Commonsense Knowledge Graphs

Filip Ilievski

Common sense is the common knowledge about the world that is possessed by every schoolchild and the methods for making obvious inferences from this knowledge.

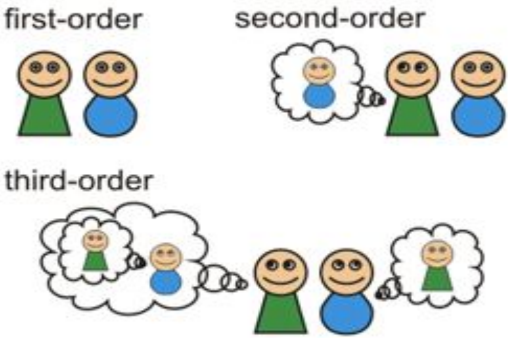
Davis, E. (2014). Representations of commonsense knowledge.

Commonsense knowledge includes the basic facts about events (including actions) and their effects, facts about knowledge and how it is obtained, facts about beliefs and desires. It also includes the basic facts about material objects and their properties.

McCarthy, J. (1989). Artificial intelligence, logic and formalizing common sense.

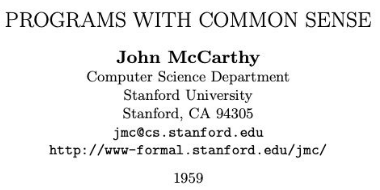
Aspects of common sense

Intuitive physics



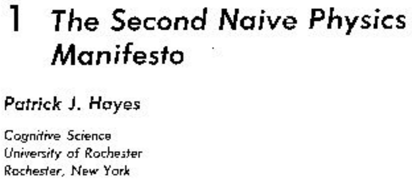
Commonsense psychology

Common facts


PROGRAMS WITH COMMON SENSE

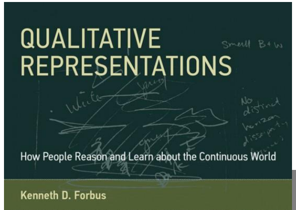
John McCarthy
Computer Science Department
Stanford University
Stanford, CA 94305
jmc@cs.stanford.edu
http://www-formal.stanford.edu/jmc/

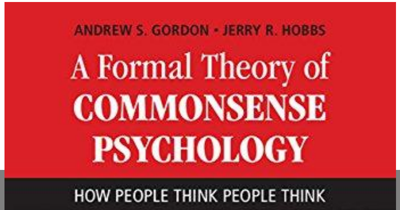
1959

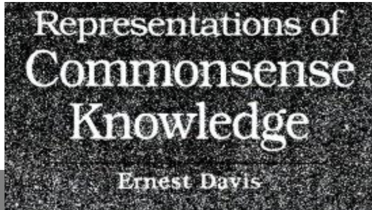

1 The Second Naive Physics Manifesto

Patrick J. Hayes

Cognitive Science
University of Rochester
Rochester, New York







Common Sense Knowledge Graphs

- COMET [Bosselut et al., 2019]
- Atomic [Sap et al., 2019]
- WebChild [Tandon et al., 2014]
- WebChild 2.0 [Tandon et al., 2017]
- Open Mind Common Sense [Minski, Singh, Havasi,1999]
- ConceptNet [Liu, Singh, 2004]
- ConceptNet 5.5 [Speer et al., 2017]
- NELL [Carlson et al., 2010]
- NELL [Mitchell et al., 2015]
- Wikidata [Vrandečić, 2012]
- Cyc [Lenat et al., 1984]
- OpenCyc 4.0 [Lenat 2012]

Considerations

Representation

- symbolic
- natural language
- neural

Creation method

- expert input
- crowdsourcing
- information extraction, machine learning

Knowledge type

- entities and actions
- inferential/rules

Topic

- general
- social

COMET

Atomic

WebChild

ConceptNet

NELL

Wikidata

OpenCyc

Representation Method

Representation

- symbolic: frisbee, dog
- natural language: "PersonX throws a frisbee"
- neural: <black box>

Creation method

- expert input
- crowdsourcing
- information extraction, machine learning

Knowledge type

- entities and actions
- inferential/rules

Topic

- general
- social

COMET

Atomic

WebChild

ConceptNet

NELL

Wikidata

OpenCyc

Creation Method

Representation

- symbolic
- natural language
- neural

Creation method

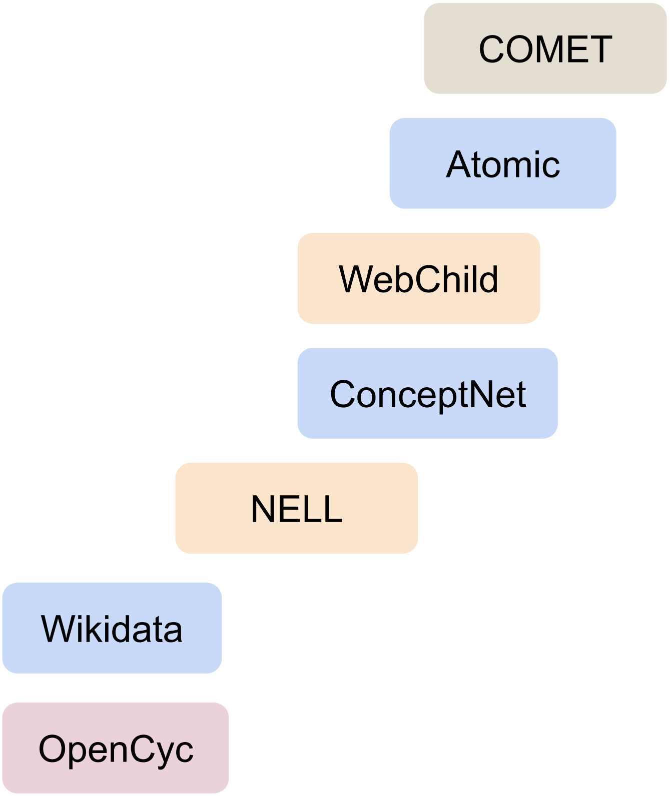
- expert input
- crowdsourcing
- information extraction, machine learning

Knowledge type

- entities and actions
- inferential/rules

Topic

- general
- social



Knowledge Type

Representation

- symbolic
- natural language
- neural

Creation method

- expert input
- crowdsourcing
- information extraction, machine learning

Knowledge type

- entities and actions: frisbee, dog, throw, catch
- inferential/rules:
  PersonX throws frisbee, as a result
  others then, catches frisbee

Topic

- general
- social

COMET

Atomic

WebChild

ConceptNet

NELL

Wikidata

OpenCyc

Topic

Representation
- symbolic
- natural language
- neural

Creation method
- expert input
- crowdsourcing
- information extraction, machine learning

Knowledge type
- entities and actions
- inferential/rules

Topic
- general
- social

COMET

Atomic

WebChild

ConceptNet

NELL

Wikidata

OpenCyc

Design Approach

Representation

- symbolic
- natural language
- neural

Creation method

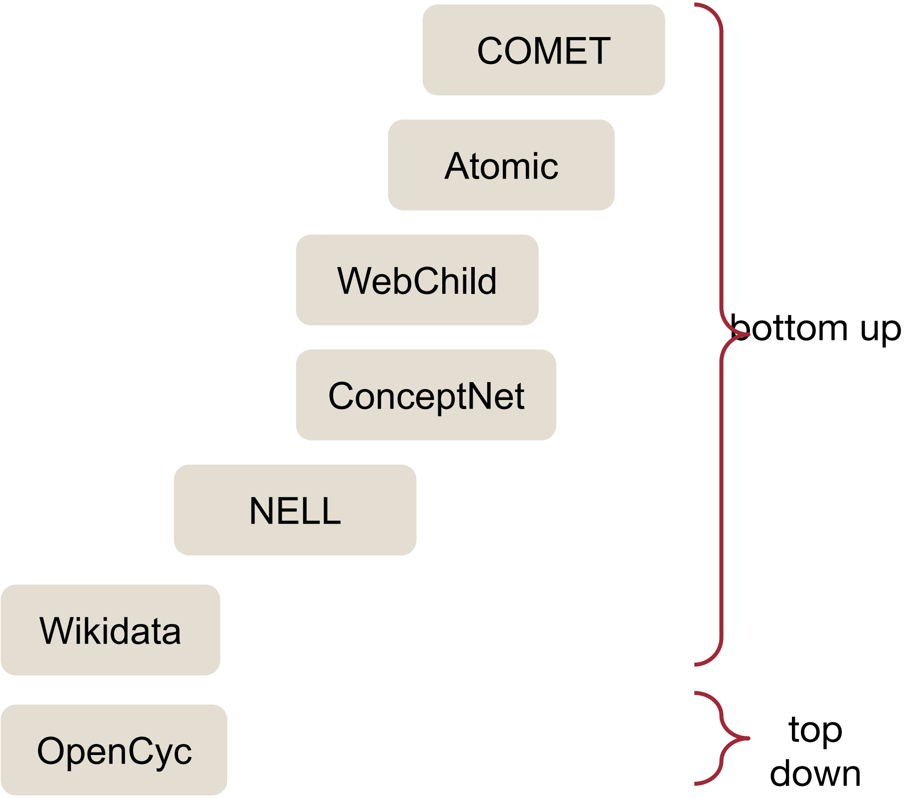
- expert input
- crowdsourcing
- information extraction, machine learning

Knowledge type

- entities and actions
- inferential/rules

Topic

- general
- social



Cyc

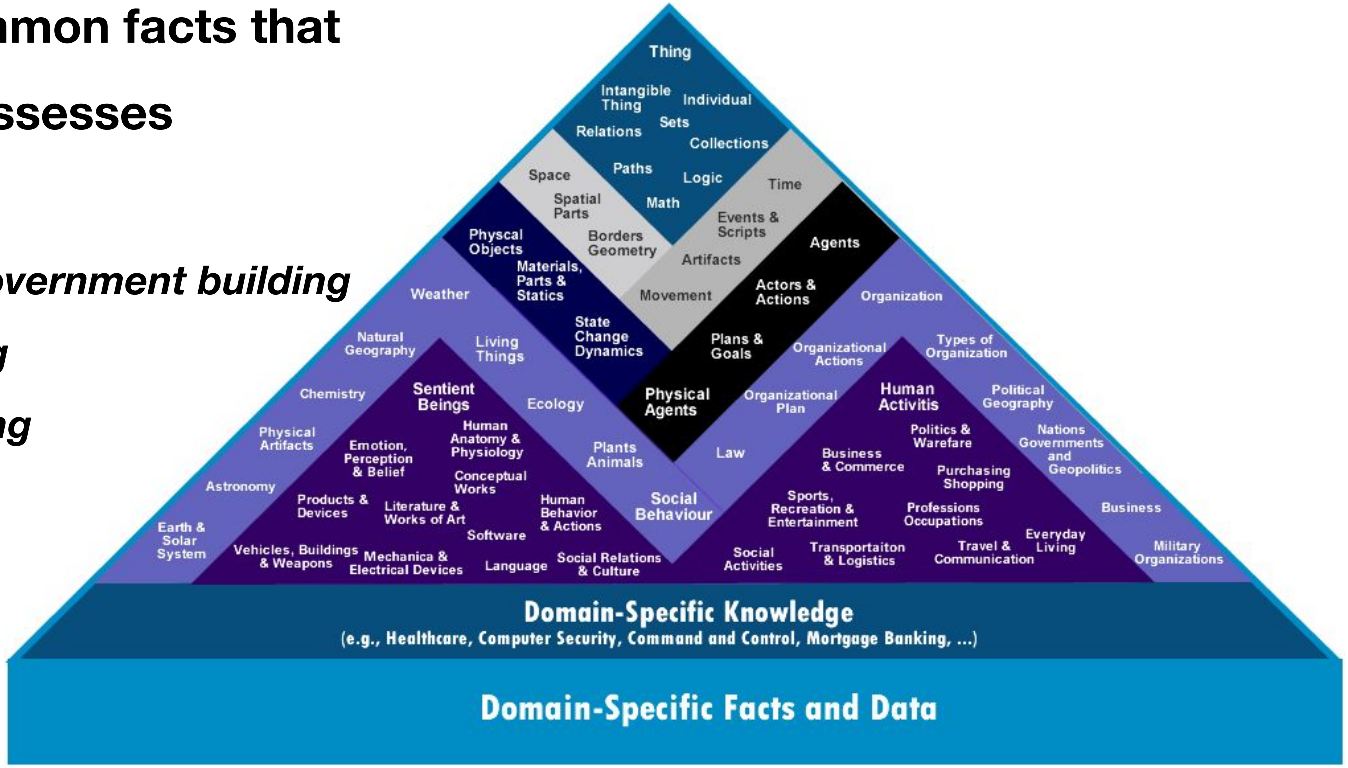
[Lenat, 1995]

