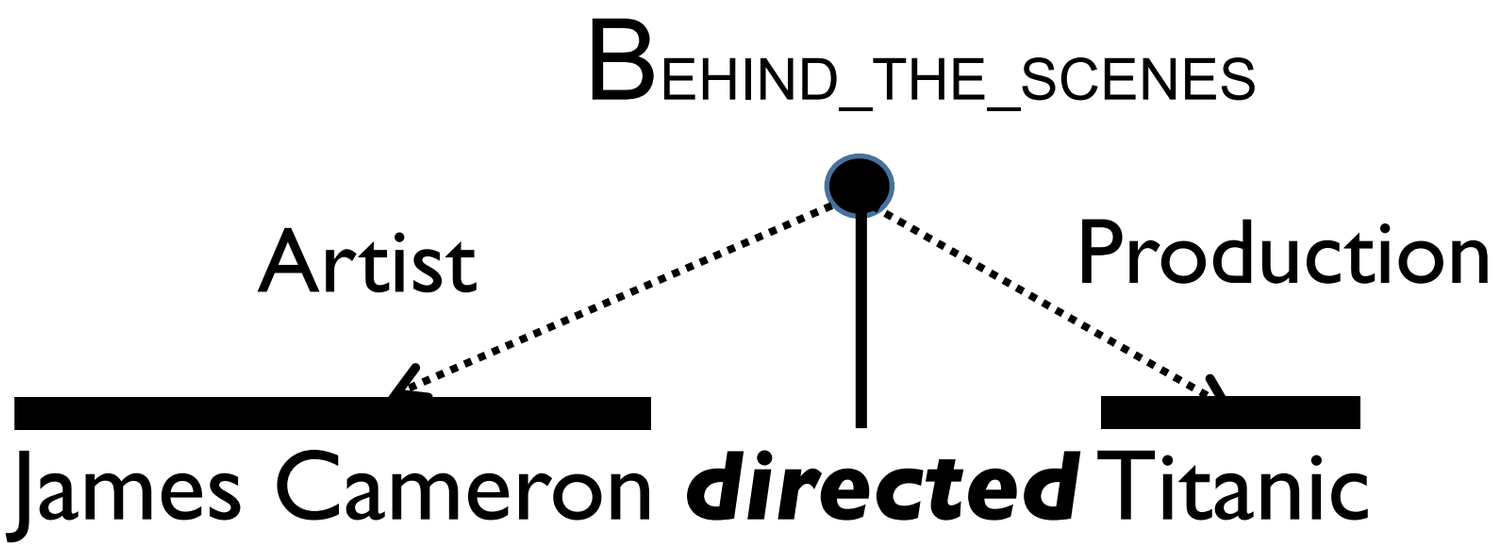
Motivation

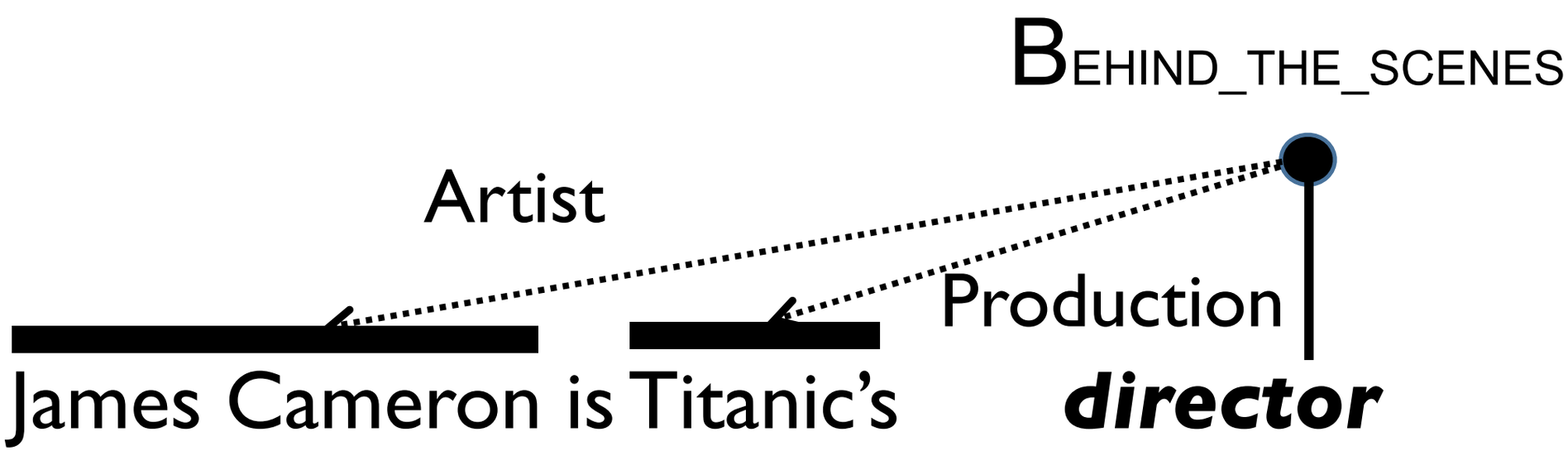
Research at Google

Motivation

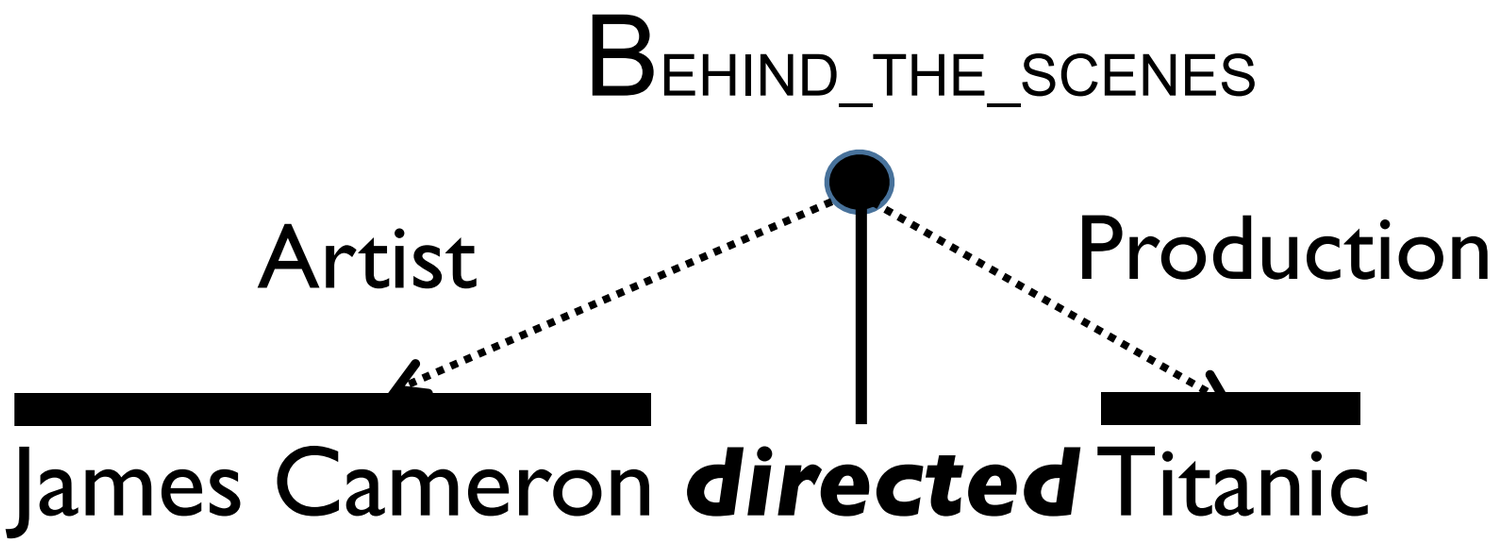


Motivation

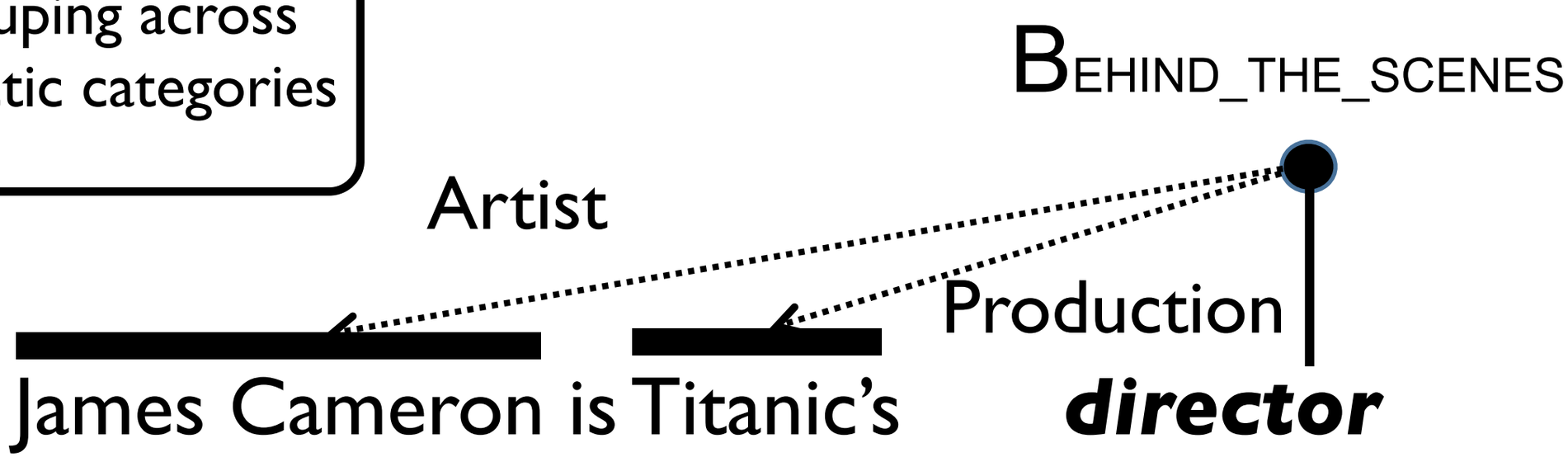

BEHIND_THE_SCENES
Artist
Production
James Cameron **directed** Titanic


BEHIND_THE_SCENES
Artist
Production
James Cameron is Titanic's **director**

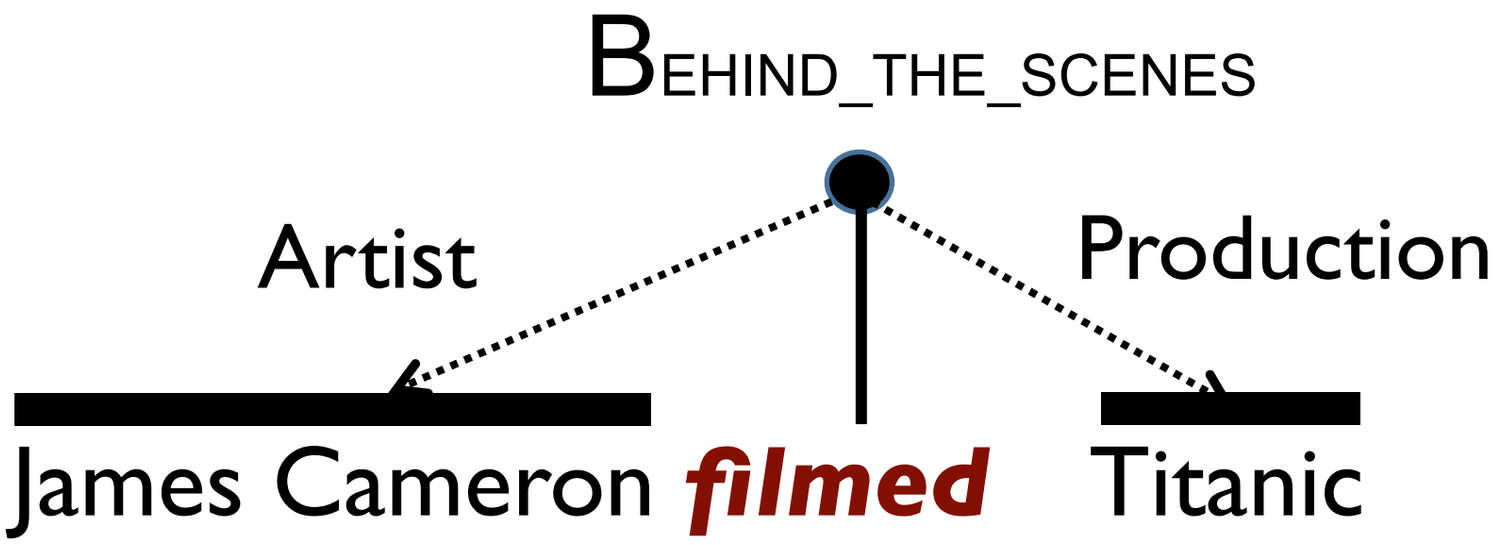
Motivation



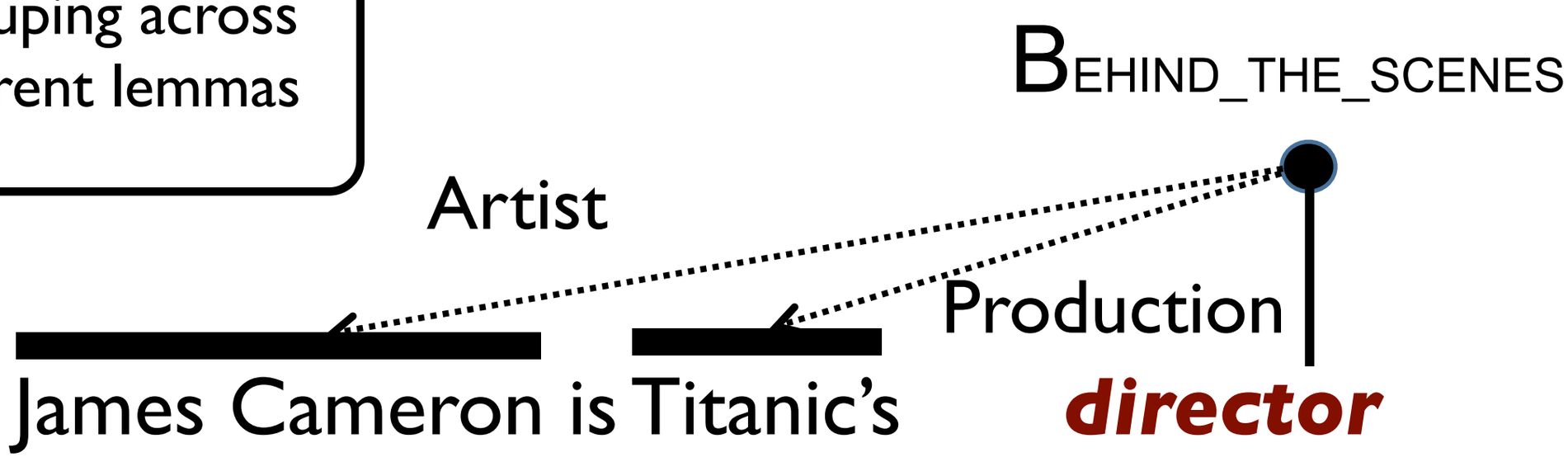
Grouping across syntactic categories



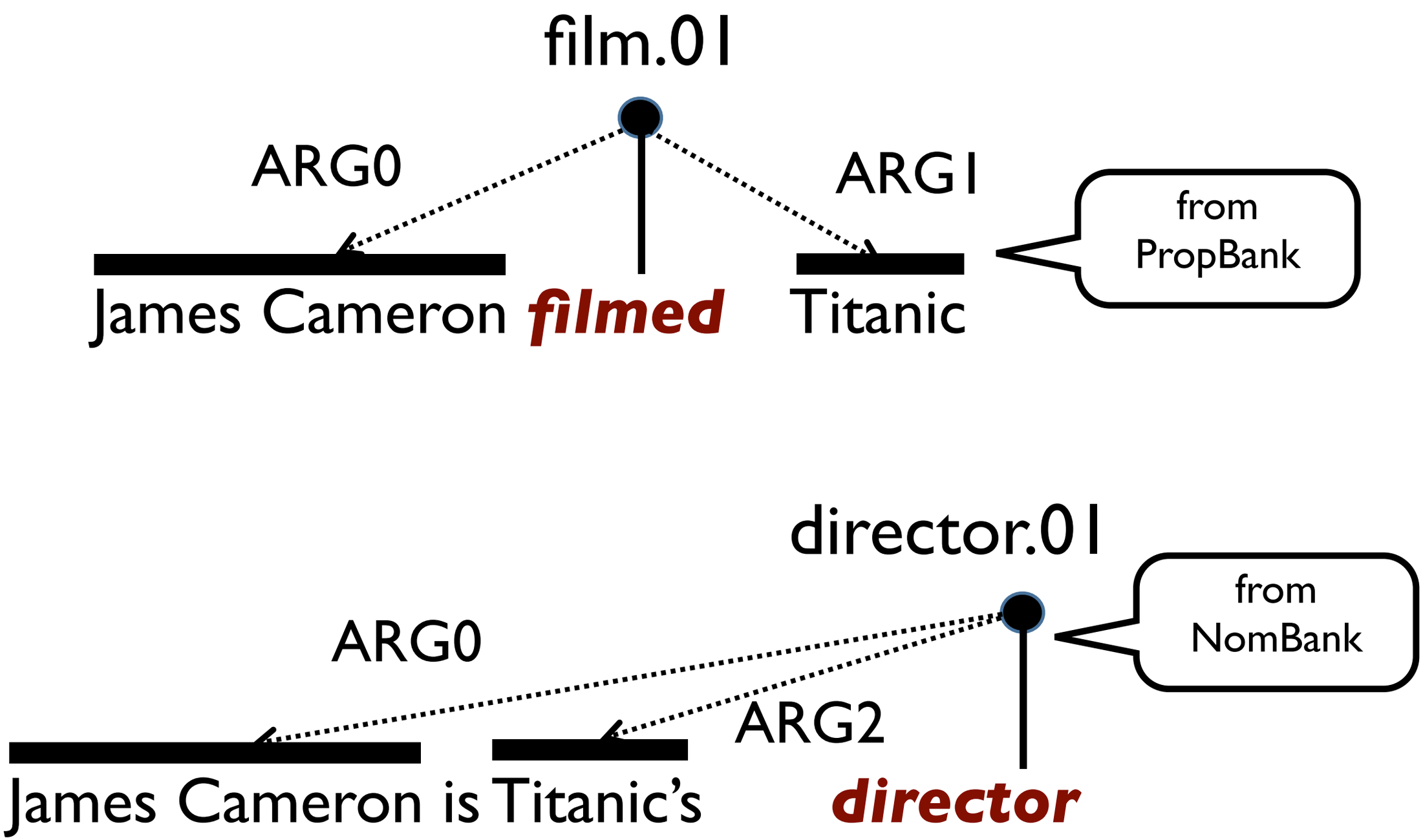
Motivation



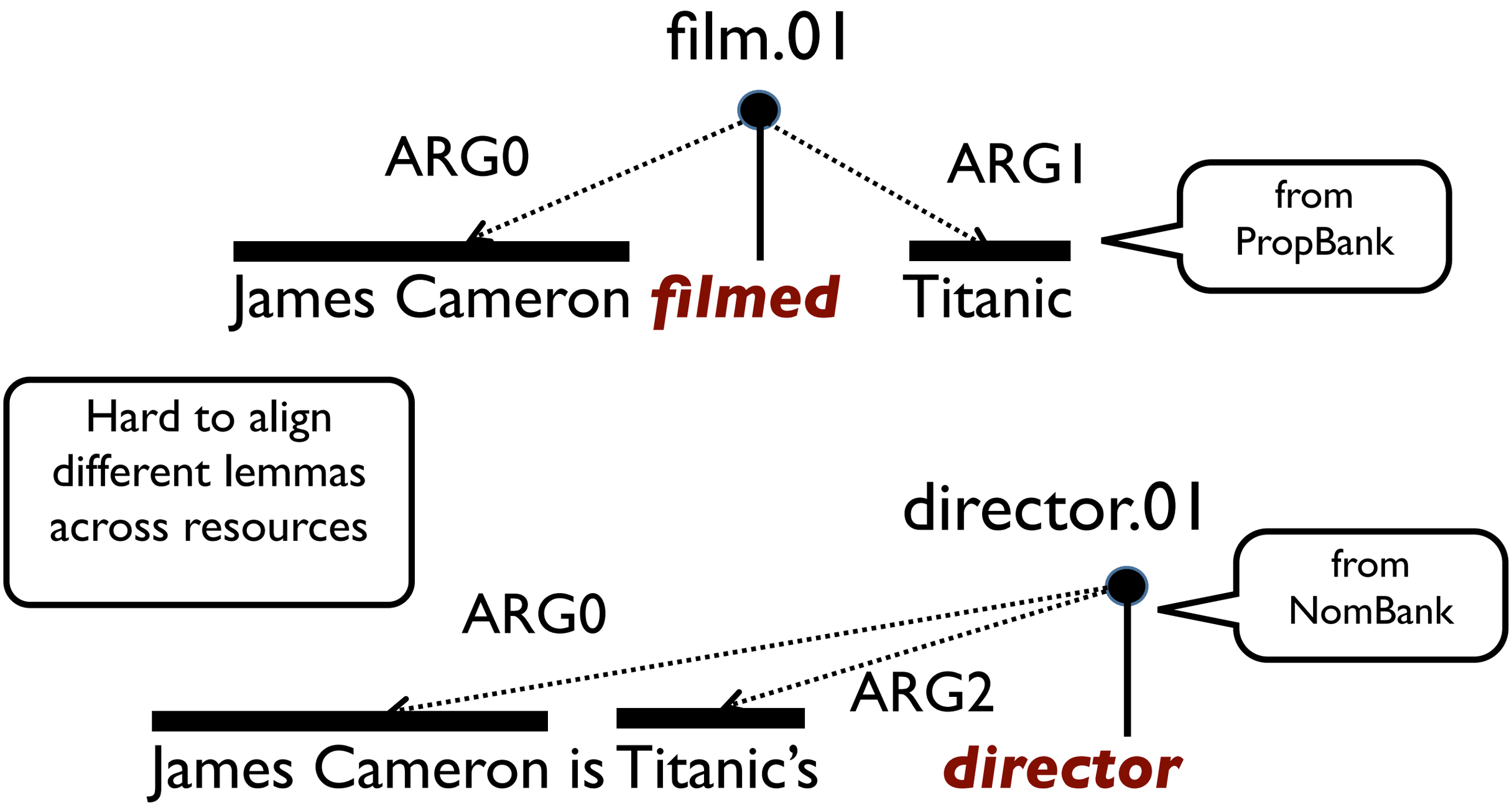
Grouping across different lemmas



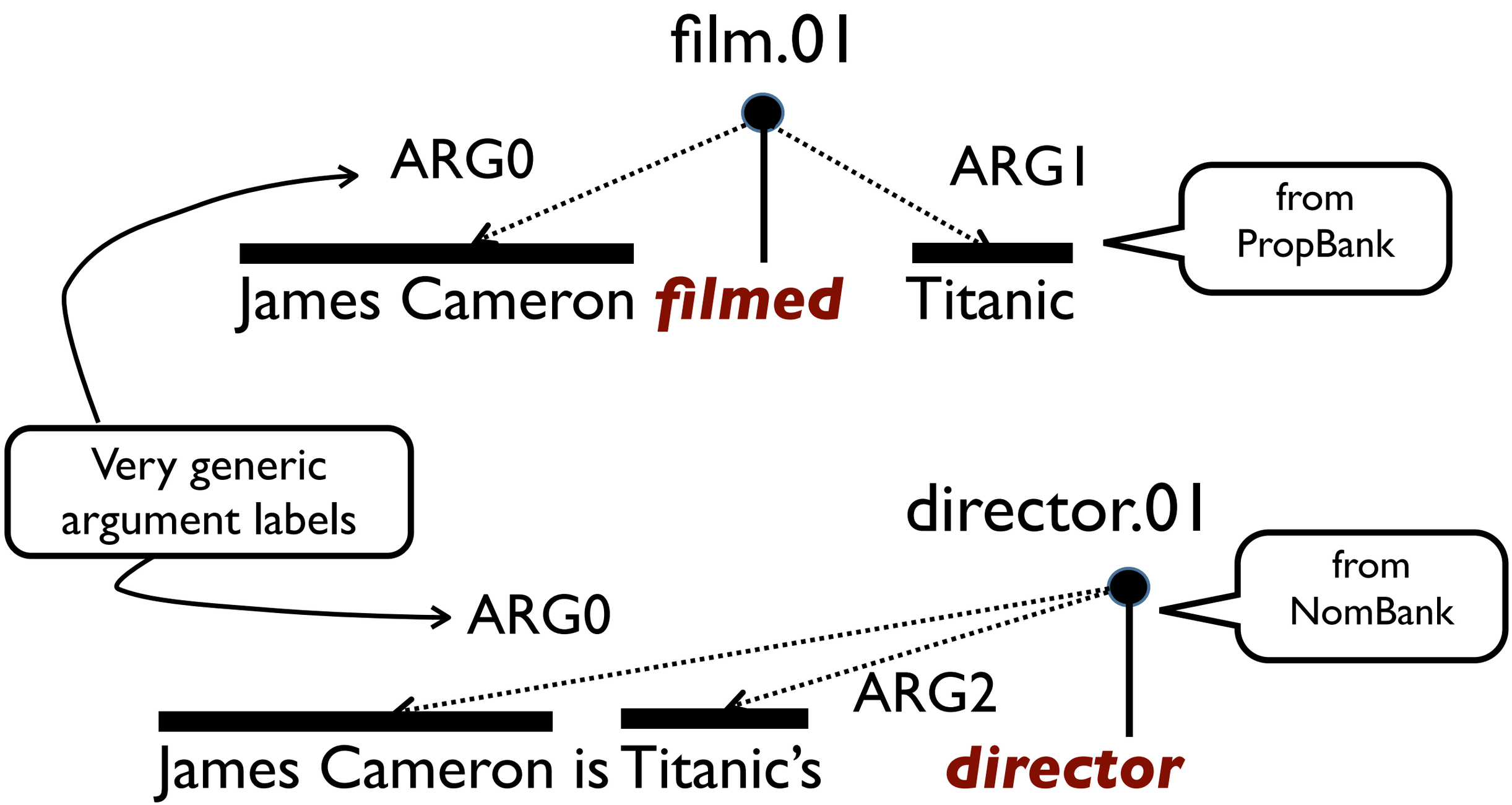
Motivation


film.01
ARG0
ARG1
from PropBank
James Cameron filmed Titanic
director.01
from NomBank
ARG0
ARG2
James Cameron is Titanic's director

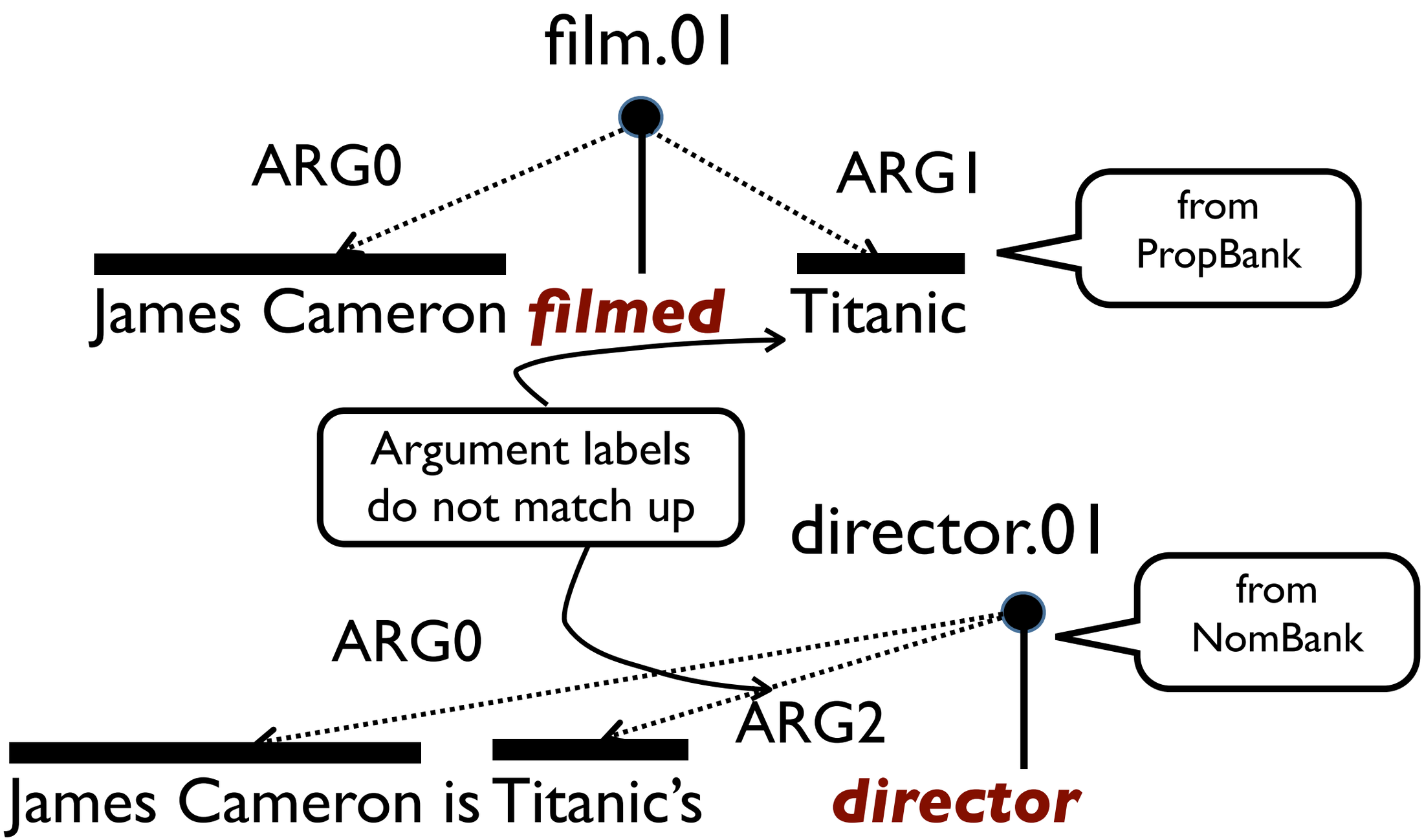
Motivation


film.01
ARG0
ARG1
from PropBank
James Cameron filmed Titanic
Hard to align different lemmas across resources
director.01
from NomBank
ARG0
ARG2
James Cameron is Titanic's director

Motivation

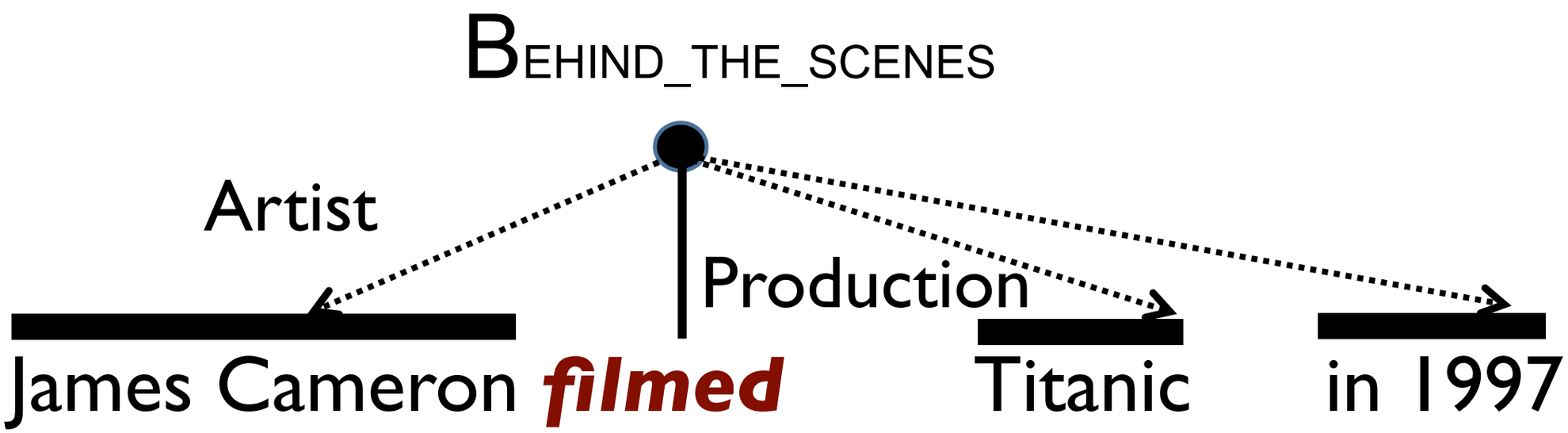

film.01
ARG0
ARG1
from PropBank
James Cameron filmed Titanic
Very generic argument labels
director.01
from NomBank
ARG0
ARG2
James Cameron is Titanic's director

Motivation

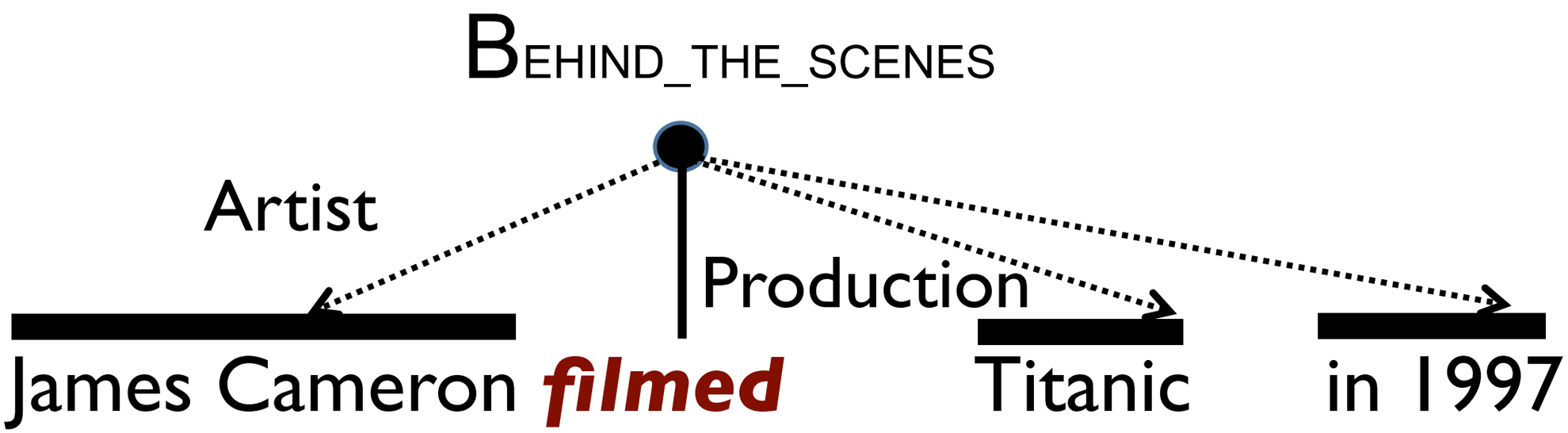

film.01
ARG0
ARG1
from PropBank
James Cameron filmed Titanic
Argument labels do not match up
director.01
from NomBank
ARG0
ARG2
James Cameron is Titanic's director

