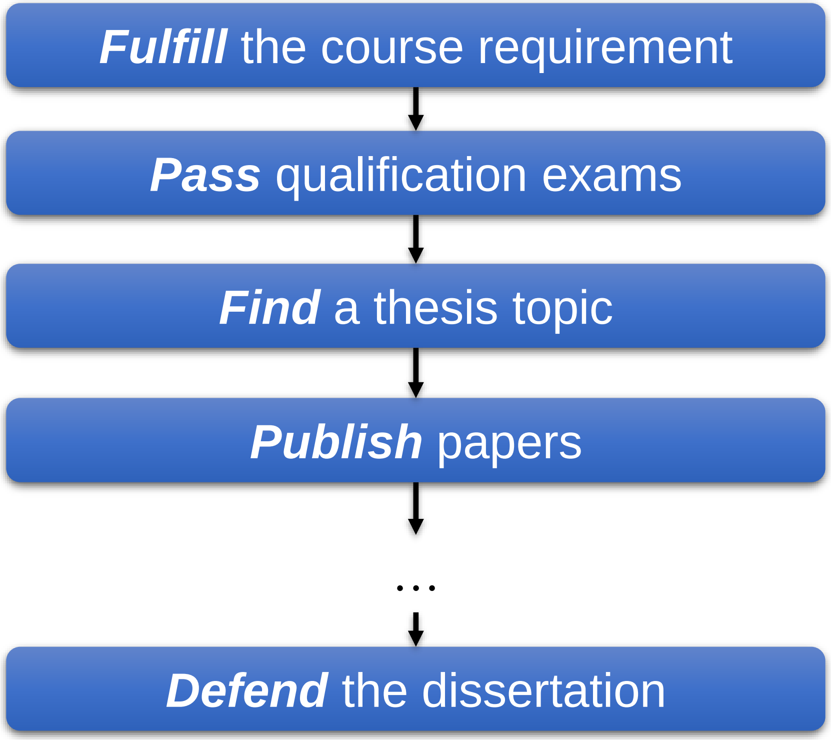
"What Are You Trying To Do" Semantic Typing of Event Processes

Muhao Chen[1,2], Hongming Zhang[2], Haoyu Wang[2] & Dan Roth[2]
[1]Information Sciences Institute, USC
[2]Department of Computer and Information Science, UPenn

Understanding Event Processes

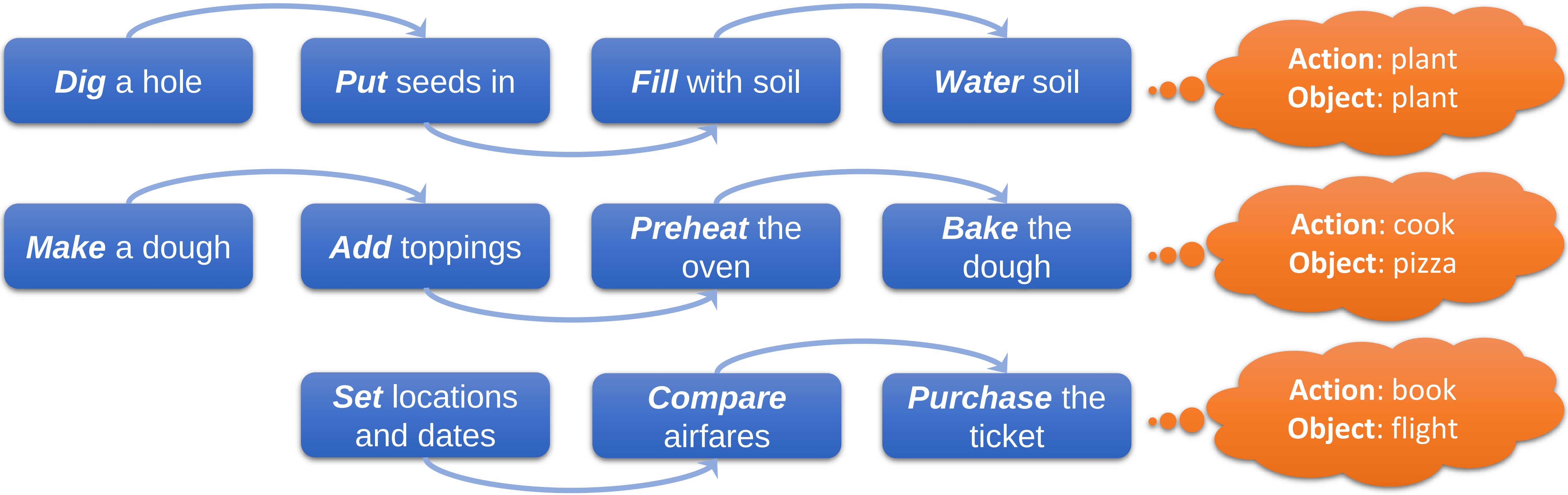
Natural language always involves descriptions of event processes.

