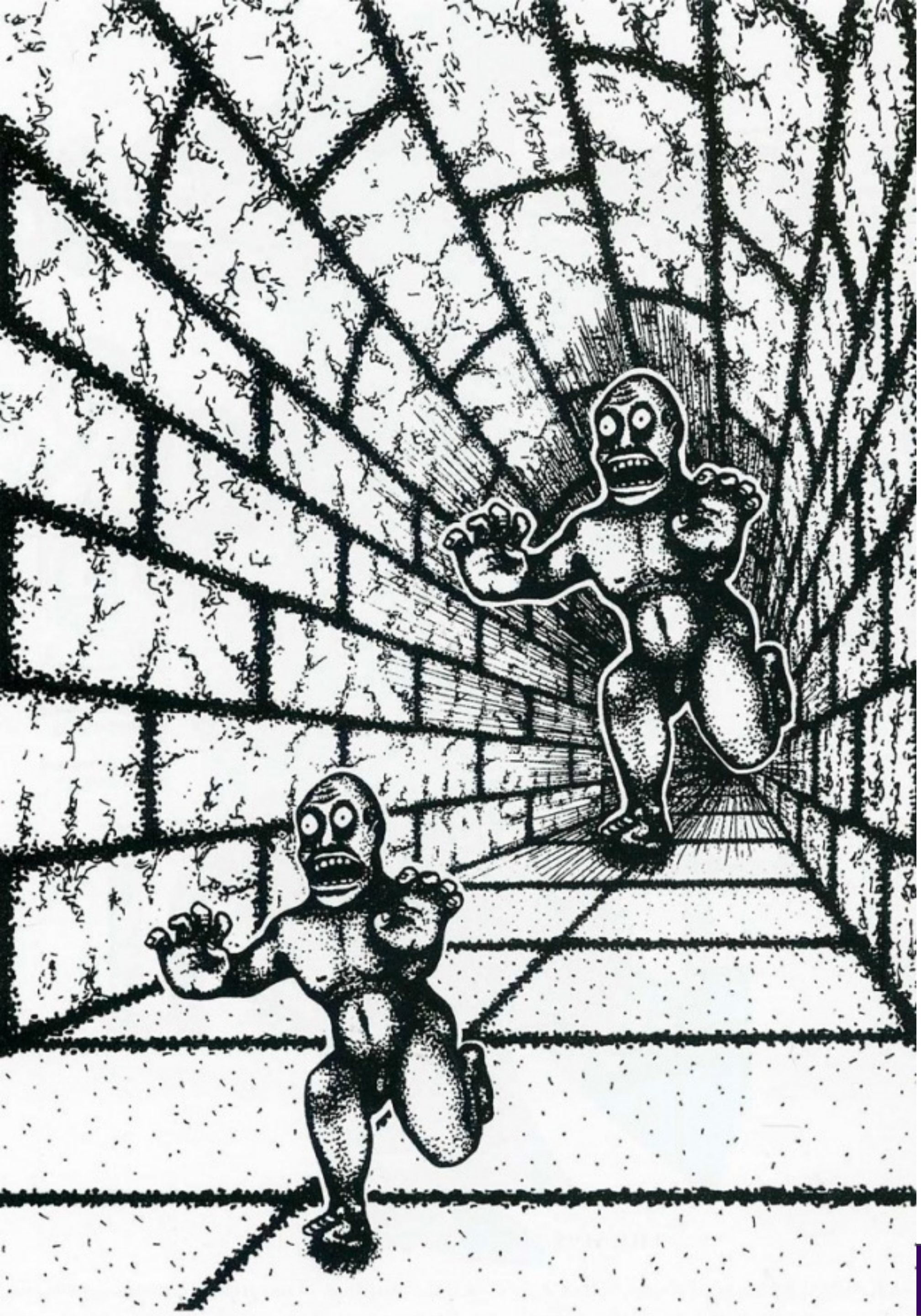
CIS 700: Interactive Fiction and Text Generation

Commonsense Reasoning

These slides are from the ACL 2020 Commonsense Tutorial by Yejin Choi, Vered Shwartz, Maarten Sap, Antoine Bosselut, and Dan Roth