Research at Google

Motivation

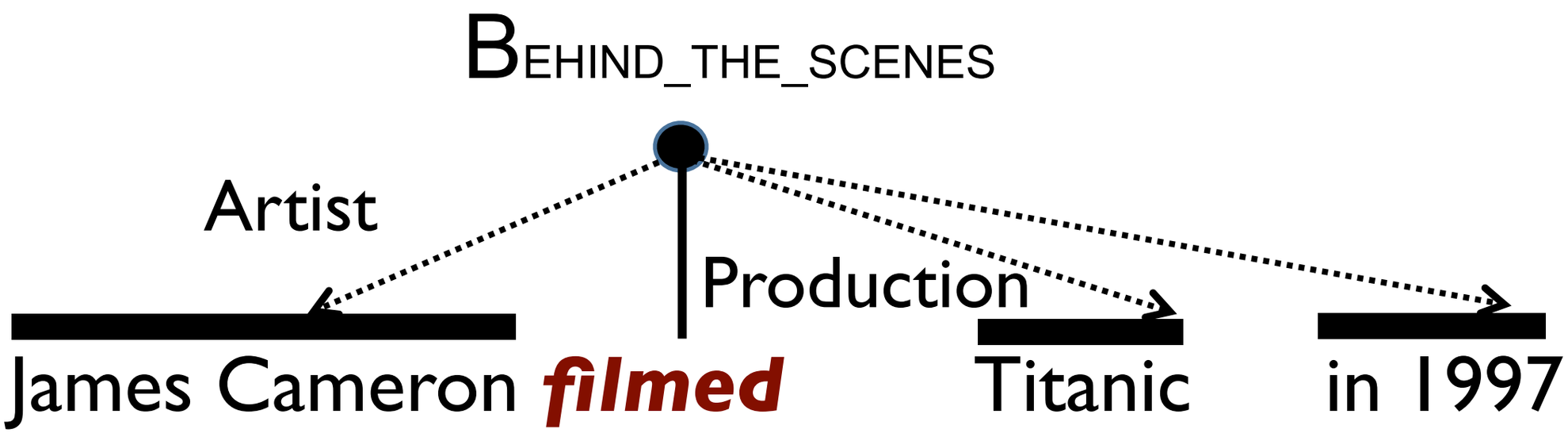


Motivation



$$\lambda e.\text{filmed.arg1}(e, \text{James\_Cameron}) \wedge \text{filmed.arg2}(e, \text{Titanic}) \wedge \text{filmed.in}(e, 1997)$$

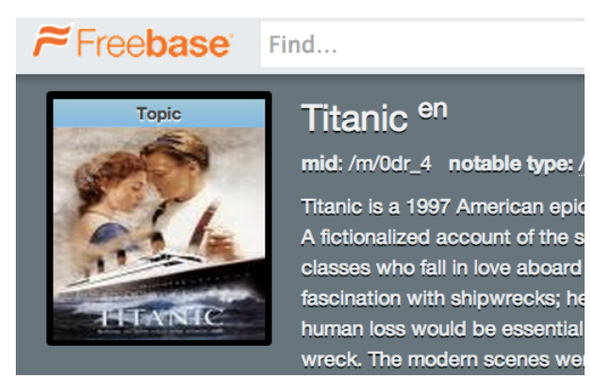
Motivation



$$\lambda e.\text{filmed.arg1}(e, \text{James\_Cameron}) \wedge \text{filmed.arg2}(e, \text{Titanic}) \wedge \text{filmed.in}(e, 1997)$$

Freebase

$$\lambda e.\text{filmed.arg1}(e, /\text{m}/03\_\text{gd}) \wedge \text{filmed.arg2}(e, /\text{m}/0\text{dr}\_4) \wedge \text{filmed.in}(e, 1997)$$



Motivation

$$\lambda e.\text{filmed.arg1}(e, \text{James\_Cameron}) \wedge \text{filmed.arg2}(e, \text{Titanic}) \wedge \text{filmed.in}(e, 1997)$$

Freebase

$$\lambda e.\text{filmed.arg1}(e, /\text{m}/03\_\text{gd}) \wedge \text{filmed.arg2}(e, /\text{m}/0\text{dr}\_4) \wedge \text{filmed.in}(e, 1997)$$

Motivation

$$\lambda e.\text{filmed.arg1}(e, \text{James\_Cameron}) \wedge \text{filmed.arg2}(e, \text{Titanic}) \wedge \text{filmed.in}(e, 1997)$$

Freebase

$$\lambda e.\text{filmed.arg1}(e, /\text{m}/03\_\text{gd}) \wedge \text{filmed.arg2}(e, /\text{m}/0\text{dr}\_4) \wedge \text{filmed.in}(e, 1997)$$

FrameNet

$$\lambda e.\textsc{Behind\_the\_scenes}.\text{Artist}(e, /\text{m}/03\_\text{gd}) \wedge \textsc{Behind\_the\_scenes}.\text{Production}(e, /\text{m}/0\text{dr}\_4) \wedge \textsc{Behind\_the\_scenes}.\text{Time}(e, 1997)$$

Research at Google

Applications of Frame-Semantic Parsing

Research at Google

Stance Classification

- Frame Semantics for Stance Classification
  Hasan and Ng (CoNLL 2013)
  - Two sided debates in an online forum
  - Classification of stance
  - Improvement over a baseline that uses bag of words and dependencies

Research at Google

Dialog Systems

- Unsupervised Induction and Filling of Semantic Slots for Spoken Dialogue Systems Using Frame-Semantic Parsing
  Chen, Wang and Rudnicky (ASRU 2013)
  - Annotation of dialog transcripts with frame-semantic structures
  - Uses only a subset of frames
  - Uses these annotations for slot induction

Research at Google

Stock Price Movement

- Semantic Frames to Predict Stock Price Movement
  Xie et al. (ACL 2013)
  - Predict the change in stock price from financial news
  - Lots of features along with features based on frames and roles
  - Shows improvements over other features

Research at Google

Summarization

- Generating Automated Meeting Summaries
  Thomas Kleinbauer (PhD thesis, Saarland University)
  - Part of a large system for generating meeting summaries

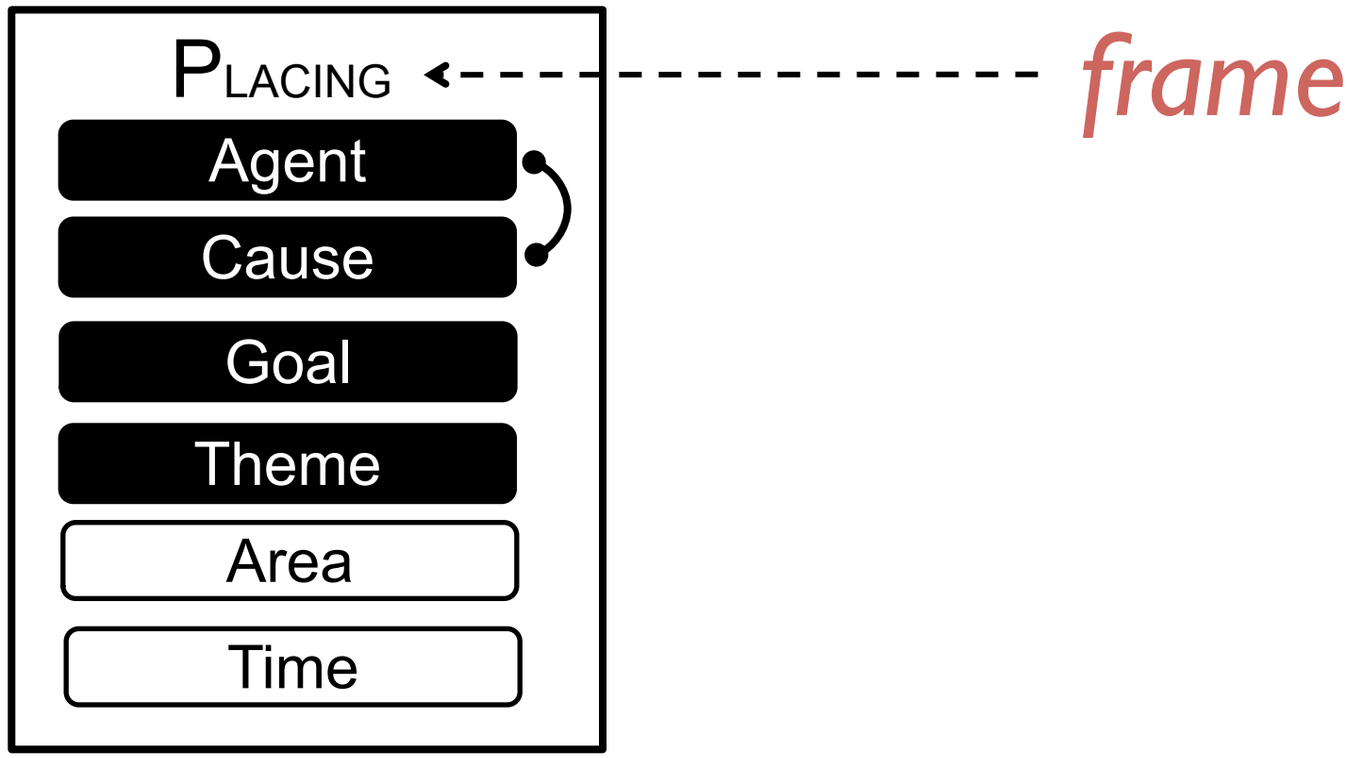
Research at Google

Outline

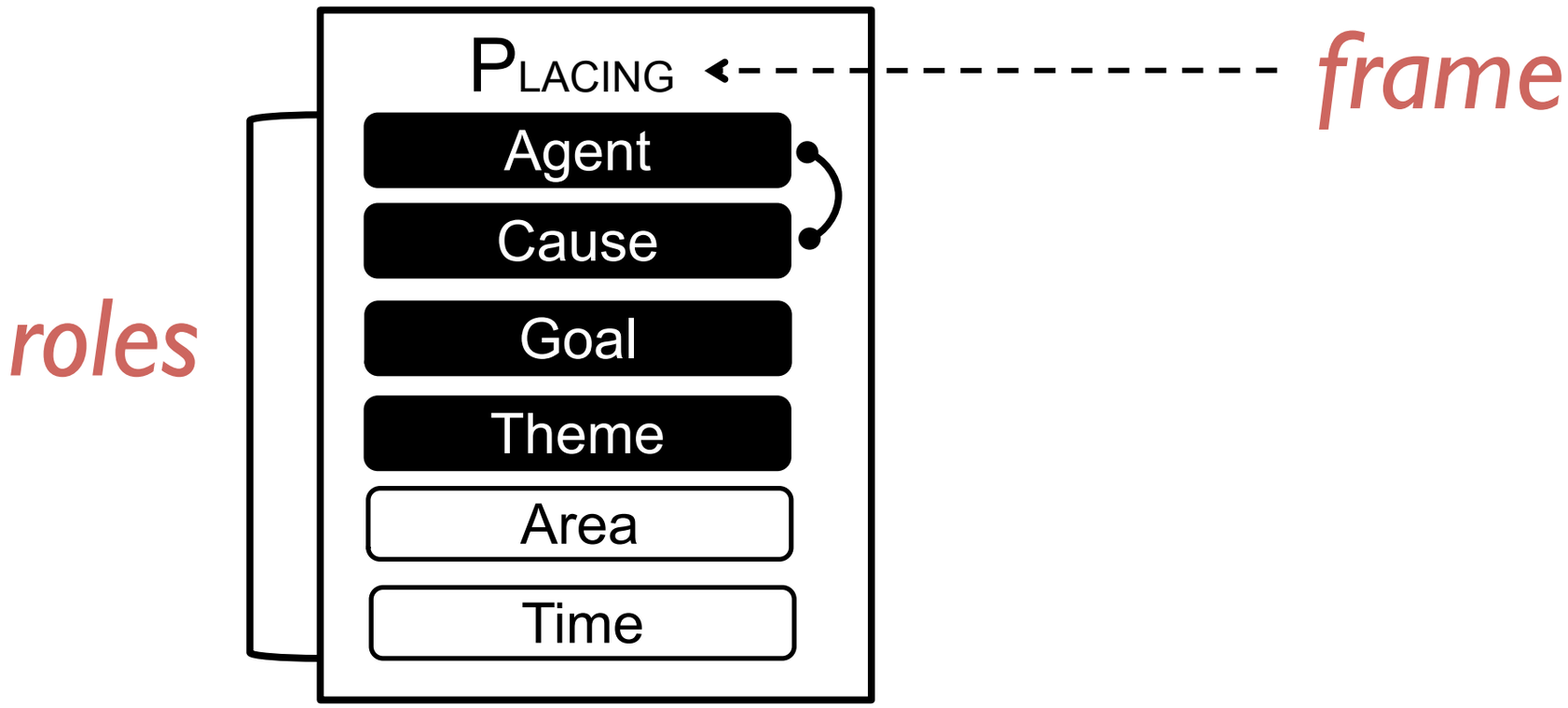
- Why should we build statistical models for frame semantics?
- Overview of statistical methods
- Future directions

Structure of Lexicon and Data

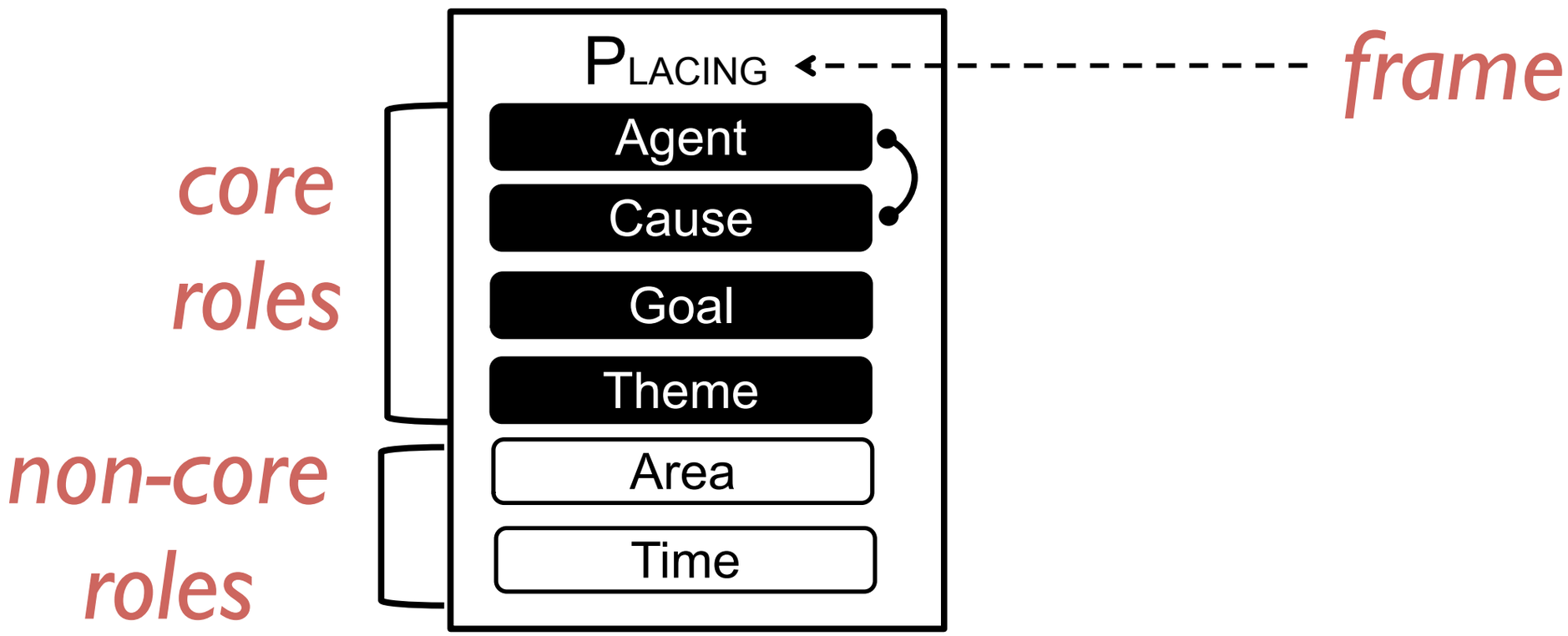
Structure of Lexicon and Data



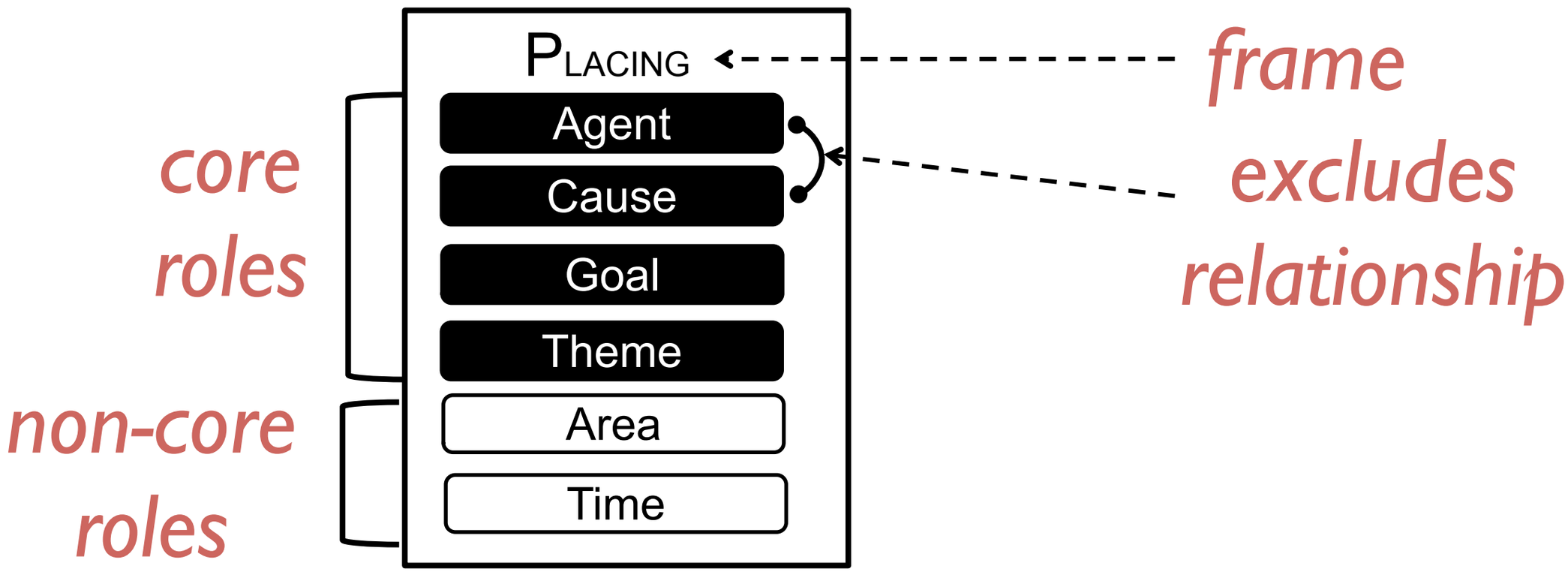
Structure of Lexicon and Data



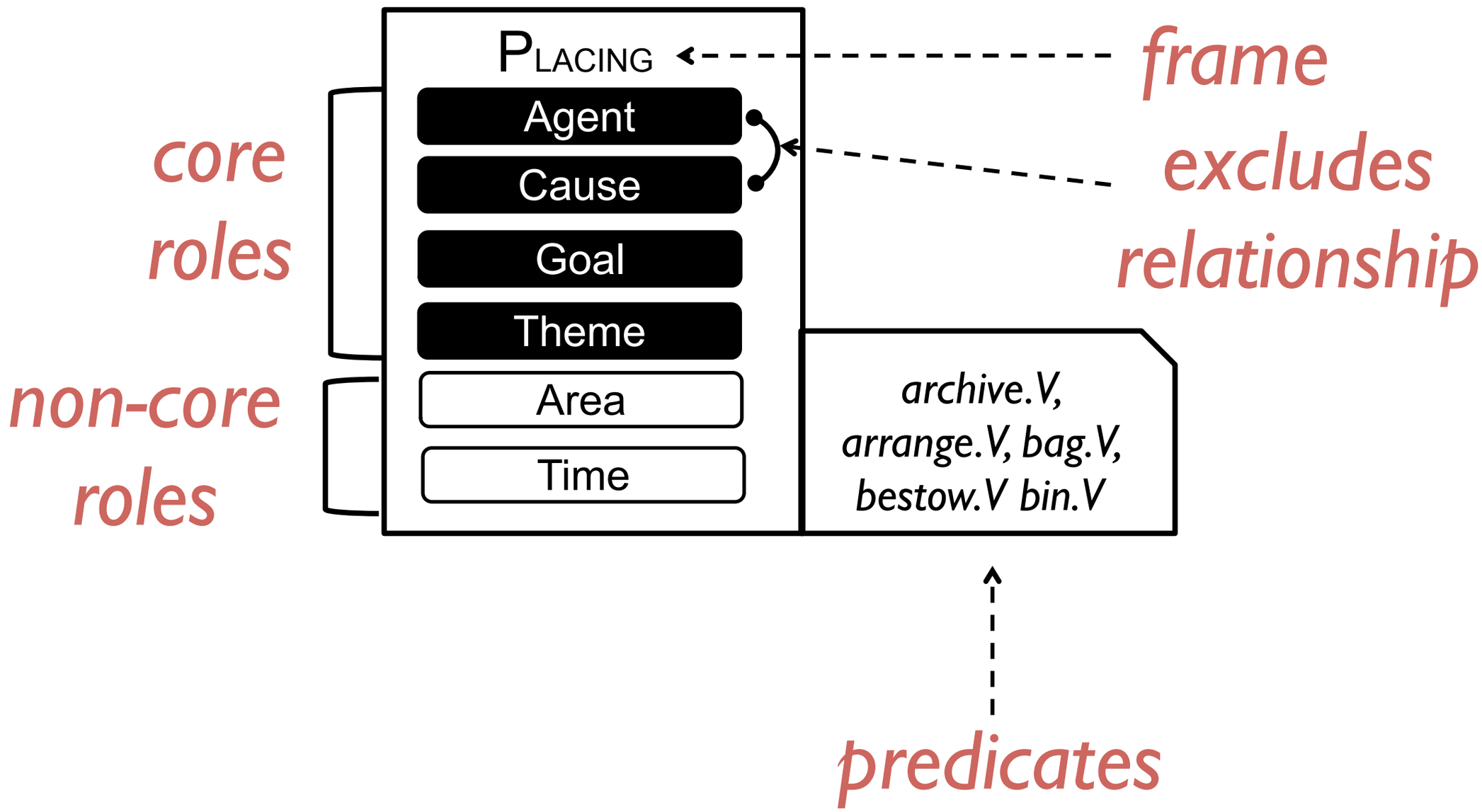
Structure of Lexicon and Data



Structure of Lexicon and Data

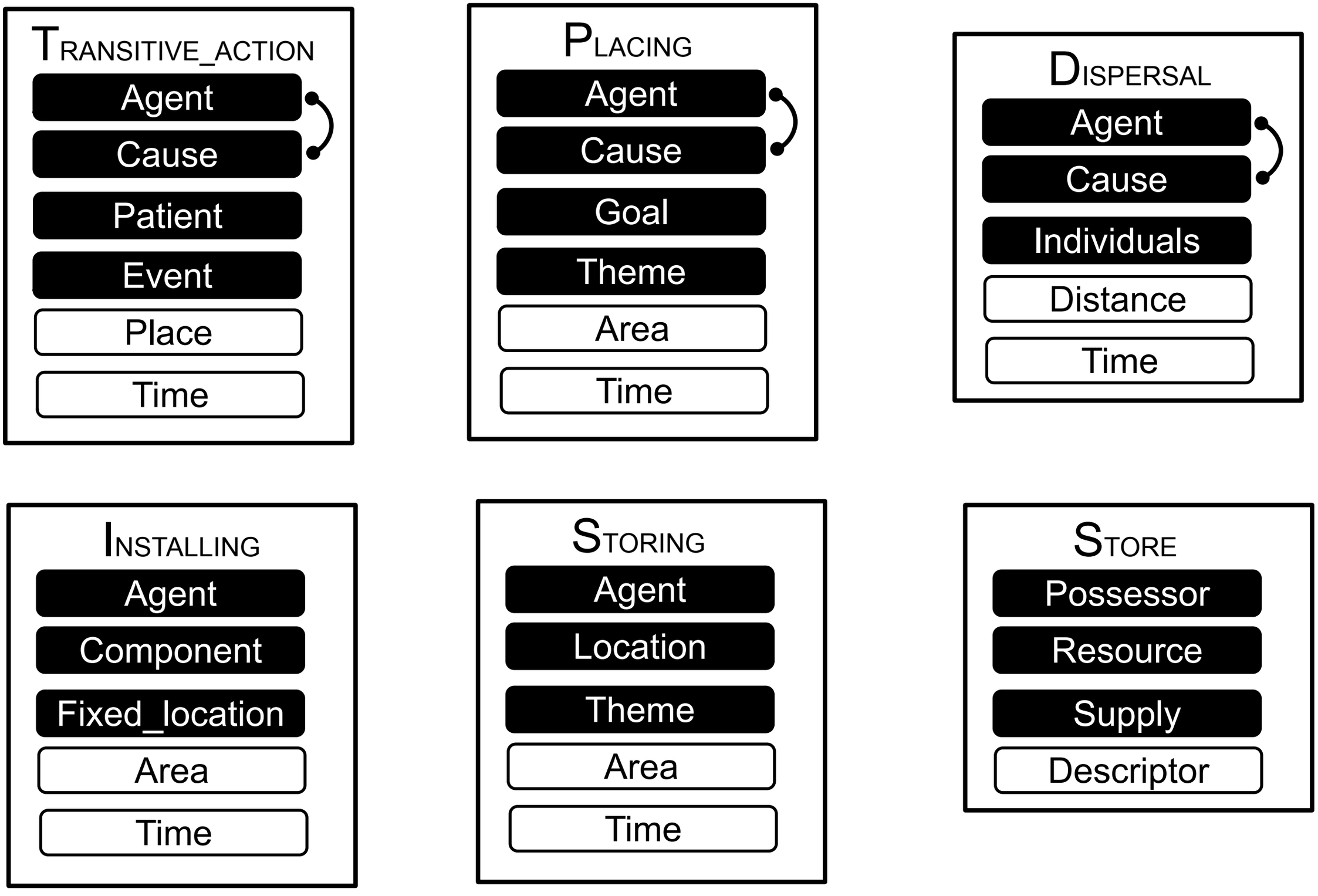


Structure of Lexicon and Data



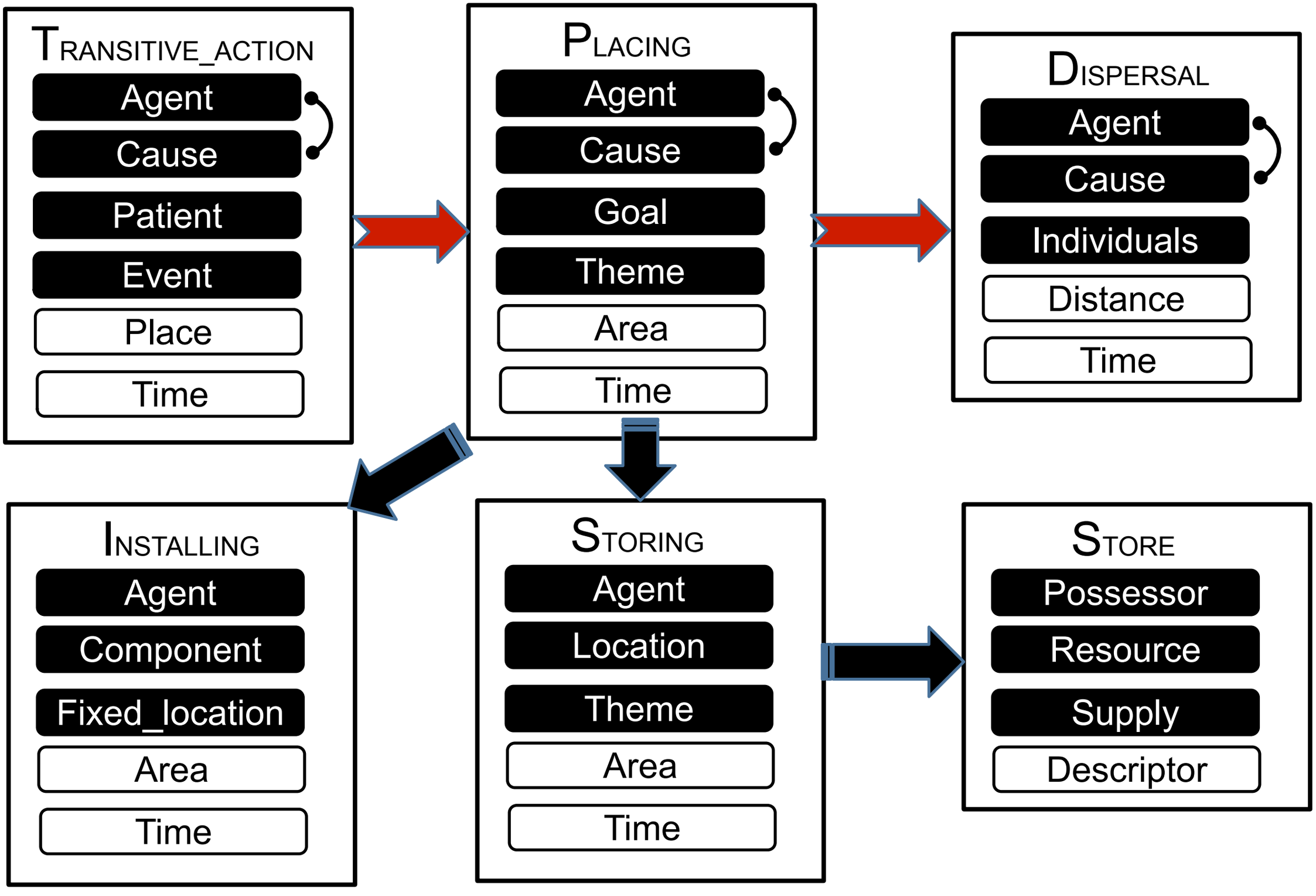
Research at Google

Structure of Lexicon and Data



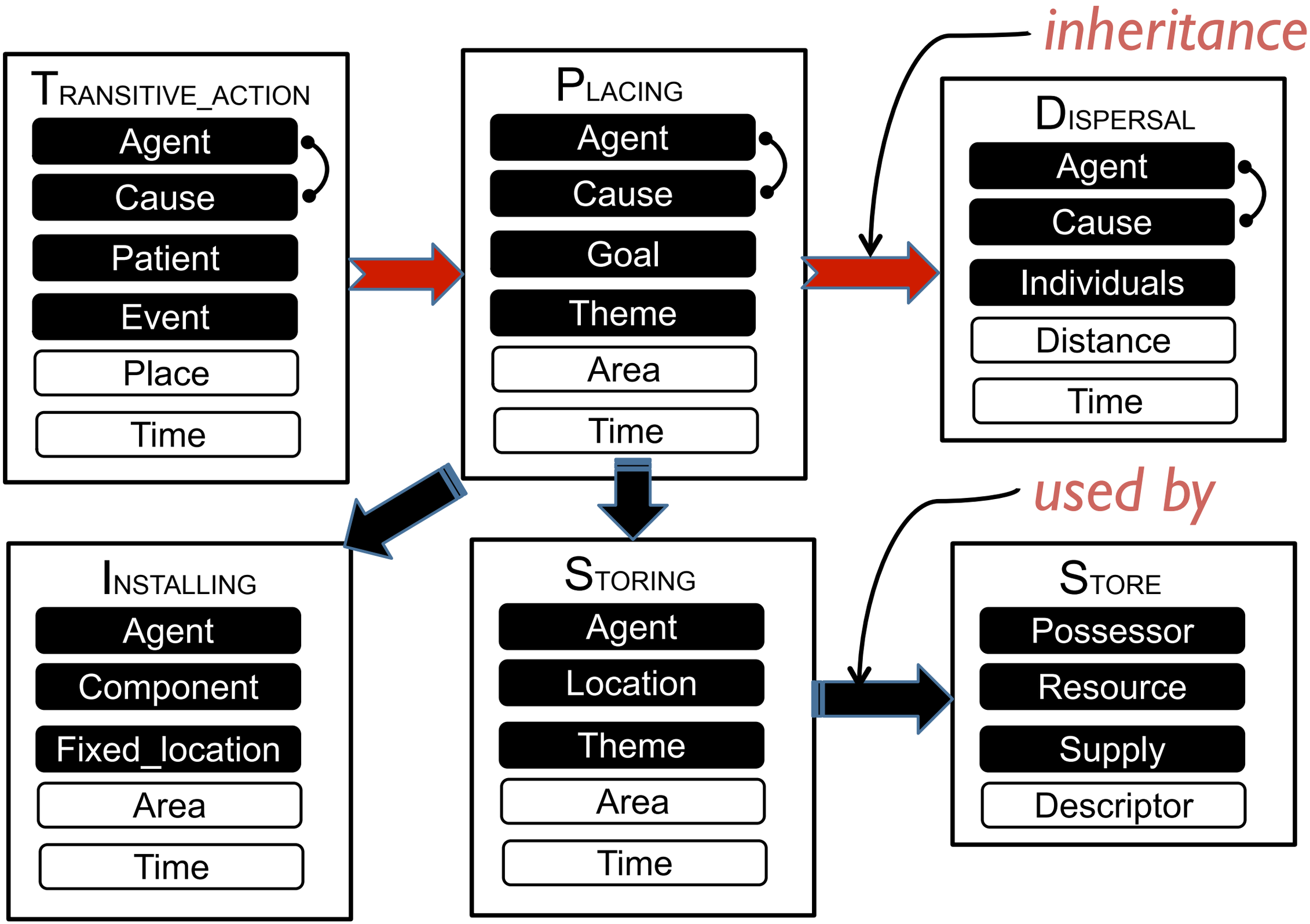
Research at Google

Structure of Lexicon and Data



Research at Google

Structure of Lexicon and Data



Research at Google

Datasets

Benchmark Dataset
(SemEval 2007)

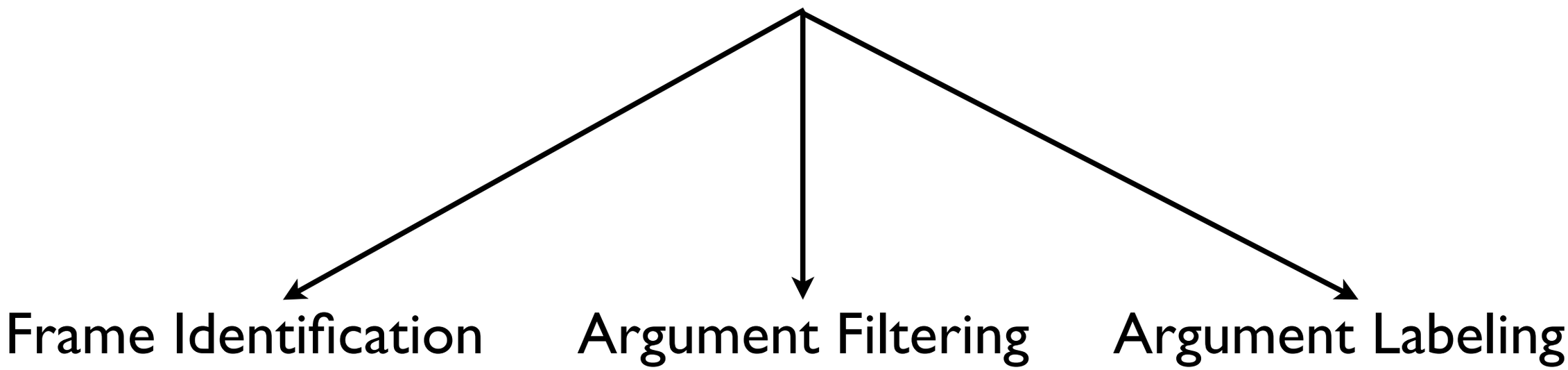
665 frames
720 role labels
8.4K unique predicate types

