An Overview of Commonsense Knowledge Graph Construction and Reasoning at HKUST

Yangqiu Song

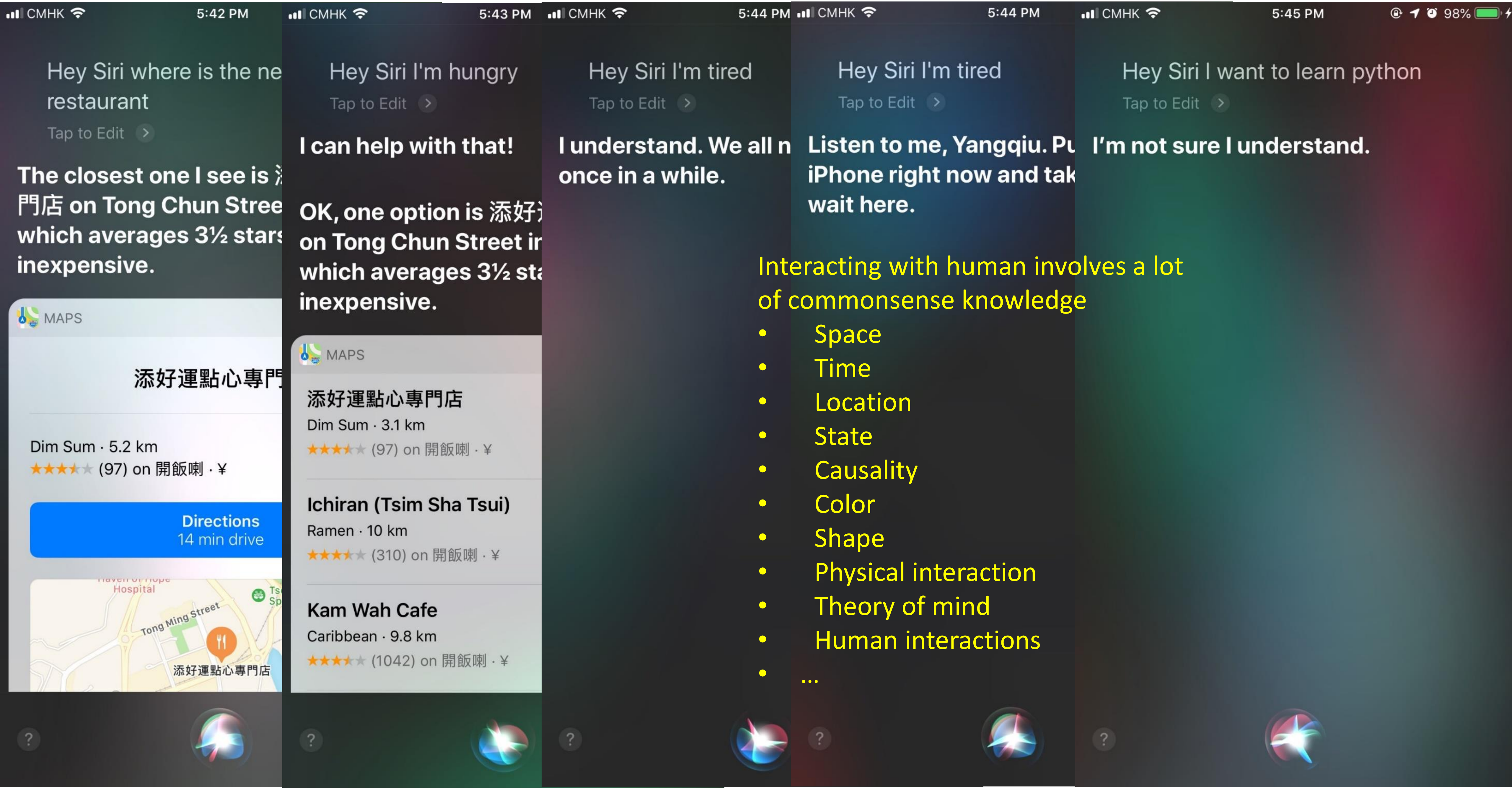
Department of CSE, HKUST, Hong Kong

香 港 科 技 大 學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Understanding human’s language requires complex knowledge

- "Crucial to comprehension is the knowledge that the reader brings to the text. The construction of meaning depends on the reader's knowledge of the language, the structure of texts, a knowledge of the subject of the reading, and a broad-based background or world knowledge.” (Day and Bamford, 1998)
  - Pragmatics: Contexts and knowledge contributes to the meanings

https://www.thoughtco.com/world-knowledge-language-studies-1692508



Interacting with human involves a lot of commonsense knowledge

- Space
- Time
- Location
- State
- Causality
- Color
- Shape
- Physical interaction
- Theory of mind
- Human interactions
- …

Judy Kegl, The boundary between word knowledge and world knowledge, TINLAP3, 1987
Ernie Davis, Building AIs with Common Sense, Princeton Chapter of the ACM, May 16, 2019

Social-Chemistry-101 (UW)

- Understanding law-related documents needs social understanding

“段某某又打电话给糜某2问明其所在位置后，于当晚21时许携带空啤酒瓶，带领袁某（已判决）以及携带三、四十厘米长刀具的易某（已判决）等人来到三期宿舍对面的麻将馆找糜某2。糜某2接到段某某电话后，从麻将馆的厨房找来一把菜刀放在口袋里以备段某某来打架。段某某、袁某、易某等人找到糜某2后，段某某又与糜某2发生言语冲突，继而双方发生打架。糜某2被段某某持啤酒瓶、易某持刀致伤后逃脱，后被糜某1等人送至湖口县中医院进行治疗。”

Maxwell Forbes, Jena D. Hwang, Vered Shwartz, Maarten Sap, Yejin Choi: Social Chemistry 101: Learning to Reason about Social and Moral Norms. EMNLP (1) 2020: 653-670

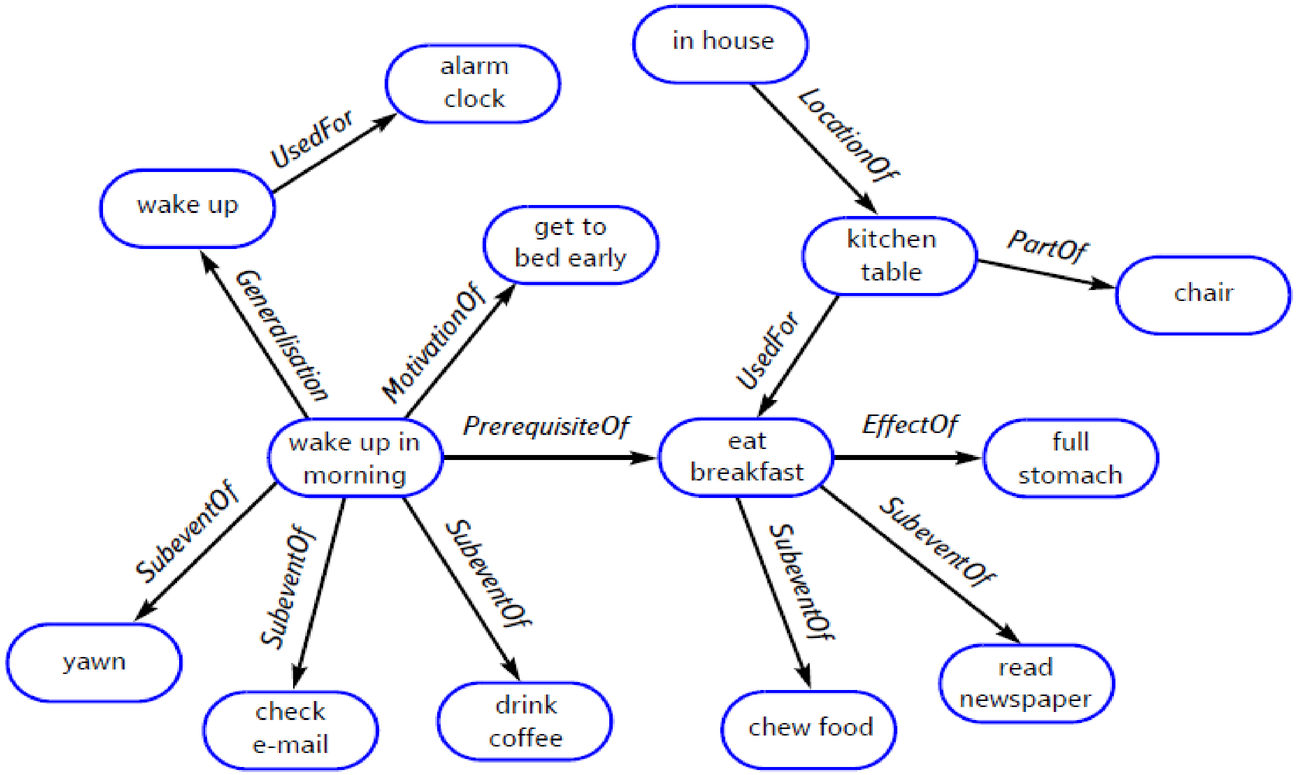
Commonsense Knowledge is the Key

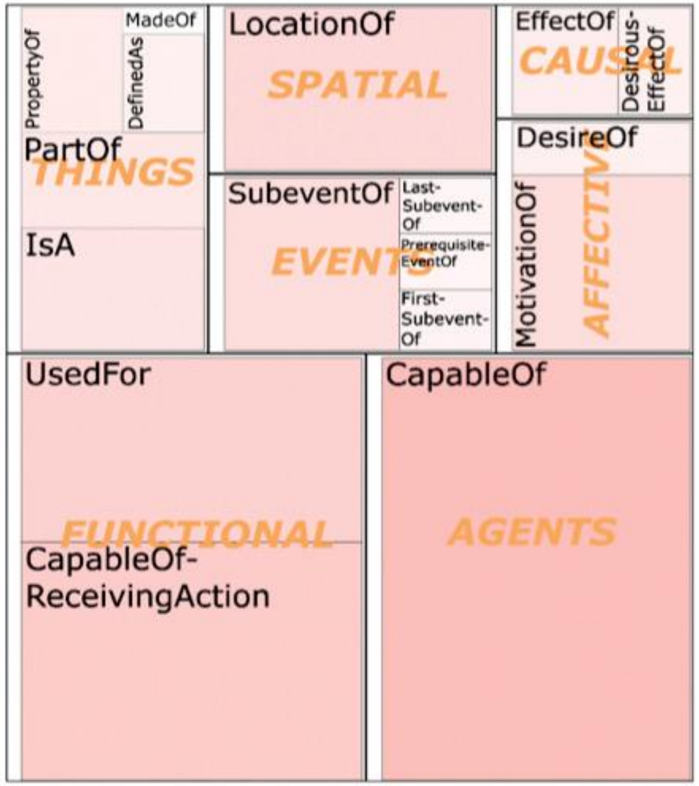
- How to define commonsense knowledge? (Liu & Singh, 2004)
  - “While to the average person the term ‘commonsense’ is regarded as synonymous with ‘good judgement’, ”
  - “in the AI community it is used in a technical sense to refer to the millions of basic facts and understandings possessed by most people.”
  - “Such knowledge is typically omitted from social communications”, e.g.,
    - If you forget someone’s birthday, they may be unhappy with you.

H Liu and P Singh, ConceptNet - a practical commonsense reasoning tool-kit, BTTJ, 2004

How to collect commonsense knowledge?

- ConceptNet5 (Speer and Havasi, 2012)
  - Core is from Open Mind Common Sense (OMCS) (Liu & Singh, 2004)





  - Essentially a crowdsourcing based approach + text mining

The Scale

- "A founder of AI, Marvin Minsky, once estimated that '...commonsense is knowing maybe 30 or 60 million things about the world and having them represented so that when something happens, you can make analogies with others'." (Liu & Singh, 2004)

- ConceptNet
  - 2004: 1.6 million relations among 300,000 nodes
  - 2017: 21 million edges over 8 million nodes
    - 1.5 million nodes are English

Slides credit: Haixun Wang

What contribute to ConceptNet5.5 (21 million edges and over 8 million nodes)?

- Facts acquired from Open Mind Common Sense (OMCS) (Singh 2002) and sister projects in other languages (Anacleto et al. 2006)
- Information extracted from parsing Wiktionary, in multiple languages, with a custom parser (“Wikiparsec”)
- “Games with a purpose” designed to collect common knowledge (von Ahn, Kedia, and Blum 2006) (Nakahara and Yamada 2011) (Kuo et al. 2009)
- Open Multilingual WordNet (Bond and Foster 2013), a linked-data representation ofWordNet (Miller et al. 1998) and its parallel projects in multiple languages
- JMDict (Breen 2004), a Japanese-multilingual dictionary
- OpenCyc, a hierarchy of hypernyms provided by Cyc (Lenat and Guha 1989), a system that represents commonsense knowledge in predicate logic
- A subset of DBPedia (Auer et al. 2007), a network of facts extracted from Wikipedia infoboxes

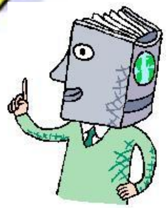
Most of them are entity-centric knowledge, there are only
116,097 edges among
74,989 nodes about
events

Speer, Chin, and Havasi, ConceptNet 5.5: An Open Multilingual Graph of General Knowledge. AAAI 2017.

Most Existing KBs are Entity-centric

- Many large-scale knowledge graphs about entities and their attributes (property-of) and relations (thousands of different predicates) have been developed
  - Millions of entities and concepts
  - Billions of relationships

WIKIPEDIA The Free Encyclopedia · BabelNet · Freebase Labs · NELL · yago select knowledge · DBpedia · ProBase



Google Knowledge Graph (2012)
570 million entities and 18 billion facts

However,

- Semantic meaning in our language can be described as 'a finite set of mental primitives and a finite set of principles of mental combination (Jackendoff, 1990)'.
- The primitive units of semantic meanings include
  - Thing (or Object, Entity, Concept, Instance, etc.),
  - Property,
  - Place,
  - Path,
  - Amount,
  - Activity,
  - State,
  - Event,
  - etc.