Penn Engineering
UNIVERSITY of PENNSYLVANIA



Monsters in a Tunnel

- **Two monsters are running** (rather than standing still on one foot)
- **One is chasing another** (rather than trying to copy his movements)
- **The chaser has hostile intentions and the chased is afraid** (even though two faces are identical)

Important Observations:

- A great deal of **intuitive inferences** are **commonsense inferences**, which can be described in **natural language**.
- None of these inferences is absolutely true. The inferences are **stochastic** in nature. Everything is **defeasible** with additional context.
- Commonsense inferences are about **predicting new information** that **is likely to be true** based on partially available information.

Claims of AI systems reaching a “human level”

GOOGLE \ WEB \ APPS

Google's AI translation system is approaching human-level accuracy

But there's still significant work to be done

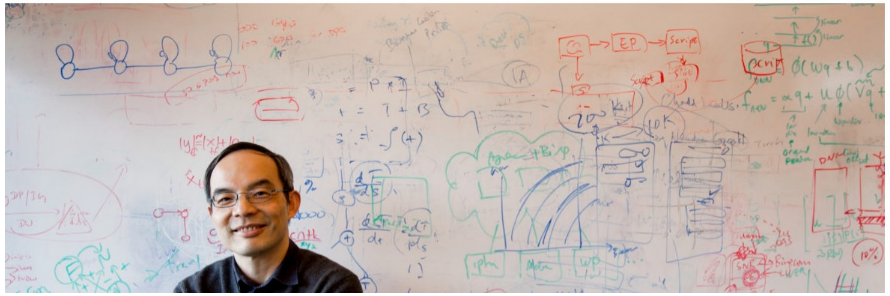
By Nick Statt | @nickstatt | Sep 27, 2016, 2:07pm EDT

SHARE

Microsoft claims new speech recognition record, achieving a super-human 5.1% error rate

BY TODD BISHOP on August 20, 2017 at 7:44 pm

5 Comments | Share 820 | Tweet | Share | Reddit | Email



Microsoft, Google Beat Humans at Image Recognition

Deep learning algorithms compete at ImageNet challenge

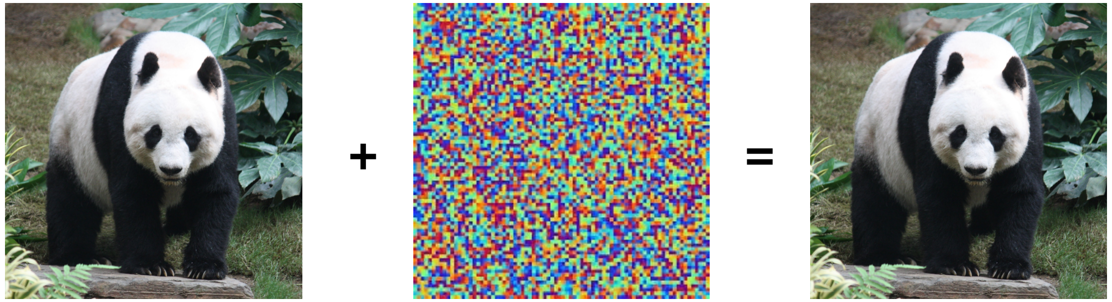
R. Colin Johnson
2/18/2015 08:15 AM EST
14 comments

Alibaba and Microsoft AI beat human scores on Stanford reading test

Neural networks edged past human scores on the measure of machine reading.

Rob LeFebvre, @roblef
01.15.18 in Personal Computing

10 Comments | 937 Shares



Giant panda

Gibbon

Object Recognition

Szegedy et al, 2014....

VQA

Jabri et al, 2017

We may be "solving" datasets rather than the underlying "task"

A horse standing in the grass.

Captioning

MacLeod et al, 2017

a Tesla moved to Prague in 1880. … Tadakatsu moved to Chicago in 1881.