Training set:
2.2K sentences
11.2K predicate tokens

Test set:
120 sentences
1. 1K predicate tokens

LTH Frame-Semantic Parser

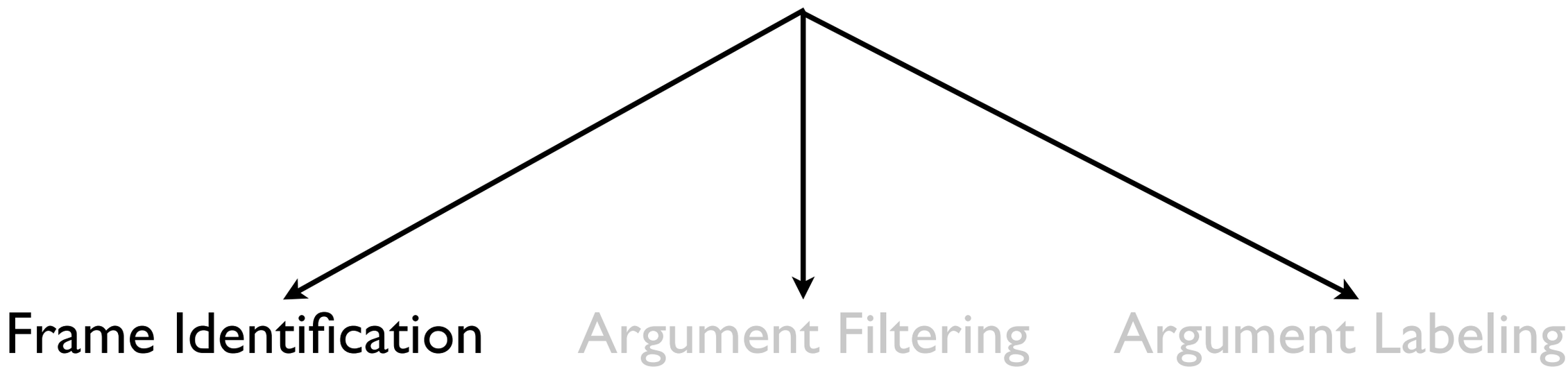
Johansson and Nugues (2007)



Research at Google

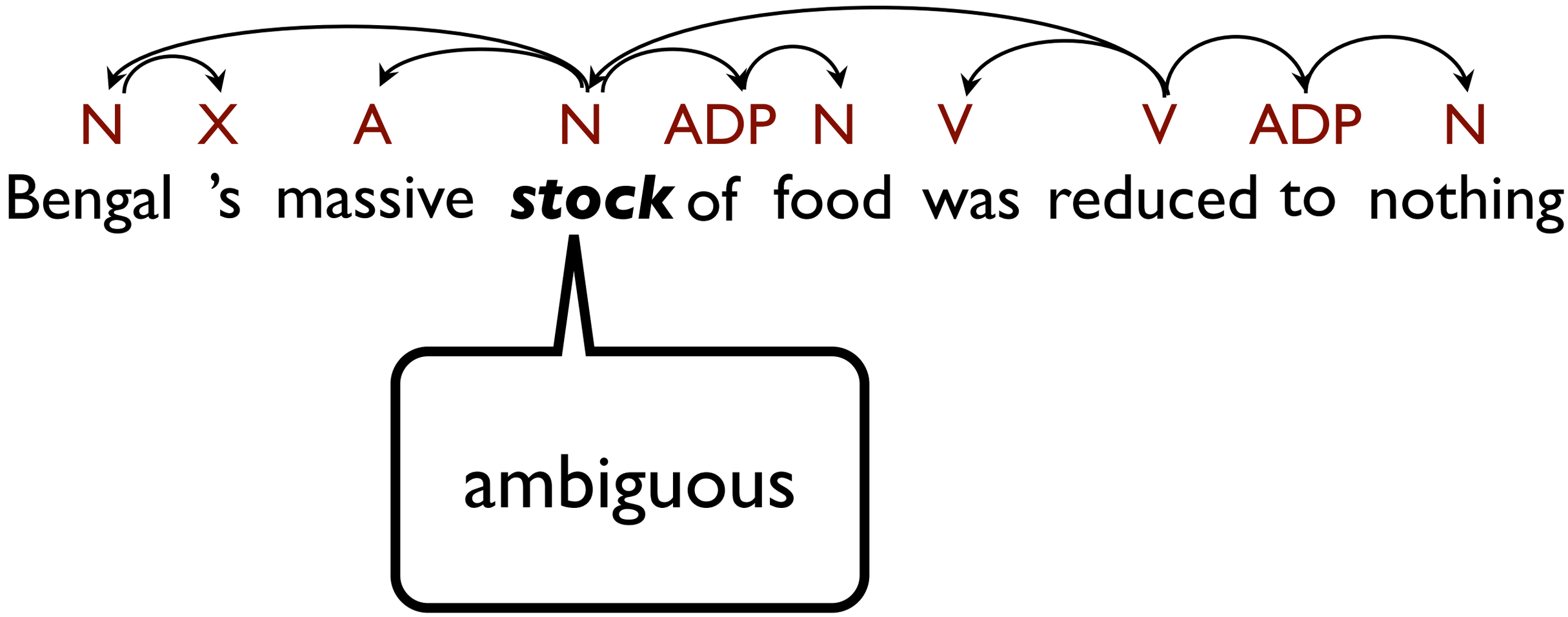
LTH Frame-Semantic Parser

Johansson and Nugues (2007)



LTH Frame-Semantic Parser

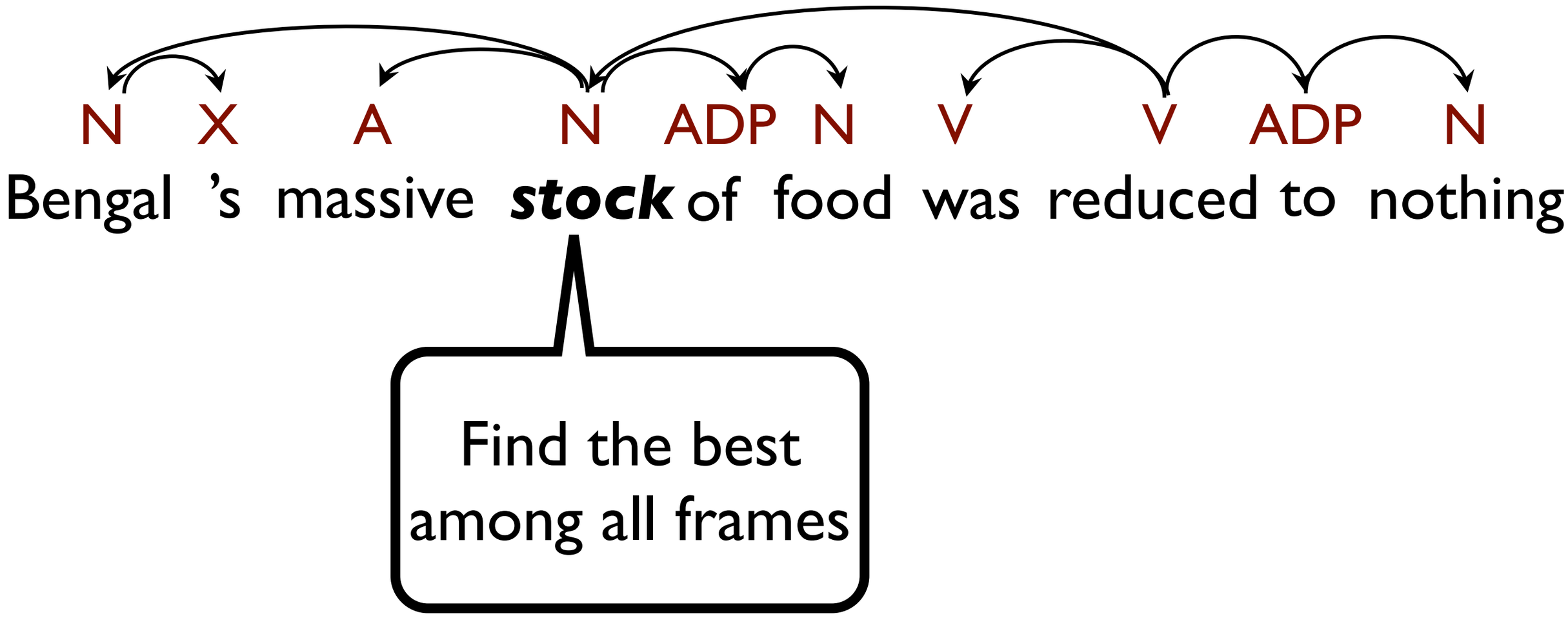
Johansson and Nugues (2007)



Research at Google

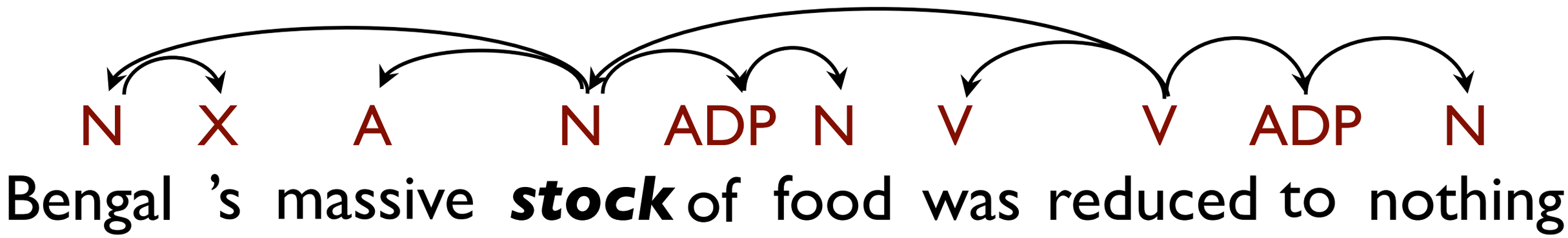
LTH Frame-Semantic Parser

Johansson and Nugues (2007)



LTH Frame-Semantic Parser

Johansson and Nugues (2007)

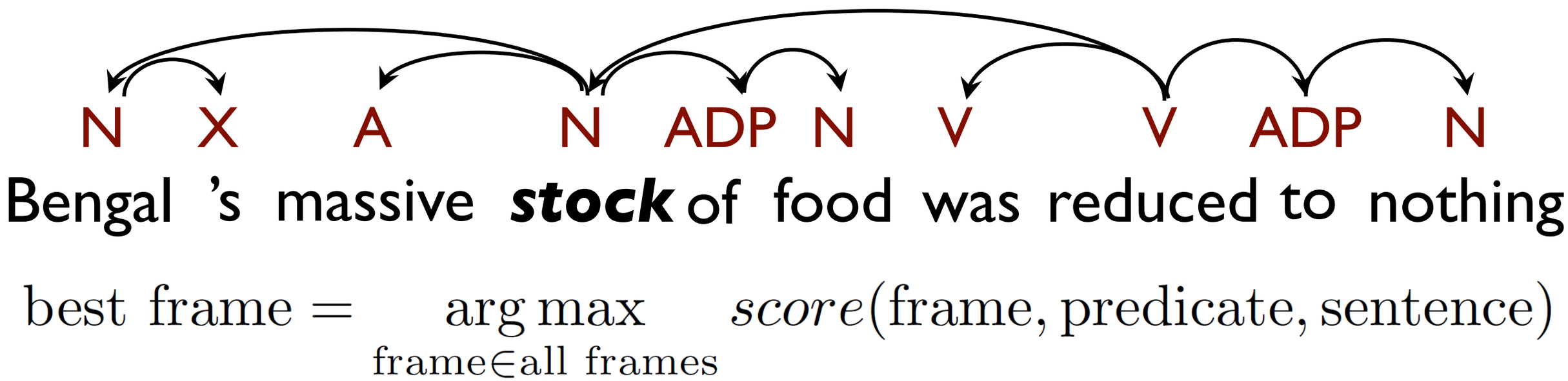


$$\text{best frame} = \underset{\text{frame} \in \text{all frames}}{\arg\max} \; score(\text{frame}, \text{predicate}, \text{sentence})$$

Taken from an SVM classifier trained on ambiguous predicates

LTH Frame-Semantic Parser

Johansson and Nugues (2007)

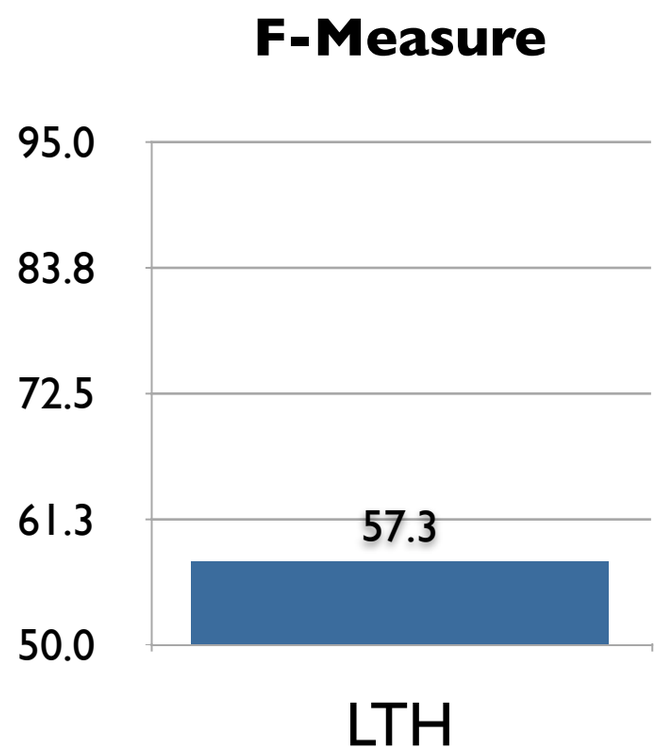


$$\text{best frame} = \underset{\text{frame} \in \text{all frames}}{\arg\max} \; score(\text{frame}, \text{predicate}, \text{sentence})$$

To increase coverage, potential predicates were extracted from WordNet and automatically frame labels were selected for them.

LTH Frame-Semantic Parser

Johansson and Nugues (2007)

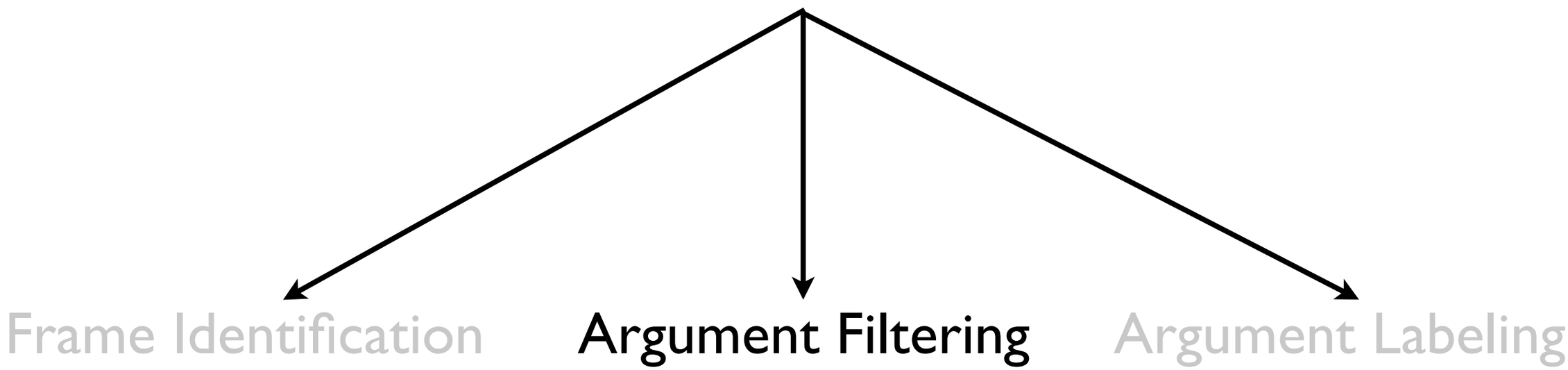


Frame Identification Accuracy

Research at Google

LTH Frame-Semantic Parser

Johansson and Nugues (2007)



Argument Filtering

STORE

Bengal 's massive ***stock*** of food was reduced to nothing

Research at Google

Argument Filtering

STORE

Bengal 's massive ***stock*** of food was reduced to nothing

Potential Arguments

Bengal 's | Bengal | massive stock | of food

food | massive | Bengal 's massive | to nothing

to | of | 's

Research at Google

Argument Filtering

STORE

Bengal 's massive ***stock*** of food was reduced to nothing

Potential Arguments

Binary SVM Classification

Bengal 's ✓

Bengal ✗

massive stock ✗

of food ✓

food ✓

massive ✓

Bengal 's massive ✗

to nothing ✗

to ✗

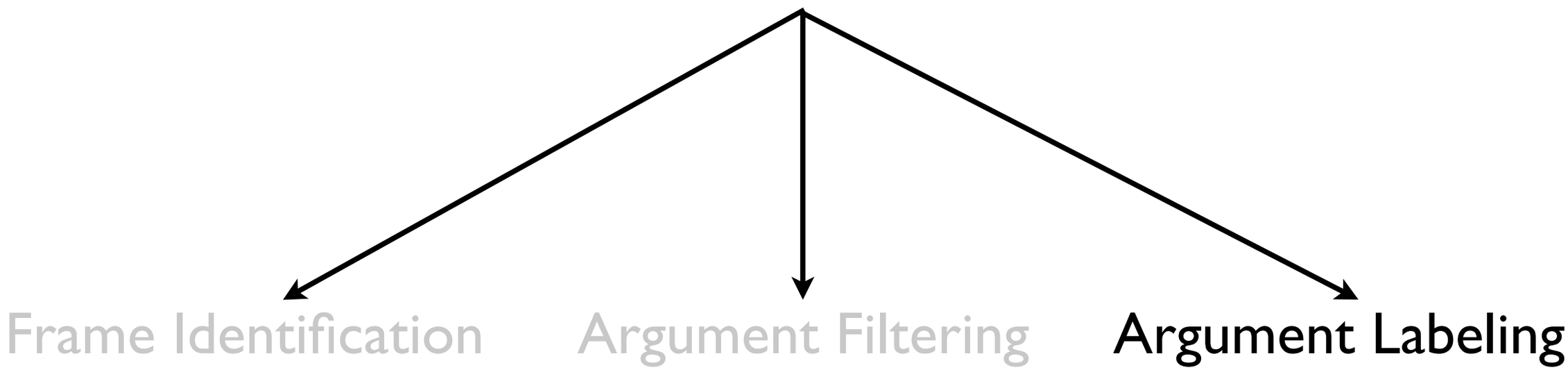
of ✗

's ✗

Research at Google

LTH Frame-Semantic Parser

Johansson and Nugues (2007)



Research at Google

Argument Filtering

STORE

Bengal 's massive ***stock*** of food was reduced to nothing

Multiclass SVM Classification

Research at Google

Argument Filtering

STORE

Bengal 's massive ***stock*** of food was reduced to nothing

Potential Arguments

Multiclass SVM Classification

Bengal 's | Bengal | massive stock | of food

food | massive | Bengal 's massive | to nothing

to | of | 's

Research at Google

Argument Filtering

STORE

Bengal 's massive ***stock*** of food was reduced to nothing

Potential Arguments

Multiclass SVM Classification

Bengal 's — Possessor

Bengal

massive stock

of food — Resource

food — Ø

massive — Descriptor

Bengal 's massive

to nothing

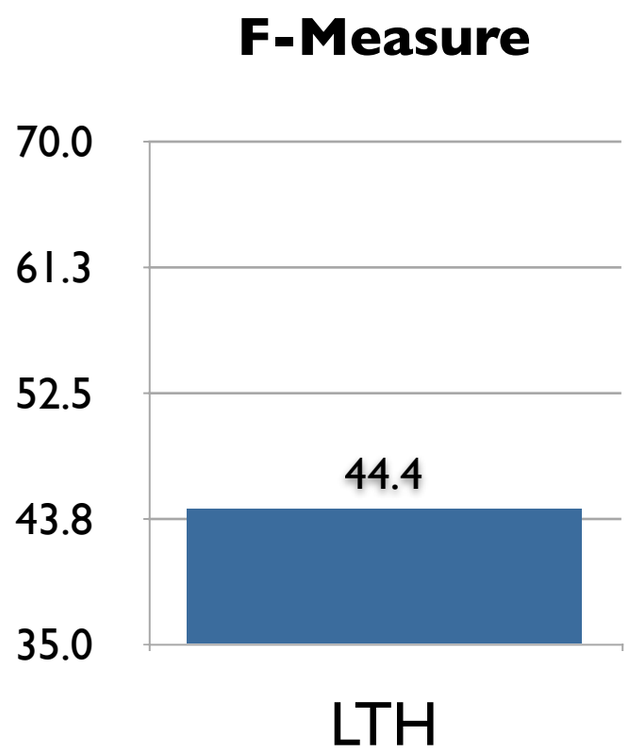
to

of

's

LTH Frame-Semantic Parser

Johansson and Nugues (2007)



Full Frame-Semantic Structure Prediction

Research at Google

SEMAFOR

Das, Chen, Martins, Schneider, Smith (2014)

Frame Identification

Argument Identification

Research at Google

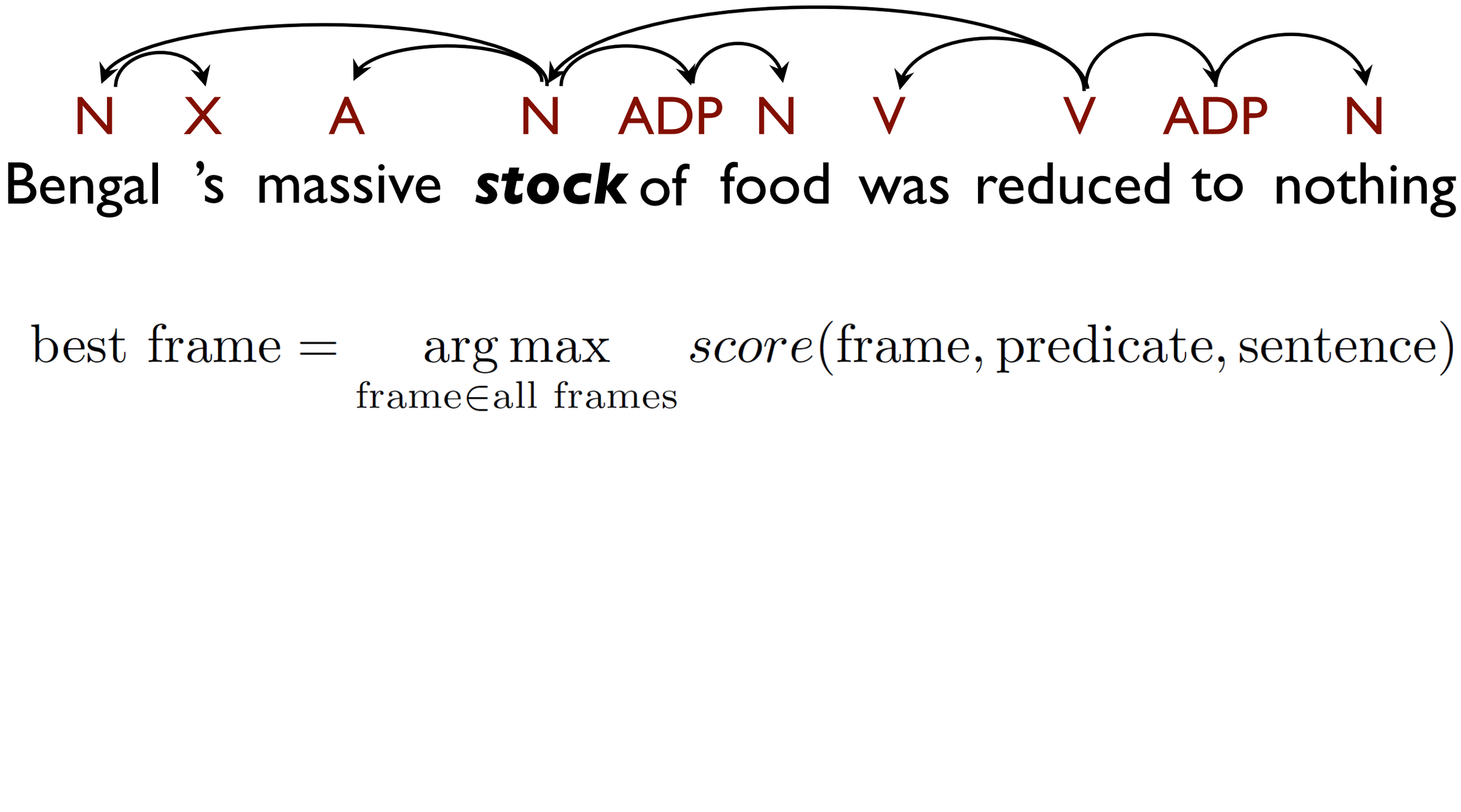
SEMAFOR

Das, Chen, Martins, Schneider, Smith (2014)

Frame Identification

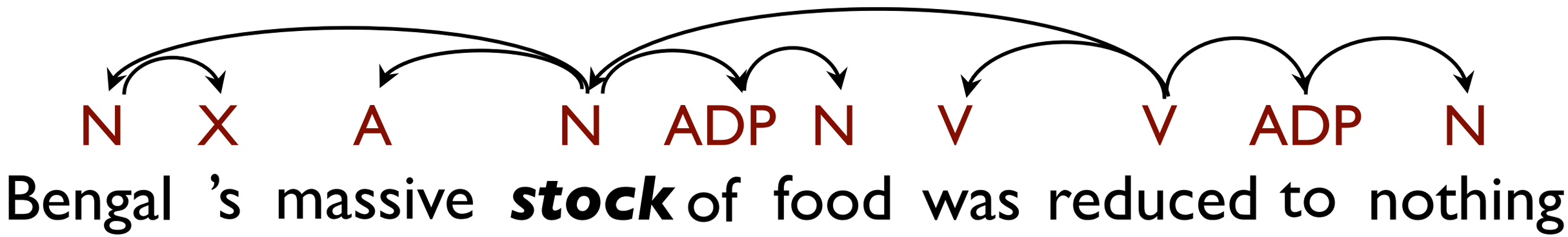
Argument Identification

SEMAFOR: Frame Identification



$$\text{best frame} = \underset{\text{frame} \in \text{all frames}}{\arg\max} \; score(\text{frame}, \text{predicate}, \text{sentence})$$

SEMAFOR: Frame Identification



$$\text{best frame} = \underset{\text{frame} \in \text{all frames}}{\arg\max} \; score(\text{frame}, \text{predicate}, \text{sentence})$$

$$= \log p(\text{frame} \mid \text{predicate}, \text{sentence})$$

SEMAFOR: Frame Identification

$$\text{best frame} = \underset{\text{frame} \in \text{all frames}}{\arg\max} \; score(\text{frame}, \text{predicate}, \text{sentence})$$

$$= \log p(\text{frame} \mid \text{predicate}, \text{sentence})$$

Logistic regression
with a latent variable

$$p(\text{frame} \mid \text{predicate}, \text{sentence}) =$$

SEMAFOR: Frame Identification

$$\text{best frame} = \underset{\text{frame} \in \text{all frames}}{\arg\max} \; score(\text{frame}, \text{predicate}, \text{sentence})$$

$$= \log p(\text{frame} \mid \text{predicate}, \text{sentence})$$

Logistic regression
with a latent variable

$$p(\text{frame} \mid \text{predicate}, \text{sentence}) =$$

$$\frac{1}{Z} \sum_{\substack{\text{proto-predicates} \\ \text{evoking frame}}} \exp \boldsymbol{w} \cdot \boldsymbol{f}(\text{frame}, \text{proto-predicate}, \text{sentence}, \text{lexical-semantic relations between predicate and proto-predicate})$$

SEMAFOR: Frame Identification

$$\text{best frame} = \underset{\text{frame} \in \text{all frames}}{\arg\max} \; score(\text{frame}, \text{predicate}, \text{sentence})$$

$$= \log p(\text{frame} \mid \text{predicate}, \text{sentence})$$

Logistic regression
with a latent variable

$$p(\text{frame} \mid \text{predicate}, \text{sentence}) =$$

$$\frac{1}{Z} \sum_{\substack{\text{proto-predicates} \\ \text{evoking frame}}} \exp \boldsymbol{w} \cdot \boldsymbol{f}(\text{frame}, \text{proto-predicate}, \text{sentence}, \text{lexical-semantic relations between predicate and proto-predicate})$$

Predicates evoking a frame in supervised data, e.g.

cargo.N, inventory.N, reserve.N, stockpile.N, store.N, supply.N

evoke STORE

SEMAFOR: Frame Identification

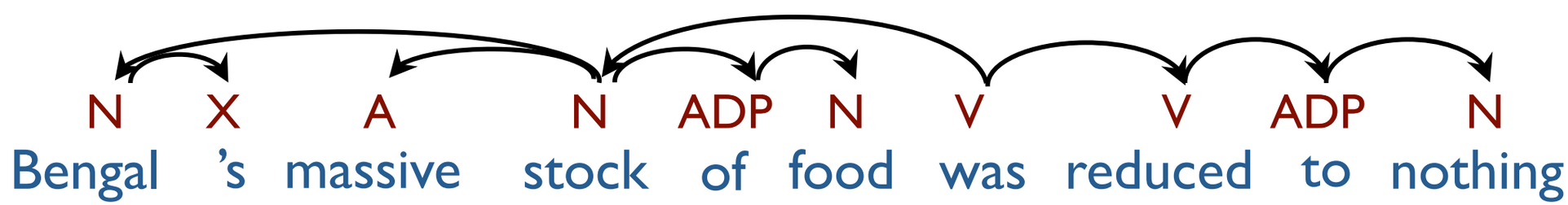
$$p(\text{frame} \mid \text{predicate}, \text{sentence}) = \frac{1}{Z} \sum_{\substack{\text{proto-predicates} \\ \text{evoking frame}}} \exp \boldsymbol{w} \cdot \boldsymbol{f}(\text{frame}, \text{proto-predicate}, \text{sentence}, \text{lexical-semantic relations between predicate and proto-predicate})$$

frame = STORE

predicate = *stock.N*

proto-predicate = *stockpile.N*

sentence =



SEMAFOR: Frame Identification

$$p(\text{frame} \mid \text{predicate}, \text{sentence}) =$$

$$\frac{1}{Z} \sum_{\substack{\text{proto-predicates} \\ \text{evoking frame}}} \exp \boldsymbol{w} \cdot \boldsymbol{f}(\text{frame}, \text{proto-predicate}, \text{sentence}, \text{lexical-semantic relations between predicate and proto-predicate})$$

frame = STORE

predicate = *stock.N*

proto-predicate = *stockpile.N*

LexSem = {*synonym*}

sentence =

N	X	A	N	ADP	N	V	V	ADP	N
Bengal	's	massive	stock	of	food	was	reduced	to	nothing

SEMAFOR: Frame Identification

$$p(\text{frame} \mid \text{predicate}, \text{sentence}) = \frac{1}{Z} \sum_{\substack{\text{proto-predicates} \\ \text{evoking frame}}} \exp \boldsymbol{w} \cdot \boldsymbol{f}(\text{frame}, \text{proto-predicate}, \text{sentence}, \text{lexical-semantic relations between predicate and proto-predicate})$$

frame = STORE

predicate = *stock.N*

proto-predicate = *stockpile.N*

LexSem = {*synonym*}

sentence =

N	X	A	N	ADP	N	V	V	ADP	N
Bengal	's	massive	stock	of	food	was	reduced	to	nothing

$f_{10245} = 1$ If

frame = STORE $\wedge$
proto-predicate = *stockpile.N*
synonym $\in$ **LexSem**

SEMAFOR: Frame Identification

$$p(\text{frame} \mid \text{predicate}, \text{sentence}) =$$

$$\frac{1}{Z} \sum_{\substack{\text{proto-predicates} \\ \text{evoking frame}}} \exp \boldsymbol{w} \cdot \boldsymbol{f}(\text{frame}, \text{proto-predicate}, \text{sentence}, \text{lexical-semantic relations between predicate and proto-predicate})$$

frame = STORE

predicate = *stock.N*

proto-predicate = *stockpile.N*

LexSem = {*synonym*}

comes from ***WordNet****!*

sentence =

N	X	A	N	ADP	N	V	V	ADP	N
Bengal	's	massive	stock	of	food	was	reduced	to	nothing

$f_{10245} = 1$ If

frame = STORE $\wedge$

proto-predicate = *stockpile.N*

synonym $\in$ **LexSem**

Research at Google

Datasets

Benchmark Dataset
(SemEval 2007)

665 frames
720 role labels
8.4K unique predicate types

Training set:
2.2K sentences
11.2K predicate tokens

Test set:
120 sentences
1. 1K predicate tokens

New Data
(FrameNet 1.5, 2010)

877 frames
1068 role labels
9.3K unique predicate types

Training set:
3.3K sentences
19.6K predicate tokens

Test set:
2420 sentences
4.5K predicate tokens

Research at Google

Datasets

Benchmark Dataset
(SemEval 2007)

665 frames
720 role labels
8.4K unique predicate types

Training set:
2.2K sentences
11.2K predicate tokens

Test set:
120 sentences
1. 1K predicate tokens

New Data
(FrameNet 1.5, 2010)

877 frames
1068 role labels
9.3K unique predicate types

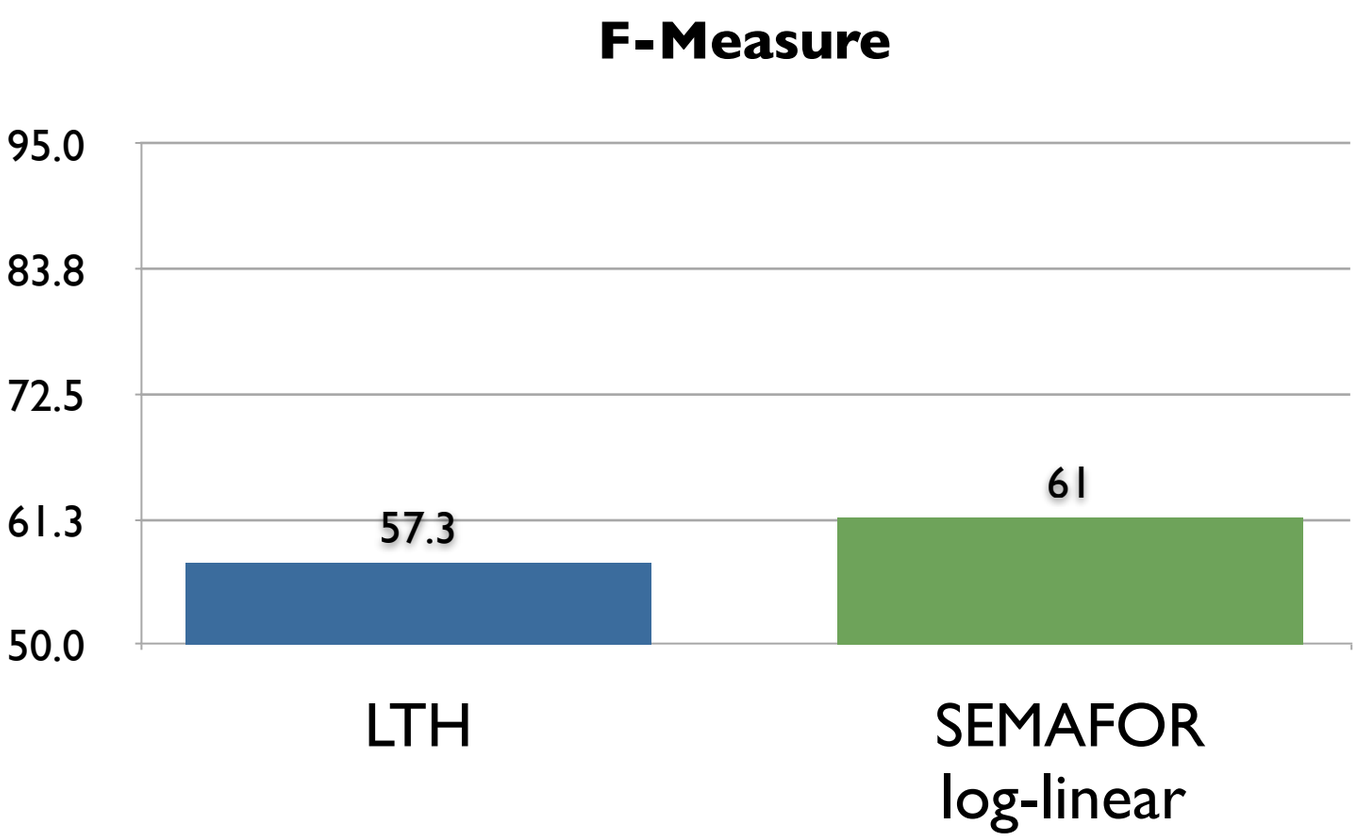
Training set:
3.3K sentences
19.6K predicate tokens

