Decompositional Semantics

Rachel Rudinger

April 19 2021

A story about semantic annotation…

Traditional Semantic Annotation

Who did what to whom?

AGENT PATIENT

Alex shattered the window.

AGENT *Participant that performs the action.*

PATIENT *Participant that undergoes the action.*

Traditional Semantic Annotation

AGENT PATIENT ???

Alex shattered the window with a hammer.

AGENT	*Participant that performs the action.*
PATIENT	*Participant that undergoes the action and changes state.*

Traditional Semantic Annotation

AGENT PATIENT INSTRUMENT

Alex shattered the window with a hammer.

AGENT	*Participant that performs the action.*
PATIENT	*Participant that undergoes the action and changes state.*
INSTRUMENT	*Participant used to carry out the action.*

Traditional Semantic Annotation

??? PATIENT

The cold air shattered the window.

AGENT	*Participant that performs the action.*
PATIENT	*Participant that undergoes the action and changes state.*
INSTRUMENT	*Participant used to carry out the action.*

Traditional Semantic Annotation

FORCE PATIENT

The cold air shattered the window.

AGENT — *Participant that performs the action with intent.*

FORCE — *Participant that causes the action without intent.*

PATIENT — *Participant that undergoes the action and changes state.*

INSTRUMENT — *Participant used to carry out the action.*

Traditional Semantic Annotation

AGENT? FORCE?

???

PATIENT

Alex *accidentally* shattered the window.

AGENT	*Participant that performs the action with intent.*	FORCE	*Participant that causes the action without intent.*
PATIENT	*Participant that undergoes the action and changes state.*		
INSTRUMENT	*Participant used to carry out the action.*		

Traditional Semantic Annotation

FORCE? INSTRUMENT?

???

PATIENT

Alex's singing shattered the window.

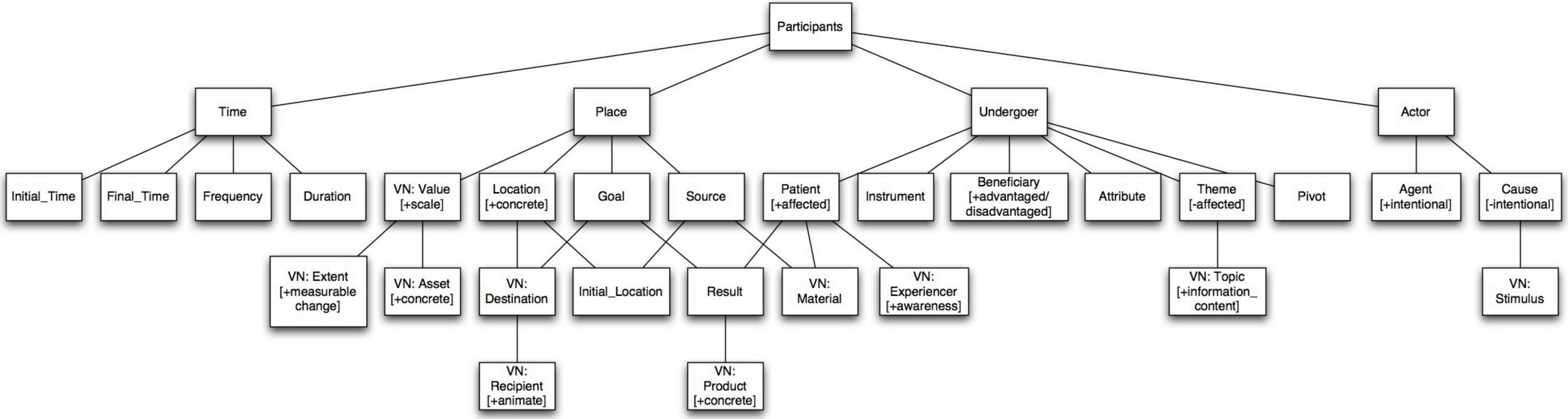
AGENT — *Participant that performs the action with intent.*

PATIENT — *Participant that undergoes the action and changes state.*

INSTRUMENT — *Participant used to carry out the action.*

FORCE — *Participant that causes the action without intent.*

VerbNet Role Hierarchy

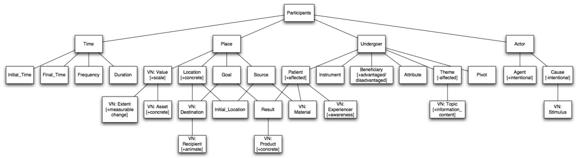


JOHNS HOPKINS UNIVERSITY

A hierarchical unification of LIRICS and VerbNet semantic roles. Bonial, Corvey, Palmer, Petukhova, and Bunt. ICSC. 2011.

Practical Challenges

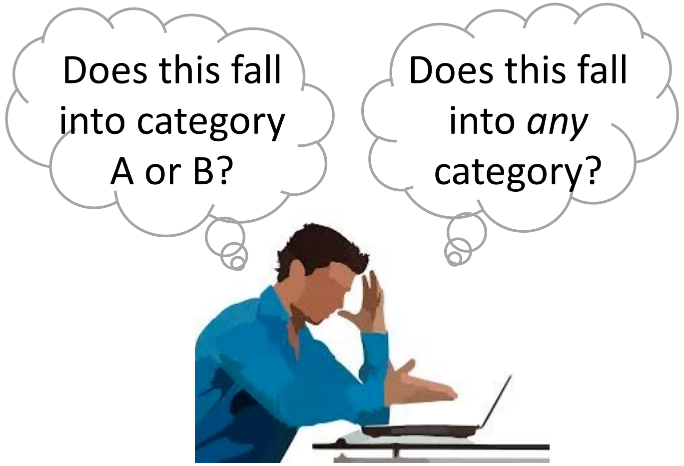
Establish ontology.



Train expert annotators.

Annotate.

Annotation challenges.

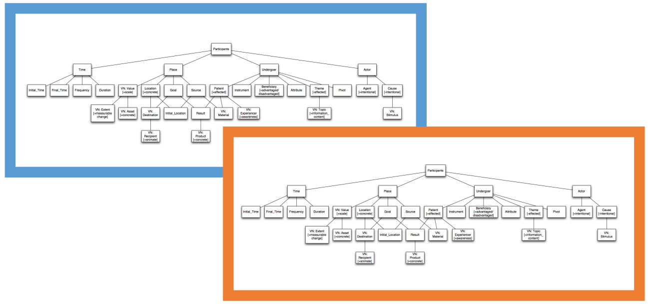


Modify ontology.

Retrain?

Re-annotate?

Mapping between ontologies?



JOHNS HOPKINS UNIVERSITY

Dowty (1991)

"…and as soon as we try to be precise about exactly what Agent, Patient, etc., 'mean', it is all too subject to difficulties and apparent counterexamples."

"…we may have a hard time pinning down the traditional role type because role types are simply not discrete categories at all, but rather are cluster concepts"

Thematic proto-roles and argument selection. David Dowty. Language. 1991.

Dowty's Proto-Agent and Proto-Patient Properties ("Semantic Proto-Roles")

Proto-Agent properties	Proto-Patient properties
Volitional involvement in the event or state	Undergoes change of state
Sentience (and/or perception)	Incremental theme
Causing an event or change of state in another participant	Causally affected by another participant
Movement (relative to another participant)	Stationary relative to movement of another participant
Exists independently of the event named by the verb	Does not exist independently of the event, or not at all

JOHNS HOPKINS UNIVERSITY

Thematic proto-roles and argument selection. David Dowty. Language. 1991.

The Decompositional Approach

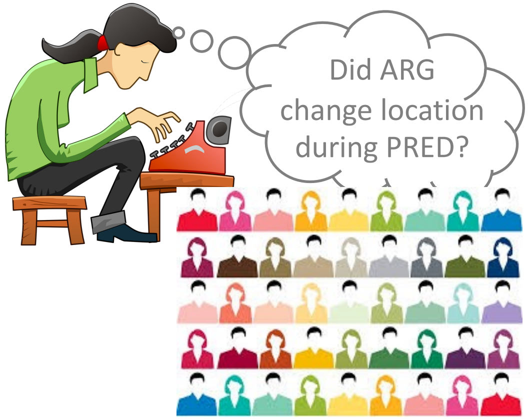
Identify properties of interest.

Translate properties into templatic English questions.

Pose each question independently to non-expert annotators.

Extend inventory of properties.

Make new annotations (but keep the old)!



JOHNS HOPKINS UNIVERSITY

Decompositional Semantics Initiative

"Rapid, simple, commonsensical annotations of meaning"

1. Target aspects of meaning at the phrase- or sentence level.
2. Simple, linguistically- or cognitively-motivated properties.
3. Many independent labels.
4. Straightforward questions for crowd workers.

JOHNS HOPKINS UNIVERSITY

UNIVERSITY of ROCHESTER

http://decomp.io

Decompositional Semantics Initiative

"Rapid, simple, commonsensical annotations of meaning"

