Activity (or Process), State, and Event-based Knowledge Graphs

Yangqiu Song

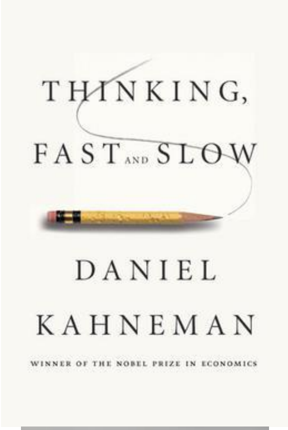
Department of CSE, HKUST

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

THE DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING 計算機科學及工程學系

KnowComp Group
Understanding the World by Computational Knowledge

Why Do We (Still) Need Knowledge Graphs in the Era of Deep Learning?



- Deep Learning for **System II** Processing, as proposed by Yoshua Bengio
  - We need to equip machine learning systems with "slow, logical, sequential, conscious, linguistic, algorithmic, planning, reasoning"
    - Particularly, such a system requires the "understanding of how actions" (including events and activities/processes) "interact with changes in distribution" which can be reflected by states.
    - Need "new priors to help systematic & OOD generalization"
- Language models need sensory grounding for meaning and understanding, as argued by Yann LeCun
- Multi-objective federated learning, proposed by Qiang Yang
  - One of the utility sources to make sure federation follows common sense
- Abductive learning, proposed by Zhihua Zhou
  - KB plays an important role to provide weak labels

http://www.iro.umontreal.ca/~bengioy/AAAI-9feb2020.pdf
https://wp.nyu.edu/consciousness/do-large-language-models-need-sensory-grounding-for-meaning-and-understanding/

Knowledge Graphs

- Many large-scale knowledge graphs about entities and their attributes (property-of) and relations (thousands of different predicates) have been developed since Google released its knowledge graph in 2012
  - Millions of entities and concepts
  - Billions of relationships

Are they enough to characterize our mental world?

Graph (2012)
entities and 18 billion facts

Challenge 1: How to Grow a Mind?
--Statistics, Structure, and Abstraction

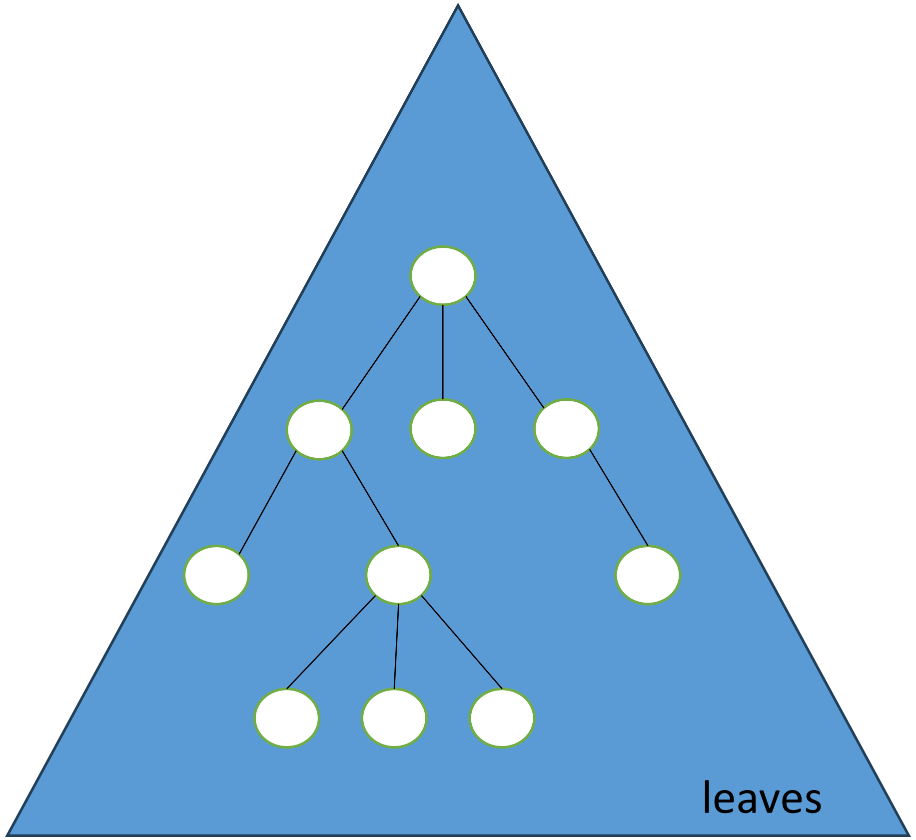
- "In coming to understand the world—in learning concepts, acquiring language, and grasping causal relations—our minds make inferences that appear to go far beyond the data available."

- The ability of performing powerful abstraction is the key

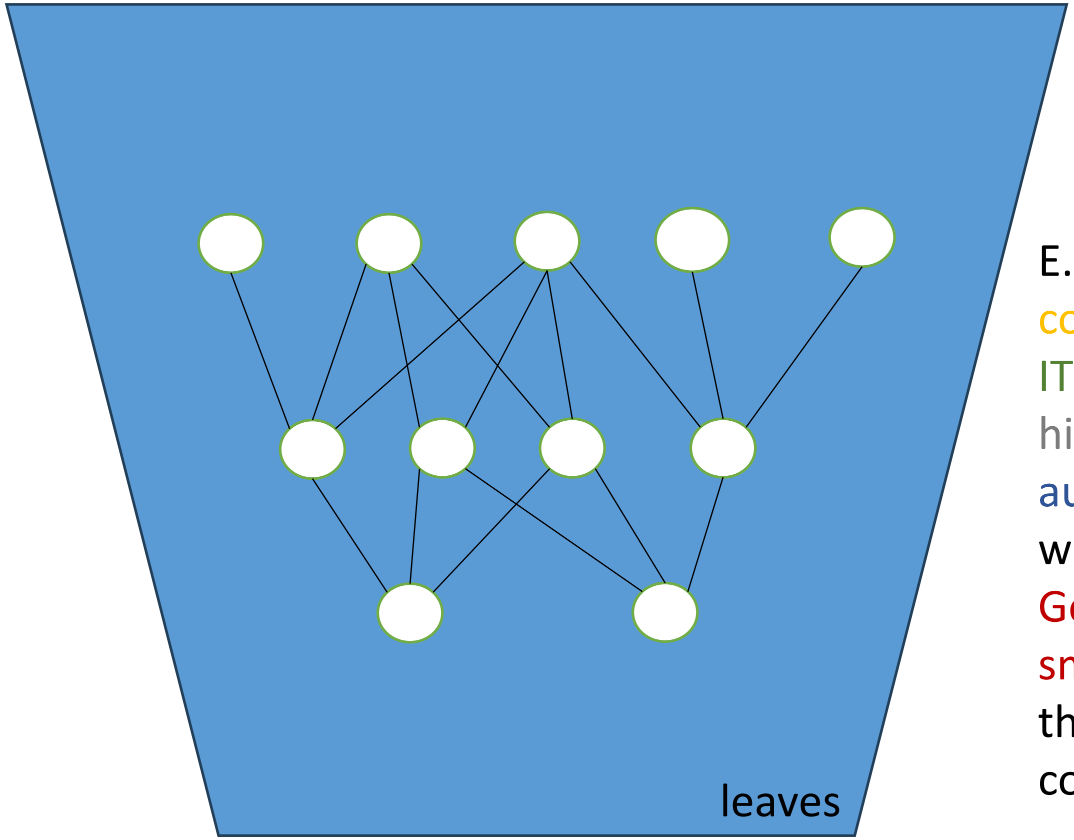
How to grow a mind: statistics, structure, and abstraction. Science. Joshua B Tenenbaum, Charles Kemp, Thomas L Griffiths, Noah D Goodman. 2011.

The True Structure of Conceptualization

What you may think of



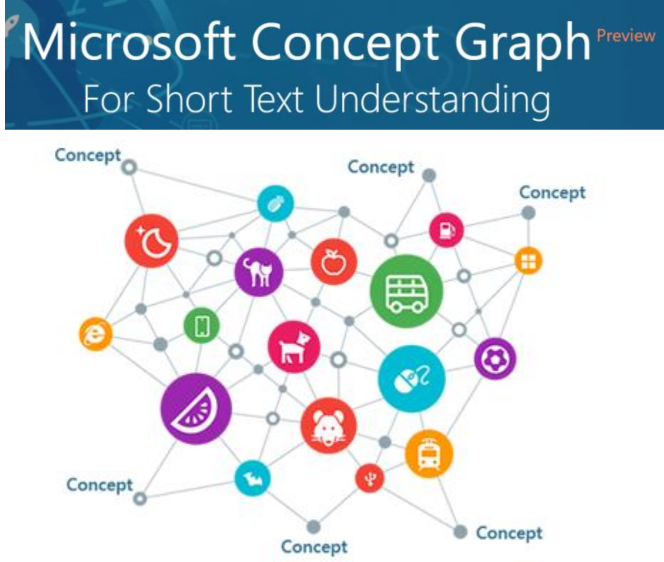
What it really likes

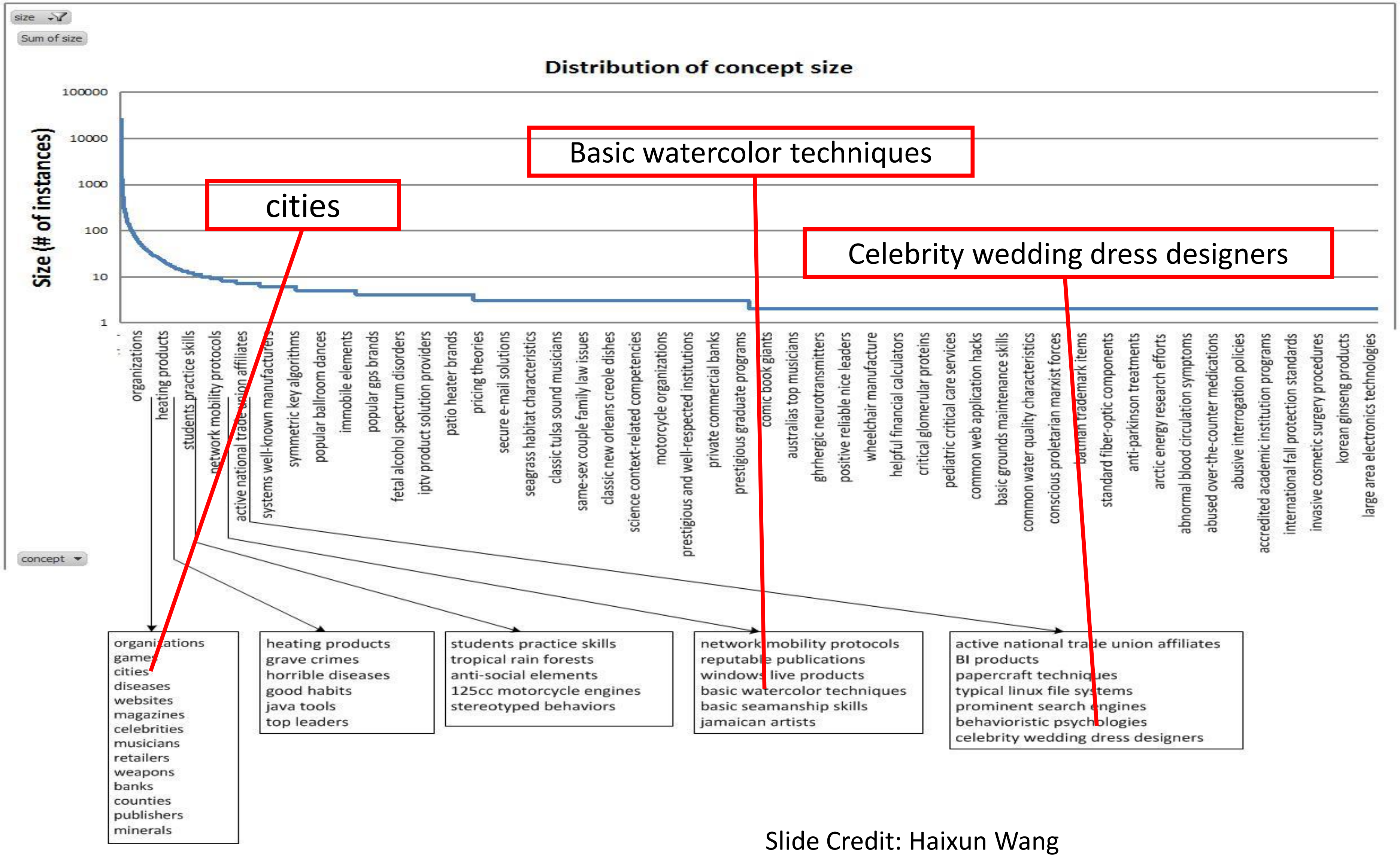


E.g., mapping Tesla to a company, big company, IT company, AI company, high-tech company, automobile company, when comparing it with Google, Toyota, some small company, needs the right level of comparison

The nature of conceptual abstraction (aggregation of objects' attributes) and the (dynamic) compositionality make the number of concepts grow vastly comparable to entities we have in the world

ProBase (2011 at MSRA)





Slide Credit: Haixun Wang

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

Challenge 2: Knowledge Graph or More General Mind Graphs to Describe Mental World?