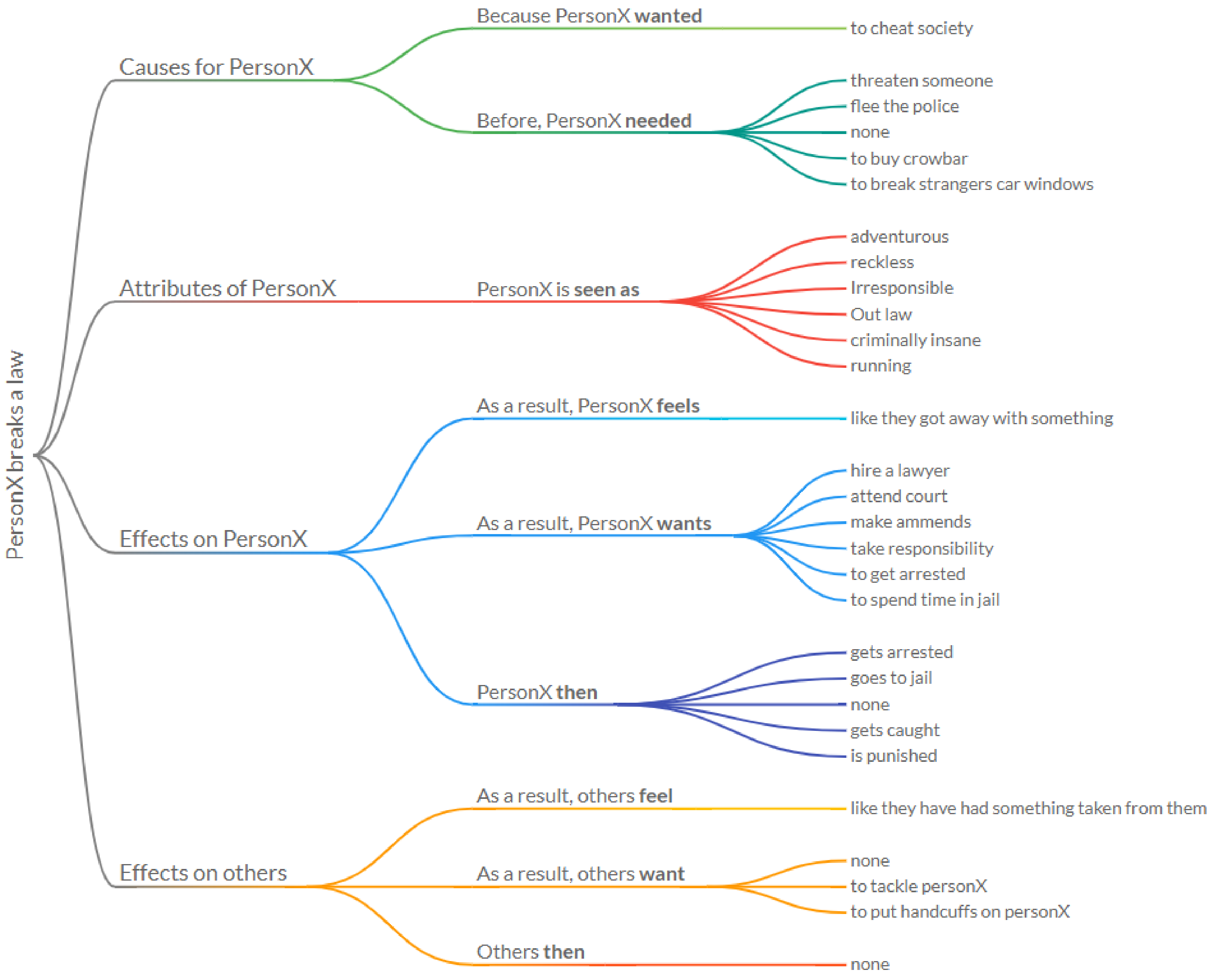
Eventuality

How to collect more knowledge about eventualities rather than entities and relations?

Ray Jackendoff. (Ed.). (1990). Semantic Structures. Cambridge, Massachusetts: MIT Press.

ATOMIC

- Crowdsourcing 9 Types of IF-THEN relations
- All personal entity information has been removed to reduce ambiguity
- Arbitrary texts



Maarten Sap, Ronan LeBras, Emily Allaway, Chandra Bhagavatula, Nicholas Lourie, Hannah Rashkin, Brendan Roof, Noah A. Smith, Yejin Choi: ATOMIC: An Atlas of Machine Commonsense for If-Then Reasoning. AAAI, 2019.

KnowlyWood

- Perform information extraction from free text
  - Mostly movie scripts and novel books
- Four relations: previous, next, parent, similarity
- Only verb+object

Niket Tandon, Gerard de Melo, Abir De, Gerhard Weikum: Knowlywood: Mining Activity Knowledge From Hollywood Narratives. CIKM 2015: 223-232

How to define and scale up the commonsense knowledge acquisition and inference?

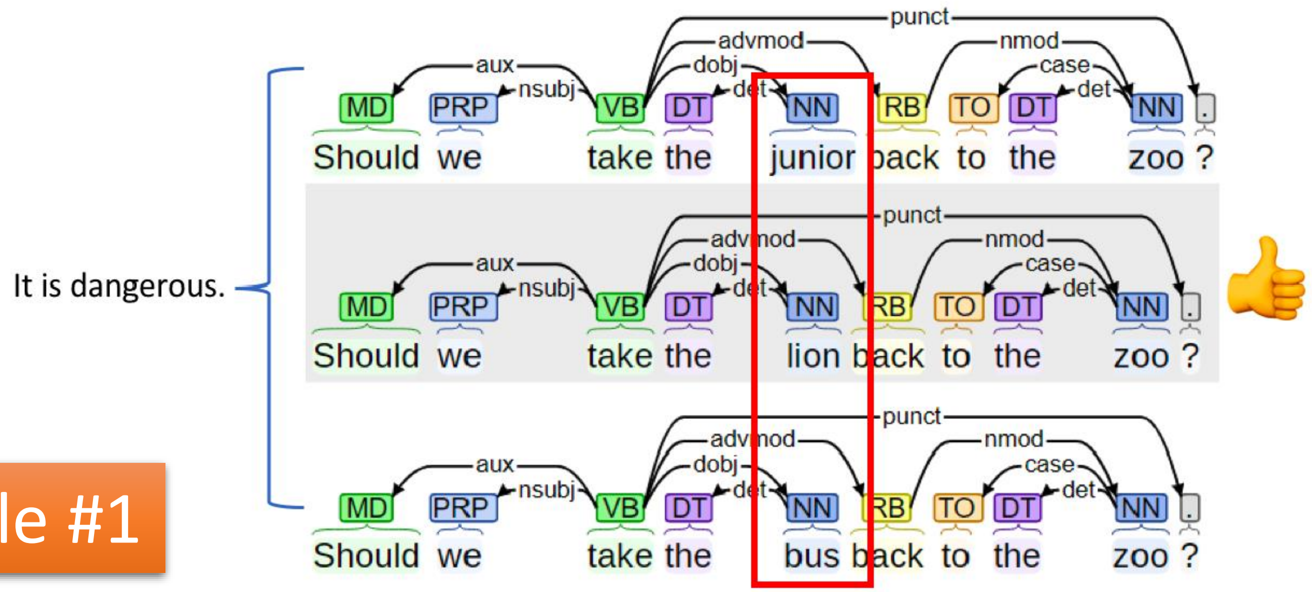
Outline

- Motivation: NLP and commonsense knowledge
- Consideration: selectional preference
- New proposal: large-scale and higher-order selectional preference
- Application on the Winograd Schema Challenge
- Extensions

“Linguistic description – grammar = semantics”
The lower bound of a semantic theory (Katz and Fodor, 1963)

- Disambiguation needs both “the speaker's knowledge of his language and his knowledge about the world” (Katz and Fodor, 1963)
- Compare semantic meanings by fixing grammar
  - Syntactically unambiguous

Principle #1



Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. Language, 39(2), 170–210.

Selectional Preference (SP)

Principle #2

- The need of language inference based on '**partial information** (in John MaCarthy's phrase)' (Wilks, 1975)
  - The soldiers fired at the women, and we saw several of them fall.
  - The needed partial information: hurt things tending to fall down
    - "not invariably true"
    - "tend to be of a very high degree of generality indeed"

(hurt, X) connection (X, fall)

- Selectional preference (Resnik, 1993)
  - A relaxation of selectional restrictions (Katz and Fodor, 1963) and as syntactic features (Chomsky, 1965)
  - Applied to isA hierarchy in WordNet and verb-object relations

Yorick Wilks. 1975. An intelligent analyzer and understander of English. Communications of the ACM, 18(5):264–274.
Katz, J. J., & Fodor, J. A. (1963). The structure of a semantic theory. Language, 39(2), 170–210.
Noam Chomsky. 1965. Aspects of the Theory of Syntax. MIT Press, Cambridge, MA.
Philip Resnik. 1993. Selection and information: A class-based approach to lexical relationships. Ph.D. thesis, University of Pennsylvania.

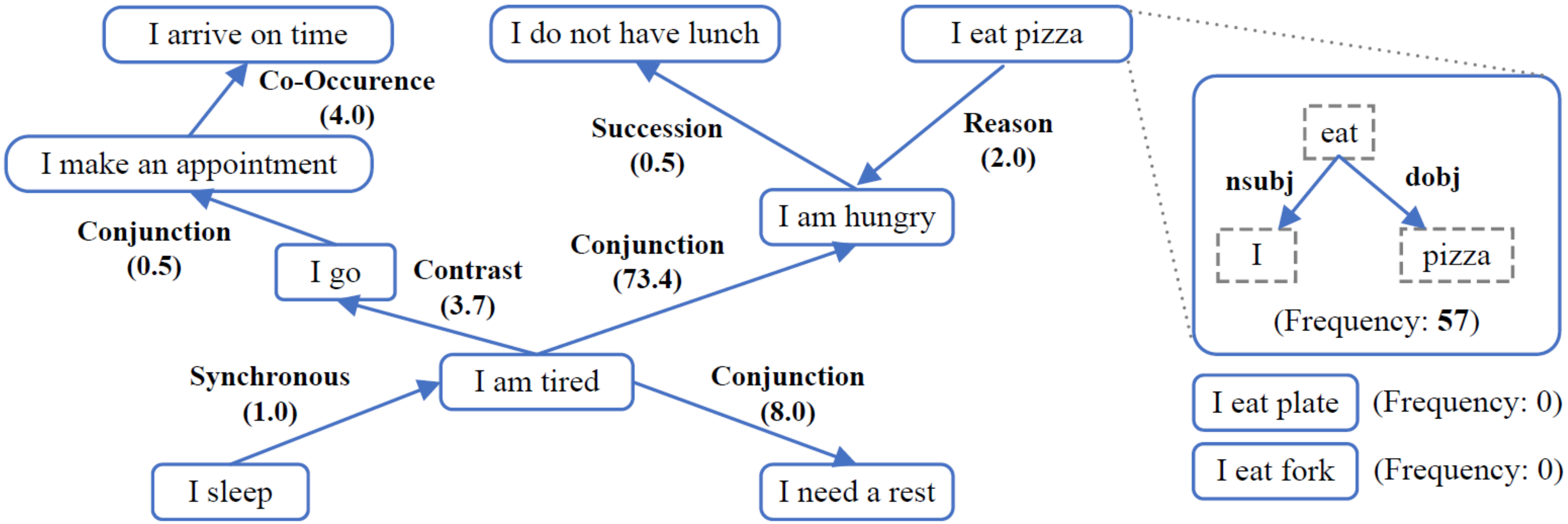
Outline

- Motivation: NLP and commonsense knowledge
- Consideration: selectional preference
- New proposal: large-scale and higher-order selectional preference
- Extensions

A New Eventuality Knowledge Graph: ASER

Activities, **S**tates, **E**vents, and their **R**elations

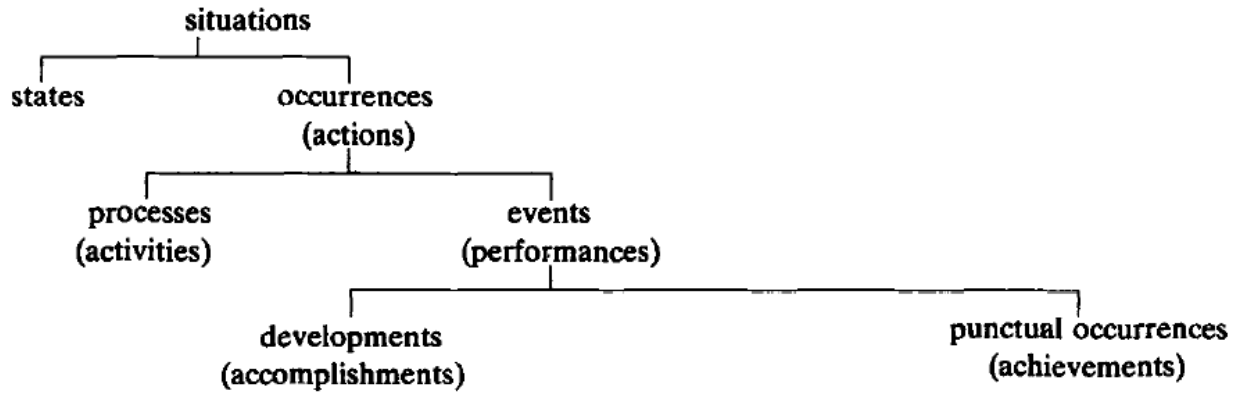
- Use verb-centric patterns from dependency parsing
  - Principle #1: to compare semantics by fixing syntax (Katz and Fodor, 1963)
- Maintain a set of key tags and a set of auxiliary tags
  - Principle #2: to obtain frequent 'partial information' (Wilks, 1975)



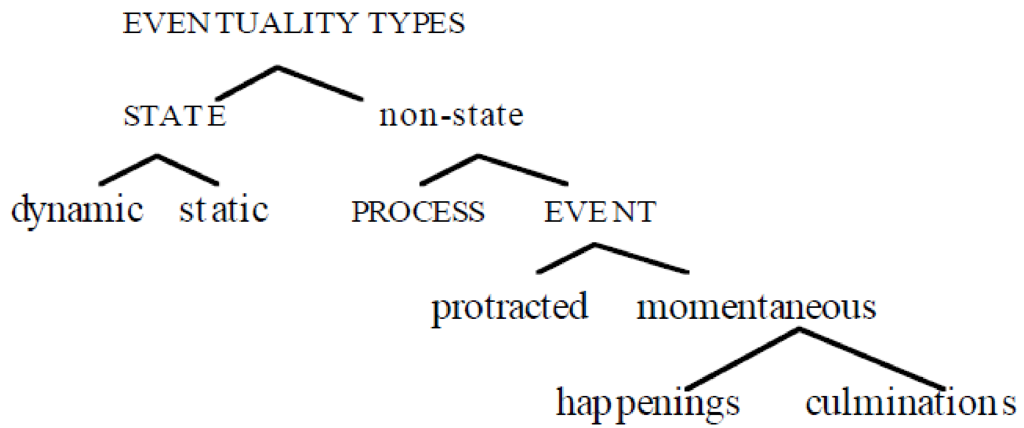
A hybrid graph of
- Each eventuality is a hyper-edge of words
- Heterogeneous edges among eventualities

ASER
Activities, **S**tates, **E**vents, and their **R**elations

Mourelatos' taxonomy (1978)



Bach's taxonomy (1986)