Semantic Proto-Roles

Genericity

Time

Event Factuality

PredPatt

Decomp Toolkit

Word Sense

Diverse Natural Language Inference

Cross-lingual Decompositional Semantic Parsing

Common Sense Inference

ParaBank 1 & 2

JOHNS HOPKINS UNIVERSITY

UNIVERSITY of ROCHESTER

http://decomp.io

Dataset 1: Semantic Proto-Roles

Dataset 2: Event Factuality

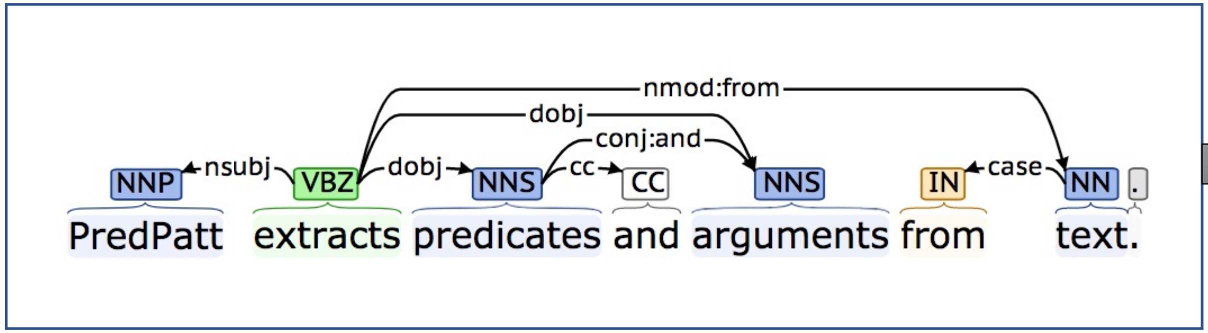
Dataset 3: Temporal Relations

Dataset 4: Genericity

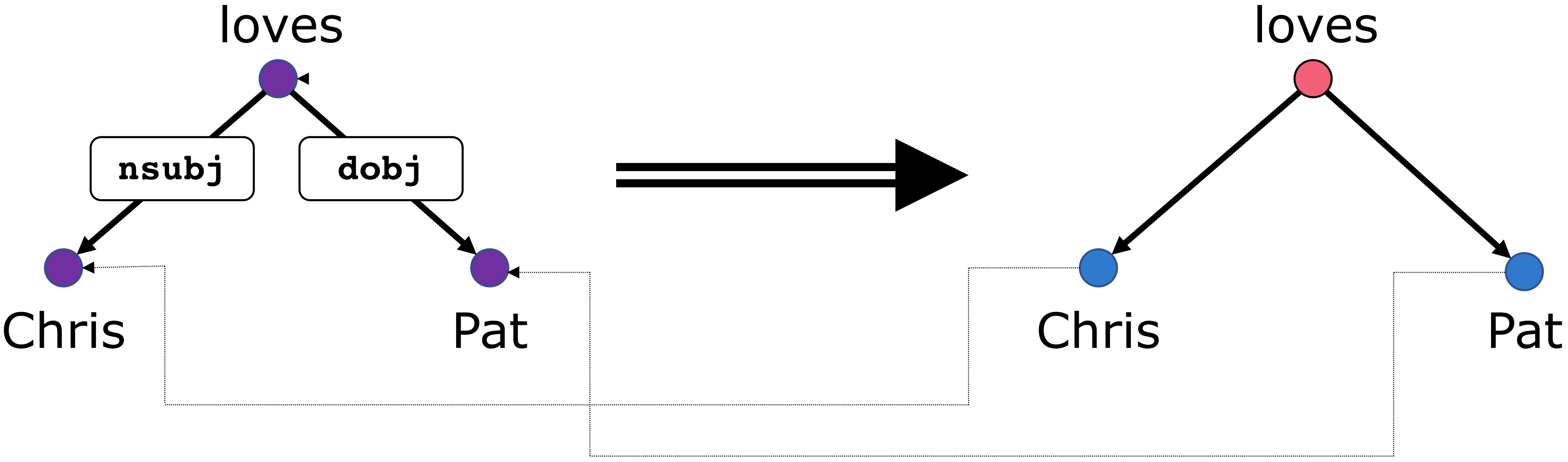
Before we dive into the data…

Predicate-Argument Identification with PREDPATT

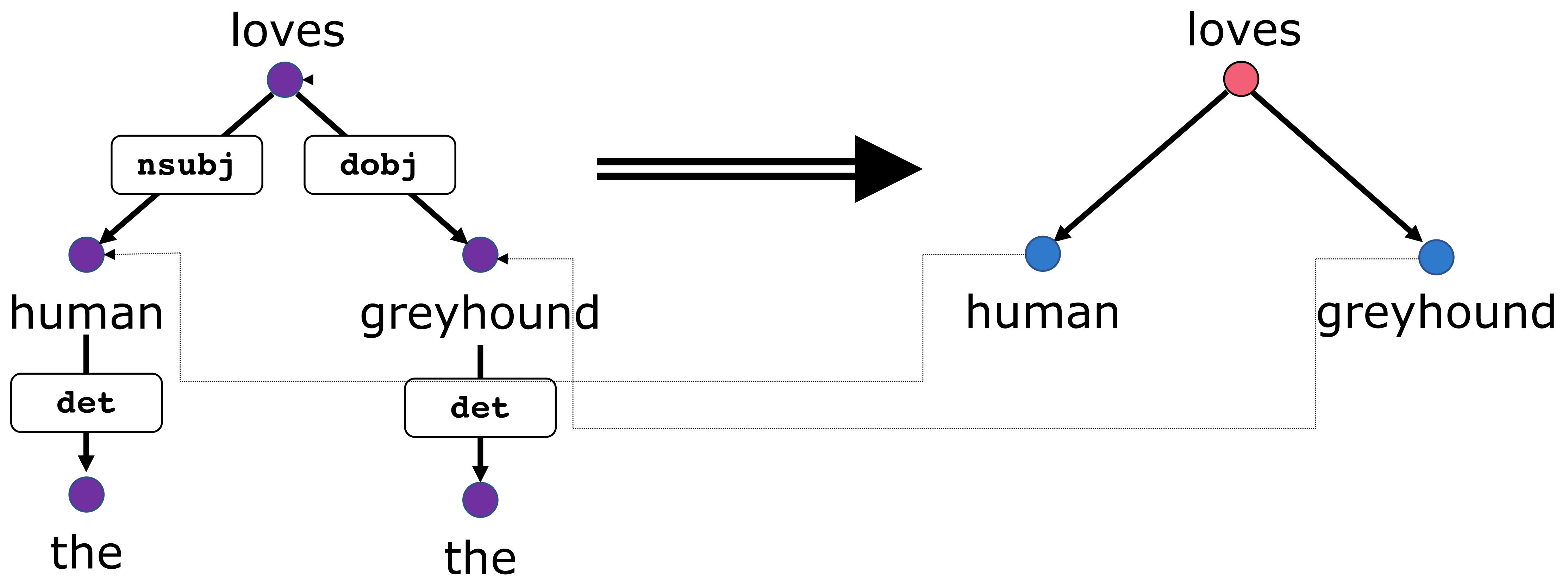
- Decomp annotation protocols rely on predicate-argument structure.
- PredPatt: series of rules to map Universal Dependencies (UD) parse to unlabeled predicate-argument structure.
- Scalability and (potential) Multilinguality: Piggy-backing on UD resources.



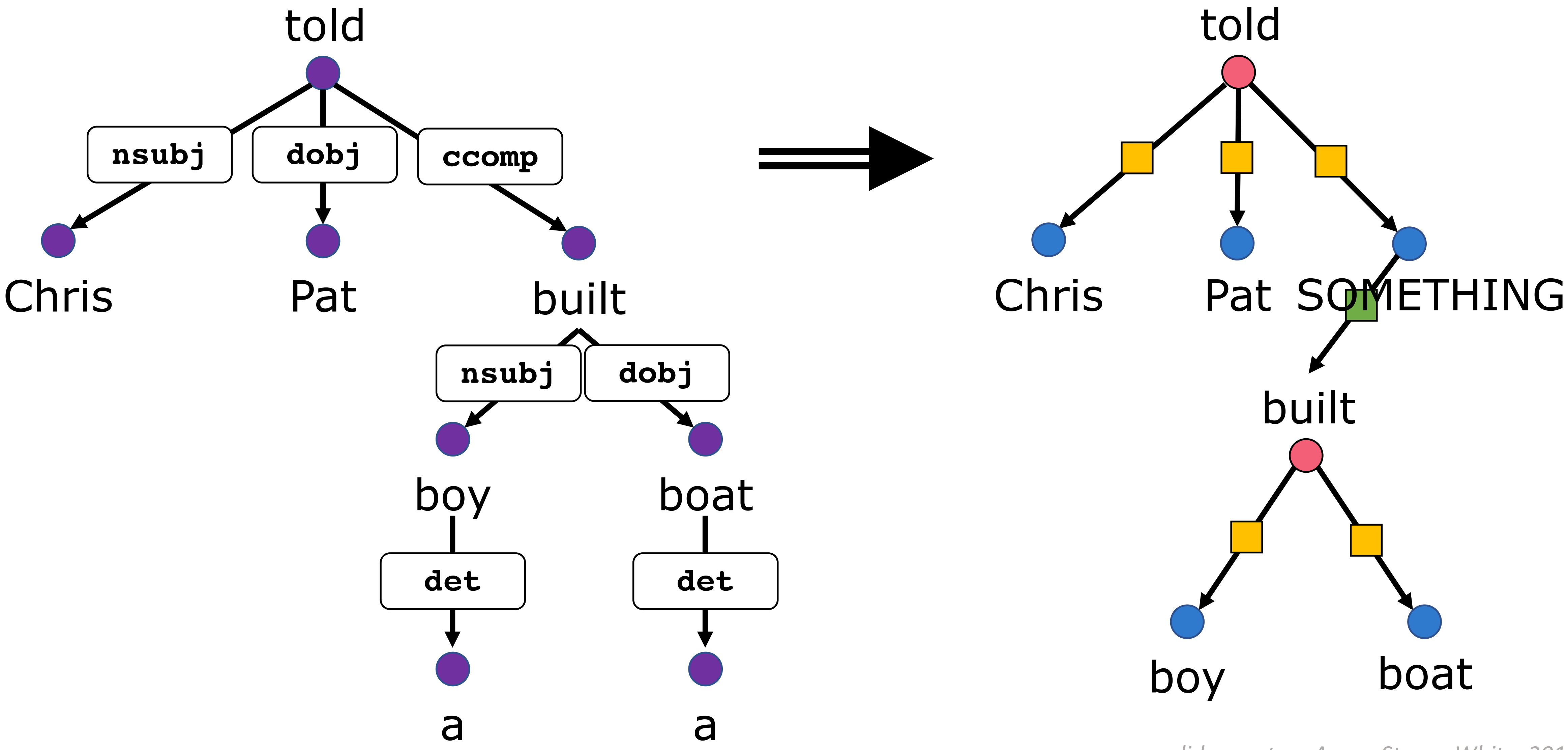

```
?a extracts ?b from ?c
       ?a:   PredPatt
       ?b:   predicates
       ?c:   text
?a extracts ?b from ?c
       ?a:   PredPatt
       ?b:   arguments
       ?c:   text
```



slide courtesy Aaron Steven White, 2019



slide courtesy Aaron Steven White, 2019



slide courtesy Aaron Steven White, 2019

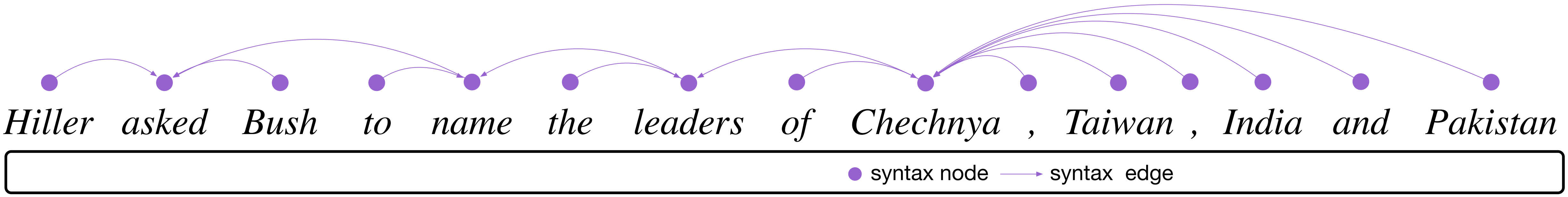
Important note

No typing beyond:

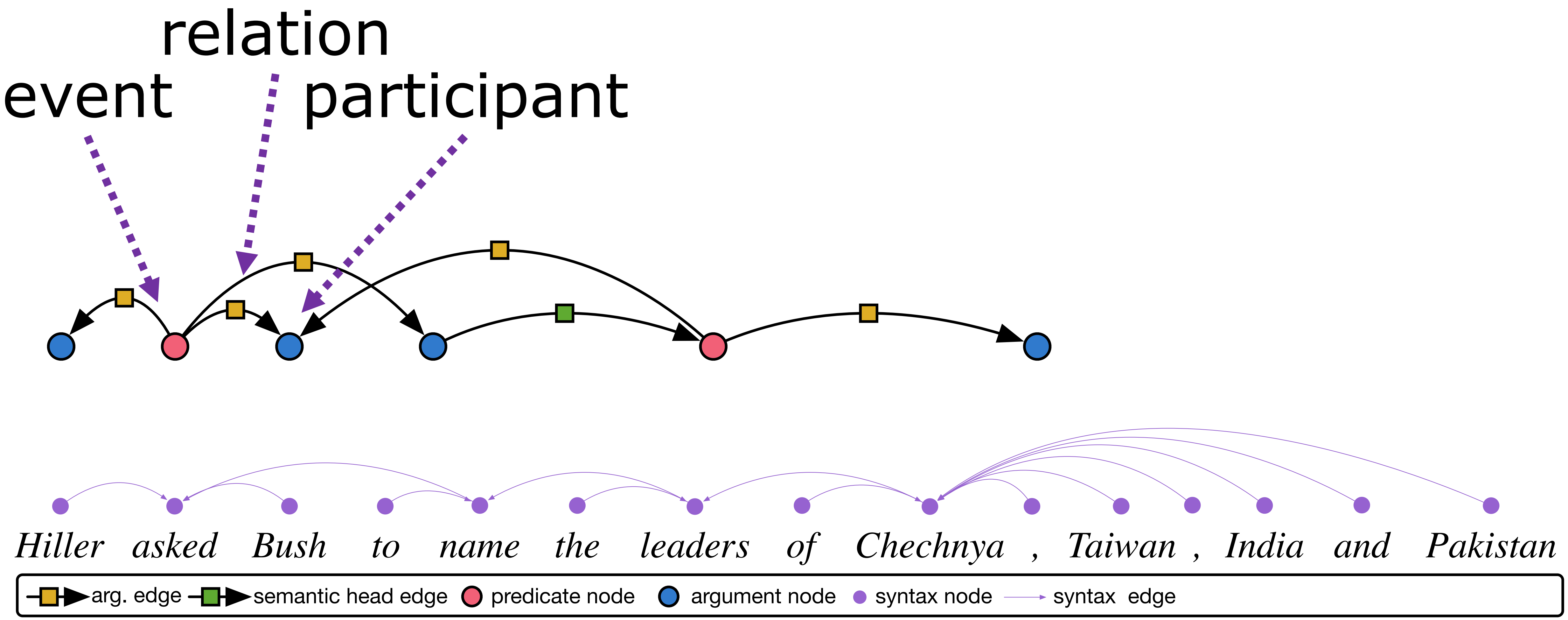
- event v. participant

- argument v. head

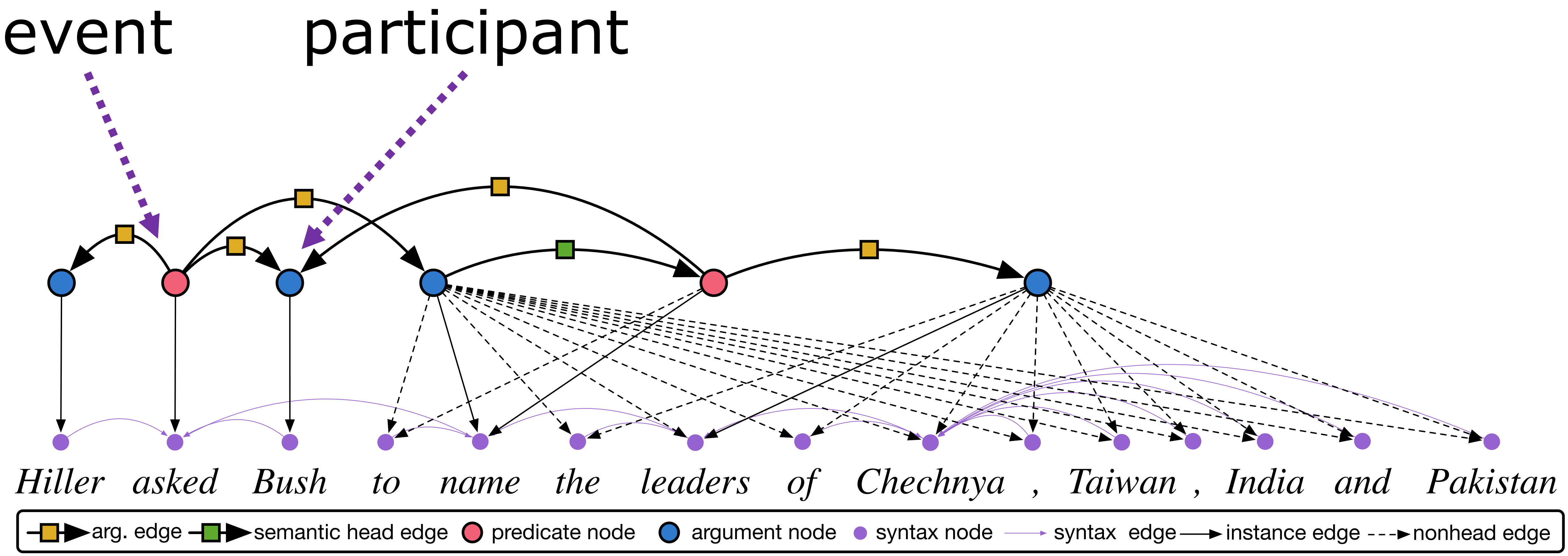
Hiller asked Bush to name the leaders of Chechnya , Taiwan , India and Pakistan