- What is our knowledge?
  - Knowledge is often defined as justified true belief
- What makes us take actions?
  - Beliefs and desires are mediated by intentions which in turn controls human's actions (Kashima et al., 1998)
- Theory of Mind
  - Needs the reasoning about intents, feelings, mental states, and realities (actions or reactions) (Sap et al., 2022)

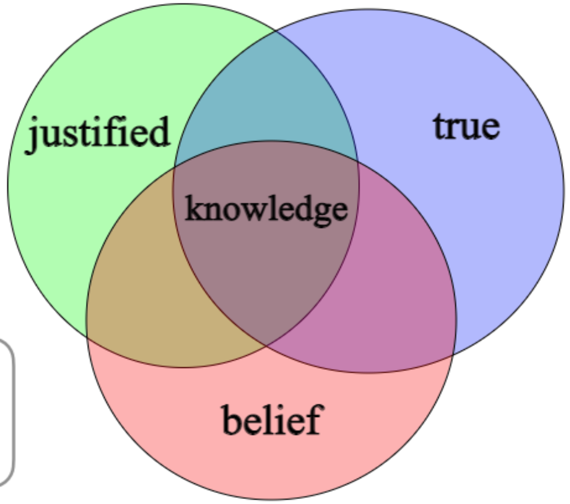


Figure taken from Wikipedia

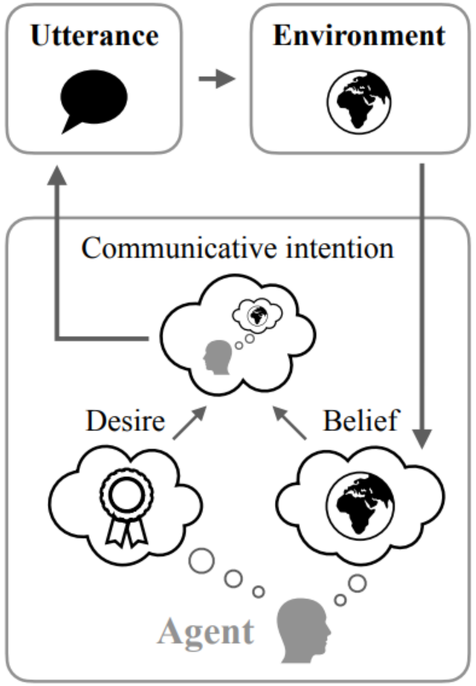


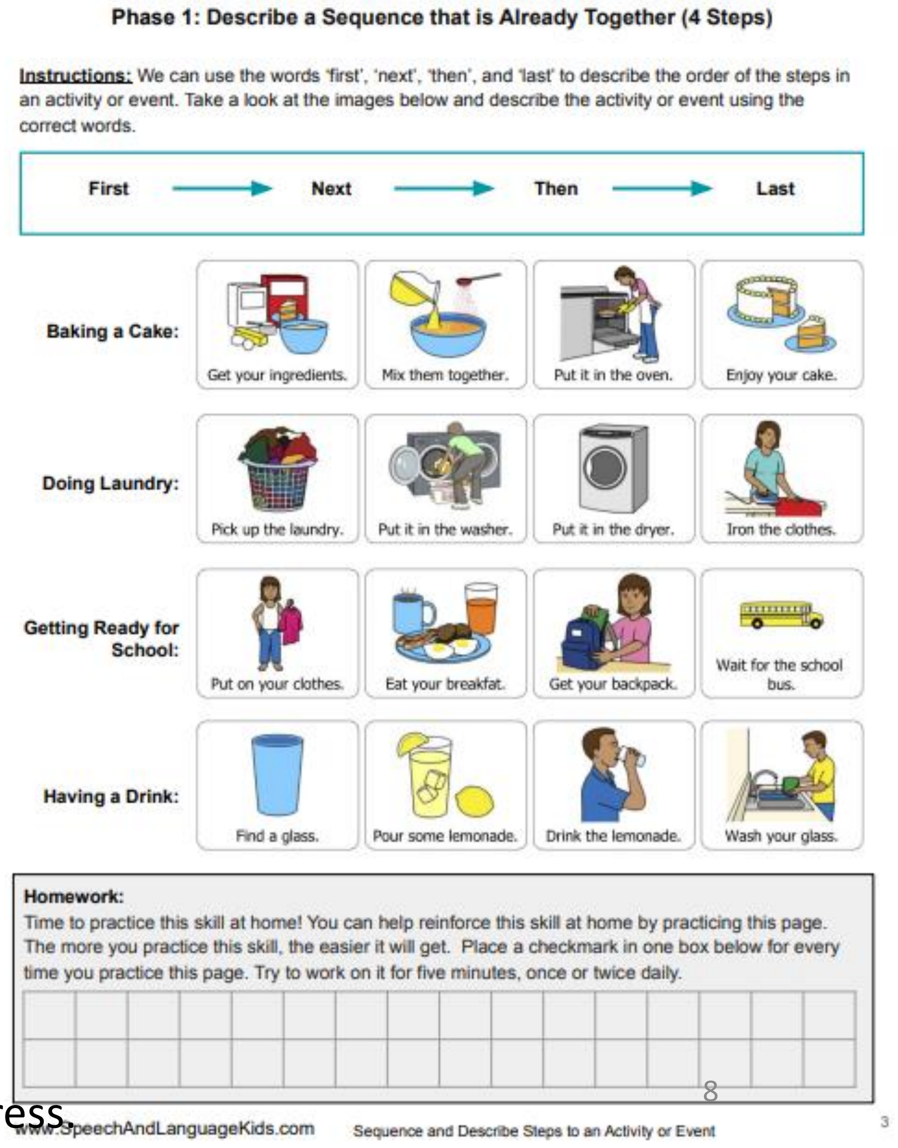
Figure taken from Andreas (2022)

https://en.wikipedia.org/wiki/Knowledge
Jacob Andreas: Language Models as Agent Models. EMNLP (Findings) 2022: 5769-5779
Kashima, Yoshihisa, Allison McKintyre, and Paul Clifford. "The Category of the Mind: Folk Psychology of Belief, Desire, and Intention." Asian Journal Of Social Psychology 1, no. 3 (December 1998): 289–313.
Maarten Sap, Ronan Le Bras, Daniel Fried, Yejin Choi: Neural Theory-of-Mind? On the Limits of Social Intelligence in Large LMs. EMNLP 2022: 3762-3780

Challenge 3: Primitive Semantic Units in Our Mind

- 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),
  - Activity (or Process),
  - State,
  - Event,
  - Place,
  - Path,
  - Property,
  - Amount,
  - etc.

Ray Jackendoff, Rumelhart Prize winner, studied under linguists Noam Chomsky



Phase 1: Describe a Sequence that is Already Together (4 Steps)

Instructions: We can use the words 'first', 'next', 'then', and 'last' to describe the order of the steps in an activity or event. Take a look at the images below and describe the activity or event using the correct words.

	First	Next	Then	Last
Baking a Cake:	Get your ingredients.	Mix them together.	Put it in the oven.	Enjoy your cake.
Doing Laundry:	Pick up the laundry.	Put it in the washer.	Put it in the dryer.	Iron the clothes.
Getting Ready for School:	Put on your clothes.	Eat your breakfat.	Get your backpack.	Wait for the school bus.
Having a Drink:	Find a glass.	Pour some lemonade.	Drink the lemonade.	Wash your glass.

Homework:
Time to practice this skill at home! You can help reinforce this skill at home by practicing this page. The more you practice this skill, the easier it will get. Place a checkmark in one box below for every time you practice this page. Try to work on it for five minutes, once or twice daily.

www.SpeechAndLanguageKids.com Sequence and Describe Steps to an Activity or Event 3

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

Activities and Events can also be Conceptualized

What activities do you find relaxing?

25 Calming Activities to Do in Your House

- Take a hot bath or shower.
- Lay down with your legs elevated and watch your breathing rise and fall.
- Doodle or color.
- Look out a window or let in fresh air.
- Have a warm drink without caffeine or alcohol.
- Engage in a visualization of a place that makes you feel safe.

More items... • 22 Mar 2020

https://www.psychologytoday.com › blog › friendship-20

25 Calming Activities to Do in Your House - Psychology Today

Search for: What activities do you find relaxing?

OutofStress
https://www.outofstress.com › things-stress-relief

86 Fun Activities To Relax and De-Stress

23 May 2022 — 65 deeply **relaxing** activities that you can do anyplace and anytime. ... write about past **events** that have been on your mind.

Live Bold and Bloom
https://liveboldandbloom.com › self-improvement › re...

29 Relaxing Things To Do To DeStress And Recharge

16 Oct 2022 — 1. Read a book. 2. Take a hike. 3. Practice meditation. 4. Write something. 5. Listen to music. 6. Get a massage from a family member. 7. Play ...

20 ways to switch off the stress

8 Sept 2016

1. Sydney Float Centre 2. Hang upside down 3. Walk the Royal Na... 4. Get your shit toget... 5. Connect to countr... 6. Take a bath at ...

Eventbrite
https://www.eventbrite.co.uk › blog › 10-unusual-effe...

Pre-Event Chill: 10 Best Ways to Relax When Traditional ...

10 Apr 2021 — 2. Indulge in dark chocolate · 3. Repeat a mantra · 4. Smell an orange · 6. Play a **relaxing** video game · 8. Clean your desk · 9. Inflate a balloon.

pro-activ.com
https://www.pro-activ.com › ... › Active Lifestyle

How to relieve stress: 10 fun & relaxing activities - Pro-Activ

10 ways to relieve stress: · Breathe deep: If your breathing is quick and shallow, rest a hand on your chest and watch it rise and fall. · Try a massage · Doodle ...

Apartment Therapy
https://www.apartmenttherapy.com › rest-relaxation-id...

83 Ways to Rest and Recharge, Whether You Have 5 ...

18 May 2021 — 1. Go on a walk. · 2. Do a body scan meditation. · 3. Focus on a hobby. · 4. Curate a book list. · 5. Start a dance party. · 6. Stand outside or near ...

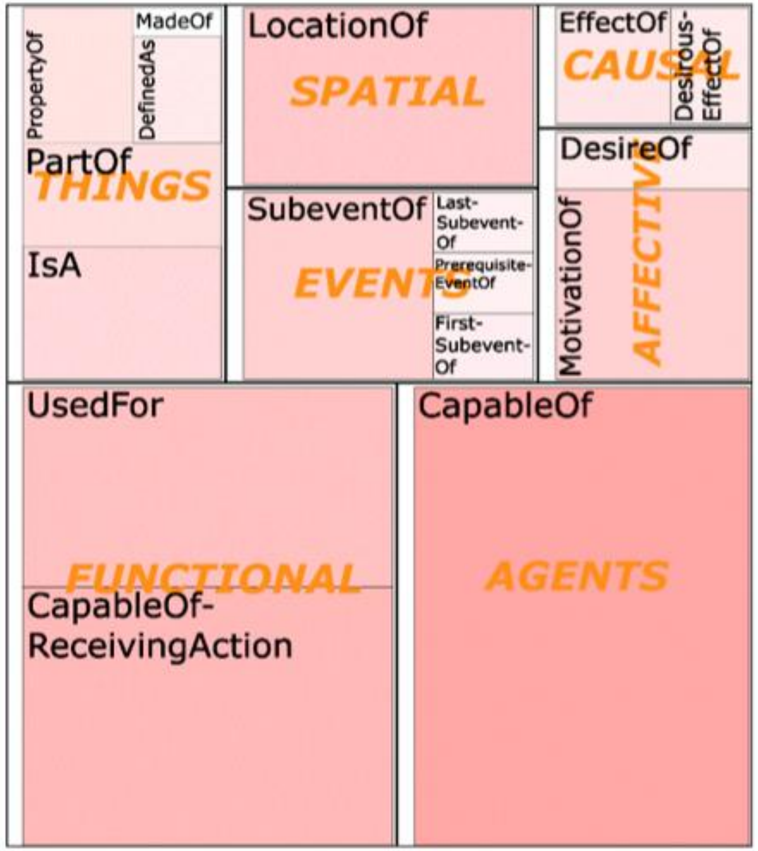
Declutter The Mind
https://declutterthemind.com › Blog