Knowledge of the common facts that an average person possesses

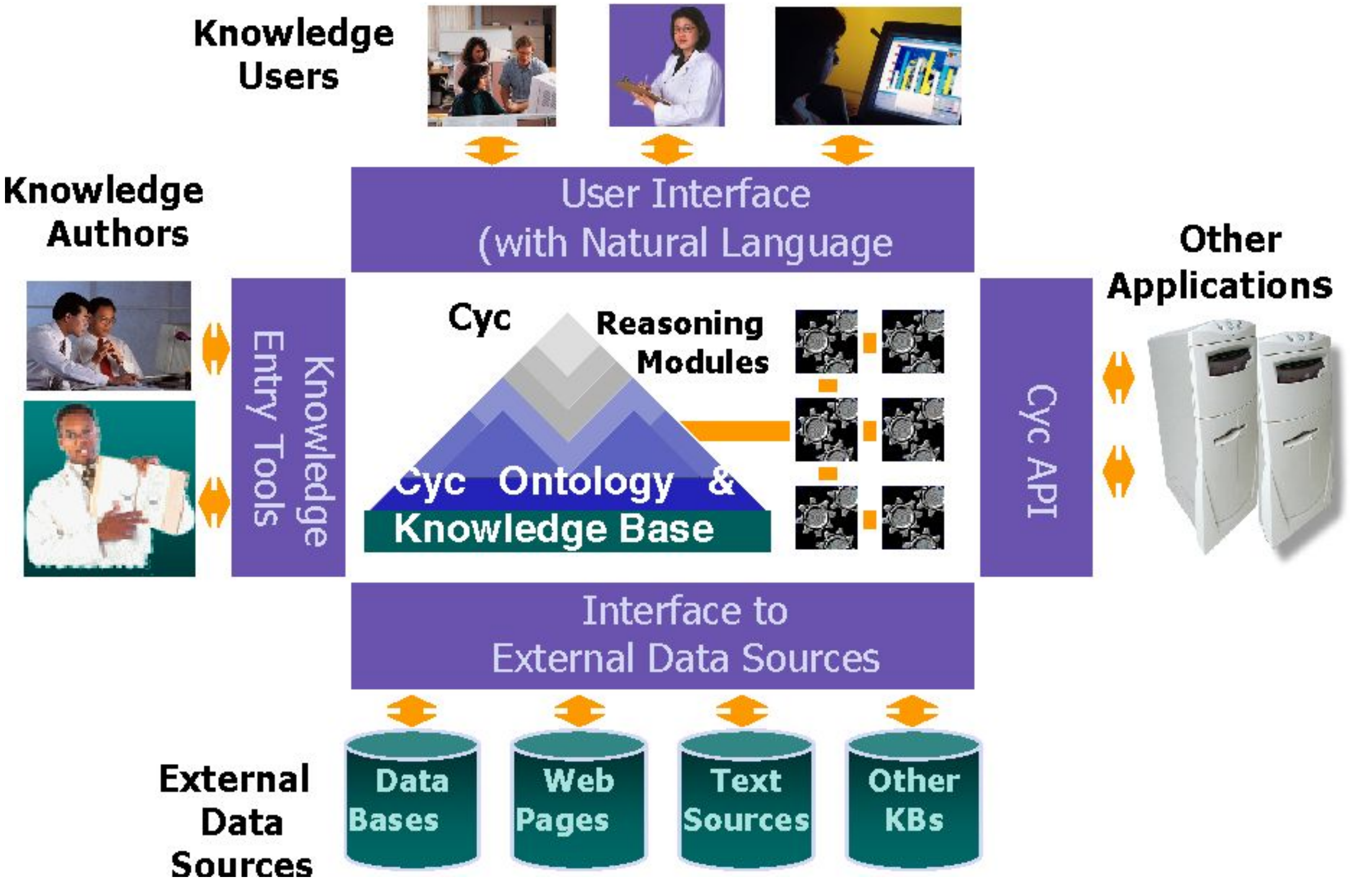
An embassy is a type of government building

Knives are used for cutting

Dogs are capable of barking



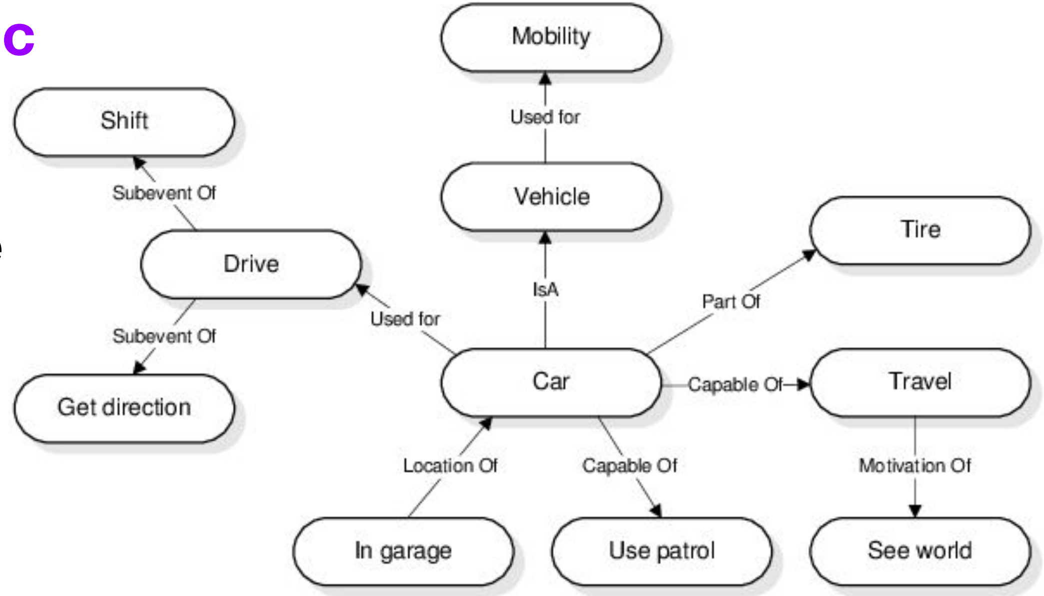
Evolution of Cyc



ConceptNet: An introduction

"a freely-available **semantic network**, designed to help computers understand the meanings of words that people use"

"an open, multi-lingual **knowledge graph**"



https://www.conceptnet.io/

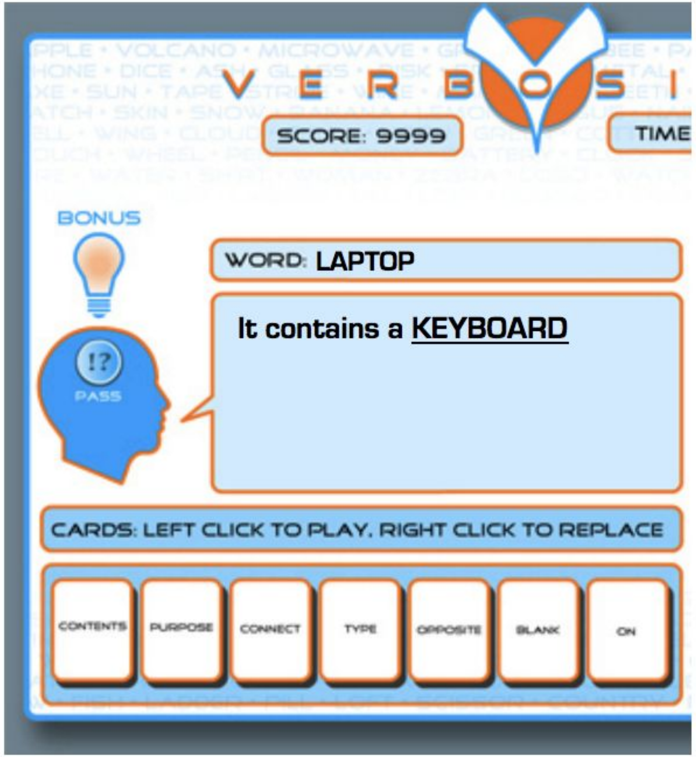
Sources of knowledge

- Similar to previous versions, relational knowledge contributed to Open Mind Common Sense and its sister projects in other languages
- Subset of DBpedia
- Wiktionary (a dominant source)
  - Dictionary-style information also used from Open Multilingual WordNet
- High-level ontology from OpenCyc

Human-generated knowledge: Games with a purpose (GWAP)

"multi-player online game that is designed to be fun and accomplish tasks that are easy for humans but beyond the capability of today's computers."

https://www.cmu.edu/homepage/computing/2008/summer/games-with-a-purpose.shtml



ATOMIC:
inferential knowledge in *natural language* form

https://mosaickg.apps.allenai.org/kg_atomic

Slides adapted from Sap et al. https://homes.cs.washington.edu/~msap/acl2020-commonsense/

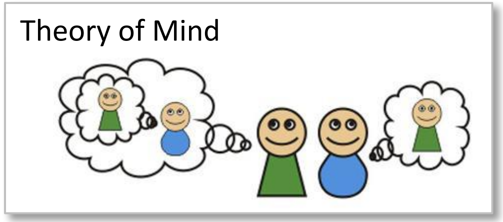
Knowledge of causes and effects

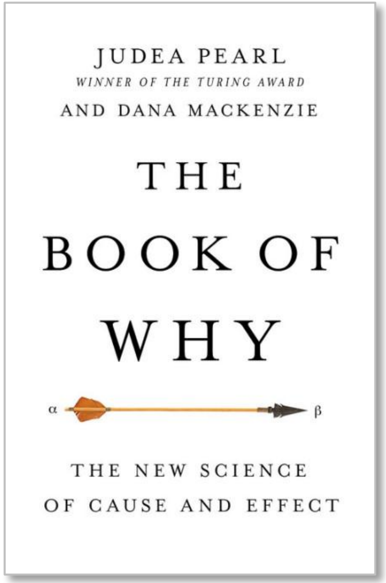
- Humans have **theory of mind**, allowing us to
  - make inferences about **people's mental states**
  - understand **likely events** that precede and follow

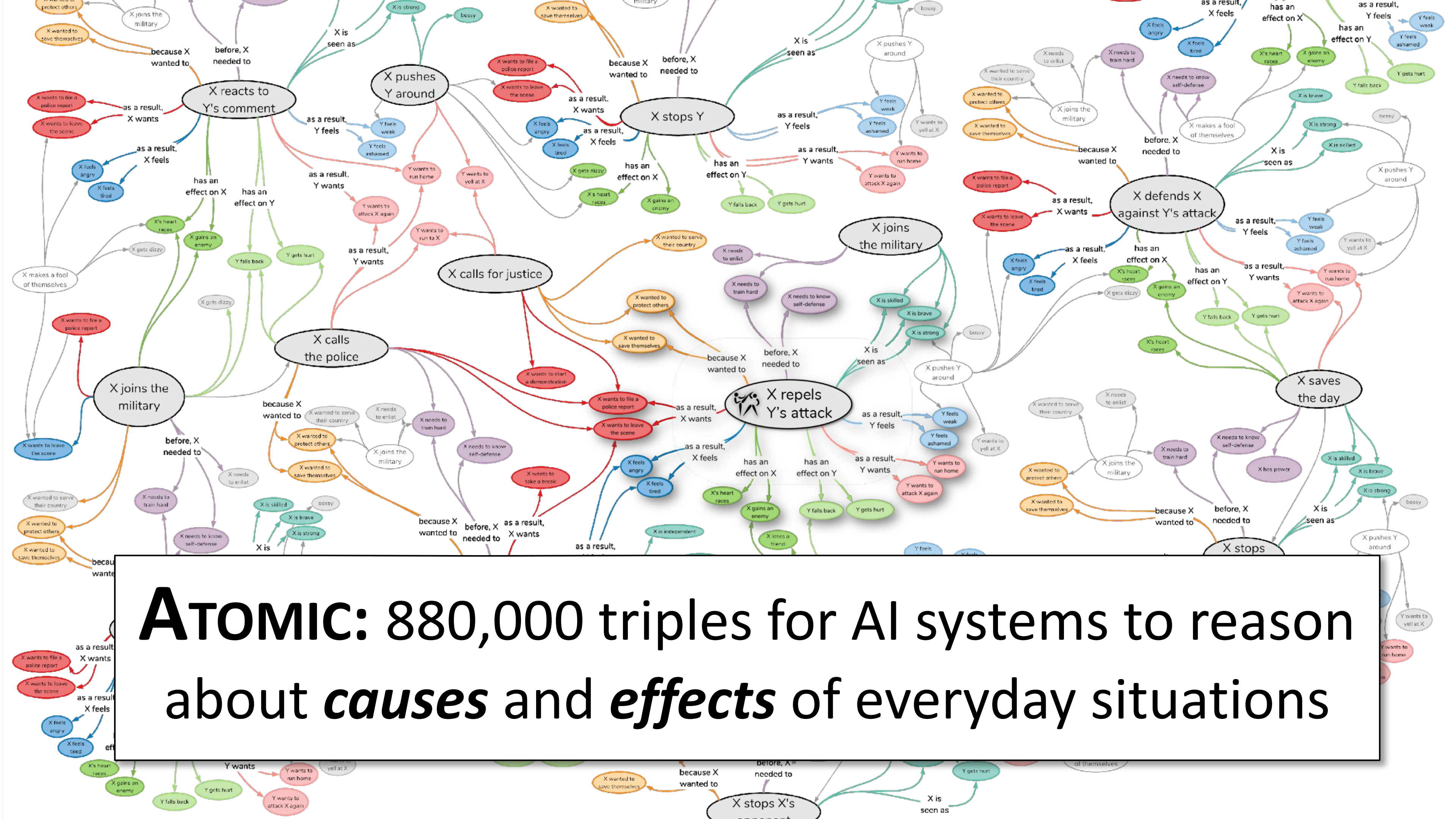
(Moore, 2013)

- AI systems struggle with ***inferential*** reasoning
  - only find **complex correlational patterns** in data
  - **limited to the domain** they are trained on

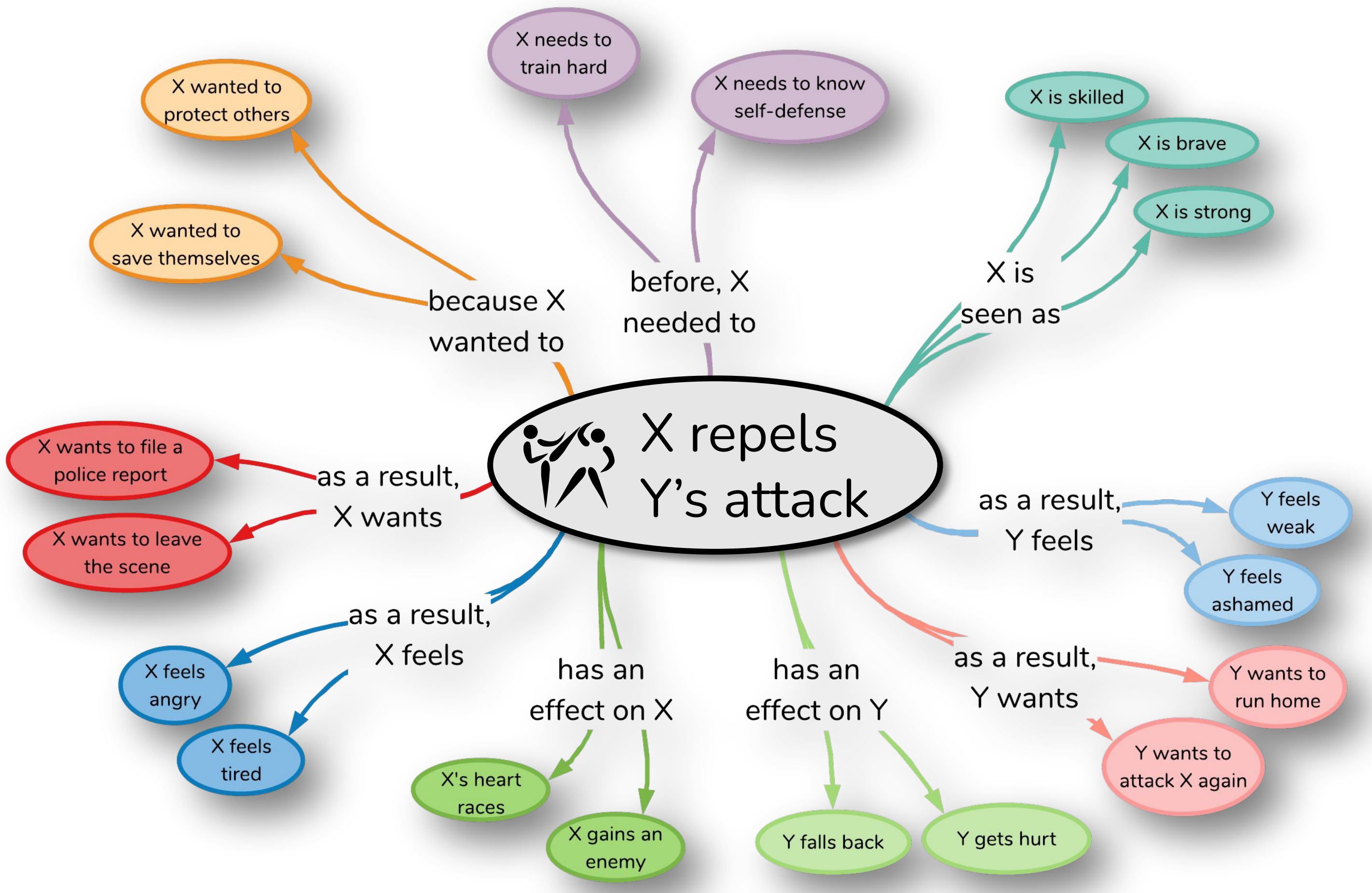
(Pearl; Davis and Marcus 2015; Lake et al. 2017; Marcus 2018)



