



# Understanding Event Processes in Natural Language

## Event-Centric Natural Language Understanding (Part III)

Muhao Chen

Viterbi School of Engineering

University of Southern California

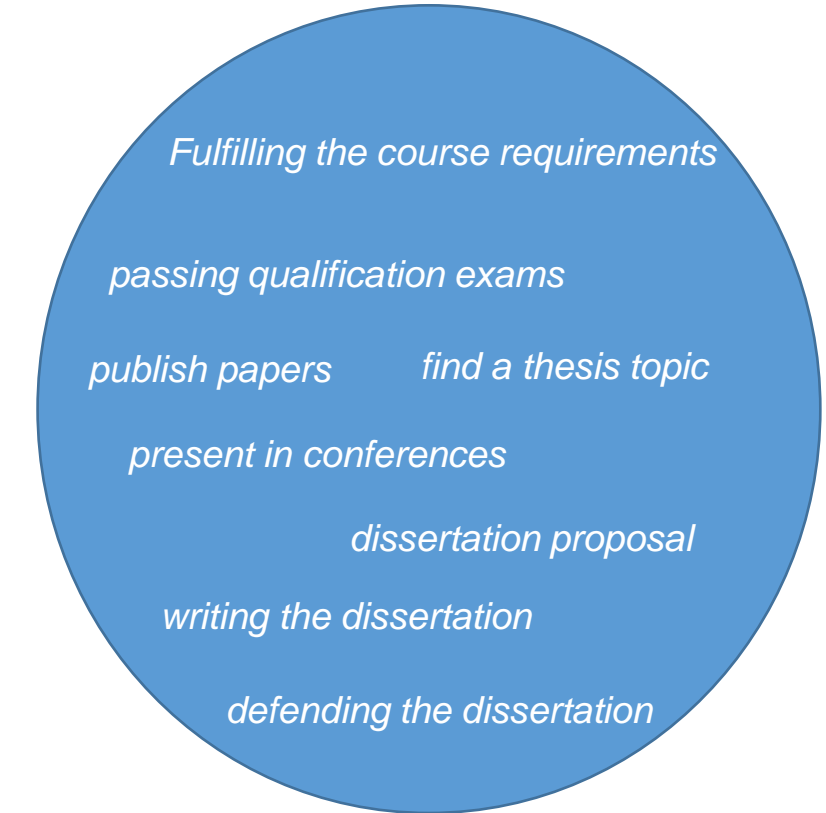
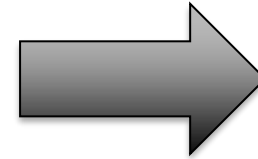
**Feb 2021**

**AAAI Tutorials**

**Event-Centric Natural Language Understanding**

How Do Machines Understand the *Evolution* of Events?

Earning a PhD in Computer Science typically takes around 5 years. It first involves ***fulfilling the course requirements*** and ***passing qualification exams***. Then within several years, the student is expected to ***find a thesis topic***, ***publish several papers*** about the topic and ***present them in conferences***. The last one or two years are often about ***completing the dissertation proposal***, ***writing*** and ***defending the dissertation***.



## Event-centric IE has helped machines in

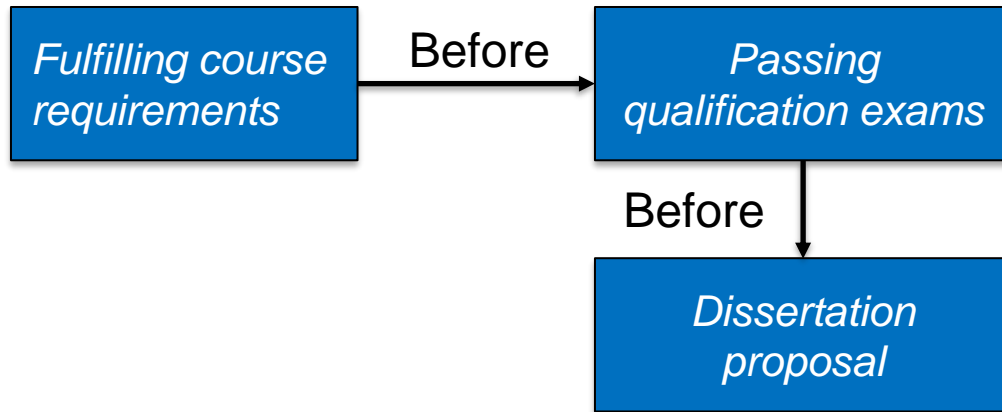
- Recognition and typing events
- Inducing the relations between two events
- Identifying the equivalence of events

# Event Process Understanding And Prediction

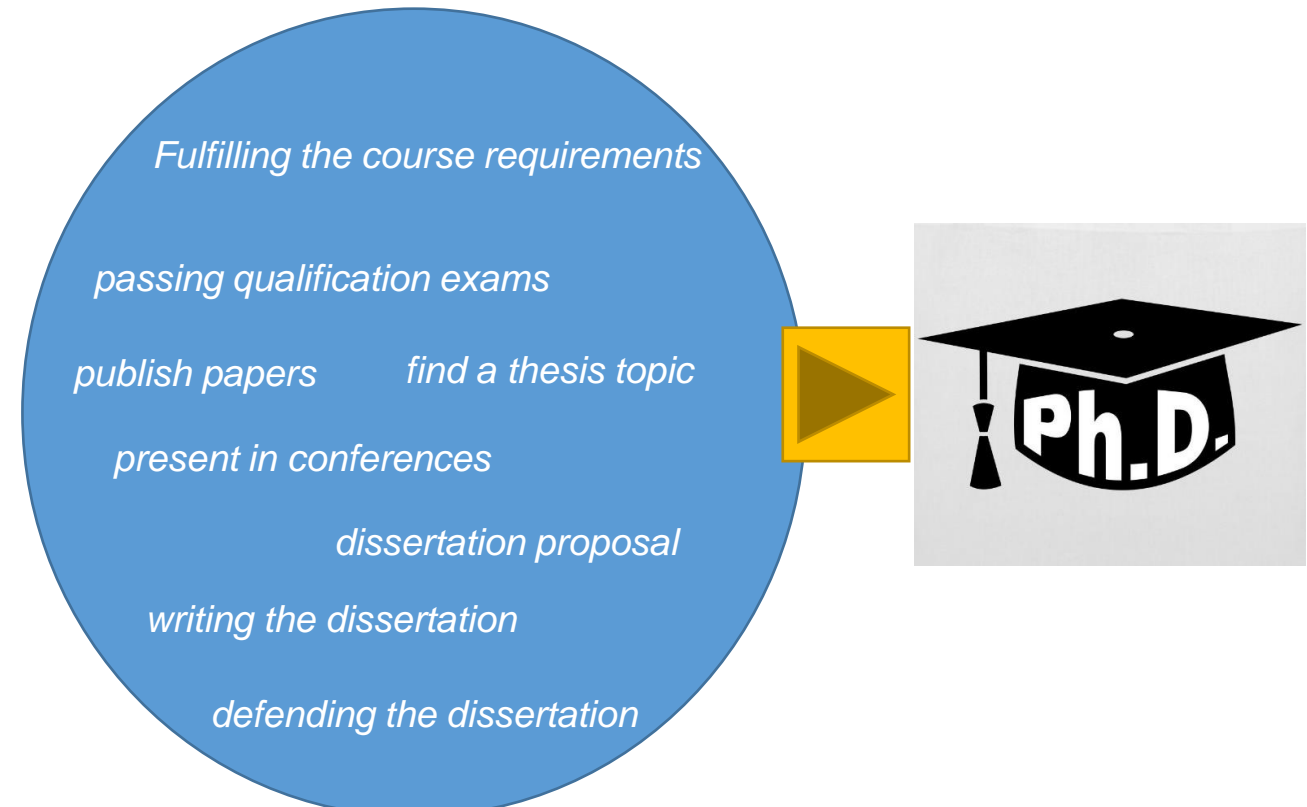
Extraction only is not enough.

Events are **NOT simple, static** predicates.

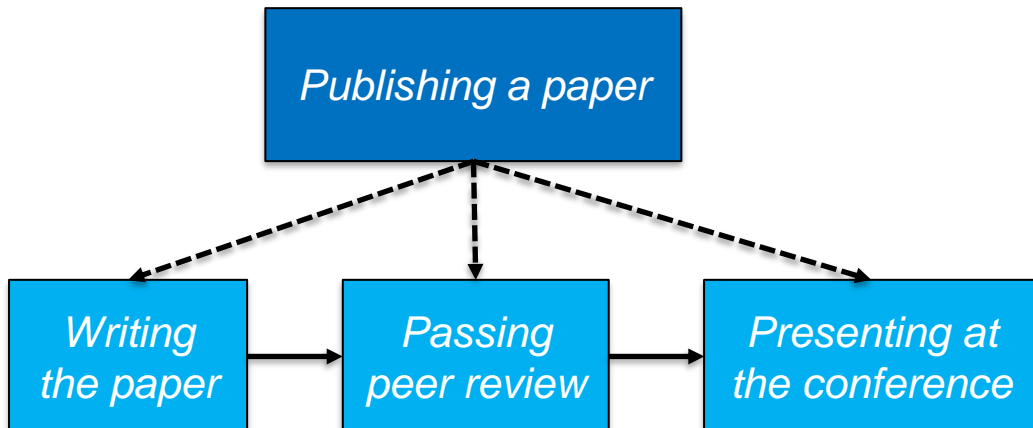
They evolve,



and are always directed by specific intents or central goals [Zacks et al. *Nature Neuroscience*, 2001]

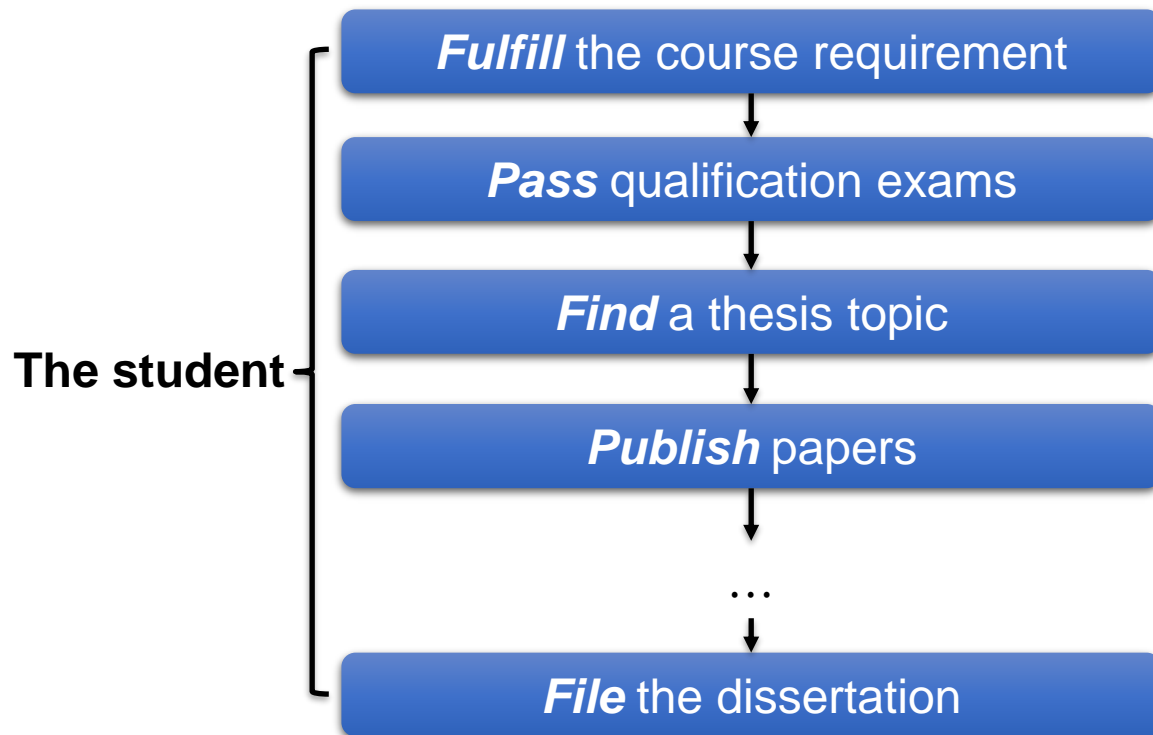


are described in different granularities,



## An event process (or event chain)

- Partially ordered events that are centered around common protagonists [Chambers et al., ACL-08]



## Prediction problems on event processes

### Event process completion

- What happens next?