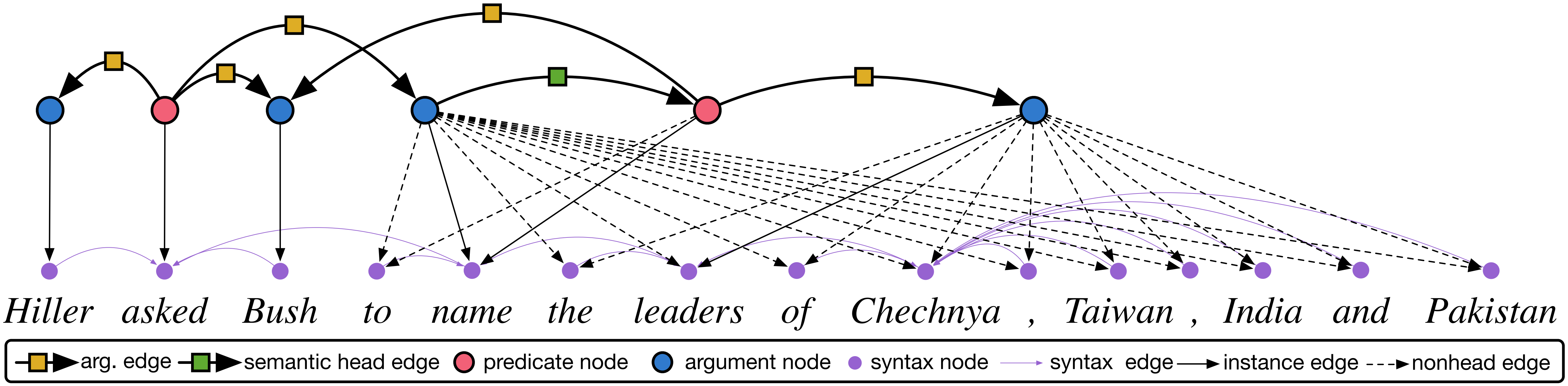
slide courtesy Aaron Steven White, 2019



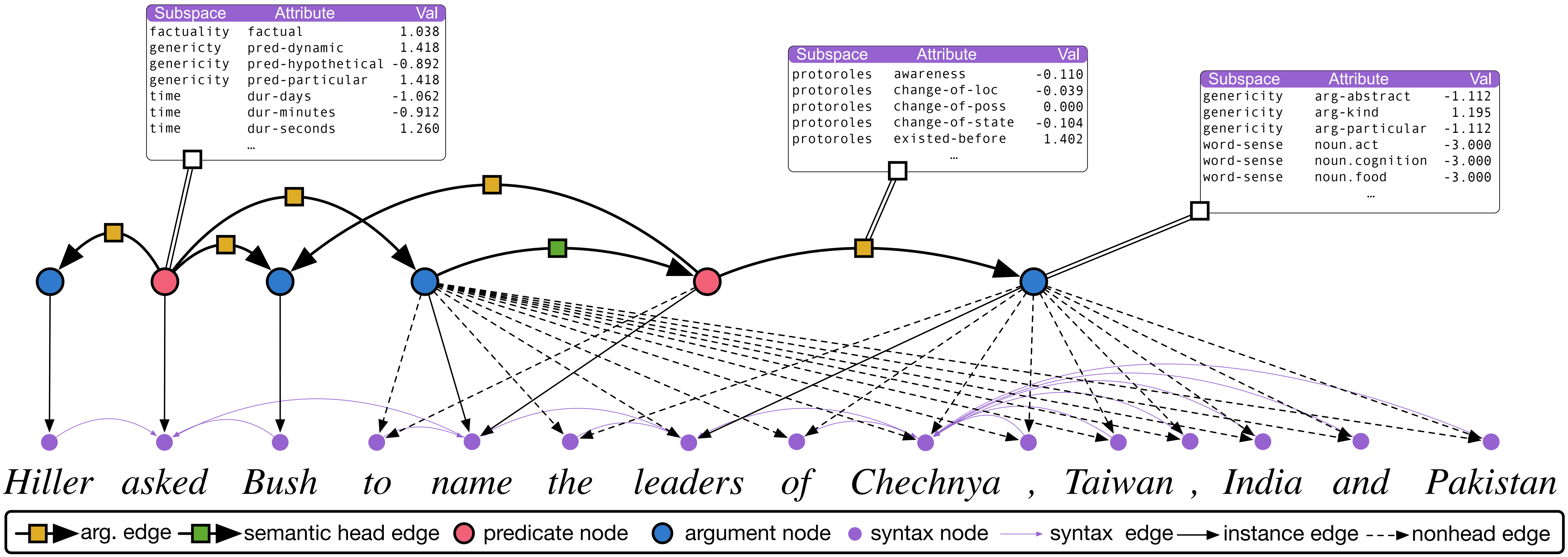
slide courtesy Aaron Steven White, 2019



slide courtesy Aaron Steven White, 2019



slide courtesy Aaron Steven White, 2019



slide courtesy Aaron Steven White, 2019

Diving into the data…

Dataset 1: Semantic Proto-Roles

Dataset 2: Event Factuality

Dataset 3: Temporal Relations

Dataset 4: Genericity

Traditional Semantic Role Labeling

AGENT PATIENT INSTRUMENT

Alex shattered the window with a hammer.

AGENT	*Participant that performs the action.*
PATIENT	*Participant that undergoes the action and changes state.*
INSTRUMENT	*Participant used to carry out the action.*
FORCE	*Participant that causes the action without intent.*

Etc…

Dowty (1991)

"…and as soon as we try to be precise about exactly what Agent, Patient, etc., 'mean', it is all too subject to difficulties and apparent counterexamples."

"…we may have a hard time pinning down the traditional role type because role types are simply not discrete categories at all, but rather are cluster concepts"

Thematic proto-roles and argument selection. David Dowty. Language. 1991.

Dowty's Proto-Agent and Proto-Patient Properties ("Semantic Proto-Roles")

Proto-Agent properties	Proto-Patient properties
Volitional involvement in the event or state	Undergoes change of state
Sentience (and/or perception)	Incremental theme
Causing an event or change of state in another participant	Causally affected by another participant
Movement (relative to another participant)	Stationary relative to movement of another participant
Exists independently of the event named by the verb	Does not exist independently of the event, or not at all

Thematic proto-roles and argument selection. David Dowty. Language. 1991.

The Decompositional Approach

Identify properties of interest.

Instigated
Awareness
Physical
…

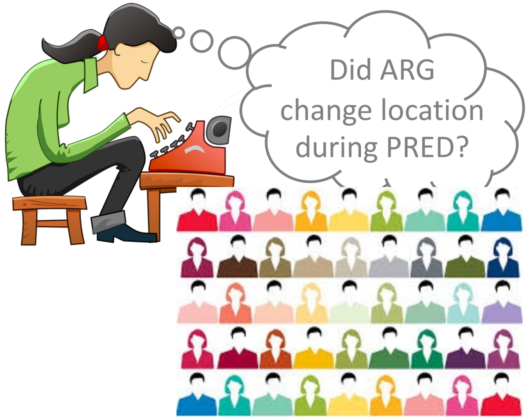
Translate properties into templatic English questions.

Pose each question independently to non-expert annotators.

Extend inventory of properties.

Instigated
Awareness
Physical
Sentient
Moved
Destroyed
…

Make new annotations (but keep the old)!



JOHNS HOPKINS
UNIVERSITY

Semantic Proto-Role Properties

INSTIGATION
VOLITION
AWARENESS
SENTIENT
PHYSICALLY EXISTED
EXISTED BEFORE
EXISTED DURING
EXISTED AFTER

CREATED
DESTROYED
CHANGED
CHANGED STATE
CHANGED POSSESSION
CHANGED LOCATION
CHANGED STATE CONTINUOUS
WAS FOR BENEFIT

STATIONARY
LOCATION
PHYSICAL CONTACT
MANIPULATED
WAS USED
PARTITIVE
...AND MORE?

Semantic Proto-Roles. Reisinger, Rudinger, Ferraro, Harman, Rawlins, and Van Durme. TACL. 2015.

Crowdsourcing Proto-Role Annotations

The antibody then kills the cell.

How likely or unlikely is it that the antibody is aware of being involved in the killing?

very unlikely	somewhat unlikely	not enough information	somewhat likely	very likely
1	2	3	4	5

JOHNS HOPKINS UNIVERSITY

Semantic Proto-Roles. Reisinger, Rudinger, Ferraro, Harman, Rawlins, and Van Durme. TACL. 2015.

Semantic Proto-Roles

Does the property apply to the argument with respect to the underlined event?

5	VOLITION
5	INSTIGATION
4	AWARE
5	PHYSICALLY EXIST
4	CHANGED STATE
1	DESTROYED
1	MANIPULATED
	…

1	VOLITION
1	INSTIGATION
3	AWARE
5	PHYSICALLY EXIST
5	CHANGED STATE
5	DESTROYED
2	MANIPULATED
	…

1	VOLITION
1	INSTIGATION
1	AWARE
5	PHYSICALLY EXIST
2	CHANGED STATE
1	DESTROYED
3	MANIPULATED
	…

5 = very likely
4 = somewhat likely
3 = not enough info.
2 = somewhat unlikely
1 = very unlikely

The cat **<u>ate</u>** the rat (with its sharp teeth).

JOHNS HOPKINS UNIVERSITY

Semantic Proto-Roles. Reisinger, Rudinger, Ferraro, Harman, Rawlins, and Van Durme. TACL. 2015.

Semantic Proto-Roles

Does the property apply to the argument with respect to the underlined event?

+ VOLITION	- VOLITION	- VOLITION
+ INSTIGATION	- INSTIGATION	- INSTIGATION
+ AWARE	- AWARE	- AWARE
+ PHYSICALLY EXIST	+ PHYSICALLY EXIST	+ PHYSICALLY EXIST
- CHANGED STATE	+ CHANGED STATE	+ CHANGED STATE
- DESTROYED	+ DESTROYED	- DESTROYED
- MANIPULATED	- MANIPULATED	+ MANIPULATED
…	…	…

4 or 5 → +

1, 2, or 3 → -

The cat <u>**ate**</u> the rat (with its sharp teeth).

Semantic Proto-Roles. Reisinger, Rudinger, Ferraro, Harman, Rawlins, and Van Durme. TACL. 2015.

Task: Semantic Proto-Role Labeling (SPRL)

A multi-label task.

Input (X): A sentence; a predicate-argument pair in the sentence.

Output (Y): A score for each SPR property. (Binary or Scalar 1-5)

Y:

5	VOLITION
5	INSTIGATION
4	AWARE
5	PHYSICALLY EXIST
4	CHANGED STATE
1	DESTROYED
1	MANIPULATED
	…

X: The cat **ate** the rat (with its sharp teeth).

Task: Semantic Proto-Role Labeling (SPRL)

A multi-label task.

Input (X): A sentence; a predicate-argument pair in the sentence.

Output (Y): A score for each SPR property. (Binary or Scalar 1-5)

Y:

1 VOLITION
1 INSTIGATION
3 AWARE
5 PHYSICALLY EXIST
5 CHANGED STATE
5 DESTROYED
2 MANIPULATED
...

X: The cat **ate** the rat (with its sharp teeth).

Task: Semantic Proto-Role Labeling (SPRL)

A multi-label task.

Input (X): A sentence; a predicate-argument pair in the sentence.