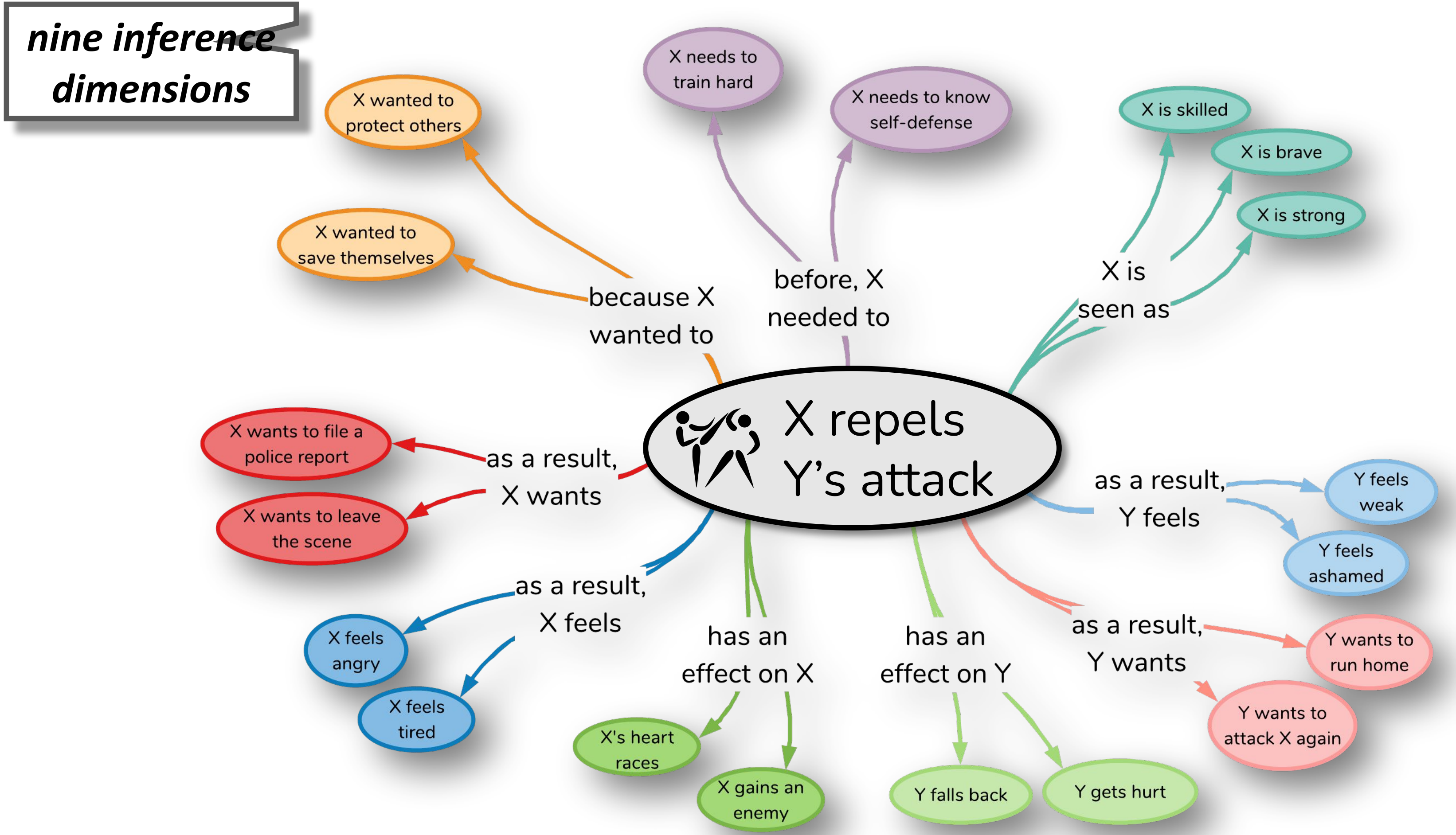


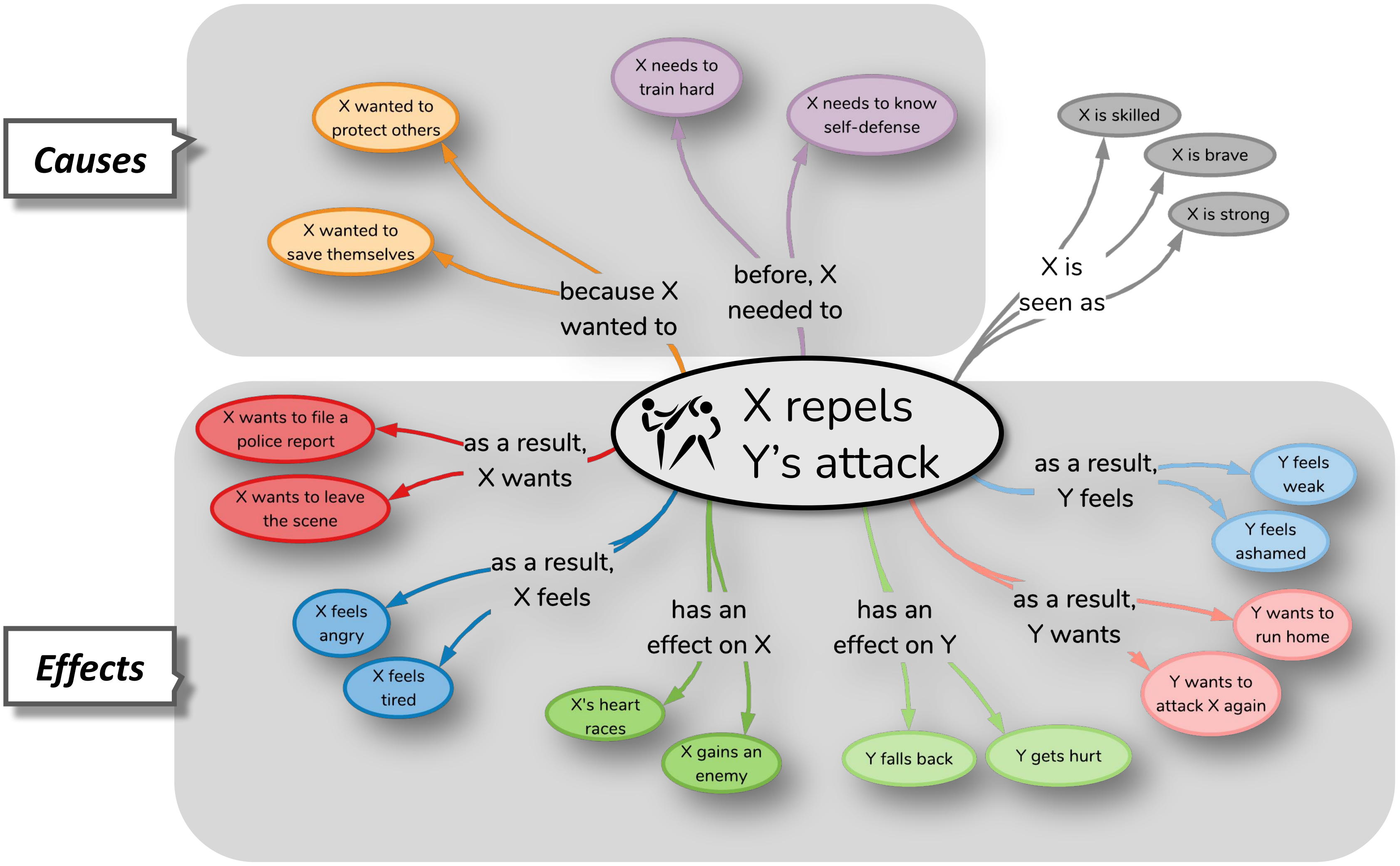


ATOMIC: 880,000 triples for AI systems to reason about ***causes*** and ***effects*** of everyday situations

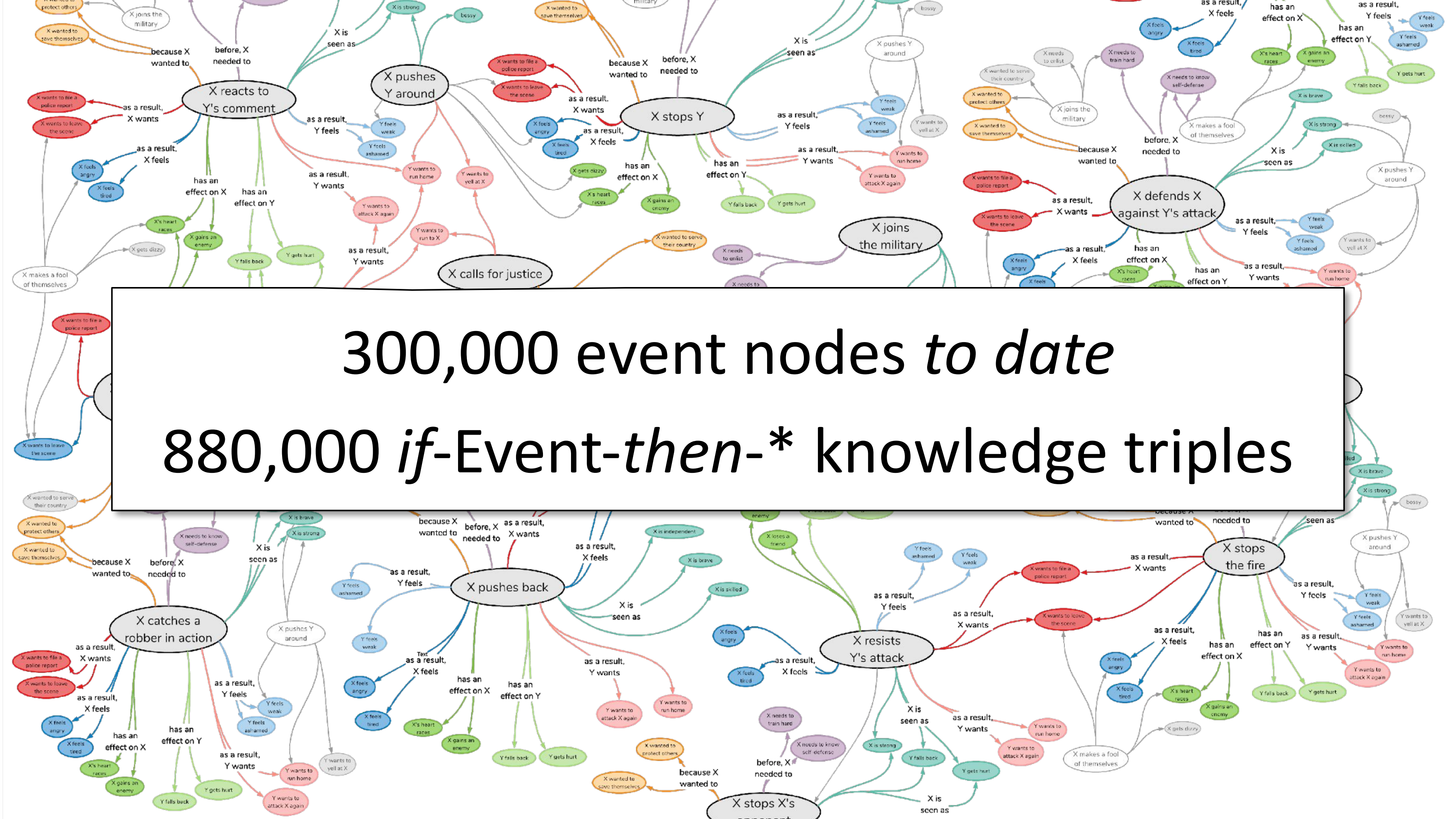


X repels Y's attack
because X wanted to
X wanted to protect others
X wanted to save themselves
before, X needed to
X needs to train hard
X needs to know self-defense
X is seen as
X is skilled
X is brave
X is strong
as a result, X wants
X wants to file a police report
X wants to leave the scene
as a result, Y feels
Y feels weak
Y feels ashamed
as a result, X feels
X feels angry
X feels tired
has an effect on X
X's heart races
X gains an enemy
has an effect on Y
Y falls back
Y gets hurt
as a result, Y wants
Y wants to run home
Y wants to attack X again


nine inference dimensions
X repels Y's attack
because X wanted to
X wanted to protect others
X wanted to save themselves
before, X needed to
X needs to train hard
X needs to know self-defense
X is seen as
X is skilled
X is brave
X is strong
as a result, X wants
X wants to file a police report
X wants to leave the scene
as a result, Y feels
Y feels weak
Y feels ashamed
as a result, X feels
X feels angry
X feels tired
has an effect on X
X's heart races
X gains an enemy
has an effect on Y
Y falls back
Y gets hurt
as a result, Y wants
Y wants to run home
Y wants to attack X again



Causes
X wanted to protect others
X wanted to save themselves
because X wanted to
X needs to train hard
X needs to know self-defense
before, X needed to
X is skilled
X is brave
X is strong
X is seen as
X repels Y's attack
as a result, X wants
X wants to file a police report
X wants to leave the scene
as a result, X feels
X feels angry
X feels tired
has an effect on X
X's heart races
X gains an enemy
has an effect on Y
Y falls back
Y gets hurt
as a result, Y wants
Y wants to run home
Y wants to attack X again
as a result, Y feels
Y feels weak
Y feels ashamed
Effects