Where did Tesla move in 1880? Chicago

QA

Jia et al, 2017

Theory of Core Knowledge

Domain	Description
Objects	supports reasoning about objects and the laws of physics that govern them
Agents	supports reasoning about agents that act autonomously to pursue goals
Places	supports navigation and spatial reasoning around an environment
Number	supports reasoning about quality and how many things are present
Forms	supports representation of shapes and their affordances
Social Beings	supports reasoning about Theory of Mind and social interaction

Developmental psychologists have shown that children develop the ability to reason about these domains early in life. Such reasoning is important for later learning.

Definition of Common Sense

The basic level of **practical knowledge** and **reasoning** concerning **everyday situations** and **events** that are **commonly** shared among **most** people.

It's OK to keep the closet door open

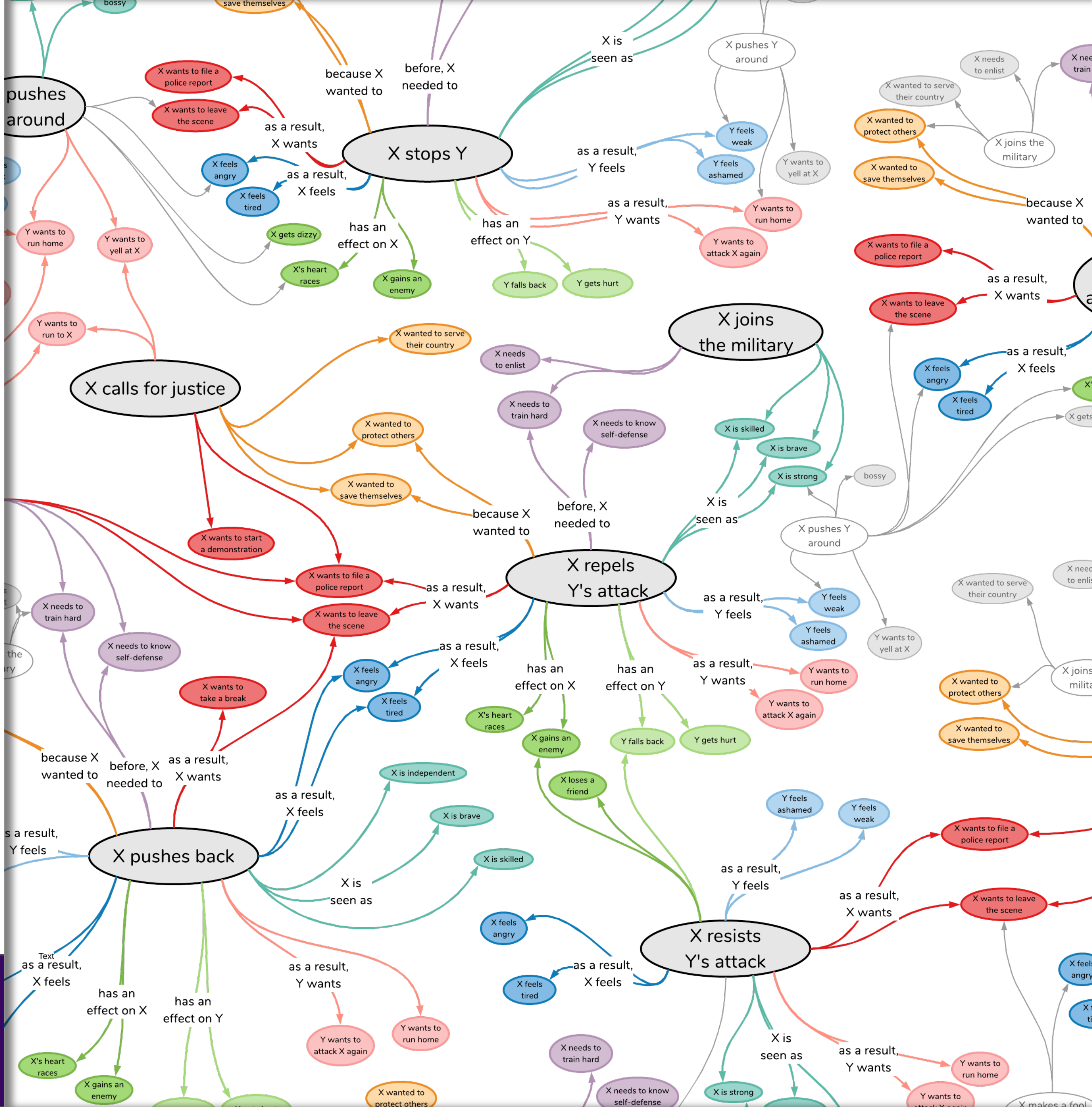
It's not OK to keep the refrigerator door open because the food might go bad