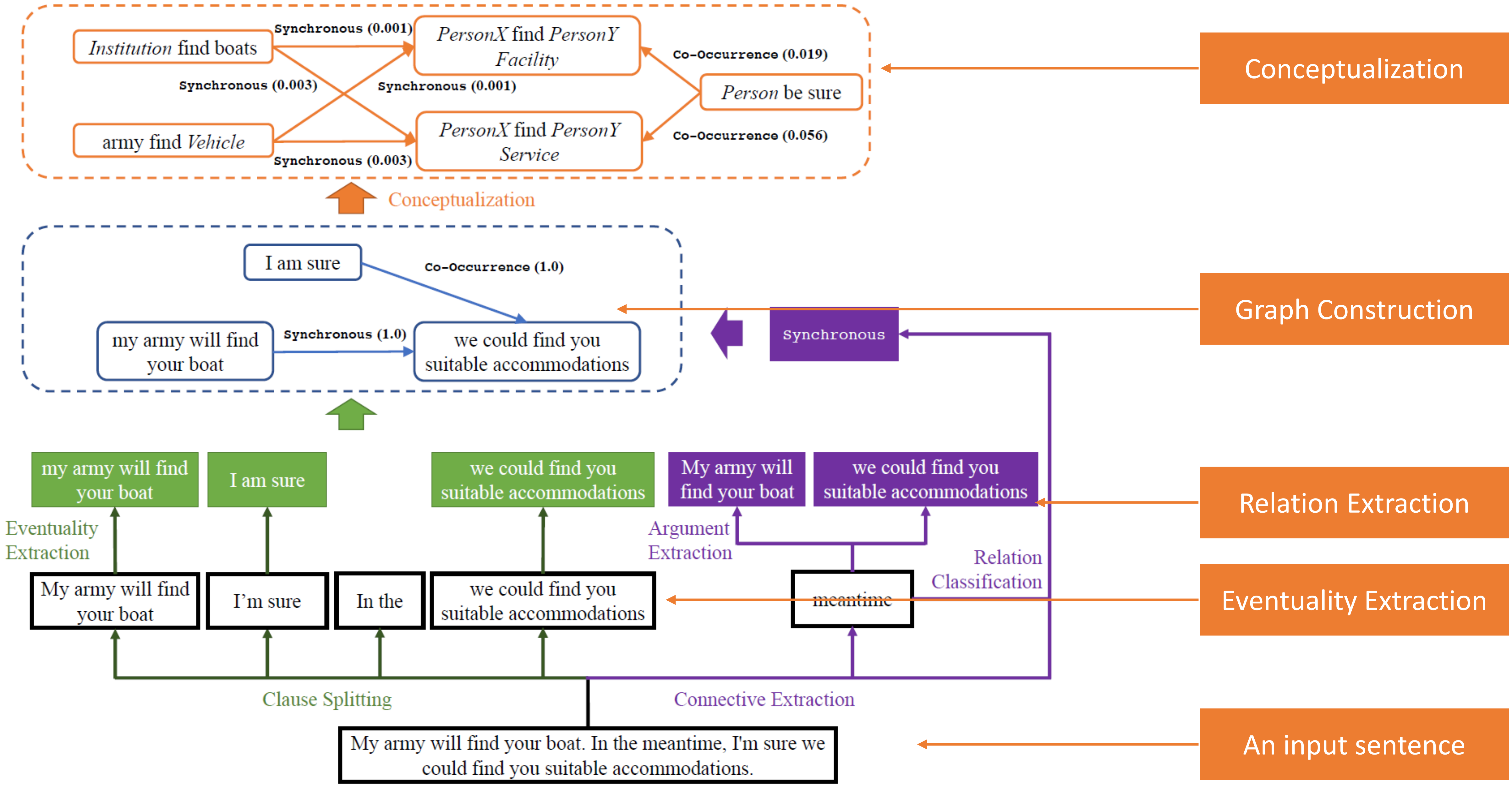


- **State**: The air smells of jasmine.
- **Process**: It's snowing.
- **Development**: The sun went down.
- **Punctual occurrence**: The cable snapped. He blinked. The pebble hit the water.

- **Static states**: be in New York, love (one's cat);
- **Dynamic states**: sit, stand, drunk, present, sick;
- **Processes**: walk, push a cart, sleep;
- **Protracted events**: build (a cabin), eat a sandwich, polish a shoe, walk to Boston;
- **Culminations**: take off; arrive, leave, depart;
- **Happenings**: blink, flash, knock, kick, hit, pat, wink;

Alexander P. D. Mourelatos. Events, processes, and states. Linguistics and Philosophy, 2, 415-434. 1978.
Emmon Bach. The algebra of events. Linguistics and philosophy, 9 (1), 5-16. 1986.

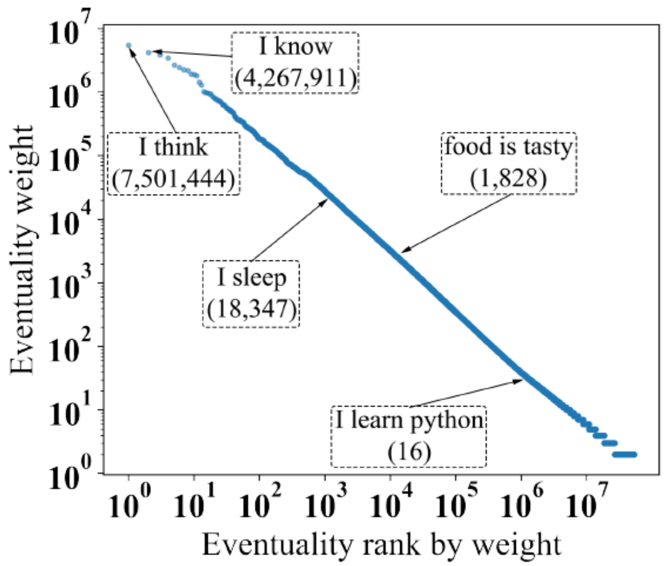
A Running Example



Eventualities

- Using patterns to collect partial information
- Six relations are also kept but treated as auxiliary edges
  - advmod,
  - amod,
  - nummod,
  - aux,
  - compound,
  - neg

Pattern	Code	Example
n1-nsubj-v1	s-v	`The dog barks'
n1-nsubj-v1-dobj-n2	s-v-o	`I love you'
n1-nsubj-v1-xcomp-a	s-v-a	`He felt ill'
n1-nsubj-(v1-iobj-n2)-dobj-n3	s-v-o-o	`You give me the book'
n1-nsubj-a1-cop-be	s-be-a	`The dog is cute'
n1-nsubj-v1-xcomp-a1-cop-be	s-v-be-a	`I want to be slim'
n1-nsubj-v1-xcomp-n2-cop-be	s-v-be-o	`I want to be a hero'
n1-nsubj-v1-xcomp-v2-dobj-n2	s-v-v-o	`I want to eat the apple'
n1-nsubj-v1-xcomp-v2	s-v-v	`I want to go'
(n1-nsubj-a1-cop-be)-nmod-n2-case-p1	s-be-a-p-o	`It' cheap for the quality'
n1-nsubj-v1-nmod-n2-case-p1	s-v-p-o	`He walks into the room'
(n1-nsubj-v1-dobj-n2)-nmod-n3-case-p1	s-v-o-p-o	`He plays football with me'
n1-nsubjpass-v1	spass-v	`The bill is paid'
n1-nsubjpass-v1-nmod-n2-case-p1	spass-v-p-o	`The bill is paid by me'



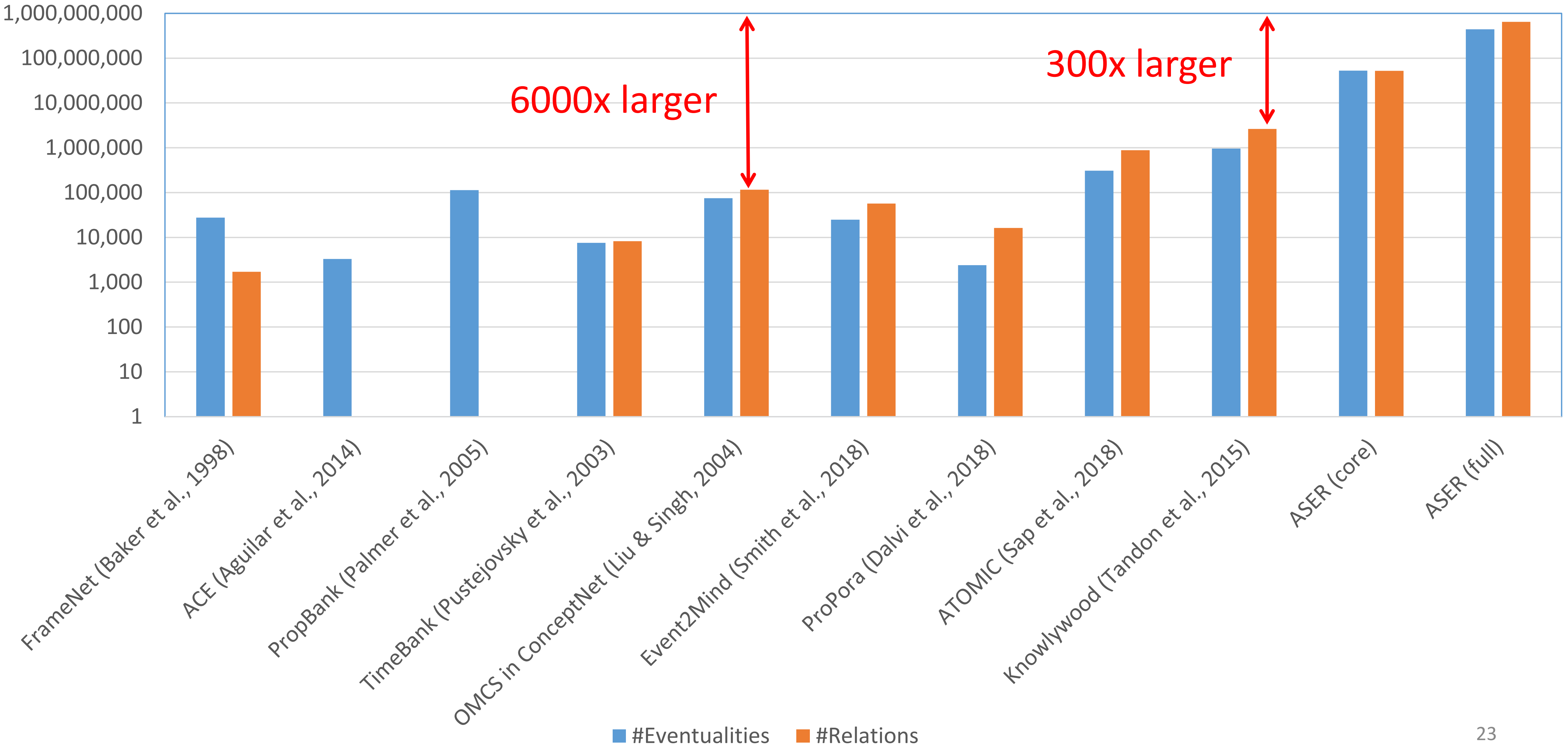
Eventuality Relations

- 14 relations taking from CoNLL shared task
  - More frequent relations
- Less ambiguous connectives
  - 'so that' 31 times only in 'Result' relations
- Some are ambiguous
  - 'while': Conjunction 39 times, Contrast 111 times, Expectation 79 times, and Concession 85 times
- Classifiers trained on Penn Discourse Treebank (PDTB) (Prasad et al., 2007)

Relation Type	Examples
Precedence	E1 **before** E2; E1 , **then** E2; E1 **till** E2; E1 **until** E2
Succession	E1 **after** E2; E1 **once** E2
Synchronous	E1, **meanwhile** E2; E1 **meantime** E2; E1, **at the same time** E2
Reason	E1, **because** E2
Result	E1, **so** E2; E1, **thus** E2; E1, **therefore** E2; E1, **so that** E2
Condition	E1, **if** E2; E1, **as long as** E2
Contrast	E1, **but** E2; E1, **however** E2; E1**, by contrast** E2; E1, **in contrast** E2; E1 , **on the other hand**, E2; E1, **on the contrary**, E2
Concession	E1, **although** E2
Conjunction	E1 **and** E2; E1, **also** E2
Instantiation	E1, **for example** E2; E1, **for instance** E2
Restatement	E1, **in other words** E2
Alternative	E1 **or** E2; E1, **unless** E2; E1, **as an alternative** E2; E1, **otherwise** E2
ChosenAlternative	E1, E2 **instead**
Exception	E1, **except** E2

Prasad, R., Miltsakaki, E., Dinesh, N., Lee, A., Joshi, A., Robaldo, L., & Webber, B. L. (2007). The penn discourse treebank 2.0 annotation manual.
Nianwen Xue, Hwee Tou Ng, Sameer Pradhan, Rashmi Prasad, Christopher Bryant, Attapol T. Rutherford. The CoNLL-2015 Shared Task on Shallow Discourse Parsing.

Scales of Verb Related Knowledge Graphs



Partial Information Aggregation

- “hurt things tending to fall down”

(hurt, X) connection (X, fall)

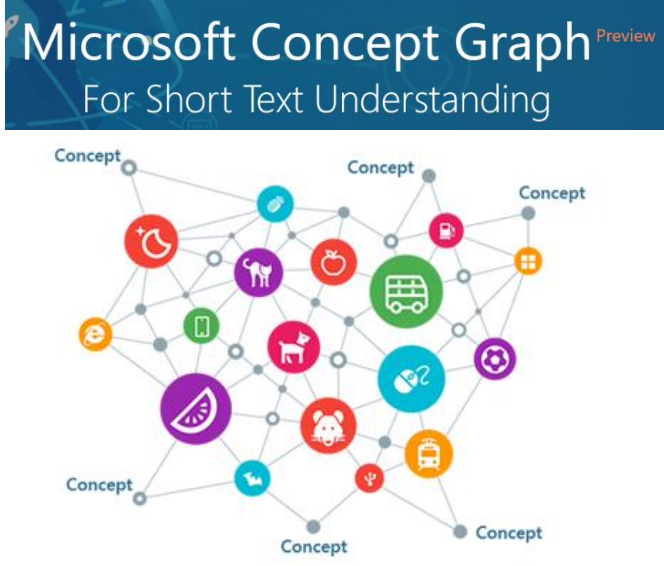
- “stocks price may increase when a company acquires a start-up”