Output (Y): A score for each SPR property. (Binary or Scalar 1-5)

Y:

1	VOLITION
1	INSTIGATION
1	AWARE
5	PHYSICALLY EXIST
2	CHANGED STATE
1	DESTROYED
3	MANIPULATED
	…

X: The cat **ate** the rat (with its sharp teeth).

Dataset 1: Semantic Proto-Roles

Dataset 2: Event Factuality

Dataset 3: Temporal Relations

Dataset 4: Genericity

What is event factuality?

Did the event mentioned in text happen or not?

Example: Did the <u>watering</u> event happen?

HAPPENED!

Pat $\boxed{\text{watered}}$ the plants.

DIDN'T HAPPEN!

Pat did not $\boxed{\text{water}}$ the plants.

Why is event factuality a hard problem?

Event factuality can be influenced by words from diverse syntactic and semantic categories.

negation | adverbs | quantifiers | modal auxiliaries | clause-embedding verbs | nouns

HAPPENED!
Pat watered the plants.

DIDN'T HAPPEN!
Pat almost watered the plants.

UNCERTAIN?
Pat might have watered the plants.

HAPPENED!
Pat managed to water the plants.

DIDN'T HAPPEN!
Pat did not water the plants.

DIDN'T HAPPEN!
Pat watered none of the plants.

DIDN'T HAPPEN!
Pat failed to water the plants.

DIDN'T HAPPEN!
Pat's watering the plants was a hallucination.

Collecting Data

New Dataset: It Happened (UDS-IH2)

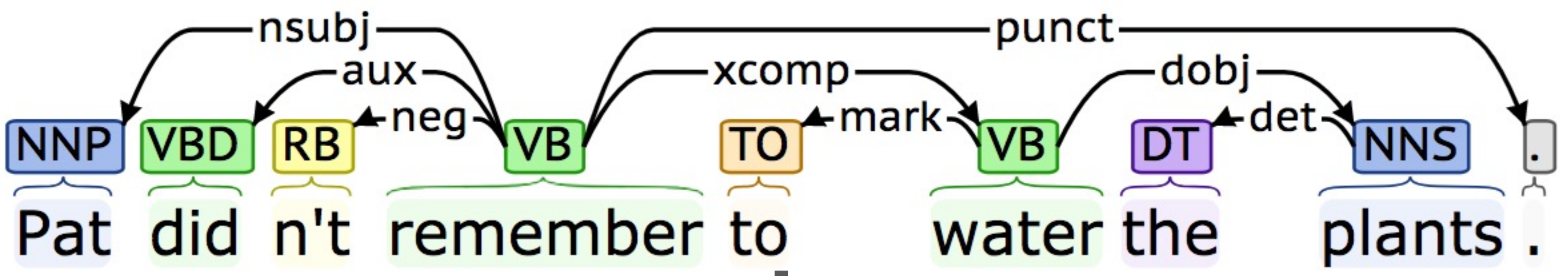
- Largest English factuality dataset to date
  - 27,289 predicates extracted with PredPatt White et al. 2016
- Covers all of Universal Dependencies English Web Treebank v1.2 (extends White et al. 2016)
  - User-generated text: **weblogs, reviews, question-answers, newsgroups, email**
  - ~17K sentences
  - Gold syntactic dependency parses (Universal Dependencies)

Linguistic Data Consortium

https://catalog.ldc.upenn.edu/LDC2012T13

https://github.com/UniversalDependencies/UD_English-EWT

Event Identification



Rule-based Predicate-Argument Extraction from Syntactic Dependencies (PredPatt)

Pat didn’t remember to water the plants.

https://github.com/hltcoe/PredPatt

Collecting "It Happened" Dataset (UDS-IH2)

Al - Zaman : American forces killed Shaikh Abdullah al - Ani , the preacher at the mosque in the town of Qaim , near the Syrian border .

The sentence [-------] understandable, and killed [-------] refer to a predicate.

Collecting "It Happened" Dataset (UDS-IH2)

Al - Zaman : American forces killed Shaikh Abdullah al - Ani , the preacher at the mosque in the town of Qaim , near the Syrian border .

The sentence is understandable, and killed does refer to a predicate.

According to the author, the situation referred to by killed -------, and you are ------- about that.

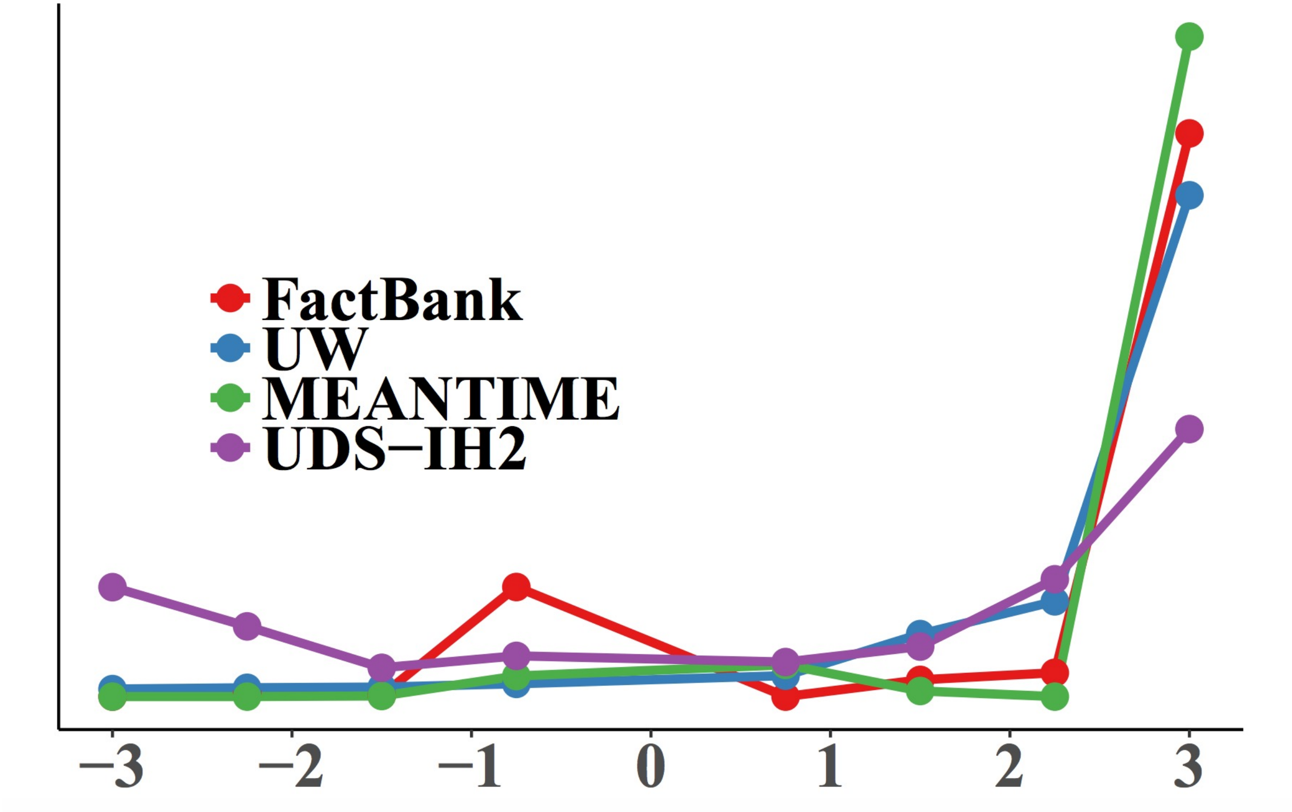
Collecting “It Happened” Dataset (UDS-IH2)

Al - Zaman : American forces killed Shaikh Abdullah al - Ani , the preacher at the mosque in the town of Qaim , near the Syrian border .

The sentence is understandable, and killed does refer to a predicate.

According to the author, the situation referred to by killed had happened or was happening, and you are totally confident about that.

Relative Frequency of Factuality Labels



It-Happened shows more entropy in the distribution of labels

Higher entropy likely due to better genre distribution: **weblogs, reviews, newsgroups, emails**

Examples from UDS-IH2

DIDN'T HAPPEN! DIDN'T HAPPEN!

Give me a call Tuesday afternoon to discuss
(gone to Kelowna golfing for the weekend)

HAPPENED! HAPPENED!

Examples from UDS-IH2

I <3 Max's

Dataset 1: Semantic Proto-Roles

Dataset 2: Event Factuality

Dataset 3: Temporal Relations

Dataset 4: Genericity

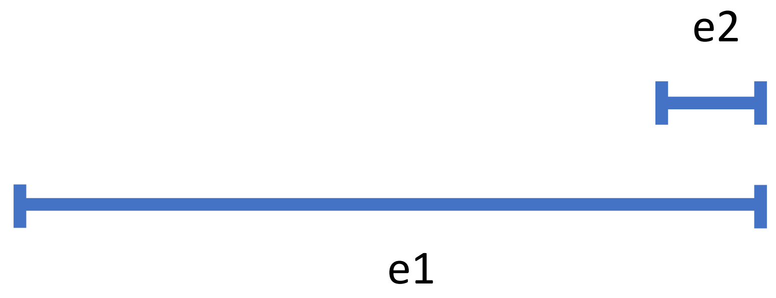
Temporal Interpretation of Events in Text

We were <u>looking over the menu [e1]</u> when Jo <u>knocked her water over [e2]</u>.

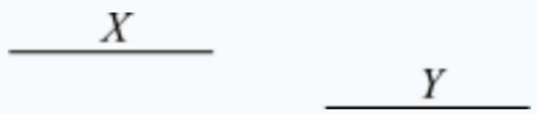
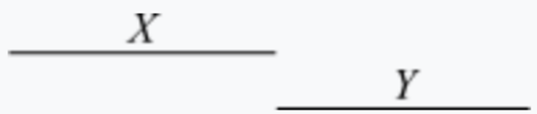
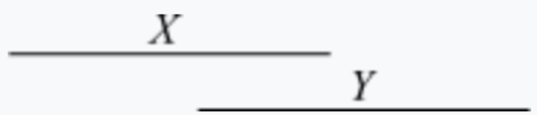
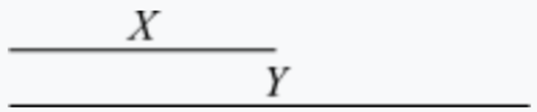
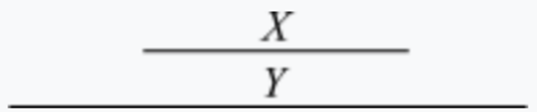
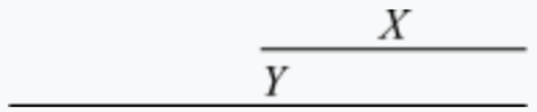
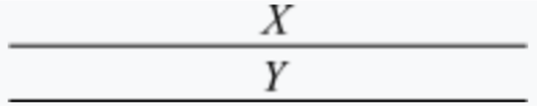
What order do events e1 and e2 happen in? (e1 < e2)

How long does each event last? (e1 minutes; e2 seconds)

Can we construct a timeline of the events?



Categorical Temporal Relations

Relation	Illustration	Interpretation
$X < Y$ $Y > X$		X takes place before Y
$X \mathbf{m} Y$ $Y \mathbf{mi} X$		X meets Y (*i* stands for *inverse*)
$X \mathbf{o} Y$ $Y \mathbf{oi} X$		X overlaps with Y
$X \mathbf{s} Y$ $Y \mathbf{si} X$		X starts Y
$X \mathbf{d} Y$ $Y \mathbf{di} X$		X during Y
$X \mathbf{f} Y$ $Y \mathbf{fi} X$		X finishes Y
$X = Y$		X is equal to Y

…but what about duration?

Allen, James F. "Towards a general theory of action and time." *Artificial intelligence* 23.2 (1984): 123-154.

Approach

Capture absolute and relative duration

UDS-T

- Dataset: Universal Decompositional Semantics – Time (UDS-T)
- Covers English Web Treebank
- # Events: 32,302
- # Event-Event Relations: 70,368

Vashishtha, S., B. Van Durme, & A.S. White. 2019. Fine-Grained Temporal Relation Extraction. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (ACL 2019), Florence, Italy, July 29-31, 2019.

http://decomp.io/projects/time/

What to [1] feed my dog after gastroenteritis ? My dog has [2] been [2] sick [2] for about 3 days [2] now .

[1]feed

Range: 49 - 66

The situation lasted for hours and you are totally confident about that.