Essential for humans to live and interact with each other in a reasonable and safe way

Essential for AI to understand human needs and actions better

CIS 700: Interactive Fiction and Text Generation

Commonsense resources



Grandma’s glasses

Tom’s grandma was reading a new book, when she dropped her glasses.

She couldn’t pick them up, so she called Tom for help.

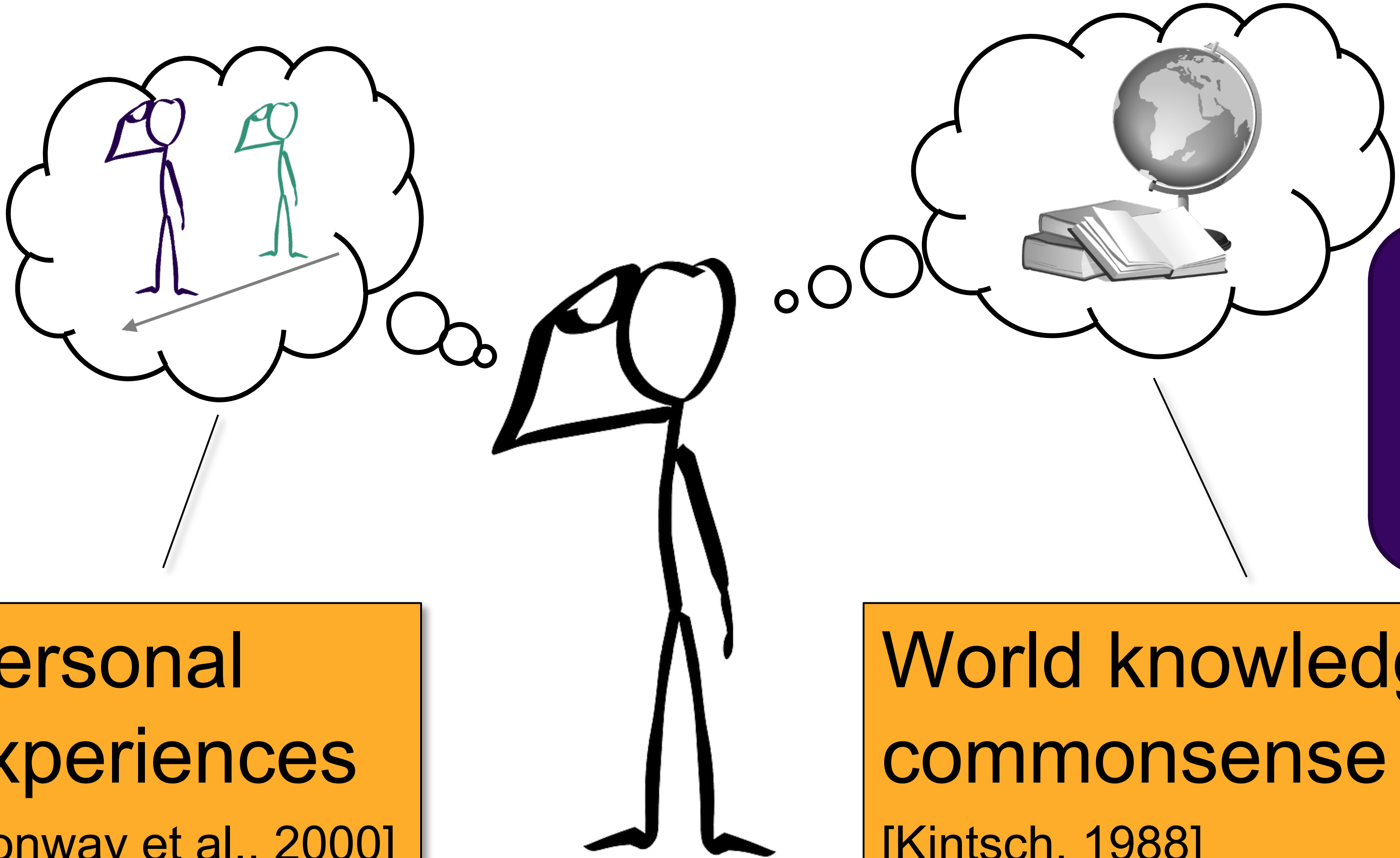
Tom rushed to help her look for them, they heard a loud crack.