300,000 event nodes *to date*

880,000 *if*-Event-*then*-* knowledge triples

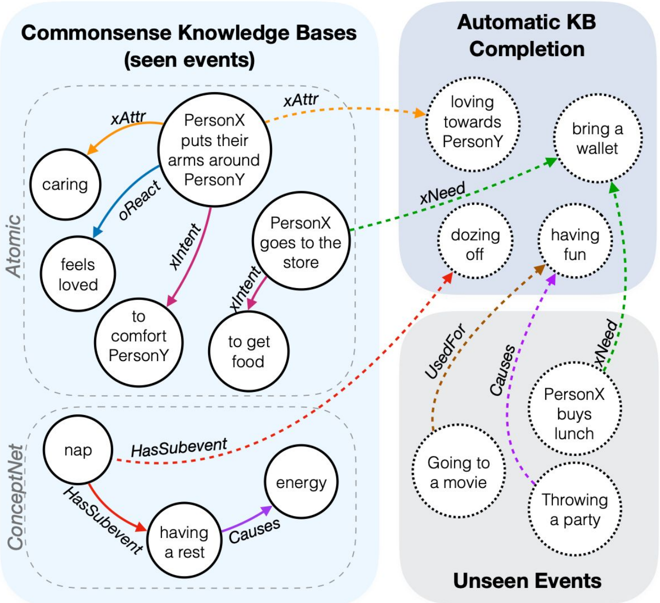
Event	Type of relations	Inference examples	Inference dim.
"PersonX pays PersonY a compliment"	If-Event-Then-Mental-State	PersonX wanted to be nice	xIntent
		PersonX will feel good	xReact
		PersonY will feel flattered	oReact
	If-Event-Then-Event	PersonX will want to chat with PersonY	xWant
		PersonY will smile	oEffect
		PersonY will compliment PersonX back	oWant
	If-Event-Then-Persona	PersonX is flattering	xAttr
		PersonX is caring	xAttr
"PersonX makes PersonY's coffee"	If-Event-Then-Mental-State	PersonX wanted to be helpful	xIntent
		PersonY will be appreciative	oReact
		PersonY will be grateful	oReact
	If-Event-Then-Event	PersonX needs to put the coffee in the filter	xNeed
		PersonX gets thanked	xEffect
		PersonX adds cream and sugar	xWant
	If-Event-Then-Persona	PersonX is helpful	xAttr
		PersonX is deferential	xAttr
"PersonX calls the police"	If-Event-Then-Mental-State	PersonX wants to report a crime	xIntent
		Others feel worried	oReact
	If-Event-Then-Event	PersonX needs to dial 911	xNeed
		PersonX wants to explain everything to the police	xWant
		PersonX starts to panic	xEffect
		Others want to dispatch some officers	oWant
	If-Event-Then-Persona	PersonX is lawful	xAttr
		PersonX is responsible	xAttr

COMET: Commonsense Transformers for Automatic Knowledge Graph Construction

Antoine Bosselut [◇♠] **Hannah Rashkin** [◇♠] **Maarten Sap** [◇♠] **Chaitanya Malaviya** [◇]
Asli Celikyilmaz [♣] **Yejin Choi** [◇♠]
[◇]Allen Institute for Artificial Intelligence, Seattle, WA, USA
[♠]Paul G. Allen School of Computer Science & Engineering, Seattle, WA, USA
[♣]Microsoft Research, Redmond, WA, USA

Abstract

We present the first comprehensive study on automatic knowledge base construction for two prevalent commonsense knowledge graphs: ATOMIC (Sap et al., 2019) and ConceptNet (Speer et al., 2017). Contrary to many conventional KBs that store knowledge with canonical templates, commonsense KBs only store loosely structured open-text descriptions of knowledge. We posit that an important step toward automatic commonsense completion is the development of *generative* models of commonsense knowledge, and propose ***COMmonsEnse Transformers*** (COMET) that learn to generate rich and



Building Common Sense KGs Is Hard

- Commonsense knowledge is **immeasurably vast**, making it **impossible to manually enumerate**
- Commonsense knowledge is often implicit, and often **can't be directly extracted from text**

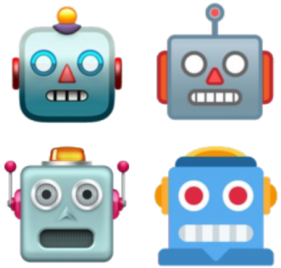
Slide by Antoine Bosselut

Traditional KB Completion

Gather training set of knowledge tuples

(person, CapableOf, buy)

Learn relationships among entities

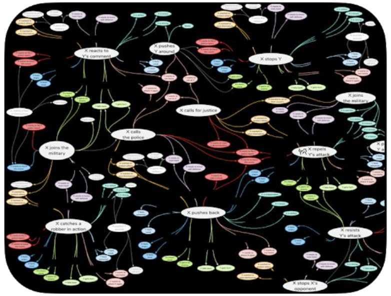


(Socher et al., 2013)
(Bordes et al., 2013)
(Riedel et al., 2013)
(Toutanova et al., 2015)
(Yang et al., 2015)
(Trouillon et al., 2016)
(Nguyen et al., 2016)
(Dettmers et al., 2018)

Predict new relationships

(person, CapableOf, **?**)

Store in knowledge graph



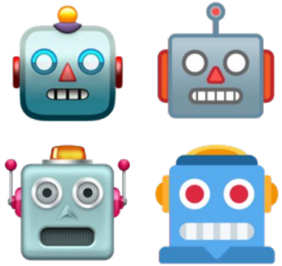
Slide by Antoine Bosselut

COMET Idea

Gather training set of knowledge tuples

(person, CapableOf, buy)

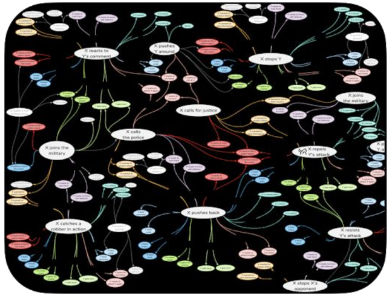
Learn relationships among entities



Predict new relationships

(person, CapableOf, ?)

Store in knowledge graph

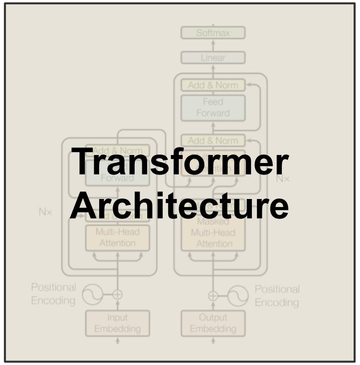


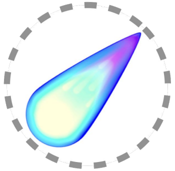
ATOMIC Input Template and ConceptNet Relation-only Input Template

s tokens	mask tokens	*r* token	*o* tokens
PersonX goes to the mall	[MASK]	<xIntent>	to buy clothes

ConceptNet Relation to Language Input Template

s tokens	mask tokens	*r* tokens	mask tokens	*o* tokens
go to mall	[MASK] [MASK]	has prerequisite	[MASK]	have money





COMET

Symbolic Knowledge Graph

Knowledge stored as triples

Knowledge is not contextualized

Knowledge is incomplete

Symbolic Knowledge Graph

Knowledge stored as triples

Knowledge is not contextualized

Knowledge is incomplete

COMET Knowledge Base Transformer

Knowledge generated dynamically

Input format is natural language

Kai knew that things were getting out of control and managed to keep his temper in check

- Kai wants to avoid trouble
- Kai intends to be calm
- Kai stays calm
- Kai is viewed as cautious

Randomly selected novel generations from ATOMIC **with COMET**

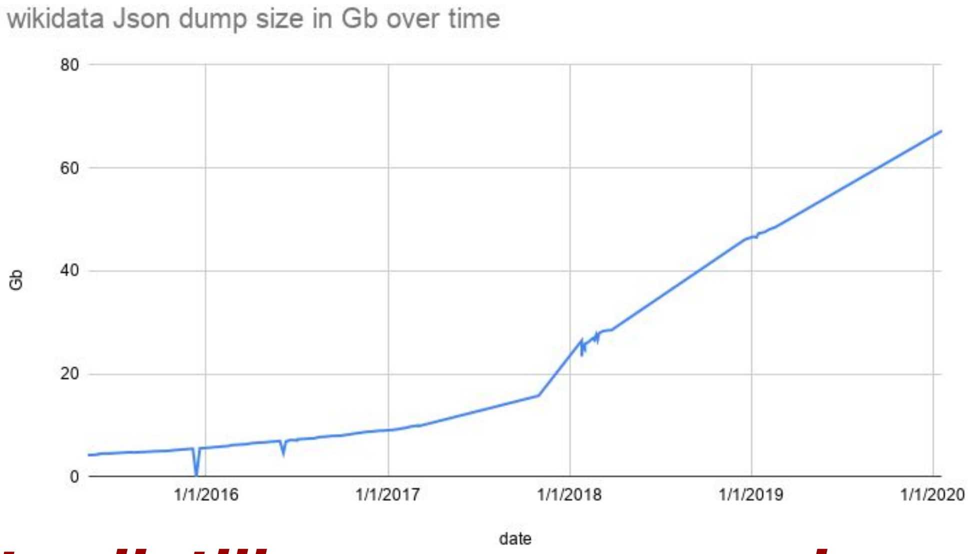
Seed Concept	Relation	Generated	Plausible
X holds out X's hand to Y	xAttr	helpful	✓
X meets Y eyes	xAttr	intense	✓
X watches Y every ___	xAttr	observant	✓
X eats red meat	xEffect	gets fat	✓
X makes crafts	xEffect	gets dirty	✓
X turns X's phone	xEffect	gets a text	
X pours ___ over Y's head	oEffect	gets hurt	✓
X takes Y's head off	oEffect	bleeds	✓
X pisses on Y's bonfire	oEffect	gets burned	
X spoils somebody rotten	xIntent	to be mean	
X gives Y some pills	xIntent	to help	✓
X provides for Y's needs	xIntent	to be helpful	✓
X explains Y's reasons	xNeed	to know Y	✓
X fulfils X's needs	xNeed	to have a plan	✓
X gives Y everything	xNeed	to buy something	✓
X eats pancakes	xReact	satisfied	✓
X makes ___ at work	xReact	proud	✓
X moves house	xReact	happy	✓
X gives birth to the Y	oReact	happy	✓
X gives Y's friend ___	oReact	grateful	✓
X goes ___ with friends	oReact	happy	✓
X gets all the supplies	xWant	to make a list	✓
X murders Y's wife	xWant	to hide the body	✓
X starts shopping	xWant	to go home	✓
X develops Y theory	oWant	to thank X	✓
X offer Y a position	oWant	to accept the job	✓
X takes ___ out for dinner	oWant	to eat	✓

Randomly selected novel generations from ConceptNet **with COMET**

Seed	Relation	Completion	Plausible
piece	`PartOf`	machine	✓
bread	`IsA`	food	✓
planet	`AtLocation`	space	✓
dust	`AtLocation`	fridge	
puzzle	`AtLocation`	your mind	🤔
college	`AtLocation`	town	✓
dental chair	`AtLocation`	dentist	✓
finger	`AtLocation`	your finger	
sing	`Causes`	you feel good	✓
doctor	`CapableOf`	save life	✓
post office	`CapableOf`	receive letter	✓
dove	`SymbolOf`	purity	✓
sun	`HasProperty`	big	✓
bird bone	`HasProperty`	fragile	✓
earth	`HasA`	many plant	✓
yard	`UsedFor`	play game	✓
get pay	`HasPrerequisite`	work	✓
print on printer	`HasPrerequisite`	get printer	✓
play game	`HasPrerequisite`	have game	✓
live	`HasLastSubevent`	die	✓
swim	`HasSubevent`	get wet	✓
sit down	`MotivatedByGoal`	you be tire	✓
all paper	`ReceivesAction`	recycle	✓
chair	`MadeOf`	wood	✓
earth	`DefinedAs`	planet	✓

Wikidata

90M nodes **>1.1B edges** **8k properties**



How to distill commonsense knowledge?

Ilievski et al. (2020). [Commonsense Knowledge in Wikidata](). Wikidata Workshop at ISWC 2020