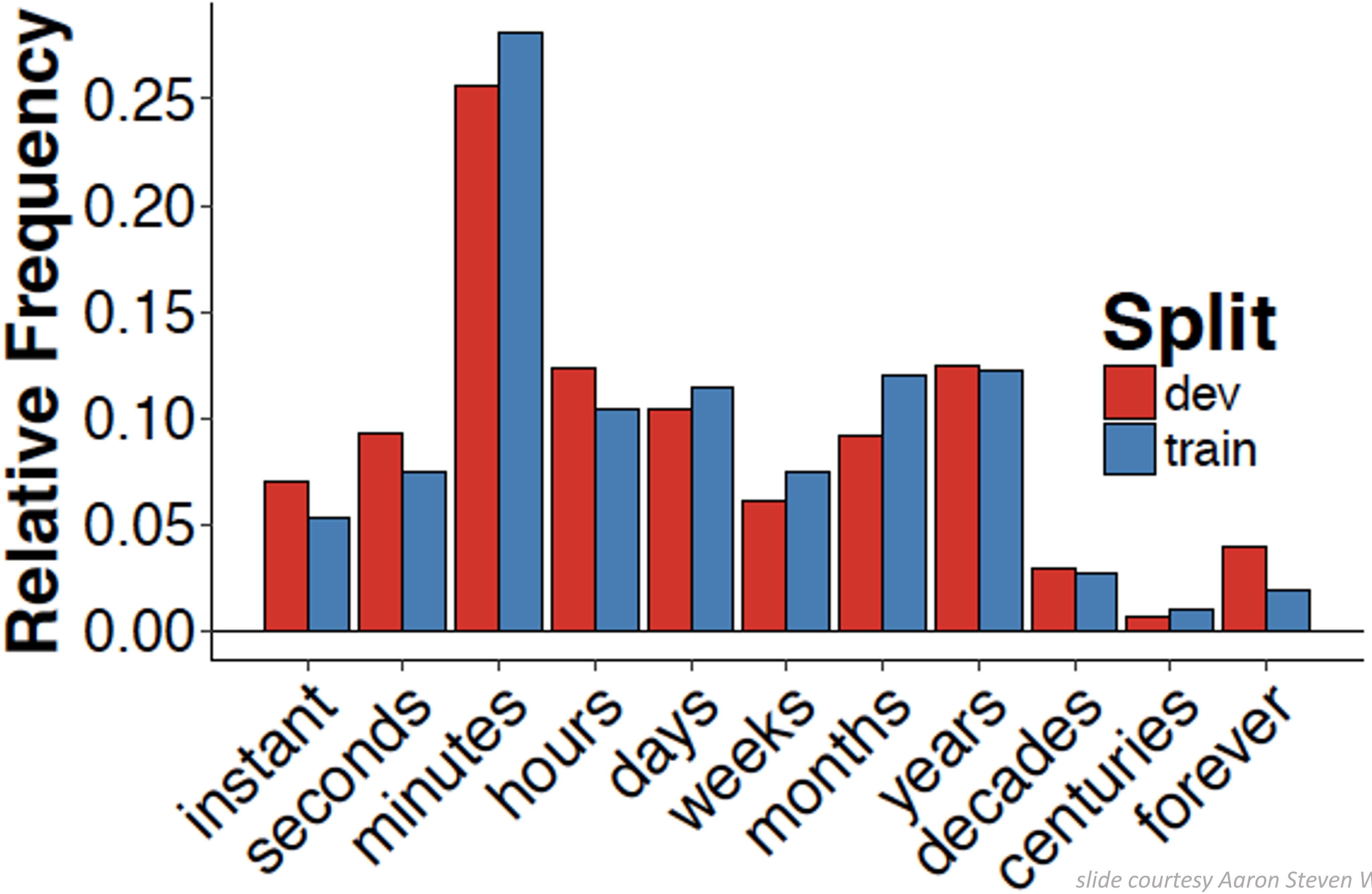
[2]been sick for now

Range: 12 - 49

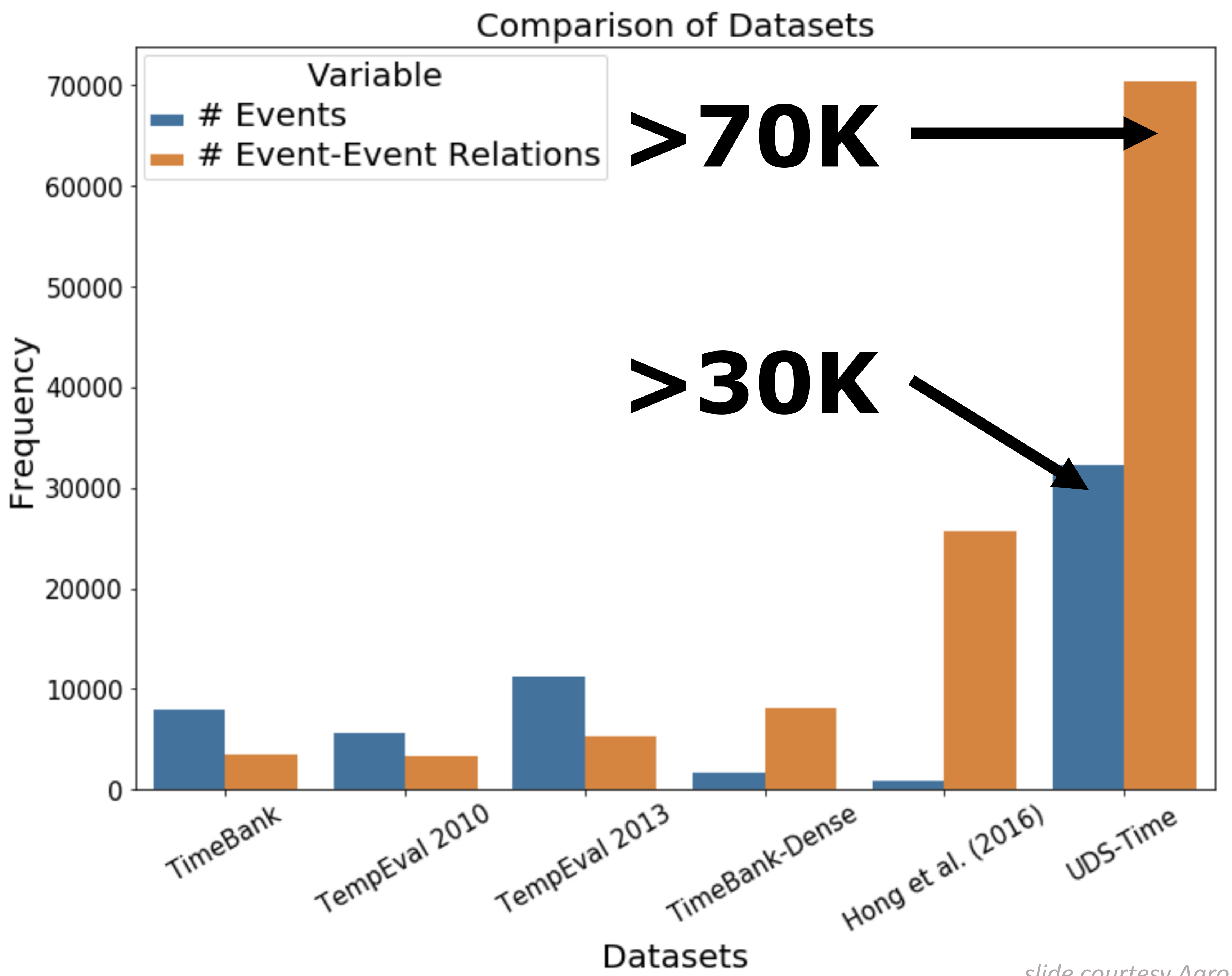
The situation lasted for days and you are totally confident about that.

You are totally confident about the chronology you provided.

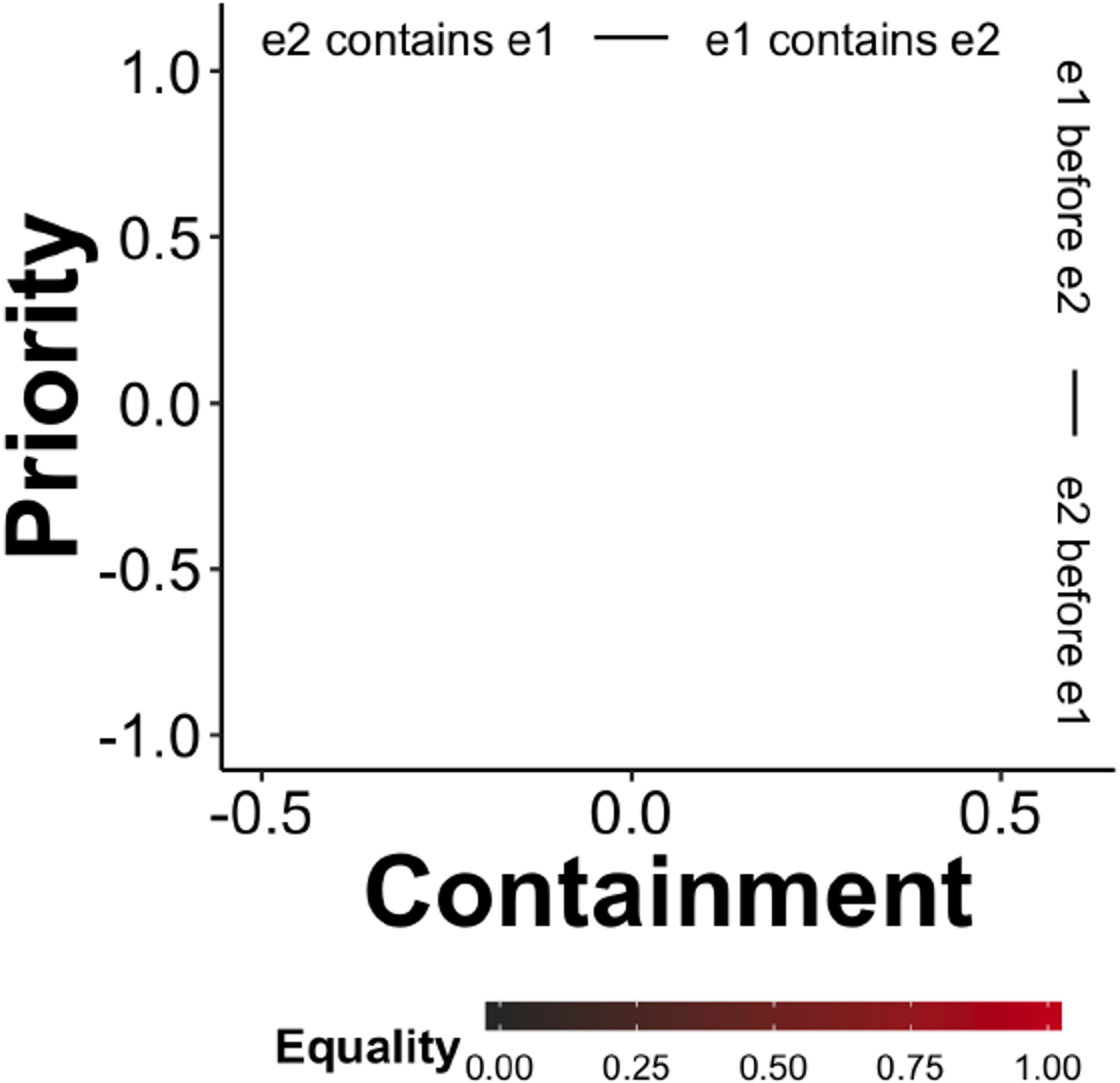
slide courtesy Aaron Steven White, 2019



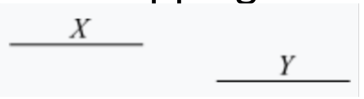
slide courtesy Aaron Steven White, 2019



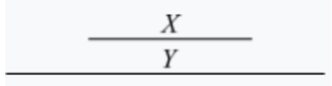
slide courtesy Aaron Steven White, 2019



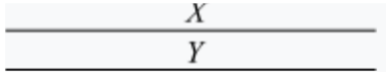
Priority: Positive if e1 come strictly before e2; negative if vice-versa; close to zero if overlapping.