They realized that Tom broke her glasses by stepping on them.

Promptly, his grandma yelled at Tom to go get her a new pair.

Humans reason about the world with **mental models**

[Graesser, 1994]



Personal experiences

[Conway et al., 2000]

World knowledge and commonsense

[Kintsch, 1988]

Commonsense resources aim to be a bank of knowledge for machines to be able to reason about the world in tasks

Tom’s grandma was reading a new book, when she dropped her glasses.

She couldn’t pick them up, so she called Tom for help.

Tom rushed to help her look for them, they heard a loud crack.

They realized that Tom broke her glasses by stepping on them.

Promptly, his grandma yelled at Tom to go get her a new pair.

ConceptNet

ATOMIC

usedFor

Tom’s grandma was reading a new book, when she dropped her glasses.

She couldn’t pick them up, so she called Tom for help.

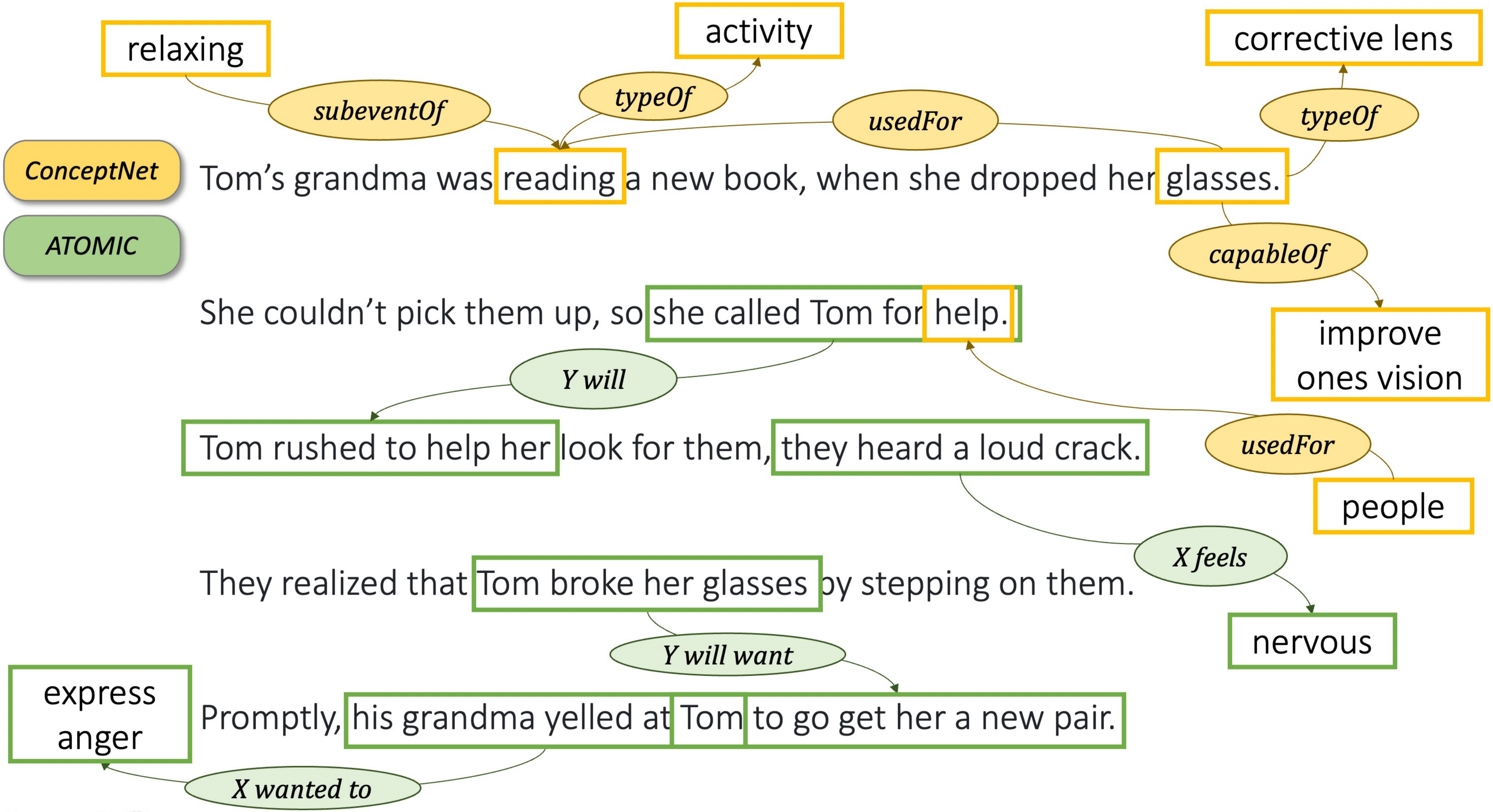
Y will

Tom rushed to help her look for them, they heard a loud crack.

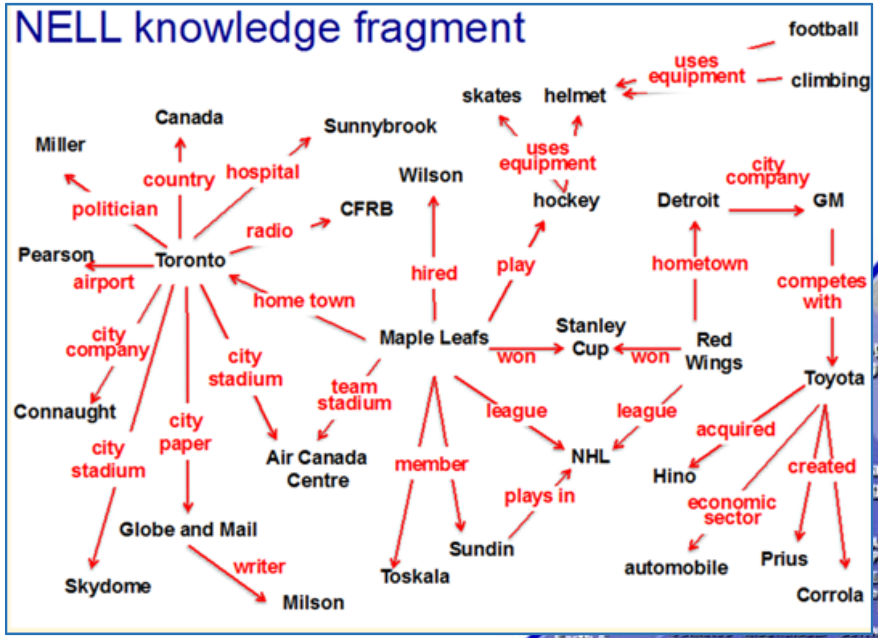
They realized that Tom broke her glasses by stepping on them.