Principles of Commonsense Knowledge

P1: Concepts, not entities

houses have rooms

Versailles Palace has 700 rooms

WD guidelines on entity capitalization

Principles of Commonsense Knowledge

P1: Concepts, not entities

houses have rooms

Versailles Palace has 700 rooms

WD guidelines on entity capitalization

P2: Common concepts

Container used for storage

Noma subclass of aphthous stomatitis

Corpus frequency

Principles of Commonsense Knowledge

P1: Concepts, not entities

houses have rooms

Versailles Palace has 700 rooms

WD guidelines on entity capitalization

P2: Common concepts

Container used for storage

Noma subclass of aphthous stomatitis

Corpus frequency

P3: General-domain relations

wheel is part of a car

cholesterol has component cell membrane

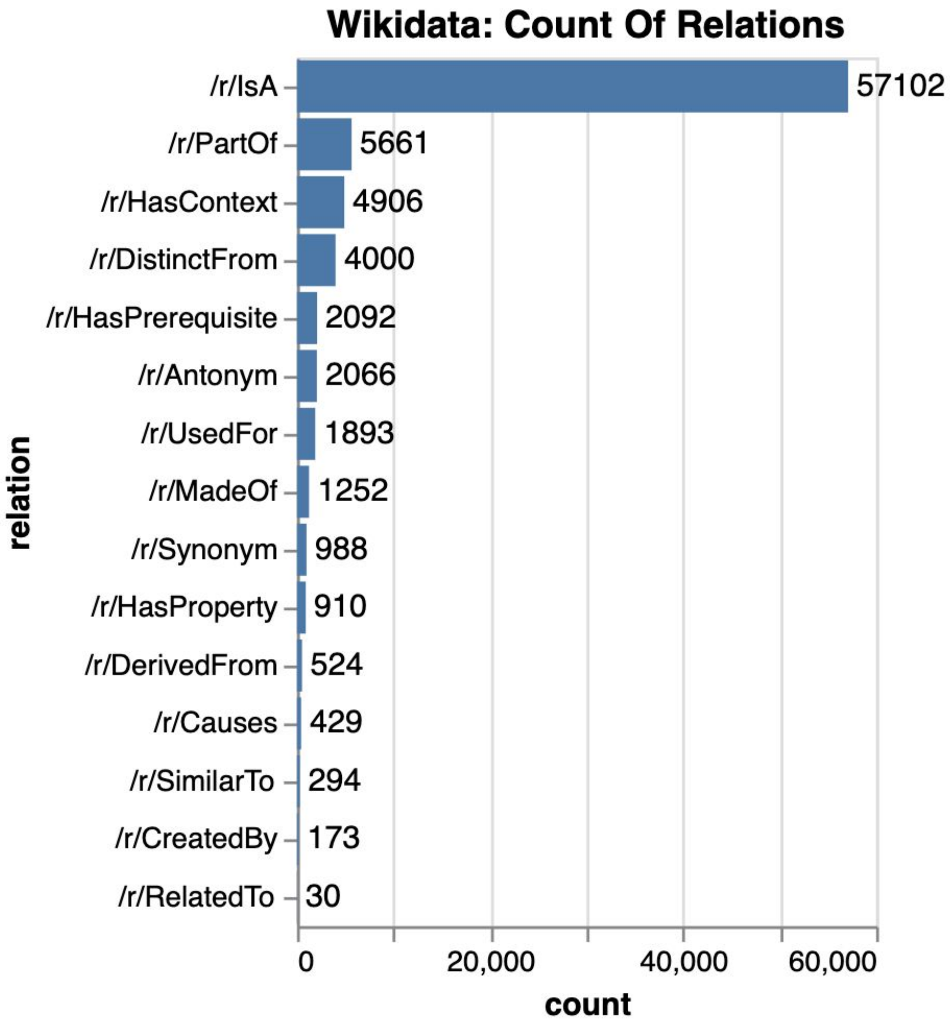
Mapping to ConceptNet

Mapping general-domain relations to ConceptNet

Category	ConceptNet	Wikidata
distinctness	/r/DistinctFrom	different from (P1889)
antonymy	/r/Antonym	opposite of (P461)
synonymy	/r/Synonym	said to be the same as (P460)
similarity	/r/SimilarTo	partially coincident with (P1382)
derivation	/r/DerivedFrom	named after (P138), fictional analog of (P1074)
inheritance	/r/IsA	instance of (P31), subclass of (P279), subproperty of (P1647)
meronymy	/r/PartOf	part of (P361), *has part (P527), *has parts of the class (P2670)
material	/r/MadeOf	material used (P186), is a list of (P360), *has list (P2354)
attribution	/r/CreatedBy	*product or material produced (P1056)
utility	/r/UsedFor	use (P366), *uses (P2283), used by (P1535)
properties	/r/HasProperty	color (P462), has quality (P1552), properties of this type (P1963), Wikidata property (P1687), sex or gender (P21)
causation	/r/Causes	*has cause (P828), has effect (P1542), symptoms (P780)
ordering	/r/HasPrerequisite	*followed by (P156), follows (P155)
context	/r/HasContext	facet of (P1269), sport (P641), field of this occupation (P425), health specialty (P1995), competition class (P2094), genre (P136), studied by (P2579), field of work (P101), afflicts (P689), *practiced by (P3095), depicts (P180), main subject (P921)
other	/r/RelatedTo	see also (P1659), subject item of this property (P1629)

*Wikidata-CS = 0.01% * Wikidata*

	Wikidata-CS	**Wikidata**	**Ratio**
# nodes	71,243	84 million	***0.08%***
# edges	***101,771***	1.5 billion	***0.01%***


Wikidata: Count Of Relations
/r/IsA
/r/PartOf
/r/HasContext
/r/DistinctFrom
/r/HasPrerequisite
/r/Antonym
/r/UsedFor
/r/MadeOf
/r/Synonym
/r/HasProperty
/r/DerivedFrom
/r/Causes
/r/SimilarTo
/r/CreatedBy
/r/RelatedTo
57102
5661
4906
4000
2092
2066
1893
1252
988
910
524
429
294
173
30
relation
0
20,000
40,000
60,000
count

Commonsense Knowledge in Wikidata

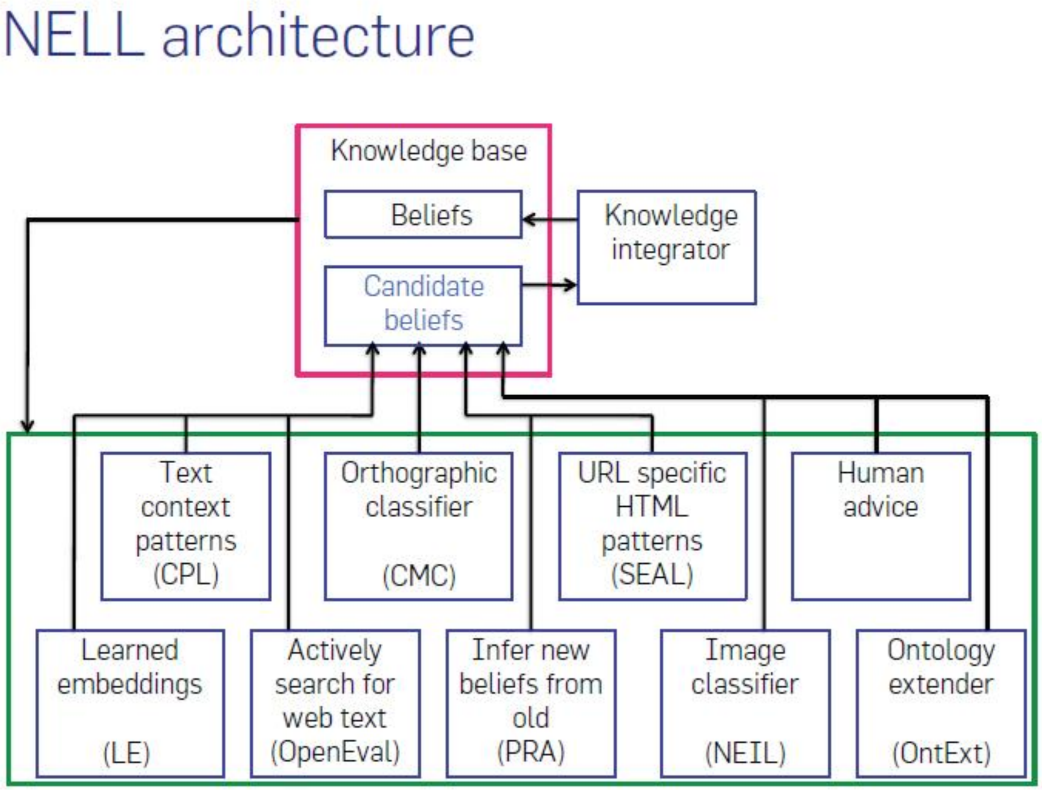
shower part of bathroom

reading uses written work

queen follows jack

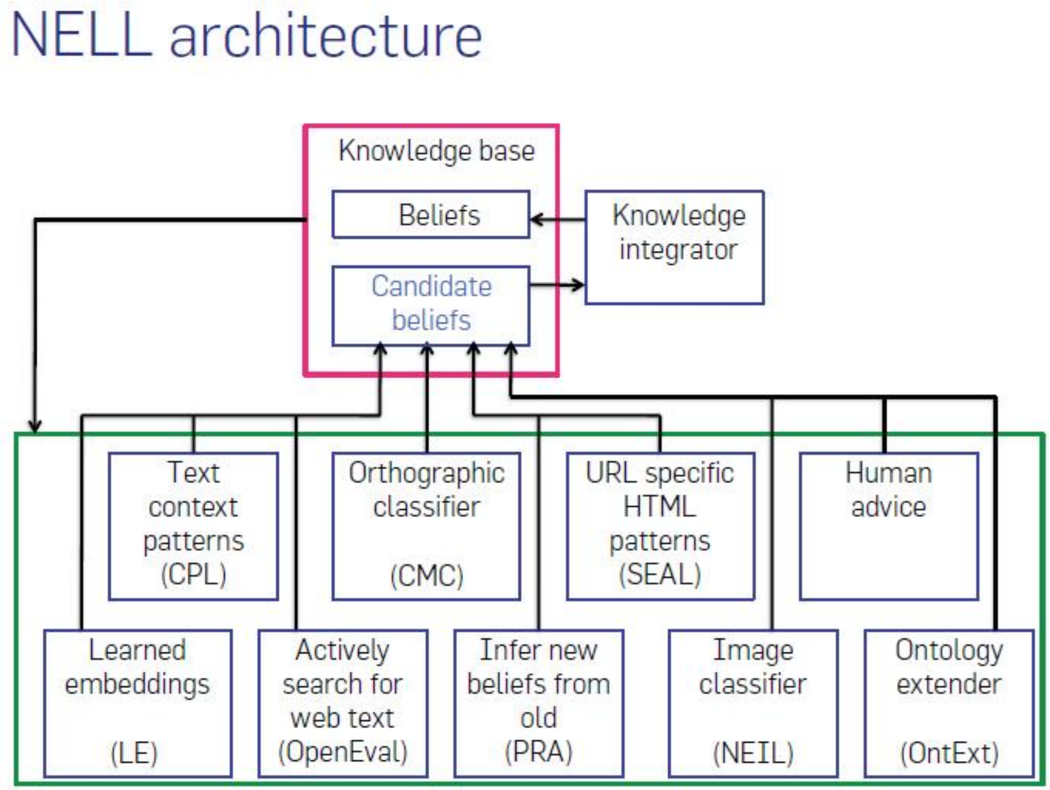
political opposition opposite of government

Never-Ending Language Learning (NELL)



https://cacm.acm.org/magazines/2018/5/227193-never-ending-learning/fulltext

NELL statistics



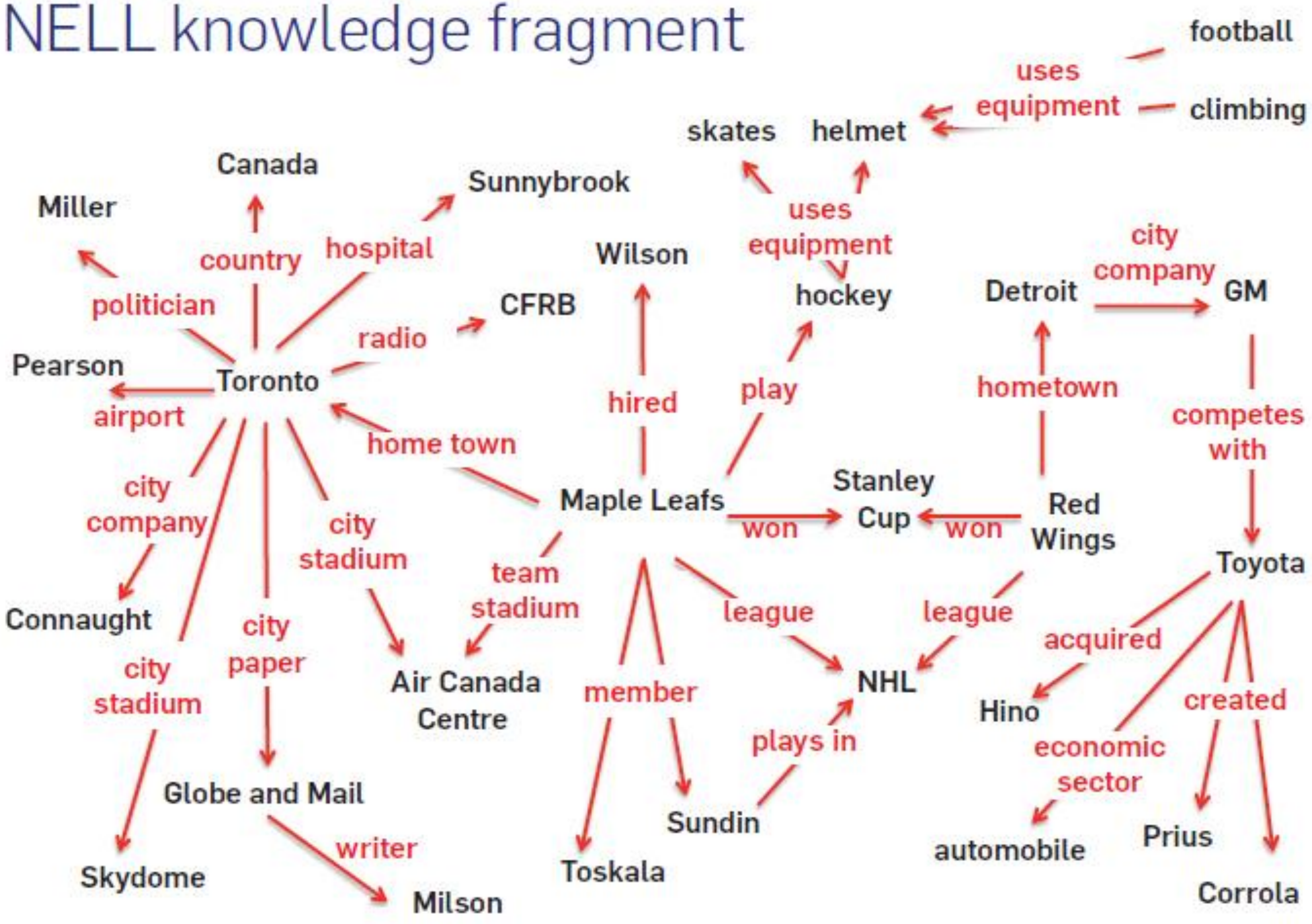
100M candidate beliefs

3M high-confidence facts

~3K predicates

https://cacm.acm.org/magazines/2018/5/227193-never-ending-learning/fulltext

NELL knowledge fragment



https://cacm.acm.org/magazines/2018/5/227193-never-ending-learning/fulltext

Latest learned facts

Recently-Learned Facts twitter

Refresh

instance	iteration	date learned	confidence
translucent_paper is an office supply	1111	06-jul-2018	93.4
the_barbirolli_string_quartet is a musical artist	1111	06-jul-2018	99.2
private_support is an event outcome	1111	06-jul-2018	99.8
vancouver_olympic_games is an instance of the olympics	1111	06-jul-2018	95.2
eddie_mathews is a person	1111	06-jul-2018	98.9
roswell_road is a street in the city atlanta	1116	12-sep-2018	93.8
james_madison is a U.S. politician who holds the office of secretary	1115	03-sep-2018	98.4
rice was born in the city orleans	1116	12-sep-2018	100.0
republic is a country also known as china	1111	06-jul-2018	100.0
dodge is a specific automobile maker dealer in utah	1115	03-sep-2018	93.8

WebChild

Automatic acquisition and organization of common sense

>18M assertions

>2M disambiguated concepts and activities

Tandon et al. (2017). [Webchild 2.0: Fine-grained commonsense knowledge distillation](). ACL 2017