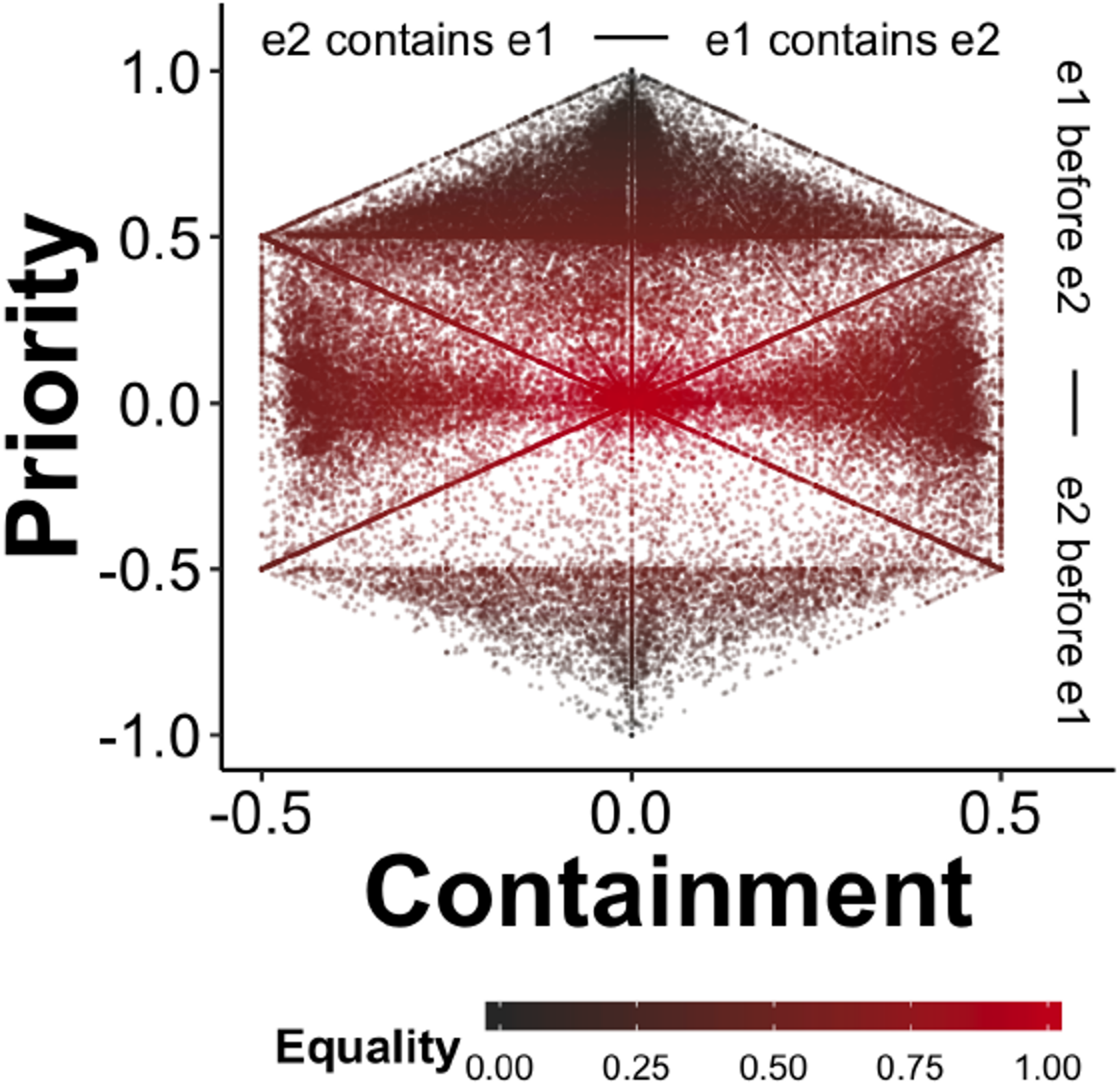


Containment: Positive if e1 contains e2 (i.e. e2 happens entirely during e1); negative if e2 contains e1; close to zero if neither contains the other.

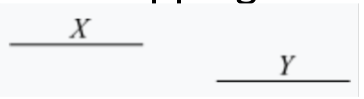


Equality: Do e1 and e2 occur at the same time and duration; i.e. do e1 and e2 contain each other.

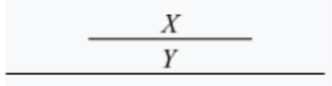




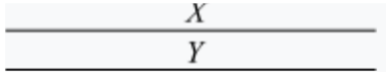
Priority: Positive if e1 come strictly before e2; negative if vice-versa; close to zero if overlapping.



Containment: Positive if e1 contains e2 (i.e. e2 happens entirely during e1); negative if e2 contains e1; close to zero if neither contains the other.

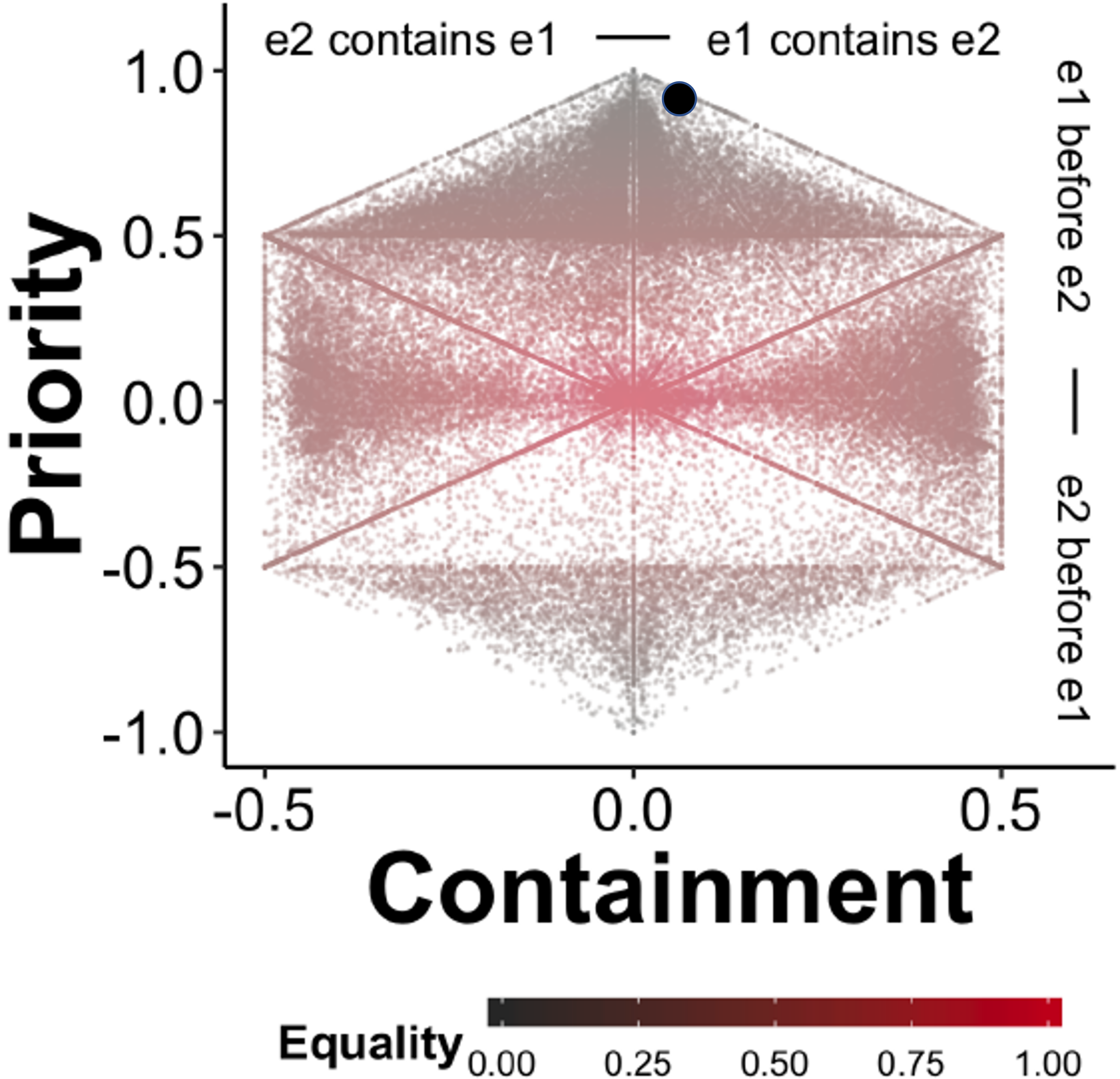


Equality: Do e1 and e2 occur at the same time and duration; i.e. do e1 and e2 contain each other.

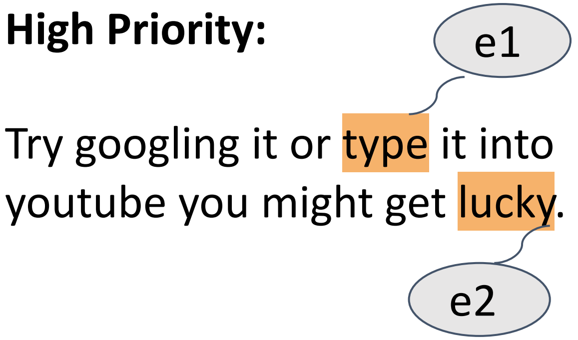


Note 1: the triangle at top and bottom because extreme priority precludes overlap/containment.

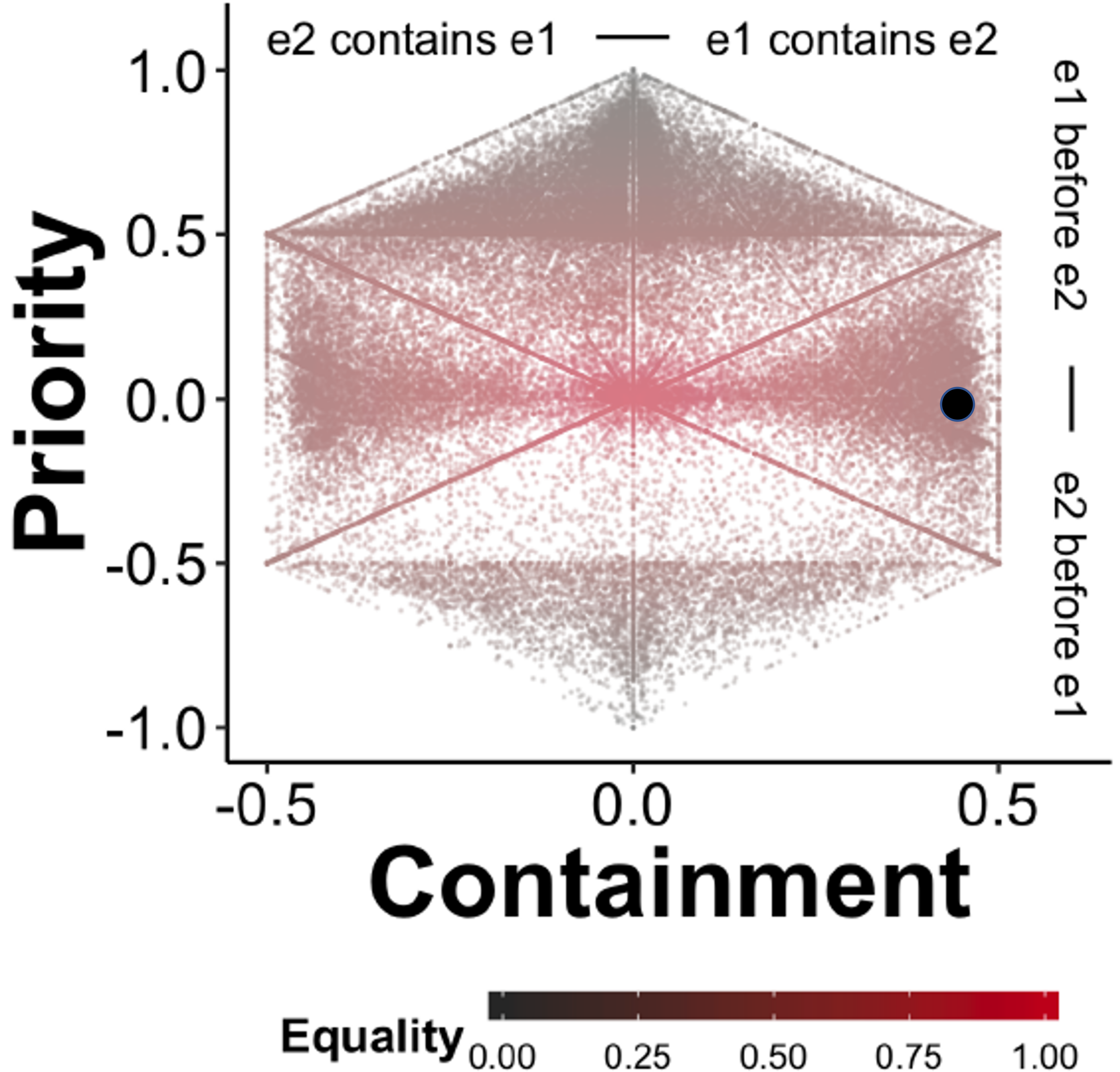
Note 2: center is red because high equality means low priority (neither comes before the other).