Test set:
2420 sentences
4.5K predicate tokens

SEMAFOR: Frame Identification

Results

Benchmark

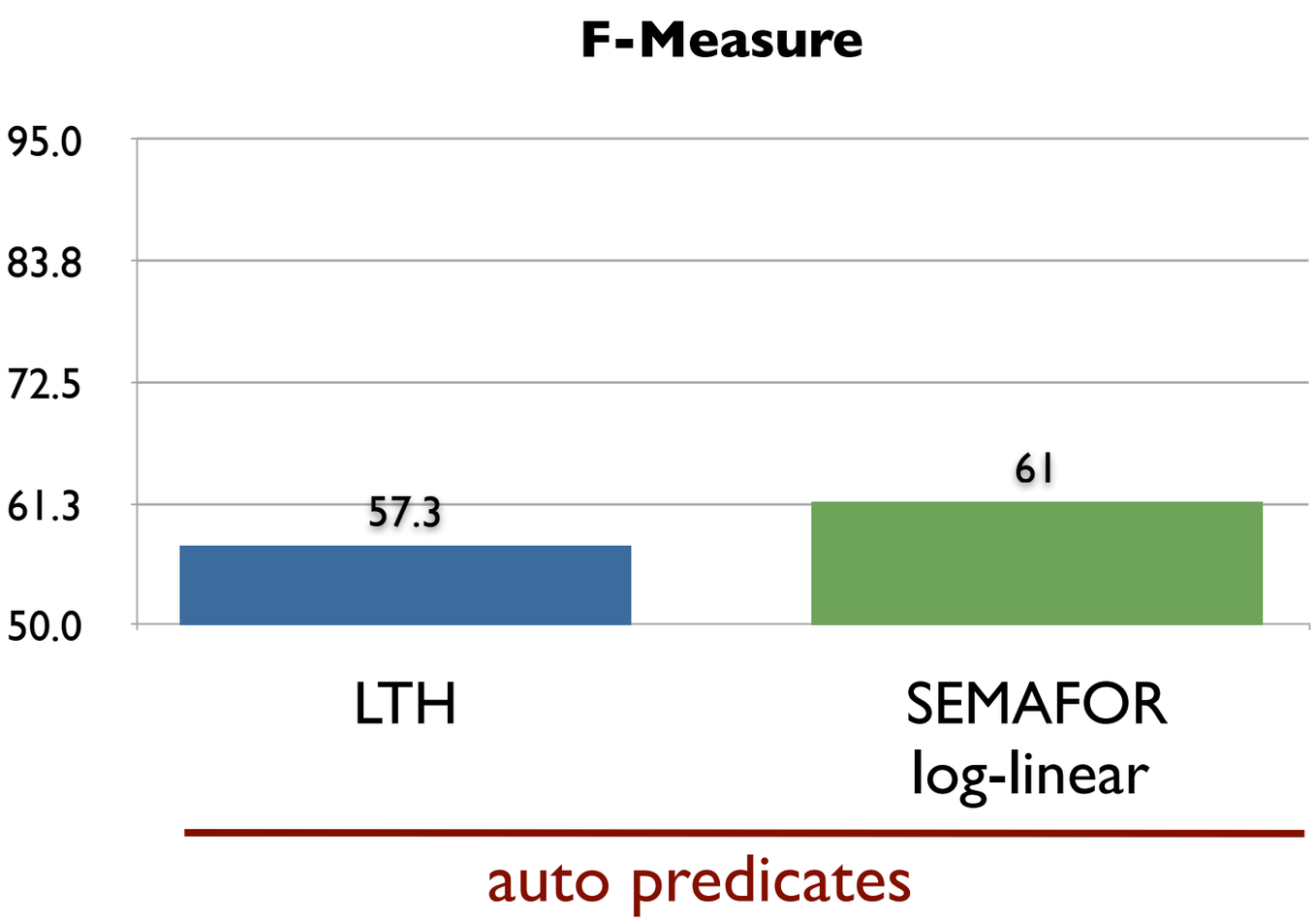


New Data

SEMAFOR: Frame Identification

Results

Benchmark

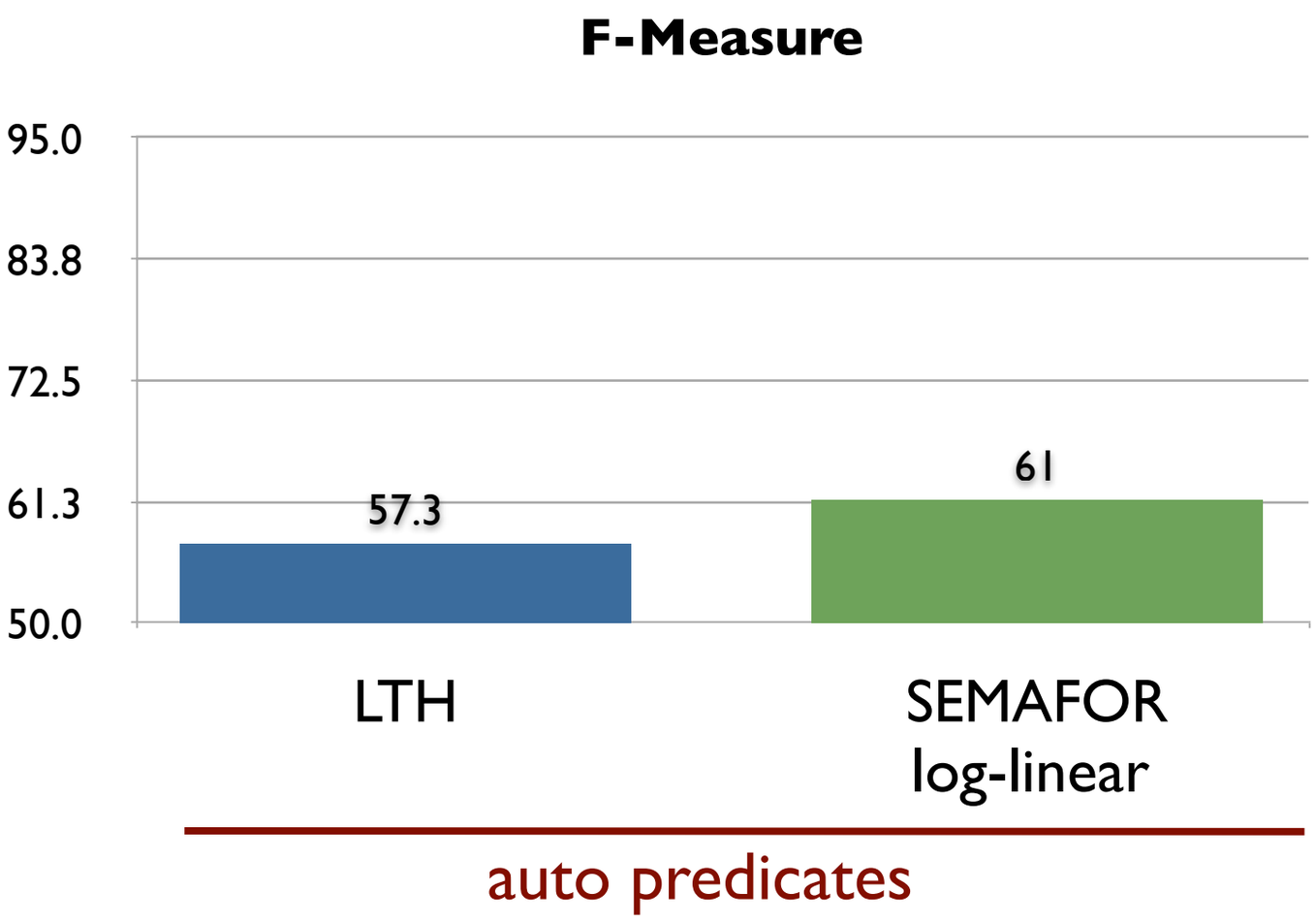


New Data

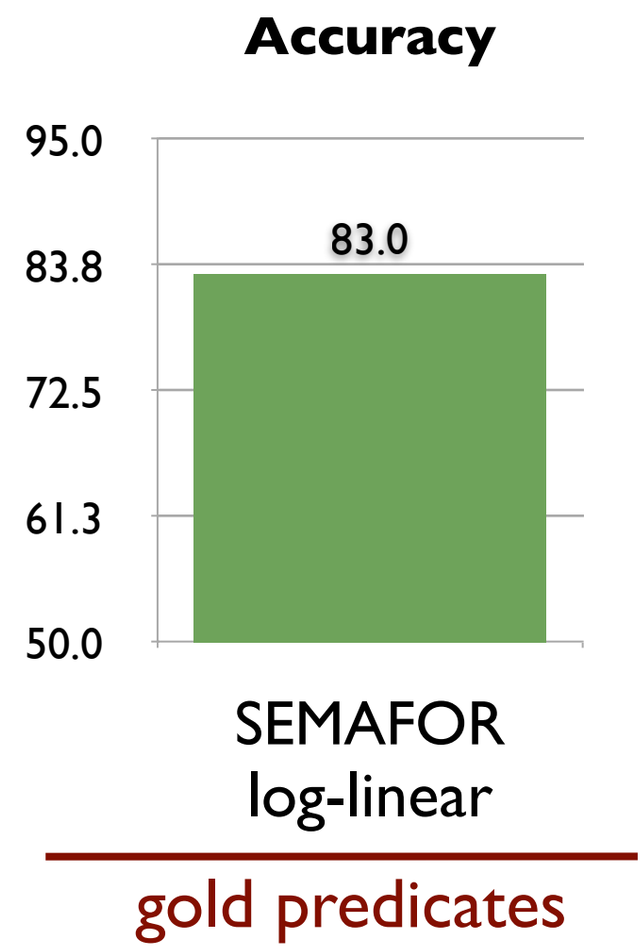
SEMAFOR: Frame Identification

Results

Benchmark

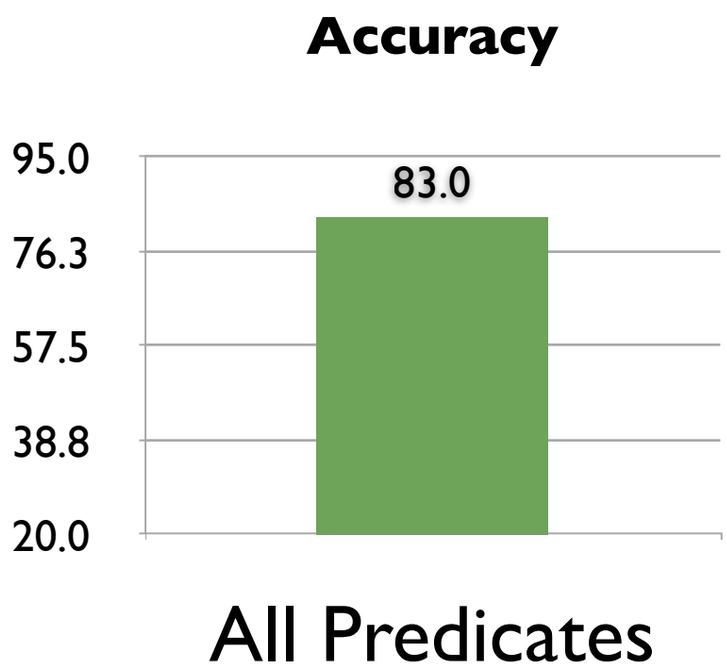


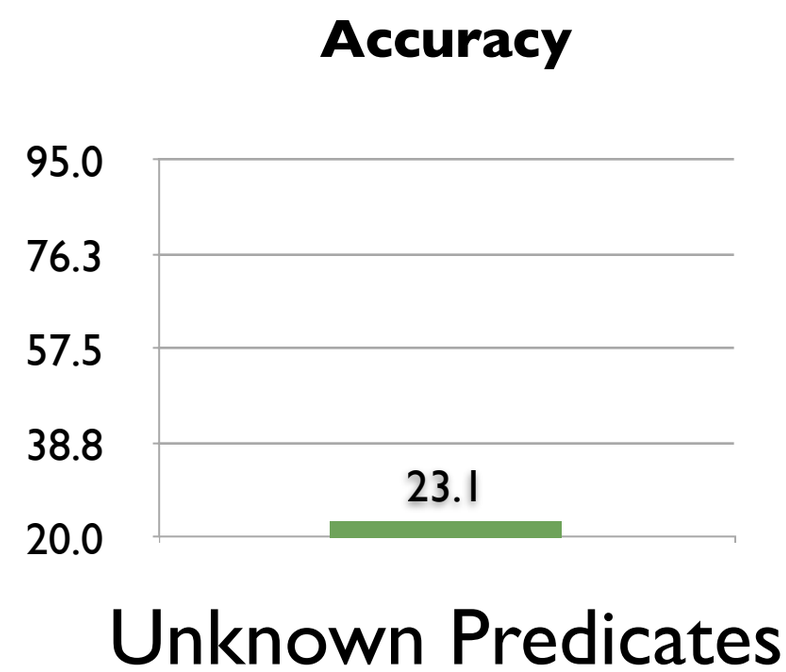
New Data



SEMAFOR: Frame Identification

Frame Identification





SEMAFOR: Handling Unknown Predicates

Knowledge of only 9,263 predicates in supervised data

SEMAFOR: Handling Unknown Predicates

Knowledge of only 9,263 predicates in supervised data

However, English has lot more potential predicates
(~65,000 in newswire English)

SEMAFOR: Handling Unknown Predicates

Knowledge of only 9,263 predicates in supervised data

However, English has lot more potential predicates
(~65,000 in newswire English)

Lexicon expansion using
graph-based semi-supervised learning

Research at Google

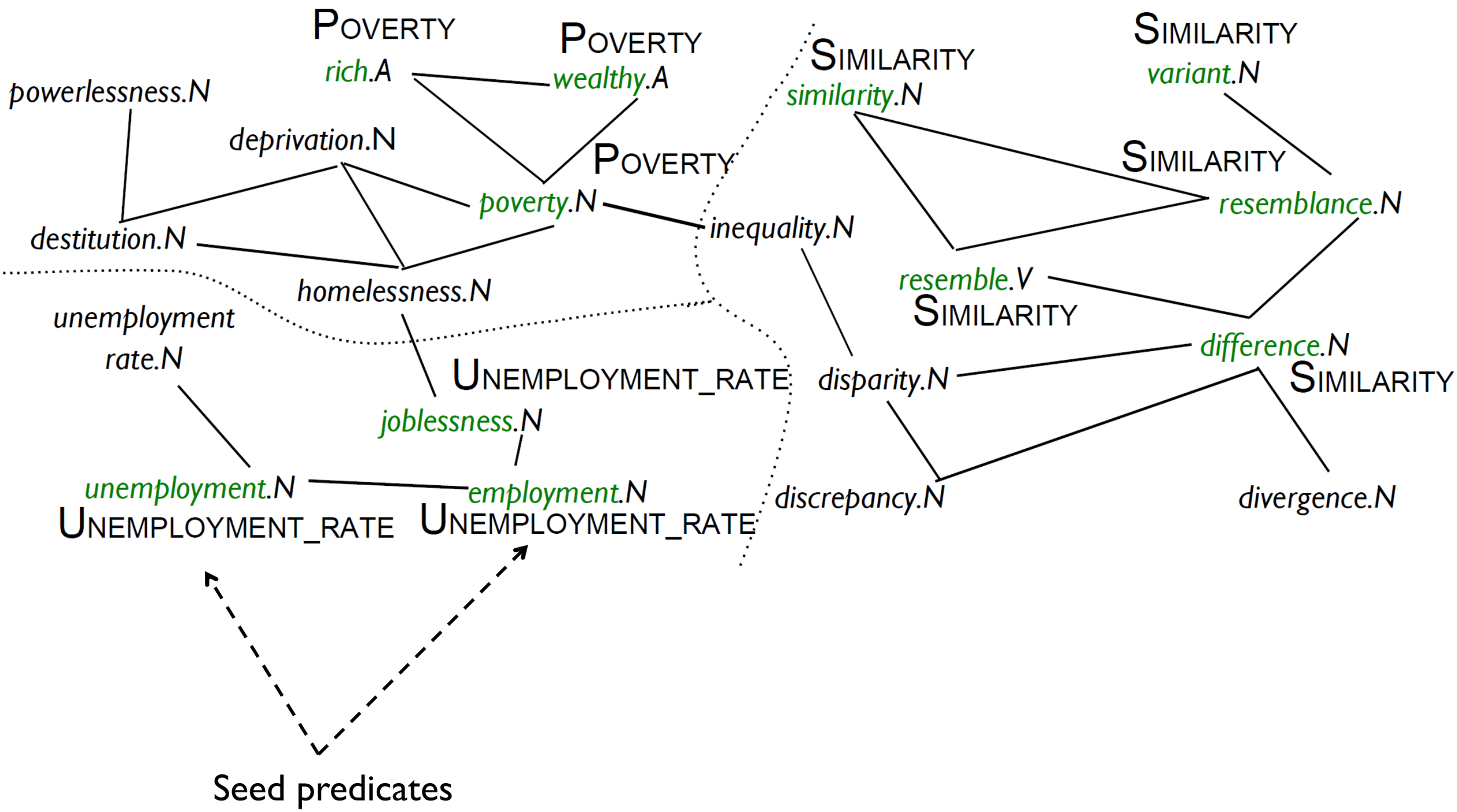
How can label propagation help?

- Build a graph over potential predicates as vertices
  - compute similarity matrix using co-occurrence statistics
- Label distribution at each vertex

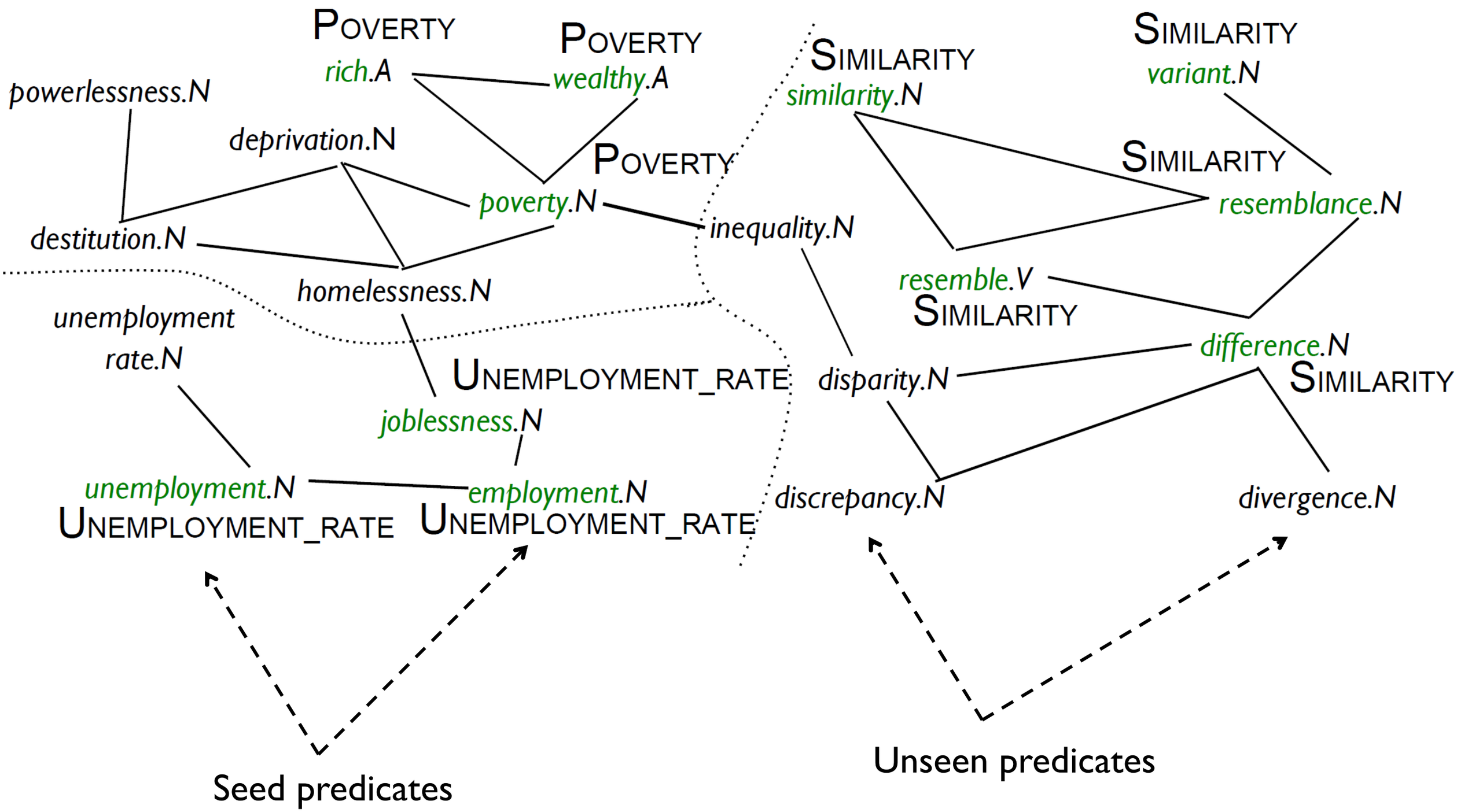
  $\approx$ distribution over frames that the predicate can evoke

Research at Google

Example Graph

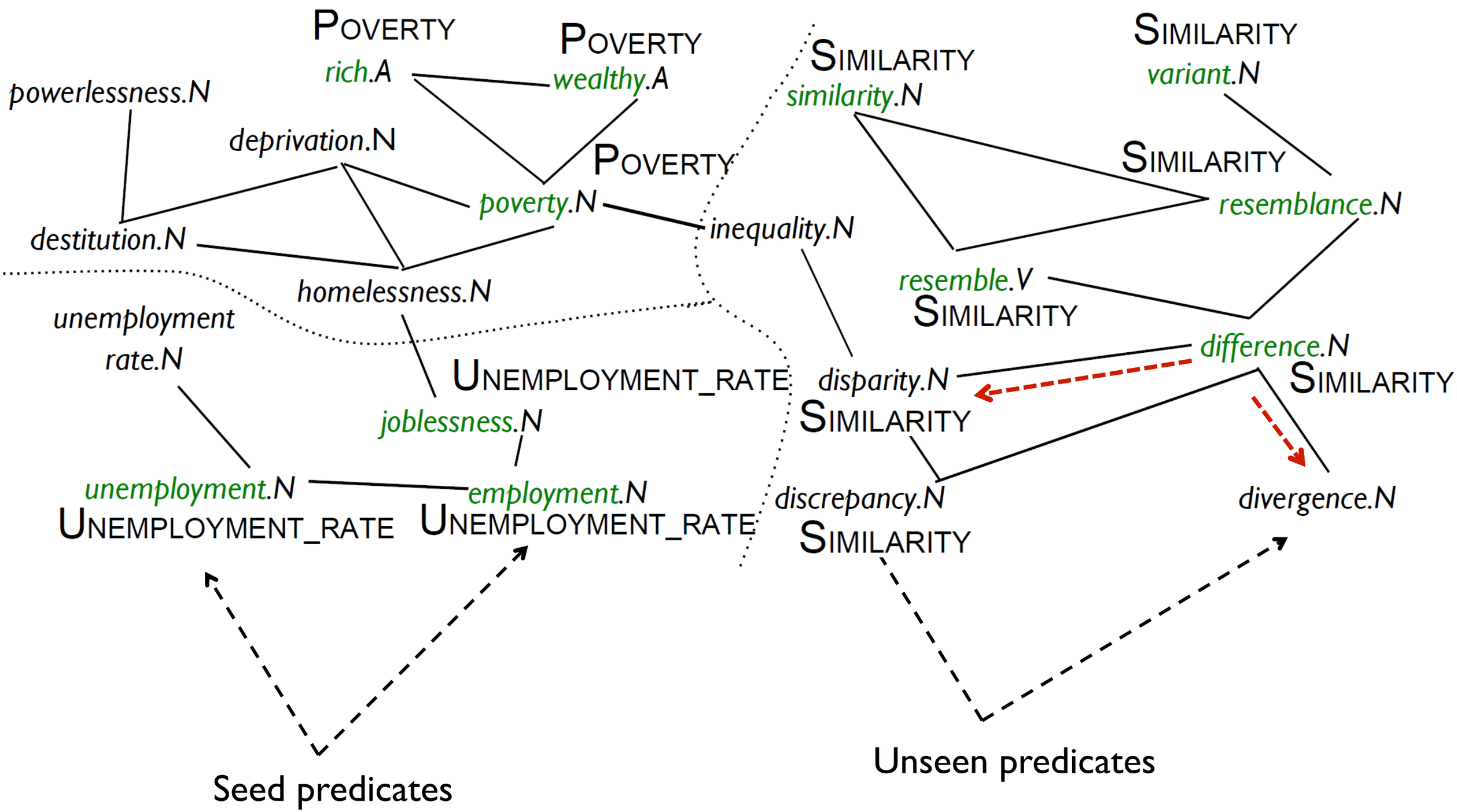


Example Graph



Research at Google

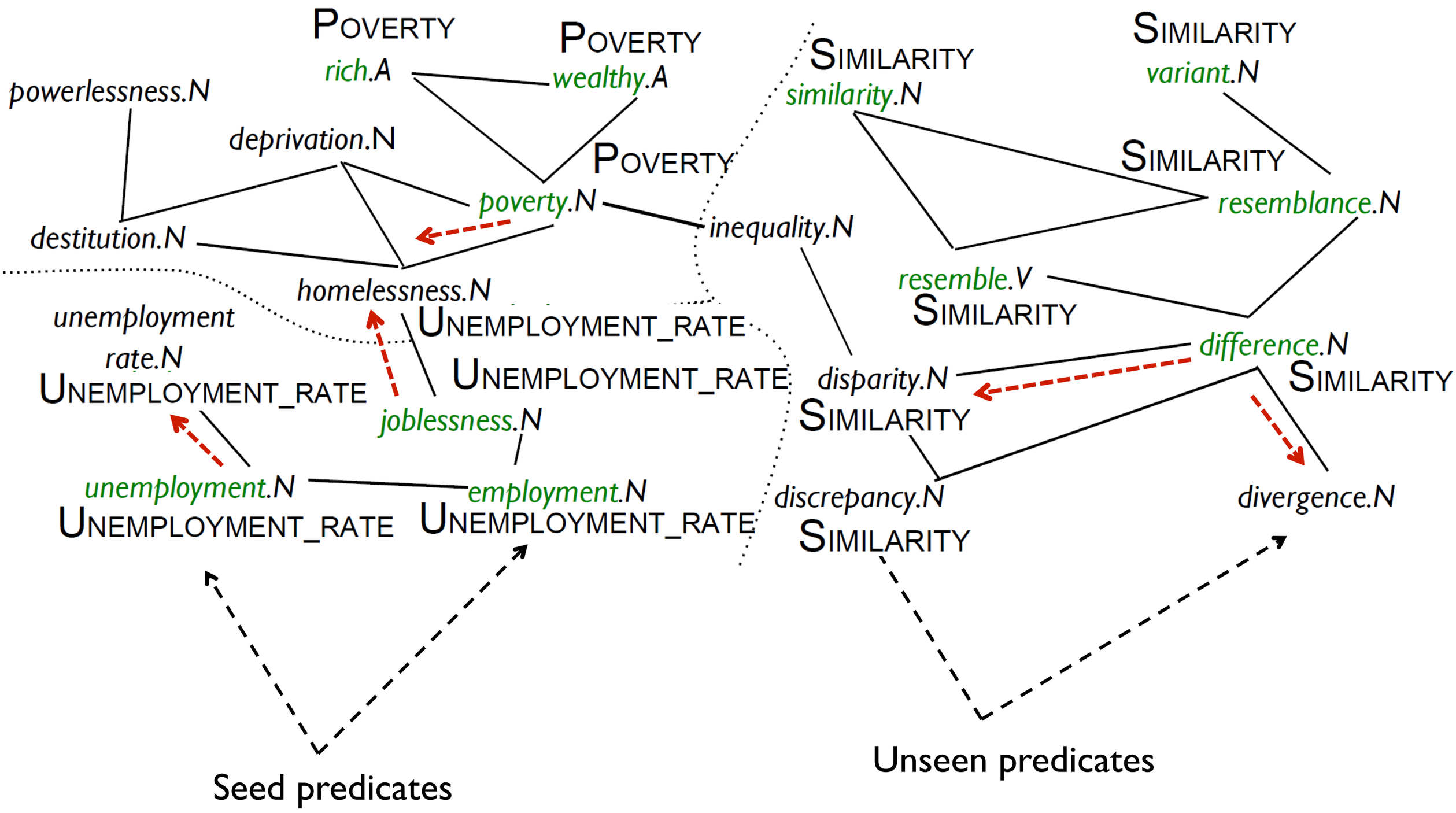
Example Graph



Graph Propagation

Research at Google

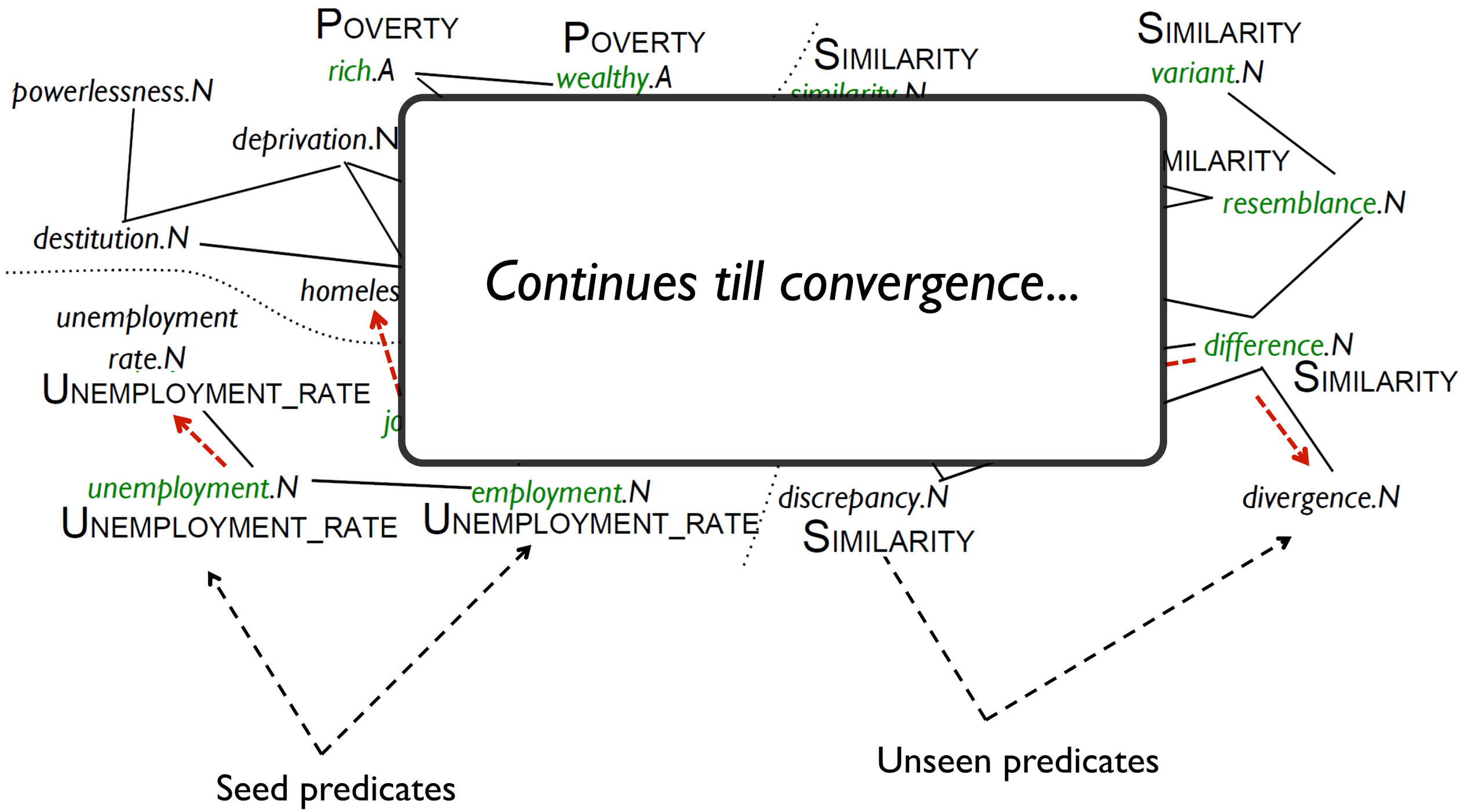
Example Graph



Graph Propagation

Research at Google

Example Graph

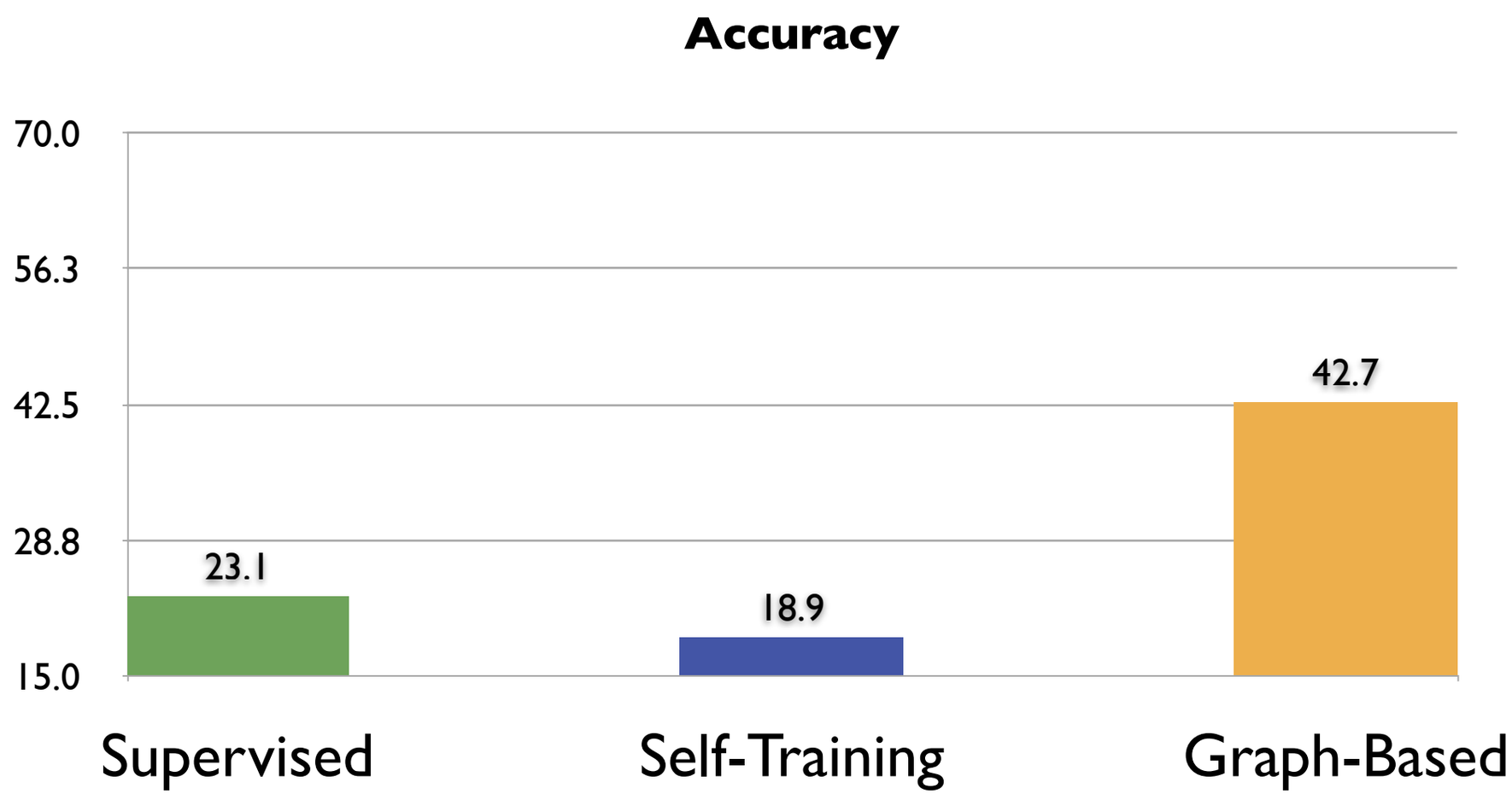


Continues till convergence...