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***.



An event process: a chain of events that happen sequentially.

Understanding 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.

A New Task: Multi-axis Event Process Typing

Three Contributions of This Work

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**

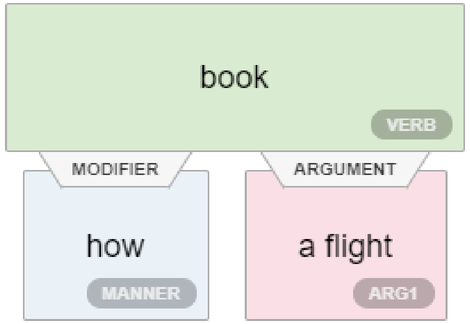
A Large Event Process Typing Dataset

wikiHow: an online wiki-style community containing professionally edited **how-to guideline articles.**

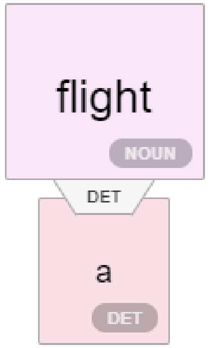
wikiHow to do anything...

How to Book a Flight

SRL



Dependency Parsing on ARG1



Action: ***book*** *(VERB)*   Object: ***flight*** *(head of ARG1)*

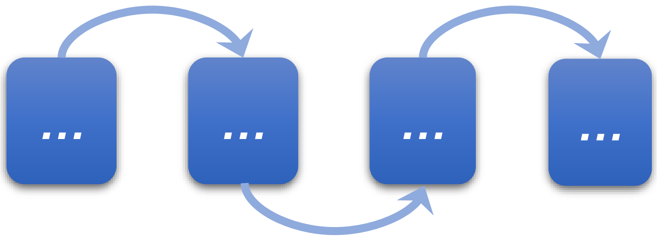
1 **Outline your tentative travel plans.** Just as with online booking, think about where you plan or would like to travel, the dates you would like to go, if you just want to book flights or maybe a package deal.

2 **Contact a travel agent or an airline representative.** You can call either traditional travel agents or airline reps to help you find the best flight booking.

3 **Compare prices from different agents.** Call several travel agents and ask them for price quotes. By comparing what different agents offer, you will get the best flight deal.

4 **Purchase your ticket.** Once you decide on the right flight offer for your upcoming travel, it's time to buy your ticket.

SRL on Section Titles



An Event Process

A Large Event Process Typing Dataset

A large dataset of typed event processes

- 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 froms)
- **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.

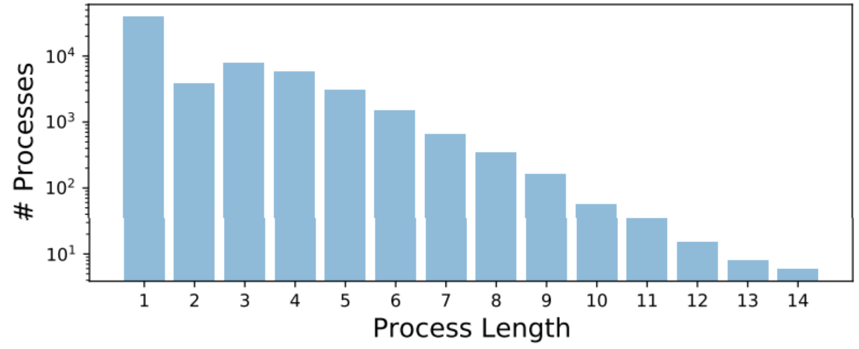


Figure 2: Distribution of process lengths.