### Intention prediction

- What is the goal of “*digging a hole, putting some seeds in the hole and filling it with soil*”?

### Membership prediction

- What are the steps of “*buying a car*”?

### Salience prediction

- Is *defending the dissertation* more important than *doing an internship*?

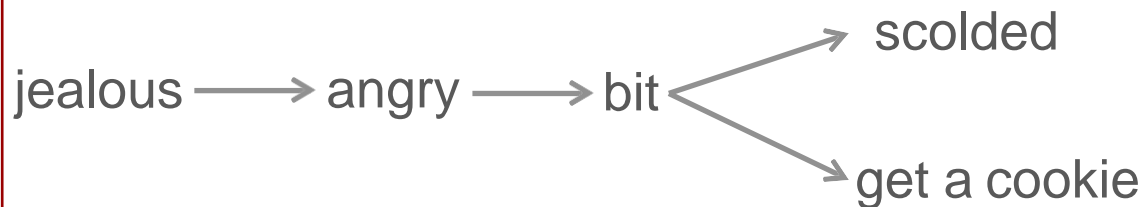
## Narrative prediction

One day Wesley's auntie came over to visit. He was happy to see her, because he liked to play with her. When she started to give his little sister attention, he got **jealous**. He got **angry** at his auntie and **bit** his sister's hand when she wasn't looking.

Then what might happen?

**O1:** He was **scolded**. ✓

**O2:** She **gave him a cookie** for being so nice. ✗



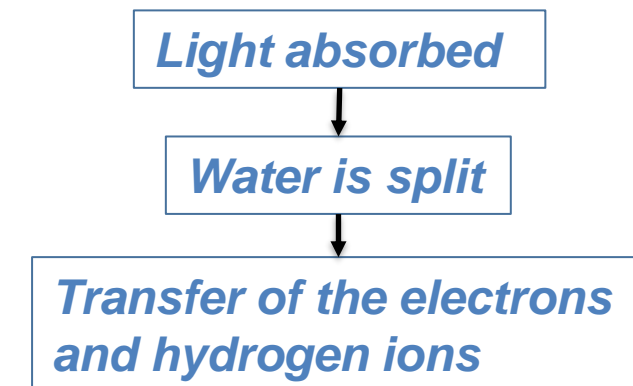
## Machine comprehension

**Water is split**, providing a source of electrons and protons (hydrogen ions, H<sup>+</sup>) and giving off O<sub>2</sub> as a by-product. **Light absorbed** by chlorophyll drives a **transfer of the electrons and hydrogen ions** from water to an acceptor called NADP<sup>+</sup>.

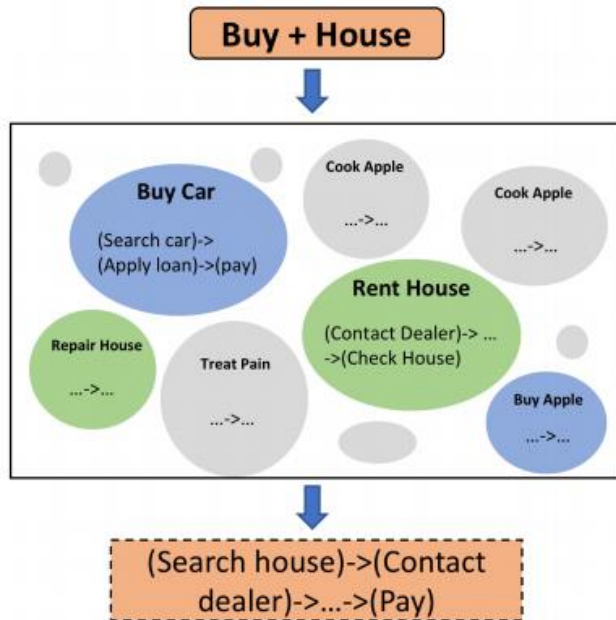
What can the splitting of water lead to?

**A:** Light absorption

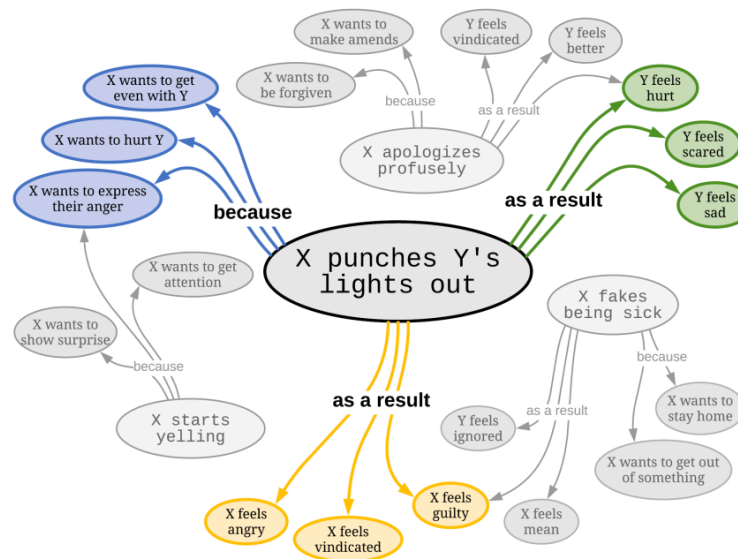
**B:** Transfer of ions



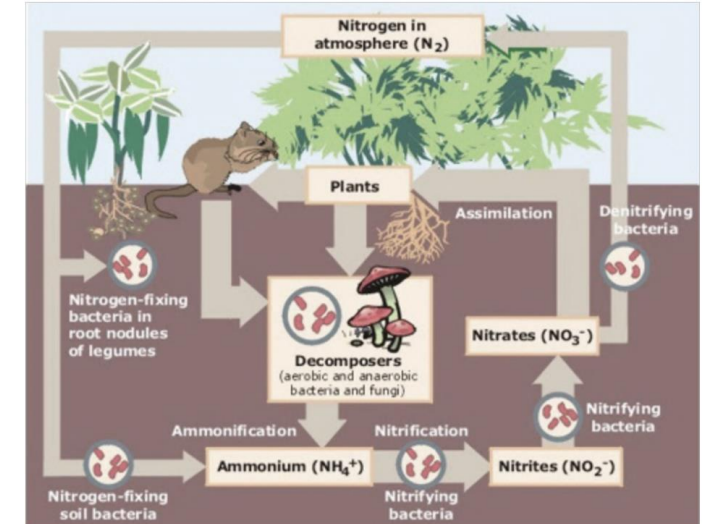
## 1. Event process completion



## 2. Event intention prediction



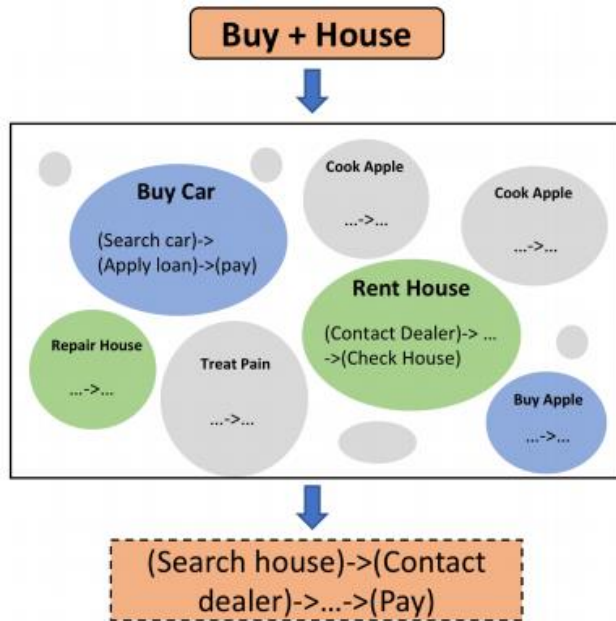
## 3. Event processes in downstream NLU tasks