WebChild relations

1. object properties

hasTaste, hasShape, evokesEmotion

2. comparative

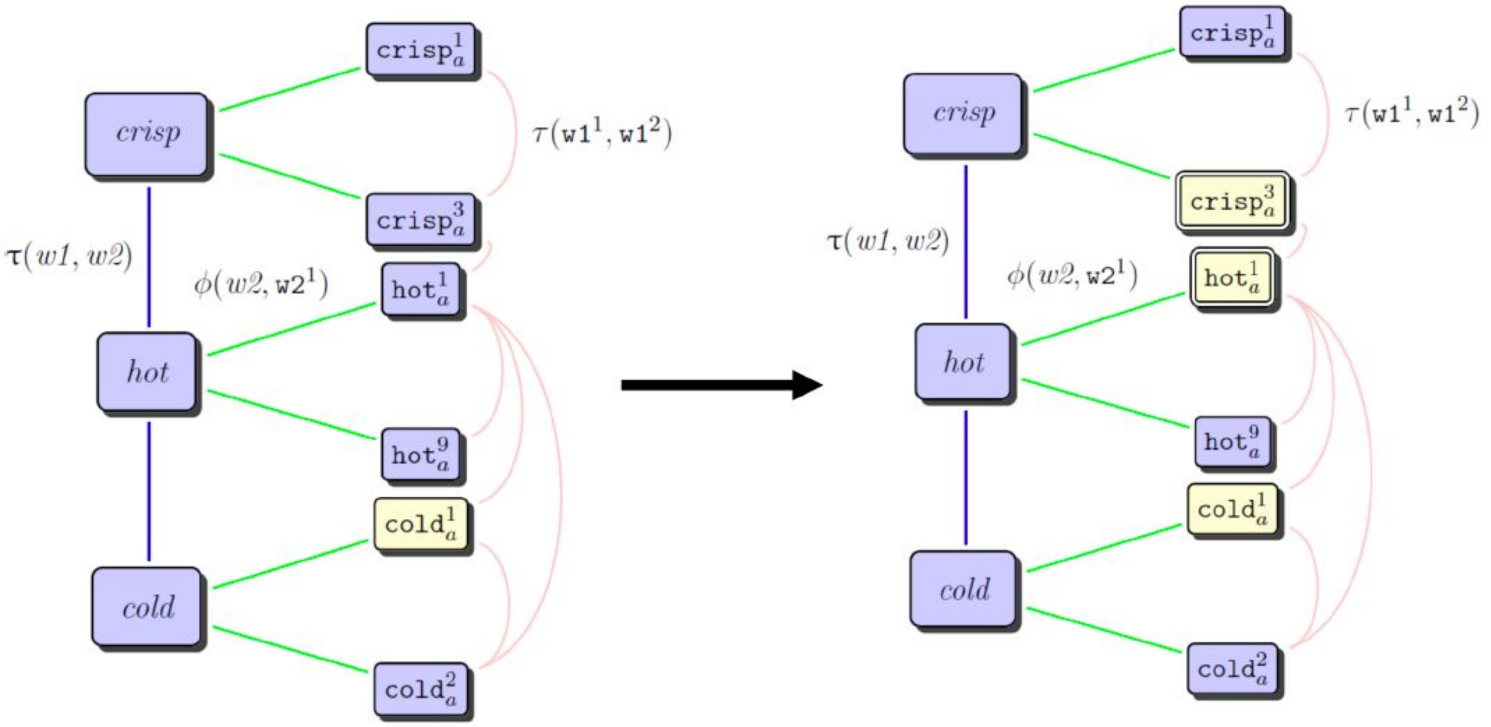
fasterThan, smallerThan

3. part-of

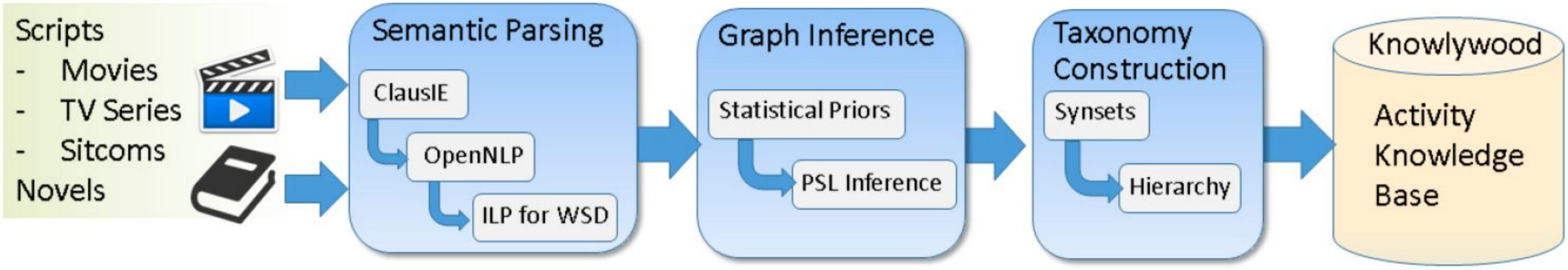
member of, physical part of, substance of

4. activities

WebChild label propagation



WebChild activity extraction



Example

mountain: *a land mass that projects well above its surroundings; higher than a hill*

TYPE OF	natural elevation
	size to **object**, under the category of **mountaineering**
PHYSICAL PROPERTIES	large, high, heavy, cold, hard, More
ABSTRACT PROPERTIES	elegant, old, safe, holy, risky, More
COMPARABLES	mountain,hill, mountain,mount, mountain, high hill, valley,mountain, More
HAS PHYSICAL PARTS	mountain peak, mountainside, slope, tableland, hill, More
HAS SUBSTANCE	mixture, metallic element, material, page, wood, More
IN SPATIAL PROXIMITY	coast, tunnel, lake, sea, river, More
ACTIVITIES	climb mountain, cross mountain, move mountain, see mountain, ascend mountain

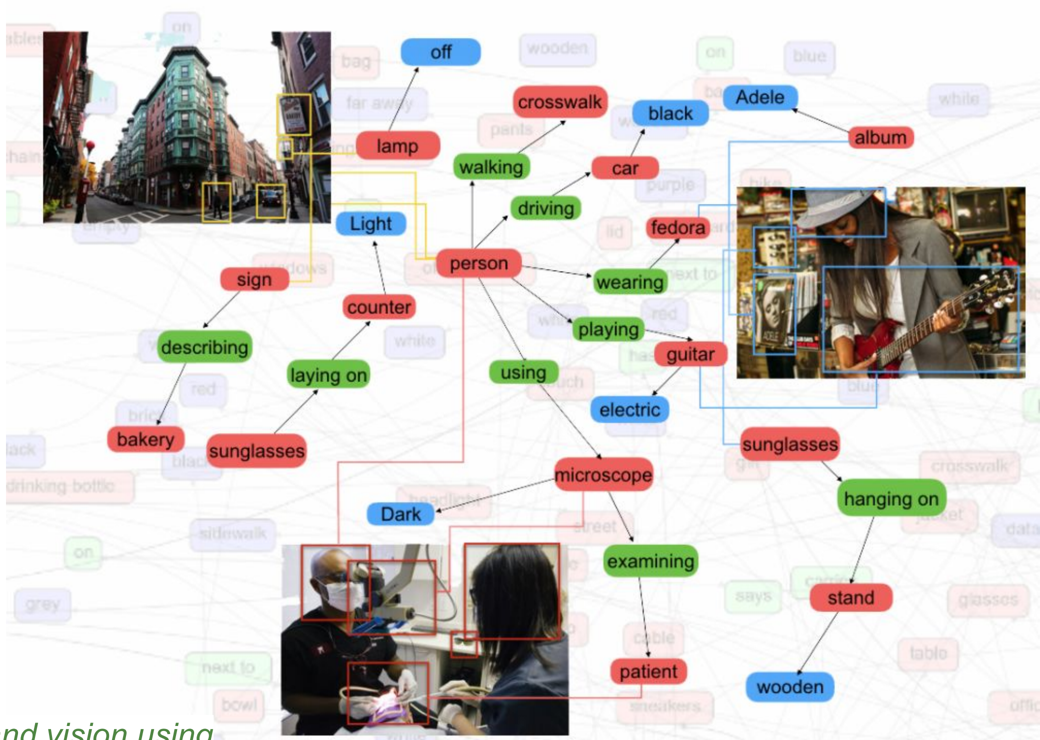
[Demo]

Visual Genome

108k images

annotated with scene graphs

canonicalized to WordNet senses



Krishna et al. (2017). [Visual genome: Connecting language and vision using crowdsourced dense image annotations.](#) *International journal of computer vision*

Demos

Dog and Frisbee

Wikidata:

https://sqid.toolforge.org/#/view?id=Q144 (dog)
https://sqid.toolforge.org/#/view?id=Q131689 (frisbee)

ConceptNet:

https://www.conceptnet.io/c/en/dog
https://www.conceptnet.io/c/en/dogs
http://conceptnet.io/c/en/frisbee
https://www.conceptnet.io/c/en/dogs_catching_frisbees

VisualGenome

https://visualgenome.org/VGViz/explore?query=throwing%20frisbee%20dog

ATOMIC:

https://mosaickg.apps.allenai.org/kg_atomic/?l=PersonX%20throws%20a%20frisbee

COMET:

comet dog
https://mosaickg.apps.allenai.org/comet_atomic/?l=PersonX%20throws%20frisbee

DICE

https://dice.mpi-inf.mpg.de/subject/dog

Catch and Throw

Wikidata:

- https://sqid.toolforge.org/#/view?id=Q17144564 (throw)
- https://sqid.toolforge.org/#/view?id=Q91553195 (catch)

ConceptNet:

- https://www.conceptnet.io/c/en/throw
- https://www.conceptnet.io/c/en/catch

VisualGenome

- https://visualgenome.org/VGViz/explore?query=catch%20frisbee

Solving tasks with CSKGs

On stage, a woman takes a seat at the piano. She

1. sits on a bench as her sister plays with the doll.
2. smiles with someone as the music plays.
3. is in the crowd, watching the dancers.
4. <u>nervously sets her fingers on the keys.</u>

(Zellers et al., 2018)

piano is used for...

- en performing music
- en music
- en accompanying an orchestra

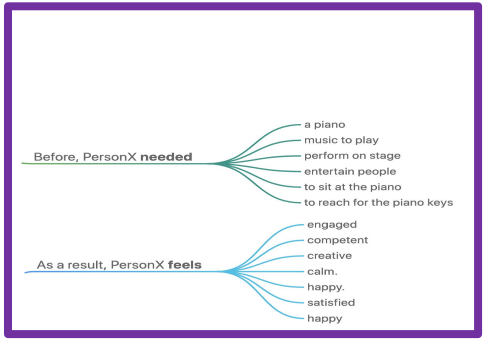
Things located at piano

- en keys
- en black keys
- en hammers
- en a keyboard

ConceptNet: pianos have keys, are used to perform music

- S: (n) **piano**, pianoforte, forte-piano (a keyboard instrument that is played by depressing keys that cause hammers to strike tuned strings and produce sounds)

WordNet: pianos are played by pressing keys



ATOMIC: to play piano, a person needs to sit at it, on stage and reach for the keys; feelings

On stage, a woman takes a seat at the piano. She

1. sits on a bench as her sister plays with the doll.
2. smiles with someone as the music plays.
3. is in the crowd, watching the dancers.
4. nervously sets her fingers on the keys.

FrameNet: performer entertains audience

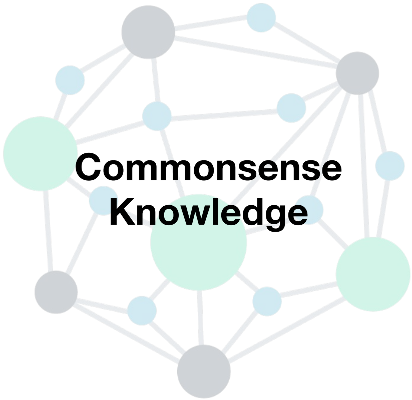
Audience [Aud]	The Audience experiences the Performance.
Medium [Medium]	Medium is the physical entity or channel used by the Performer to transmit the Performance to the Audience.
Performance [Perance]	The Performers generates the Performance which the Audience perceives.
Performer [Perfer]	The Performer provides an experience for the Audience.

Visual Genome: person can play a piano while sitting, his hands are on the keyboard

- man plays piano
- keys ON piano
- woman watches man
- pillow ON couch
- light ON wall
- window IN room
- person playing piano
- guy ON bench
- hands ON keyboard

How to enhance a natural-language agent with commonsense knowledge?

+



Commonsense knowledge sources are heterogeneous

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

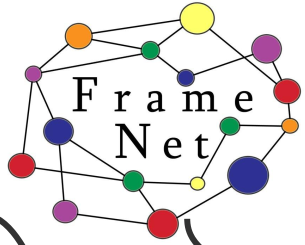
ConceptNet

NELL

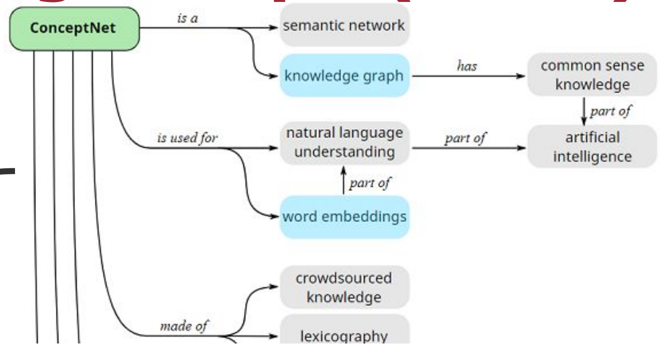
Wikidata

OpenCyc

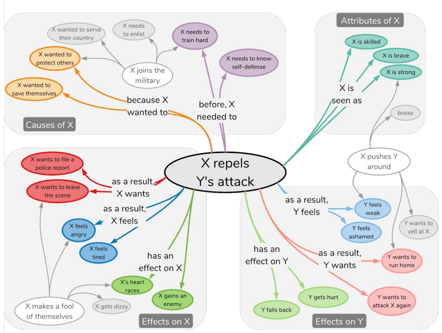
The Commonsense Knowledge Graph (CSKG)



FrameNet (Baker et al., 1998)



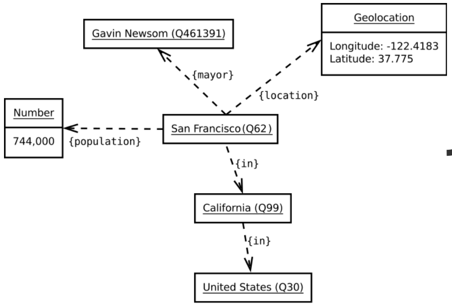
ConceptNet (Speer, Chin and Havasi 2017)