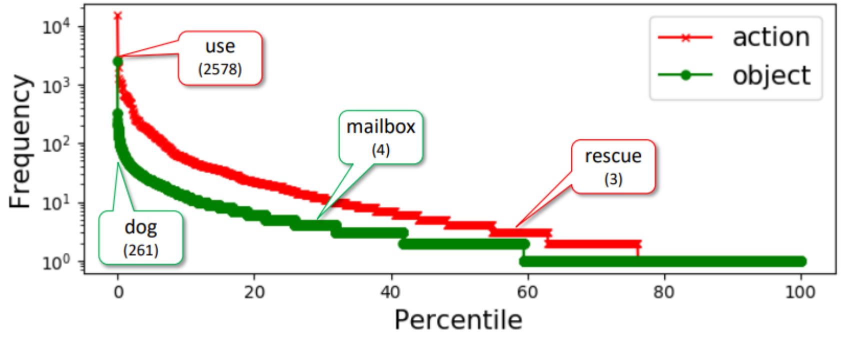
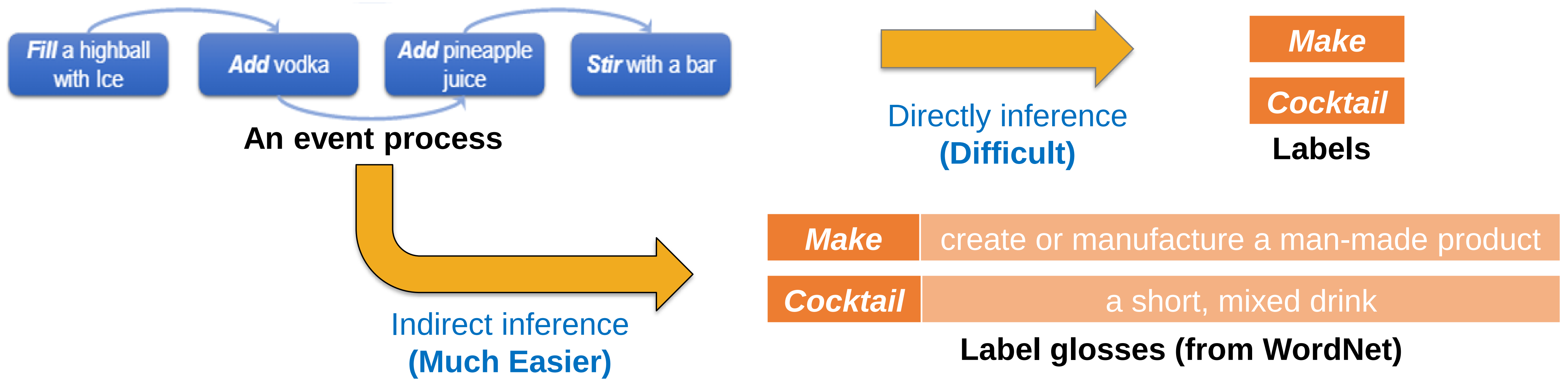


Figure 3: Distribution of actions and objects. Number of frequencies are shown in the brackets.