High Priority:



slide courtesy Aaron Steven White, 2019



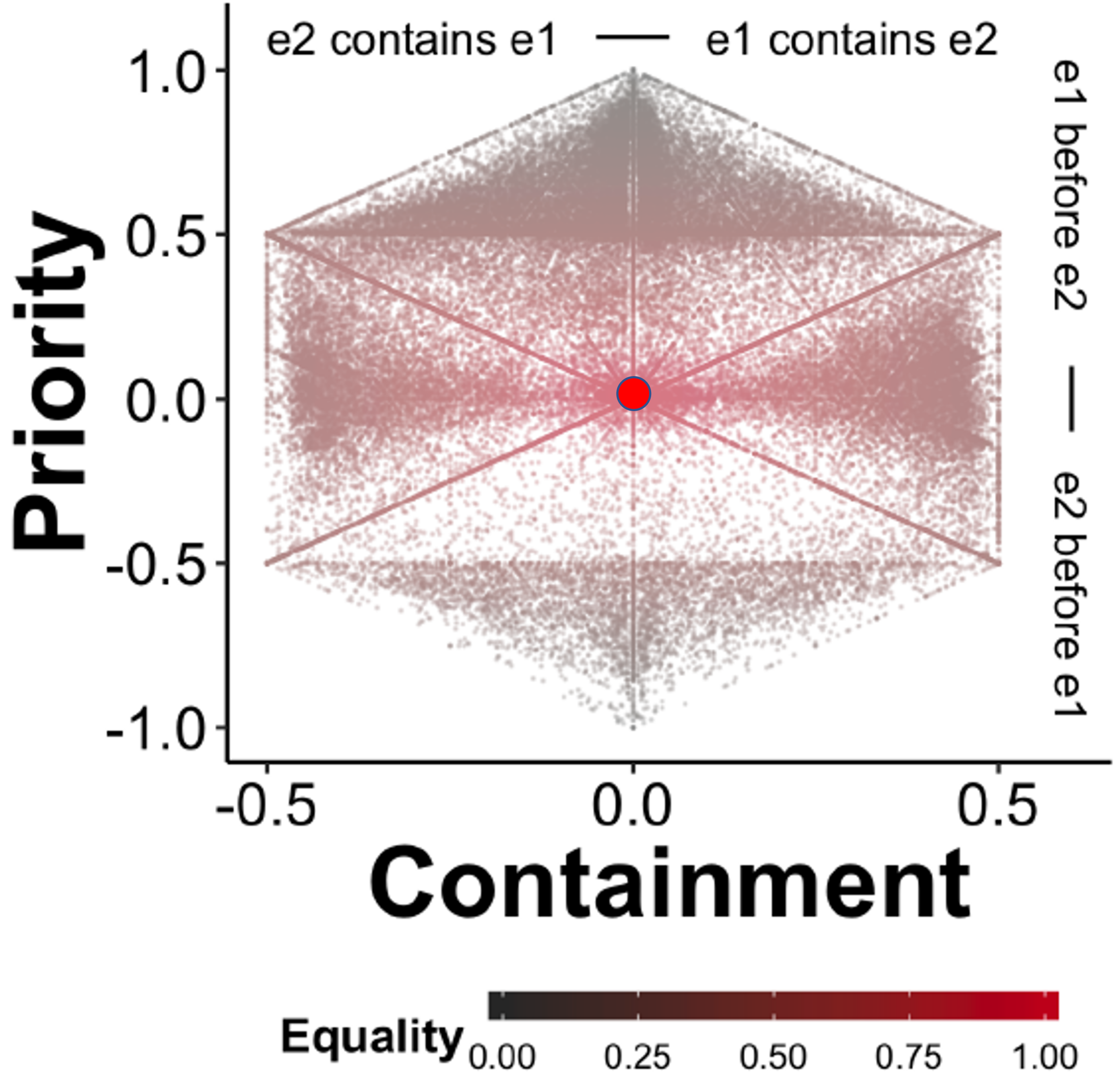
High Containment:

Both Tina and Vicky are excellent. I will definitely refer my friends and family.

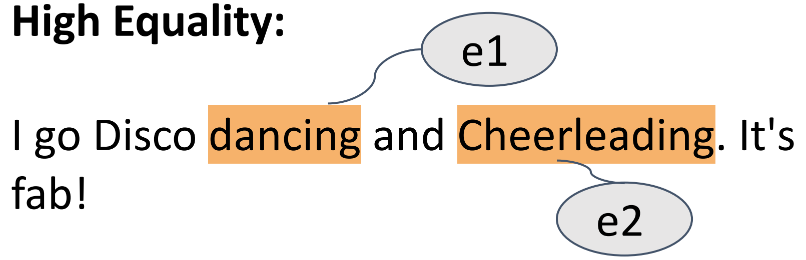
e1

e2

slide courtesy Aaron Steven White, 2019



High Equality:



slide courtesy Aaron Steven White, 2019

Dataset 1: Semantic Proto-Roles

Dataset 2: Event Factuality

Dataset 3: Temporal Relations

Dataset 4: Genericity

Linguistic Generalization: NPs/Entities

Individuals vs. Kinds

Ind Ind

Pat ate a wedge of cheese.

Ind Knd

Pat loves cheese.

Ind Knd? Ind?

My grocer carries three cheeses.

Knd? Ind? Knd? Ind?

Trader Joe's carries twelve cheeses.

Linguistic Generalization: Clauses/Events

Episodics

Mary ate oatmeal for breakfast today.

Pat carried the basket of eggs into the house.

Events that are spatio-temporally bounded.

Habituals

Mary eats oatmeal for breakfast.

Pat's chicken lays green eggs.

Recurring event with individual participant.

Generics

Oatmeal grows in temperate climates.

Chickens lay eggs.

Generic event AND generic participant.

A Decompositional Approach to Genericity

“In our framework, prototypical episodics, habituals, and generics correspond to **sets of properties** that the referents of a clause’s head predicate and arguments have—namely, clausal **categories are built up from properties of the predicates that head them along with those predicates’ arguments**.”

Govindarajan et al., 2019

A Decompositional Approach to Genericity

- Discard mutually exclusive categories (e.g. EPISODIC/HABITUAL/GENERIC)
- Independently annotate for 3 Properties for Arguments/Participants
  - Particular
  - Kind
  - Abstract
- Independently annotate for 3 Properties for Predicates/Events
  - Particular
  - Dynamic
  - Hypothetical

Govindarajan et al., 2019

Each property:

- Independent binary choice [*does/doesn't*]
- 5-point confidence scale
  - 5: totally confident
  - 4: very confident
  - 3: somewhat confident
  - 2: not very confident
  - 1: not at all confident

Govindarajan et al., 2019

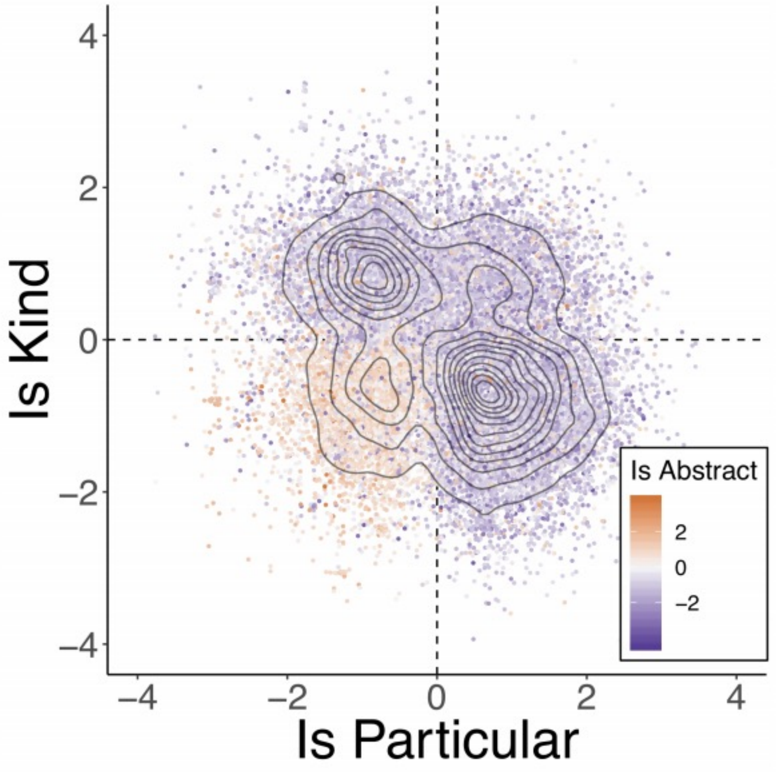
UDS-G Dataset

- Universal Decompositional Semantics -- Genericity
- Covers entire English Web Treebank (Universal Dependencies)
- Size
  - Args: 37,146
  - Pred: 33,114

Govindarajan et al., 2019

Label Distributions

ARGUMENT

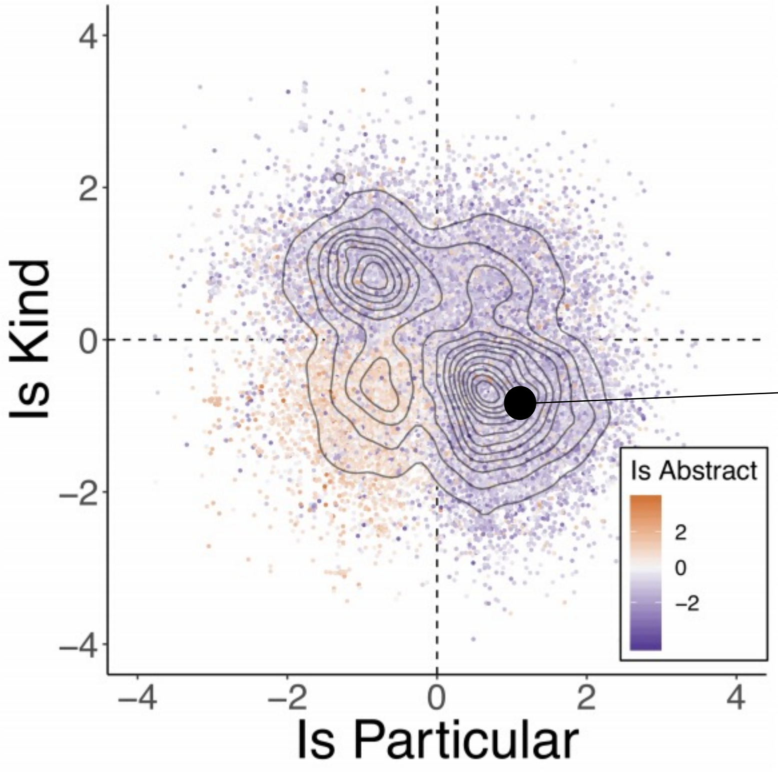


- Kind and Particular are negatively correlated (pearson correlation = -0.33)

Govindarajan et al., 2019

Label Distributions

ARGUMENT



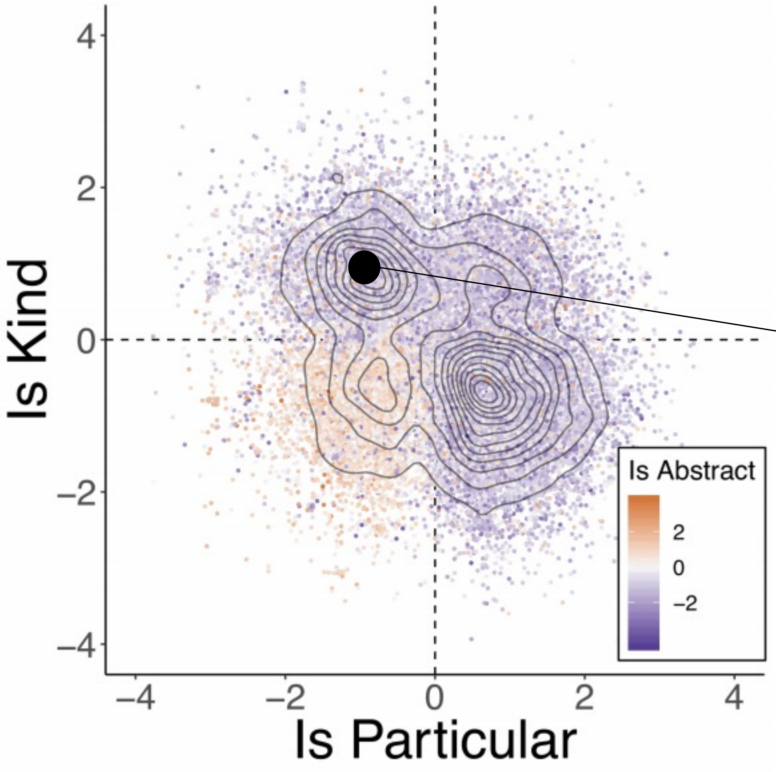
- Kind and Particular are negatively correlated (pearson correlation = -0.33)

"I think this <u>place</u> is probably really great especially judging by the reviews on here." [particular, not kind]