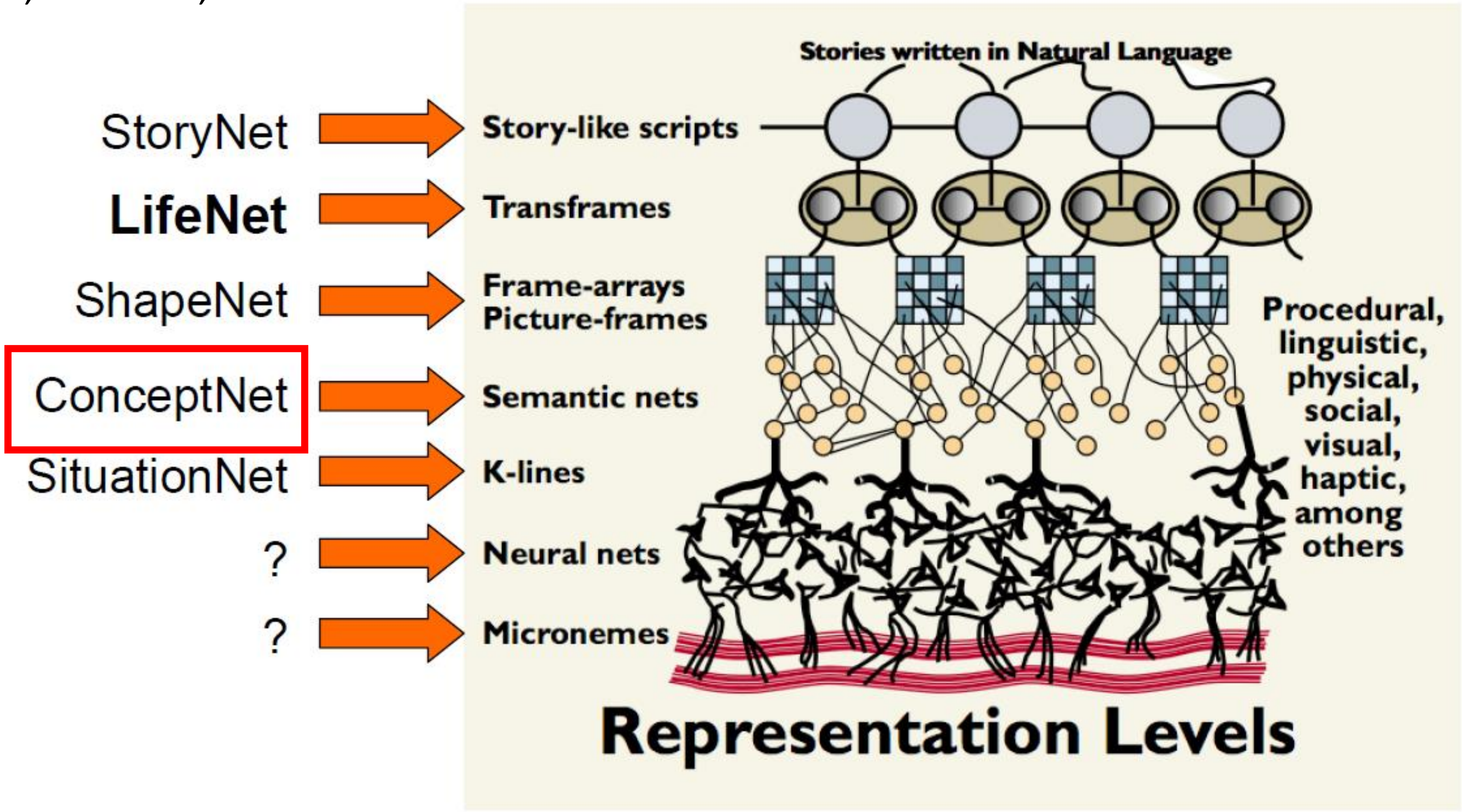
11 Relaxing Activities Before Bed to Fall Asleep Fast

1. Listen to ASMR · 2. Practice meditation · 3. Listen to sleep hypnosis audio recordings · 4.

That's Why ConceptNet was Designed in Such a Way

- Knowledge in ConceptNet
  - Things, Spatial, Location, Events, Causal, Affective, Functional, Agents

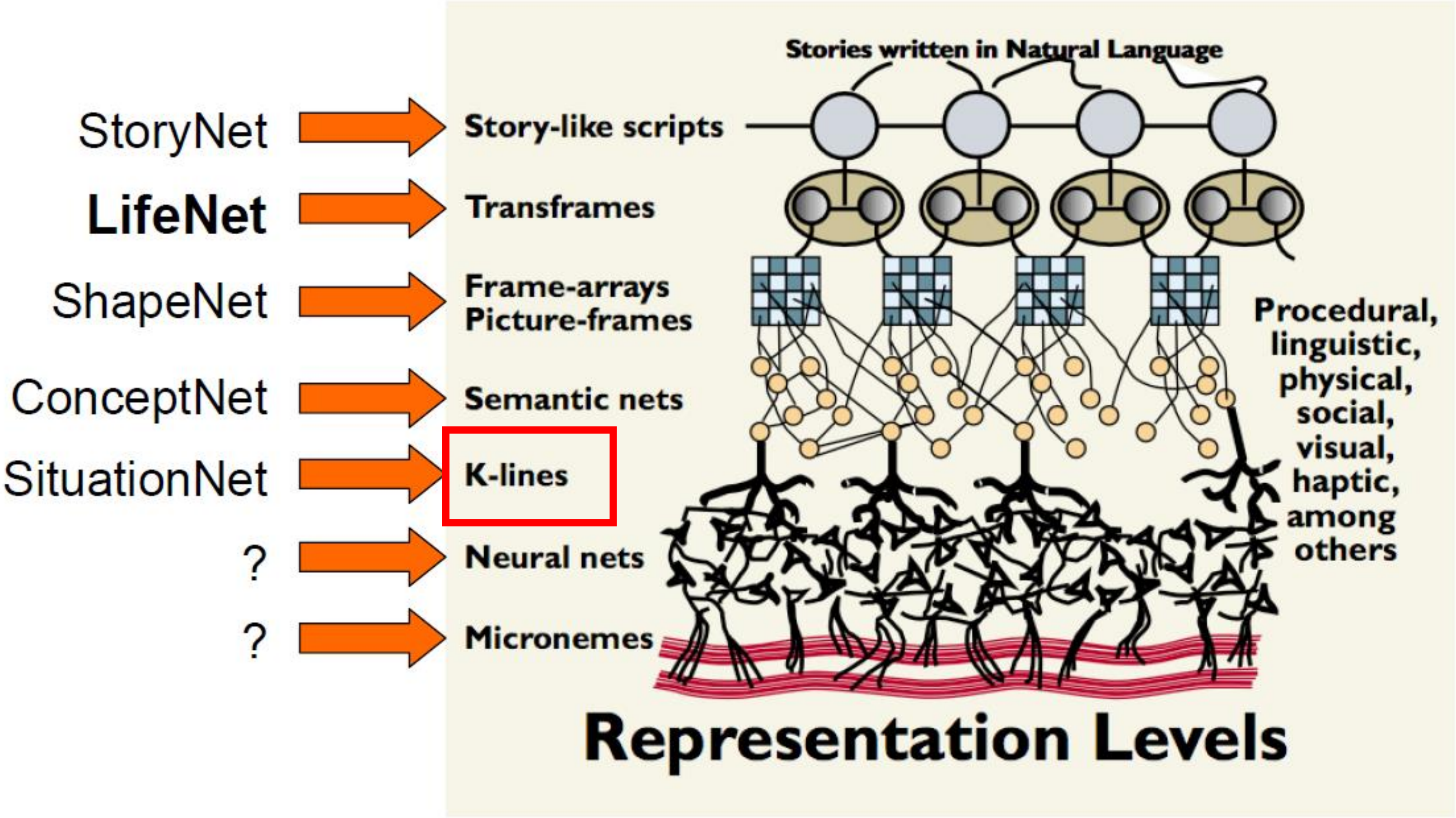




Singh, Barbara Barry, and Hugo Liu (2004). **Teaching machines about everyday life.** *BT Technology Journal*, 22(4):227-240.
Figure taken from: https://ocw.mit.edu/courses/media-arts-and-sciences/mas-961-ambient-intelligence-spring-2005/lecture-notes/week4_push_singh.pdf

A More Fundamental Layer is Called K-Line

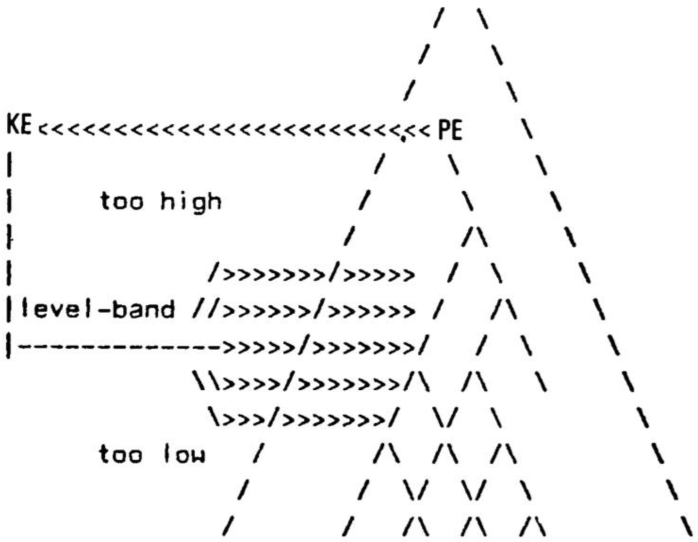
- Encode memories in "abstract" form.
- Search all memory for the "nearest match."
- Use prototypes with detachable defaults.



M. Minsky, "K-Lines: A theory of Memory," Cognitive Science 4 (1980). 117-133.
Figure taken from: https://ocw.mit.edu/courses/media-arts-and-sciences/mas-961-ambient-intelligence-spring-2005/lecture-notes/week4_push_singh.pdf

The Implementation of K-Line Theory

- We need the right level and right perspective of conceptualization of events
  - Different levels of abstractness
    - "PersonX drinks coca cola" → "[drinking coca cola]," "[drinking beverage]," "[event]"
  - Different perspectives
    - "Coca cola" → "[sugary beverage]," "[phosphate containing beverage]," "[iced drink]," not in a strict hierarchical taxonomy
      - PersonX drinks [iced drink], xReact, refreshed
      - PersonX drinks [sugary beverage], xEffect, gain weight

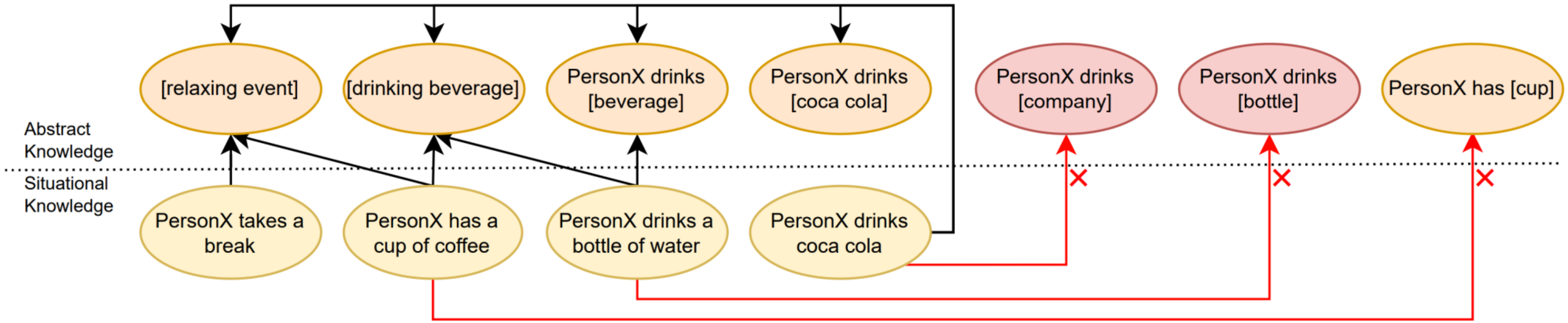


Attach a K-node (a mental state, KE) to a "Pyramid" agent (PE) at a certain level

The pyramid is a tree structure that we use to conceptualize the world

The mapping has a lower-band limit and a higher-band limit, to compare the right common, non-conflicting properties

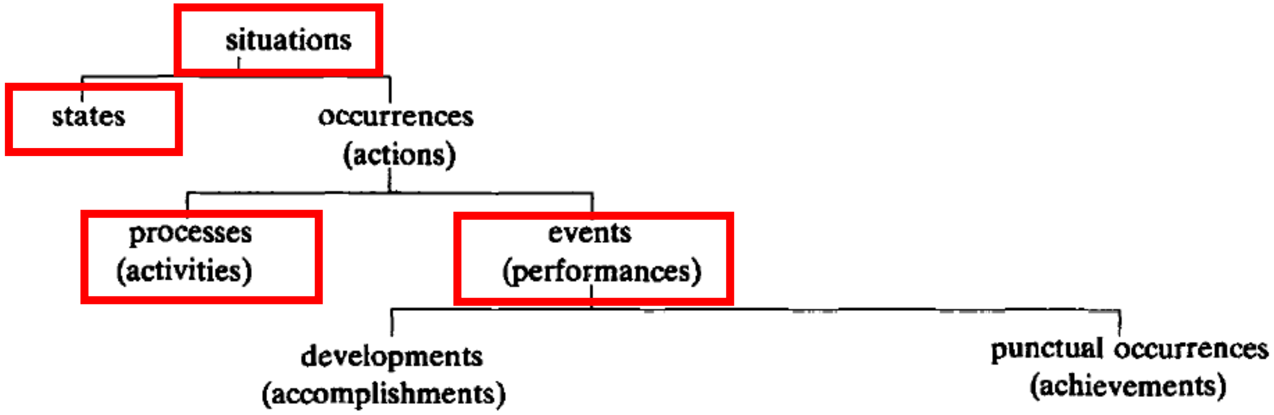
Figure taken from:
M. Minsky (1980)



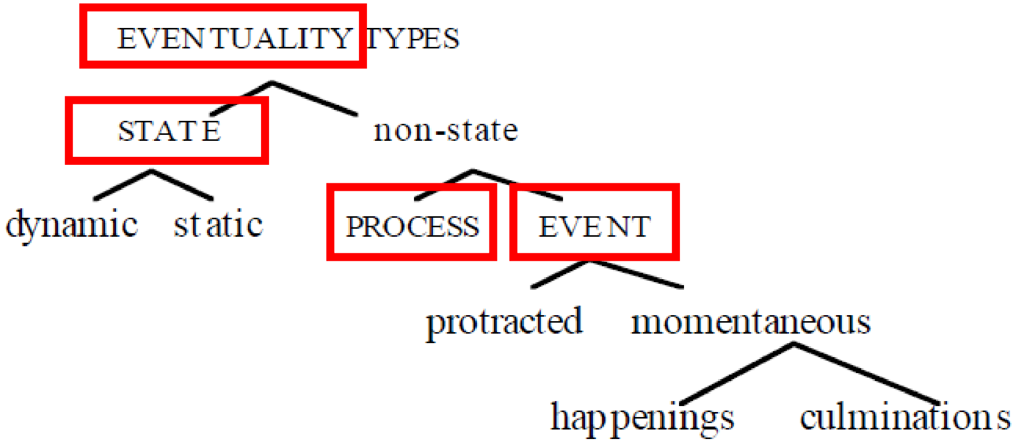
M. Minsky, "K-Lines: A theory of Memory," Cognitive Science 4 (1980). 117-133.
Mutian He, Tianqing Fang, Weiqi Wang, and Yangqiu Song. Acquiring and Modelling Abstract Commonsense Knowledge via Conceptualization. 2022.