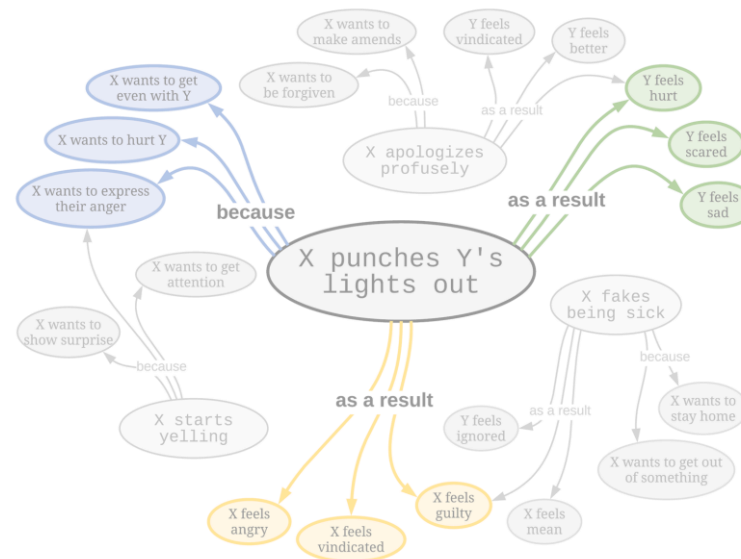
## 4. Open Research Directions



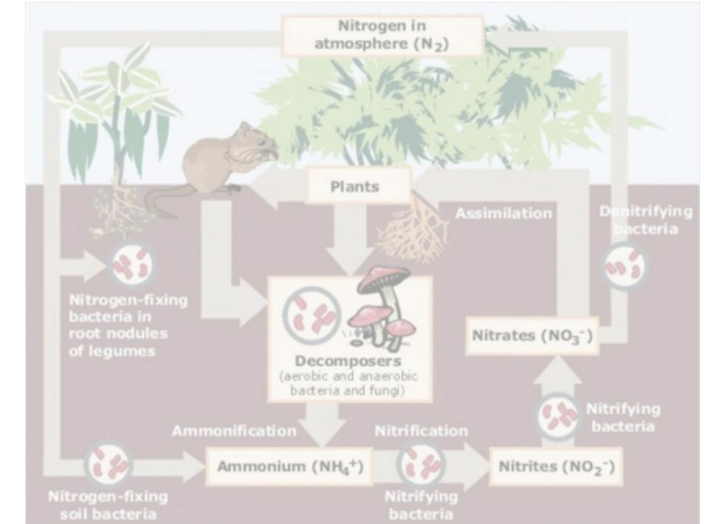
## 1. Event process completion



## 2. Event intention prediction



## 3. Event processes in downstream NLU tasks



## 4. Open Research Directions



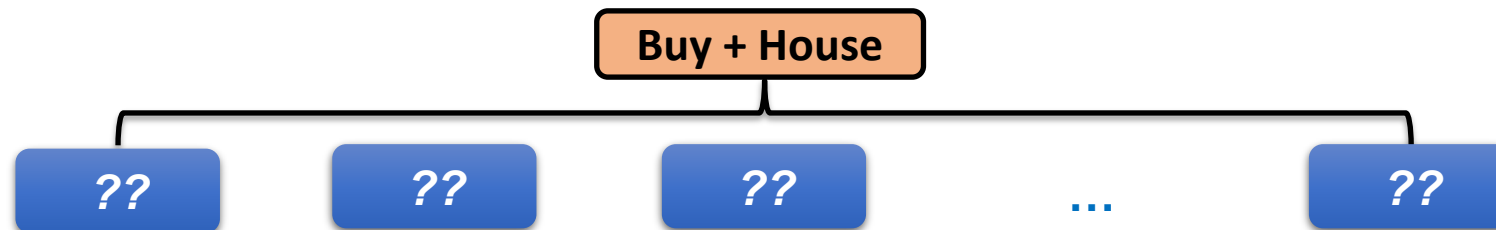


## Two forms of process prediction

1. Predicting steps of the process



2. Inducing the entire process from scratch.



Chambers and Jurafsky. Unsupervised Learning of Narrative Event Chains. ACL-08

Unsupervised event process completion can be done using corpus statistics

- Capturing the co-occurrence of events using pointwise mutual information

$$pmi(e(w, d), e(v, g))$$

- The next most likely forthcoming event can be found by maximizing the accumulated PMI

$$\max_{j:0 < j < m} \sum_{i=0}^n pmi(e_i, f_j)$$

( $n$ : #events in the process;  $m$ : #events in the vocabulary.)

## Known events:

(pleaded subj), (admits subj), (convicted obj)

## Likely Events:

sentenced obj	0.89	indicted obj	0.74
paroled obj	0.76	fined obj	0.73
fired obj	0.75	denied subj	0.73

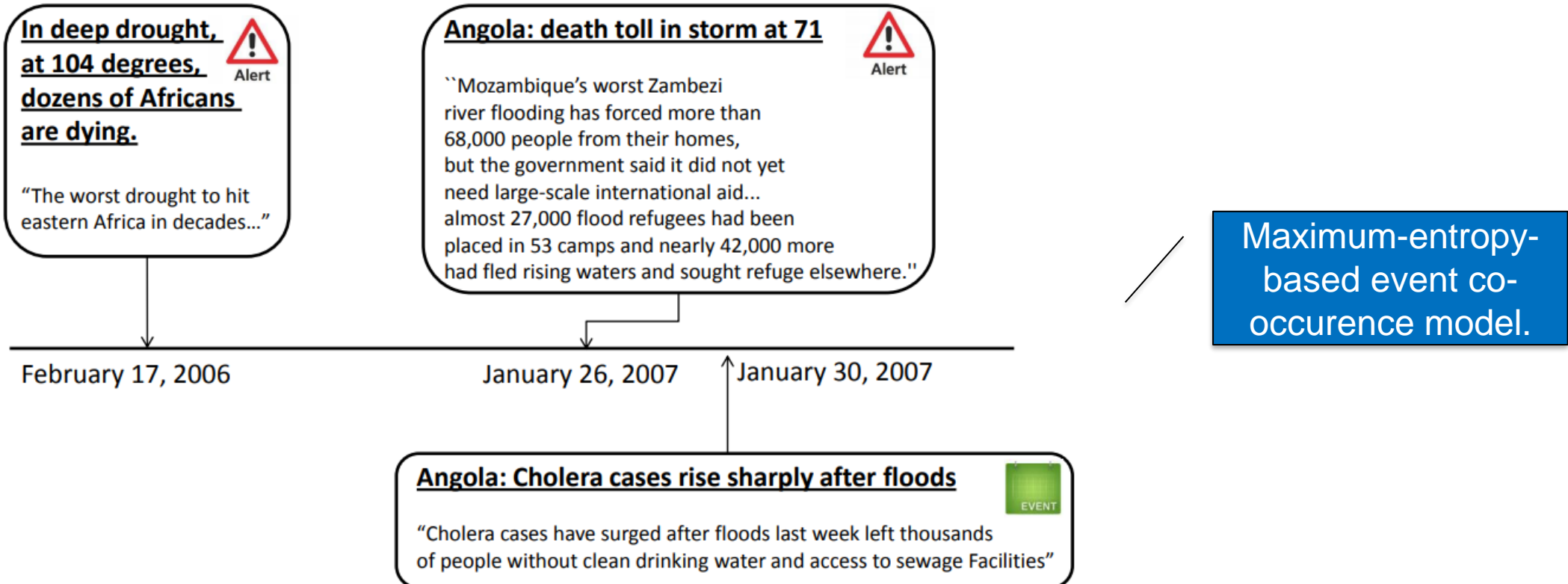


Improves narrative cloze tests (36% improvement on NYT Narrative Cloze).

# Event Process Completion

Radinsky and Horvitz. Mining the Web to Predict Future Events. WSDM, 2013

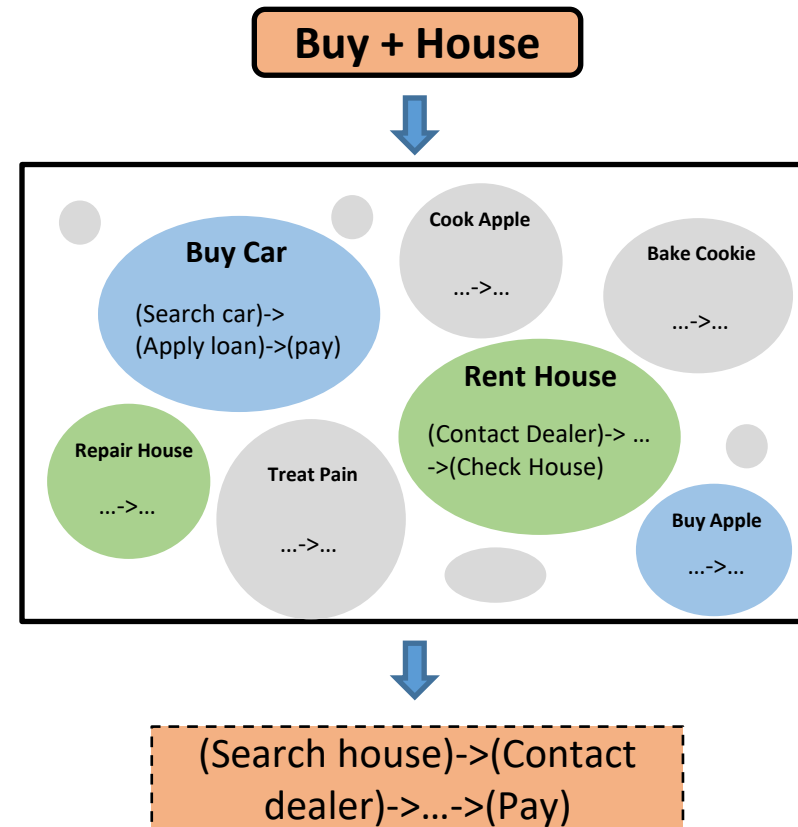
Extension of the event chain model on multiple **dated** and **topically cohesive** documents.



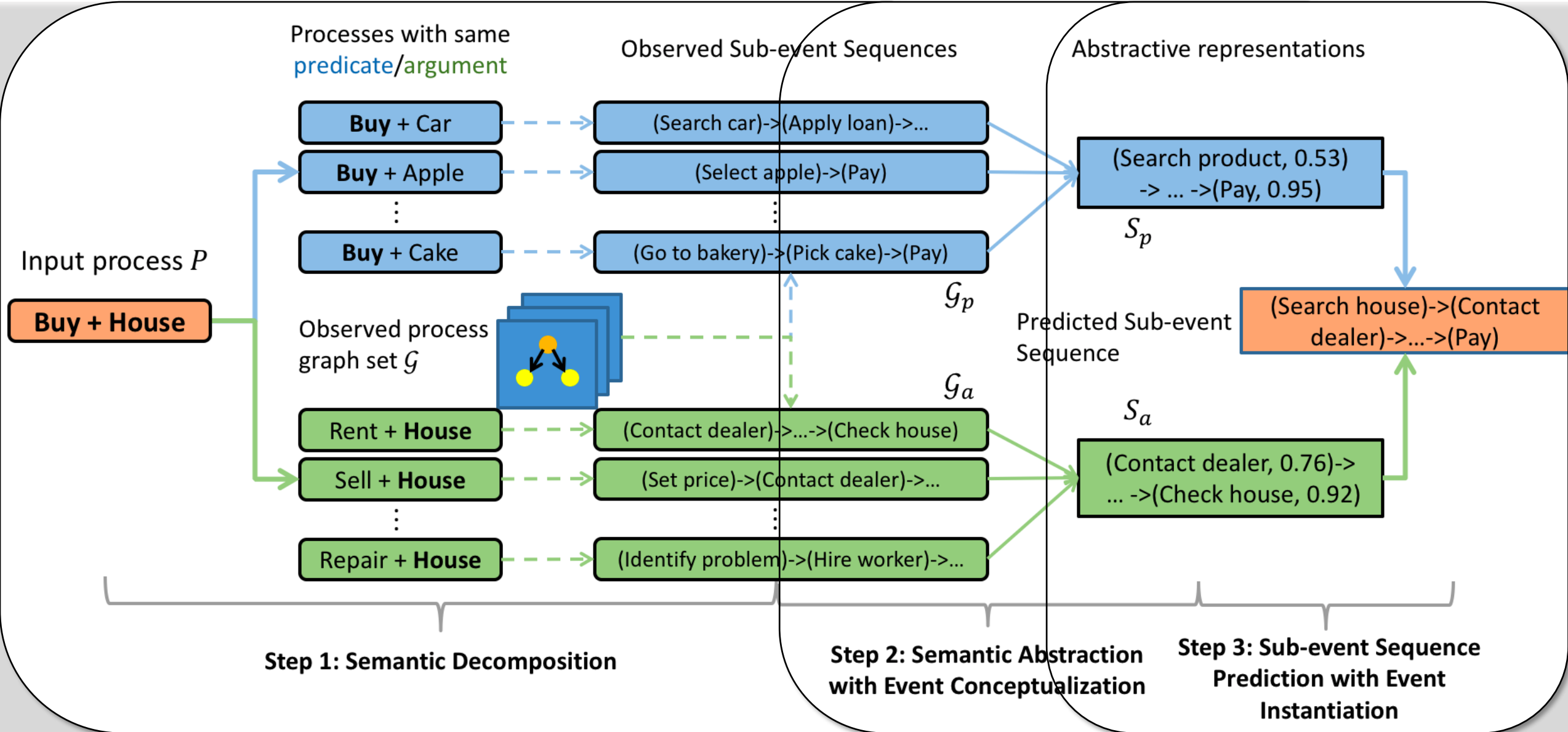
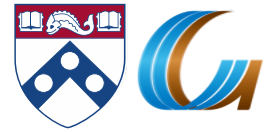
The likelihood of **cholera rising** is predicted **high** after a **drought followed by storms** in Angola (*based on corpus statistics*).