Govindarajan et al., 2019

Label Distributions

ARGUMENT



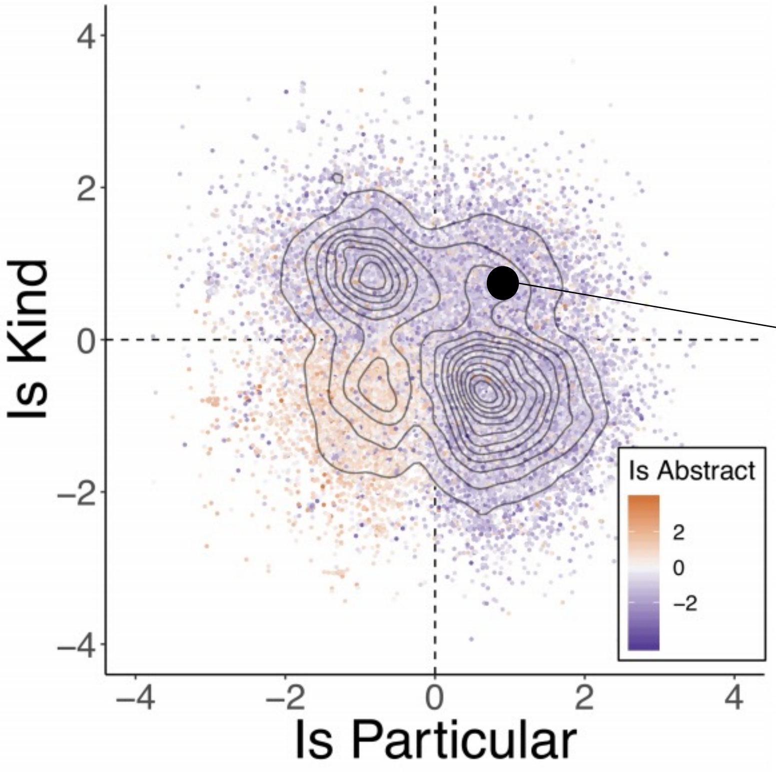
- Kind and Particular are negatively correlated (pearson correlation = -0.33)

“What made it perfect was that they only offered <u>transportation</u> so that…”
[kind, not particular]

Govindarajan et al., 2019

Label Distributions

ARGUMENT



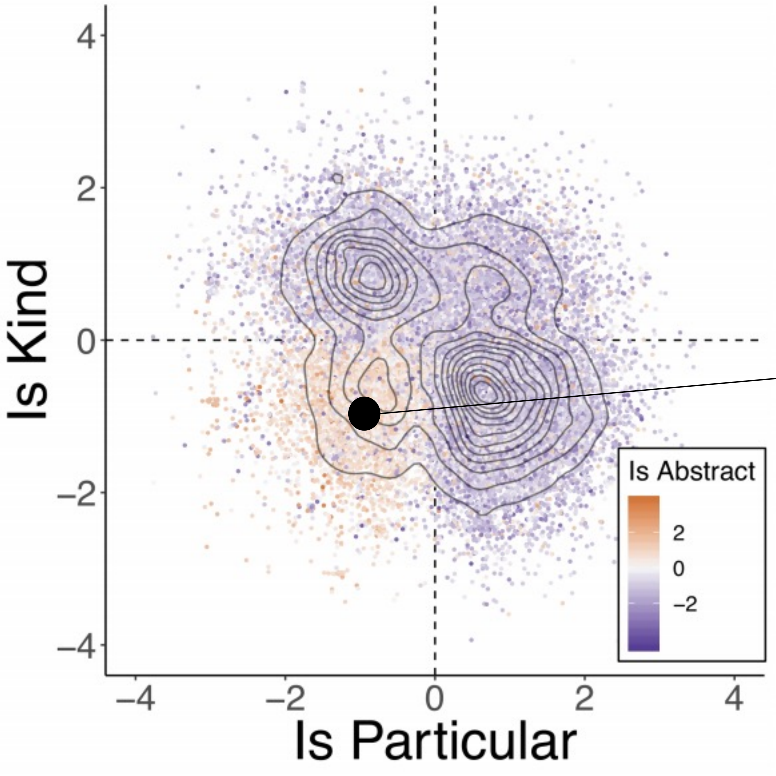
- Kind and Particular are negatively correlated (pearson correlation = -0.33)

"Some places do the registration right at the hospital…"
[kind, particular]

Govindarajan et al., 2019

Label Distributions

ARGUMENT



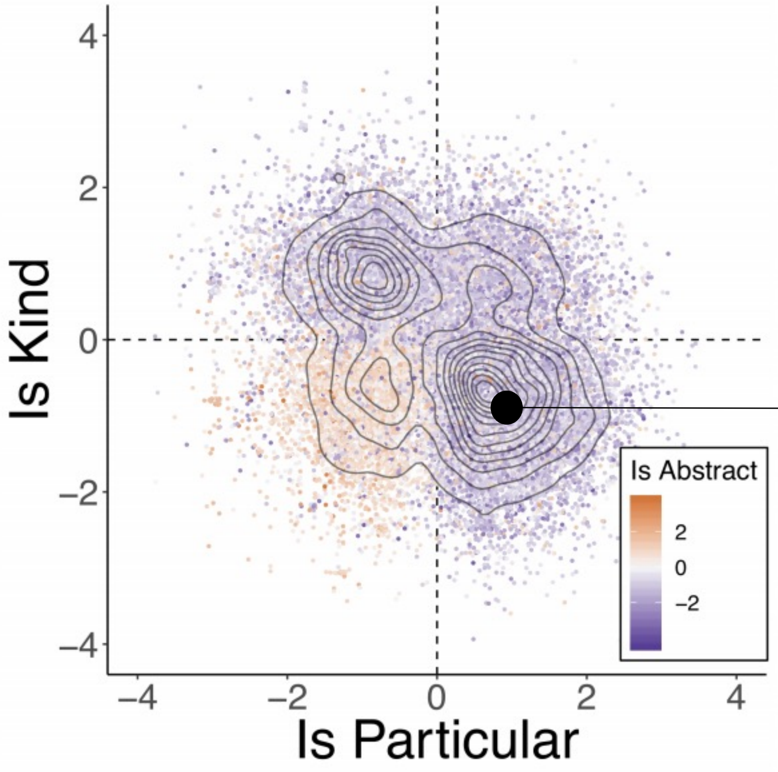
- Abstract is negatively correlated with both Particular (corr = -0.28) and Kind (corr = -0.11)

“Power be where power lies.”
[abstract, not kind, not particular]

Govindarajan et al., 2019

Label Distributions

ARGUMENT



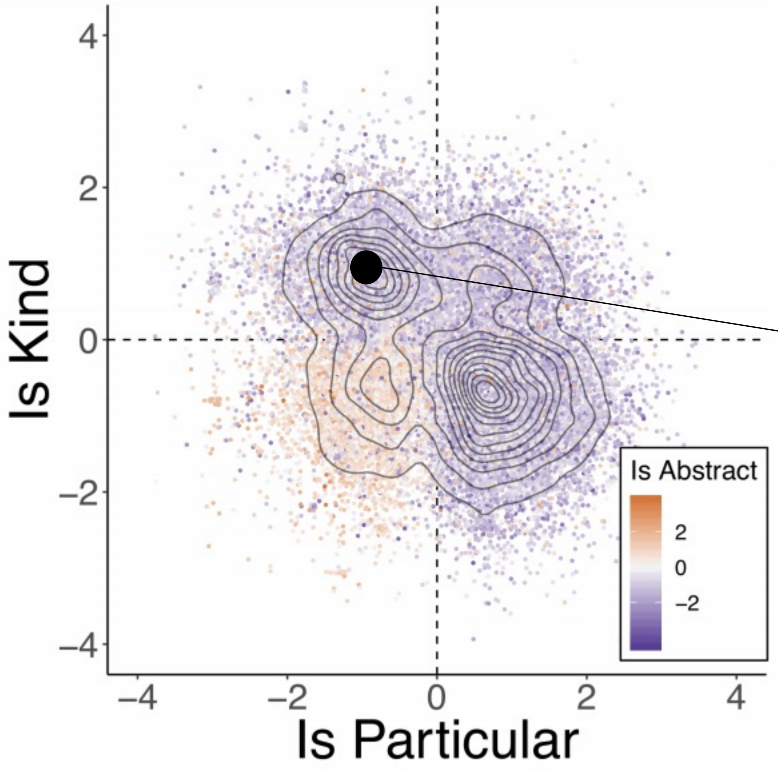
- Abstract is negatively correlated with both Particular (corr = -0.28) and Kind (corr = -0.11)

"Meanwhile, his reputation seems to be improving, although Bangs noted a 'pretty interesting social dynamic.'"
[abstract, particular, not kind]

Govindarajan et al., 2019

Label Distributions

ARGUMENT



- Abstract is negatively correlated with both Particular (corr = -0.28) and Kind (corr = -0.11)

"The Pew researchers tried to transcend the economic <u>argument</u>."
[abstract, kind, not particular]

Govindarajan et al., 2019

Predictive Models

	Feature sets				Is.Particular		Is.Kind		Is.Abstract		All
	Type	Token	GloVe	ELMO	ρ	R1	ρ	R1	ρ	R1	wR1
ARGUMENT	+	-	-	-	42.4	7.4	30.2	4.9	51.4	11.7	8.1
	-	+	-	-	50.6	13.0	41.5	8.8	33.8	4.8	8.7
	-	-	+	-	44.5	8.3	33.4	4.6	45.2	7.7	6.9
	-	-	-	+	57.5	17.0	48.1	13.3	55.7	14.9	15.1
	+	+	-	-	55.3	14.1	46.2	11.6	52.6	13.0	12.9
	-	+	-	+	**58.6**	15.6	48.6	**13.7**	**56.8**	14.2	14.5
	+	+	-	+	58.3	16.3	47.8	13.2	56.3	**15.2**	14.9
	+	+	+	+	58.1	**17.0**	**48.9**	13.2	56.1	15.1	**15.1**
					Is.Particular		**Is.Hypothetical**		**Is.Dynamic**		
PREDICATE	+	-	-	-	14.0	0.8	13.4	0.0	32.5	5.6	2.0
	-	+	-	-	22.3	2.8	37.7	7.3	31.7	5.1	5.1
	-	-	+	-	20.6	2.2	23.4	2.4	29.7	4.6	3.0
	-	-	-	+	26.2	3.6	43.1	10.0	37.0	6.8	6.8
	-	-	+	+	26.8	4.0	42.8	8.9	37.3	7.3	6.7
	+	+	-	-	24.0	3.3	37.9	7.6	37.1	7.6	6.1
	-	+	-	+	**27.4**	4.1	43.3	10.1	**38.6**	**7.8**	**7.4**
	+	-	-	+	27.1	4.0	43.0	10.1	37.5	7.6	7.2
	+	+	+	+	26.8	**4.1**	**43.5**	**10.3**	37.1	7.2	7.2

Best models so far use combination of ELMo and hand-engineered lexical features.

Govindarajan et al., 2019

Some practical stuff…

The Decomp Toolkit

Decomp Toolkit

- Access labels from all UDS datasets (e.g. 4 datasets described above)
- Navigate predicate-argument graph structure, decorated with semantic attributes
- Aligned with Universal Dependencies syntax
- https://github.com/decompositional-semantics-initiative/decomp

Selected Citations

Reisinger, D., Rudinger, R., Ferraro, F., Harman, C., Rawlins, K., & Van Durme, B. (2015). Semantic proto-roles. *Transactions of the Association for Computational Linguistics*, *3*, 475-488.

Teichert, A., Poliak, A., Van Durme, B., & Gormley, M. R. (2017, February). Semantic proto-role labeling. In *Thirty-First AAAI Conference on Artificial Intelligence*.

Rudinger, R., Teichert, A., Culkin, R., Zhang, S., & Van Durme, B. (2018). Neural-Davidsonian Semantic Proto-role Labeling. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 944-955).

Opitz, J., & Frank, A. (2019, June). An Argument-Marker Model for Syntax-Agnostic Proto-Role Labeling. In *Proceedings of the Eighth Joint Conference on Lexical and Computational Semantics (* SEM 2019)* (pp. 224-234).

White, A. S., Reisinger, D., Sakaguchi, K., Vieira, T., Zhang, S., Rudinger, R., ... & Van Durme, B. (2016, November). Universal decompositional semantics on universal dependencies. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing* (pp. 1713-1723).

Rudinger, R., White, A. S., & Van Durme, B. (2018). Neural models of factuality. In *Proceedings of NAACL-HLT* (pp. 731-744).

White, A. S., Rudinger, R., Rawlins, K., & Van Durme, B. (2018, January). Lexicosyntactic Inference in Neural Models. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*.

Jiang, N., & de Marneffe, M. C. (2019, July). Do you know that Florence is packed with visitors? Evaluating state-of-the-art models of speaker commitment. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 4208-4213).

Vashishtha, S., Van Durme, B., & White, A. S. (2019, July). Fine-Grained Temporal Relation Extraction. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 2906-2919).

Govindarajan, V., Durme, B. V., & White, A. S. (2019). Decomposing Generalization: Models of Generic, Habitual, and Episodic Statements. *Transactions of the Association for Computational Linguistics*, *7*, 501-517.

Stengel-Eskin, E., White, A. S., Zhang, S., & Van Durme, B. (2019). Transductive Parsing for Universal Decompositional Semantics. *arXiv preprint arXiv:1910.10138*.

Poliak, A., Haldar, A., Rudinger, R., Hu, J. E., Pavlick, E., White, A. S., & Van Durme, B. (2018). Collecting Diverse Natural Language Inference Problems for Sentence Representation Evaluation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 67-81).

Find pointers to everything at decomp.io