Agenda

08:30 PST	10 mins	**Introduction to commonsense knowledge (Filip)**
08:40 PST	25 min	**Part I - Axiomatization of commonsense knowledge (Mayank)**
09:05 PST	40 min	**Part II - Consolidating commonsense knowledge (Filip)**
09:45 PST	15 min	Break
10:00 PST	45 min	**Part III - Extracting and contextualizing commonsense knowledge (Simon)**
10:45 PST	45 min	**Part IV - Language models, QA, and evaluation challenges (Antoine)**
11:30 PST	15 min	**Way forward: KGs+LMs+axioms? (Filip)**

The way forward

Filip Ilievski

Why axiomatize commonsense?

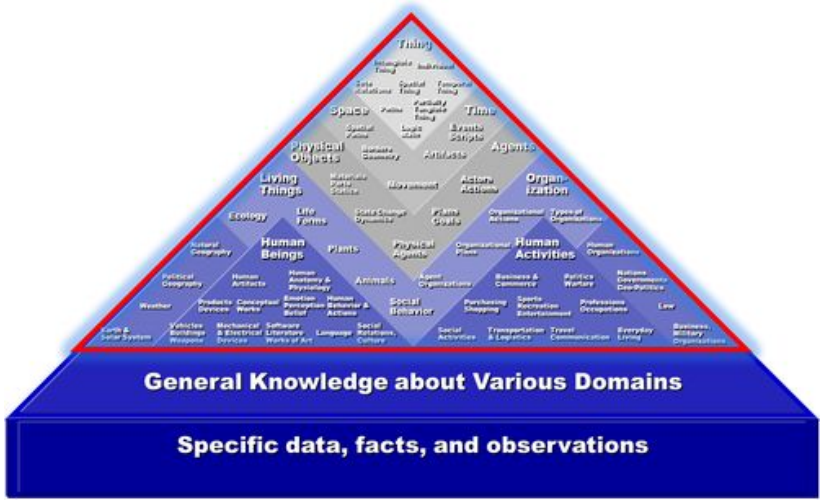
- **Fundamental reasons:** is our conception of common sense sound and complete? Put another way, are there examples of common sense that can't be modeled by one or more of the proposed axioms?

- Axiomatization can provide **explainability** and also help us think about common sense from a cognitive-science perspective

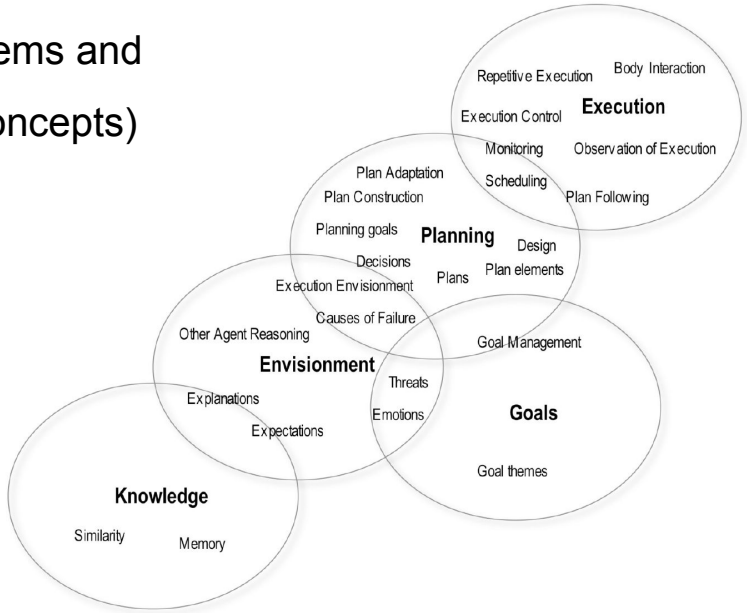
- Axiomatization is a type of **top-down knowledge** that has become increasingly necessary to complement bottom-up knowledge

Top-down axiomatization

- Scales, time, spaces and dimension, material, causal connections, (in other domains) force, shape, systems and functionality, hitting, abrasion, wear (and related concepts)
- Competency vs. coverage theories
- Naive physics vs. psychology theories