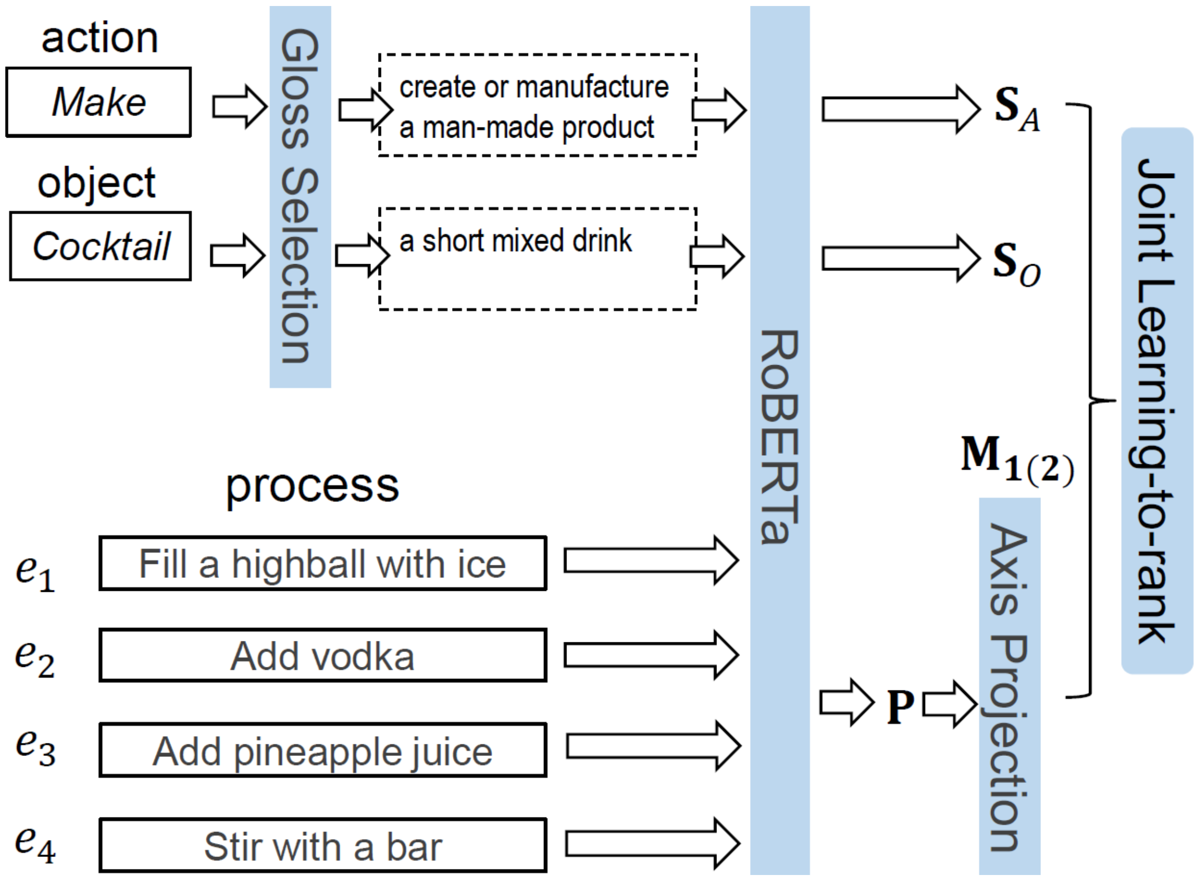
P2GT: Typing with Indirect Supervision from Gloss Knowledge



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)