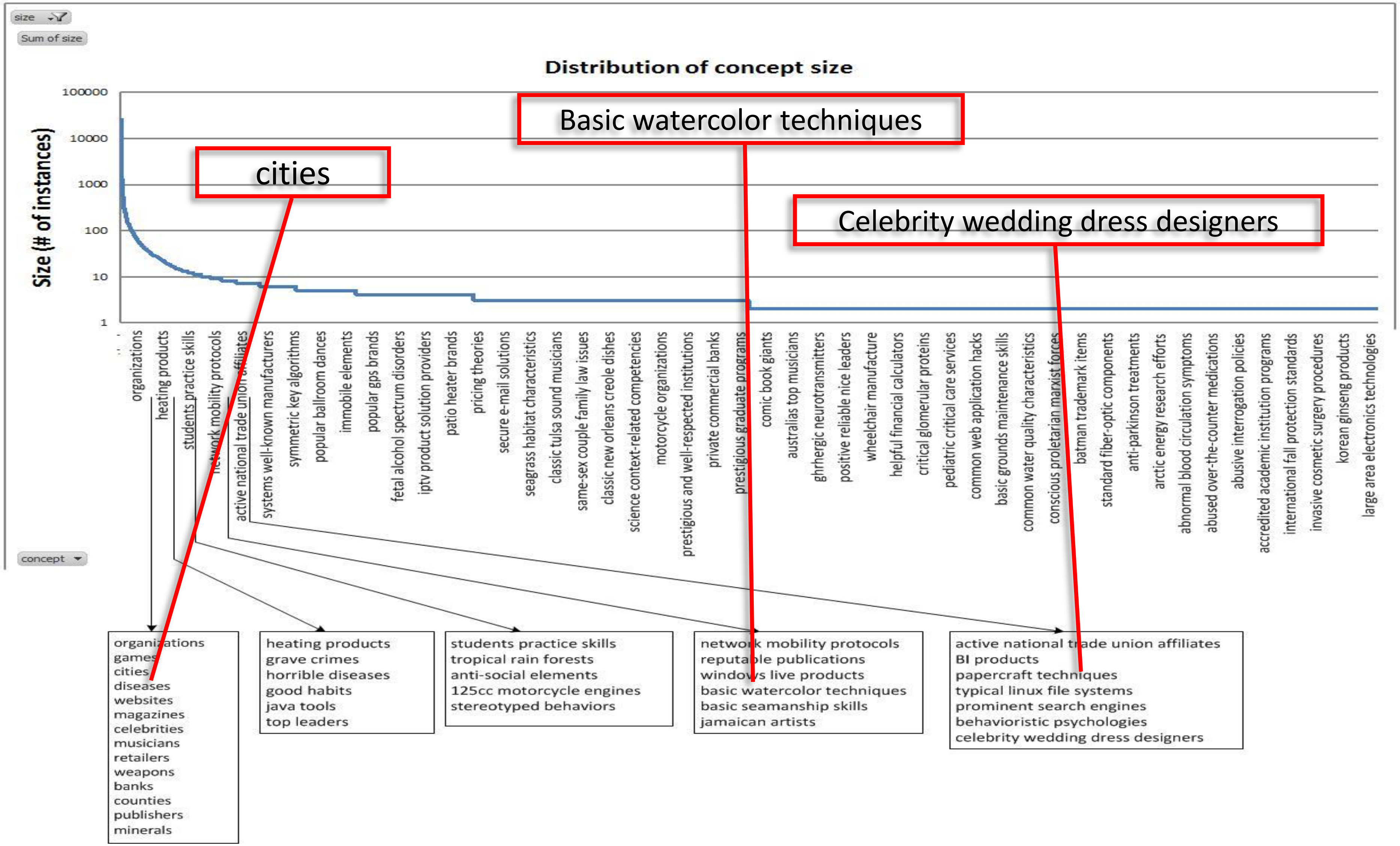
(company, acquire, start-up) result-in (stock, increase)

Normalization

			Probability
He, she, I, Bob, …	→	PERSON	1.0
1996, 2020, 1949, …	→	YEAR	1.0
23, 20, 333, ….	→	DIGIT	1.0
www.google.com, …	→	URL	1.0

Conceptualization with ProBase





Concepts are the glue that holds our mental world together.

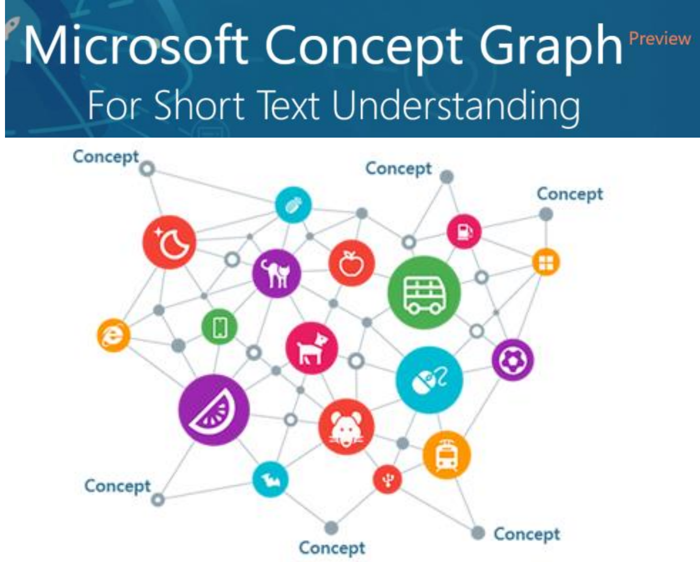
Gregory L. Murphy
NYU

Probase is a *large, universal, probabilistic* knowledge base with **an extremely large concept space**

Data are available at https://concept.research.microsoft.com/

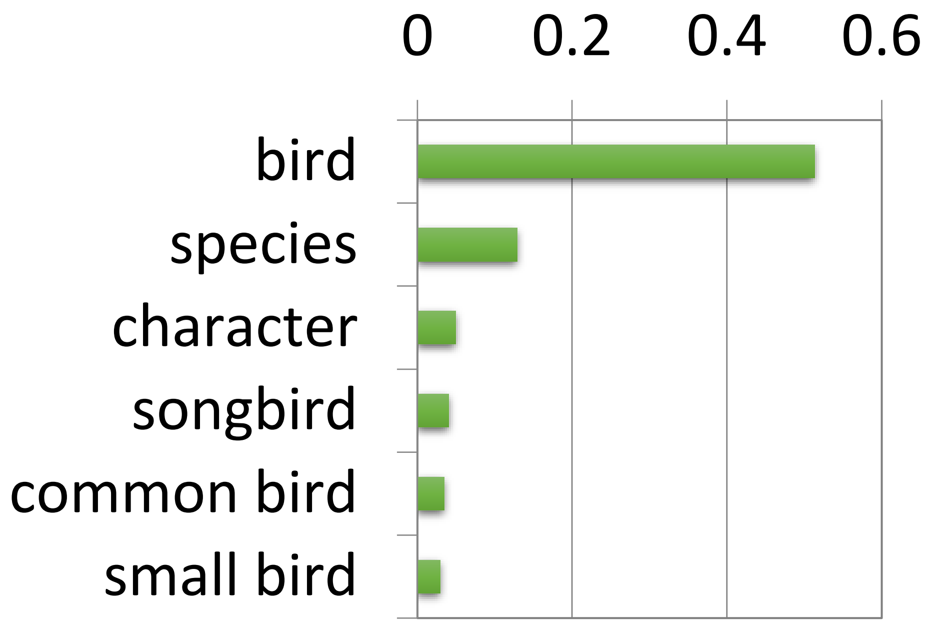
Wentao Wu, Hongsong Li, Haixun Wang, Kenny Qili Zhu: Probase: a probabilistic taxonomy for text understanding. SIGMOD Conference 2012: 481-492

Conceptualization with ProBase

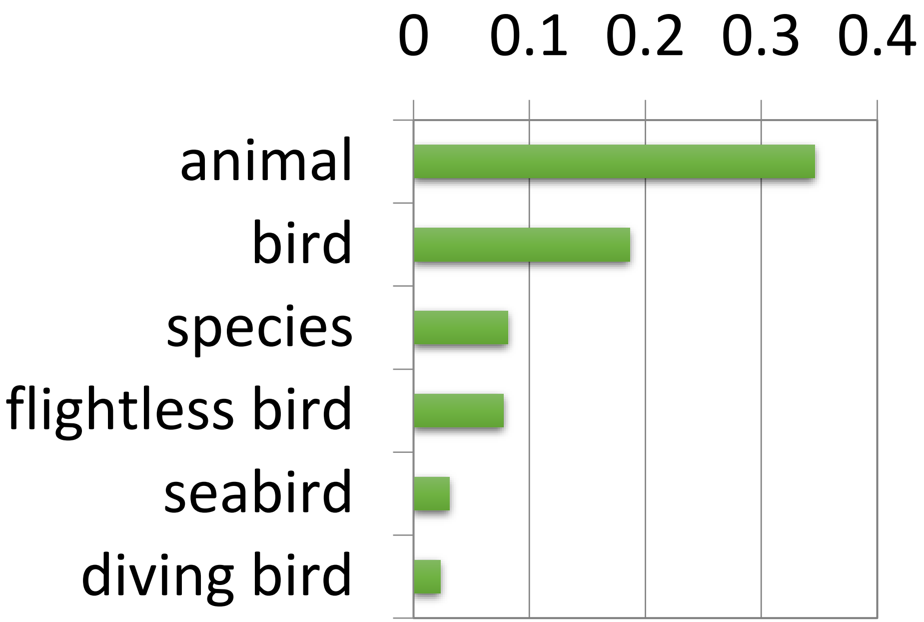


Typicality $$P(concept \mid instance) = \frac{\#(concept, instance)}{\#(instance)}$$

- Robin



- Penguin



Data are available at https://concept.research.microsoft.com/

Wentao Wu, Hongsong Li, Haixun Wang, Kenny Qili Zhu: Probase: a probabilistic taxonomy for text understanding. SIGMOD Conference 2012: 481-492

Yangqiu Song, Haixun Wang, Zhongyuan Wang, Hongsong Li, Weizhu Chen: Short Text Conceptualization Using a Probabilistic Knowledgebase. IJCAI 2011: 2330-2336

A Running Example

Obama

(politician, 0.0855)
(democrat, 0.0560)
(liberal, 0.0560)

dog

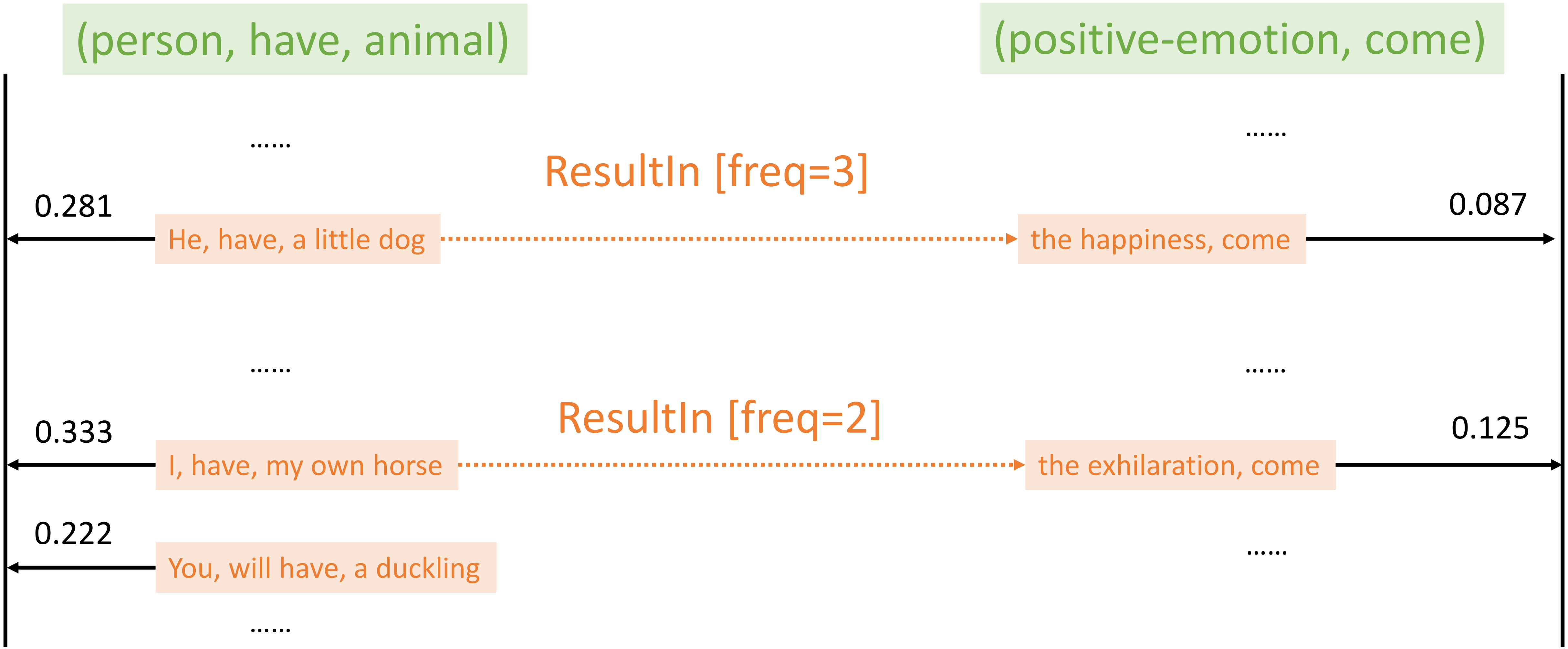
(animal, 0.2811)
(pet, 0.1377)
(domestic animal, 0.0525)

(Obama, have, dog)

(obama have animal, 0.2811)
(obama have pet, 0.1377)
(politician have dog, 0.0855)
(democrat have dog, 0.05604)
...
(politician have animal, 0.0240)
(democrat have animal, 0.01575)
...

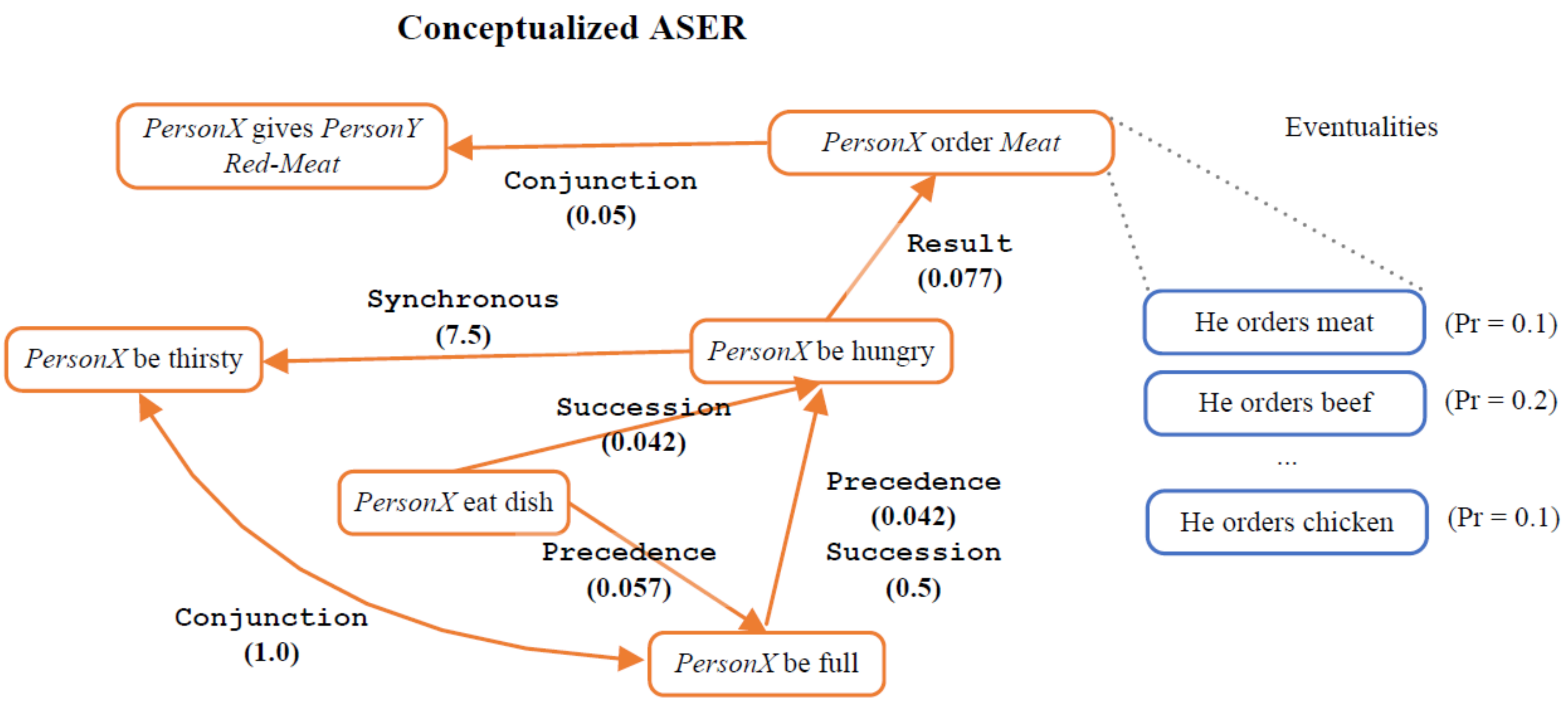
$$\prod_{i=1}^{N} P\left(C_{i,k} \middle| E_i\right)$$