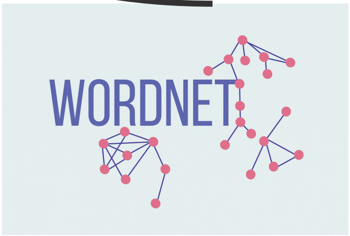
ATOMIC (Sap et al. 2019)

CSKG

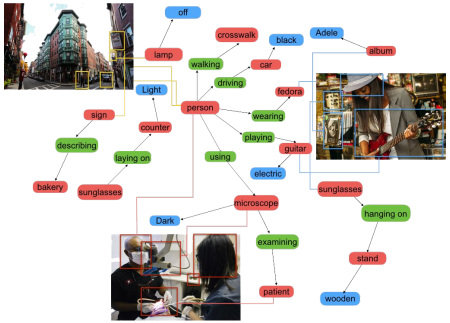
CSKG (Ilievski et al. 2021)



Wikidata (Vrandecic and Krotzsch 2014)



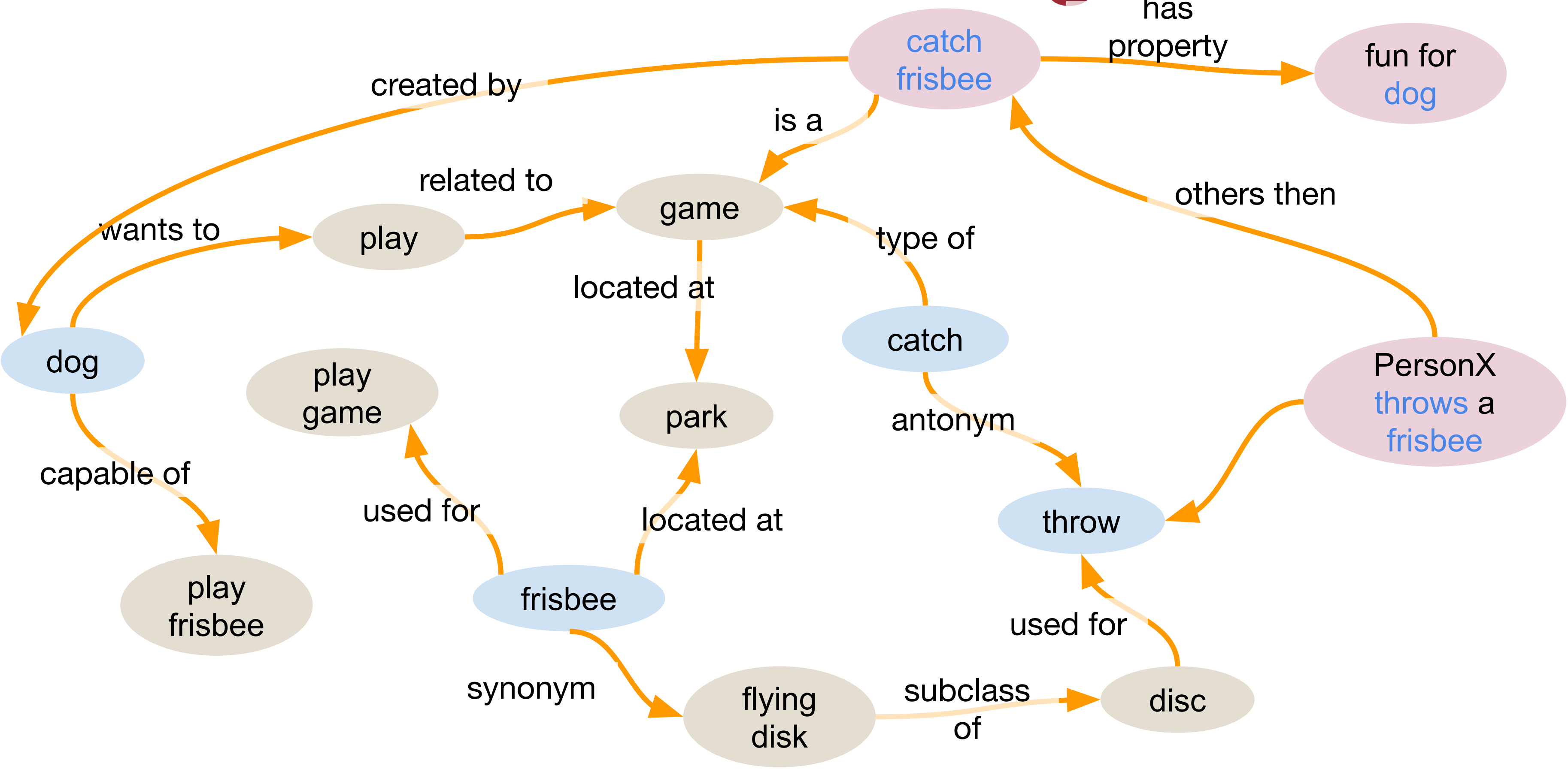
WordNet (Miller 1995)



Visual Genome (Krishna et al. 2017)

F Ilievski, P Szekely, B Zhang (2021). CSKG: The CommonSense Knowledge Graph. ESWC

Consolidated knowledge

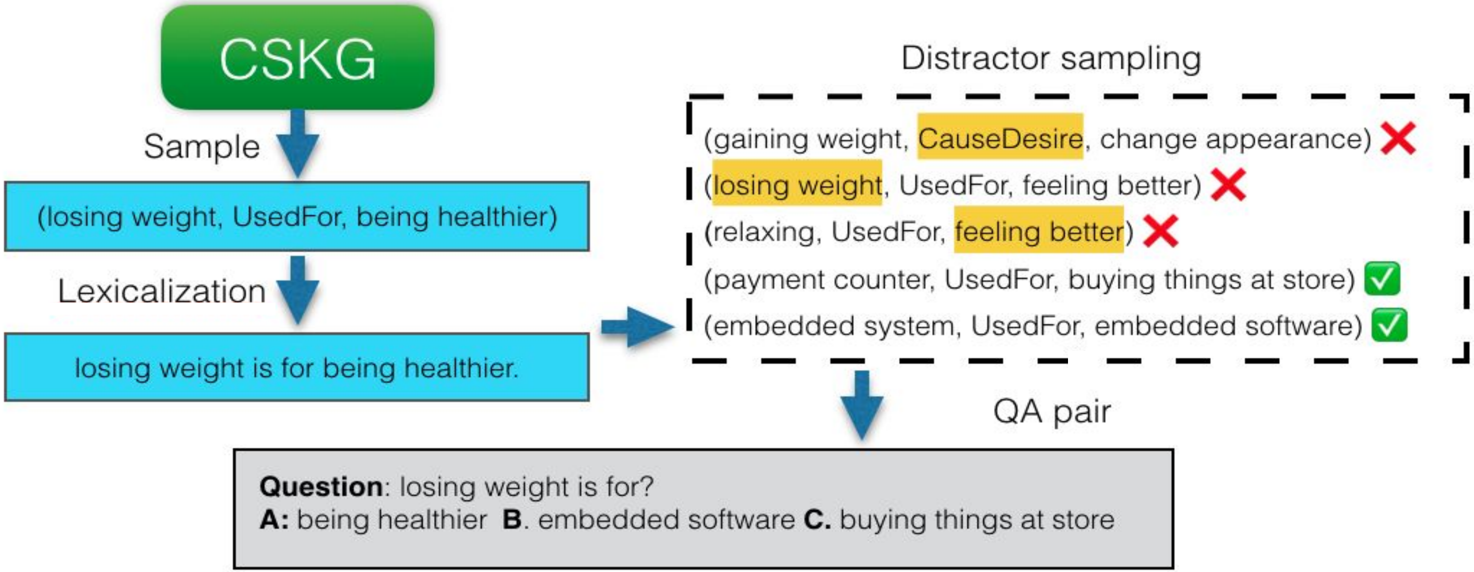


Generating questions with CSKG

[Ma et al., 2021]

Pretrain LMs with synthetic MCQA sets generated from CSKG

Answer commonsense questions on unseen datasets



Pre-training LMs with CSKG questions

SocialIQA test data

CSQA test data

PhysicalIQA test data

Winogrande test data

aNLI test data

CSKG

+

Question: losing weight is for?
A: being healthier **B**. embedded software **C.** buying things at store

=

Experimental setup

CSKG

KG subsets
- *ATOMIC*
- *CWWV (ConceptNet, WordNet, Wikidata, Visual Genome)*
- *full CSKG*

Language models
- *GPT-2*
- *RoBERTa*

+

Question: losing weight is for?
A: being healthier **B**. embedded software **C.** buying things at store

=

SocialIQA test data

CSQA test data

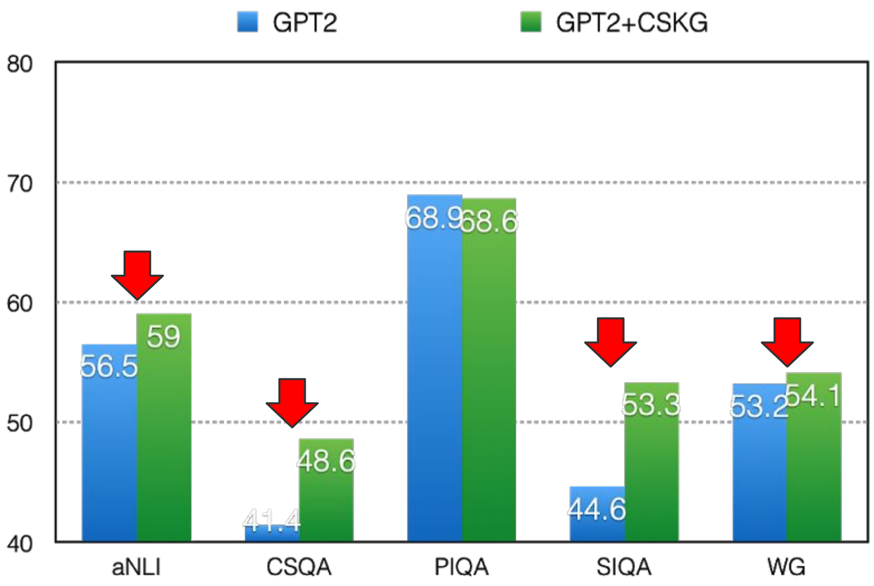
PhysicalIQA test data

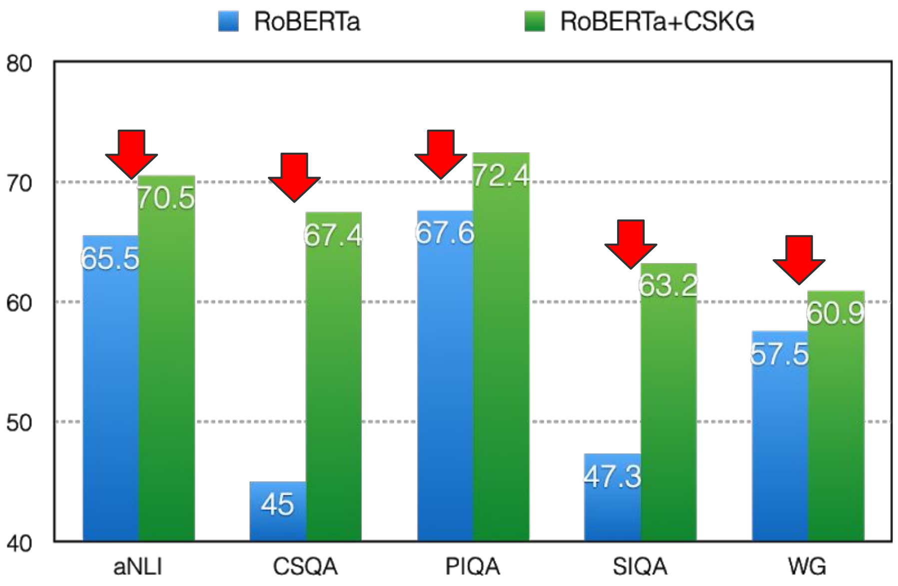
Winogrande test data

aNLI test data

5 different tasks

Pretraining on CSKG MCQA sets helps accuracy





Adding knowledge improves accuracy

Model	KG	aNLI	CSQA	PIQA	SIQA	WG
Majority	-	50.8	20.9	50.5	33.6	50.4
GPT2-L	-	56.5	41.4	68.9	44.6	53.2
RoBERTa-L	-	65.5	45.0	67.6	47.3	57.5
Self-talk (Shwartz et al. 2020)	-	-	32.4	70.2	46.2	54.7
COMET-CGA (Bosselut and Choi 2019)	ATOMIC	-	-	-	49.6	-
SMLM (Banerjee and Baral 2020)	ATOMIC	-	-	-	48.5	-
GPT2-L (MR)	ATOMIC	59.2($\pm$0.3)	48.0($\pm$0.9)	67.5($\pm$0.7)	53.5($\pm$0.4)	54.7($\pm$0.6)
GPT2-L (MR)	CWWV	58.3($\pm$0.4)	46.2($\pm$1.0)	68.6($\pm$0.7)	48.0($\pm$0.7)	52.8($\pm$0.9)
GPT2-L (MR)	CSKG	59.0($\pm$0.5)	48.6($\pm$1.0)	68.6($\pm$0.9)	53.3($\pm$0.5)	54.1($\pm$0.5)
RoBERTa-L (MR)	ATOMIC	**70.8**(**$\pm$1.2**)	64.2($\pm$0.7)	72.1($\pm$0.5)	63.1($\pm$1.5)	59.6($\pm$0.3)
RoBERTa-L (MR)	CWWV	70.0($\pm$0.3)	**67.9**(**$\pm$0.8**)	72.0($\pm$0.7)	54.8($\pm$1.2)	59.4($\pm$0.5)
RoBERTa-L (MR)	CSKG	70.5($\pm$0.2)	67.4($\pm$0.8)	**72.4**(**$\pm$0.4**)	**63.2**(**$\pm$0.7**)	**60.9**(**$\pm$0.8**)
RoBERTa-L (supervised)	-	85.6	78.5	79.2	76.6	79.3
Human	-	91.4	88.9	94.9	86.9	94.1

The impact of more knowledge depends on the KG-task alignment

Model	KG	aNLI	CSQA	PIQA	SIQA	WG
Majority	-	50.8	20.9	50.5	33.6	50.4
GPT2-L	-	56.5	41.4	68.9	44.6	53.2
RoBERTa-L	-	65.5	45.0	67.6	47.3	57.5
Self-talk (Shwartz et al. 2020)	-	-	32.4	70.2	46.2	54.7
COMET-CGA (Bosselut and Choi 2019)	ATOMIC	-	-	-	49.6	-
SMLM (Banerjee and Baral 2020)	ATOMIC	-	-	-	48.5	-
GPT2-L (MR)	ATOMIC	59.2(±0.3)	48.0(±0.9)	67.5(±0.7)	53.5(±0.4)	54.7(±0.6)
GPT2-L (MR)	CWWV	58.3(±0.4)	46.2(±1.0)	68.6(±0.7)	48.0(±0.7)	52.8(±0.9)
GPT2-L (MR)	CSKG	59.0(±0.5)	48.6(±1.0)	68.6(±0.9)	53.3(±0.5)	54.1(±0.5)
RoBERTa-L (MR)	ATOMIC	**70.8**(±**1.2**)	64.2(±0.7)	72.1(±0.5)	63.1(±1.5)	59.6(±0.3)
RoBERTa-L (MR)	CWWV	70.0(±0.3)	**67.9**(±**0.8**)	72.0(±0.7)	54.8(±1.2)	59.4(±0.5)
RoBERTa-L (MR)	CSKG	70.5(±0.2)	67.4(±0.8)	**72.4**(±**0.4**)	**63.2**(±**0.7**)	**60.9**(±**0.8**)
RoBERTa-L (supervised)	-	85.6	78.5	79.2	76.6	79.3
Human	-	91.4	88.9	94.9	86.9	94.1

Takeaways

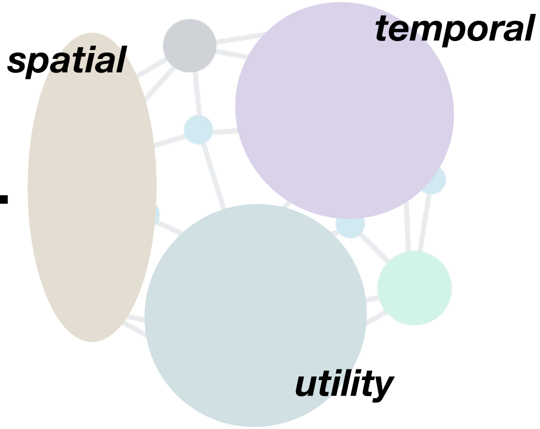
A single pre-trained model can learn to perform all the tasks

Pretraining with CSKG knowledge brings large and consistent improvement over vanilla LMs

More knowledge helps, IF well-aligned with the task

What is the role of different types of commonsense knowledge?

