#### Supersense and Sensibility Proxy Tasks for Semantic Annotation of Prepositions

Luke Gessler, Shira Wein, and Nathan Schneider Georgetown University



#### LAW XIV at COLING'2020

#### **Prepositional Supersense**

A coarse-grained label describing a preposition's meaning

- (1) I rented an apartment in<sub>Locus</sub> Boston
- (2) I hope to see you in<sub>Time</sub> the future
- (3) It's gone downhill since a change **in<sub>Characteristic</sub>** ownership

# Annotating prepositional supersenses

- ... is currently hard and expensive

# Annotating prepositional supersenses

- ... is currently hard and expensive
- SNACS v2.5:
  - 50 supersenses
  - 100-page manual



#### Configuration Identity **Species** Gestalt Possessor Whole Org QuantityItem Characteristic Possession PartPortion - Stuff OrgMember **QuantityValue** - Approximator Ensemble ComparisonRef RateUnit SocialRel

# Crowdsourcing prepositional supersenses

Direct supersense annotation is too
 hard for crowdworkers

He	nas made a bold boardroom choice <b>in</b> ?? silk twill
Plea	se choose the correct supersense of the preposition "in" as used in the sentence above.
$\bigcirc$	Circumstance
ŏ	Temporal
ŏ	Time
ŏ	StartTime
ŏ	EndTime
õ	Frequency
õ	Duration
Õ	Interval
Õ	Locus
Õ	Source
Õ	Goal
Õ	Path
Ō	Direction
Ο	Extent
Ο	Means
Ο	Manner
Ο	Explanation
Ο	Purpose
Ο	Participant
Ο	Causer
Ο	Agent
Ο	Theme
Ο	Topic
Ο	Ancillary
Ο	Stimulus
Ō	Experiencer
Ο	Originator
$\bigcirc$	Recipient

# Crowdsourcing prepositional supersenses

- Direct supersense annotation is too
   hard for crowdworkers
- But: could we have crowdworkers solve a **simpler task**, and **infer** a full supersense label?

He has made a bold boardroom choice <b>in</b> ?? silk twill
Please choose the correct supersense of the preposition "in" as used in the sentence above.
O Circumstance
O Temporal
O Time
O StartTime
O EndTime
O Frequency
O Duration
O Interval
O Locus
O Source
🔘 Goal
O Path
O Direction
O Extent
O Means
O Manner
O Explanation
O Purpose
O Participant
O Causer
O Agent
O Theme
ОТоріс
O Ancillary
O Stimulus
OExperiencer
O Originator
() Recipient

# Crowdsourcing prepositional supersenses

- 2 proxy task designs
  - Intended for implementation on Amazon Mechanical Turk
  - Amazon MT already explored for crowdsourcing senses for content-words
- 4 pilot studies with a handful (5-7) of graduate student crowdworkers
- Goal: assess whether designs can work under ideal conditions

Step 1: crowdworkers generate substitute phrases



Step 1: crowdworkers generate substitute phrases



Step 2: crowdworkers **select** substitute phrases



Step 2: crowdworkers **select** substitute phrases

But	otherwise , it can feel pricey <b>for</b> what you ge
0	with
0	to get
0	having
0	considering
0	given
0	with respect to
0	Other:



Result: a distribution across substitutes, given a preposition as used in a sentence



But otherwise, it can feel pricey for what you get

Result: a **distribution across substitutes**, given a **preposition** as used in a **sentence** which we can use for **supersense inference** 



But otherwise, it can feel pricey for what you get

Distributions for in aggregated by gold supersenses



Distributions for in aggregated by gold supersenses



Distributions for in aggregated by gold supersenses



Step 1: given a target instance

I am **in**<sub>??</sub> love with a giant plate of nachos!

Step 1: given a target instance, retrieve similar instances from a gold corpus

I am **in**<sub>??</sub> love with the giant plate of nachos!

- I am now more

at<sub>Characteristic|Locus</sub> peace and my food craving is about 99% gone

- They are very rude over the phone and in <u>Circumstance</u>[Locus person
- They worked **around<sub>Time</sub>** the clock to ensure that my puppy life was saved.