Graph Propagation

SEMAFOR: Unknown Predicates

Frame Identification



Research at Google

SEMAFOR

Das, Chen, Martins, Schneider, Smith (2014)

Frame Identification

Argument Identification

SEMAFOR: Argument Identification

STORE

Bengal 's massive ***stock*** of food was reduced to nothing

SEMAFOR: Argument Identification

STORE

Bengal 's massive ***stock*** of food was reduced to nothing

Possessor

Resource

Descriptor

Use

Supply

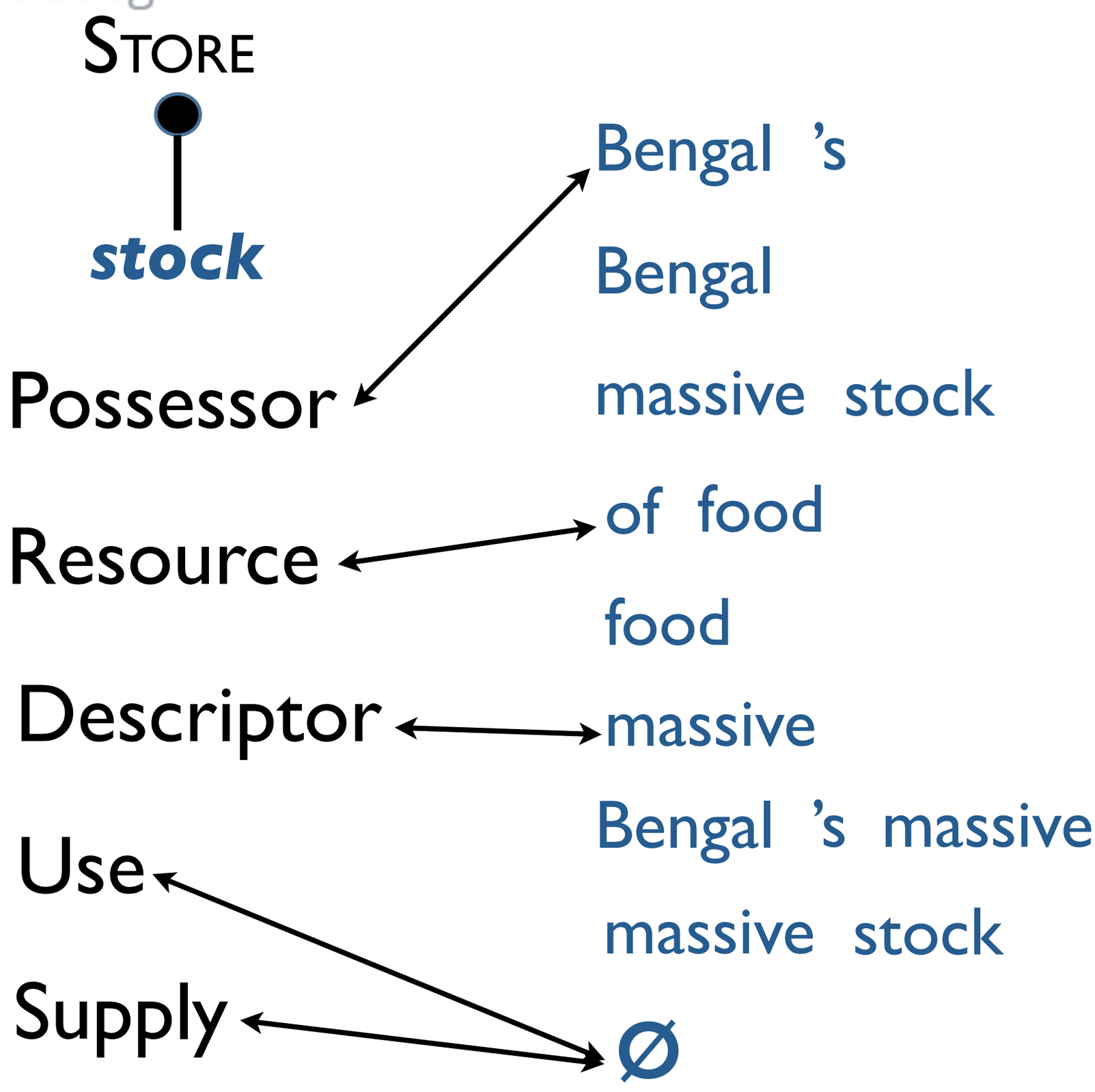
SEMAFOR: Argument Identification

STORE

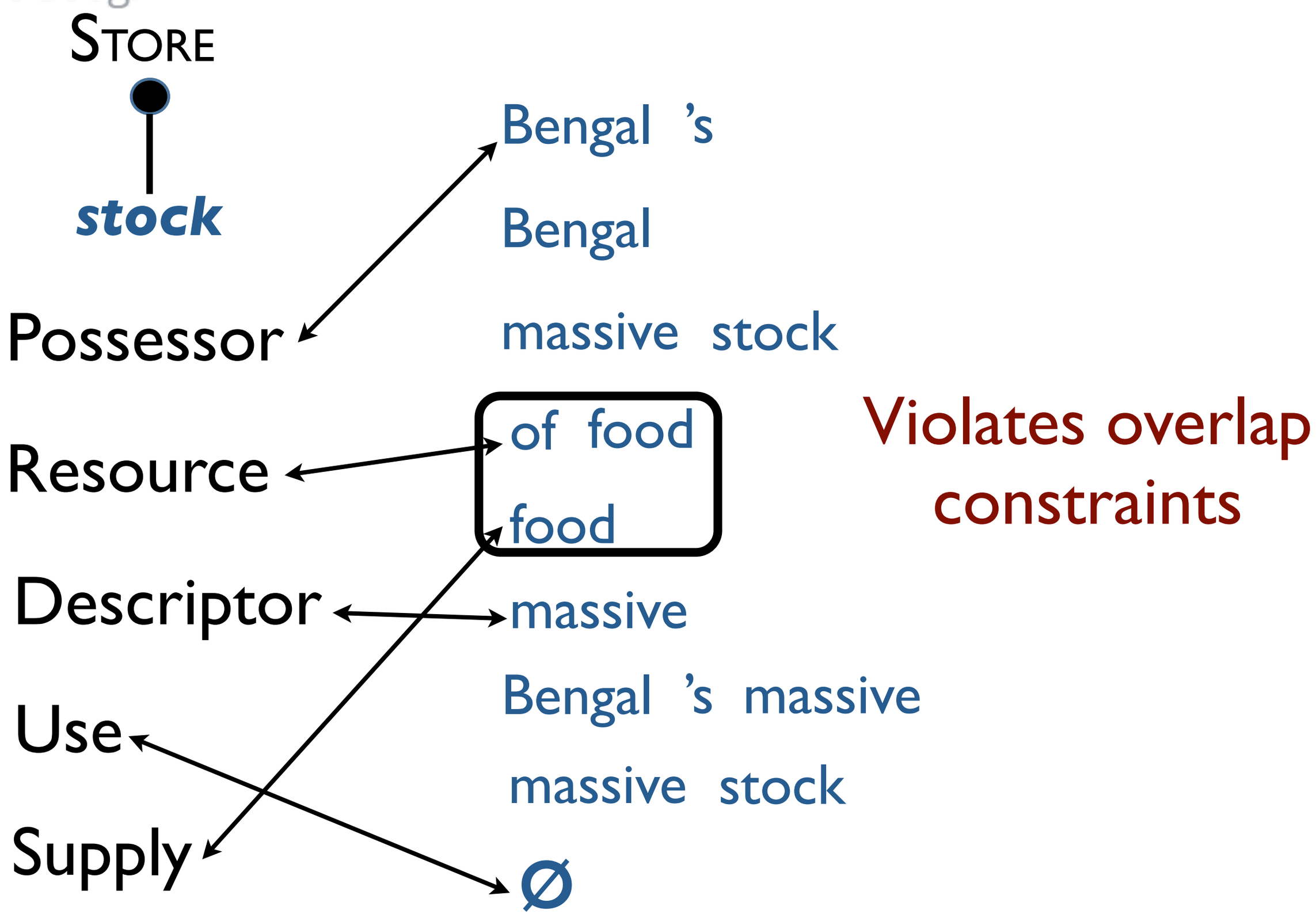
stock

Possessor	Bengal 's
Resource	Bengal
Descriptor	massive stock
Use	of food
Supply	food
	massive
	Bengal 's massive
	massive stock
	Ø