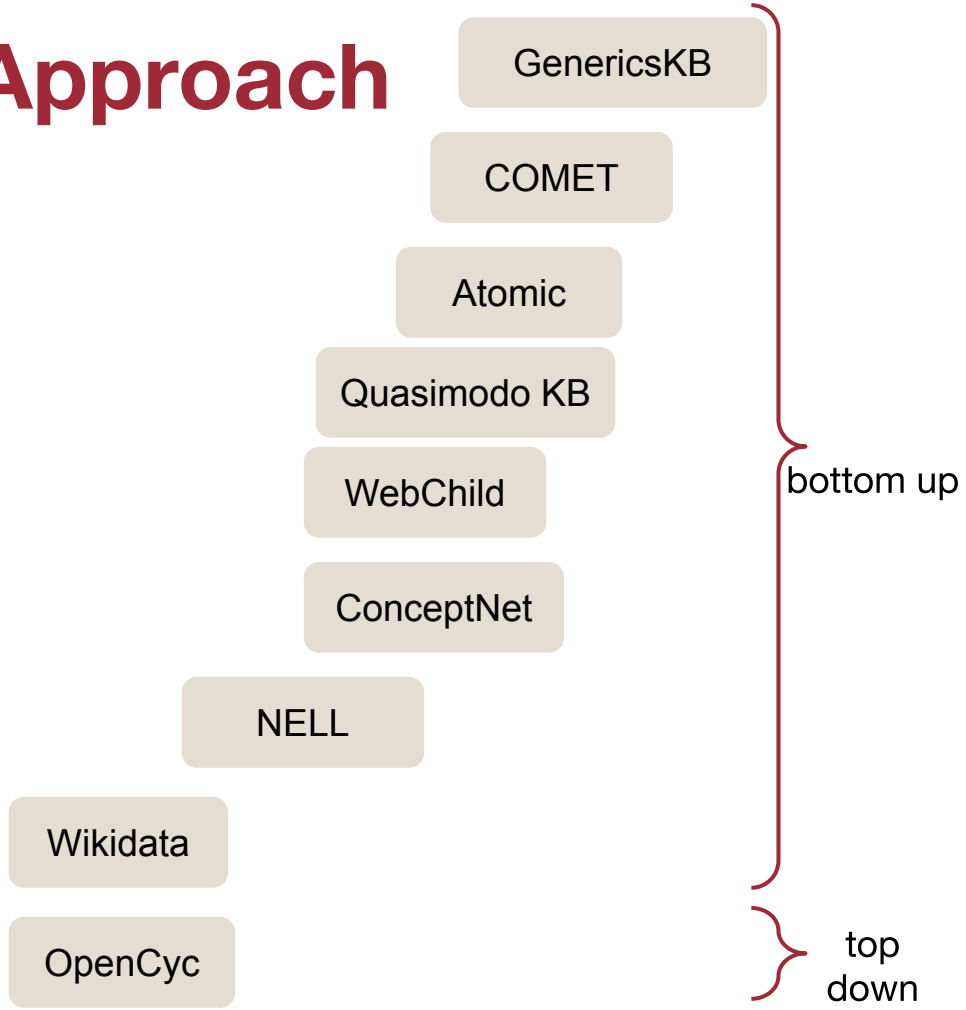


CYC (Lenat, 1995)



Gordon and Hobbs (2004)

Design Approach



Aspects Of Common Sense Knowledge

Representation

- symbolic
- natural language
- neural

Acquisition method

- expert input
- crowdsourcing
- information extraction, machine learning

Knowledge type

- entities and actions
- inferential/rules

Topic

- general
- social

GenericsKB

COMET

Atomic

Quasimodo KB

WebChild

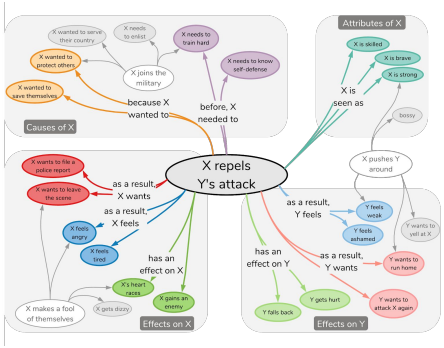
ConceptNet

NELL

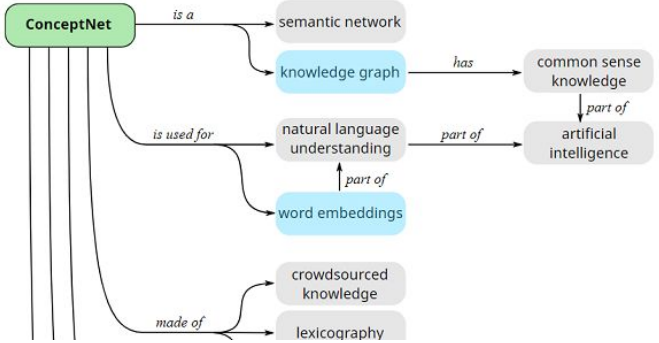
Wikidata

OpenCyc

Consolidating Knowledge Graphs



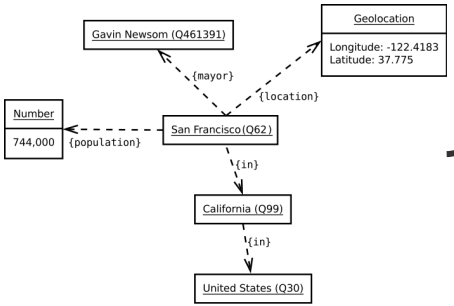
ATOMIC (Sap et al. 2019)