P2GT: Typing with Indirect Supervision from Gloss Knowledge



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

Experimental Settings

Evaluation protocol

- 60,277 event processes
- 80/10/10 train/dev/test split

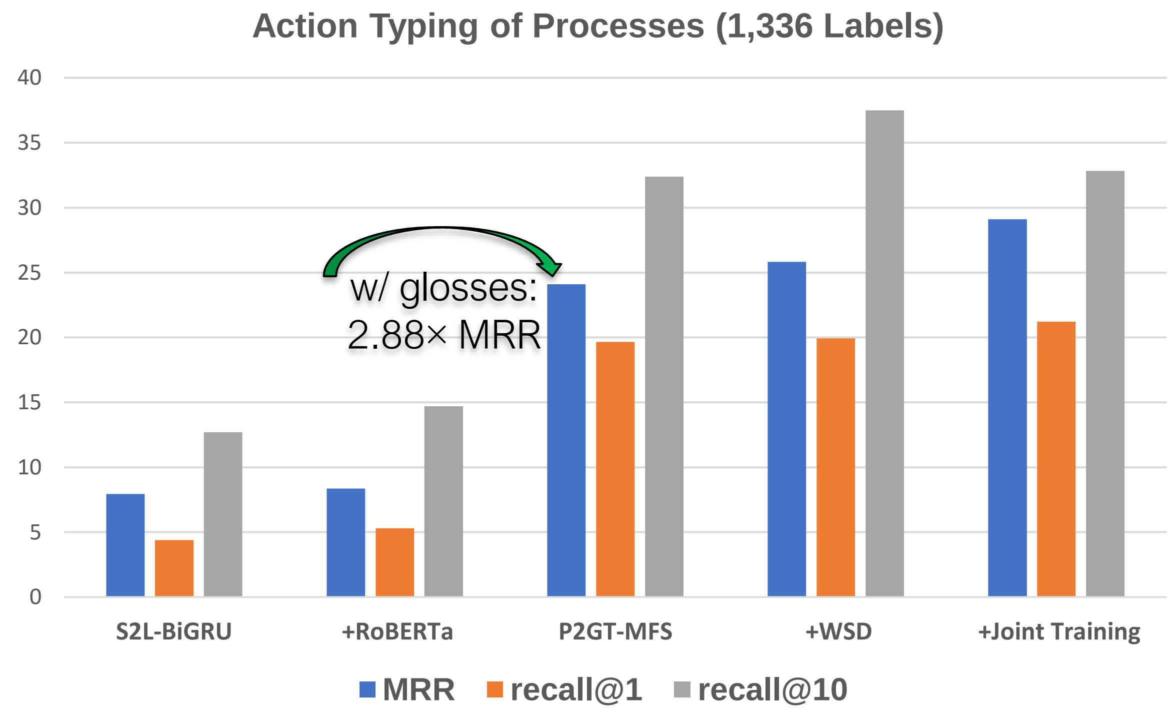
Compared methods

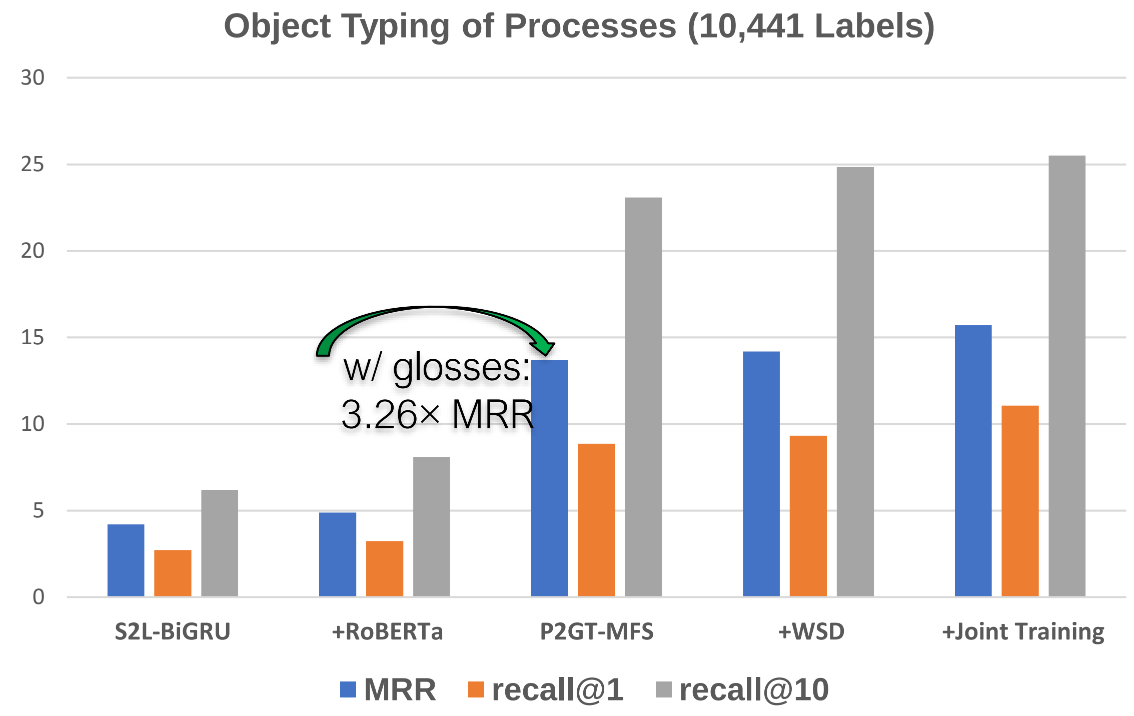
- **Sequence to label generators** (S2L) [Rashkin+, ACL-18]
  - Different encoders: pooling, BiGRU, RoBERTa
- **Variants of P2GT**
  - w/ or w/o multi-axis joint training
  - w/ or w/o WSD-based gloss selection
  - Partial information for event representation (VERB only or ARG only)

Ranking metrics

- *recall@1*, *recall@10*
- Mean Reciprocal Rank (*MRR*)