SEMAFOR: Argument Identification

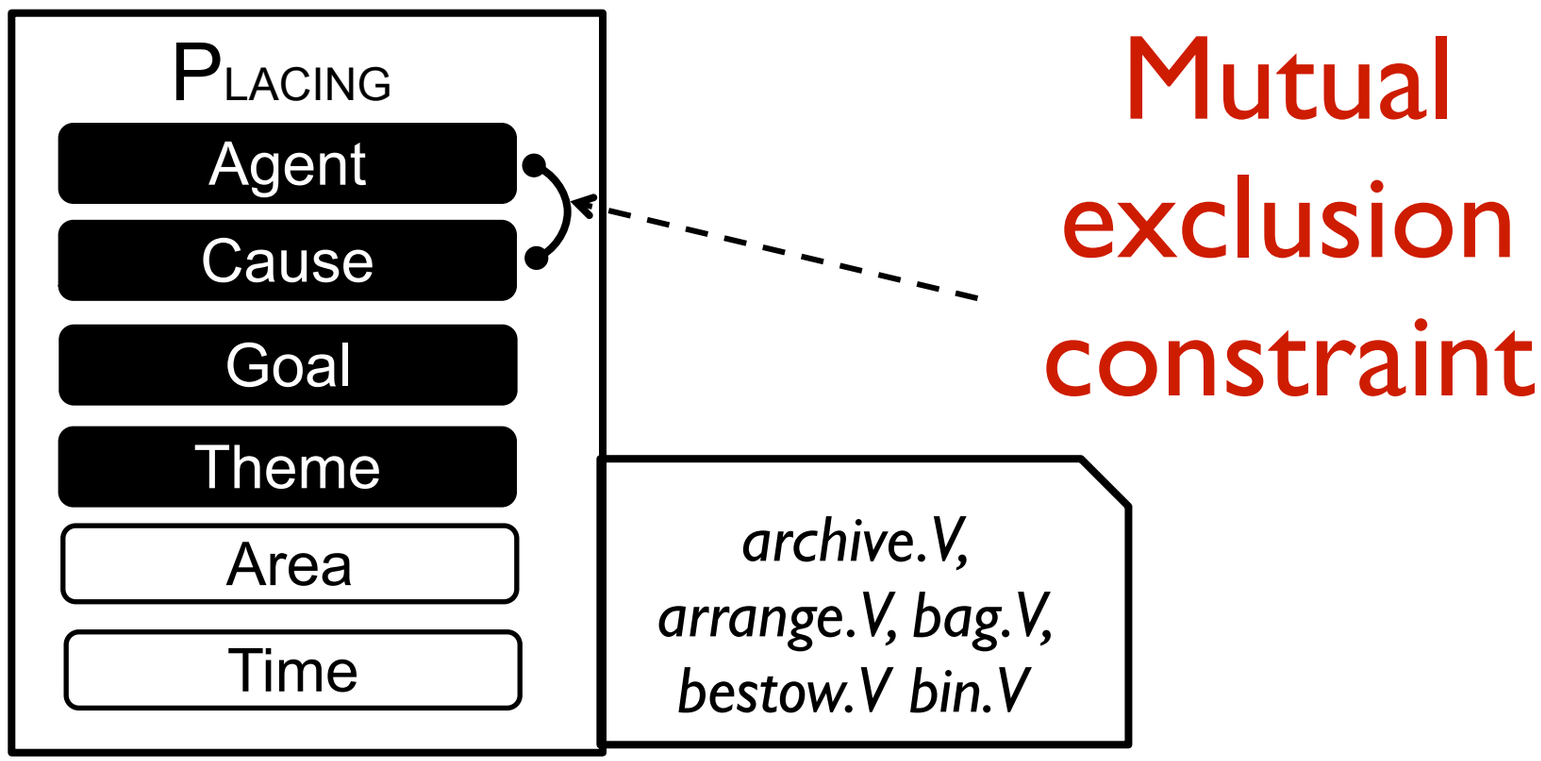


SEMAFOR: Argument Identification



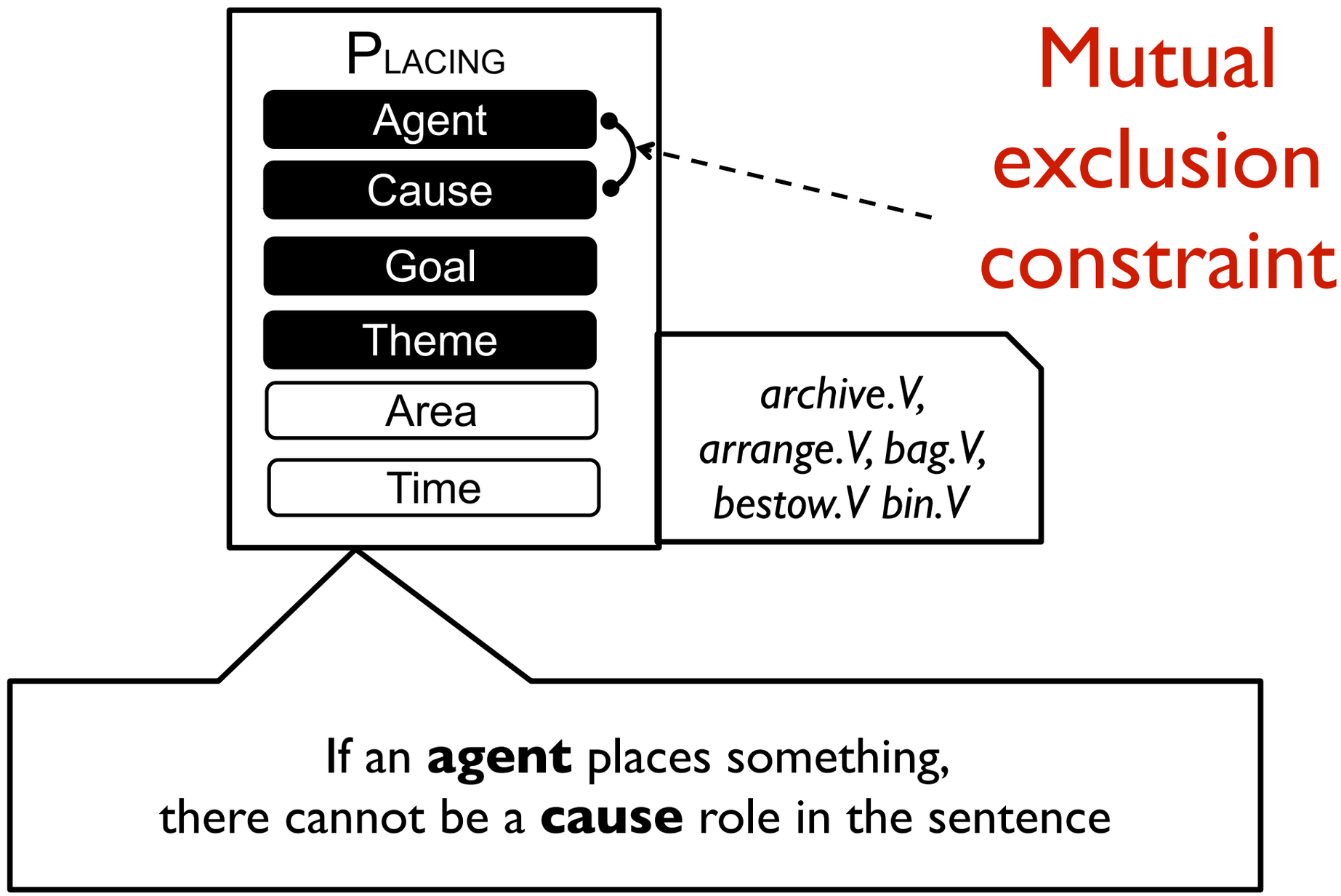
SEMAFOR: Argument Identification

Other types of *structural constraints*



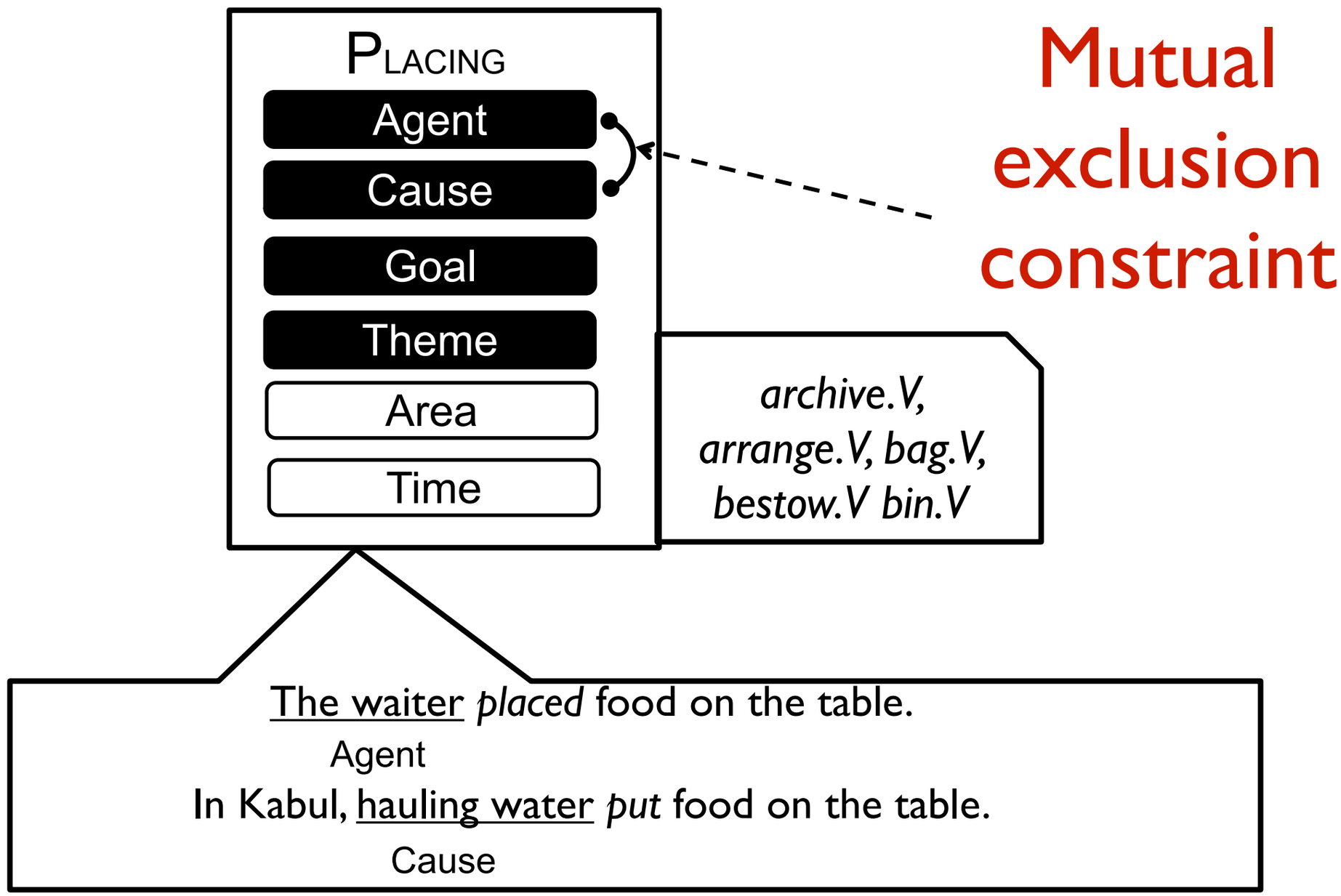
SEMAFOR: Argument Identification

Other types of *structural constraints*



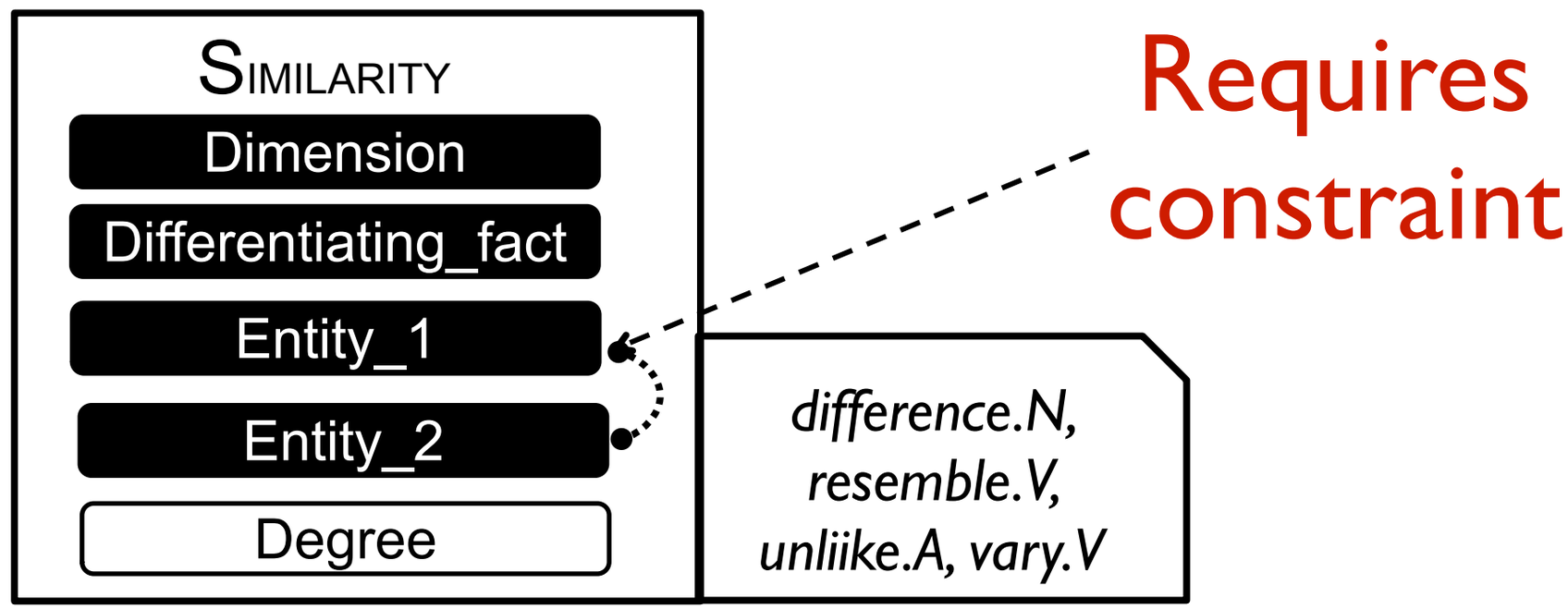
SEMAFOR: Argument Identification

Other types of *structural constraints*



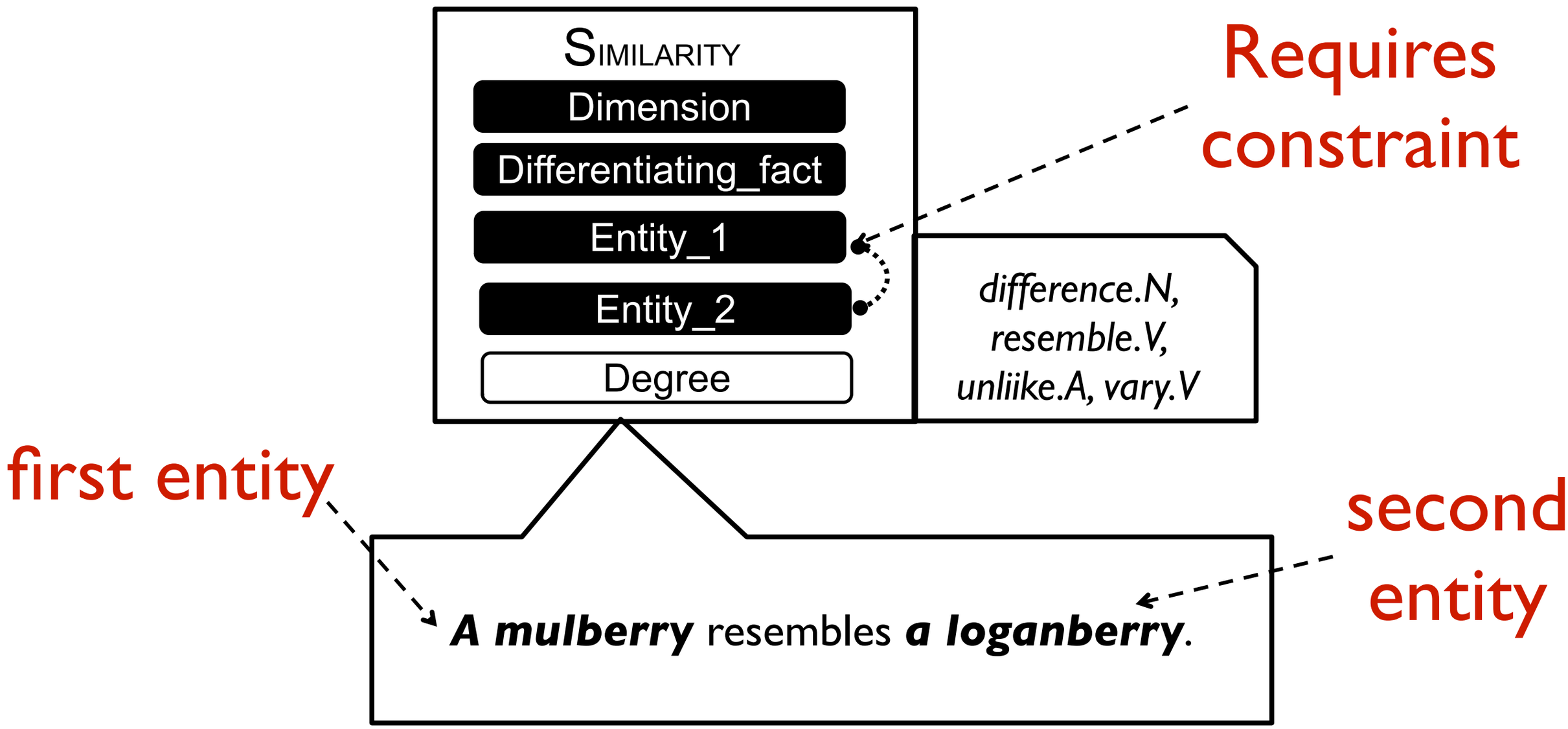
SEMAFOR: Argument Identification

Other types of *structural constraints*



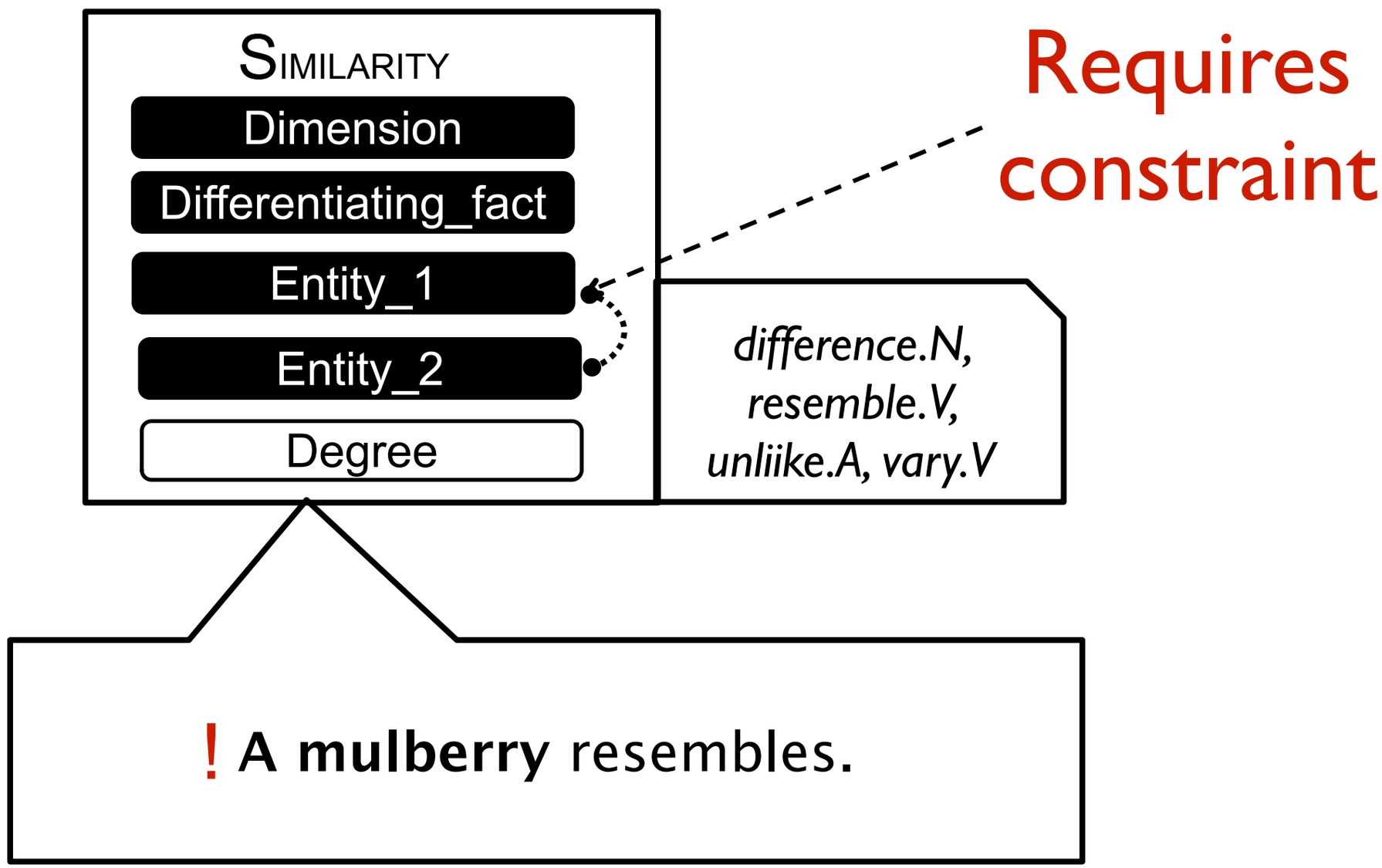
SEMAFOR: Argument Identification

Other types of *structural constraints*

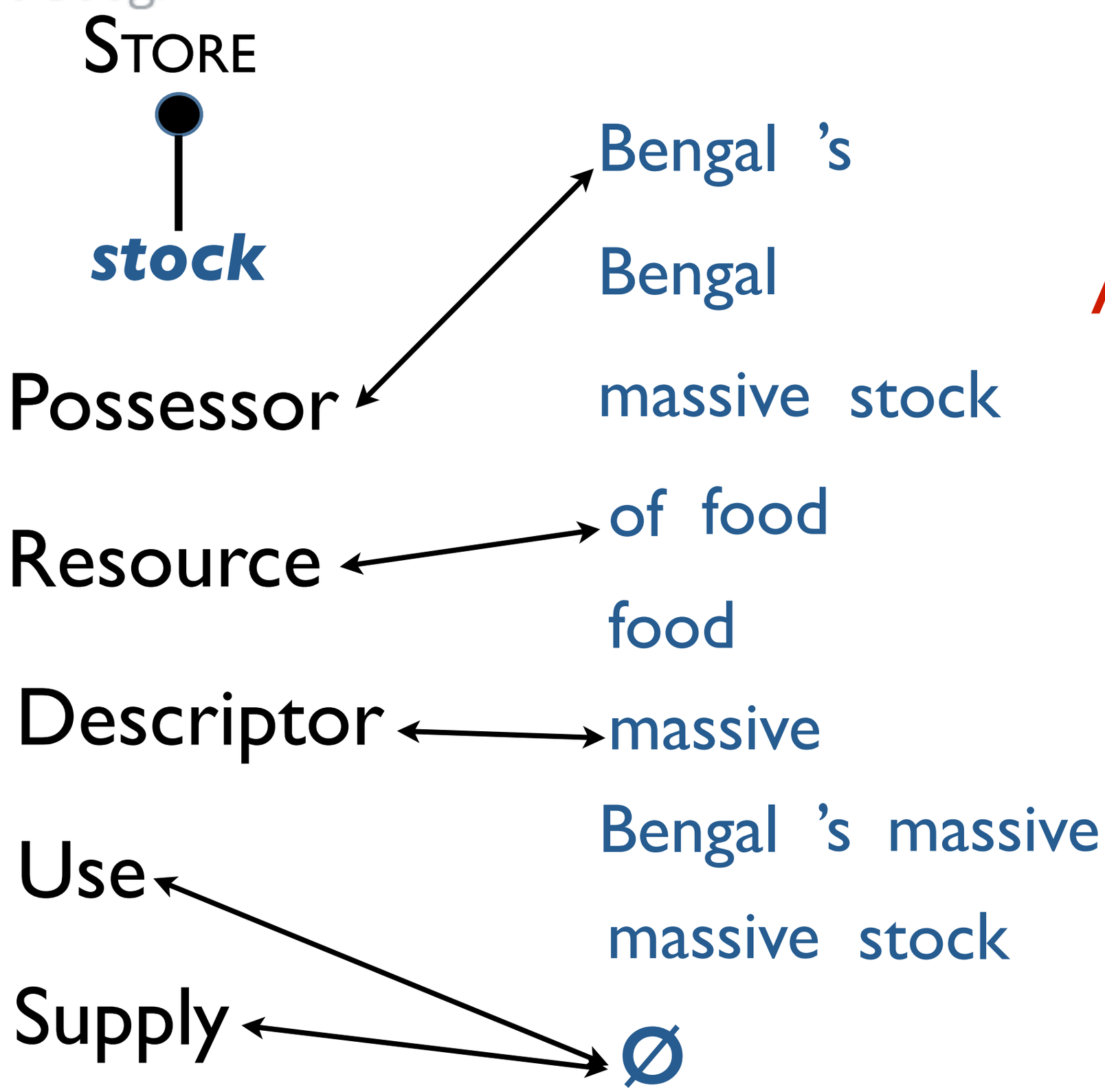


SEMAFOR: Argument Identification

Other types of *structural constraints*

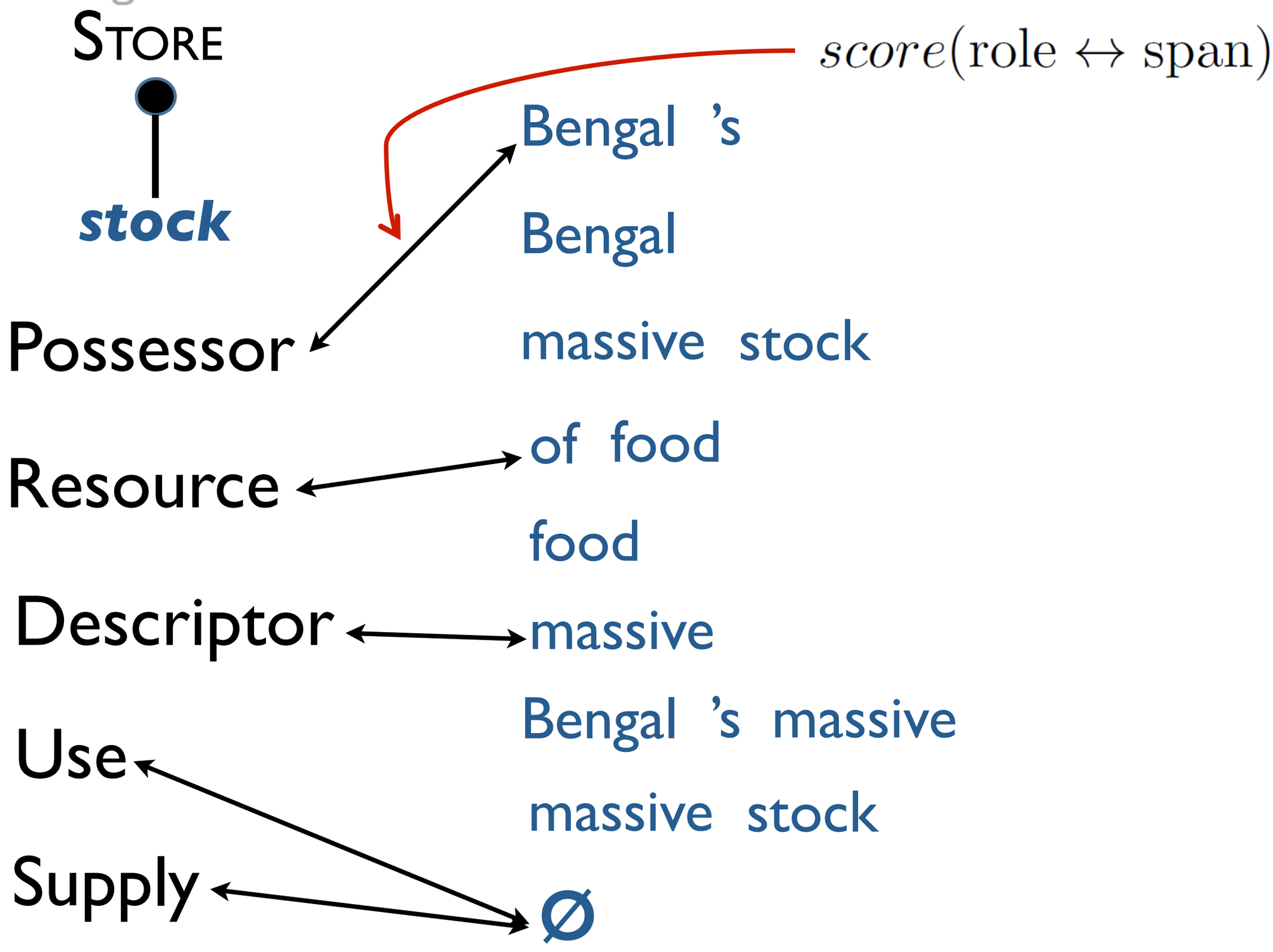


SEMAFOR: Argument Identification

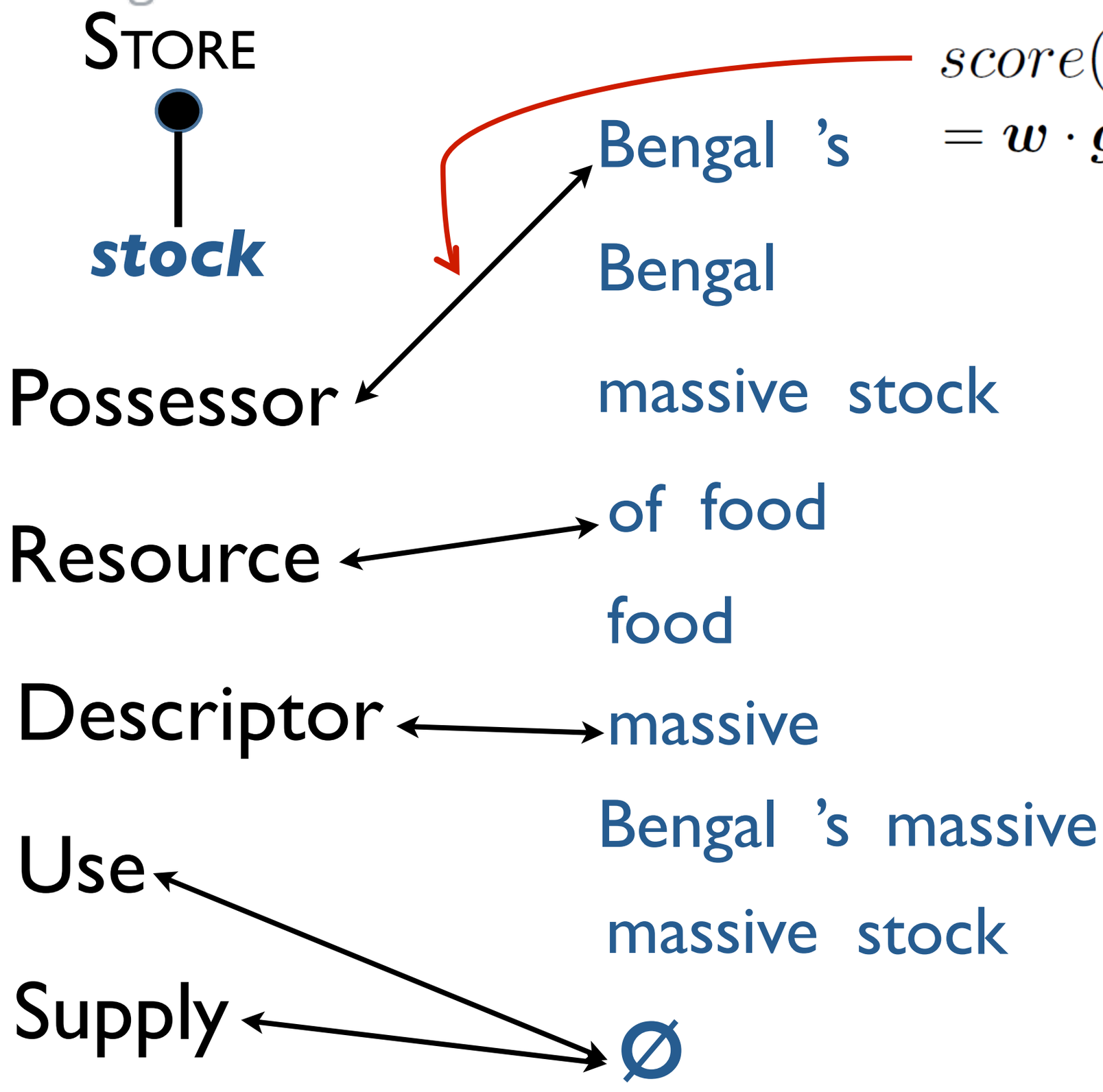


A constrained optimization problem

SEMAFOR: Argument Identification



SEMAFOR: Argument Identification



$$score(\text{role} \leftrightarrow \text{span}) = \boldsymbol{w} \cdot \boldsymbol{g}(\text{role}, \text{span}, \text{frame})$$

SEMAFOR: Argument Identification

A constrained optimization problem

$$z_{\text{role}\leftrightarrow\text{span}}$$

SEMAFOR: Argument Identification

A constrained optimization problem

$$z_{\text{role}\leftrightarrow\text{span}}$$

a binary variable for each role, span tuple

SEMAFOR: Argument Identification

A constrained optimization problem

$$z_{\text{role}\leftrightarrow\text{span}}$$

$$\boldsymbol{z} = \langle z_{\text{role}\leftrightarrow\text{span}} \rangle$$

a binary vector for all role, span tuples

SEMAFOR: Argument Identification

A constrained optimization problem

$$z_{\text{role}\leftrightarrow\text{span}}$$

$$\boldsymbol{z} = \langle z_{\text{role}\leftrightarrow\text{span}} \rangle$$

$$\text{maximize} \sum_{\text{roles,spans}} z_{\text{role}\leftrightarrow\text{span}} \cdot score(\text{role} \leftrightarrow \text{span})$$

$$\text{w.r.t.} \boldsymbol{z}$$

SEMAFOR: Argument Identification

A constrained optimization problem

$$z_{\text{role}\leftrightarrow\text{span}}$$

$$\boldsymbol{z} = \langle z_{\text{role}\leftrightarrow\text{span}} \rangle$$

$$\text{maximize} \sum_{\text{roles,spans}} z_{\text{role}\leftrightarrow\text{span}} \cdot score(\text{role} \leftrightarrow \text{span})$$

$$\text{w.r.t.} \boldsymbol{z}$$

$$\text{s.t } \forall \text{ roles}, \sum_{\text{spans}} z_{\text{role}\leftrightarrow\text{span}} = 1$$

SEMAFOR: Argument Identification

A constrained optimization problem

$$z_{\text{role}\leftrightarrow\text{span}}$$

$$\boldsymbol{z} = \langle z_{\text{role}\leftrightarrow\text{span}} \rangle$$

$$\text{maximize} \sum_{\text{roles,spans}} z_{\text{role}\leftrightarrow\text{span}} \cdot score(\text{role} \leftrightarrow \text{span})$$

$$\text{w.r.t.} \boldsymbol{z}$$

$$\text{s.t } \forall \text{ roles, } \sum_{\text{spans}} z_{\text{role}\leftrightarrow\text{span}} = 1$$

Uniqueness

SEMAFOR: Argument Identification

A constrained optimization problem

$$z_{\text{role}\leftrightarrow\text{span}}$$

$$\boldsymbol{z} = \langle z_{\text{role}\leftrightarrow\text{span}} \rangle$$

$$\text{maximize} \sum_{\text{roles,spans}} z_{\text{role}\leftrightarrow\text{span}} \cdot \mathit{score}(\text{role} \leftrightarrow \text{span})$$

$$\text{w.r.t.} \boldsymbol{z}$$

$$\text{s.t } \forall \text{ roles, } \sum_{\text{spans}} z_{\text{role}\leftrightarrow\text{span}} = 1$$

$$\forall \text{ sentence positions, } \sum_{\substack{\text{span}\\ \text{covers}\\ \text{position}}} \sum_{\text{roles}} z_{\text{role}\leftrightarrow\text{span}} \leq 1$$

Prevents overlap

SEMAFOR: Argument Identification

A constrained optimization problem

$$z_{\text{role}\leftrightarrow\text{span}}$$

$$\boldsymbol{z} = \langle z_{\text{role}\leftrightarrow\text{span}} \rangle$$

An integer linear program (ILP)

$$\text{maximize} \sum_{\text{roles,spans}} z_{\text{role}\leftrightarrow\text{span}} \cdot score(\text{role} \leftrightarrow \text{span})$$

$$\text{w.r.t.} \boldsymbol{z}$$

more structural constraints

$$\text{s.t } \forall \text{ roles}, \sum_{\text{spans}} z_{\text{role}\leftrightarrow\text{span}} = 1$$

$$\forall \text{ sentence positions}, \sum_{\substack{\text{span}\\ \text{covers}\\ \text{position}}} \sum_{\text{roles}} z_{\text{role}\leftrightarrow\text{span}} \leq 1$$

SEMAFOR: Argument Identification

A constrained optimization problem

$$z_{\text{role}\leftrightarrow\text{span}}$$

$$\boldsymbol{z} = \langle z_{\text{role}\leftrightarrow\text{span}} \rangle$$

Punyakanok, Roth and Yih (2008)

An integer linear program (ILP)

$$\text{maximize} \sum_{\text{roles,spans}} z_{\text{role}\leftrightarrow\text{span}} \cdot score(\text{role} \leftrightarrow \text{span})$$

more structural constraints

$$\text{w.r.t.} \boldsymbol{z}$$

$$\text{s.t } \forall \text{ roles}, \sum_{\text{spans}} z_{\text{role}\leftrightarrow\text{span}} = 1$$

$$\forall \text{ sentence positions}, \sum_{\substack{\text{span}\\ \text{covers}\\ \text{position}}} \sum_{\text{roles}} z_{\text{role}\leftrightarrow\text{span}} \leq 1$$

SEMAFOR: Argument Identification

A constrained optimization problem

An integer linear program (ILP)

$$\text{maximize} \sum_{\text{roles,spans}} z_{\text{role}\leftrightarrow\text{span}} \cdot score(\text{role} \leftrightarrow \text{span})$$

more structural constraints

$$\text{w.r.t.} z$$

$$\text{s.t } \forall \text{ roles, } \sum_{\text{spans}} z_{\text{role}\leftrightarrow\text{span}} = 1$$

$$\forall \text{ sentence positions, } \sum_{\substack{\text{span}\\ \text{covers}\\ \text{position}}} \sum_{\text{roles}} z_{\text{role}\leftrightarrow\text{span}} \leq 1$$

Often, very slow solutions

SEMAFOR: Argument Identification

A constrained optimization problem

An integer linear program (ILP)

$$\text{maximize} \sum_{\text{roles,spans}} z_{\text{role}\leftrightarrow\text{span}} \cdot score(\text{role} \leftrightarrow \text{span})$$

more structural constraints

$$\text{w.r.t.} z$$

$$\text{s.t } \forall \text{ roles, } \sum_{\text{spans}} z_{\text{role}\leftrightarrow\text{span}} = 1$$

$$\forall \text{ sentence positions, } \sum_{\substack{\text{span}\\ \text{covers}\\ \text{position}}} \sum_{\text{roles}} z_{\text{role}\leftrightarrow\text{span}} \leq 1$$

Fast ILP solvers proprietary

SEMAFOR: Argument Identification

A constrained optimization problem

An integer linear program (ILP)

$$\text{maximize} \sum_{\text{roles,spans}} z_{\text{role}\leftrightarrow\text{span}} \cdot score(\text{role} \leftrightarrow \text{span})$$

more structural constraints

$$\text{w.r.t.} z$$

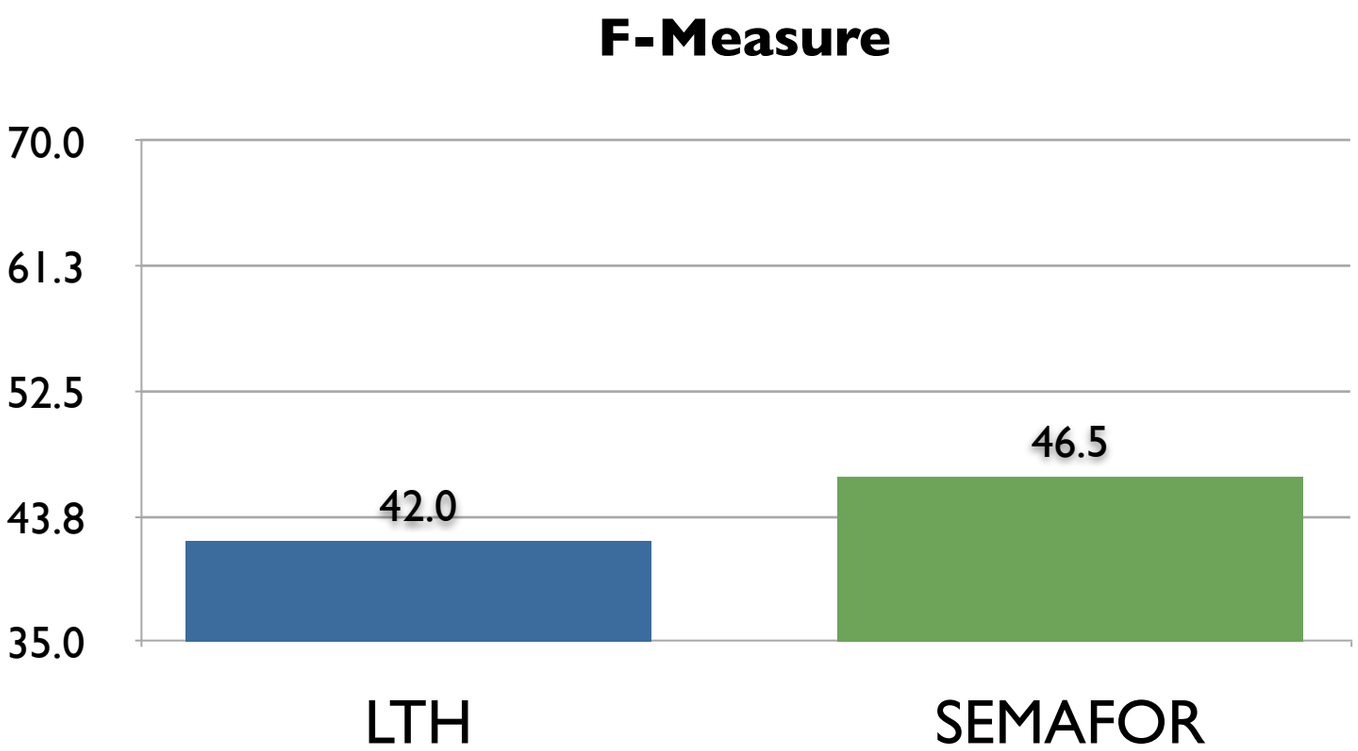
$$\text{s.t } \forall \text{ roles, } \sum_{\text{spans}} z_{\text{role}\leftrightarrow\text{span}} = 1$$

$$\forall \text{ sentence positions, } \sum_{\substack{\text{span}\\ \text{covers}\\ \text{position}}} \sum_{\text{roles}} z_{\text{role}\leftrightarrow\text{span}} \leq 1$$

Dual Decomposition

Final Results

Benchmark



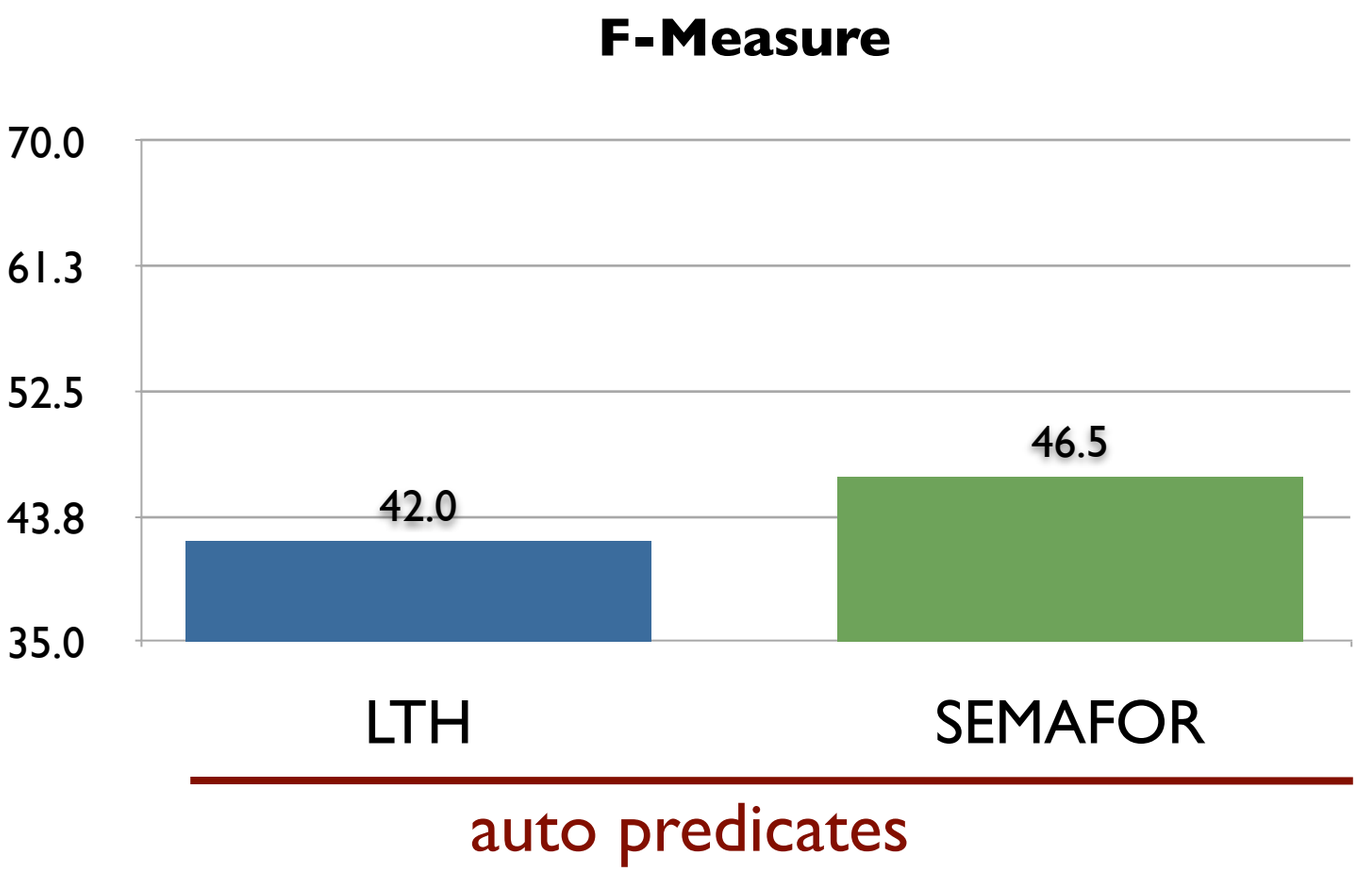
New Data

Research at Google

SEMAFOR: Frame-Semantic Parsing

Final Results

Benchmark



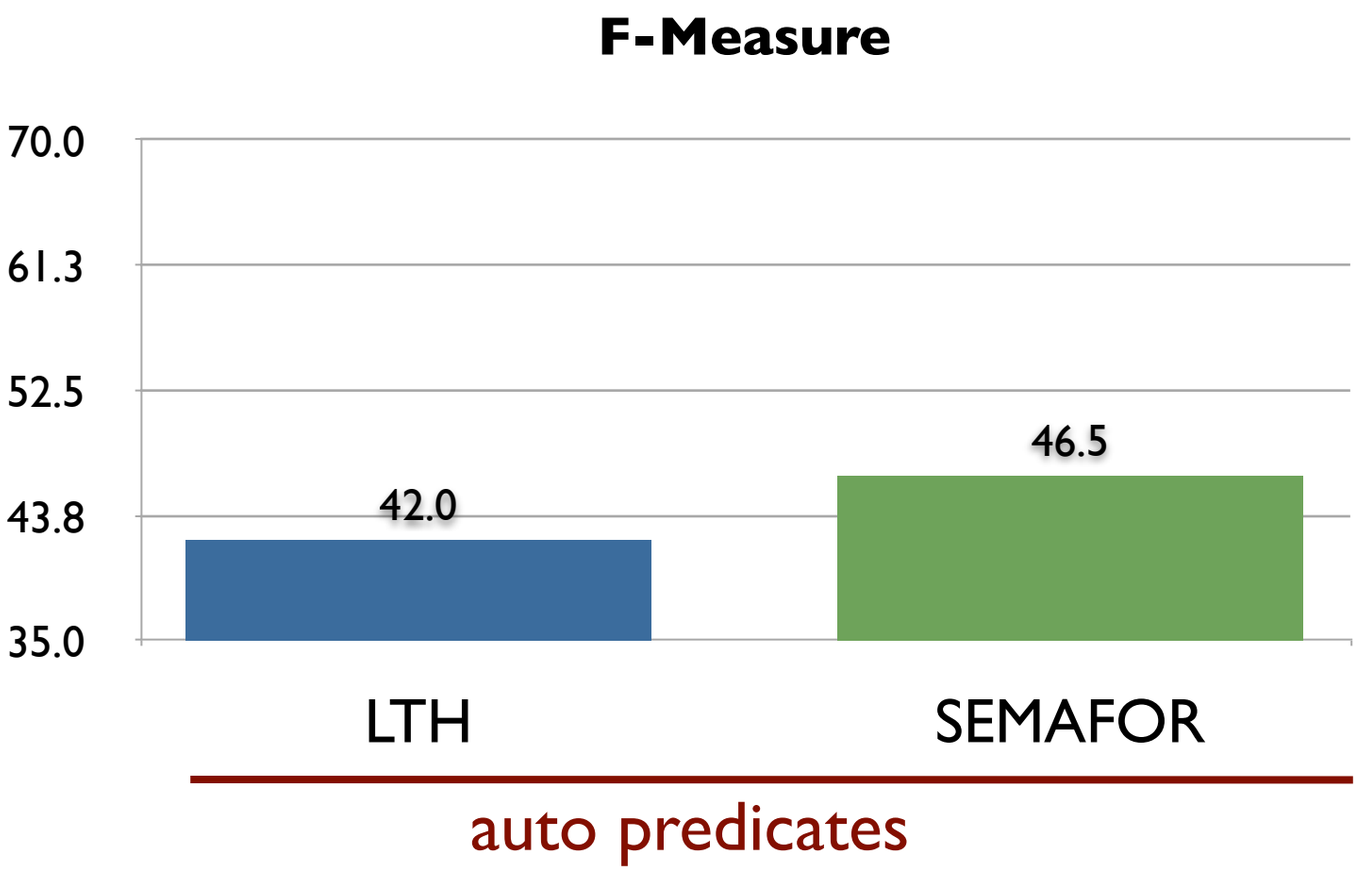
New Data

Research at Google

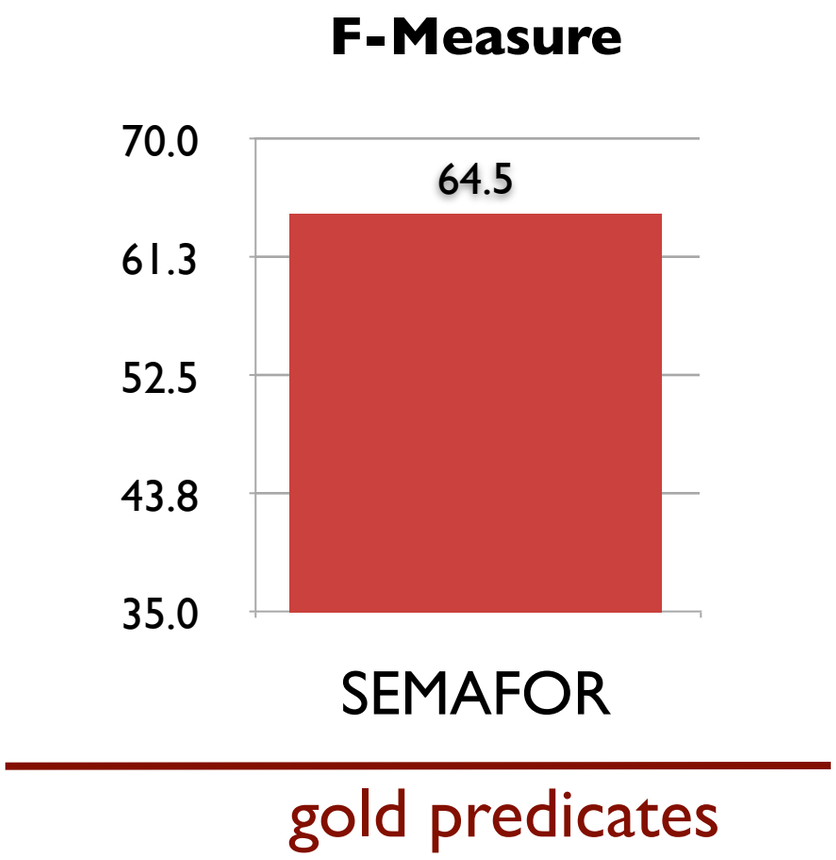
SEMAFOR: Frame-Semantic Parsing

Final Results

Benchmark



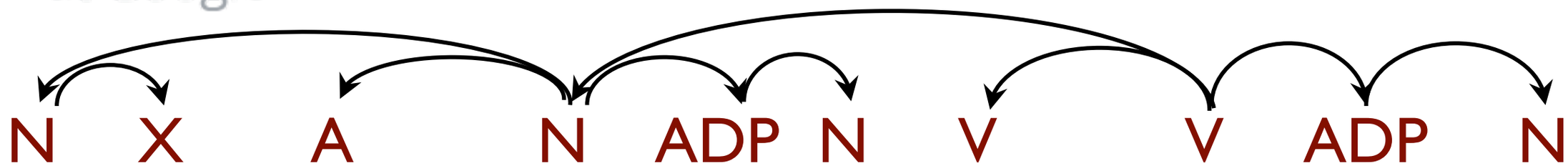
New Data



Research at Google

Hermann, Das, Weston and Ganchev
ACL 2014

Frame Identification with Embeddings

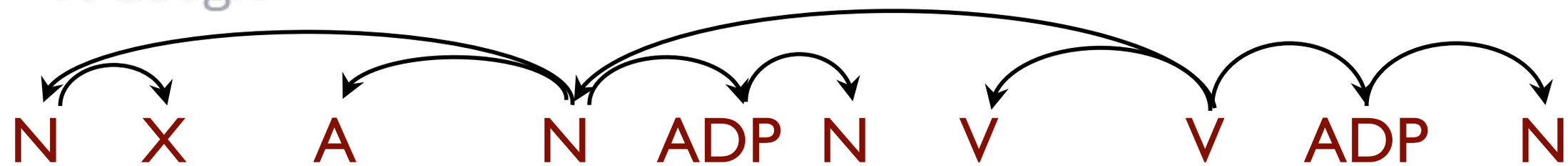


Bengal 's massive ***stock*** of food was reduced to nothing

$$\text{best frame} = \underset{\text{frame} \in \text{all frames}}{\arg\max} \; score(\text{frame}, \text{predicate}, \text{sentence})$$

$$= \log p(\text{frame} \mid \text{predicate}, \text{sentence})$$

Frame Identification with Embeddings

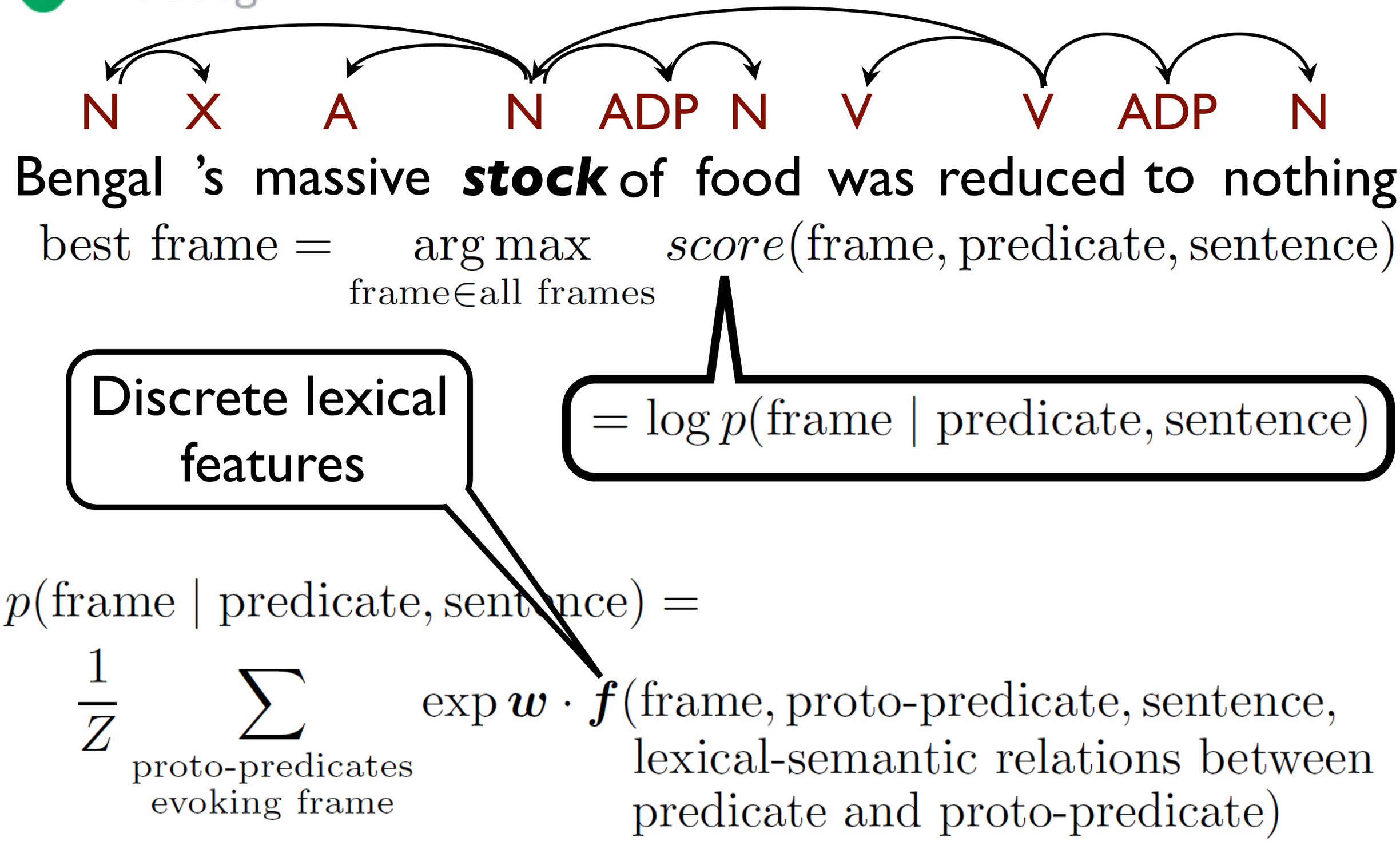


$$\text{best frame} = \underset{\text{frame} \in \text{all frames}}{\arg\max} \; score(\text{frame}, \text{predicate}, \text{sentence})$$

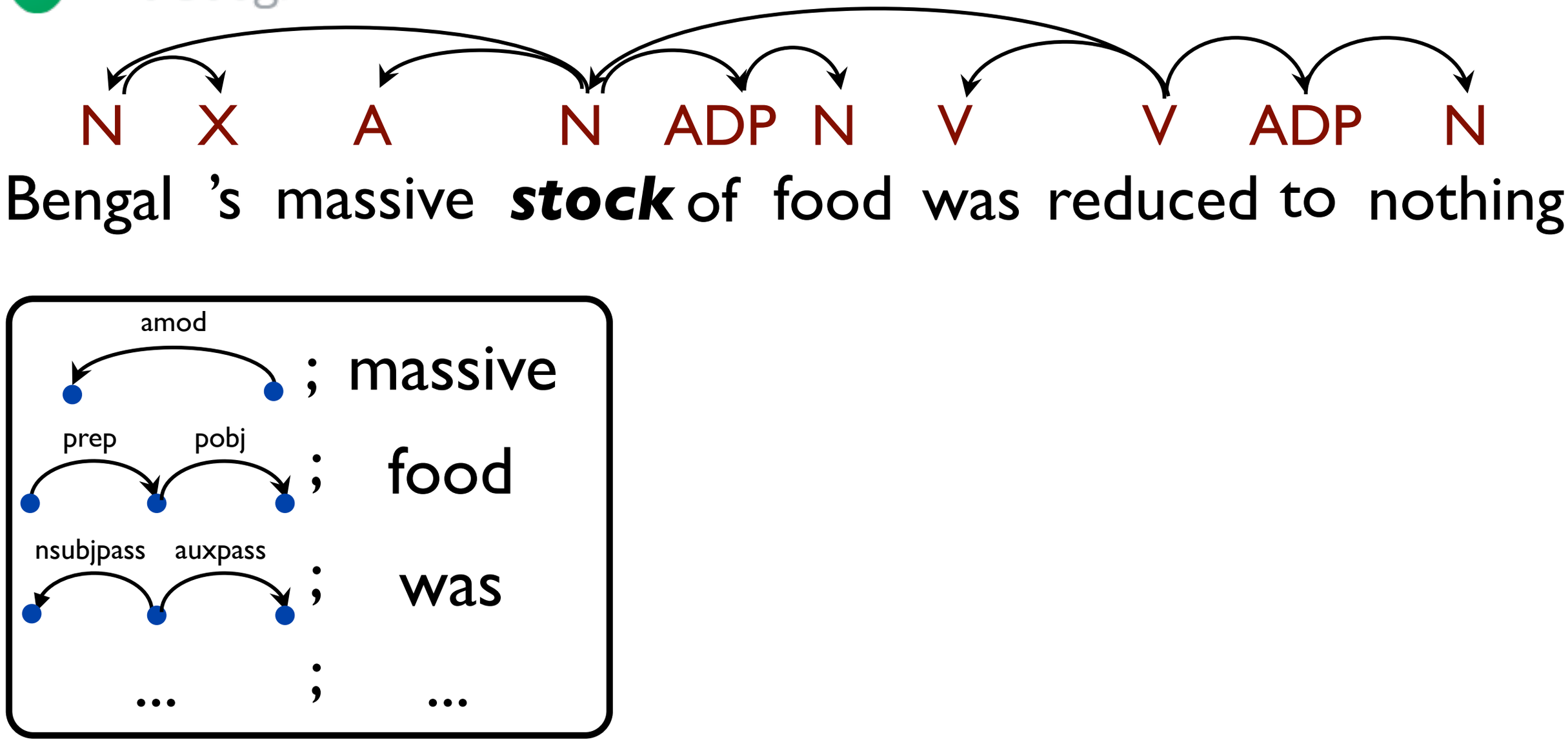
$$= \log p(\text{frame} \mid \text{predicate}, \text{sentence})$$

$$p(\text{frame} \mid \text{predicate}, \text{sentence}) = \frac{1}{Z} \sum_{\substack{\text{proto-predicates} \\ \text{evoking frame}}} \exp \boldsymbol{w} \cdot \boldsymbol{f}(\text{frame}, \text{proto-predicate}, \text{sentence}, \text{lexical-semantic relations between predicate and proto-predicate})$$