+



Approach

1. Organize the CSKG knowledge along dimensions (knowledge types)

2. Align knowledge dimensions with tasks

F Ilievski, A Oltramari, K Ma, B Zhang, DL McGuinness, P Szekely (2021). Dimensions of commonsense knowledge. Under review

Approach

1. *Organize the CSKG knowledge along dimensions (knowledge types)*
   a. *Survey relation types in relevant sources*
   b. *Manually cluster them into dimensions*

2. Align knowledge dimensions with tasks

F Ilievski, A Oltramari, K Ma, B Zhang, DL McGuinness, P Szekely (2021). Dimensions of commonsense knowledge. Under review

Overview of sources

Category	Source	Relations
Commonsense KGs	ConceptNet*	34
	ATOMIC	9
	GLUCOSE	10
	WebChild	4 (groups)
	Quasimodo	78,636
	SenticNet	4
	HasPartKB	1
Common KGs	Wikidata	6.7k
	YAGO4	116
	DOLCE*	1
	SUMO*	1,614
Lexical resources	WordNet	10
	Roget	2
	FrameNet	8 (f2f)
	MetaNet	14 (f2f)
	VerbNet	36 (roles)
Visual sources	Visual Genome	42,374
	Flickr30k	1
Corpora & LMs	GenericsKB	n/a
	GPT-2	n/a

F Ilievski, A Oltramari, K Ma, B Zhang, DL McGuinness, P Szekely (2021). Dimensions of commonsense knowledge. Under review

Annotation of dimensions (1/2)

Dimension	ATOMIC	ConceptNet	WebChild	Other	Wikidata
lexical		FormOf DerivedFrom EtymologicallyDerivedFrom		lexical_unit (FN) lemma (WN)	label
similarity		Synonym SimilarTo DefinedAs	hassimilar	reframing_mapping (FN) metaphor (FN) Synonym (RG) synonym (WN)	said to be the same as
distinctness		Antonym DistinctFrom		Antonym (RG) antonym (WN) excludes (FN)	different from opposite of
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

Annotation of dimensions (2/2)

Dimension	ATOMIC	ConceptNet	WebChild	Other	Wikidata
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
relational-other		RelatedTo HasContext EtymologicallyRelatedTo	thing agent	see_also (FN) requires (FN)	field of this occupation depicts health specialty

Dimensions of commonsense knowledge in bottom-up sources

lexical

similarity

distinctness

taxonomic

part-whole

spatial

creation

utility

desire/goal

quality

comparative

temporal

relational-other

Approach

1. Organize the CSKG knowledge along dimensions (knowledge types)
    a. Survey relation types in relevant sources
    b. Manually cluster them into dimensions

2. Align knowledge dimensions with tasks
    a. Split CSKG along its dimensions
    b. Generate pre-training data per dimension
    c. Measure the impact of each dimension against each task

F Ilievski, A Oltramari, K Ma, B Zhang, DL McGuinness, P Szekely (2021). Dimensions of commonsense knowledge. Under review

Dimension-aware LM pre-training

dimension → CSKG

+

Question: losing weight is for?
A: being healthier **B**. embedded software **C.** buying things at store

=

SocialIQA test data

CSQA test data

PhysicalIQA test data

Winogrande test data

aNLI test data

One dimension at a time to measure their impact on a task

Dimensions	Train	Dev
part-whole	87,765	4,620
taxonomic	340,609	17,927
lexical	107,861	5,677
distinctness	20,286	1,068
similarity	166,575	8,768
quality	116,593	12,492
utility	63,862	3,362
creation	304	17
temporal	312,628	31,587
relational-other	242,759	12,777
spatial	21,726	1,144
desire/goal	194,906	20,912

Pre-training language models with dimensions

CSQA = Commonsense QA

SIQA = SocialIQA

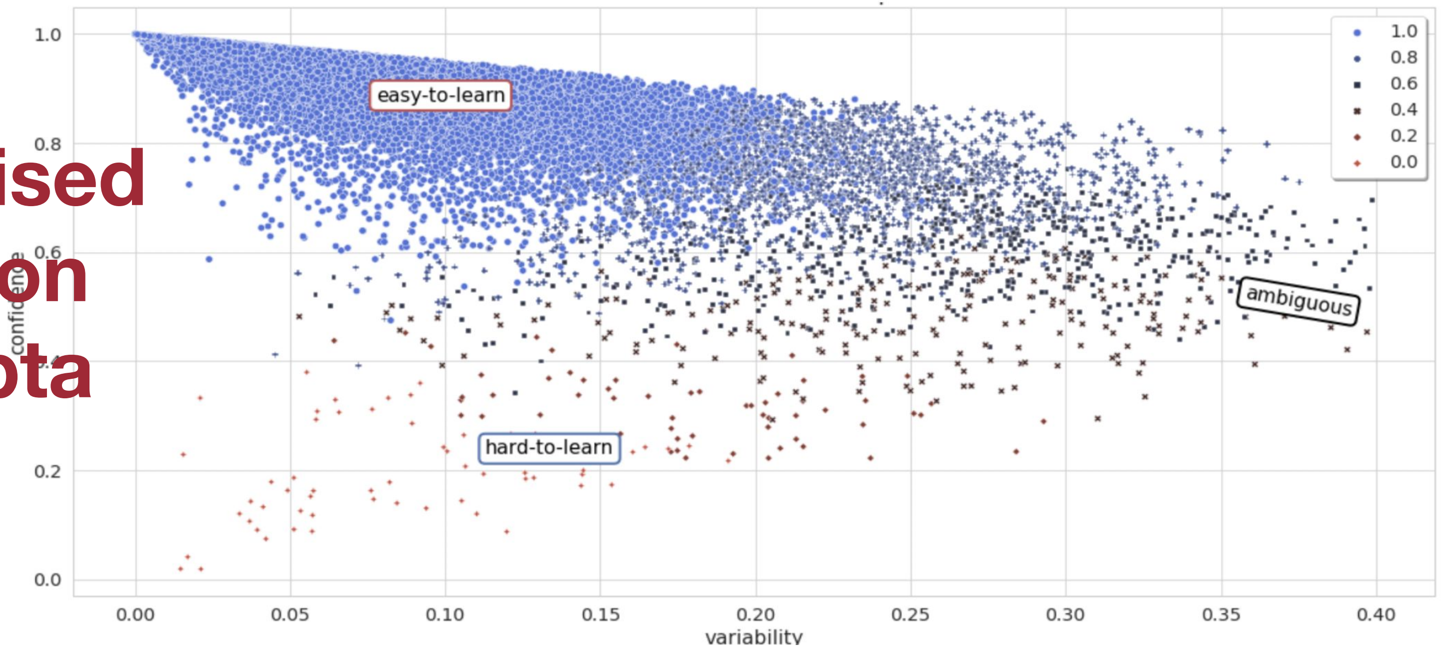
Dimensions	CSQA	SIQA
Baseline	45.0	47.3
+part-whole	63.0(±1.4)	52.6(±1.9)
+taxonomic	62.6(±1.4)	52.2(±1.6)
+lexical	49.9(±2.9)	49.0(±0.4)
+distinctness	57.2(±0.5)	50.2(±1.5)
+similarity	61.4(±0.8)	53.5(±0.6)
+quality	65.7(±0.5)	60.0(±0.7)
+utility	**67.4**(**±1.0**)	54.8(±0.7)
+creation	49.9(±1.1)	47.8(±0.2)
+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)

Novelty per dimension

Can 'vanilla' RoBERTa answer the questions without pretraining?

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

Self-supervised data selection [Swayamdipta et al., 2020]



	amb	amb+hard	amb+easy10	amb+easy30	amb+easy60	amb+easy100	all
aNLI	0.658	0.644	0.734	0.721	0.720	0.710	0.702
CSQA	0.327	0.387	0.633	0.662	0.662	0.661	0.640
PIQA	0.607	0.614	0.720	0.708	0.726	0.712	0.709
SIQA	0.432	0.439	0.625	0.633	0.631	0.621	0.610
WG	0.555	0.530	0.599	0.581	0.597	0.600	0.587

Findings

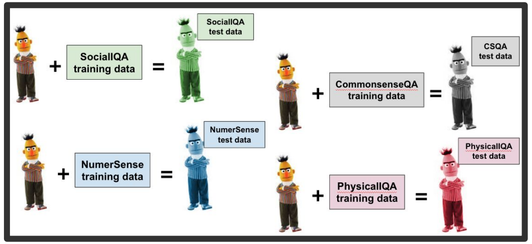
The knowledge dimensions of CSKG allow for controlled adaptation of LMs

Not all dimensions are as informative to LMs

- **lexical and distinctness knowledge is largely redundant**
- **temporal and desire/goal knowledge is both novel and useful**

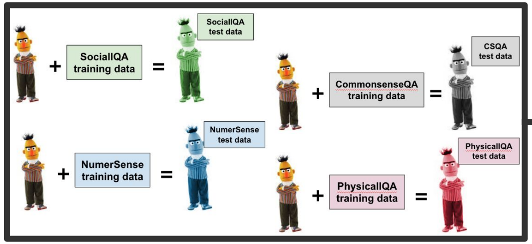
Data maps (Swayamdipta et al., 2020) could be applied for finer QA filtering

The story so far



Many dataset-specific NL commonsense agents

The story so far



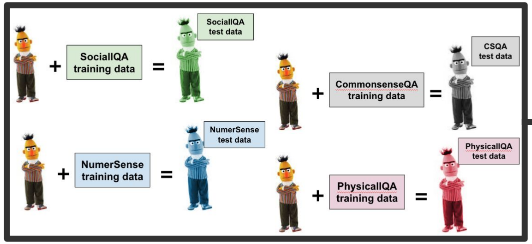
Many dataset-specific NL commonsense agents

Pre-training LMs with commonsense knowledge

KG pre-training method

CSKG

The story so far



Many dataset-specific NL commonsense agents

Pretraining LMs with all dimensions of commonsense knowledge

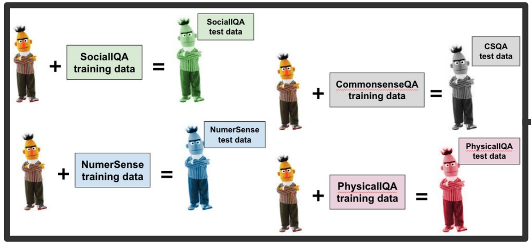
KG pre-training method

CSKG

dimensions

data maps

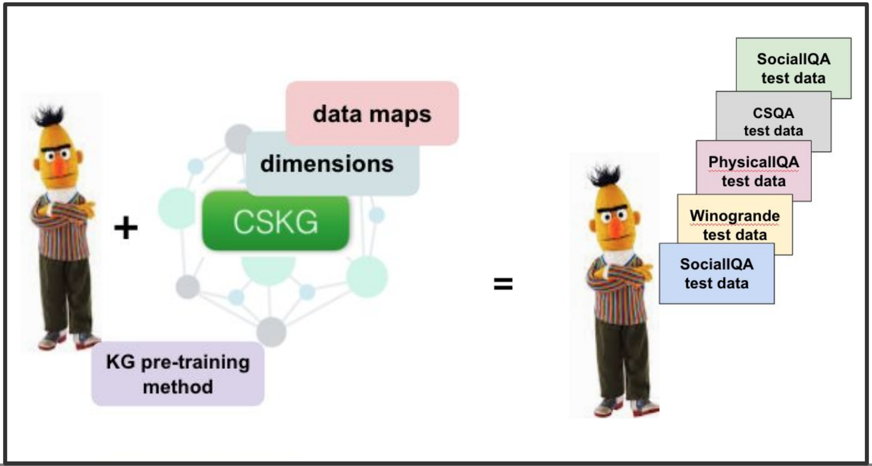
The story so far



Many dataset-specific NL commonsense agents

Pretraining LMs with all dimensions of commonsense knowledge

KG pre-training method | **CSKG** | **dimensions** | **data maps**



A single NL agent for all commonsense dimensions

What about 'perform well'?

Model	**KG**	**aNLI**	**CSQA**	**PIQA**	**SIQA**	**WG**
Majority	-	50.8	20.9	50.5	33.6	50.4
RoBERTa-L	-	65.5	45.0	67.6	47.3	57.5
RoBERTa-L (MR)	CSKG	70.5(±0.2)	67.4(±0.8)	**72.4**(±**0.4**)	**63.2**(±**0.7**)	**60.9**(±**0.8**)
RoBERTa-L (supervised)	-	85.6	78.5	79.2	76.6	79.3
Human	-	91.4	88.9	94.9	86.9	94.1