Can we perform *de novo* process induction?

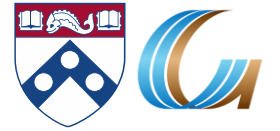
Zhang, et al. Analogous Process Structure Induction for Sub-event Sequence Prediction. EMNLP, 2020



# Analogous Event Process Induction



# Evaluation Based on wikiHow Event Processes



Model	String Match		Hypernym Allowed	
	E-ROUGE1	E-ROUGE2	E-ROUGE1	E-ROUGE2
Random	2.9165	0.4664	23.5873	8.1089
Seq2seq (GloVe)	5.0323	1.4965	27.8710	13.0946
Seq2seq (RoBERTa)	4.5455	0.4831	28.0032	12.8502
Top one similar process (Jaccard)	8.8589	5.1000	28.6548	14.6231
Top one similar process (GloVe)	9.8797	5.1452	29.4203	13.6001
Top one similar process (RoBERTa)	9.2599	4.7390	30.6599	15.8417
Analogous Process Structure Induction (APSI)	<b>14.8013</b>	<b>6.6045</b>	<b>36.1648</b>	<b>19.2418</b>
Human	29.0189	15.2542	50.4647	29.4423

(a) Basic Setting (for each sub-event, we only predict and evaluate the verb)

Model	String Match		Hypernym Allowed	
	E-ROUGE1	E-ROUGE2	E-ROUGE1	E-ROUGE2
Random	0.0000	0.0000	0.5104	0.0903
Seq2seq (GloVe)	0.1935	0.0534	0.9677	0.1069
Seq2seq (RoBERTa)	0.4870	0.0000	1.7857	0.2899
Top one similar process (Jaccard)	0.6562	0.2257	2.4797	0.5867
Top one similar process (GloVe)	0.8750	0.2106	2.8801	0.7372
Top one similar process (RoBERTa)	0.9479	0.3009	3.2811	0.9929
Analogous Process Structure Induction (APSI)	<b>3.4988</b>	<b>0.4513</b>	<b>6.1611</b>	<b>1.1885</b>
Human	11.6351	5.5905	18.0034	8.2695

(b) Advanced Setting (for each sub-event, we predict and evaluate all words)

## Quantitative results

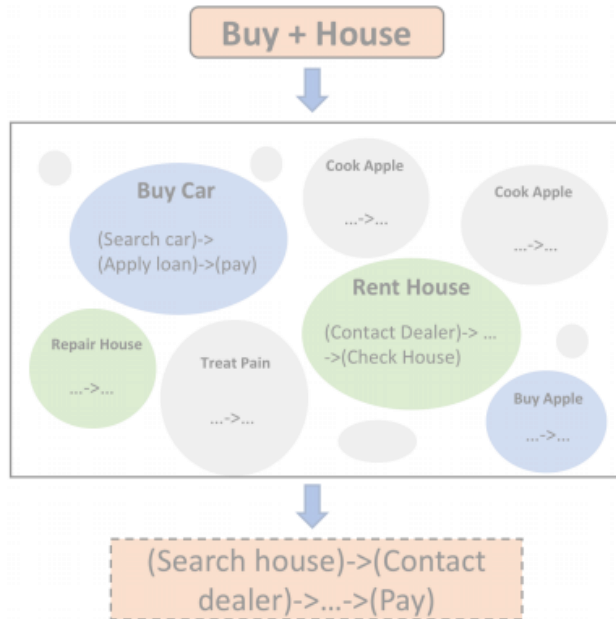
Process Name: **Treat Pain**

References: ('learn cause'->'identify symptom'->'see doctor')  
 ('identify cause'->'learn injury'->'recognize symptom'->'recognize symptom')

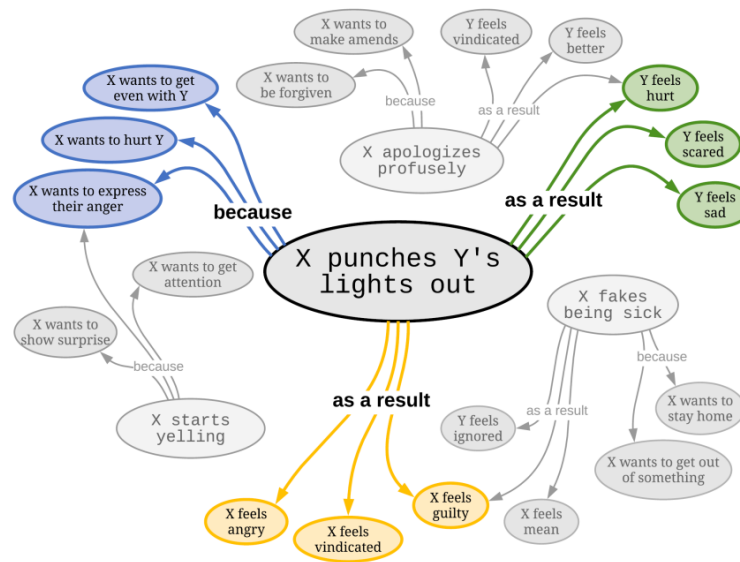
APSI Prediction: ('Identify symptom'->'see doctor'->'recognize symptom'->'take supplement')

## Qualitative results

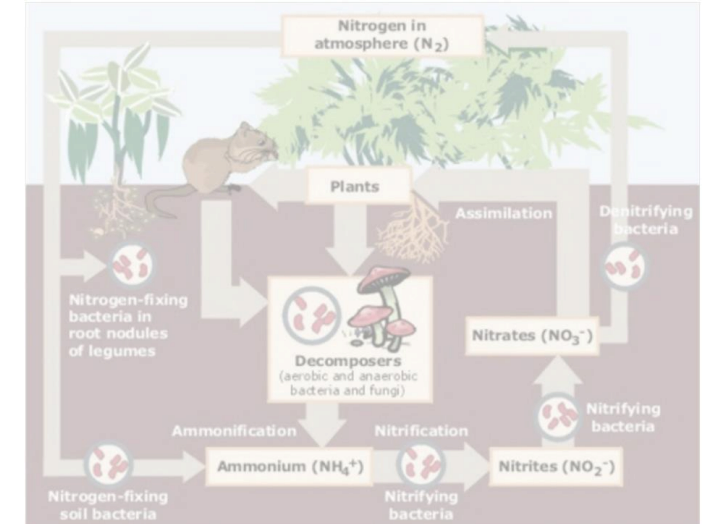
## 1. Event process completion



## 2. Event intention prediction



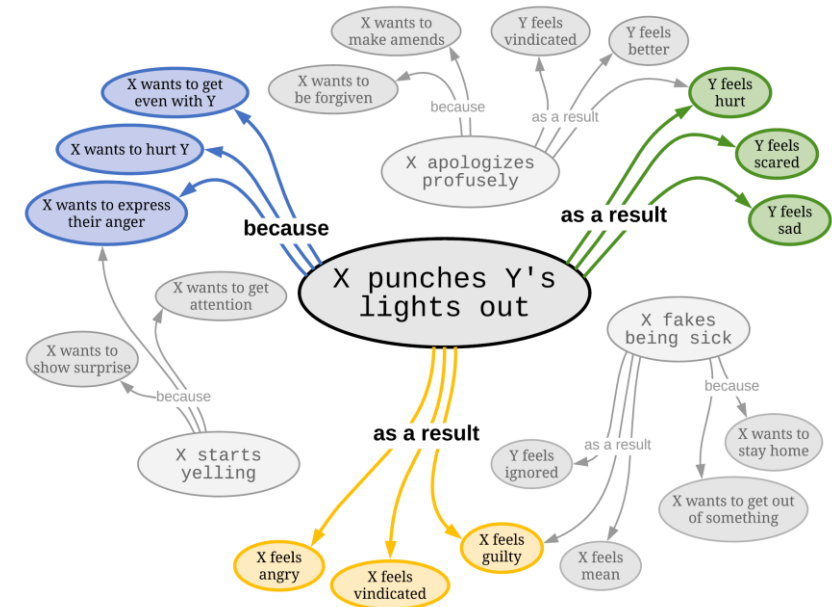
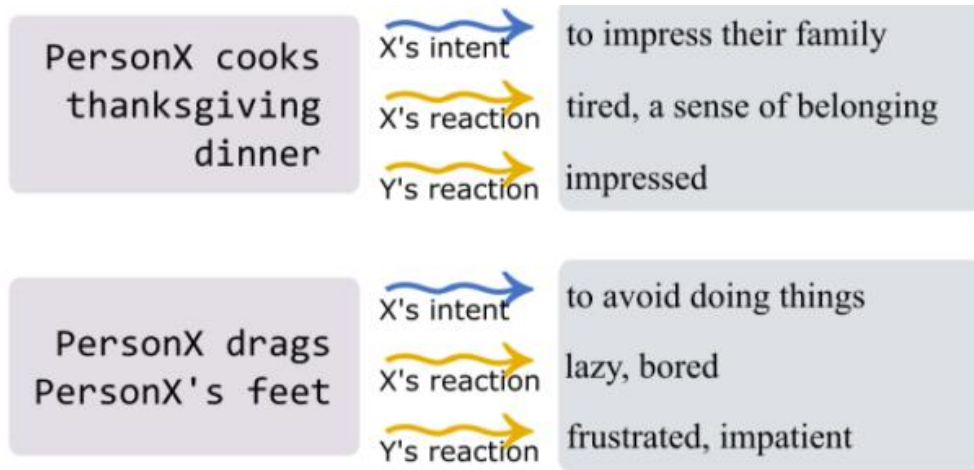
## 3. Event processes in downstream NLU tasks



## 4. Open Research Directions



People can easily anticipate the intents and possible reactions of participants in an event.



A commonsense-aware system should also perform such prediction.