Frame Identification with Embeddings

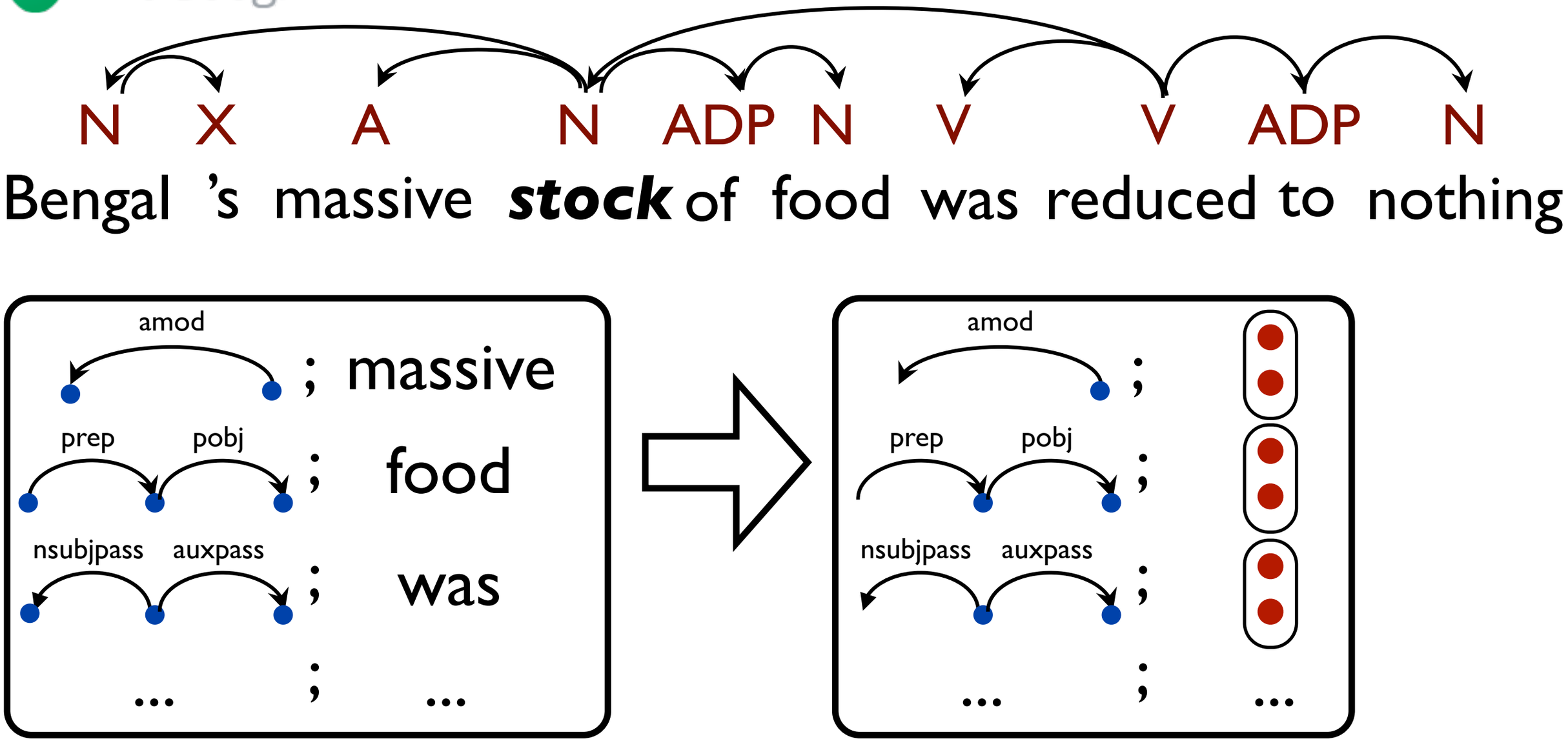


Frame Identification with Embeddings



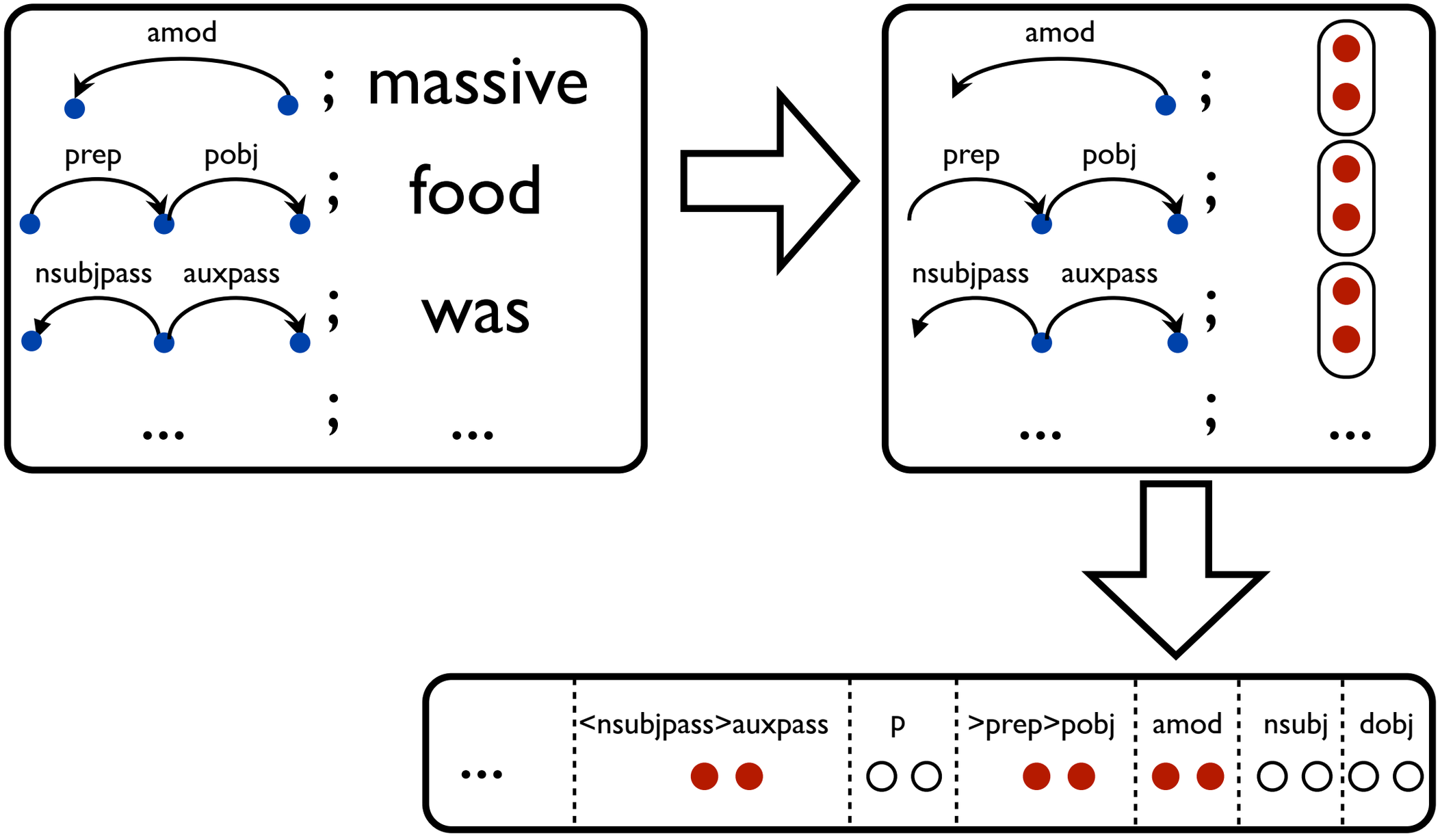
Syntactic context words connected via a set of dependency paths

Frame Identification with Embeddings



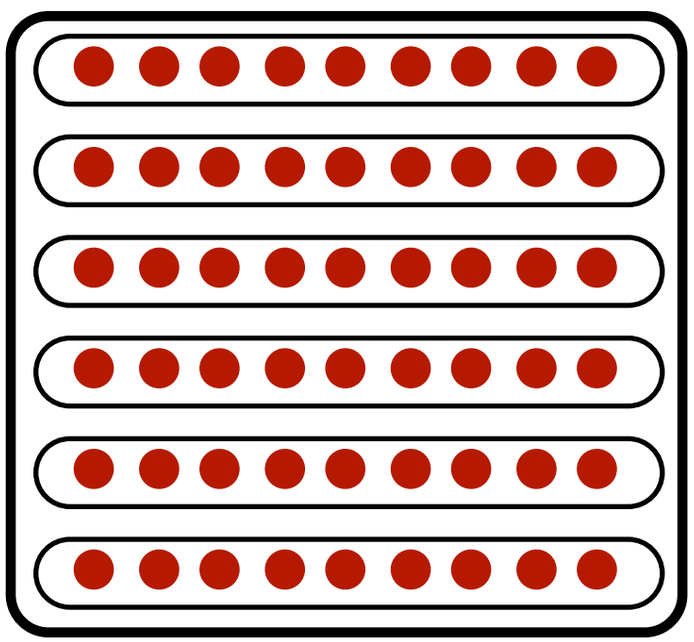
Replace context words with off-the-shelf word embeddings