Step 1: given a **target instance**, retrieve **similar instances** from a **gold corpus** (by taking **supersense tag probability vectors** for the target instance and gold-tagged instances and ranking them)



Step 2: Crowdworkers select an appropriate neighbor, the **majority winner**'s **gold tag** is used for the target instance

	I am <b>in</b> love with the giant plate of nachos ! *
	I am now more <b>at</b> peace and my food craving is about 99 % gone
	They are very rude over the phone and <b>in</b> person .
	They worked <b>around</b> the clock to ensure that my puppy life was saved .
	None None
4	

Step 2: Crowdworkers select an appropriate neighbor, the majority winner's gold tag is used for the target instance

		lams
lan	n in love with the giant plate of nachos ! *	5 respon
	I am now more <b>at</b> peace and my food craving is about 99 % gone	l ar
	They are very rude over the phone and <b>in</b> person .	The
	They worked <b>around</b> the clock to ensure that my puppy life was saved .	Т
	None	

<< love with the giant plate of nachos !



I am in, love with the giant plate of nachos! -

Step 2: Crowdworkers select an appropriate neighbor, the **majority winner**'s **gold tag** is used for the target instance

	I am >>in<< love with the giant plate of nachos !	
I am <b>in</b> love with the giant plate of nachos ! *	5 responses	
<ul> <li>I am now more at peace and my food craving is about 99 % gone</li> <li>They are very rude over the phone and in person .</li> <li>They worked around the clock to ensure that my puppy life was saved .</li> <li>None</li> </ul>	I am now more >>at<< peace and my food They worked >>around<< the clock to ens Clock to ens They are very rude over the phone and > None 0 (0%) 0 1 2 3 4 5	5 (100%
- I am in Characteristic/Locus lo nachos!	ove with the giant plate of	

Two questions

- What if no neighbor has a supersense that matches the target?

I am in (Characteristic Locus) love with the giant plate of nachos!

- Again, a great outing for the kids, a frustration for an **out<sub>Locus</sub>** of town climber
- They are very rude over the phone and in Circumstance Locus person
- They worked **around<sub>Time</sub>** the clock to ensure that my puppy life was saved.
- None

Two questions

- What if no neighbor has a supersense that matches the target?  $\rightarrow$  crowd should select none

I am in (Characteristic Locus) love with the giant plate of nachos!

- Again, a great outing for the kids, a frustration for an **out<sub>Locus</sub>** of town climber
- They are very rude over the phone and in CircumstancelLocus person
- They worked **around**<sub>Time</sub> the clock to ensure that my puppy life was saved.
- None

Two questions

- What if no neighbor has a supersense that matches the target?  $\rightarrow$  crowd should select none
- Is the quality of the neighbors seriously affected if the predicted tag for the target is wrong?

#### I am in Circumstance Locus love with the giant plate of nachos!

- I am now more at peace and my food craving is about 99% gone
- They are very rude over the phone and in Circumstance Locus person
- They worked **around**<sub>Time</sub> the clock to ensure that my puppy life was saved.
- None

Two questions

- What if no neighbor has a supersense that matches the target?  $\rightarrow$  crowd should select none
- Is the quality of the neighbors seriously affected if the predicted tag for the target is wrong?

 $\rightarrow$  hopefully, a neighbor with the appropriate tag will still be retrieved

I am in Circumstance Locus love with the giant plate of nachos!

- I am now more at characteristic peace and my food craving is about 99% gone
- They are very rude over the phone and in Circumstance Locus person
- They worked **around**<sub>Time</sub> the clock to ensure that my puppy life was saved.
- None

Case	Tagger	Crowd	"None"
1 (Tagger correct, gold present)	17/17	17/17	0/17
2 (Tagger incorrect, gold present)	0/12	6/12	5/12
3 (Tagger correct, gold absent)	3/3	0/3	2/3
4 (Tagger incorrect, gold absent)	0/8	0/8	5/8

Pilot study

- 40 instances, 5 workers
- Divide instances for analysis by (1) tag correctness, (2) presence of the target instance's gold tag among neighbors

Case	Tagger	Crowd	"None"
1 (Tagger correct, gold present)	17/17	17/17	0/17
2 (Tagger incorrect, gold present)	0/12	6/12	5/12
3 (Tagger correct, gold absent)	3/3	0/3	2/3
4 (Tagger incorrect, gold absent)	0/8	0/8	5/8

#### Pilot study

- What if no neighbor has a supersense that matches the target?
   → crowd often selects none
- Is the quality of the neighbors seriously affected if the predicted tag for the target is wrong?
  - $\rightarrow$  appropriate neighbors are still retrieved sometimes

# Conclusion

- Under **ideal conditions**, can we obtain **labels** from **proxy tasks**, by having crowdworkers solve an easy task, and then inferring the labels from their output?
  - Yes!
- Future work
  - Implementation on a platform like Amazon Mechanical Turk
  - Different ranking strategies for the neighbor selection design
  - Explore **tradeoffs** between the two designs
  - Application to **different annotation tasks**