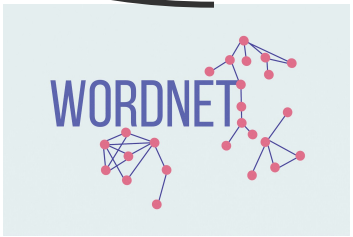
ConceptNet (Speer, Chin and Havasi 2017)

CSKG

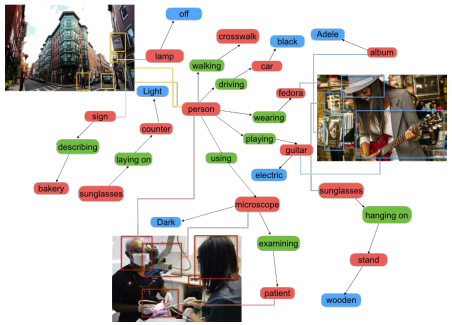
CSKG (Ilievski et al. 2020)



Wikidata (Vrandecic and Krotzsch 2014)



WordNet (Miller 1995)

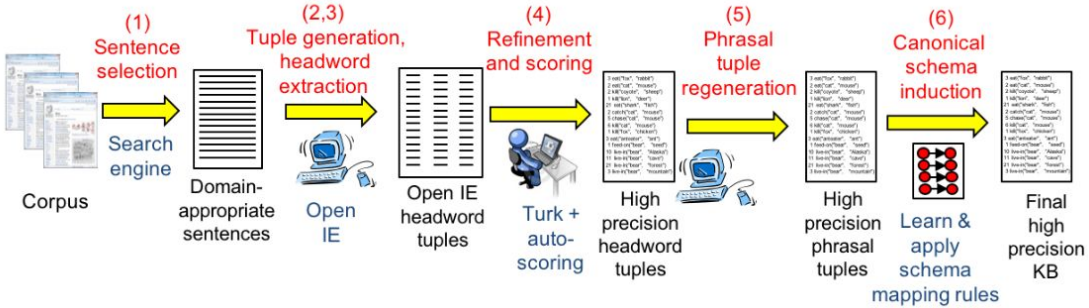


Visual Genome (Krishna et al. 2017)

Dimensions (=Bottom-up axioms!)

Dimension	ATOMIC	ConceptNet	WebChild	Other	Wikidata
taxonomic		IsA InstanceOf MannerOf	hasHypernymy	perspective_on (FN) inheritance (FN) hypernym (WN)	subClassOf instanceOf description
part-whole		PartOf HasA MadeOf AtLocation*	physicalPartOf memberOf substanceOf	HasPart (HP) meronym (WN) holonym (WN)	has part member of material used
spatial		AtLocation* LocatedNear	location spatial		location anatomical location
creation		CreatedBy			creator
utility		ReceivesAction UsedFor CapableOf ¬NotCapableOf	hassynsetmember activity participant	using (FN)	used by use uses
desire/goal	xIntent xWant oWant	CausesDesire MotivatedByGoal Desires ¬NotDesires ObstructedBy			
quality	xAttr	HasProperty ¬NotHasProperty SymbolOf	shape size color taste_property temperature	frame_element (FN)	color has quality
comparative			*6.3k relations*		
temporal	xNeed xEffect oEffect xReact oReact	HasFirstSubevent HasLastSubevent HasSubevent HasPrerequisite Causes Entails	time emotion prev next	subframe (FN) precedes (FN) inchoative_of (FN) causative_of (FN)	has cause has effect

Increasing the coverage of CSK from text



Practical human knowledge	**(car, slip on, ice)**
Problems linked to a subject	**(pen, can, leak)**
Emotions linked to events	**(divorce, can, hurt)**
Human behaviors	**(ghost, scare, people)**
Visual facts	**(road, has_color, black)**
Cultural knowledge (USA)	**(school, have, locker)**
Comparative knowledge	**(light, faster than, sound)**

Pipeline Example Outputs:

Inputs: corpus + vocabulary + types

1. Sentence selection:
"In addition, green leaves have chlorophyll."

2. Tuple Generation:
("green leaves" "have" "chlorophyll")

3. Headword Extraction:
("leaf" "have" "chlorophyll")

4. Refinement and Scoring:
("leaf" "have" "chlorophyll") @0.89 (score)

5. Phrasal tuple generation:
("leaf" "have" "chlorophyll") @0.89 (score)
("green leaf" "have" "chlorophyll") @0.89 (score)

6. Relation Canonicalization:
("leaf" "have" "chlorophyll") @0.89 (score)
("green leaf" "have" "chlorophyll") @0.89 (score)
("leaf" "contain" "chlorophyll") @0.89 (score)
("green leaf" "contain" "chlorophyll") @0.89 (score)

Facts about 'polar bear'

Click on a property for more details on the statement. Click on a column header to use it as a sorting key.