ASER (Activities, States, Events, and their Relations)

Mourelatos' taxonomy (1978)



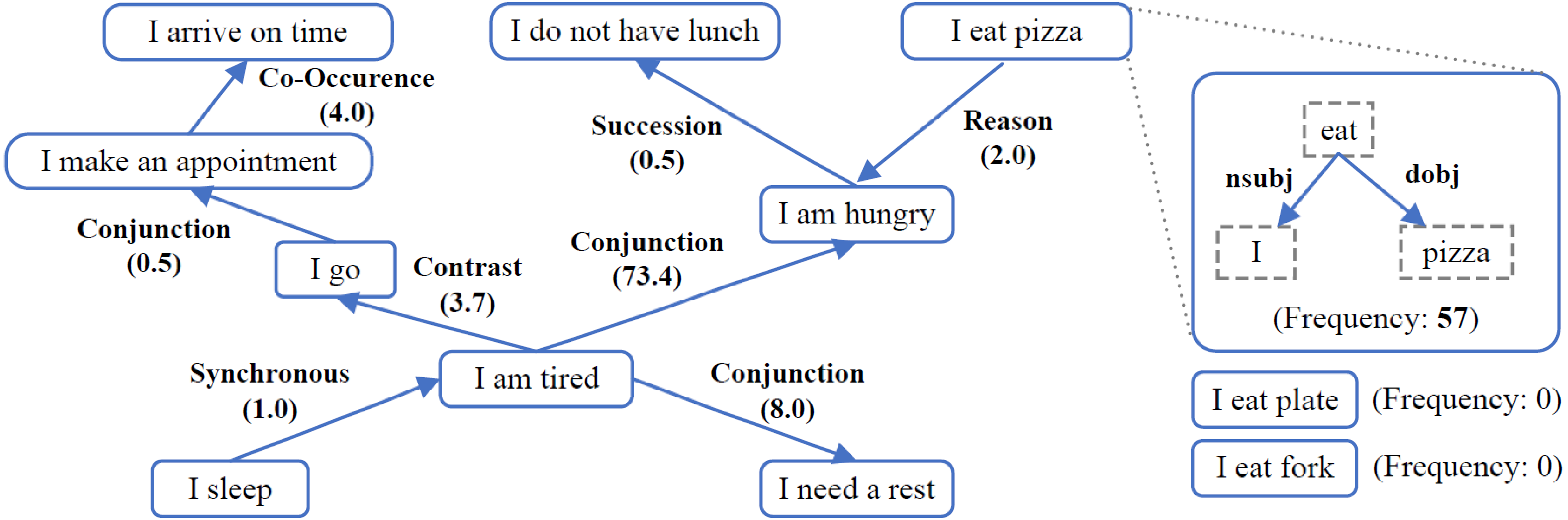
Bach's taxonomy (1986)



- **State**: A state is usually described by a stative verb and cannot be qualified as actions.
    - "The coffee machine is ready for brewing coffee."
- **Activity (or process):** Both activity and event are occurrences (actions) described by active verbs.
    - "The coffee machine is brewing coffee with following steps: …"
- **Event**: An event is defined as an occurrence that is inherently countable.
    - "The coffee machine brews a cup of coffee once more" is an event because it admits a countable noun "a cup" and cardinal count adverbials "once"

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.

ASER (Activities, States, Events, and their Relations)



Principle 1: Comparing semantic meanings by fixing grammar (Katz and Fodor, 1963)

Principle 2: The need of language inference based on 'partial information' (Wilks, 1975)

WWW'20

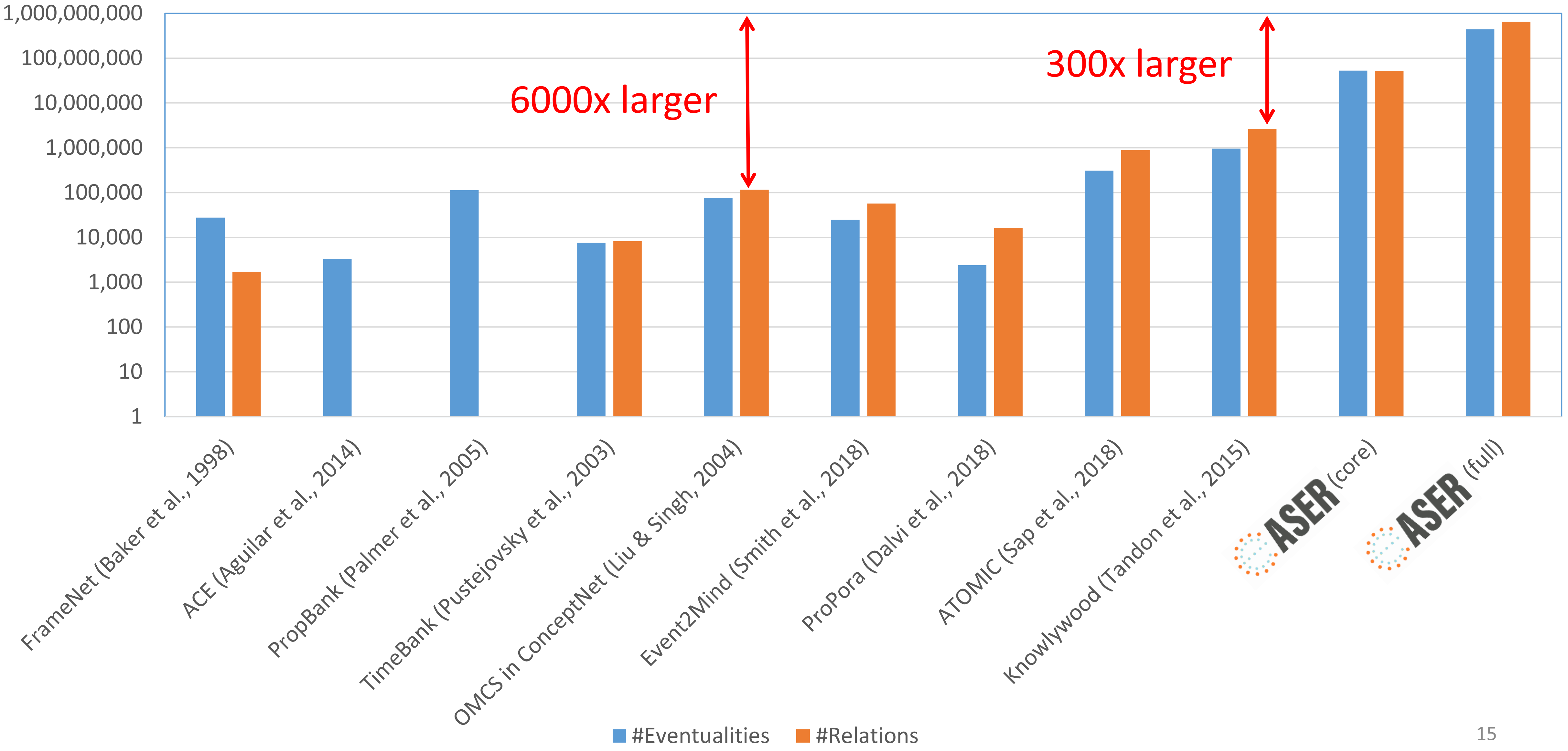
https://github.com/HKUST-KnowComp/ASER

Hongming Zhang, Xin Liu, Haojie Pan, Yangqiu Song, Cane Wing-Ki Leung: ASER: A Large-scale Eventuality Knowledge Graph. WWW 2020: 201-211

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

Yorick Wilks. 1975. An intelligent analyzer and understander of English. Communications of the ACM, 18(5):264–274.

Scales of Verb Related Knowledge Graphs



Conceptualization

(person, have, animal) (positive-emotion, come)

......

......

ResultIn [freq=3]

0.281 ← He, have, a little dog ⇢ the happiness, come → 0.087

......

......

ResultIn [freq=2]

0.333 ← I, have, my own horse ⇢ the exhilaration, come → 0.125

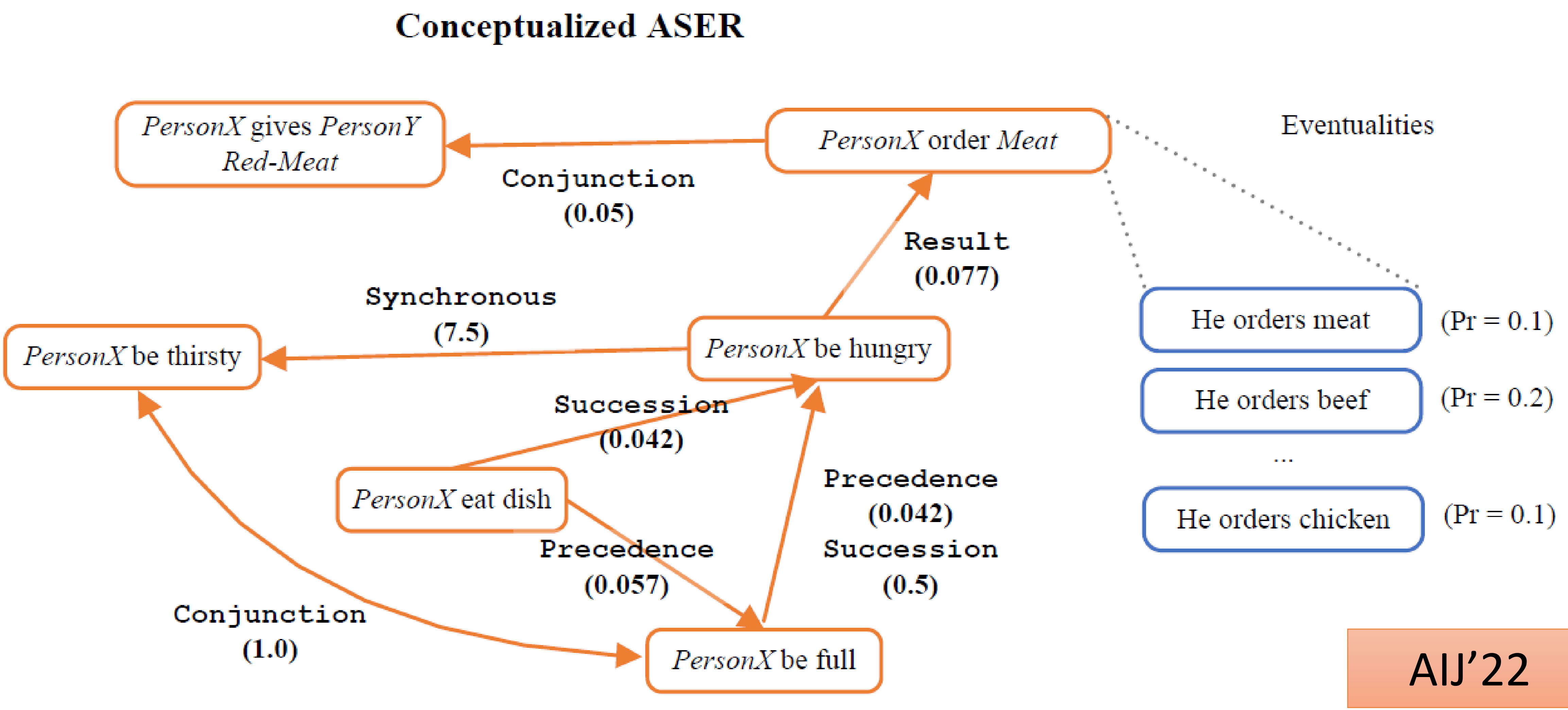
0.222 ← You, will have, a duckling

......

......