Event2Mind – A learning system that understands stereotypical intents and reactions to events (Rashkin et al. ACL-18)



Is developed based on large crowdsourced corpora:

- 25,000 events
- Free-form descriptions of their intents and reactions

Performs Seq2Ngram generation:

PersonX cooks steak

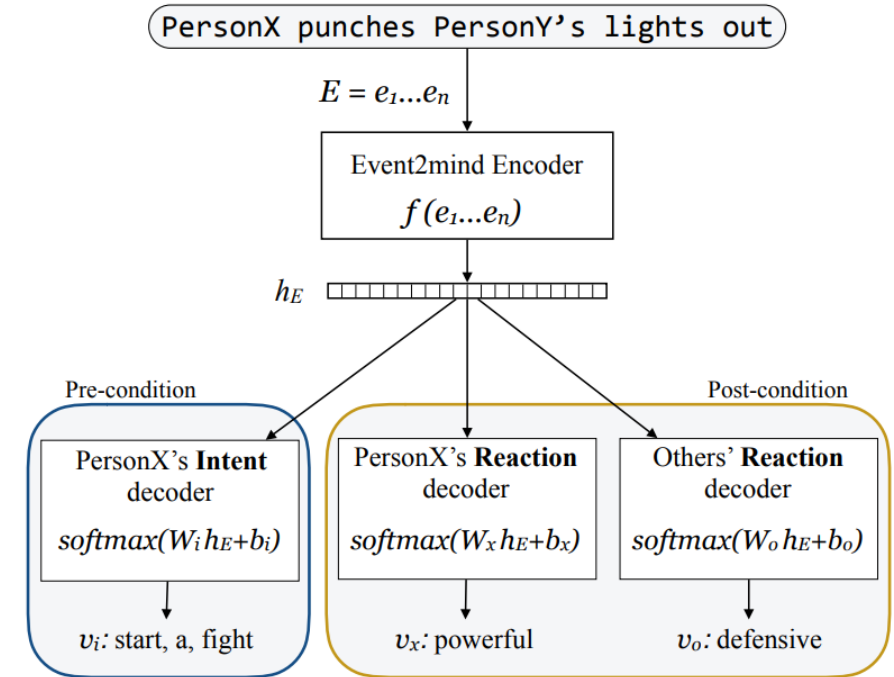
**PersonX's intent:** ["steak", "to kill their hunger", "to make dinner for the family", "to eat steak"]

**PersonX's reaction:** ["excited", "accomplished", "proud", "full"]

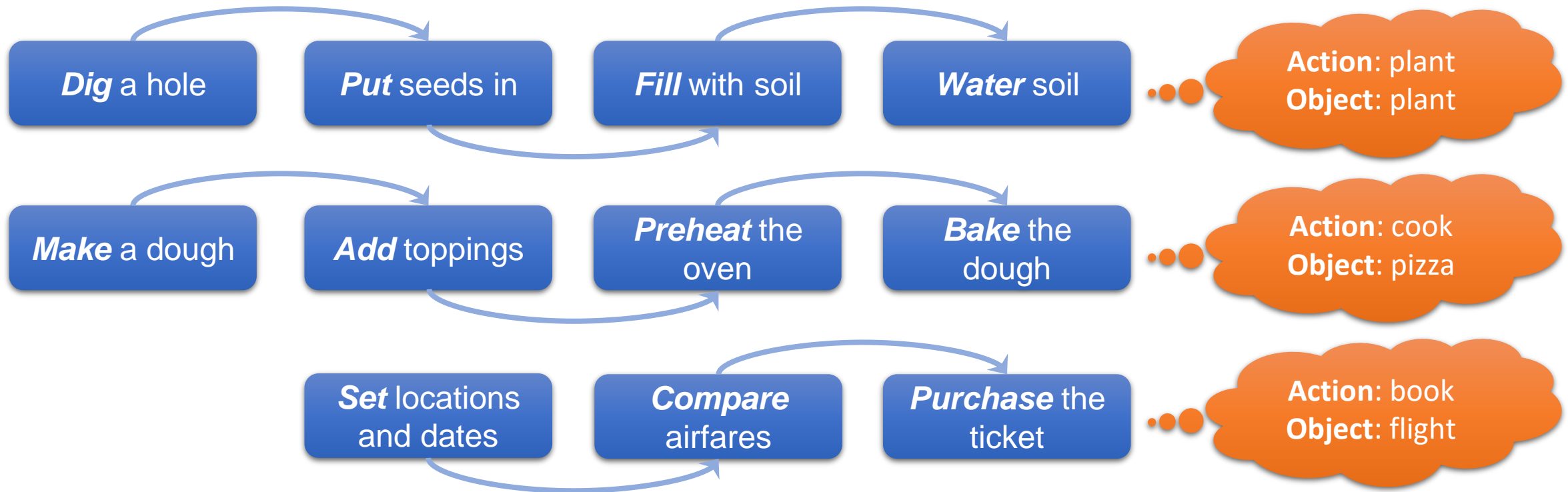
**Other people's reaction:** ["none", "happy", "person x cooked well."]

**More follow-ups of Event2Mind**

- ATOMIC: An Atlas of Machine Commonsense for If-Then Reasoning (Sap+ AACL 2019)
- COMET: Commonsense Transformers for Automatic Knowledge Graph Construction (Bosselut+, ACL-19)



# Intention Prediction for Event Processes



Event processes are directed by the **central goal**, or the **intention** of its performer [Zacks+, Nature Neuroscience 2001].

- Inherent to human's common sense.
- Missing from current computational methods.
- Important to machine commonsense reasoning, summarization, schema induction, etc.

Chen et al. “What are you trying to do?” Semantic Typing of Event Processes. CoNLL-2020

A new (cognitively motivated) **semantic typing task** for understanding event processes in natural language. Two **type axes**:

- What **action** the event process seeks to take? (**action type**)
- What type of **object(s)** it should affect? (**object type**)

This research also contributes with

- A **large dataset** of typed event processes (>60k processes)
- A **hybrid learning framework** for event process typing based on **indirect supervision**

# A Large Event Process Typing Dataset



A large dataset of typed event processes from wikiHow

- 60,277 event processes with free-form labels of action and object types

A challenging typing system

- Diversity: 1,336 action types and 10,441 object types (in free forms)
- Few-shot cases: 85.9% labels appear less than 10 times, (~half 1-shot).
- External labels: in 91.2% (84.2%) processes, the action (object) type label does not appear in the process body.

A non-trivial learning problem with **ultra fine-grained** and **extremely few-shot** labels.

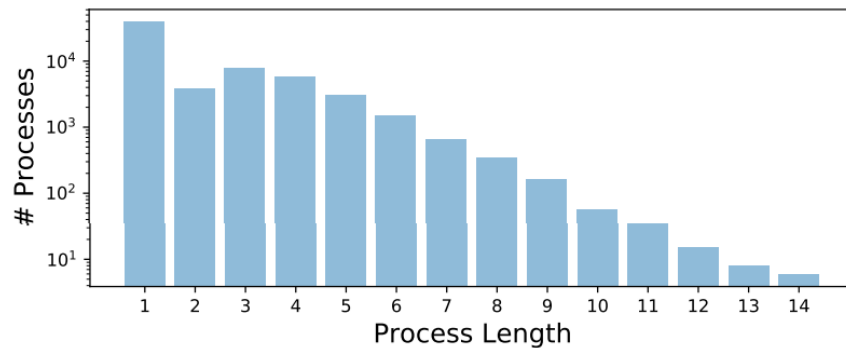


Figure 2: Distribution of process lengths.

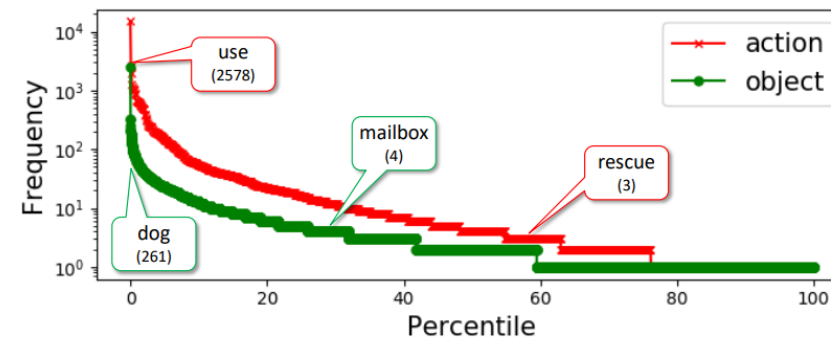
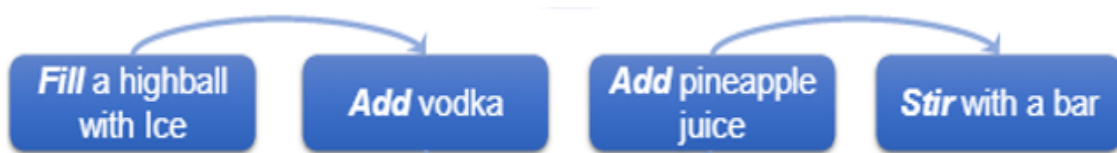


Figure 3: Distribution of actions and objects. Number of frequencies are shown in the brackets.



An event process



Indirect inference  
(Much Easier)



Directly inference  
(Difficult)

*Make*

*Cocktail*

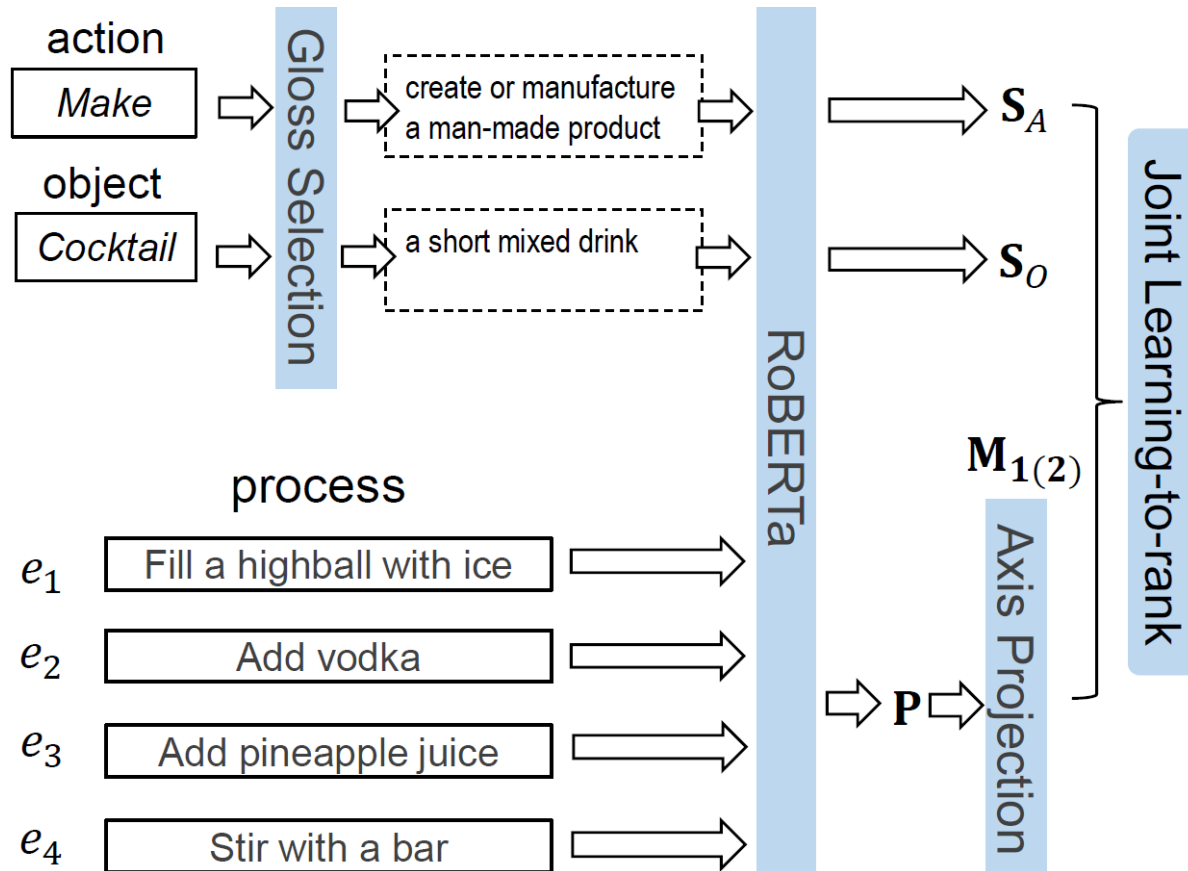
Labels

<i>Make</i>	create or manufacture a man-made product
<i>Cocktail</i>	a short, mixed drink

Label glosses (from WordNet)

## Why using label glosses?

- Semantically richer than labels themselves
- Capturing the association of a process-gloss pair (two sequences) is much easier
- Jump-starting few-shot label representations (and benefiting with fairer prediction)



## How to represent the process?

- RoBERTa encodes concatenated event contents (VERB and ARG1).