Show scores as: absolute | percentiles

Filter by source: ConceptNet | Quasimodo

Property	Score	Plausible	Typical	Remarkable	Salient	Source
adapt in summer	0.83	0.19	0.54	0.15	0.15	Quasimodo
adapt to environment	0.83	0.52	0.38	0.93	0.76	Quasimodo
adapt to tundra	0.83	0.10	0.40	0.14	0.10	Quasimodo
be at in arctic	0.67	0.17	0.29	0.93	0.18	ConceptNet
be at risk	0.83	0.62	0.54	0.88	0.80	Quasimodo
be at zoo	0.75	0.10	0.03	0.39	0.37	ConceptNet
be found in arctic	0.91	0.34	0.44	0.51	0.32	Quasimodo
be important to canada	0.92	0.43	0.70	0.27	0.29	Quasimodo
be in danger	0.82	0.91	0.93	0.77	0.97	Quasimodo
be under threat	0.83	0.83	0.80	0.85	0.95	Quasimodo
be used to snow	0.46	0.20	0.51	0.17	0.19	ConceptNet
be white	0.46	0.07	0.68	0.16	0.13	ConceptNet

Enriching CSK with context

Concept-facets dependencies: $\forall(s,p) \in \mathcal{S} \times \mathcal{P}$

$$\text{Typical}(s,p) \Rightarrow \text{Plausible}(s,p)$$

$$\text{Salient}(s,p) \Rightarrow \text{Plausible}(s,p)$$

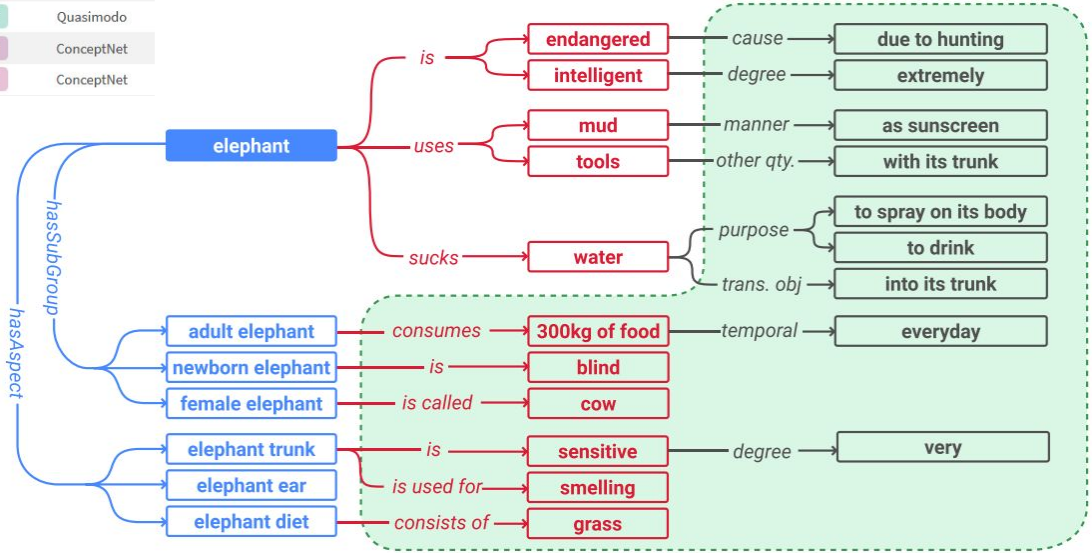
$$\text{Typical}(s,p) \wedge \text{Remarkable}(s,p) \Rightarrow \text{Salient}(s,p)$$