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

Conceptualization Examples

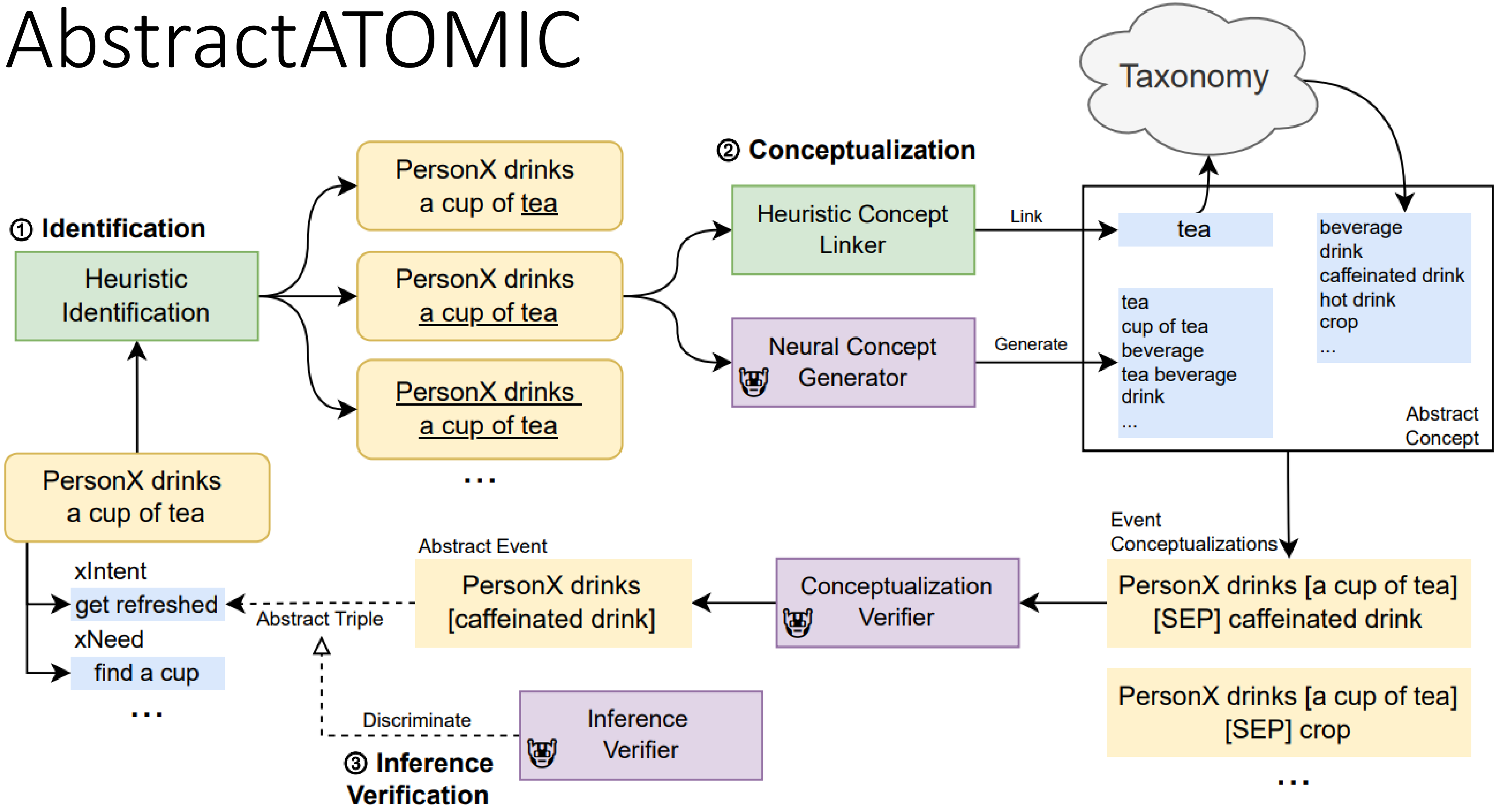


AIJ'22

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

Hongming Zhang, Xin Liu, Haojie Pan, Haowen Ke, Jiefu Ou, Tianqing Fang, Yangqiu Song: ASER: Towards large-scale commonsense knowledge acquisition via higher-order selectional preference over eventualities. Artif. Intell. 309: 103740 (2022)

Apply the Same Tech to ATOMIC: AbstractATOMIC



Under Review by AIJ

3x Larger (about 2.95M triplets) than ATOMIC

Mutian He, Tianqing Fang, Weiqi Wang, and Yangqiu Song. Acquiring and Modelling Abstract Commonsense Knowledge via Conceptualization. 2022.

Model	CSKB	a-NLI	CSQA	PIQA	SIQA	WG	Avg.
Random	-	50.0	20.0	50.0	33.3	50.0	40.7
Majority	-	50.8	20.9	50.5	33.6	50.4	41.2
RoBERTa-L (Liu et al., 2019)	-	65.5	45.0	67.6	47.3	57.5	56.6
DeBERTa-v3-L (He et al., 2023)	-	59.9	25.4	44.8	47.8	50.3	45.6
Self-talk (Shwartz et al., 2020)	-	-	32.4	70.2	46.2	54.7	-
COMET-DynGen (Bosselut et al., 2021)	ATOMIC	-	-	-	50.1	-	-
SMLM (Banerjee and Baral, 2020)	*	65.3	38.8	-	48.5	-	-
MICO (Su et al., 2022)	ATOMIC	-	44.2	-	56.0	-	-
STL-Adapter (Kim et al., 2022)	ATOMIC	71.3	66.5	71.1	64.4	60.3	66.7
Backbone: RoBERTa-Large *340M*							
RoBERTa-L (MR) (Ma et al., 2021)	ATM_{10X}	70.8	64.2	71.7	61.0	60.7	65.7
△ RoBERTa-L (MR) (Ma et al., 2021)	ATOMIC	70.8	64.2	72.1	63.1	59.2	65.9
◇ CAR-RoBERTa-L (Ours)	ATOMIC	$72.3_{\uparrow 1.5}$	$64.8_{\uparrow 0.6}$	$73.2_{\uparrow 1.1}$	$64.8_{\uparrow 1.7}$	$61.3_{\uparrow 2.1}$	$67.3_{\uparrow 1.4}$
◇ CAR-RoBERTa-L (Ours)	ATM^C	$72.7_{\uparrow 1.9}$	$66.3_{\uparrow 2.1}$	$73.2_{\uparrow 1.1}$	$64.0_{\uparrow 0.9}$	$62.0_{\uparrow 2.8}$	$67.6_{\uparrow 1.7}$
Backbone: DeBERTa-v3-Large *435M*							
DeBERTa-v3-L (MR) (Ma et al., 2021)	ATM_{10X}	74.0	65.4	73.8	59.5	73.9	69.3
△ DeBERTa-v3-L (MR) (Ma et al., 2021)	ATOMIC	76.0	67.0	78.0	62.1	76.0	71.8
◇ CAR-DeBERTa-v3-L (Ours)	ATOMIC	$78.9_{\uparrow 2.9}$	$67.2_{\uparrow 0.2}$	$\mathbf{78.6}_{\uparrow 0.6}$	$63.8_{\uparrow 1.7}$	$78.1_{\uparrow 2.1}$	$73.3_{\uparrow 1.5}$
◇ CAR-DeBERTa-v3-L (Ours)	ATM^C	$\mathbf{79.6}_{\uparrow 3.6}$	$69.3_{\uparrow 2.3}$	$\mathbf{78.6}_{\uparrow 0.6}$	$64.0_{\uparrow 1.9}$	$\mathbf{78.2}_{\uparrow 2.2}$	$\mathbf{73.9}_{\uparrow 2.1}$
Large Language Models							
GPT-3.5 (`text-davinci-003`)	-	61.8	68.9	67.8	68.0	60.7	65.4
ChatGPT (`gpt-3.5-turbo`)	-	69.3	**74.5**	75.1	**69.5**	62.8	70.2
Supervised Learning & Human Performance							
RoBERTa-L (Supervised)	-	85.6	78.5	79.2	76.6	79.3	79.8
DeBERTa-v3-L (Supervised)	-	89.0	82.1	84.5	80.1	84.1	84.0
Human Performance	-	91.4	88.9	94.9	86.9	94.1	91.2

Language models finetuned by AbstractATOMIC can significantly improve their zero-shot ability on downstream commonsense QA tasks.

ACL’23
Arxiv’23

Weiqi Wang*, Tianqing Fang*, Baixuan Xu, Chun Yi Louis Bo, Yangqiu Song, Lei Chen. CAT: A Contextualized Conceptualization and Instantiation Framework for Commonsense Reasoning. ACL, 2023

Weiqi Wang*, Tianqing Fang*, Wenxuan Ding, Baixuan Xu, Xin Liu, Yangqiu Song, Antoine Bosselut. CAR: Conceptualization-Augmented Reasoner for Zero-Shot Commonsense Question Answering. *Arxiv*, 2023.

ASER Statistics

- 1.0 (in 2019): Rule based extraction (14 Eventuality Patterns)
  - 272 millions eventualities and 206 millions relations
- 2.0 (in 2021): Discourse Parser (18 Eventuality Patterns + Wang and Lan 2015)
  - 439 millions eventualities and 649 millions relations
- Conceptualization Core (Using top 5 concepts for each detected instance):
  - Concepts: 15 millions (based on 14 millions eventualities, 1.X times)
  - Concept Relations: 224 millions (based on 53 millions eventuality relations, 4.X times)

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

Hongming Zhang, Xin Liu, Haojie Pan, Haowen Ke, Jiefu Ou, Tianqing Fang, Yangqiu Song: ASER: Towards large-scale commonsense knowledge acquisition via higher-order selectional preference over eventualities. Artif. Intell. 309: 103740 (2022)

TransOMCS: Transform ASER to ConceptNet

An open, multilingual knowledge graph

Relation: AtLocation

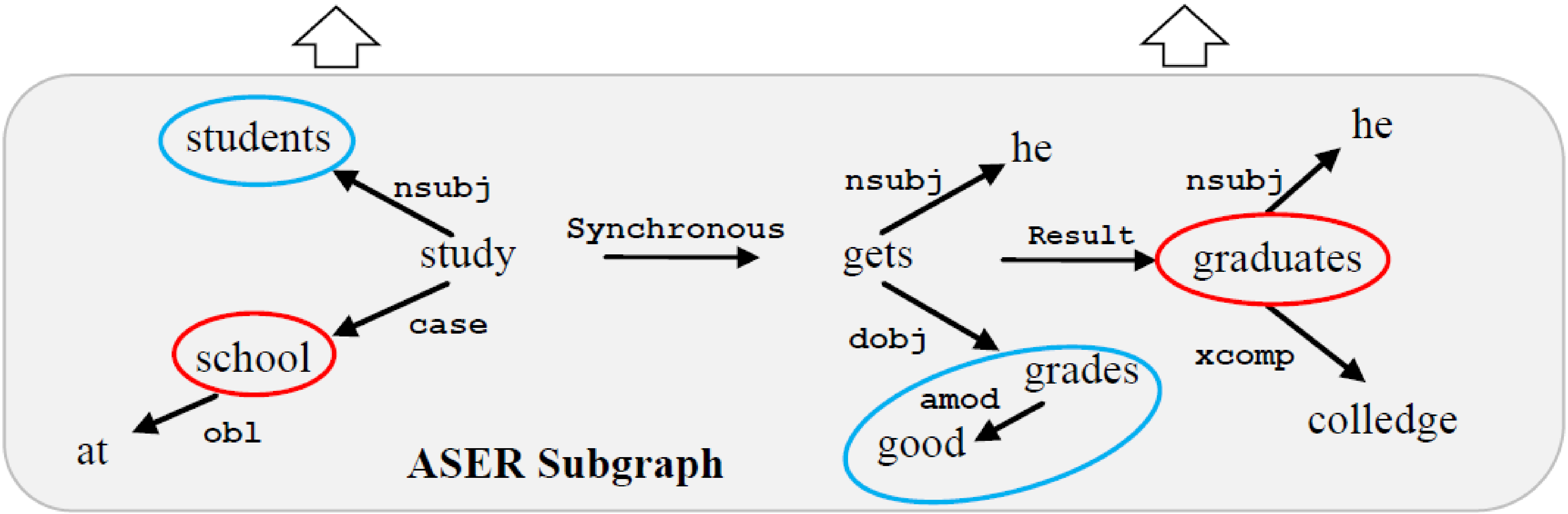
Pattern: (*H*)<-nsubj<-((*T*)-obl-(at))

Knowledge: (Student, AtLocation, School)

Relation: Causes

Pattern: (*H*)<-dobj<-()<-Result<-(*T*)

Knowledge: (Good grades, Causes, Graduate)