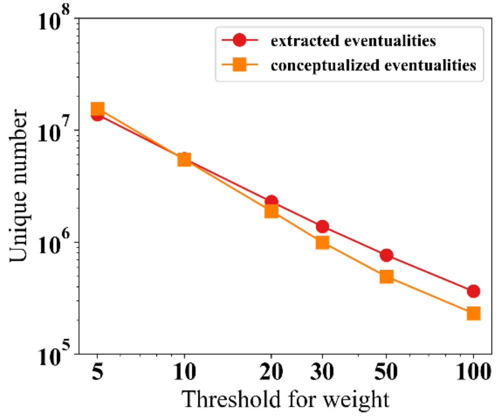
$$P(politician \mid Obama) \times P(animal \mid dog) = 0.0855 \times 0.2811 = 0.0240$$

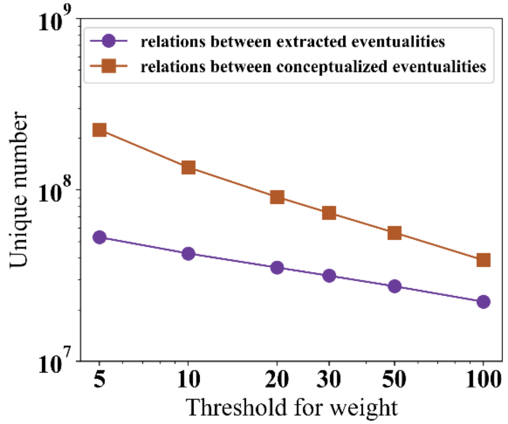


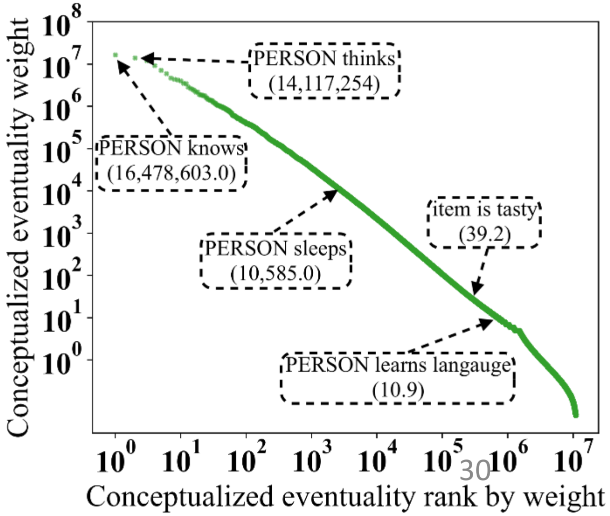
P(ResultIn | (person, have, animal) , (positive-emotion, come)) $= 0.281 \times 3 \times 0.087 + 0.333 \times 2 \times 0.125$
$= 0.157$

Conceptualization Results









ASER 2.0

- 1.0: Rule based extraction (14 Eventuality Patterns, Improved Version)

Data	#Unique Eventualities	#Unique Relations
Core	34 millions	15 millions
Full	272 millions	206 millions

- 2.0: Discourse Parser (18 Eventuality Patterns + Wang and Lan 2015)

Data	#Unique Eventualities	#Unique Relations
Core	53 millions	52 millions
Full	439 millions	649 millions

- Conceptualization Core (threshold=5):
  - Concepts: 15 millions (based on 14 millions eventualities, 1.X times)
  - Concept Relations: 224 millions (based on 53 millions eventuality relations, 4.X times)

Jianxiang Wang and Man Lan. A Refined End-to-End Discourse Parser. CONLL Shared Task 2015.

Graph Inference Examples

- One hop relations $\Pr(E_t|E_h, T_1) = \frac{w^{(r)}_{\langle E_h, T_1, E_t \rangle}}{\sum_{E'_t, s.t., (E_h, T_1, E'_t) \in \mathcal{R}} w^{(r)}_{\langle E_h, T_1, E'_t \rangle}},$

- Eventualities
  - ⟨"I drink coffee", `Reason`, "I enjoy the flavor"⟩
  - ⟨"You go to restaurant", `Precedence`, "You got sick"⟩
  - ⟨"It is a cat", `Condition`, "It is a tiger"⟩

- Concepts
  - ⟨"*Company* be *Stakeholder-Group*", `Condition`, "*PersonX* be successful"⟩
  - ⟨"*PersonX* hurt *Insect*", `Condition`, "*PersonX* help *Insect*"⟩
  - ⟨"*PersonX* be *Emotion*", `Succession`, "*PersonX* marry"⟩

Rule Mining: Eventualities

- Mine Rules using AIME+ $< E_a, T_1, E_b > \wedge < E_b, T_2, E_c > \Rightarrow < E_a, T_3, E_b >,$

Rule	$\langle E_b \xrightarrow{\texttt{Concession}} E_f \rangle \wedge \langle E_a \xrightarrow{\texttt{Result}} E_f \rangle \Rightarrow \langle E_a \xrightarrow{\texttt{Contrast}} E_b \rangle$	Concession	E1, **although** E2
Instances	⟨ I do not know → I guess ⟩ ∧ ⟨ I believe → I guess ⟩ ⇒ ⟨ I believe → I do not know ⟩ ⟨ I am not sure → I guess ⟩ ∧ ⟨ I hope so → I guess ⟩ ⇒ ⟨ I hope so → I am not sure ⟩ ⟨ I understand → I can not speak ⟩ ∧ ⟨ I am not a lawyer → I can not speak ⟩ ⇒ ⟨ I am not a lawyer → I understand ⟩		
Rule	$\langle E_f \xrightarrow{\texttt{Contrast}} E_b \rangle \wedge \langle E_a \xrightarrow{\texttt{Instantiation}} E_f \rangle \Rightarrow \langle E_a \xrightarrow{\texttt{Contrast}} E_b \rangle$		
Instances	⟨ I remember → I could not find it ⟩ ∧ ⟨ I get → I remember ⟩ ⇒ ⟨ I get → I could not find it ⟩ ⟨ I would say → I might be wrong ⟩ ∧ ⟨ I hope → I would say ⟩ ⇒ ⟨ I hope → I might be wrong ⟩ ⟨ It have been suggested → This is unlikely ⟩ ∧ ⟨ It is possible → It have been suggested ⟩ ⇒ ⟨ It is possible → This is unlikely ⟩		
Rule	$\langle E_e \xrightarrow{\texttt{ChosenAlternative}} E_b \rangle \wedge \langle E_a \xrightarrow{\texttt{ChosenAlternative}} E_e \rangle \Rightarrow \langle E_a \xrightarrow{\texttt{ChosenAlternative}} E_b \rangle$	ChosenAlternative	E1, E2 **instead**
Instances	⟨ I will not go → You come here ⟩ ∧ ⟨ I want to see → I will not go ⟩ ⇒ ⟨ I want to see → You come here ⟩ ⟨ I want → It is ⟩ ∧ ⟨ I wish → I want ⟩ ⇒ ⟨ I wish → It is ⟩ ⟨ I want → I get ⟩ ∧ ⟨ I do not get that → I want ⟩ ⇒ ⟨ I do not get that → I get ⟩		

Rule Mining: Concepts

- Mine Rules using AIME+ $< E_a, T_1, E_b > \wedge < E_b, T_2, E_c > \Rightarrow < E_a, T_3, E_b >,$

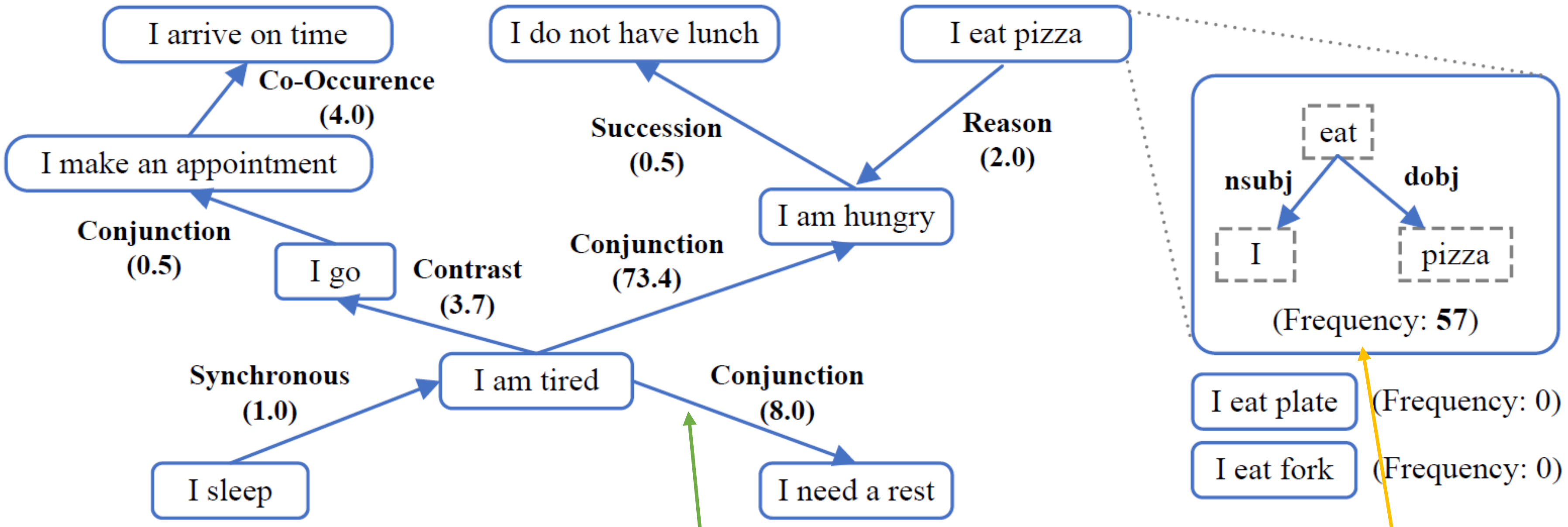
Rule	$\langle E_e \xrightarrow{\texttt{Restatement}} E_a \rangle \wedge \langle E_e \xrightarrow{\texttt{Restatement}} E_b \rangle \Rightarrow \langle E_a \xrightarrow{\texttt{Conjunction}} E_b \rangle$
Instances	⟨ *PersonX* laugh → *PersonX* smile ⟩ ∧ ⟨ *PersonX* laugh → *PersonX* open *Facial-Feature* ⟩ ⇒ ⟨ *PersonX* smile → *PersonX* open *Facial-Feature* ⟩ ⟨ *PersonX* love it → It be good ⟩ ∧ ⟨ *PersonX* love it → It be tasty ⟩ ⇒ ⟨ It be good → It be tasty ⟩ ⟨ *PersonX* wish → *PersonX* need ⟩ ∧ ⟨ *PersonX* wish → *PersonX* need ⟩ ⇒ ⟨ *PersonX* need → *PersonX* need ⟩
Rule	$\langle E_e \xrightarrow{\texttt{Instantiation}} E_a \rangle \wedge \langle E_e \xrightarrow{\texttt{Instantiation}} E_b \rangle \Rightarrow \langle E_a \xrightarrow{\texttt{Conjunction}} E_b \rangle$
Instances	⟨ *PersonX* realize → *PersonX* point out ⟩ ∧ ⟨ *PersonX* realize → PersonX have *Information* ⟩ ⇒ ⟨ *PersonX* point out → *PersonX* have *Information* ⟩ ⟨ *PersonX* have → *PersonX* get ⟩ ∧ ⟨ *PersonX* have → *PersonX* own ⟩ ⇒ ⟨ *PersonX* get → *PersonX* own ⟩ ⟨ *PersonX* know → *PersonX* be sure ⟩ ∧ ⟨ *PersonX* know → *PersonX* remember ⟩ ⇒ ⟨ *PersonX* be sure → *PersonX* remember ⟩
Rule	$\langle E_e \xrightarrow{\texttt{Concession}} E_b \rangle \wedge \langle E_e \xrightarrow{\texttt{Restatement}} E_a \rangle \Rightarrow \langle E_a \xrightarrow{\texttt{Contrast}} E_b \rangle$
Instances	⟨ *PersonX* order *Dish* → *PersonX* be hungry ⟩ ∧ ⟨ *PersonX* order *Dish* → *PersonX* order ⟩ ⇒ ⟨ *PersonX* order → *PersonX* be hungry ⟩ ⟨ *PersonX* wish → *PersonX* doubt ⟩ ∧ ⟨ *PersonX* wish → *PersonX* need ⟩ ⇒ ⟨ *PersonX* doubt → *PersonX* need ⟩ ⟨ *PersonX* love it → *PersonX* hate it ⟩ ∧ ⟨ *PersonX* love it → It be good ⟩ ⇒ ⟨ *PersonX* hate it → It be good ⟩

Instantiation	E1, **for example** E2; E1, **for instance** E2
Restatement	E1, **in other words** E2

Outline

- Motivation: NLP and commonsense knowledge
- Consideration: selectional preference
- New proposal: large-scale and higher-order selectional preference
- Extensions
  - Transform to ConceptNet
  - Transform to ATOMIC