Sibling dependencies: $\forall(s_1,p) \in \mathcal{S} \times \mathcal{P}, \forall s_2 \in \text{siblings}(s_1)$

$$\text{Remarkable}(s_1,p) \Rightarrow \neg\text{Remarkable}(s_2,p)$$

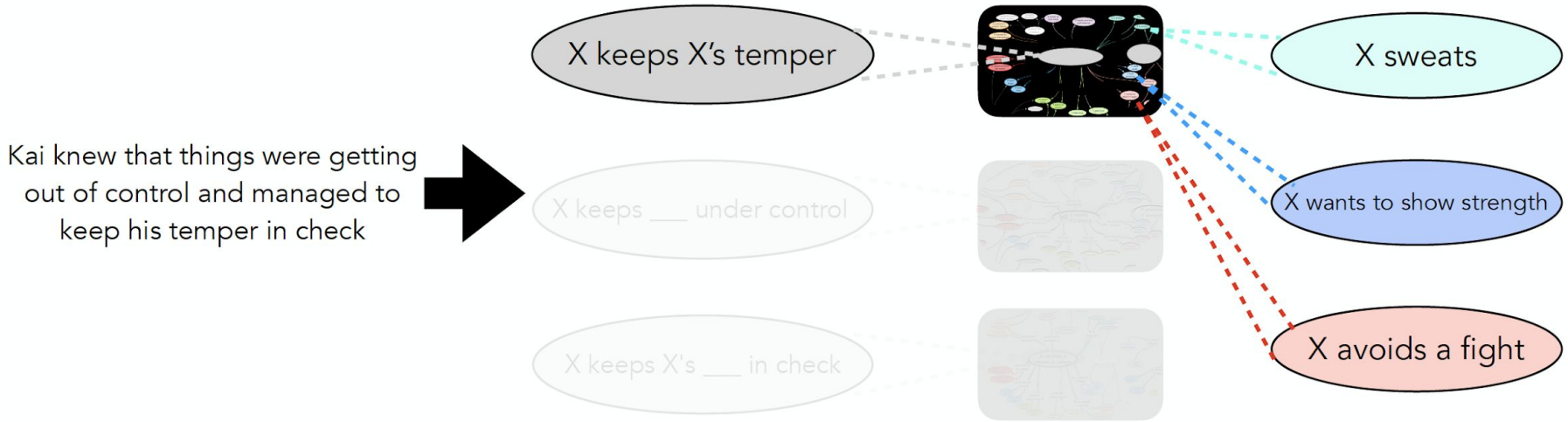
$$\text{Typical}(s_1,p) \Rightarrow \neg\text{Remarkable}(s_2,p)$$

$$\neg\text{Plausible}(s_1,p) \wedge \text{Plausible}(s_2,p) \Rightarrow \text{Remarkable}(s_2,p)$$



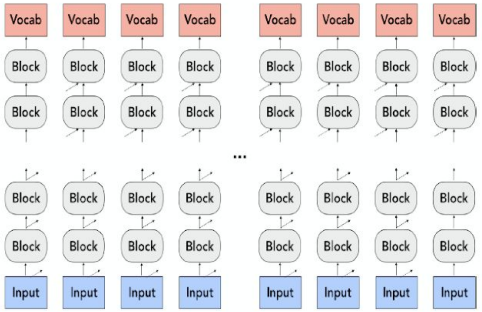
Limitations of (symbolic) CSKGs

- Situations rarely found **as-is** in commonsense knowledge graphs
- Connecting to knowledge graphs can yield **incorrect** nodes
- Suitable nodes are often **uncontextualized**



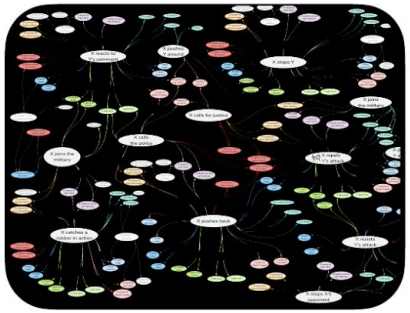
Alleviating the KG limitations with commonsense transformers

- Learn implicit knowledge at scale from language models and web-scale text
- **Learn explicit structure of knowledge** from symbolic knowledge graphs
- Resulting knowledge model **generalizes structure** to other concepts