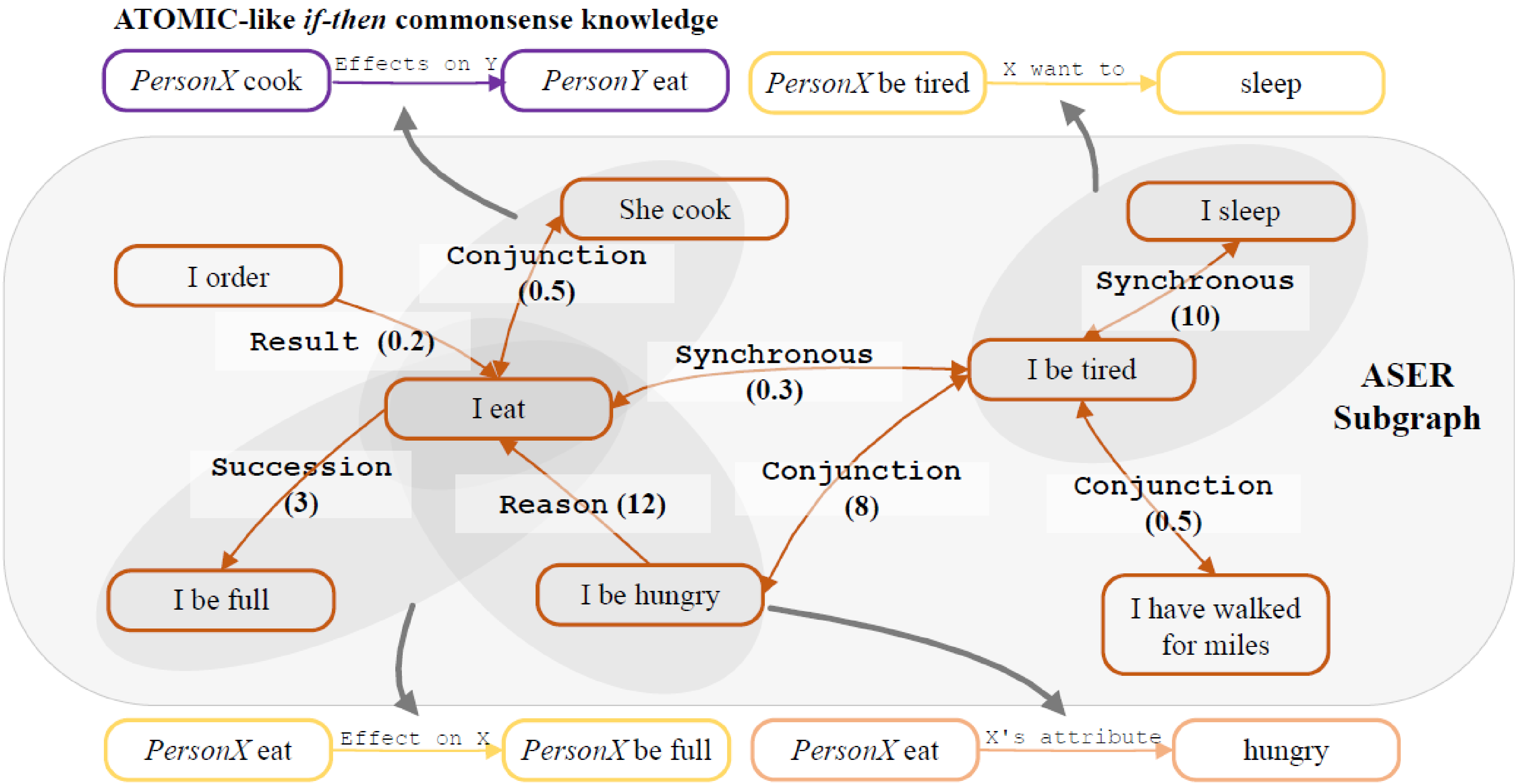
Knowledge Base Population (KBP)

IJCAI'20

100x Larger (about 18M triplets) than OMCS in ConceptNet

Hongming Zhang, Daniel Khashabi, Yangqiu Song, and Dan Roth. TransOMCS: From Linguistic Graphs to Commonsense Knowledge. International Joint Conference on Artificial Intelligence (IJCAI). 2020.

Transform ASER to ATOMIC



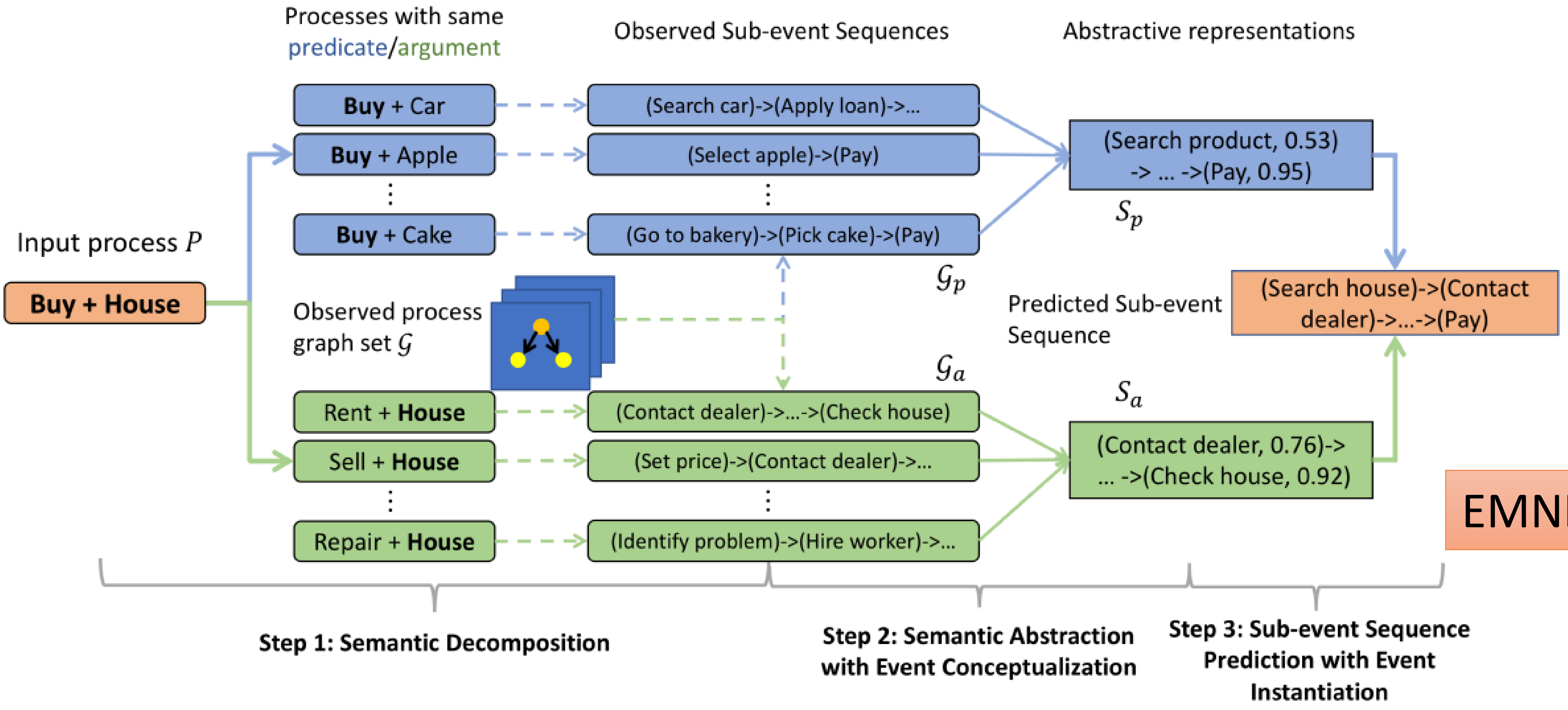
Knowledge Base Population (KBP)

WWW'21
EMNLP'21
EMNLP'22

3x Larger (about 3.4M triplets) than ATOMIC

Tianqing Fang, Quyet V. Do, Hongming Zhang, Yangqiu Song, Ginny Y. Wong and Simon See. PseudoReasoner: Leveraging Pseudo Labels for Commonsense Knowledge Base Population. Findings of EMNLP, 2022.
Tianqing Fang*, Weiqi Wang*, Sehyun Choi, Shibo Hao, Hongming Zhang, Yangqiu Song, Bin He. Benchmarking Commonsense Knowledge Base Population with an Effective Evaluation Dataset. EMNLP, 2021
Tianqing Fang, Hongming Zhang, Weiqi Wang, Yangqiu Song, and Bin He. DISCOS: Bridging the Gap between Discourse Knowledge and Commonsense Knowledge. WWW, 2021.

Incorporating More Types: **Process Induction**



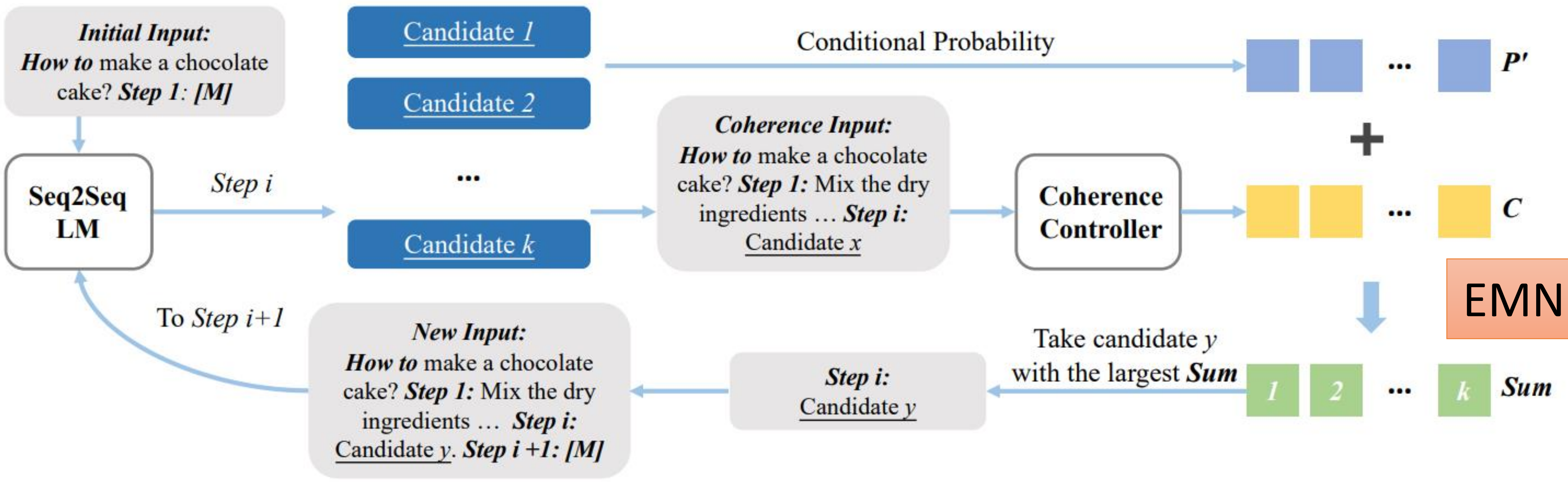
EMNLP'20

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

Incorporating More Types: **Process Generation**

Coherence controller: a candidate is coherent with the process and prior generated sub-events

- Local coherence: randomly copy a sub-event in the current process and place it at a random location.
- Global coherence: randomly choose a sub-event from other processes with a different theme and insert it at a random location

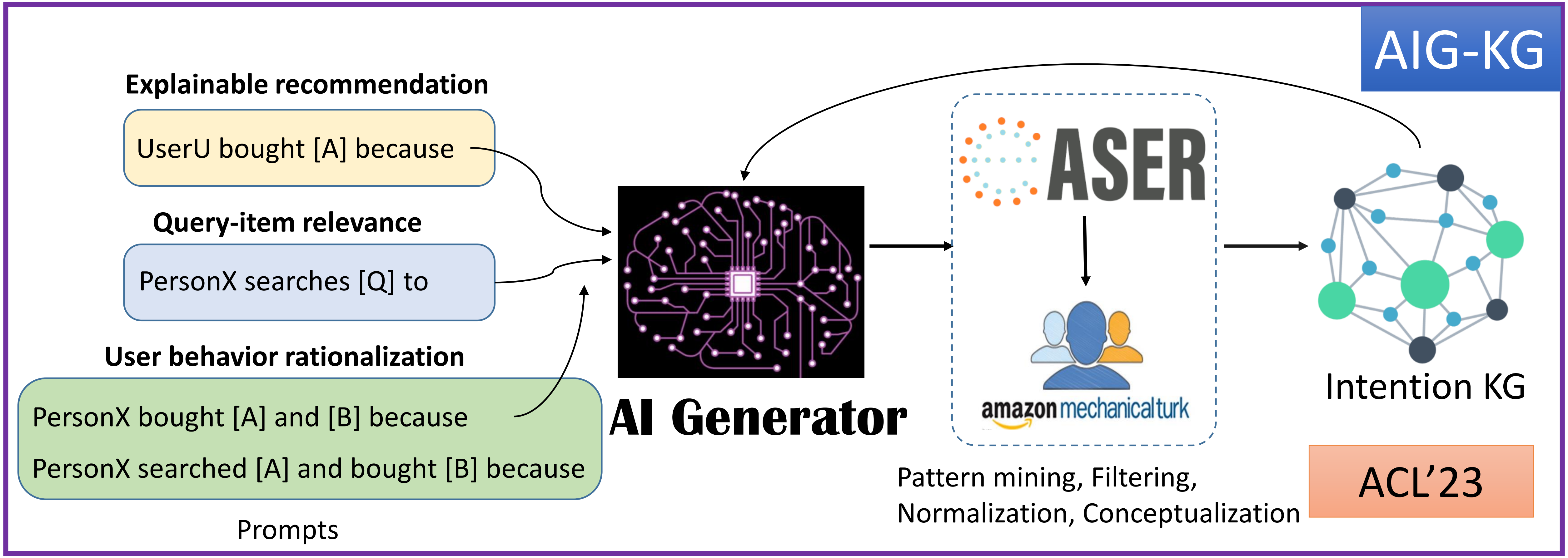


EMNLP'22

Zhaowei Wang, Hongming Zhang, Tianqing Fang, Yangqiu Song, Ginny Y. Wong, and Simon See. SubeventWriter: Iterative Sub-event Sequence Generation with Coherence Controller. EMNLP, 2022