## How to represent a label?

- The same RoBERTa encodes the label gloss

## Which gloss for a polysemous label?

- WSD [Hadiwinoto+, EMNLP-19]
- MFS (Most frequent sense)

## Learning objective?

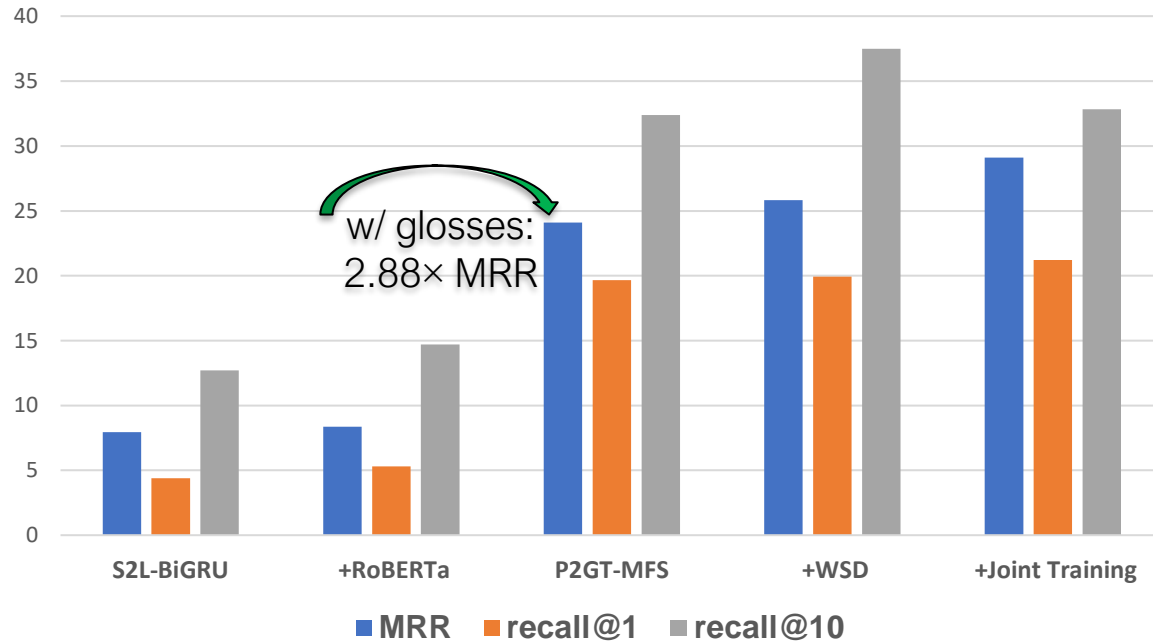
- Joint **learning-to-rank** for both type axes (different projection)

## Inference?

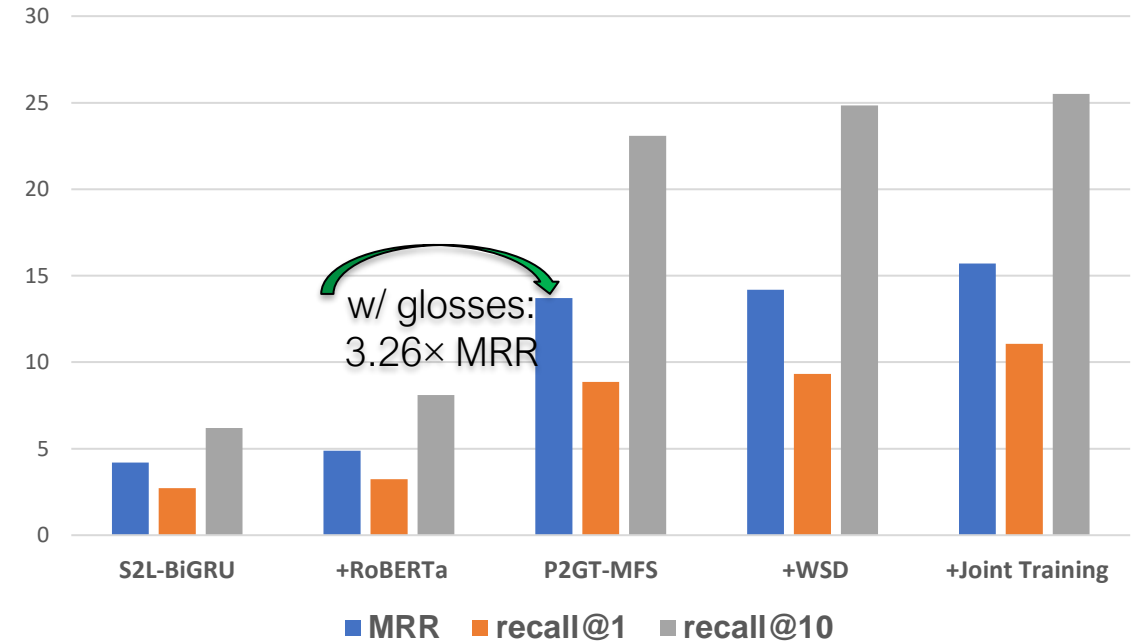
- Ranking all glosses for all labels in the vocab

# Results

### Action Typing of Processes (1,336 Labels)



### Object Typing of Processes (10,441 Labels)



- Gloss knowledge brings along the most improvement (2.88~3.26 folds of MRR)
- Joint training indicates the effectiveness of leveraging complementary supervision signals
- Sense selection (WSD) leads to lesser improvement (predominant senses are representative enough)

Event processes	Predictions
Make explosive materials ⇒ Obtain a container ⇒ Obtain shrapnel ⇒ Install a trigger	A: <b>detonate</b> , assemble, blacken O: <b>grenade</b> , blaster, <b>mine</b>
Go to DMV ⇒ Take photos ⇒ Take vision test ⇒ Take permit test ⇒ Take road test	A: <b>obtain</b> , <i>verify</i> , explore O: <b>license</b> , check, <i>visa</i>
Ignore order ⇒ Enter area ⇒ Enforce blockade ⇒ Force to retreat from area	A: <b>conquer</b> , <i>disarm</i> , <b>invade</b> O: <i>barrier</i> , <b>soldier</b> , <b>fortress</b>
Capture two opposition posts ⇒ Kill many fighters ⇒ Destroy three armed trucks ⇒ Confiscate artillery guns	A: <i>kill</i> , <b>demolish</b> , <b>fight</b> O: <i>melee</i> , <b>conflict</b> , <b>stronghold</b>
Cooperate with the counsel investigation ⇒ Open his remarks ⇒ Apologize many times ⇒ Try to restore public trust	A: <i>respond</i> , <i>disagree</i> , <b>accept</b> O: <i>apology</i> , <i>disagreement</i> , <b>slander</b>
Travel in a presidential motorcade ⇒ Be shot once in the back ⇒ Be taken to hospital ⇒ Be pronounced dead	A: <i>survive</i> , <b>die</b> , tackle O: <b>assassin</b> , crash, <i>roadkill</i>
Give advance notice ⇒ Give notice ⇒ Issue dividends	A: <b>honor</b> , pay, <b>reward</b> O: <i>finance</i> , <b>equity</b> , <b>subsidy</b>
Target quotes ⇒ Target shares quotes ⇒ Ask to clarify offer ⇒ Challenge to merge agreement ⇒ Challenge to merge businesses	A: <b>compare</b> , maximize, <b>negotiate</b> O: <i>prospectus</i> , <b>quote</b> , <b>settlement</b>
Clean windows ⇒ Buy plants ⇒ Hang pictures ⇒ Paint walls ⇒ Carpet floors	A: <b>redecorate</b> , decorate, <i>refurbish</i> O: <b>room</b> , <b>bedroom</b> , <i>makeover</i>

Table 3: Case study for typing event processes in the news domain. The predictions are given by Joint P2GT-WSD trained on our full dataset. Each case is given top 3 predictions on both axes, whereof reasonably correct ones are boldfaced, and relevant ones are italic. Few-shot labels appearing up to 10 times in our dataset are in blue.



# System Demonstration



A web demonstration of our prototype system is running at [https://cogcomp.seas.upenn.edu/page/demo\\_view/step](https://cogcomp.seas.upenn.edu/page/demo_view/step)

Examples

Decoration

Event process (choose an example or write the subevents of a process separated by '@' to get its intention)

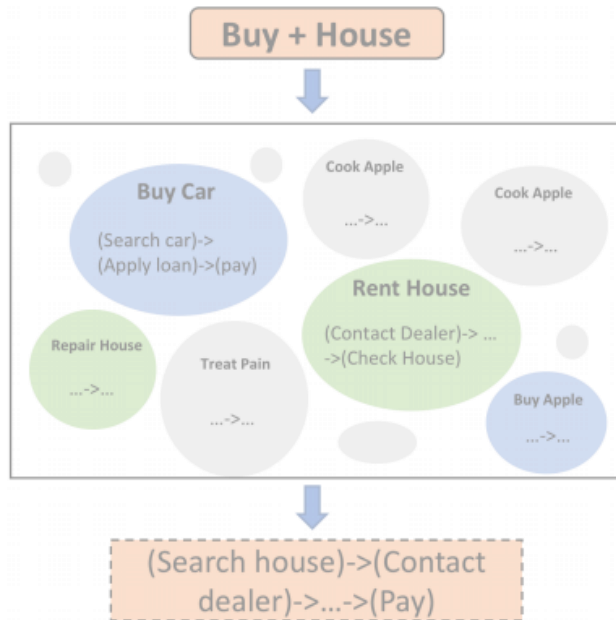
clean windows @ buy plants @ paint walls @ hang pictures @ carpet floors @ reorganize furniture

Get intention >

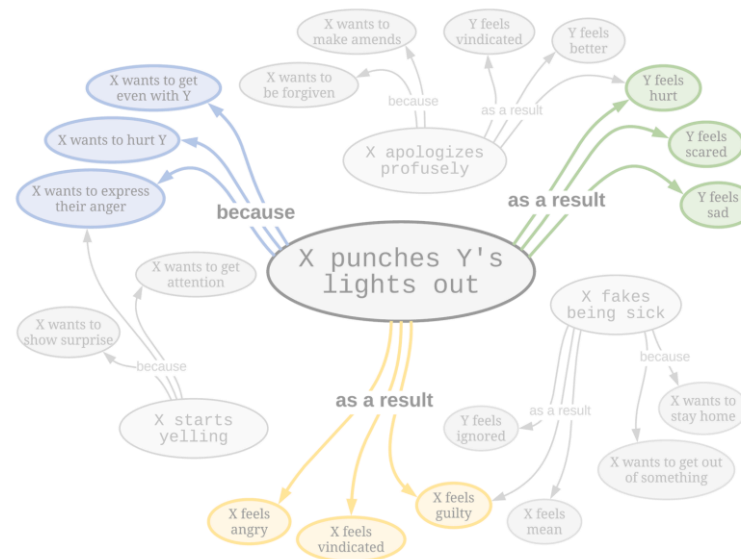
redecorate room

Cosine similarity	Action	Object	Cosine similarity
0.678	redecorate	room	0.623
0.650	stage	atmosphere	0.599
0.500	brighten	mosaic	0.589
0.427	preoccupy	suite	0.574
0.418	furnish	interior	0.573