Frame Identification with Embeddings



Input sparse embedding vector
with context blocks

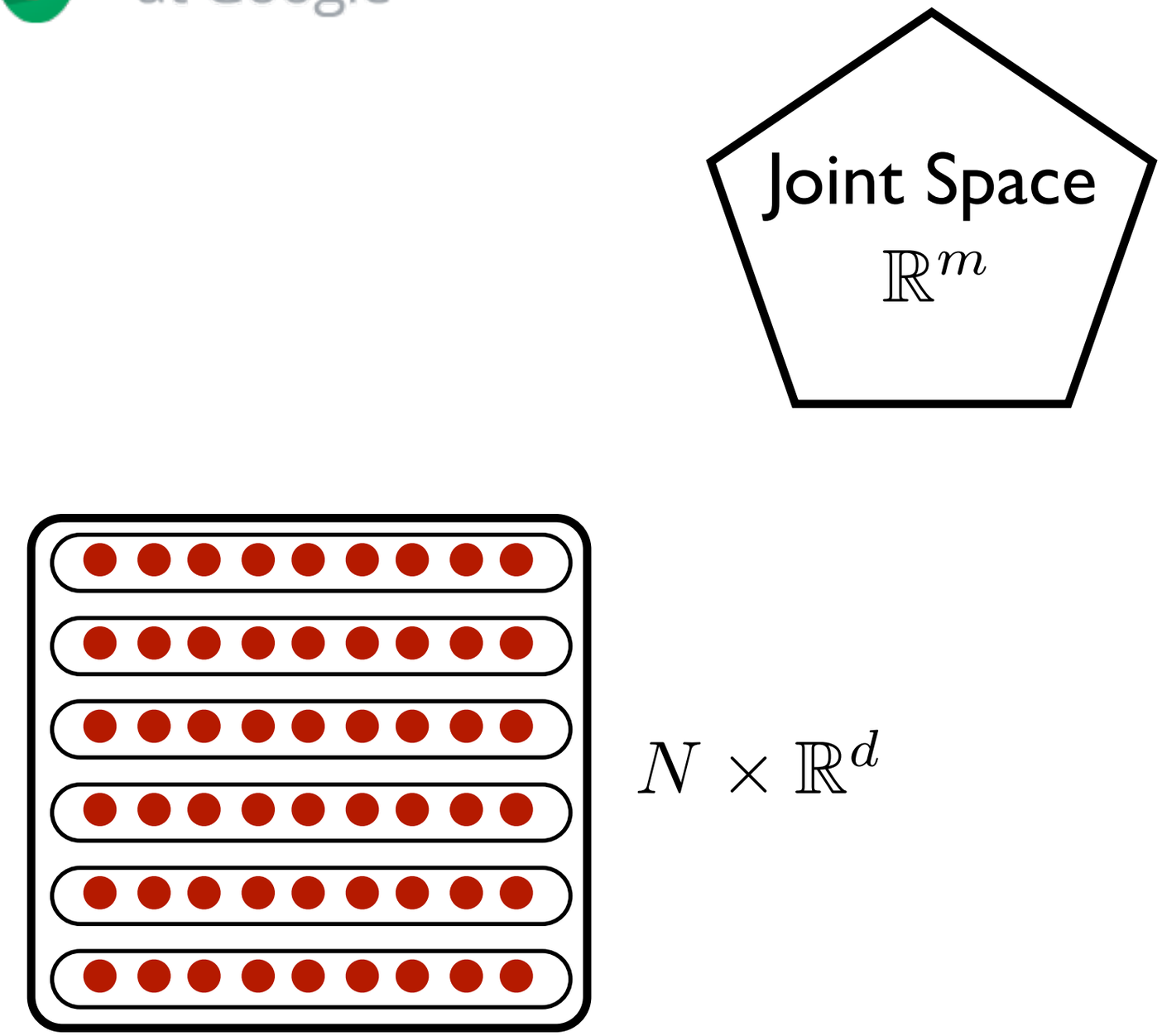
Frame Identification with Embeddings



$N \times \mathbb{R}^d$

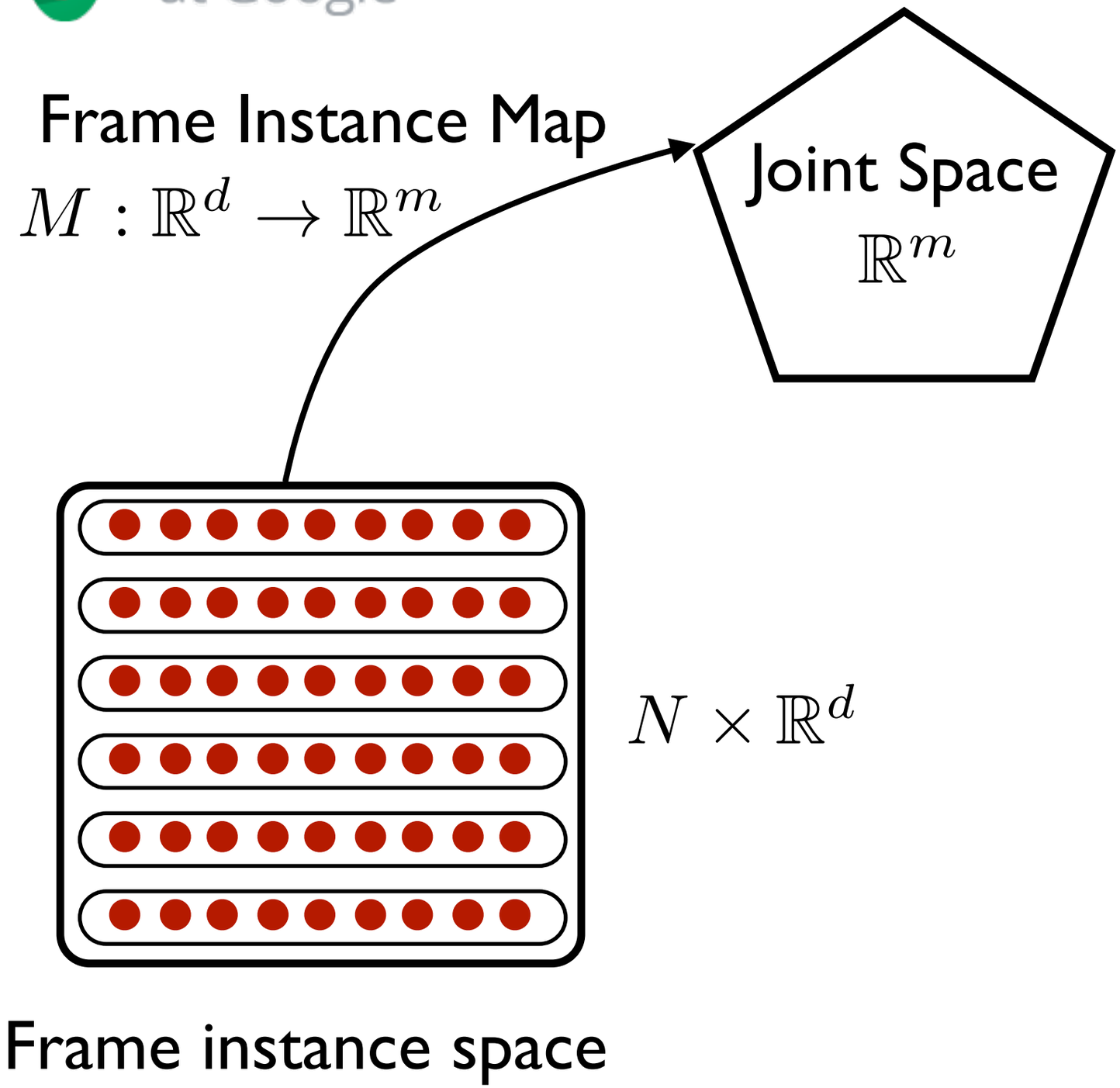
Frame instance space

Frame Identification with Embeddings

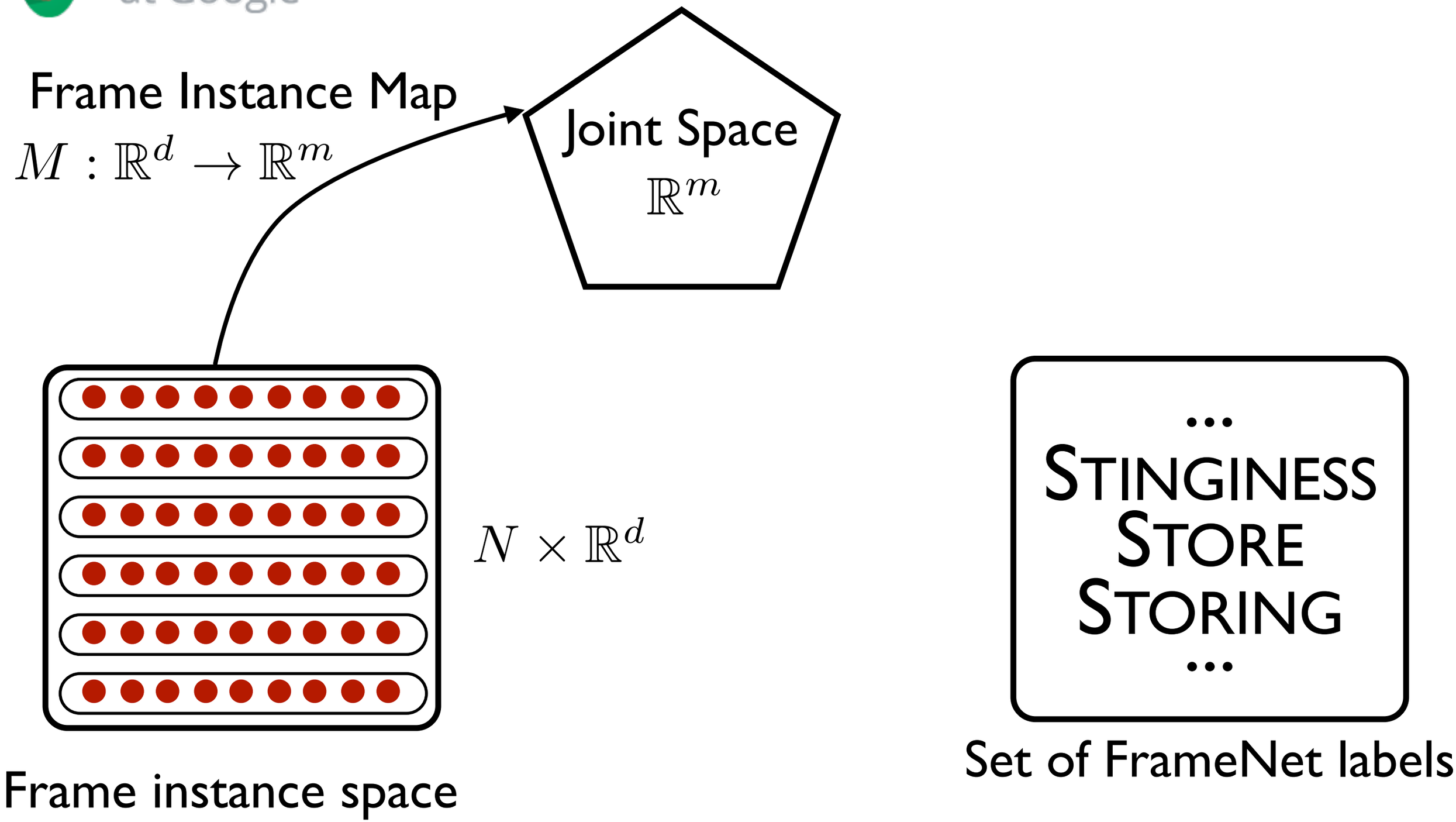


Frame instance space

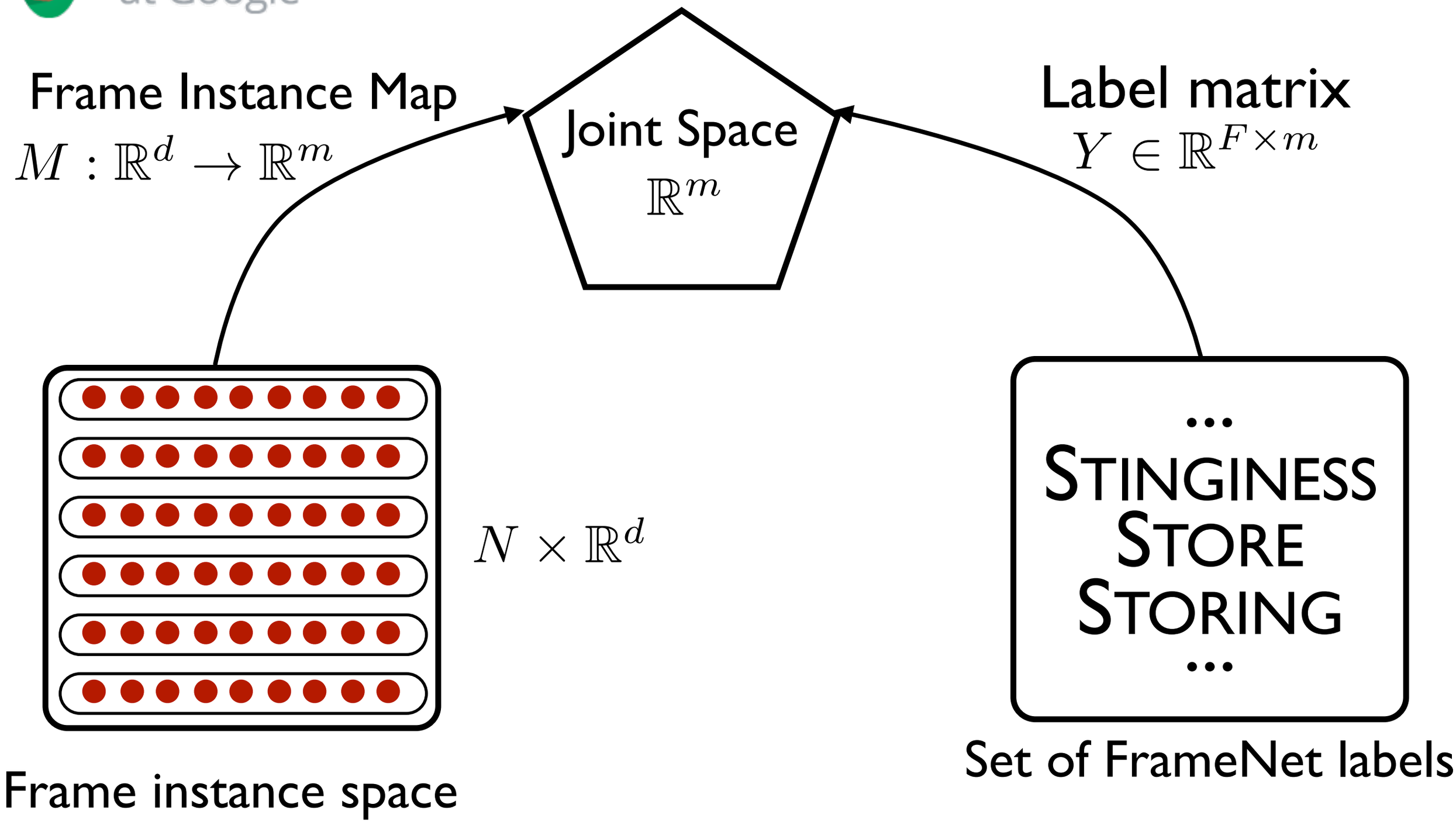
Frame Identification with Embeddings



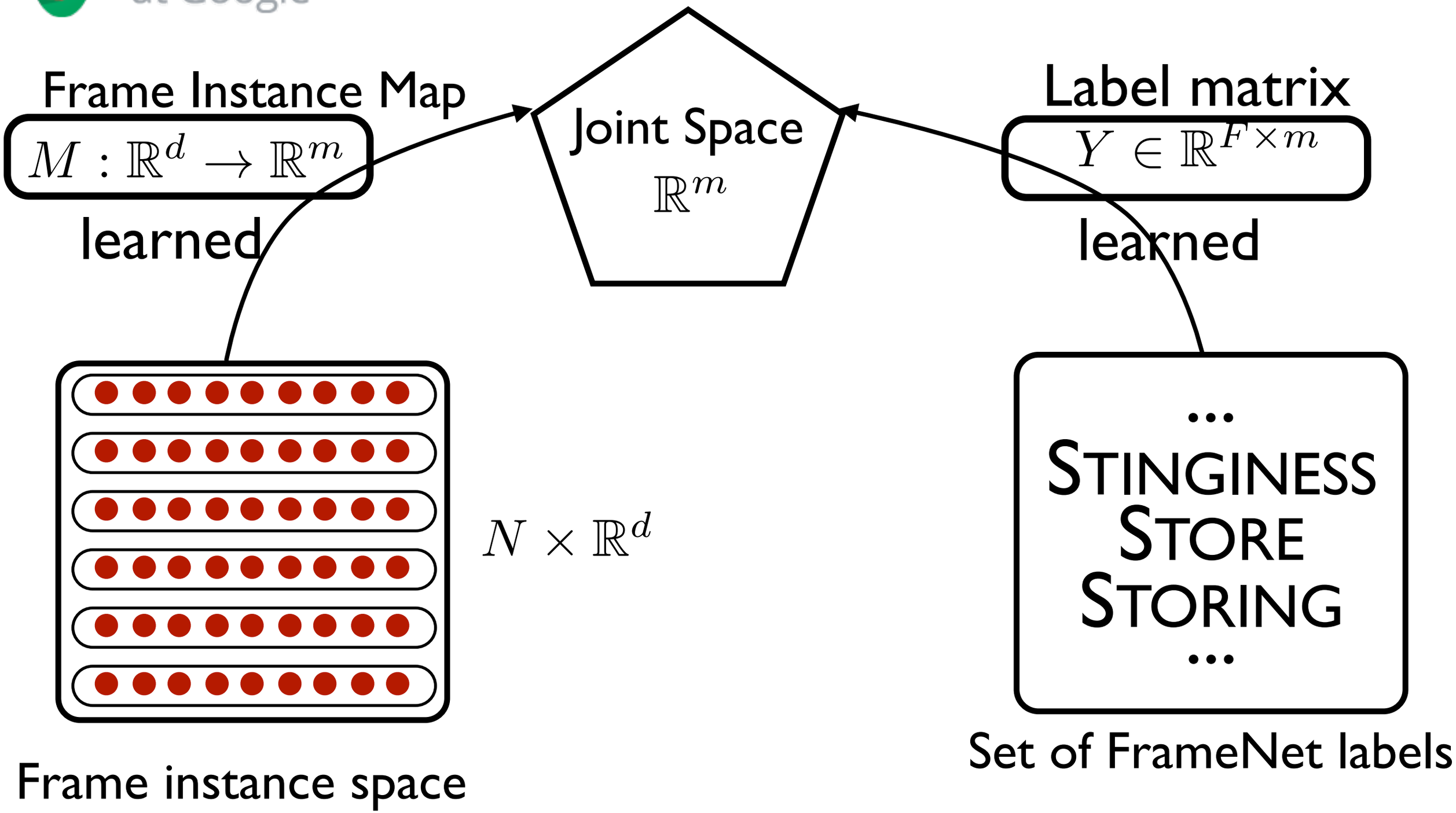
Frame Identification with Embeddings



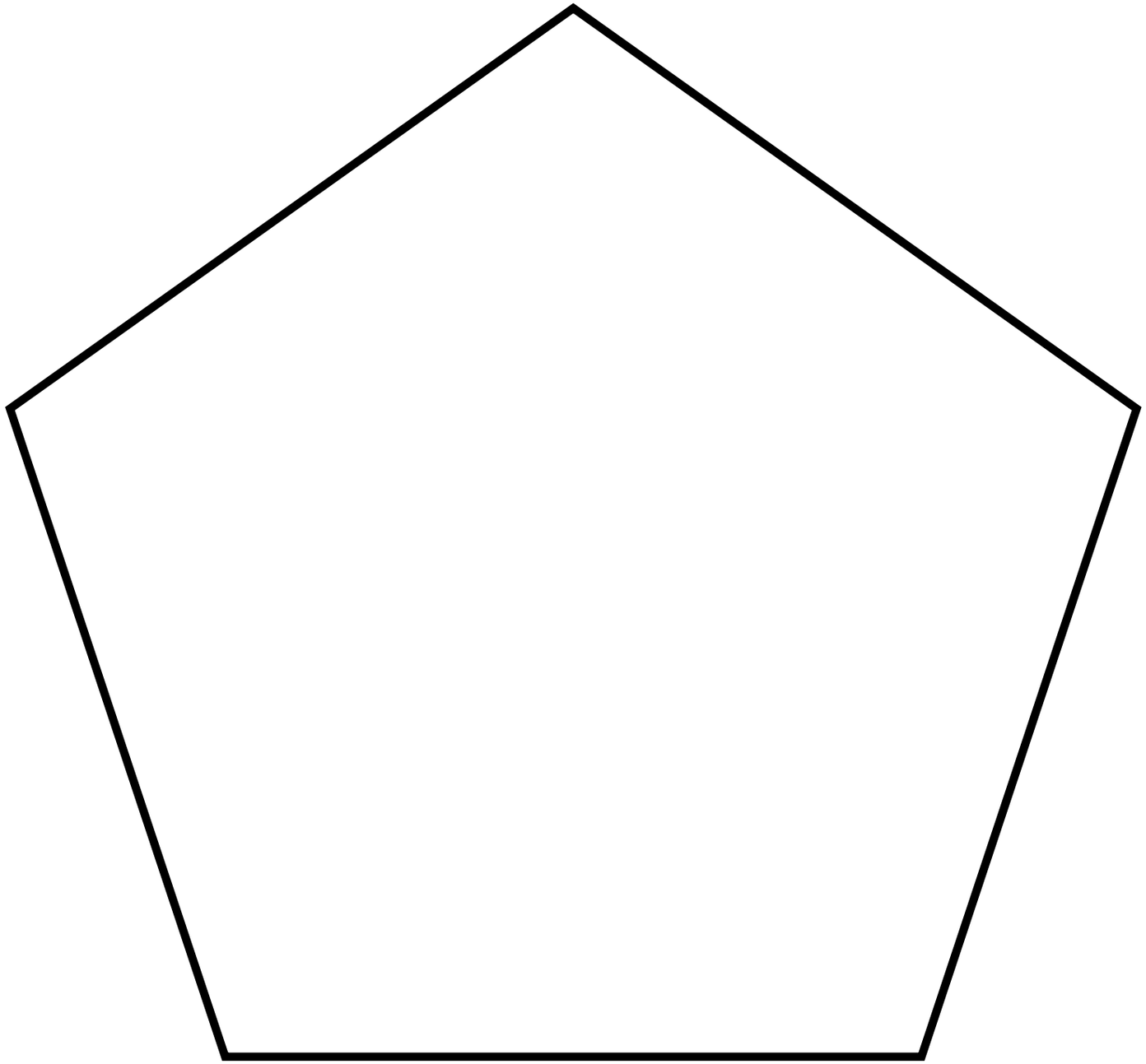
Frame Identification with Embeddings



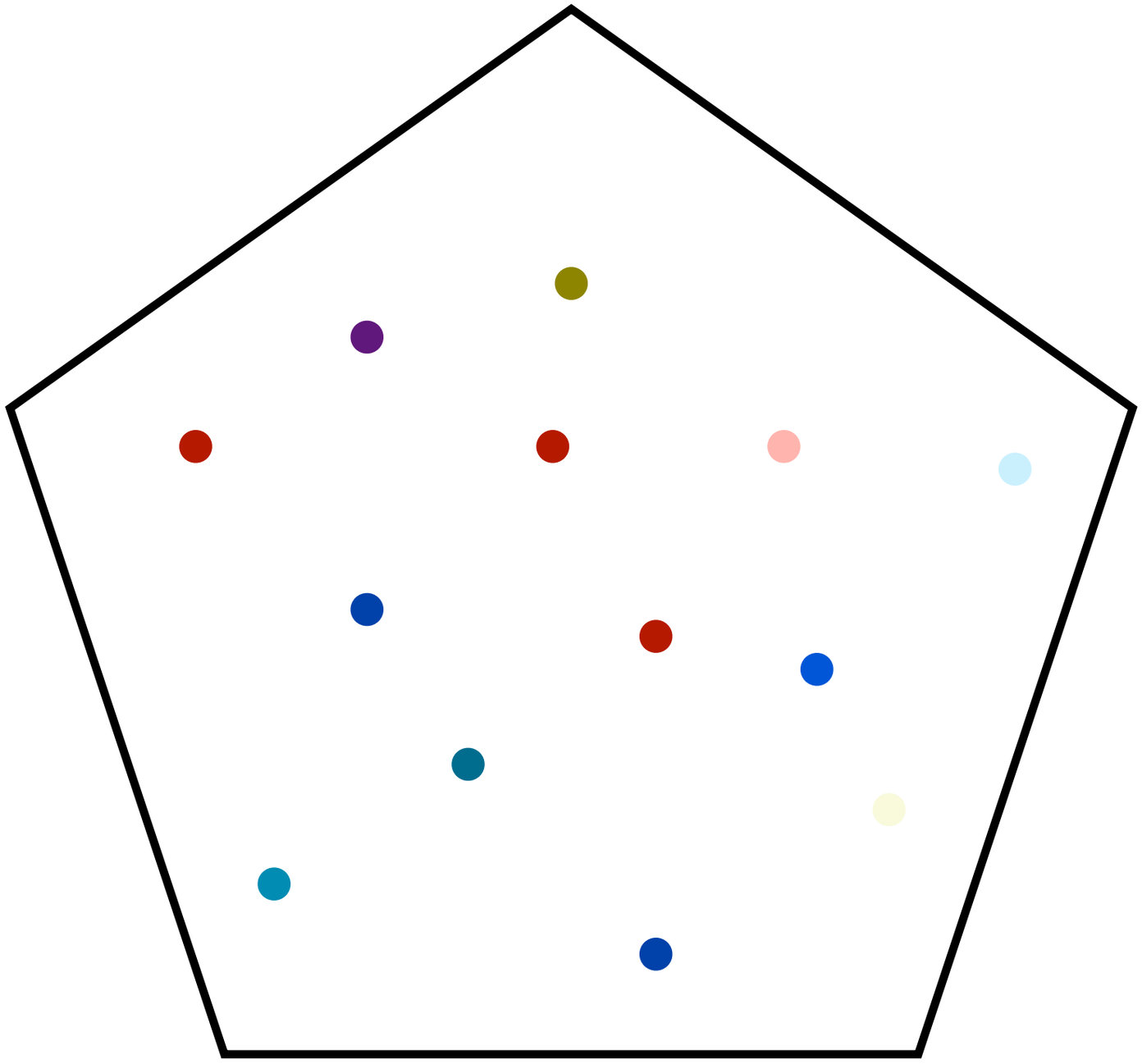
Frame Identification with Embeddings



Frame Identification with Embeddings

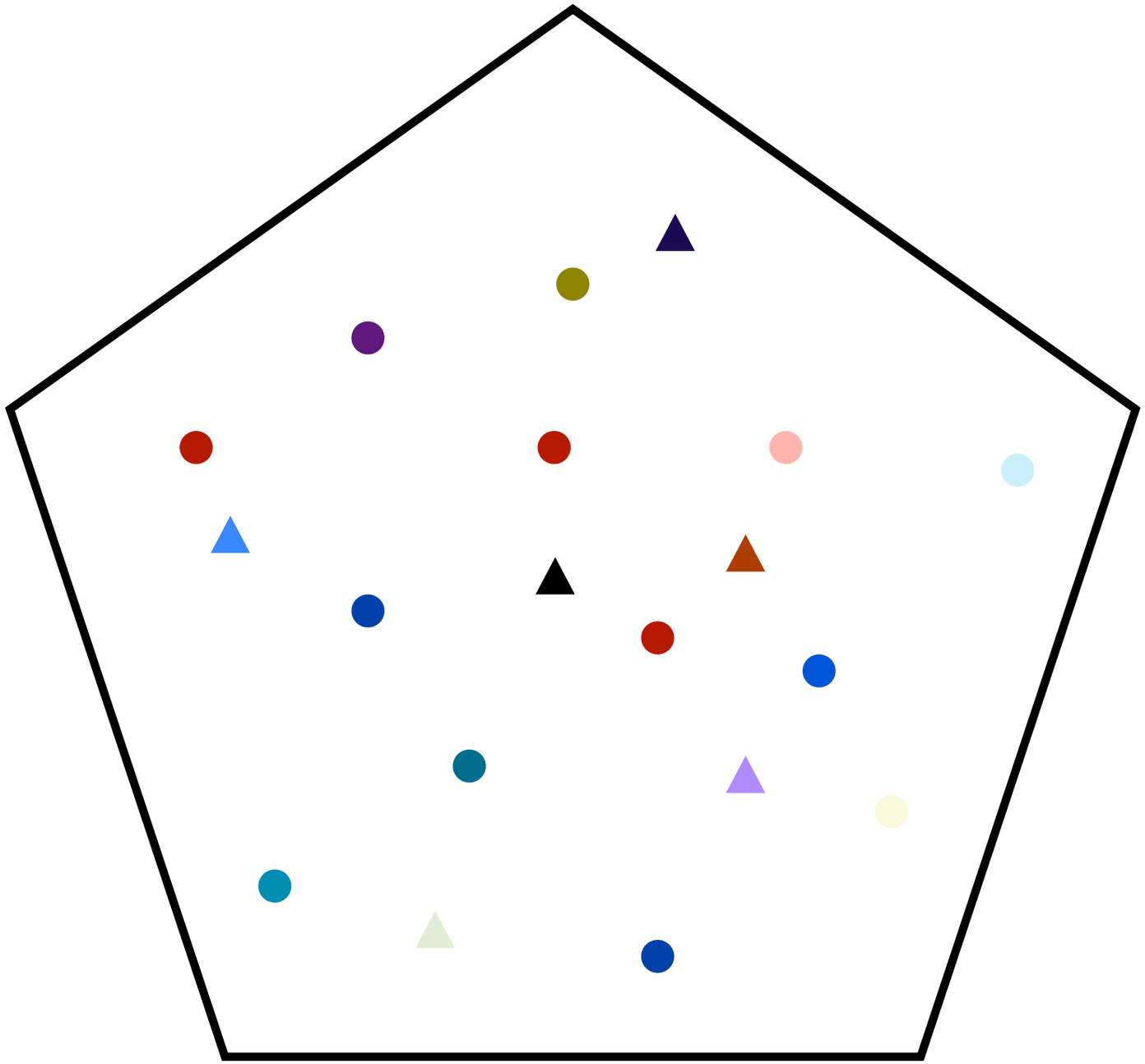


Frame Identification with Embeddings



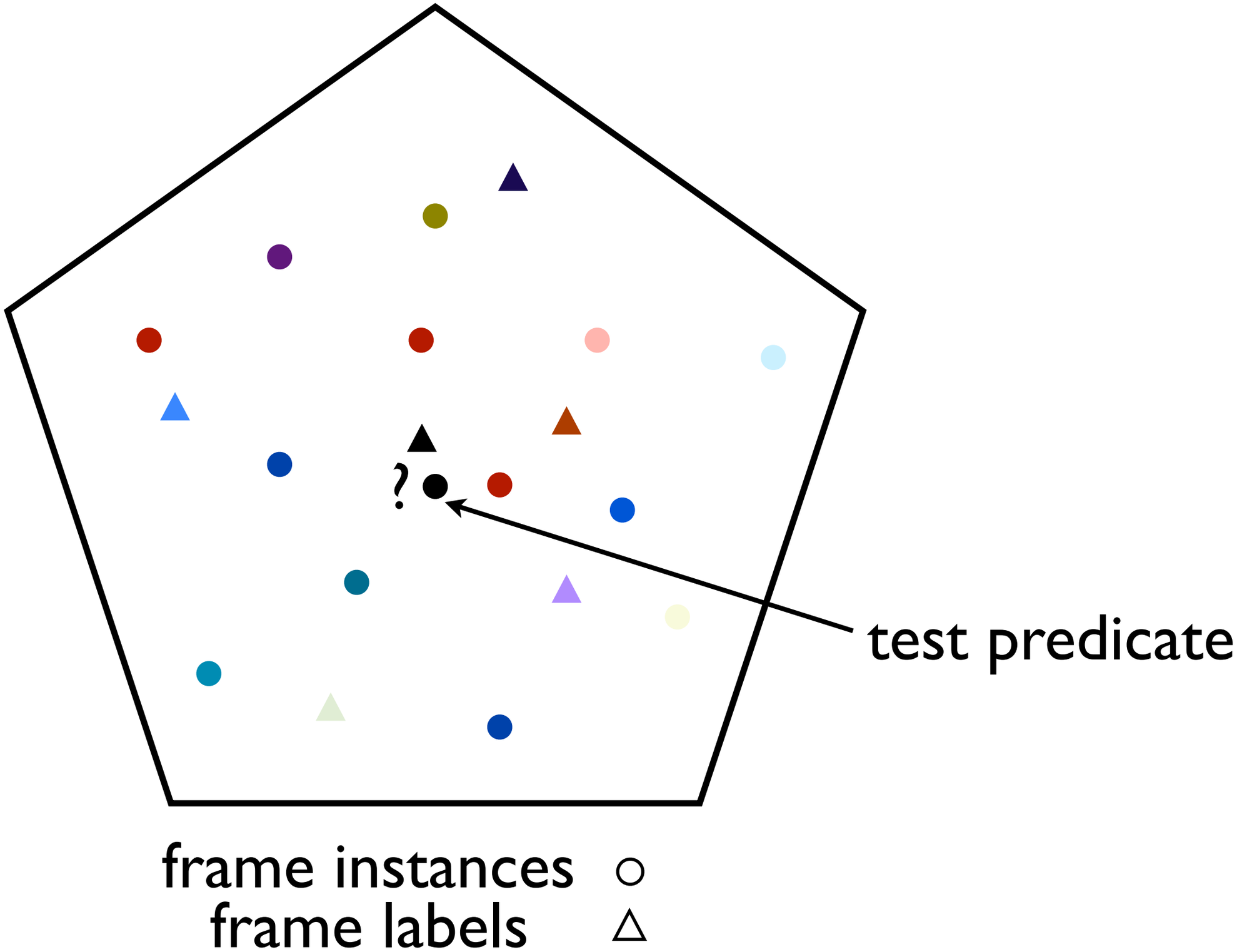
frame instances ○

Frame Identification with Embeddings

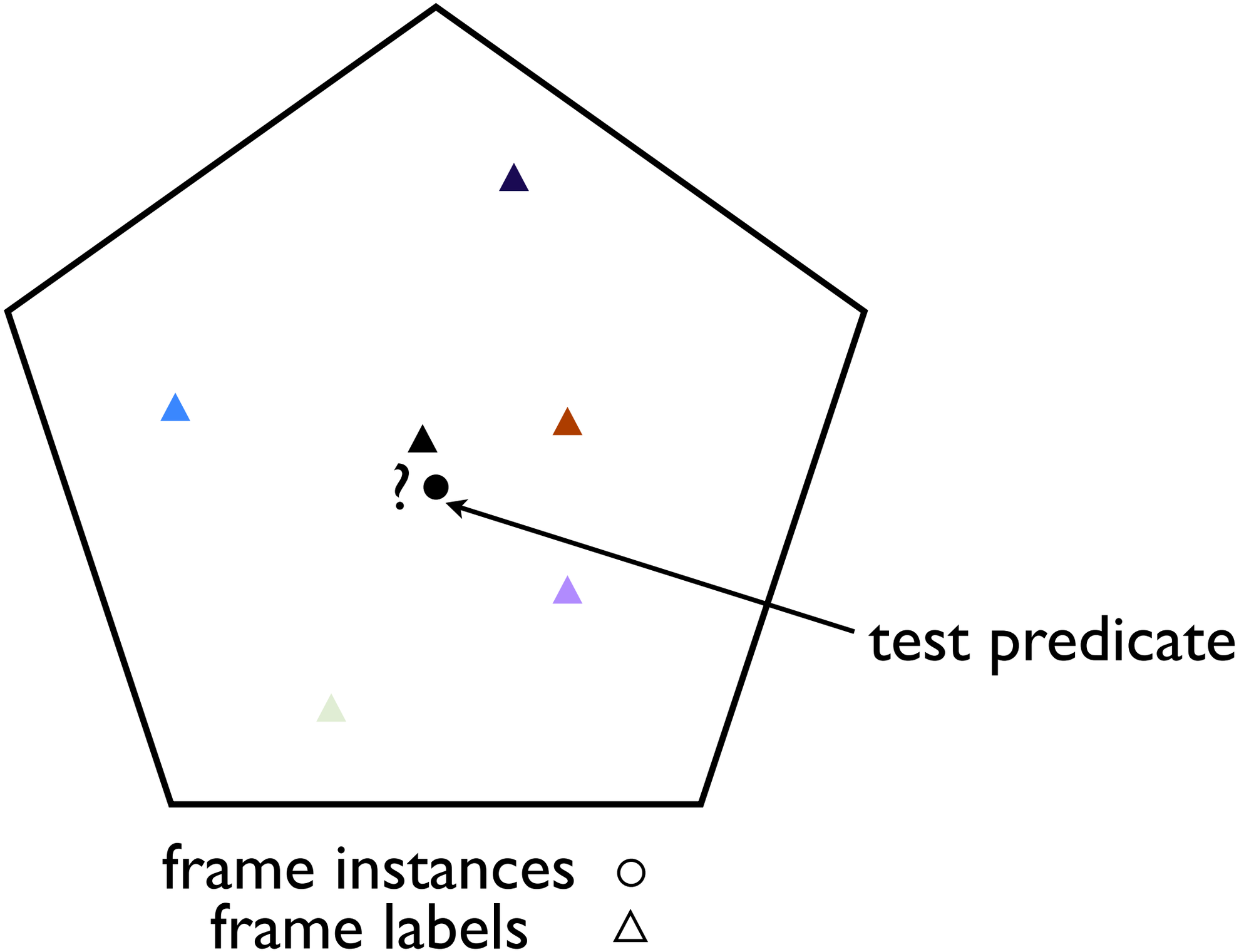


frame instances ○
frame labels △

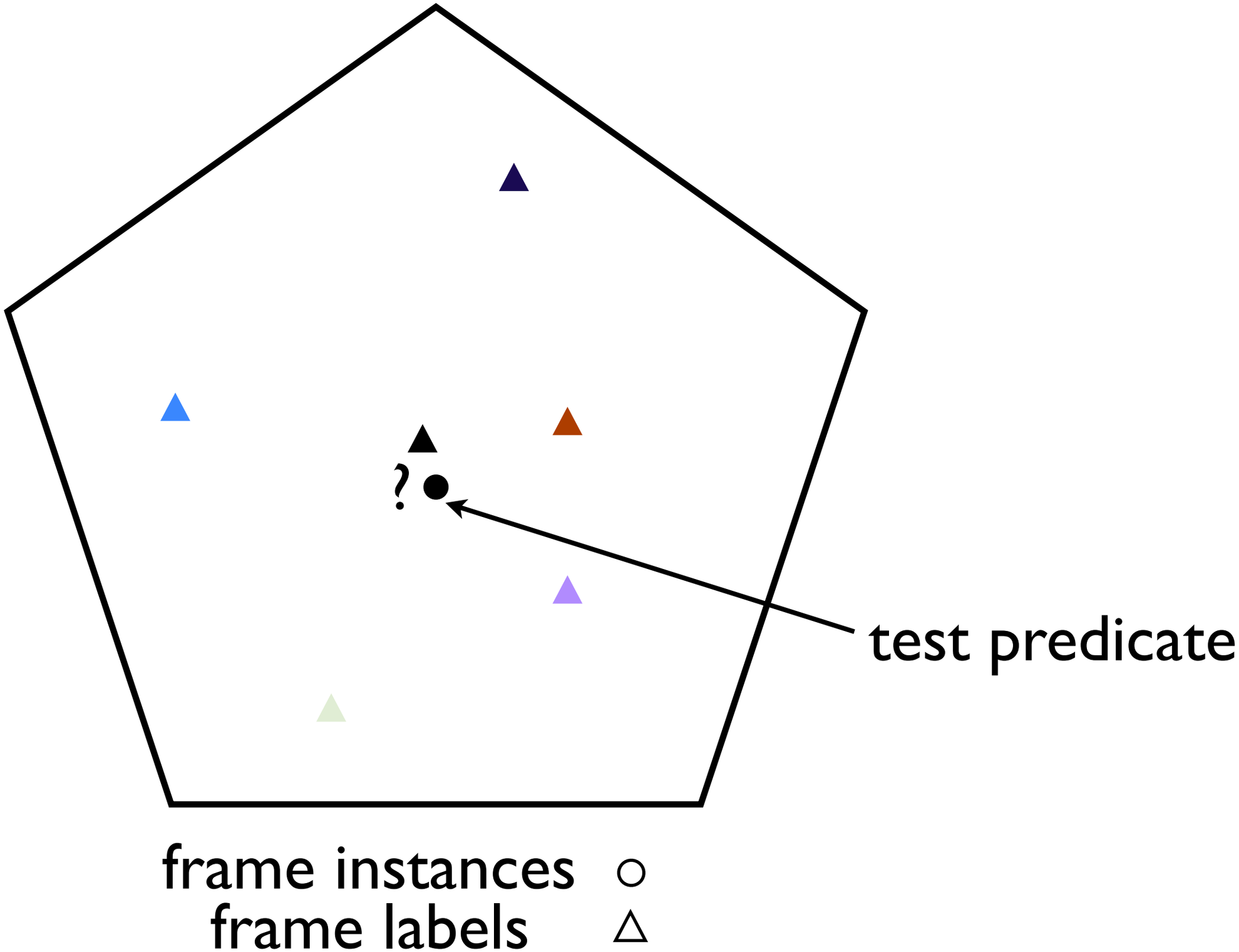
Frame Identification with Embeddings



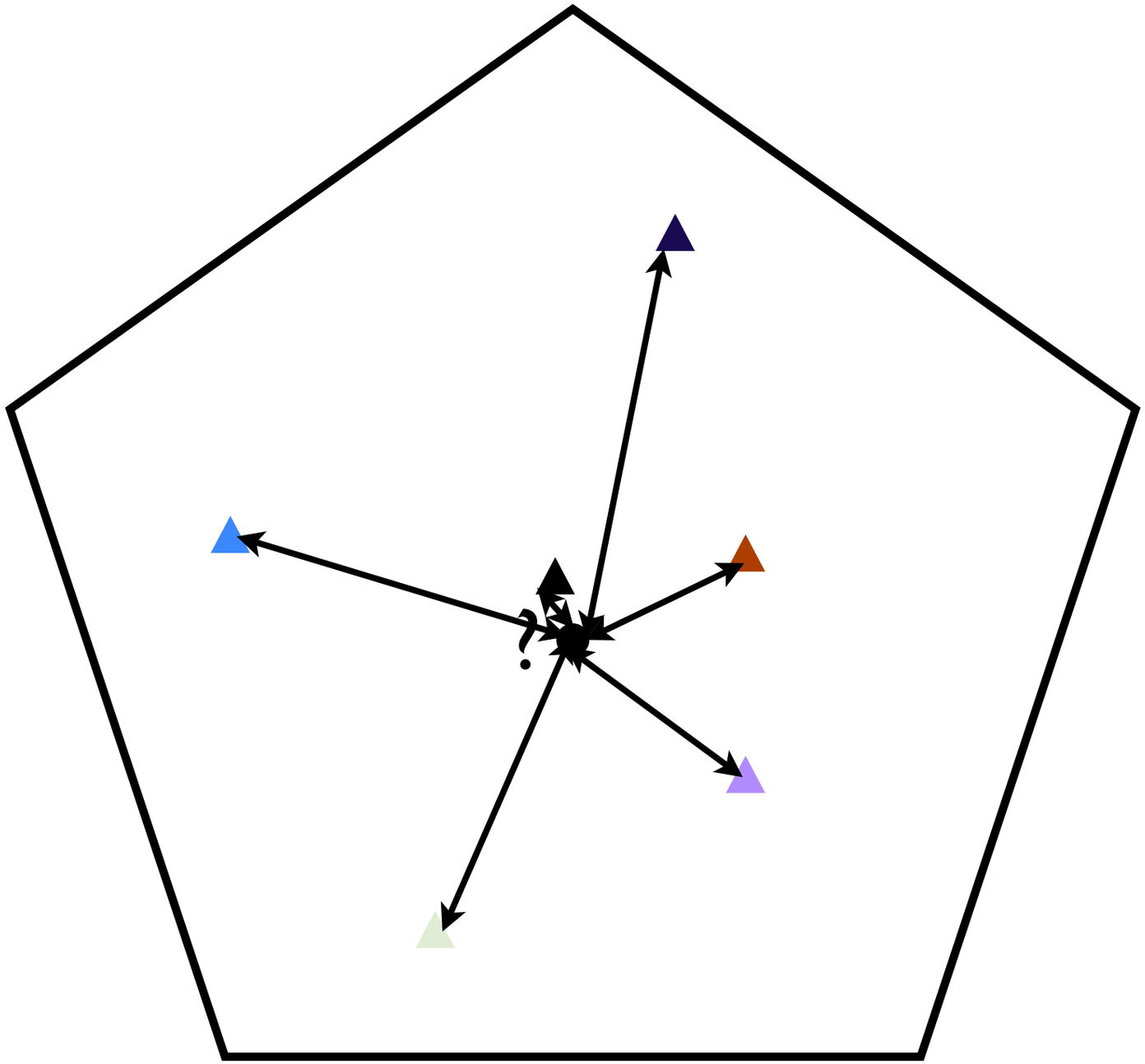
Frame Identification with Embeddings



Frame Identification with Embeddings

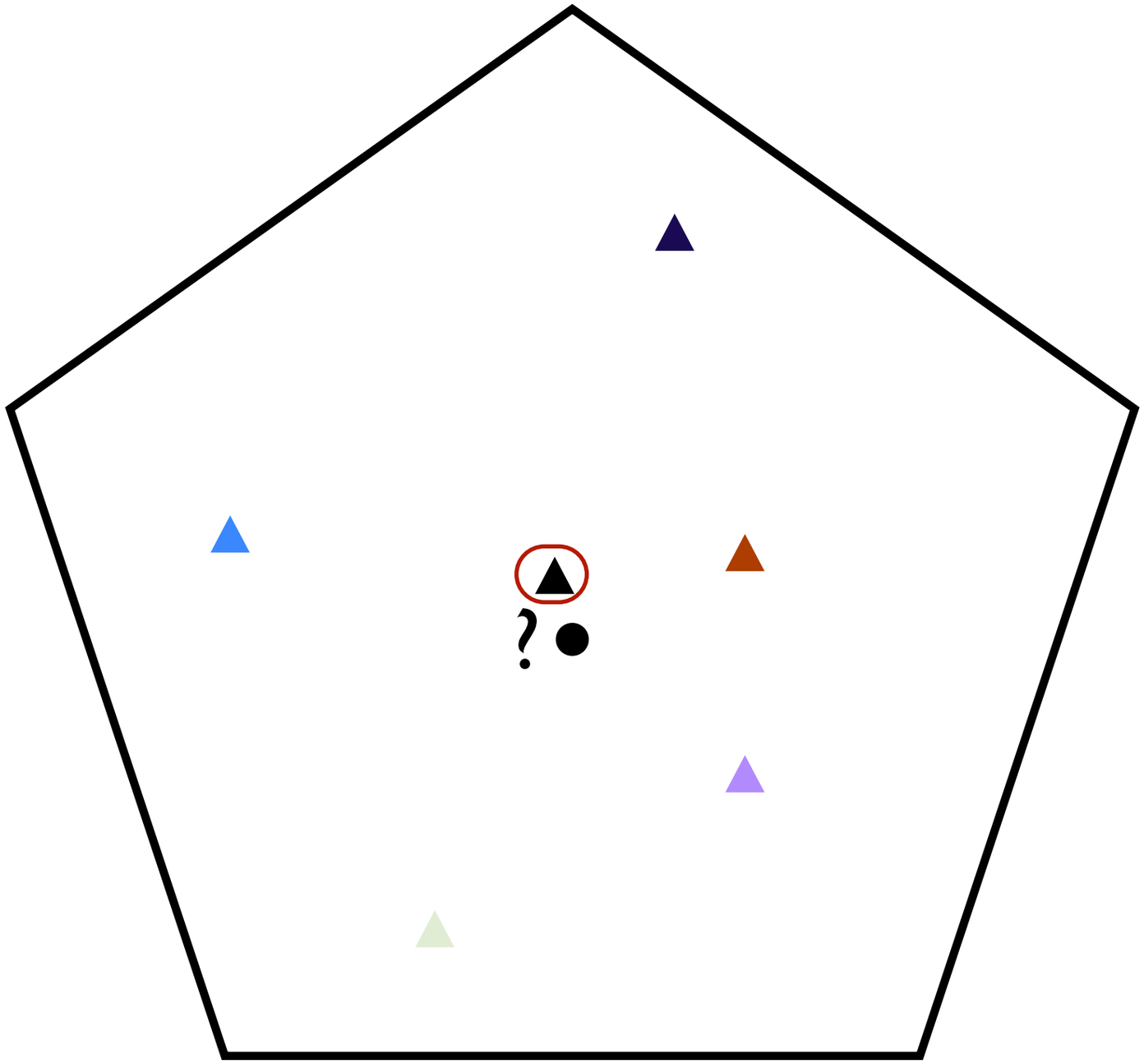


Frame Identification with Embeddings



frame instances ○
frame labels △

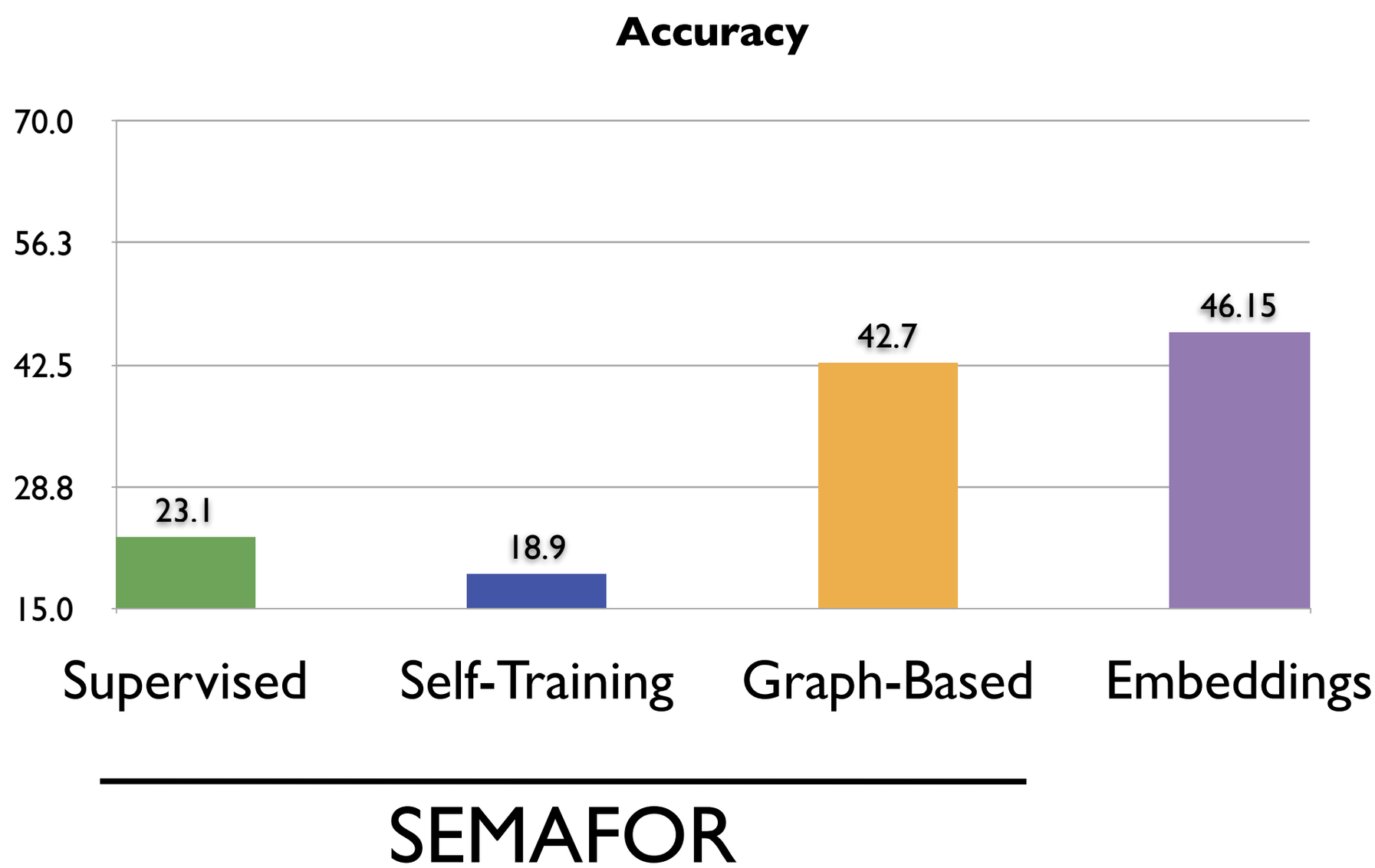
Frame Identification with Embeddings



frame instances ○
frame labels △

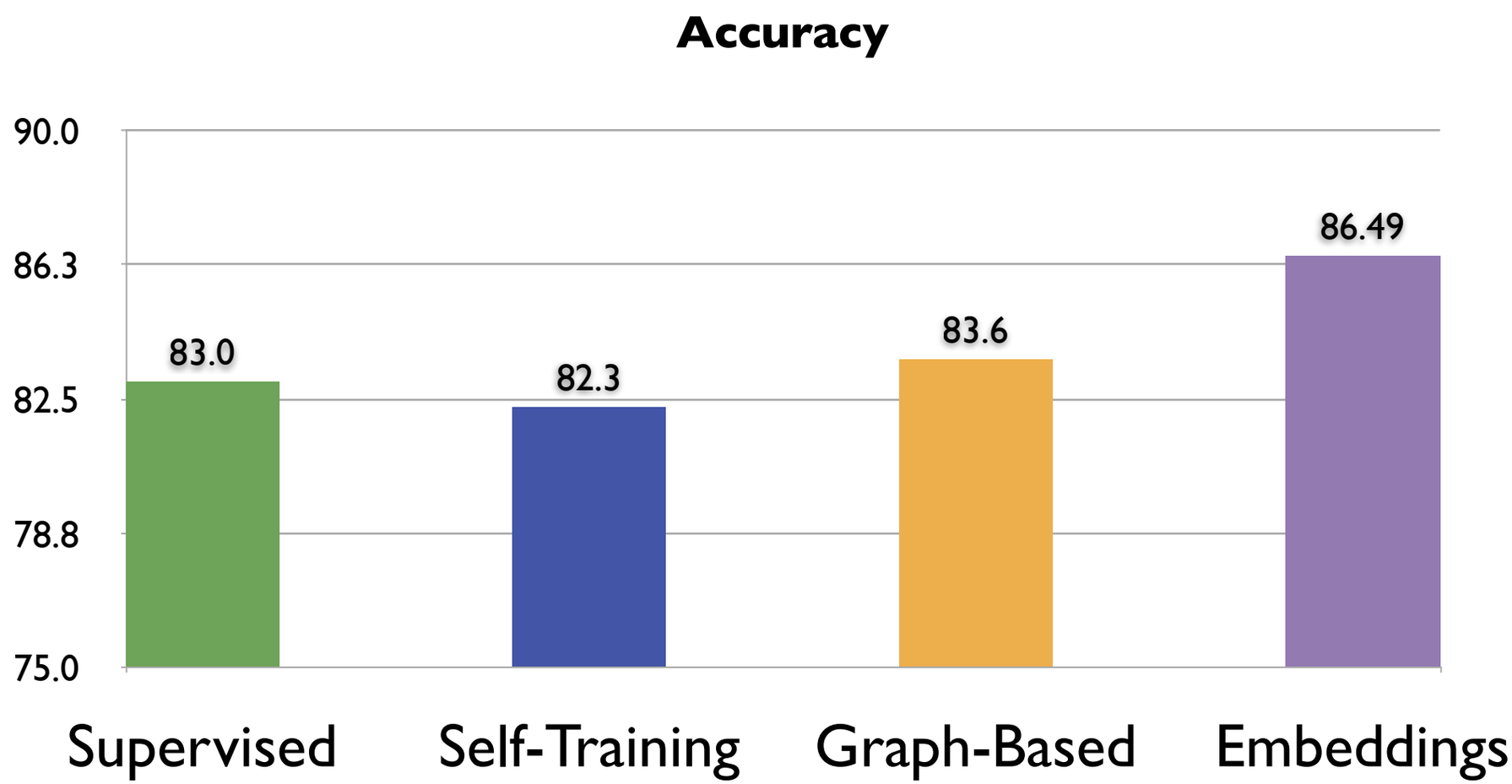
Results on Unknown Predicates

Frame Identification



Results on All Predicates

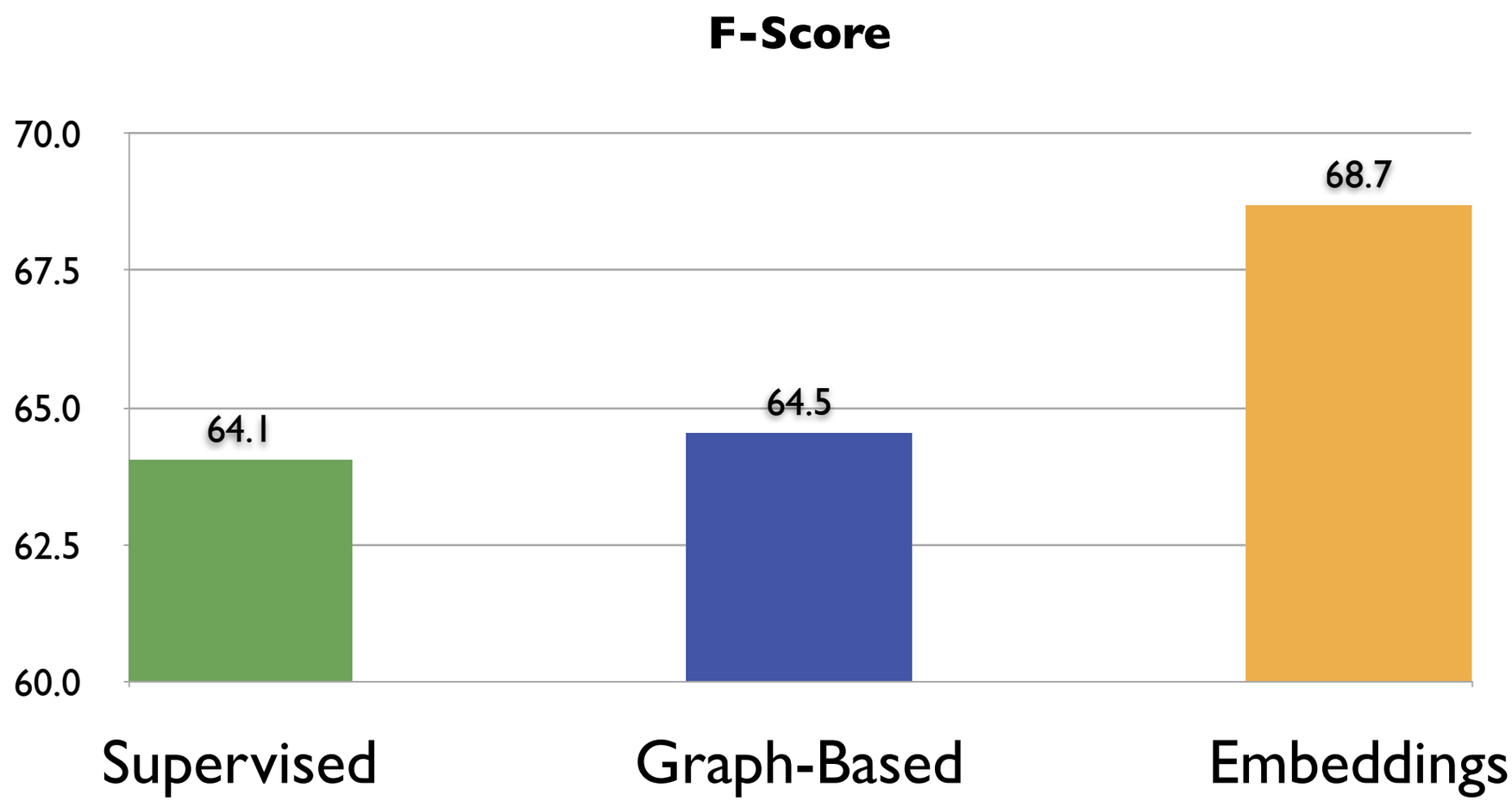
Frame Identification



SEMAFOR

Final Results

Frame-Semantic Parsing



SEMAFOR

Research at Google

Outline

- Why should we build statistical models for frame semantics?
- Overview of statistical methods
- Future directions

Research at Google

Better Models

- Very little research on argument identification
  - More non-local features
  - Using distributed representations
  - Generalization using PropBank resources

Research at Google

Data

- Number of FrameNet annotated sentences ~30 times less than PropBank/Ontonotes.

- Number of argument labels ~30 times more

- To make systems usable, we need annotations
  - inter-annotator agreement studies
  - annotation guidelines

Research at Google

Custom FrameNets

- Is a general FrameNet lexicon useful?
  - often, FrameNet frames are too general or too specific
  - is it possible to quickly build customized FrameNet lexicons for applications?
  - is it possible to use PropBank-style frames to induce FrameNet-style frames?

Research at Google

Conclusions

- Lot of exciting work in predicate-argument structure prediction
- Semi-supervised methods improve coverage
- Systems trained on small amounts of FrameNet-style data shown to be useful
- More annotations will result in usable systems

Research at Google

Thank You