ASER is Essentially a Knowledge Graph based on Linguistics



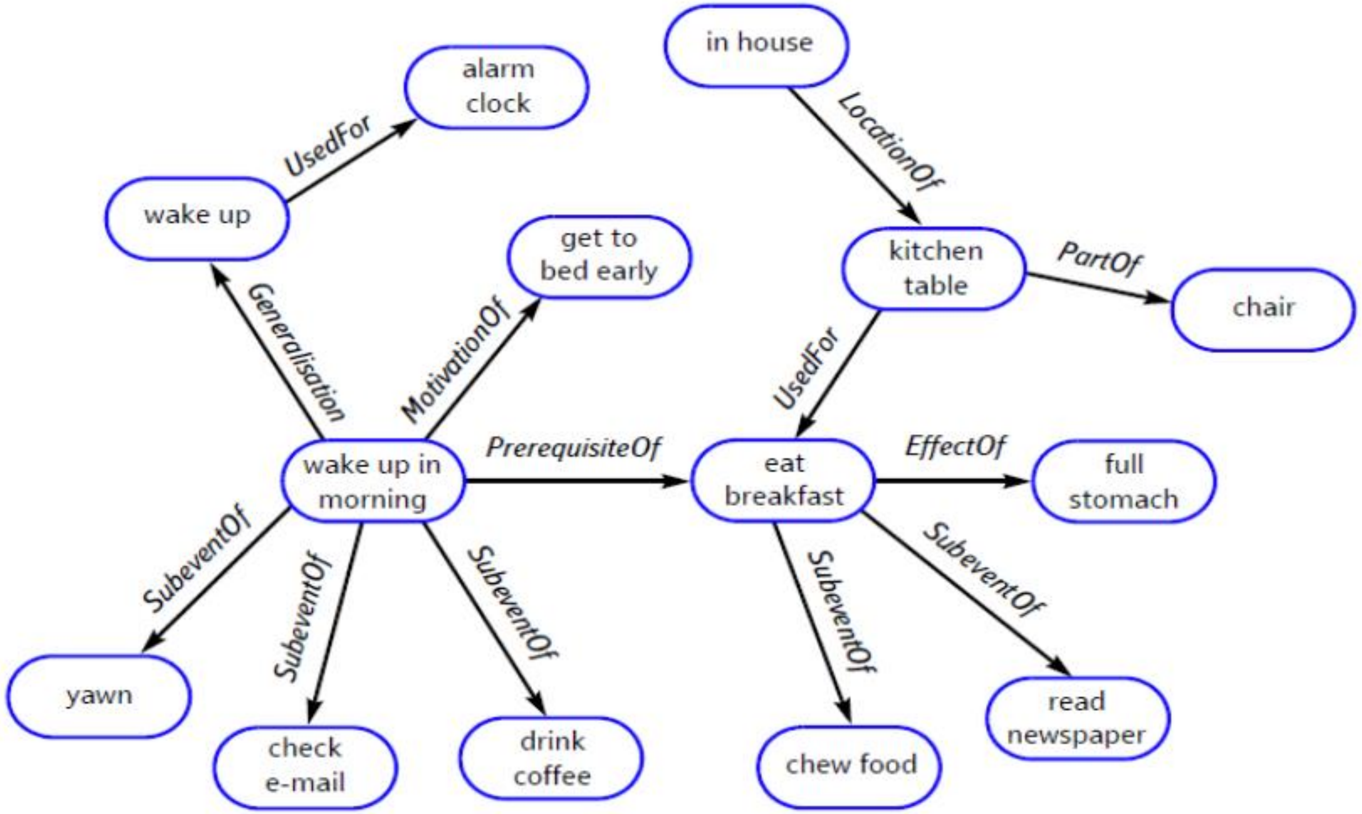
Discourse Relation

Dependency Relation

How is it transferrable from linguistic knowledge to existing definition of commonsense knowledge?

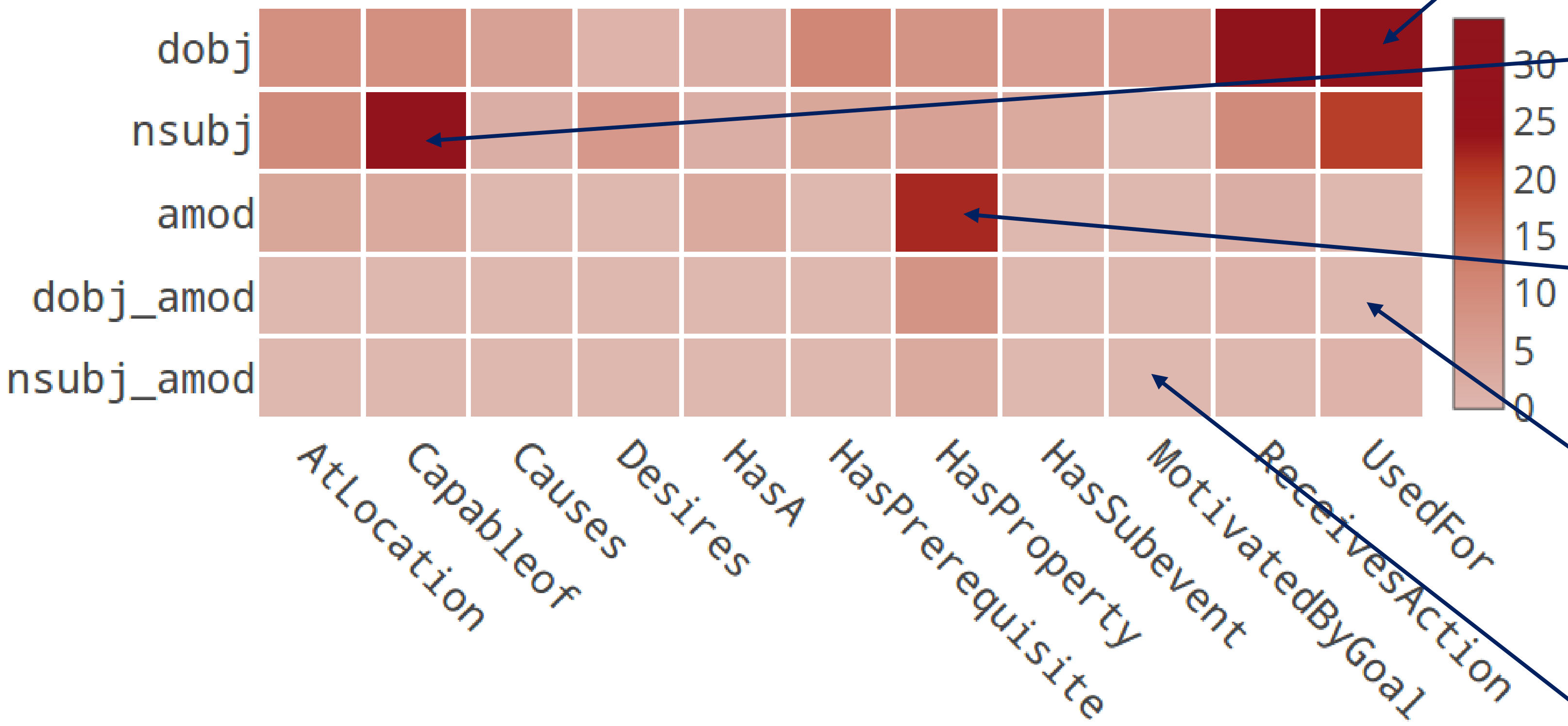
ConceptNet (Speer & Havasi, 2012)
Core is OMCS (Liu & Singh 2004)

- Commonsense knowledge base
  - Commonsense knowledge about noun-phrases, or entities.



Speer and Havasi. "Representing General Relational Knowledge in ConceptNet 5." *LREC*. 2012.

Revisit the Correlations of SP and OMCS



(sing, song) (dobj, 9.25)
(song, UsedFor, sing)

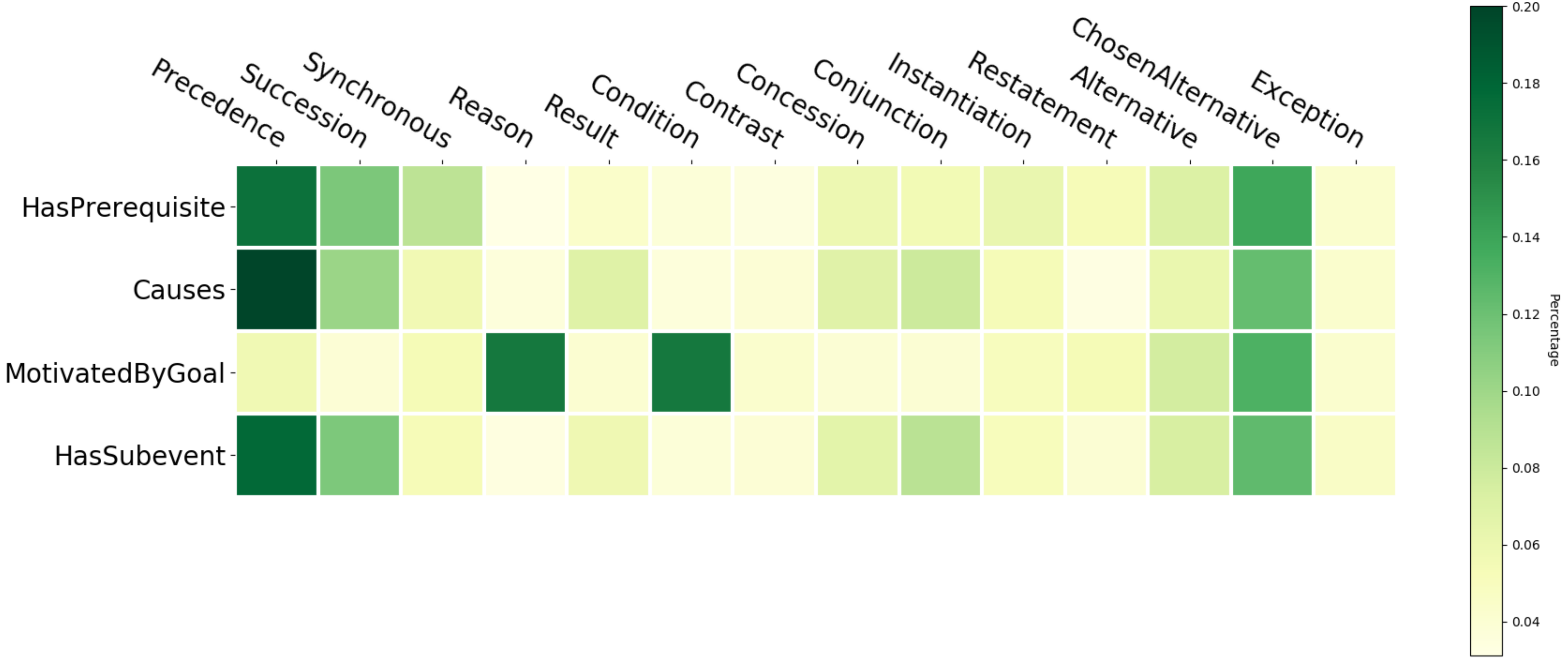
(phone, ring) (nsubj, 8.75)
(phone, CapableOf, ring)

(cold, water) (amod, 8.86)
(water, HasProperty, cold)

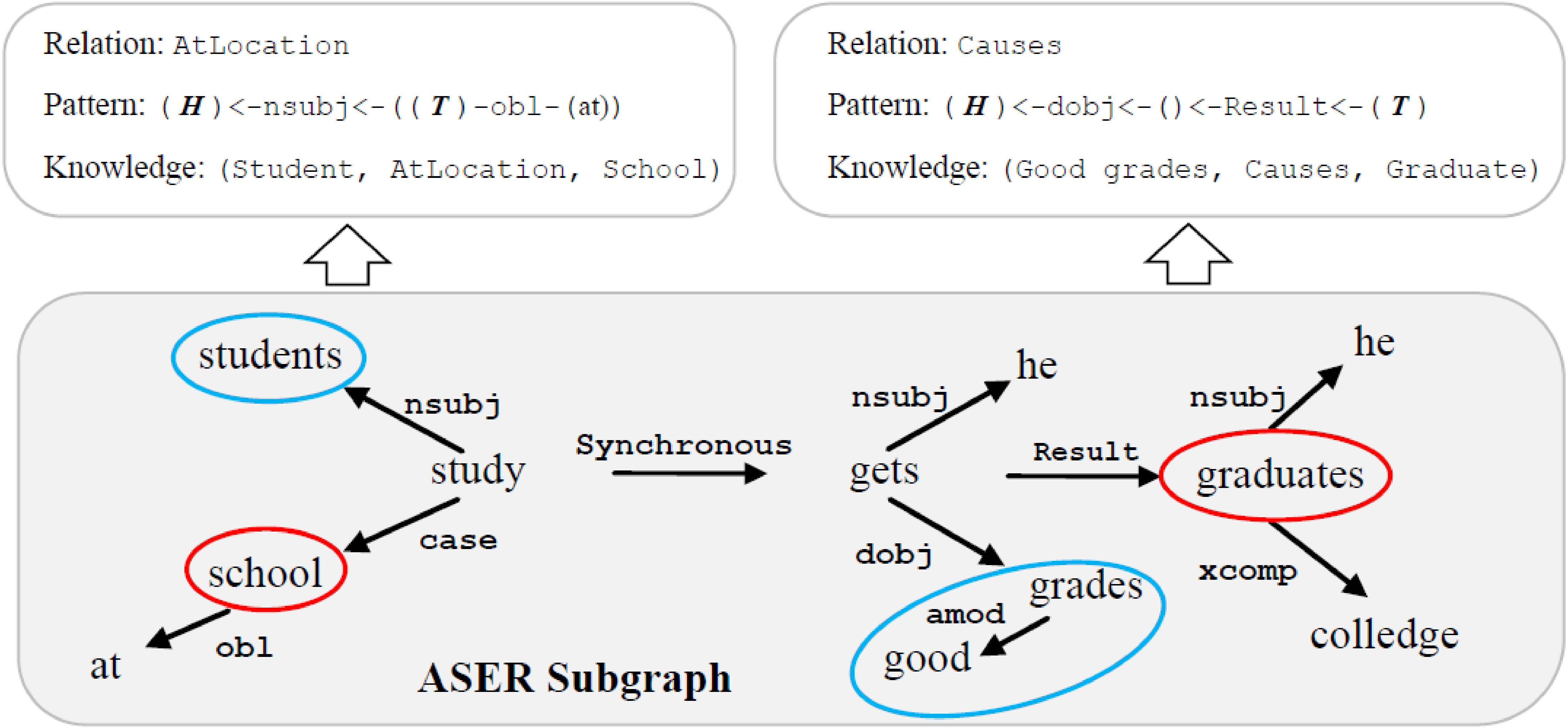
(create, new) (dobj_amod, 8.25)
(create idea, UsedFor, invent new things)

(hungry, eat) (nsubj_amod, 10.00)
(eat, MotivatedByGoal, are hungry)