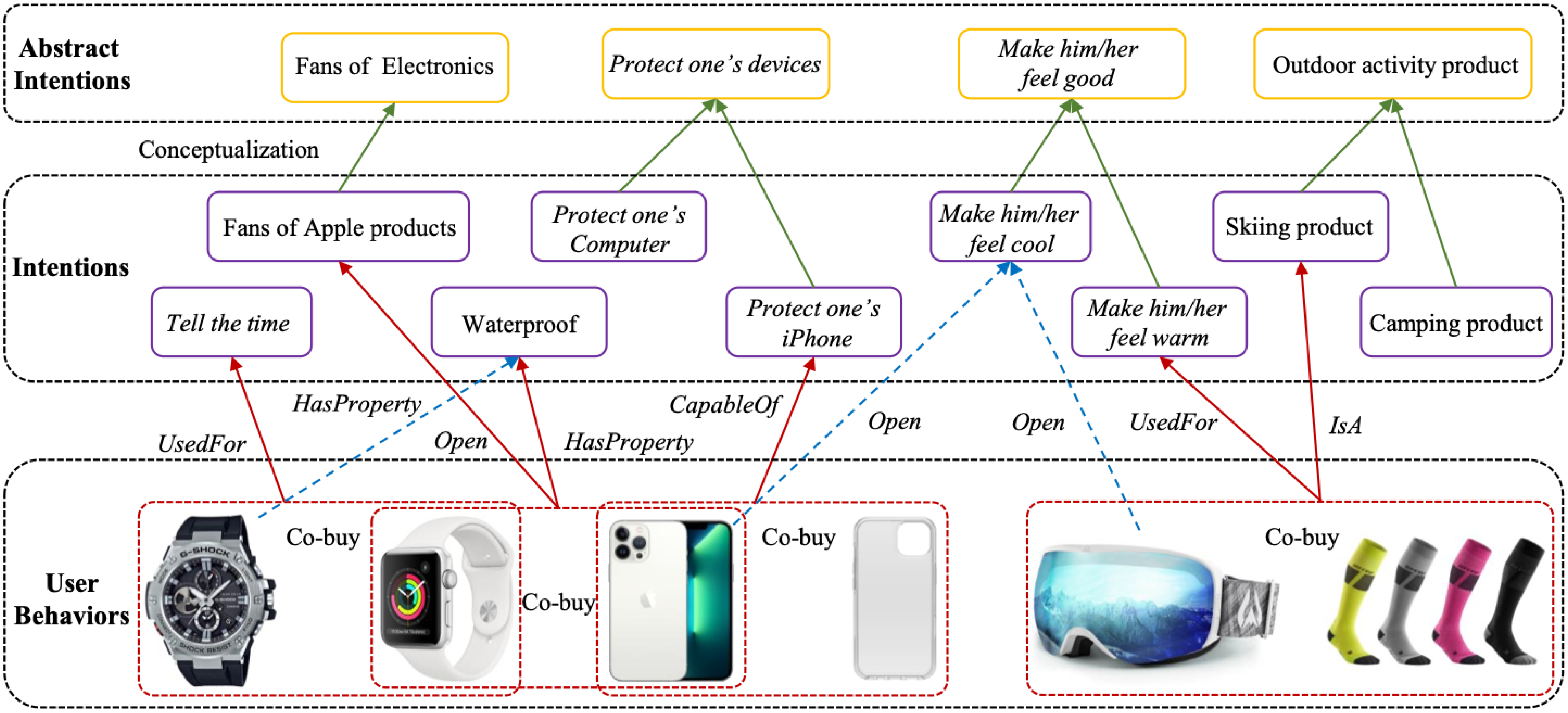
Beyond Belief/Knowledge Graph: **FolkScope**

- Intention Knowledge Graph for E-commerce Commonsense Discovery



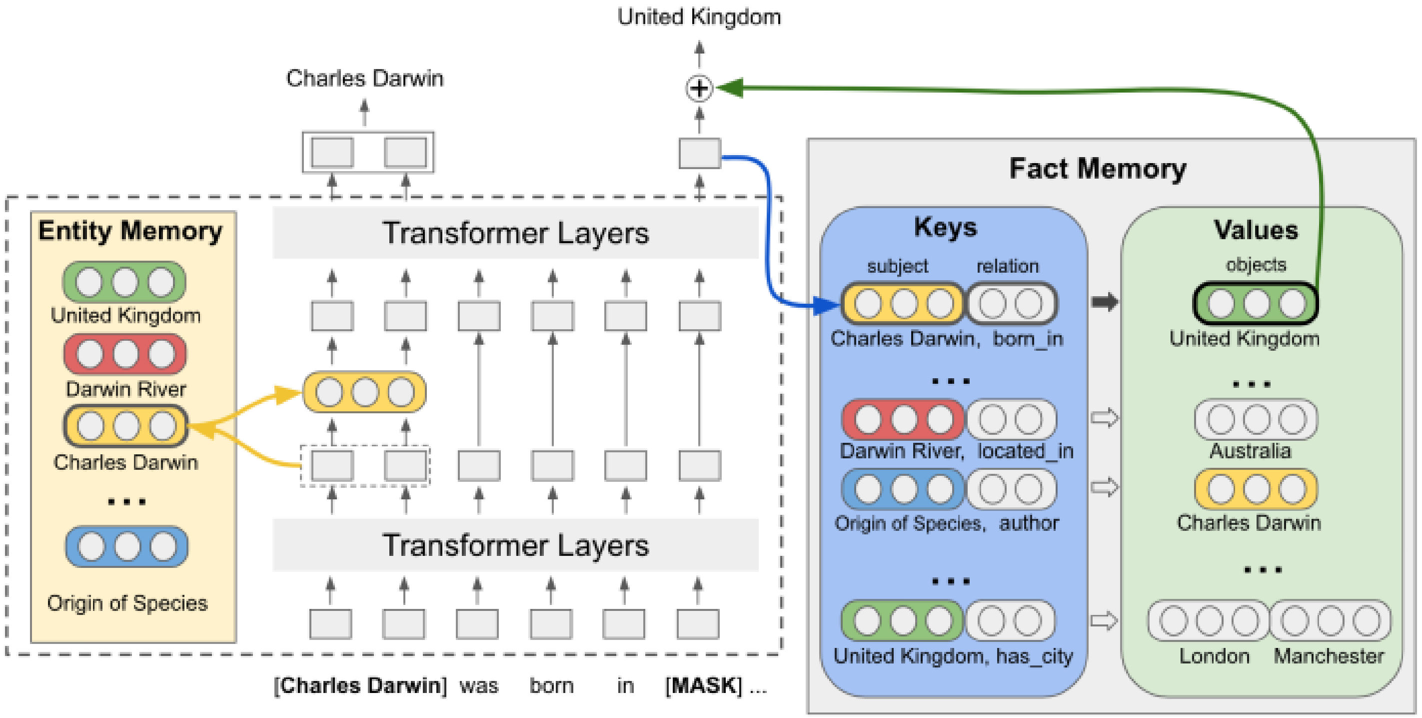
Changlong Yu, Weiqi Wang, Xin Liu, Jiaxin Bai, Yangqiu Song, Zheng Li, Yifan Gao, Tianyu Cao, and Bing Yin. FolkScope: Intention Knowledge Graph Construction for E-commerce Commonsense Discovery. Findings of ACL. 2023.

Beyond Belief/Knowledge Graph: **FolkScope**



How to help LLMs?

Entities/Facts as Memories



Neural Graph Databases

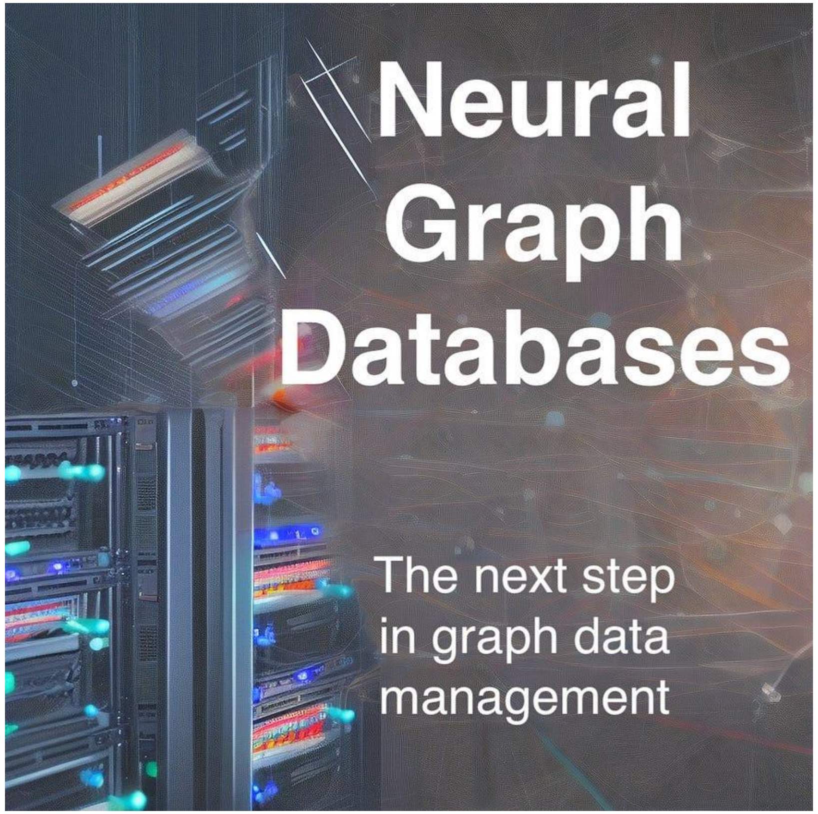


Figure taken from: Pat Verga, Haitian Sun, Livio Baldini Soares, William W. Cohen: Adaptable and Interpretable Neural Memory Over Symbolic Knowledge. NAACL-HLT 2021: 3678-3691

Figure taken from: https://towardsdatascience.com/neural-graph-databases-cc35c9e1d04f

Logical Queries over Knowledge Graphs

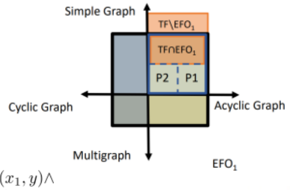
Data

Benchmarking EFO-1 (Existential **F**irst-**O**rder **Q**ueries with Single Free Variable)

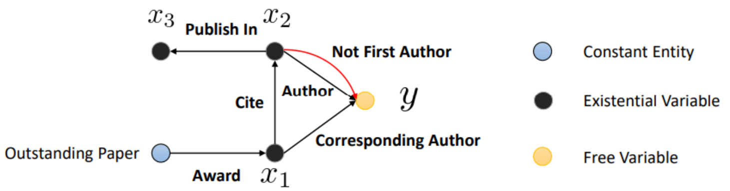
CQA Dataset	Support Operators									Support EPFO	Support EFO-1	Num. of Forms	Num. of Test Query Types
	e	p	i	I	u	U	n	d	D				
Q2B dataset [15]	✓	✓	✓	✓	✗	✗	✗	✗	✗	✓*	✗	1	9
HypE dataset [6]	✓	✓	✓	✓	✗	✗	✗	✗	✗	✓*	✗	1	9
BetaE dataset [16]	✓	✓	✓	✓	✗	✗	✓	✗	✗	✓*	✓*	2	14
EFO-1-QA (ours)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	9	301

NeurIPS'21

EFO-1 **queries with cycles**

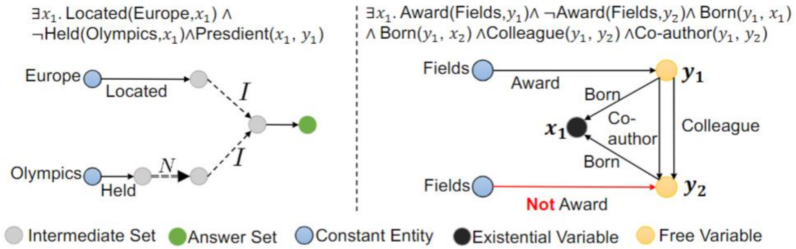


$\exists x_1 \exists x_2 \exists x_3, \mathbf{Award}(\mathrm{OutstandingPaper}, x_1) \wedge \mathbf{CorrespondingAuthor}(x_1, y) \wedge \mathbf{Cite}(x_1, x_2) \wedge \mathbf{PublishIn}(x_2, x_3) \wedge \mathbf{Author}(x_2, y) \wedge \neg\mathbf{FirstAuthor}(x_2, y)$



Arxiv'23

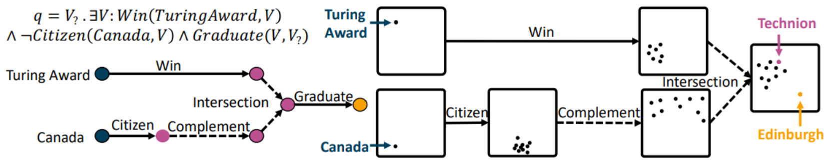
EFO-K more than one variables



Arxiv'23

Models

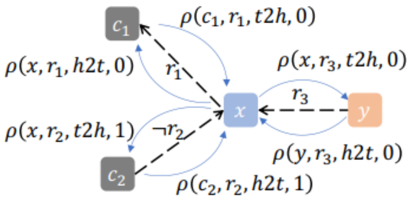
Particle filtering of logical sequential queries