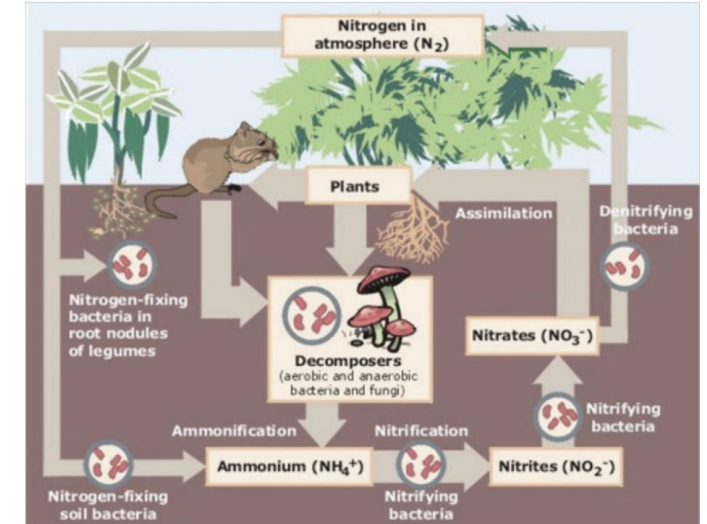
## 1. Event process completion



## 2. Event intention prediction



## 3. Event processes in downstream NLU tasks



## 4. Open Research Directions



## The ROC Story Narrative Cloze Test [Mostafazadeh+, NAACL 2016]:

One day Wesley's auntie came over to visit. He was happy to see her, because he liked to play with her. When she started to give his little sister attention, he got **jealous**. He got **angry** at his auntie and **bit** his sister's hand when she wasn't looking.

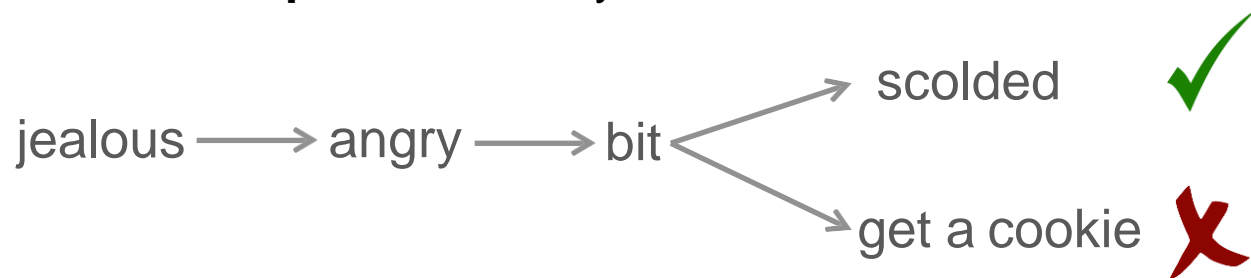
Then what might happen?

**O1:** He was **scolded**.

**O2:** She **gave him a cookie** for being so nice.

Chaturvedi, et al (EMNLP, 2017) train a language model that captures three types of sequential features:

### 1. Event sequences in 20 years of NYT data



### 2. Sentiment trajectories

### 3. Sentential topical consistency

Features	Accuracy
All	74.4%
Event-sequence	71.6%
Sentiment	64.5%
Topic	55.2%

Event sequences are most important.

## QA based on articles in biology

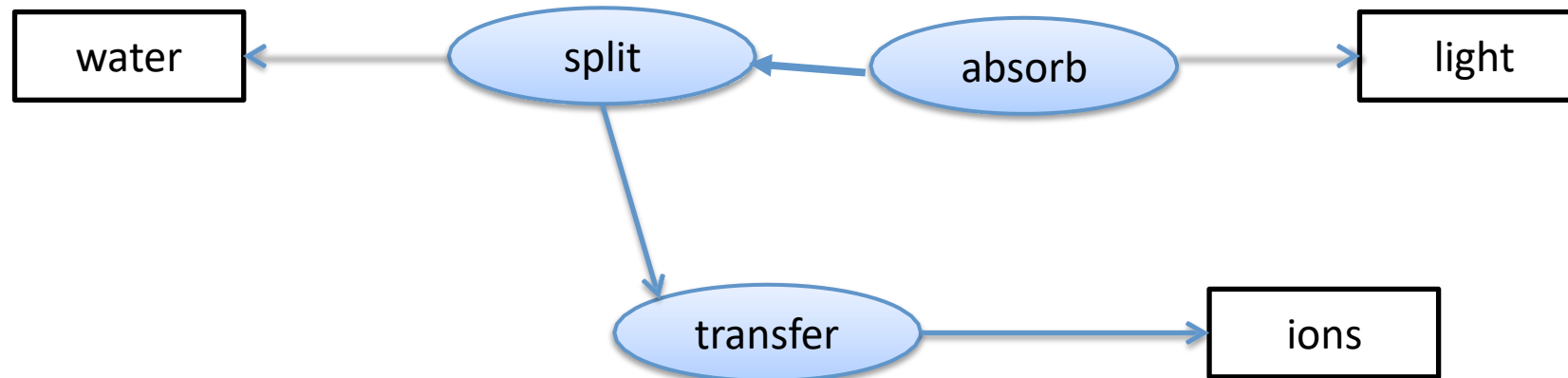
*Water is split*, providing a source of electrons and protons (hydrogen ions, H<sup>+</sup>) and giving off O<sub>2</sub> as a by-product. *Light absorbed* by chlorophyll drives a *transfer of the electrons and hydrogen ions* from water to an acceptor called NADP<sup>+</sup>.

What can the splitting of water lead to?

**A:** Light absorption

**B:** Transfer of ions

### 1. Extracting events and event-event relations from articles



### 2. Matching questions and candidate answers with extracted event processes

Events in a process as anchors of video segments.

wikiHow process: make pancakes: {add egg, add flour, ..., pour batter, remove pancake}

**Video segments:**

background

pour batter

background

remove pancake

**Video:**

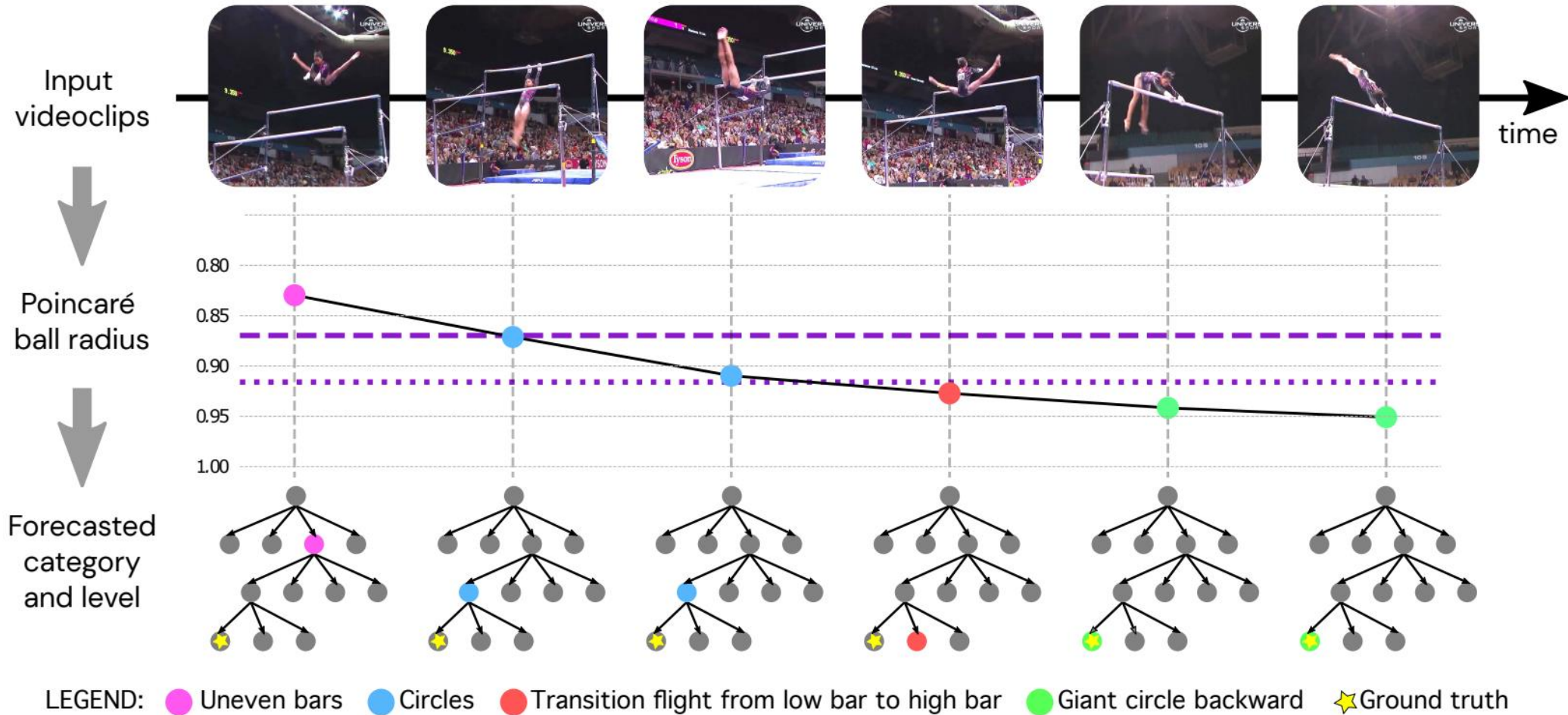


**Video narration:**

*hey folks here welcome to my kitchen ... pour a nice-sized amount ... change the angle to show ... and take it out*

Alignment learning between video narration and wikiHow event processes help action segmentation in videos.

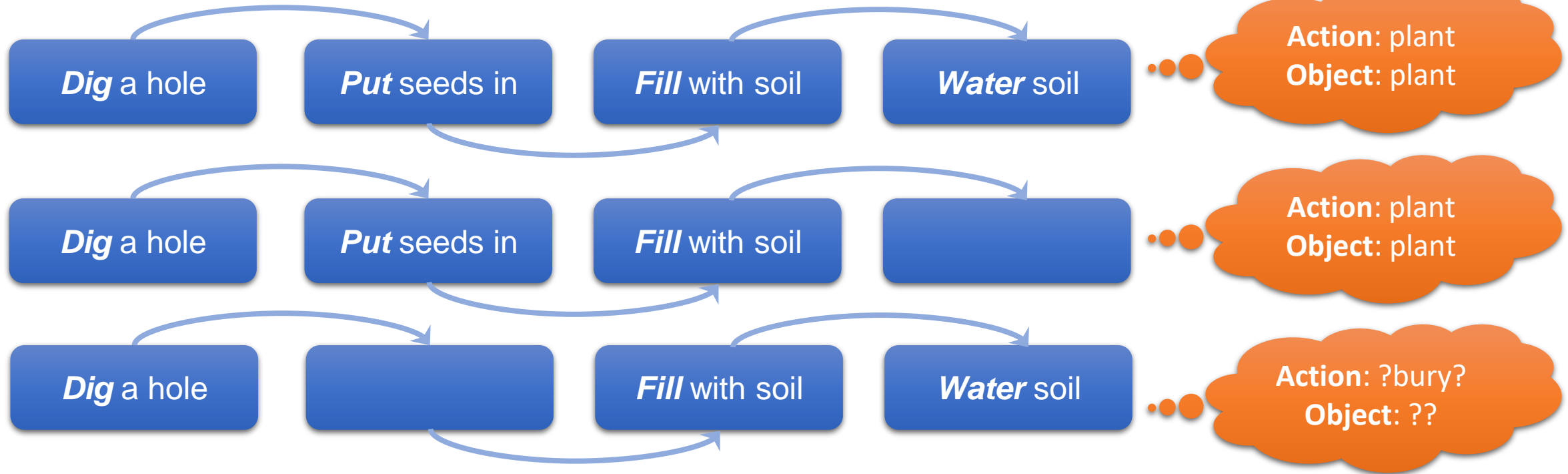
# Future Event Prediction in Videos



Hyperbolic embeddings model hierarchies of possible event evolution processes in videos.