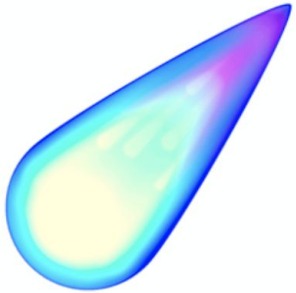
Pre-trained Language Model

+

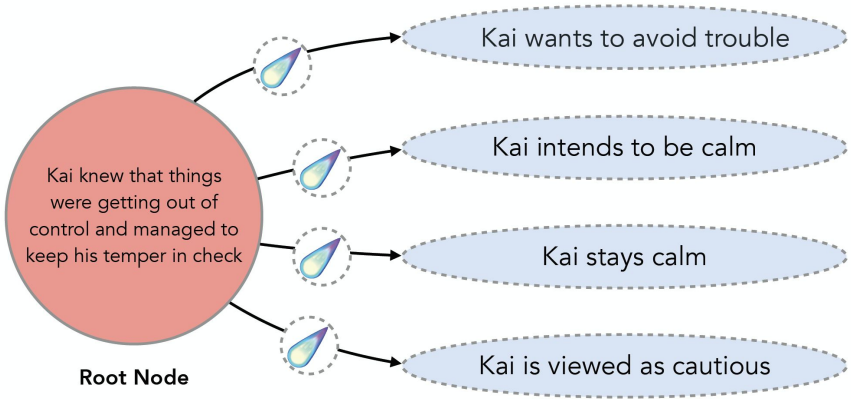


Seed Knowledge Graph Training

=

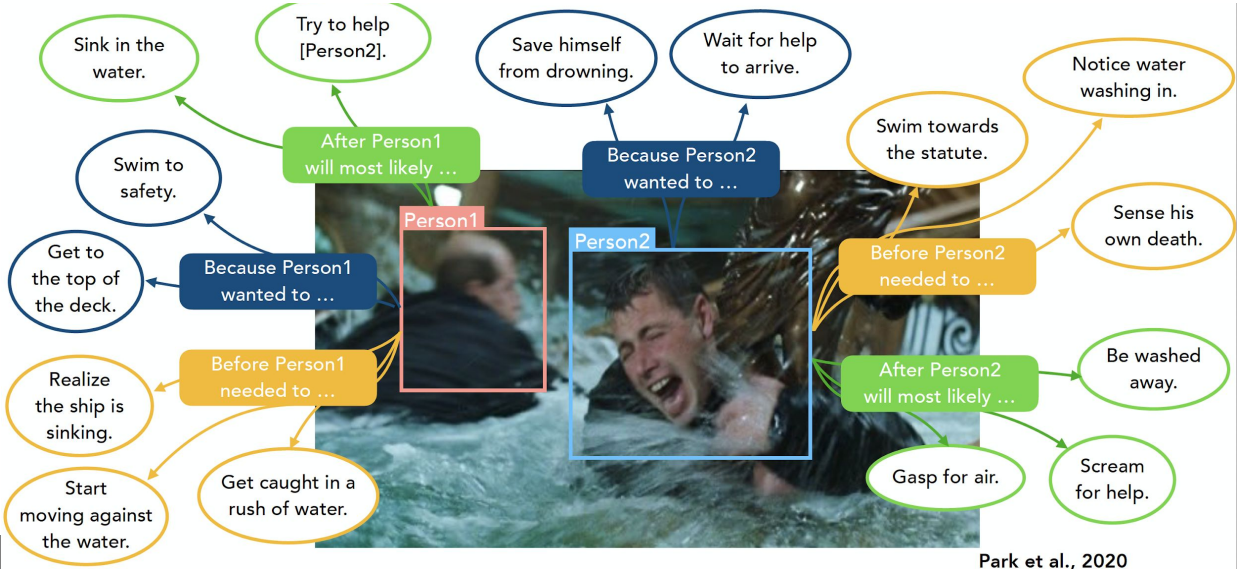


COMET



Bosselut et al., 2020

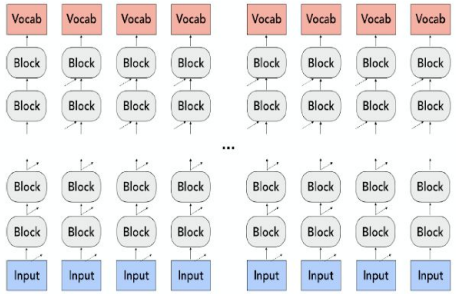
Neural CSKGs

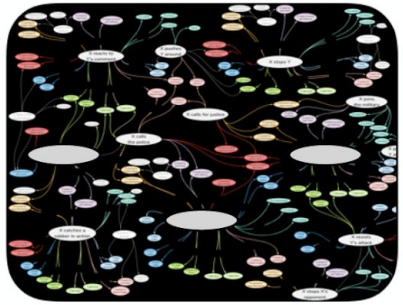


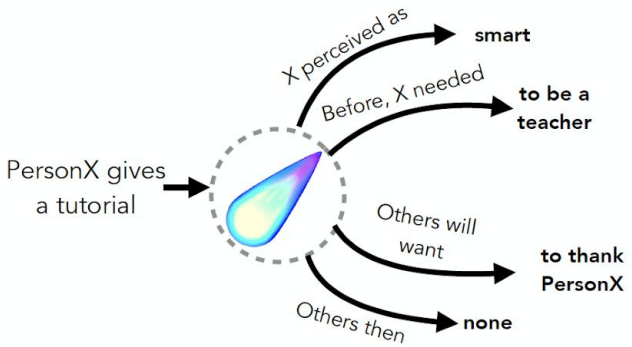
Park et al., 2020

Limitations of Language Models

- Base Self-supervised Model
  - biases in language model will be in knowledge model
- Seed Knowledge Graph
  - bias, relations, schema
- Generation Algorithm
  - diversity, mode collapse