(A) Computation Graph (B) 2-Dimensional Illustration of Query2Particles Approach

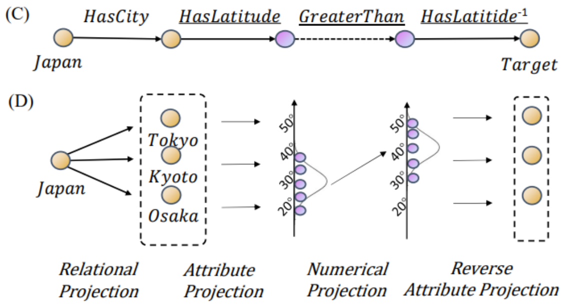
NAACL'22

Learning in the inference step as a **GNN (one-hop logical inference based MPNN)**



ICLR'23

Number and attribute queries



KDD'23

Hang Yin, Zihao Wang, and Yangqiu Song. EFO_k-CQA: Towards Knowledge Graph Complex Query Answering beyond Set Operation. Arxiv 2023
Jiaxin Bai, Chen Luo, Zheng Li, Qingyu Yin, Bing Yin, Yangqiu Song. Knowledge Graph Reasoning over Entities and Numerical Values. KDD 2023
Hang Yin, Zihao Wang, and Yangqiu Song. Rethinking Existential First Order Queries and their Inference on Knowledge Graphs. Arxiv 2023
Zihao Wang, Yangqiu Song, Ginny Y. Wong, and Simon See. Logical Message Passing Networks with One-hop Inference on Atomic Formulas. In The Eleventh International Conference on Learning Representations, ICLR 2023
Jiaxin Bai, Zihao Wang, Hongming Zhang, and Yangqiu Song. Query2Particles: Knowledge Graph Reasoning with Particle Embeddings. In Findings of the Association for Computational Linguistics: NAACL-HLT 2022, 2022
Zihao Wang, Hang Yin, and Yangqiu Song. Benchmarking the Combinatorial Generalizability of Complex Query Answering on Knowledge Graphs. In NeurIPS Datasets and Benchmarks Track, 2021

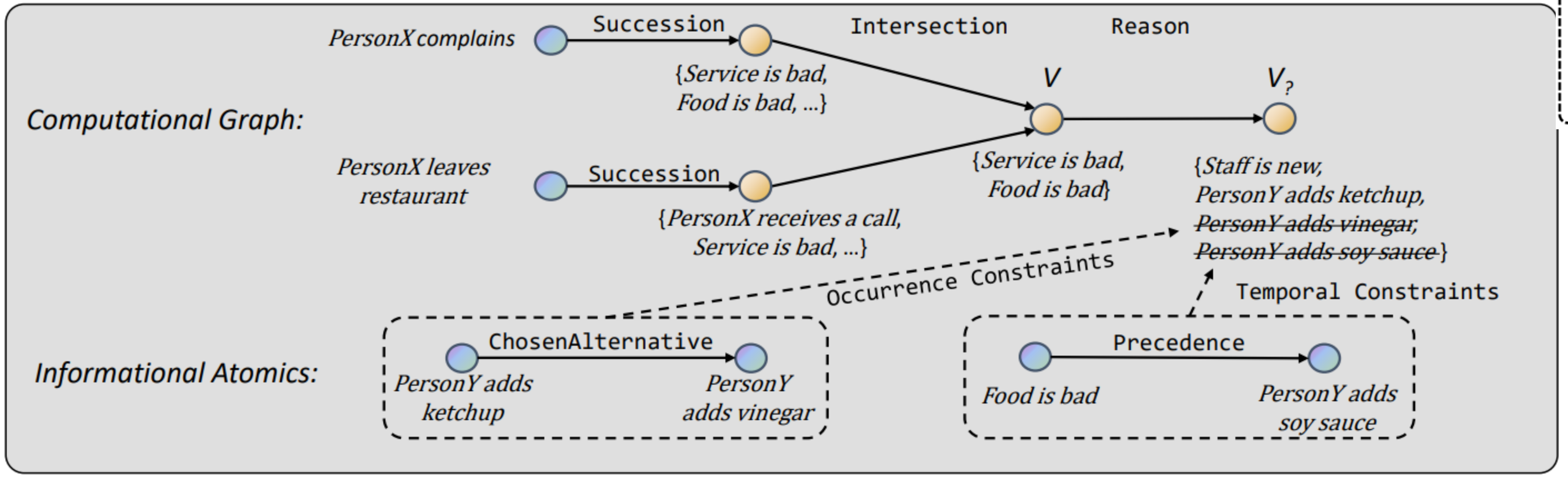
Logical Queries on ASER with Logical Constraints

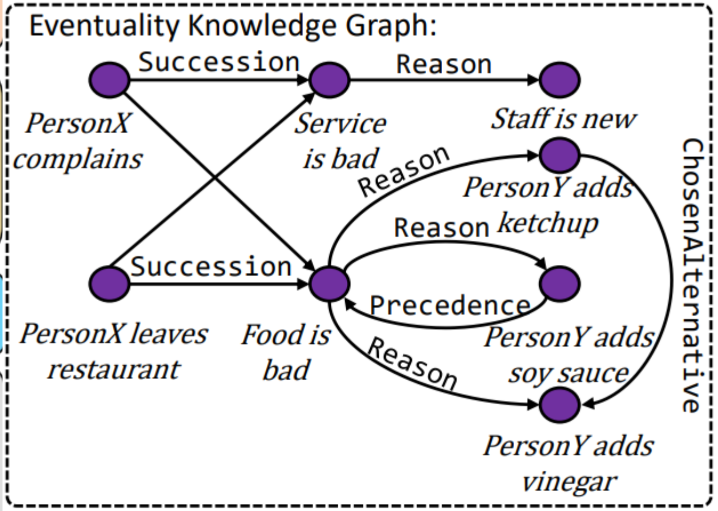
Logical Query: $q = V_?\,.\,\exists V$: Succession(*PersonX complains, V*) $\wedge$ Succession(*PersonX leaves restaurant, V*) $\wedge$ Reason(*V*, $V_?$) $\wedge$ Precedence(*Food is bad, PersonY adds soy sauce*) $\wedge$ ChosenAlternative(*PersonY adds ketchup, PersonY adds vinegar*)

Implicit Constraints:

$$\tau(V) \prec \tau(\textit{PersonX complains}) \wedge \eta(V) \wedge \eta(\textit{PersonX complains})$$
$$\wedge\, \tau(V) \prec \tau(\textit{PersonX leaves restaurant}) \wedge \eta(V) \wedge \eta(\textit{PersonX leaves restaurant})$$
$$\wedge\, \eta(V) \wedge \eta(V_?) \wedge (\eta(V_?) \rightarrow \eta(V)) \wedge \tau(V) \succ \tau(V_?)$$
$$\wedge\, \eta(\textit{Food is bad}) \wedge \eta(\textit{PersonY adds soy sauce}) \wedge \tau(\textit{Food is bad}) \prec \tau(\textit{PersonY adds soy sauce})$$
$$\wedge\, \eta(\textit{PersonY adds ketchup}) \wedge \neg\, \eta(\textit{PersonY adds vinegar})$$

Query Types: `(p,(i,(p,(e)),(p,(e))))`





Precedence(V1, V2)
V1 occurs before V2

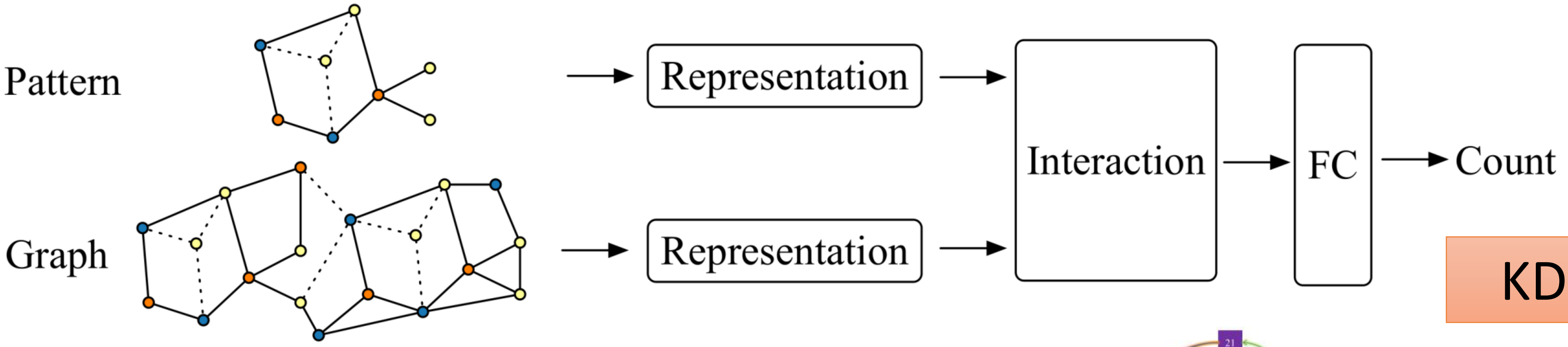
ChosenAlternative(V1, V2)
Instead of V2 occurs, V1 occurs

Arxiv'23

Jiaxin Bai, Xin Liu, Weiqi Wang, Chen Luo, Yangqiu Song. Complex Query Answering on Eventuality Knowledge Graph with Implicit Logical Constraints. Arxiv 2023

Knowledge Logical Counting or Subgraph Counting

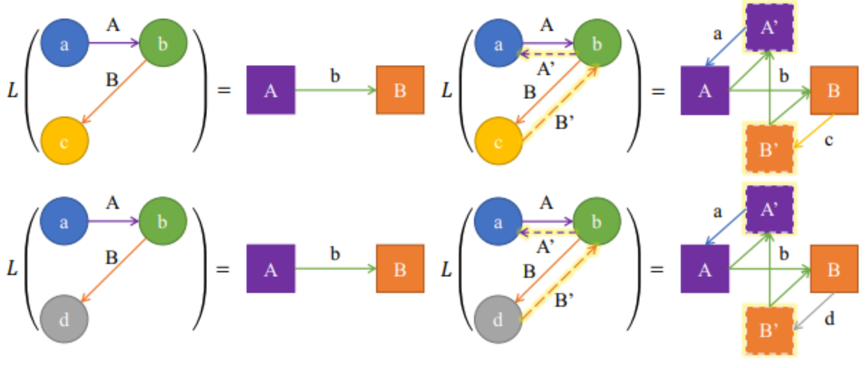
- Such KG query needs global memory and inference
- Subgraph counting is NP-Complete; Very Difficult for LLMs (Transformer) to conduct



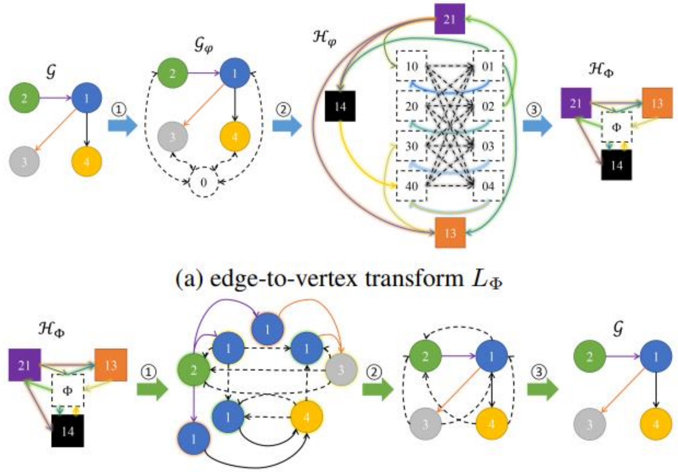
KDD'20

Dual graph counting

AAAI'22



(a) Non-injective case (b) Adding reversed edges



(a) edge-to-vertex transform L_Φ

Global dummy node

ICML'22

Xin Liu, Jiayang Cheng, Yangqiu Song, and Xin Jiang. Boosting Graph Structure Learning with Dummy Nodes. In International Conference on Machine Learning (ICML), 2022.
Xin Liu and Yangqiu Song. Graph Convolutional Networks with Dual Message Passing for Subgraph Isomorphism Counting and Matching. In the AAAI Conference on Artificial Intelligence (AAAI), 2022.
Xin Liu, Haojie Pan, Mutian He, Yangqiu Song, Xin Jiang, and Lifeng Shang. Neural Subgraph Isomorphism Counting. In ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2020.

To Conclude

- We build new types of graphs
  - Related to activity (or process), state, and event
  - Beyond knowledge and believes

- We develop data sources and packages to support
  - Knowledge grounding
  - Complex knowledge queries
  - Logical reasoning
  - Improved zero-shot learning or indirect supervision for LLMs
  - Trustworthy LLMs

Thanks to Key Contributors

Hongming Zhang

Xin Liu

Haojie Pan

Tianqing Fang

Mutian He

Weiqi Wang

Jiaxin Bai

Zhaowei Wang

Changlong Yu

Zihao Wang

Hang Yin@THU

And many other PhD/MPhil students, UG interns, visiting students, and industrial collaborators!

Thank you for your attention! ☺

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

THE DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

KnowComp Group
Understanding the World by Computational Knowledge