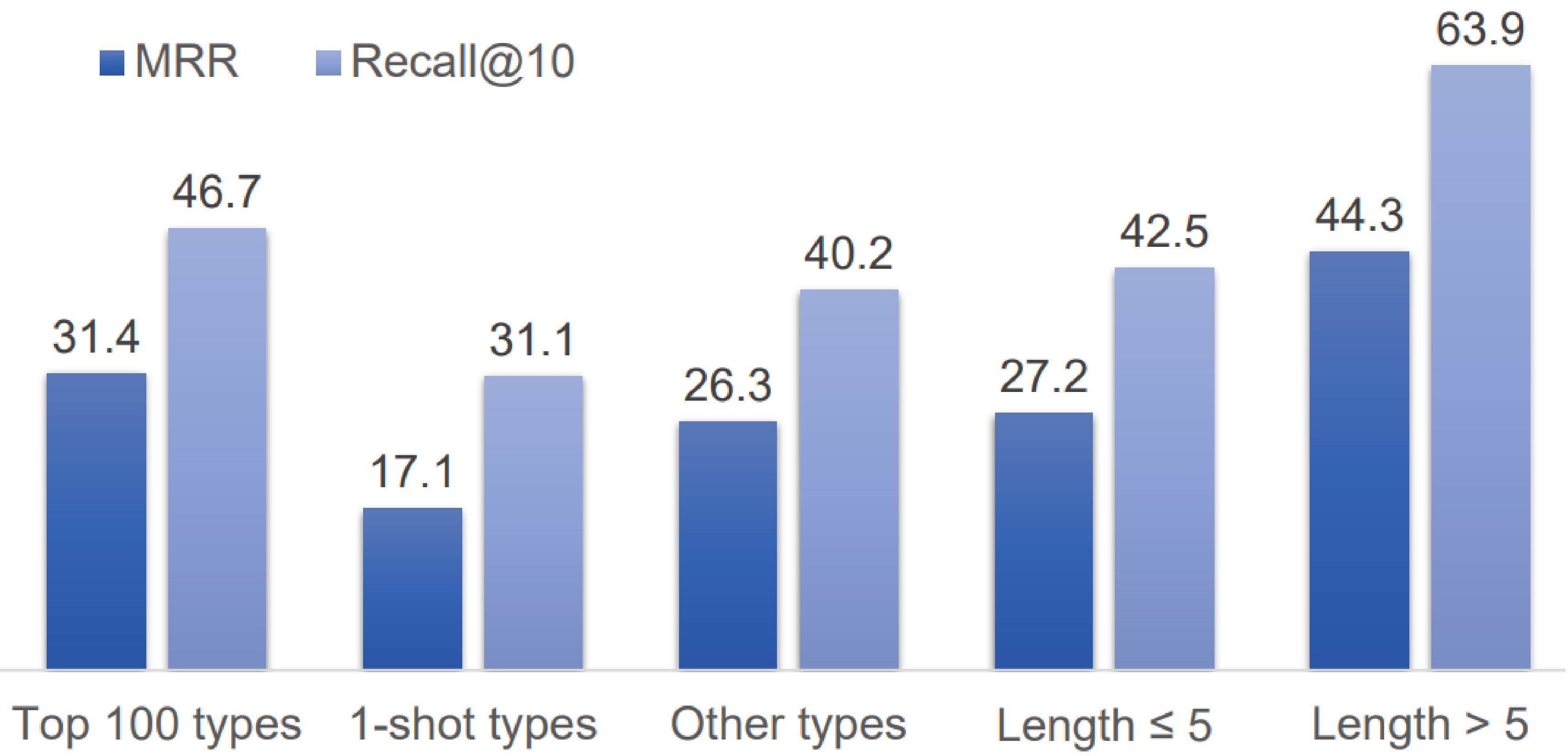
Main Results





- 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)

Error Analysis



- Performance is better on more frequent labels (as expected)
- On 1-shot cases, it performs reasonably well
- Longer processes are easier to type (w/ more contextual information of associated events)

Case Study

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

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 http://dickens.seas.upenn.edu:4035/