Revisit the Correlations of ASER and OMCS

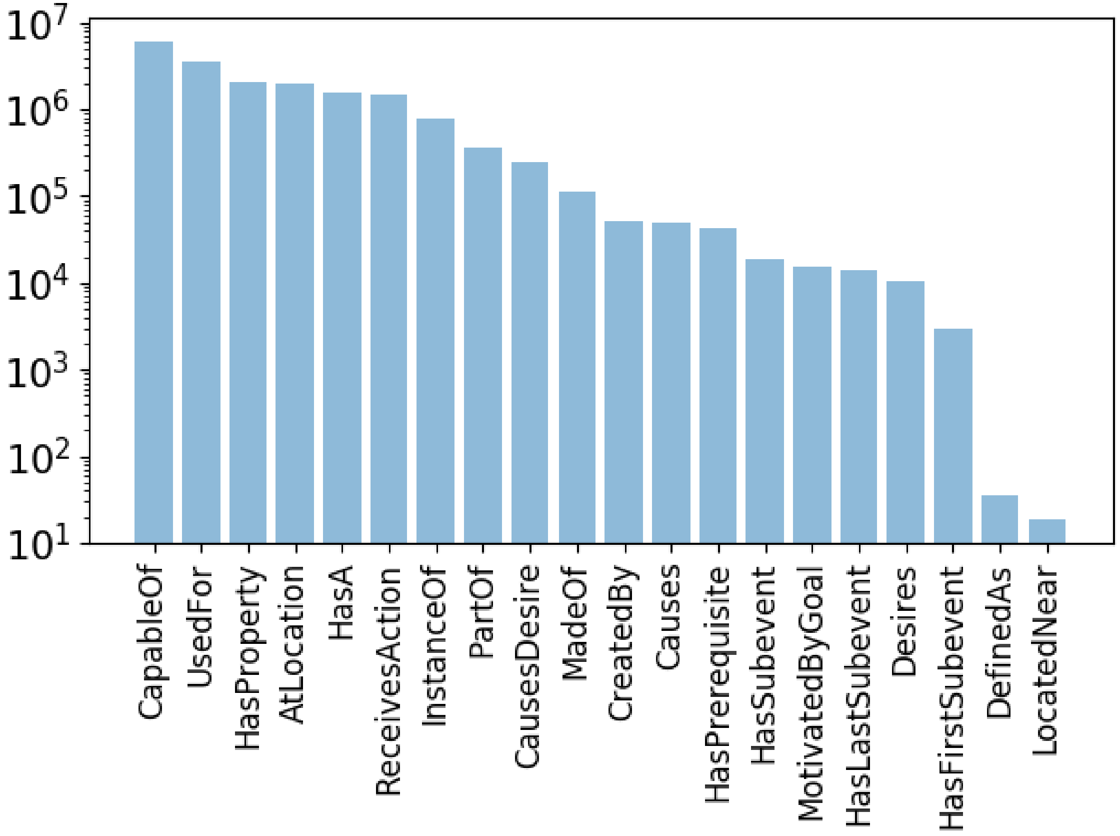


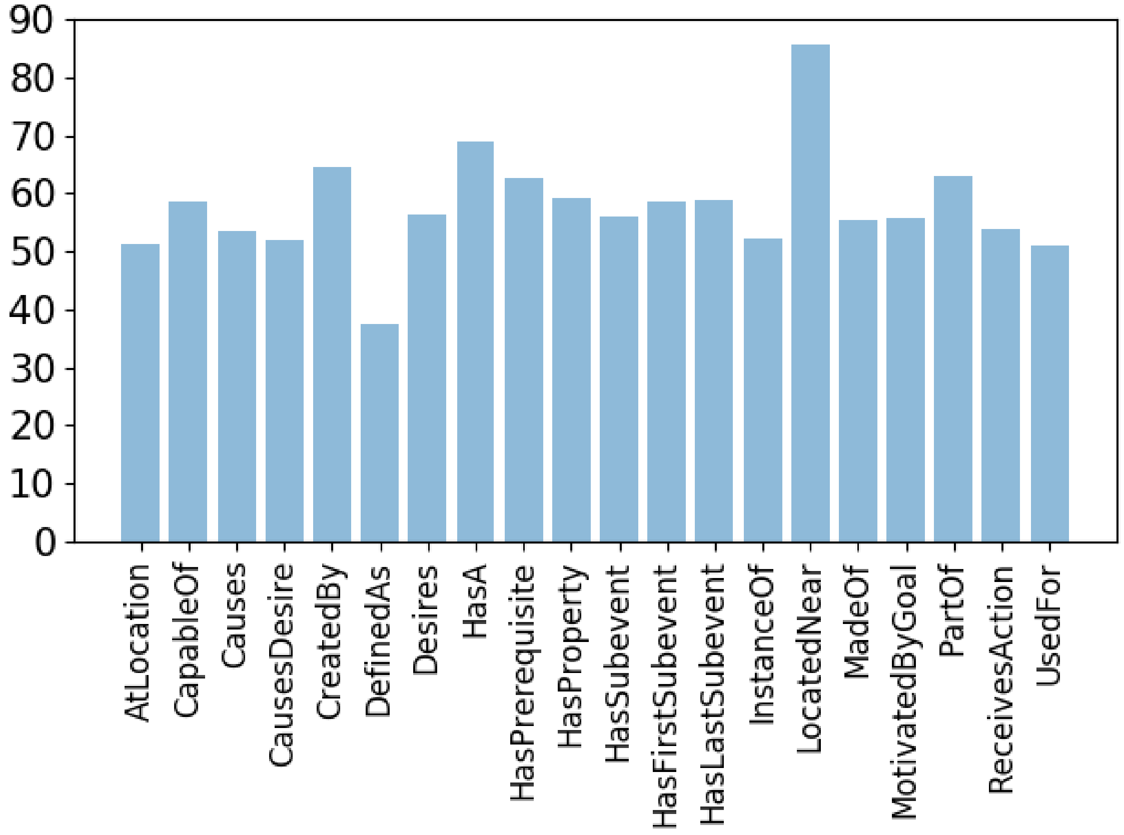
TransOMCS



Relation: AtLocation
Pattern: (H)<-nsubj<-((T)-obl-(at))
Knowledge: (Student, AtLocation, School)
Relation: Causes
Pattern: (H)<-dobj<-()<-Result<-(T)
Knowledge: (Good grades, Causes, Graduate)
students
nsubj
study
case
school
obl
at
Synchronous
ASER Subgraph
he
nsubj
gets
Result
dobj
grades
amod
good
he
nsubj
graduates
xcomp
colledge

Distribution of Relations and Accuracy

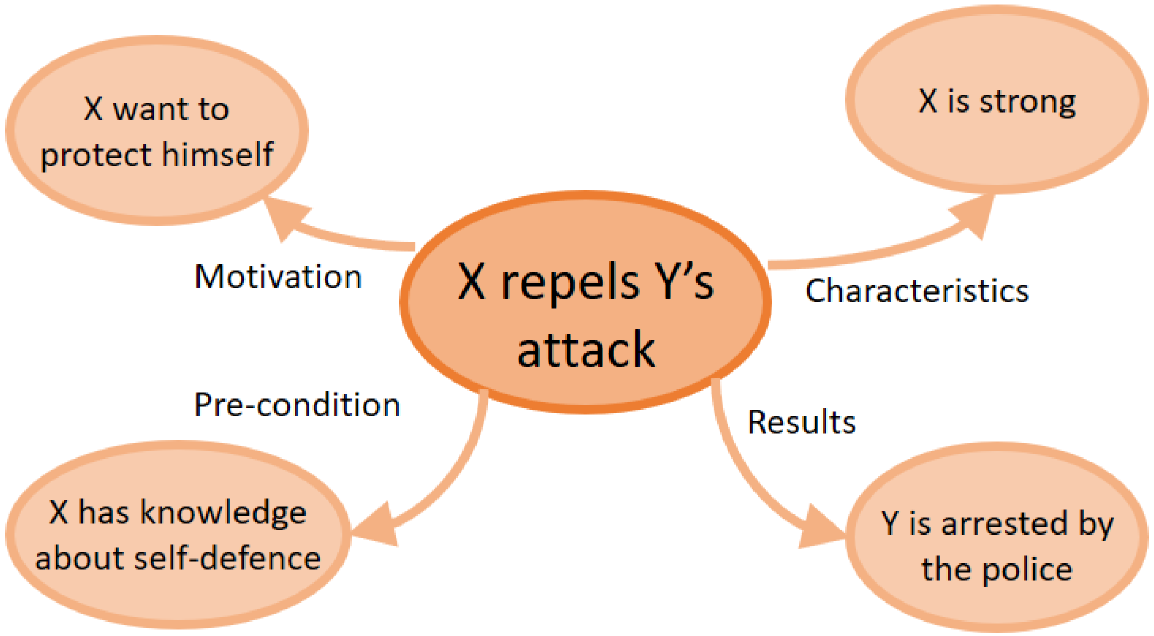




~90x Larger (about 18M triplets) than OMCS, with >50% accuracy

ATOMIC (Sap, Maarten, et al. 2019)

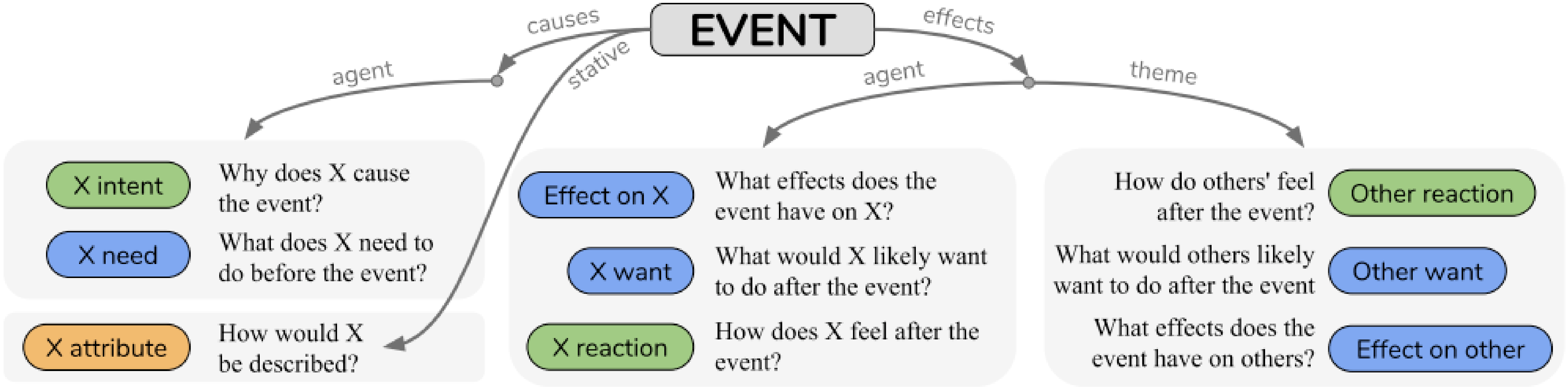
- Everyday if-then commonsense knowledge
- These are day-to-day knowledge that help us understand each other
  - If a person *X did* something, human beings are able to inference:
    - Motivation: Why person X did this.
    - Pre-conditions: What enables X to do this.
    - Characteristics: What are attributes of X.
    - Result: What will affect X/others



Sap, Maarten, et al. “Atomic: An atlas of machine commonsense for if-then reasoning.”, *AAAI* 2019.

ATOMIC (Sap, Maarten, et al. 2019)

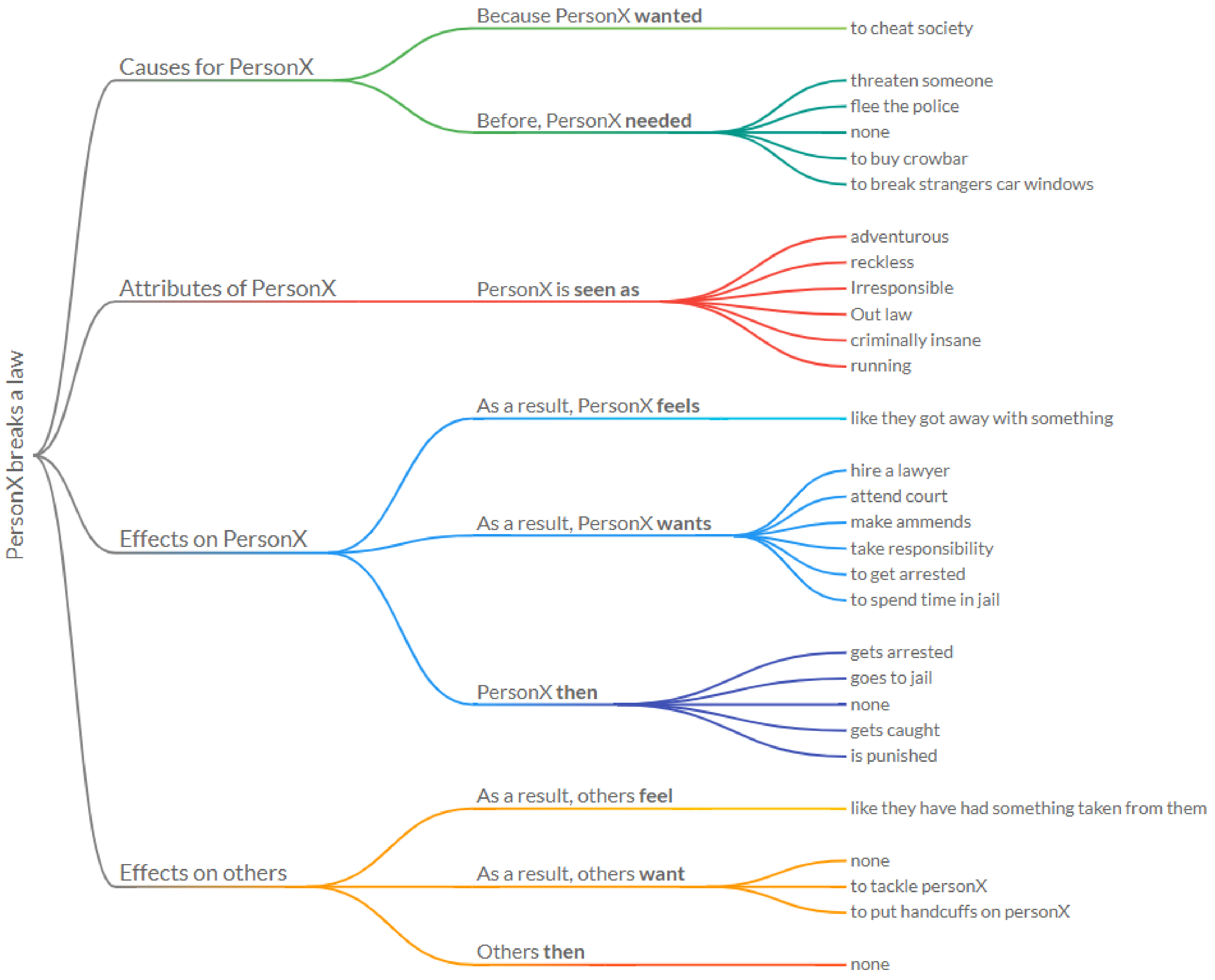
- Define 4 categories of if-then relations:
  - Causes-agent (Motivation & Pre-condition): xIntend, xNeed
  - Stative (Characteristics): xAttr
  - Effects-agent (Results on X): xWant, xReact, xEffect
  - Effects-theme (Results on others): oWant, oReact, oEffect



Sap, Maarten, et al. “Atomic: An atlas of machine commonsense for if-then reasoning.”, *AAAI* 2019.