Events in a process are not equally important



**Defending your dissertation** is essential; **Doing a TAShip** is less important; **Doing an internship** is optional...

Is there a way to automatically identify salient events in a process?  
Would that help downstream tasks such as abstractive summarization?



## Identifying the order of member events in a process is an unresolved challenge

Heavy snow is causing disruption to transport across the UK, with heavy rainfall bringing flooding to the south-west of England. Rescuers searching for a woman trapped in a landslide at her home in Looe, Cornwall, said they had found a body.

**Q1: What events have already finished?**

A: searching trapped landslide said found

**Q2: What events have begun but has not finished?**

A: snow causing disruption rainfall bringing flooding

**Q3: What will happen in the future?**

A: No answers.

warm-up

**Q4: What happened before a woman was trapped?**

A: landslide

**Q5: What had started before a woman was trapped?**

A: snow rainfall landslide

**Q6: What happened while a woman was trapped?**

A: searching

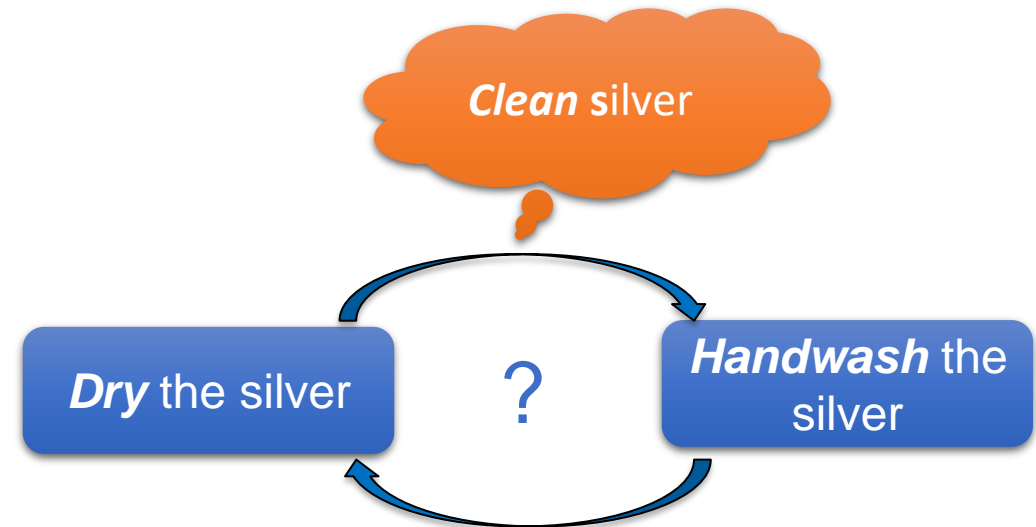
**Q7: What happened after a woman was trapped?**

A: searching said found

User-provided

Ning, et al. TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. EMNLP, 2020

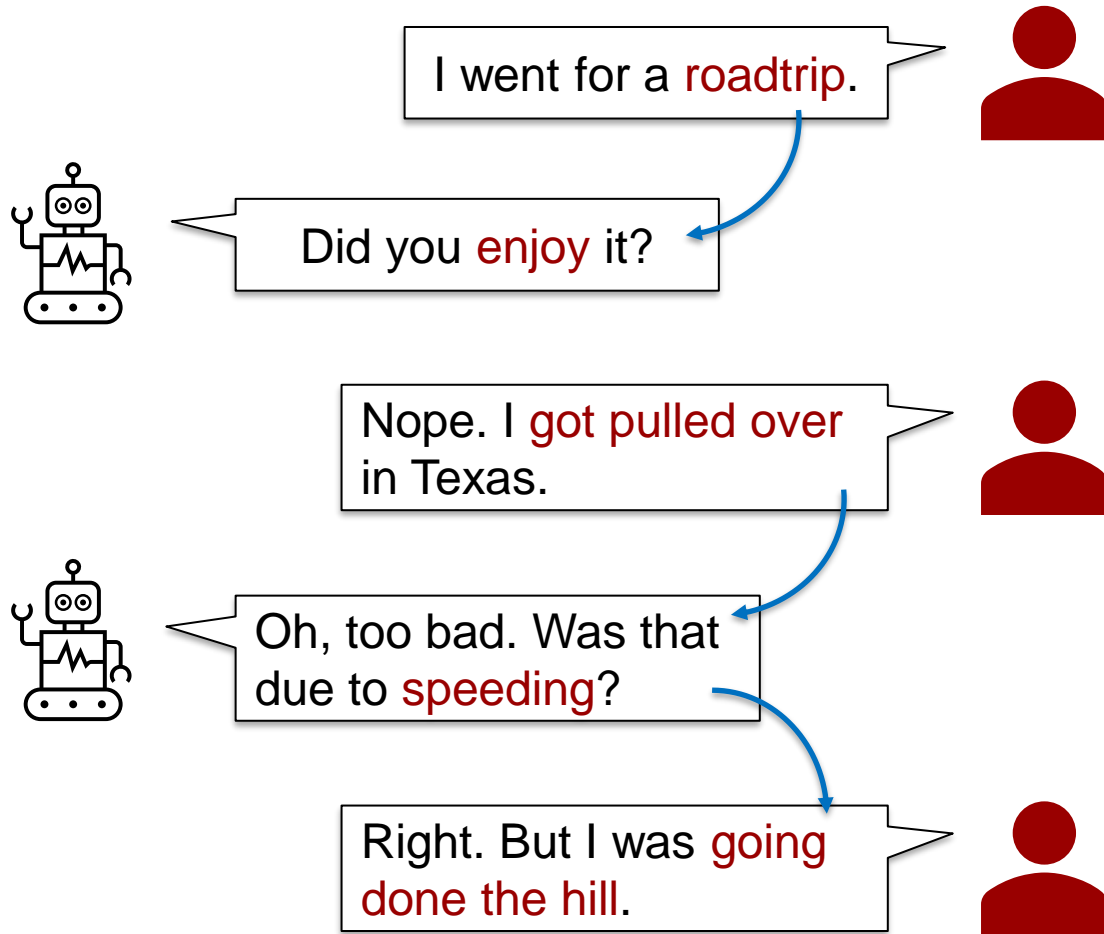
- 3.2k news snippets with 21k human-generated questions querying temporal relationships



Lyu, et al. Reasoning about Goals, Steps, and Temporal Ordering with WikiHow. EMNLP, 2020

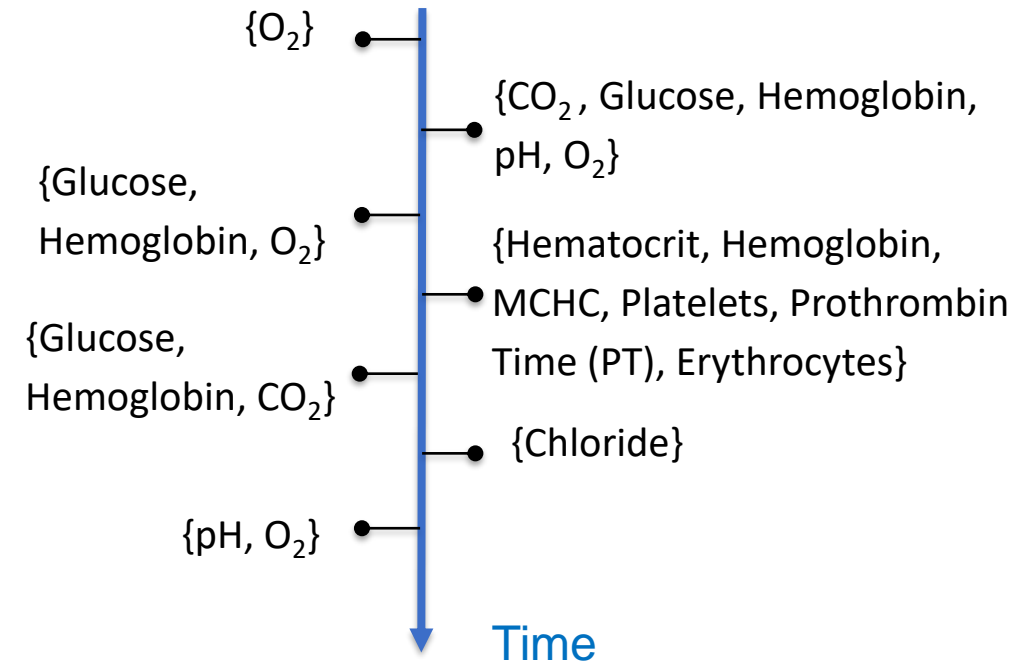
- A wikiHow-based testbed about event ordering (and more)

## Chatbots



Can event processes improve the consistency of utterance generation/retrieval?

## Understanding clinical event processes



Diagnostic prediction (Zhang et al. AIME-20), phenotype prediction, ...

- Transfer learning can be important (naturally lack of data)
- Structured prediction can be important (dependency of phenotypes, disease labels)

- Zack, et al. Human brain activity time-locked to perceptual event boundaries. *Nature neuroscience*, 4(6):651–655. 2001
- Chambers and Jurafsky. Unsupervised learning of narrative event chains. *ACL*, 2008
- Radinsky and Horvitz. Mining the Web to Predict Future Events. *WSDM*, 2013
- Berant, et al. Modeling Biological Processes for Reading Comprehension. *EMNLP*, 2014
- Chaturvedi, et al. Story comprehension for predicting what happens next. *EMNLP*, 2017
- Rashkin, et al. Event2Mind: Commonsense Inference on Events, Intents, and Reactions. *ACL*, 2018
- Liu, et al. Automatic event salience identification. *EMNLP*, 2018
- Zhukov et al. Cross-task weakly supervised learning from instructional videos. *CVPR*, 2019
- Zhang, et al. Analogous Process Structure Induction for Sub-event Sequence Prediction. *EMNLP*, 2020
- Chen, et al. “What are you trying to do?” Semantic typing of event processes. *CoNLL*, 2020
- Ning et al. TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. *EMNLP*, 2020
- Lyu, et al. Reasoning about Goals, Steps, and Temporal Ordering with WikiHow. *EMNLP*, 2020
- Jindai, et al. Is Killed More Significant than Fled? A Contextual Model for Salient Event Detection. *COLING*, 2020
- Fried, et al. Learning to Segment Actions from Observation and Narration. *ACL*, 2020
- Zhang, et al. Diagnostic Prediction with Sequence-of-sets Representation Learning for Clinical Events. *AIME*, 2020
- Surís, et al. Learning the Predictability of the Future. *arXiv:2101.01600*, 2021

**Thank You**