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

Conclusion

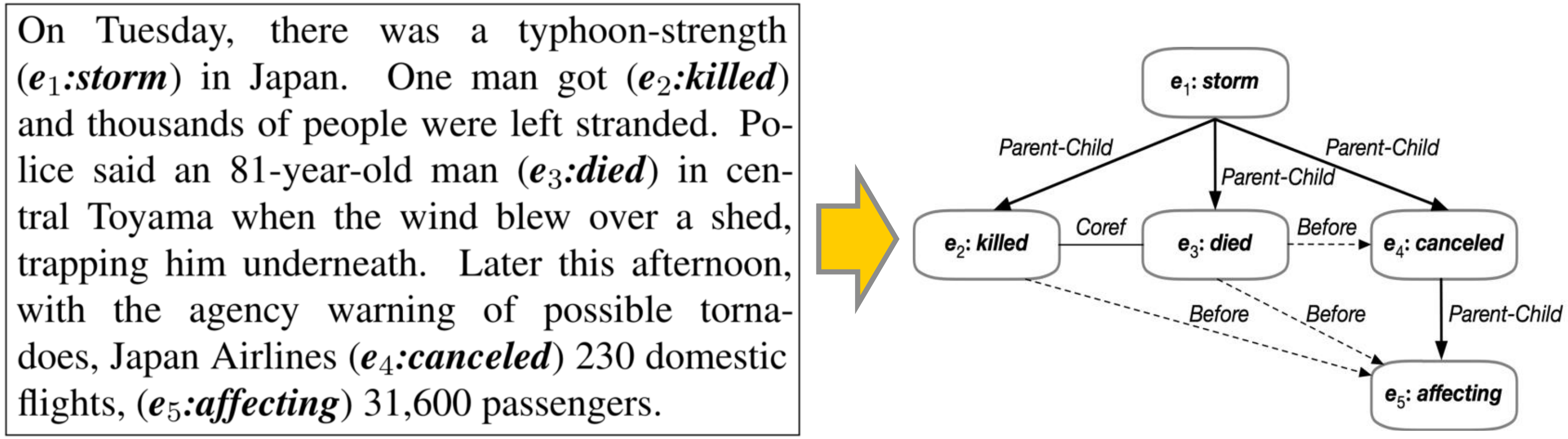
This work provided

- A new (cognitively motivated) task for event understanding, *multi-axis event process typing*, to infer the types of the overall action and affected object(s).
- A large event process dataset with ultra diverse and fine-grained type vocabularies.
- A simple yet effective method of process typing based on indirect supervision from gloss knowledge

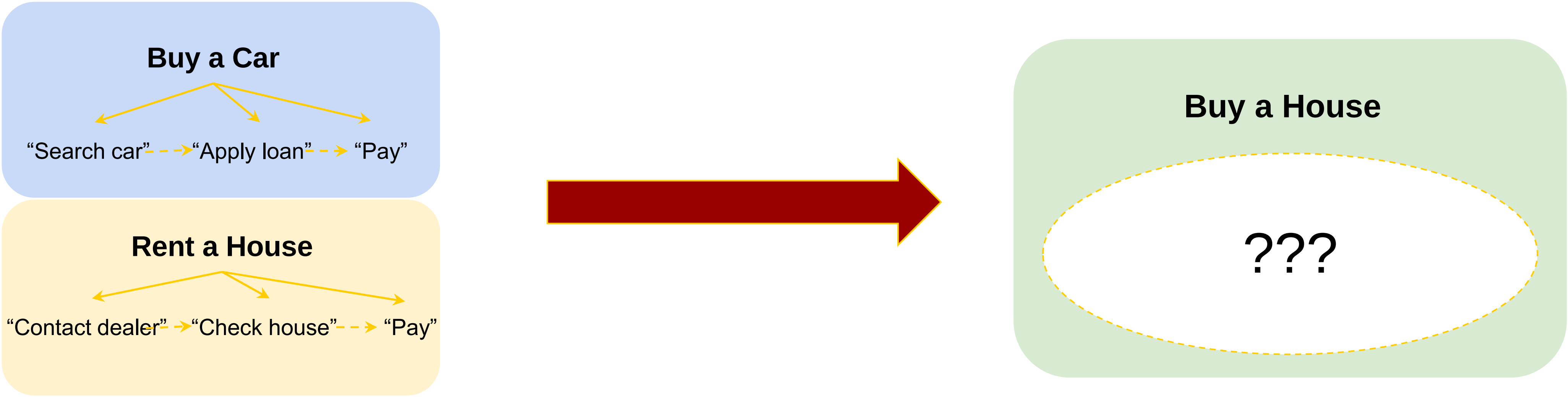
Meaningful future research

- Identifying salient events in processes
- More downstream applications of commonsense reasoning, summarization and narrative prediction
- Event schema induction and instantiation with the produced language model

Our Parallel Works About Event-centric NLU



Haoyu Wang, Muhao Chen, Hongming Zhang, Dan Roth. *Joint Constrained Learning for Event-event Relation Extraction*. **EMNLP** 2020



Hongming Zhang, Muhao Chen, Haoyu Wang, Yangqiu Song, Dan Roth. *Analogous Process Structure Induction for Sub-event Sequence Prediction*. **EMNLP** 2020

Thank You