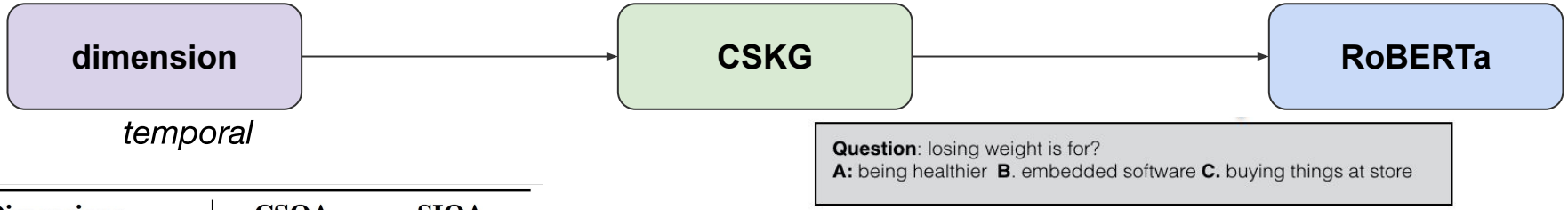




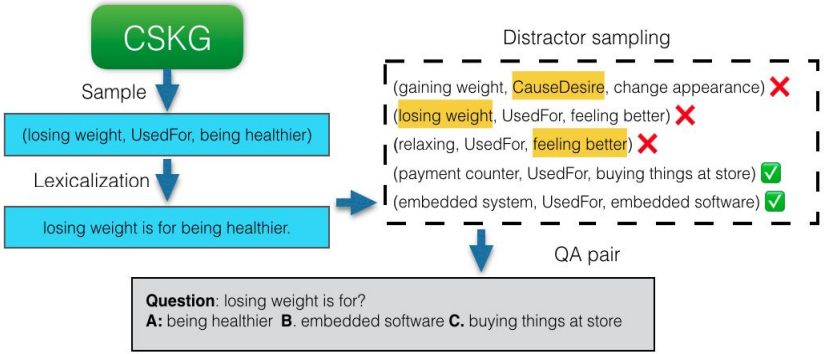


Combination: Pre-training language models with KGs and dimensions

Combination: Pre-training language models with KGs and dimensions



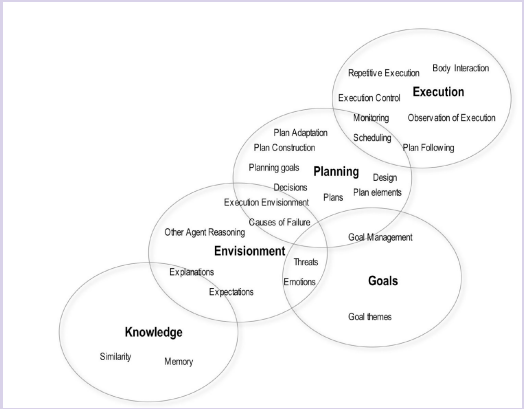
Dimensions	CSQA	SIQA
Baseline	45.0	47.3
+part-whole	63.0(±1.4)	52.6(±1.9)
+taxonomic	62.6(±1.4)	52.2(±1.6)
+lexical	49.9(±2.9)	49.0(±0.4)
+distinctness	57.2(±0.5)	50.2(±1.5)
+similarity	61.4(±0.8)	53.5(±0.6)
+quality	65.7(±0.5)	60.0(±0.7)
+utility	**67.4(±1.0)**	54.8(±0.7)
+creation	49.9(±1.1)	47.8(±0.2)
+temporal	67.3(±0.3)	**62.6(±0.9)**
+relational-other	58.2(±1.7)	51.3(±1.7)
+spatial	63.3(±0.2)	53.1(±0.3)
+desire/goal	65.0(±1.8)	60.0(±0.6)
+all	66.2(±1.4)	61.0(±0.7)