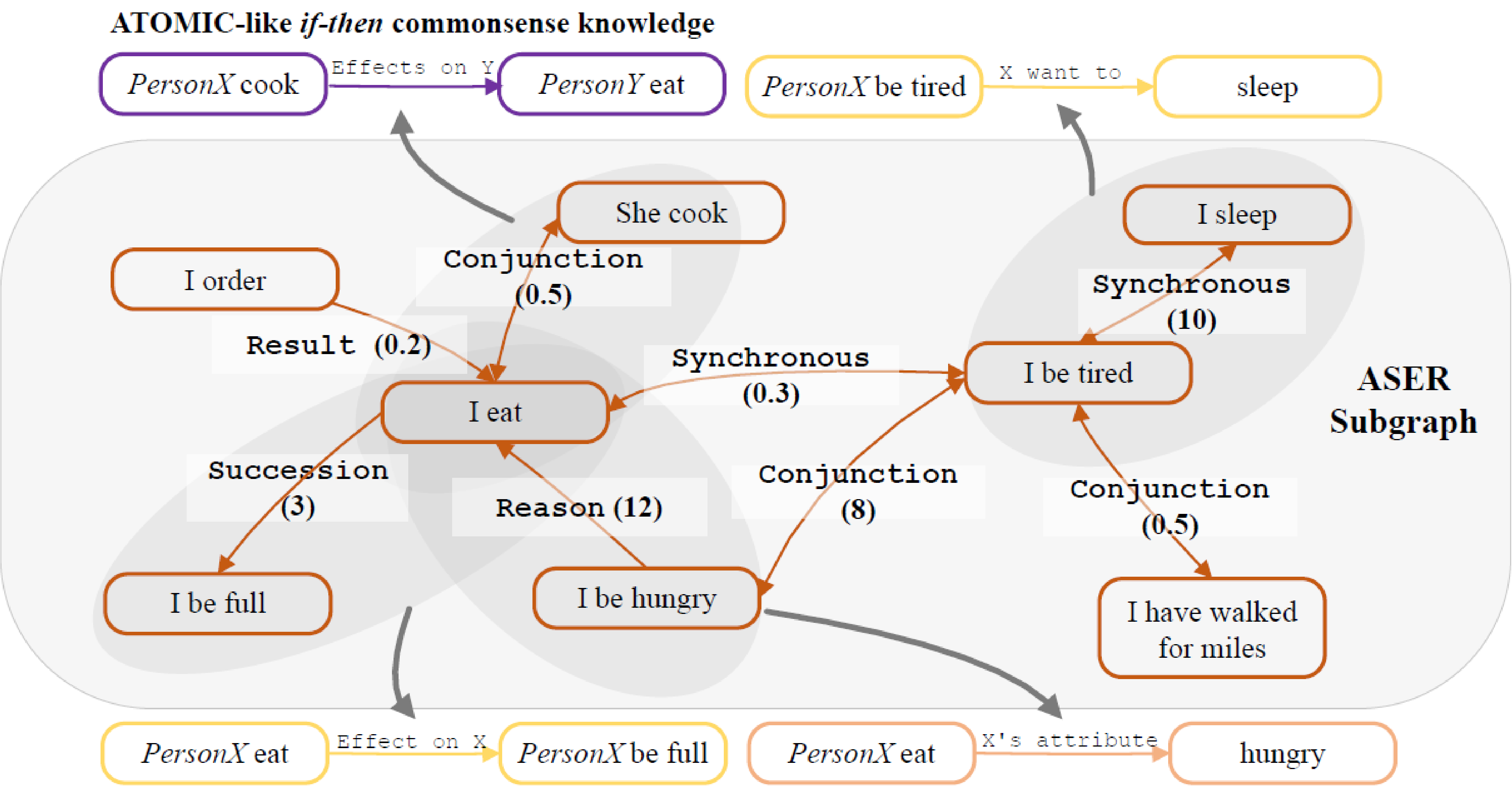
ATOMIC

- Crowdsourcing 9 Types of IF-THEN relations
- All personal entity information has been removed to reduce ambiguity
- Arbitrary texts

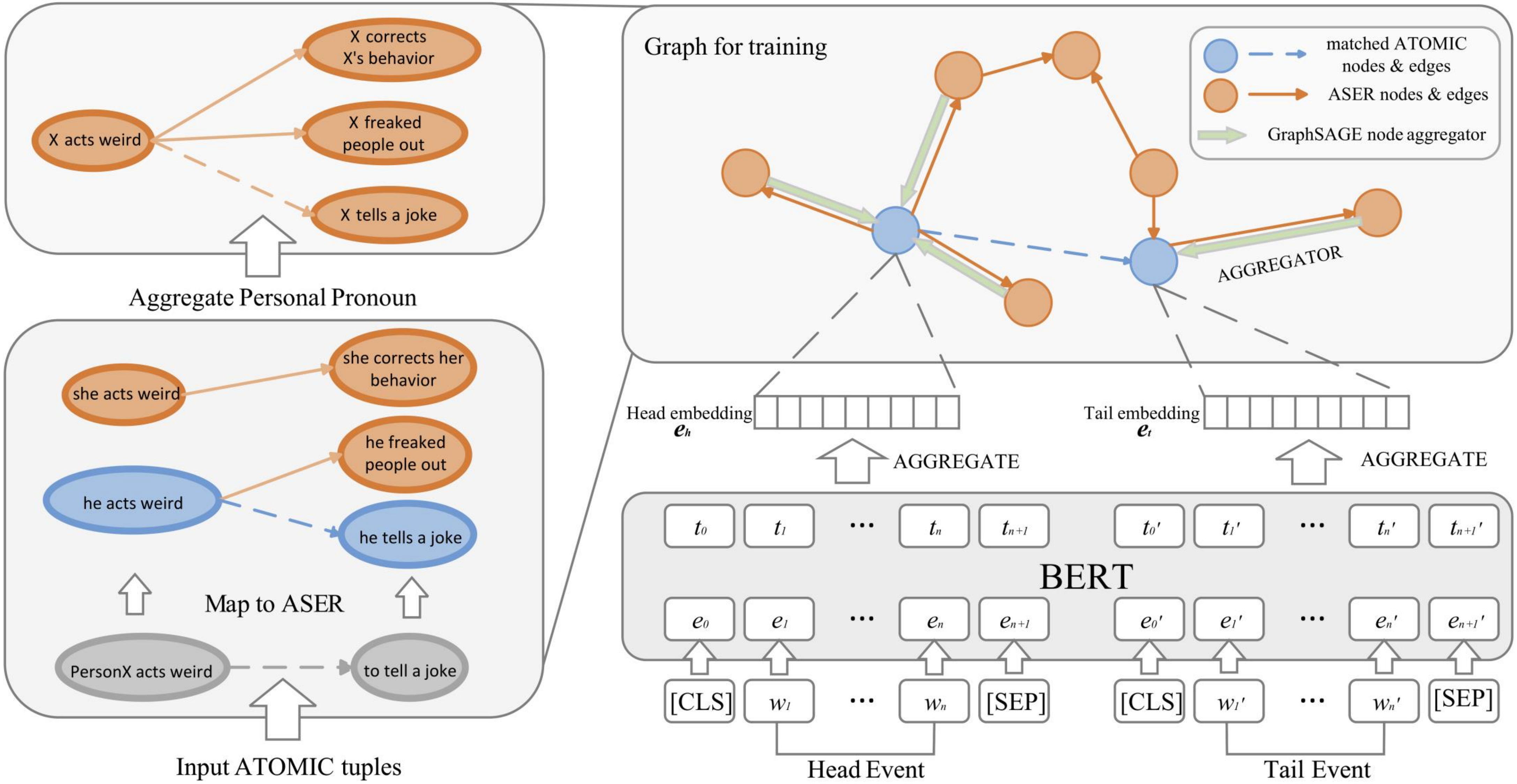


Maarten Sap, Ronan LeBras, Emily Allaway, Chandra Bhagavatula, Nicholas Lourie, Hannah Rashkin, Brendan Roof, Noah A. Smith, Yejin Choi: ATOMIC: An Atlas of Machine Commonsense for If-Then Reasoning. AAAI, 2019.

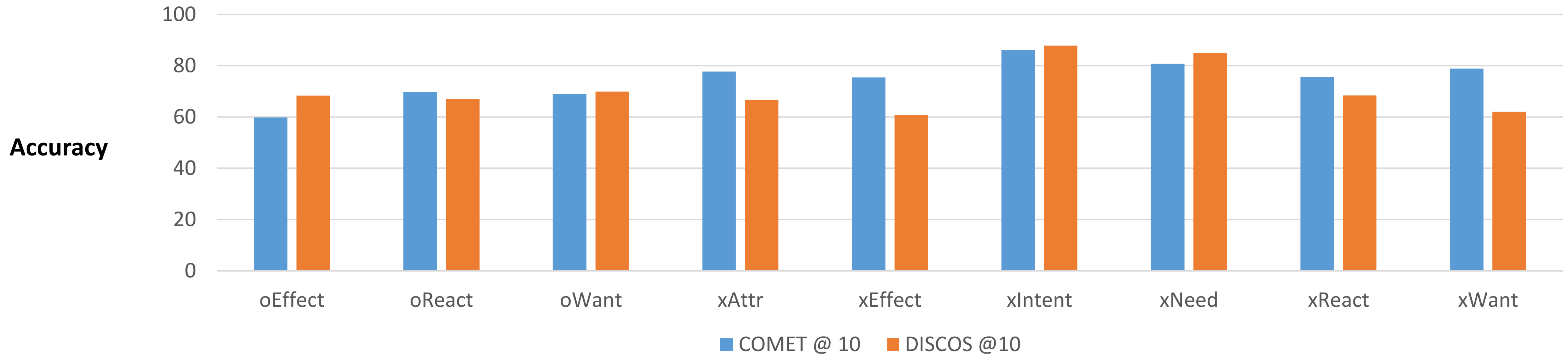
DISCOS: Transform to ATOMIC



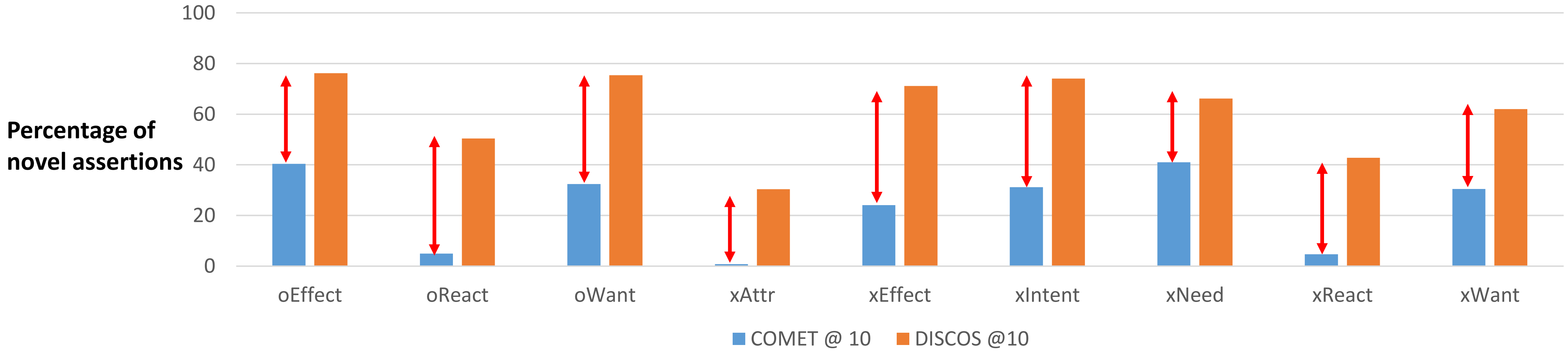
DISCOS Framework



DISCOS Result



Quality



Novelty

Conclusions and Future Work

Thank you ☺

- We extended the concept of selectional preference for commonsense knowledge acquisition
- We have proven that ASER can be transferred to other commonsense knowledge graphs:
  - OMCS/ConceptNet: TransOMCS (IJCAI 2020)
  - ATOMIC: DISCOS (WWW 21)
- We are building a commonsense knowledge population evaluation benchmark with Huawei
- We plan to build neural logical reasoning framework based on ASER
- Applications of ASER?
  - Event detection and reasoning
  - Other NLP tasks
    - Legal AI
    - …

Code and data

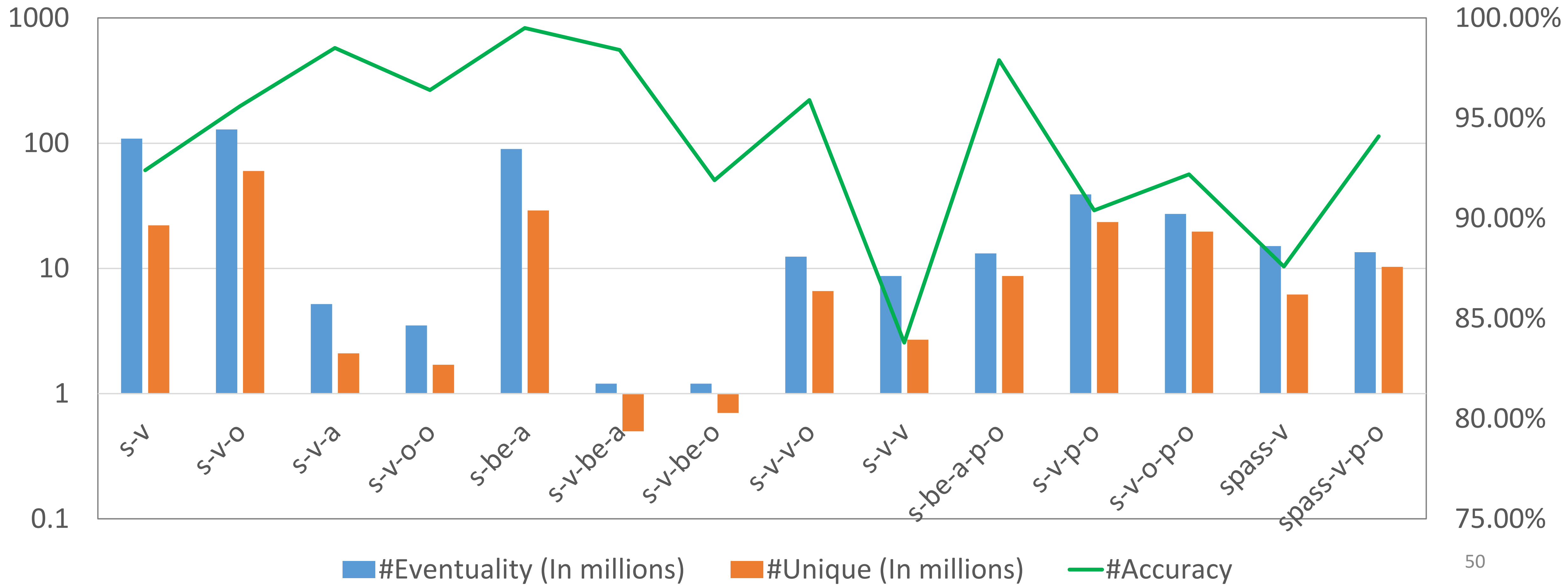
https://github.com/HKUST-KnowComp/ASER

Project Homepage

https://hkust-knowcomp.github.io/ASER/

Extraction Results

- Extract examples from 11-billion tokens from Yelp, NYT, Wiki, Reddit, Subtitles, E-books
- Evaluate about 200 examples in each pattern using Amazon Turk



Extraction Results

- Left: number of relations and overall accuracy
- Right: accuracy of each relations for the last iteration
- Each point is annotated with 200 examples by Amazon Turk