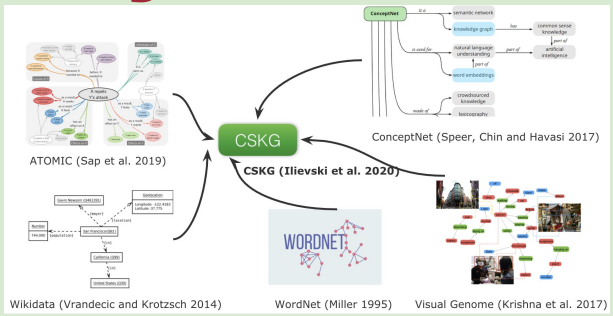


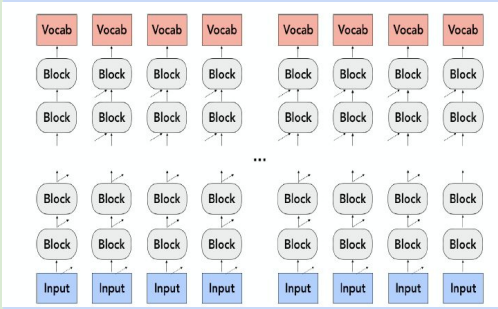
Dimensions	Dev
part-whole	67.5
taxonomic	57.0
lexical	90.1
distinctness	77.3
similarity	65.6
quality	45.5
utility	67.9
creation	82.4
temporal	47.2
relational-other	37.6
spatial	56.9
desire/goal	48.0

Ma et al. (2021), Ilievski et al. (2021)

Way Forward



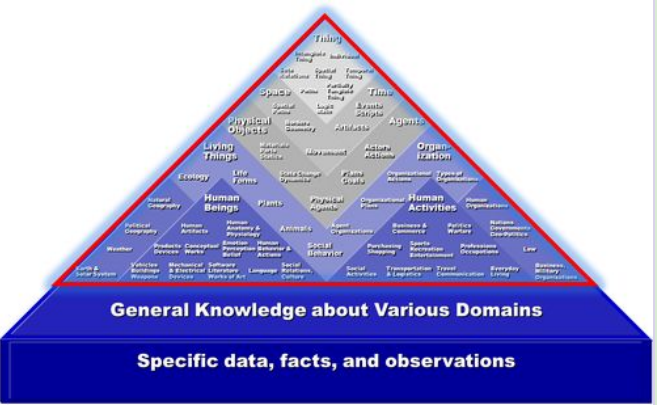


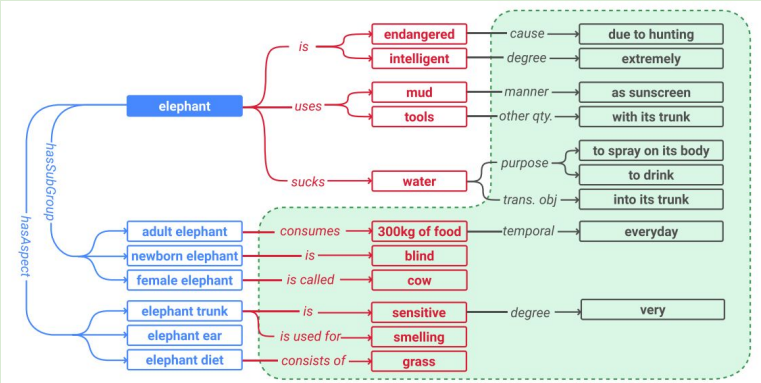


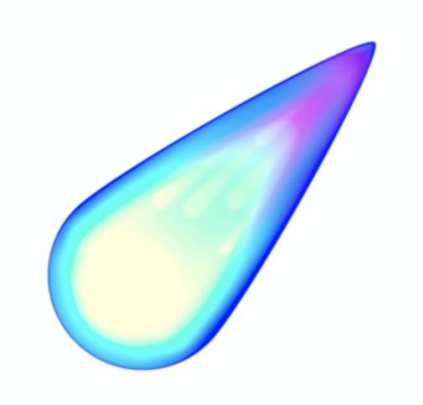
Axioms/ dimensions

+ Knowledge graphs +

Language models







Confirmed Keynote Speakers

Yejin Choi
University of Washin...

Joshua Tenenbaum
MIT

Organizers

Filip Ilievski
USC/ISI

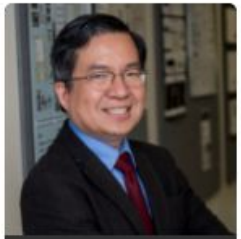
Alessandro Oltra...
Bosch Research an...

Deborah McGuin...
Rensselaer Polytech...

Pedro Szekely
USC/ISI

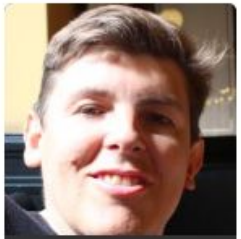
Confirmed Panelists

David Ferrucci
Elemental Cognition



Shih-Fu Chang
Colombia University

Łukasz Kaiser
Google Brain & CNRS



Tony Veale
University College D...

AAAI'21 Workshop on CSKGs