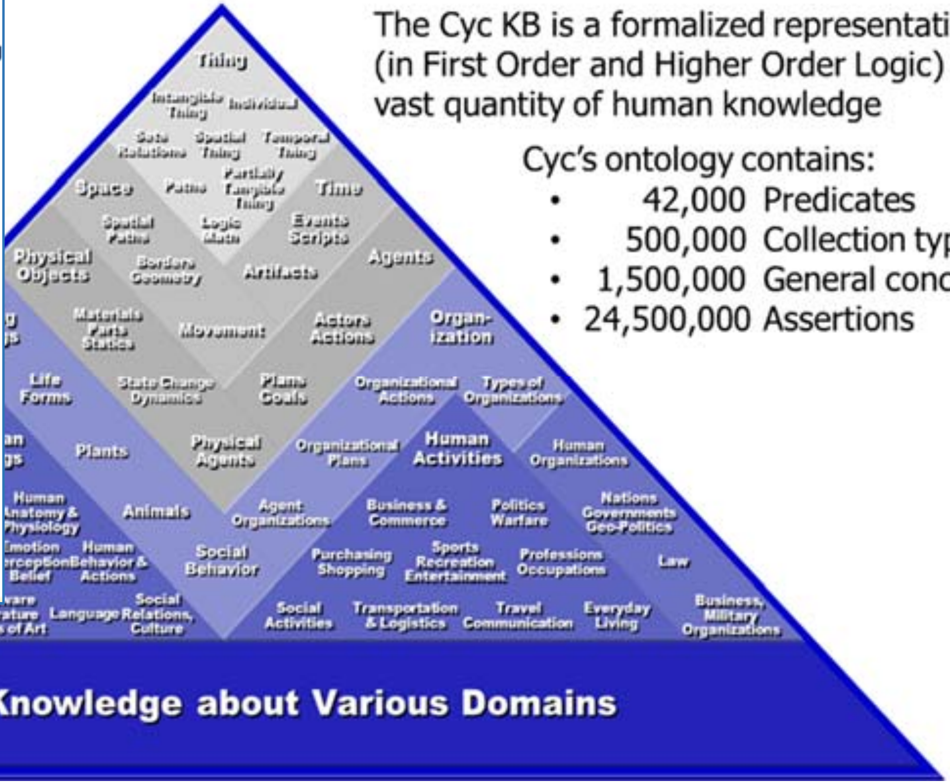
Y will want

Promptly, his grandma yelled at Tom to go get her a new pair.


ConceptNet
ATOMIC
relaxing
subeventOf
activity
typeOf
usedFor
corrective lens
typeOf
Tom's grandma was reading a new book, when she dropped her glasses.
capableOf
improve ones vision
She couldn't pick them up, so she called Tom for help.
Y will
usedFor
people
Tom rushed to help her look for them, they heard a loud crack.
X feels
nervous
They realized that Tom broke her glasses by stepping on them.
Y will want
express anger
Promptly, his grandma yelled at Tom to go get her a new pair.
X wanted to

Overview of existing resources





The Cyc KB is a formalized representation (in First Order and Higher Order Logic) of a vast quantity of human knowledge

Cyc's ontology contains:

- 42,000 Predicates
- 500,000 Collection types
- 1,500,000 General concepts
- 24,500,000 Assertions

ATOMIC (…ap et al., 2019)

…ild 2.0 (…n et al., …7)

…Net 5.5 (…al., 2017)

Open Mind Comm… (Minsky, Singh & … 1999)

Cyc (Lenat et al., 1984)

OpenCyc (Lenat, 2004)

ResearchCyc (Lenat, 2006)

OpenCyc 4.0 (Lenat, 2012)

today

How do you create a commonsense resource?

Desirable properties for a commonsense resource

Coverage

Large scale

Diverse knowledge types

Useful

High quality knowledge

Usable in downstream tasks

Multiple resources tackle different knowledge types

Creating a commonsense resource

- Representation
  - Symbolic
  - Natural language
- Knowledge type
  - Domain-specific
  - Semantic
  - Inferential

CONCEPTNET:
semantic knowledge in *natural language* form

http://conceptnet.io/

Related terms

- en book →
- en books →
- en book →

Effects of reading

- en learning →
- en ideas →
- en a headache →

reading is a type of...

- en an activity →
- en a good way to learn →
- en one way of learning →
- en one way to learn →

reading is a subevent of...

- en you learn →
- en turning a page →
- en learning →

en reading

An English term in ConceptNet 5.8

Subevents of reading

- en relaxing →
- en study →
- en studying for a subject →

Things used for reading

- en article →
- en a library →
- en literature →
- en a paper page →

Types of reading

- en browse (n, communication) →
- en bumf (n, communication) →
- en clock time (n, time) →
- en miles per hour (n, time) →

What is ConceptNet?

General commonsense knowledge

21 million edges and over 8 million nodes (as of 2017)

- Over 85 languages
- In English: over 1.5 million nodes

Knowledge covered:

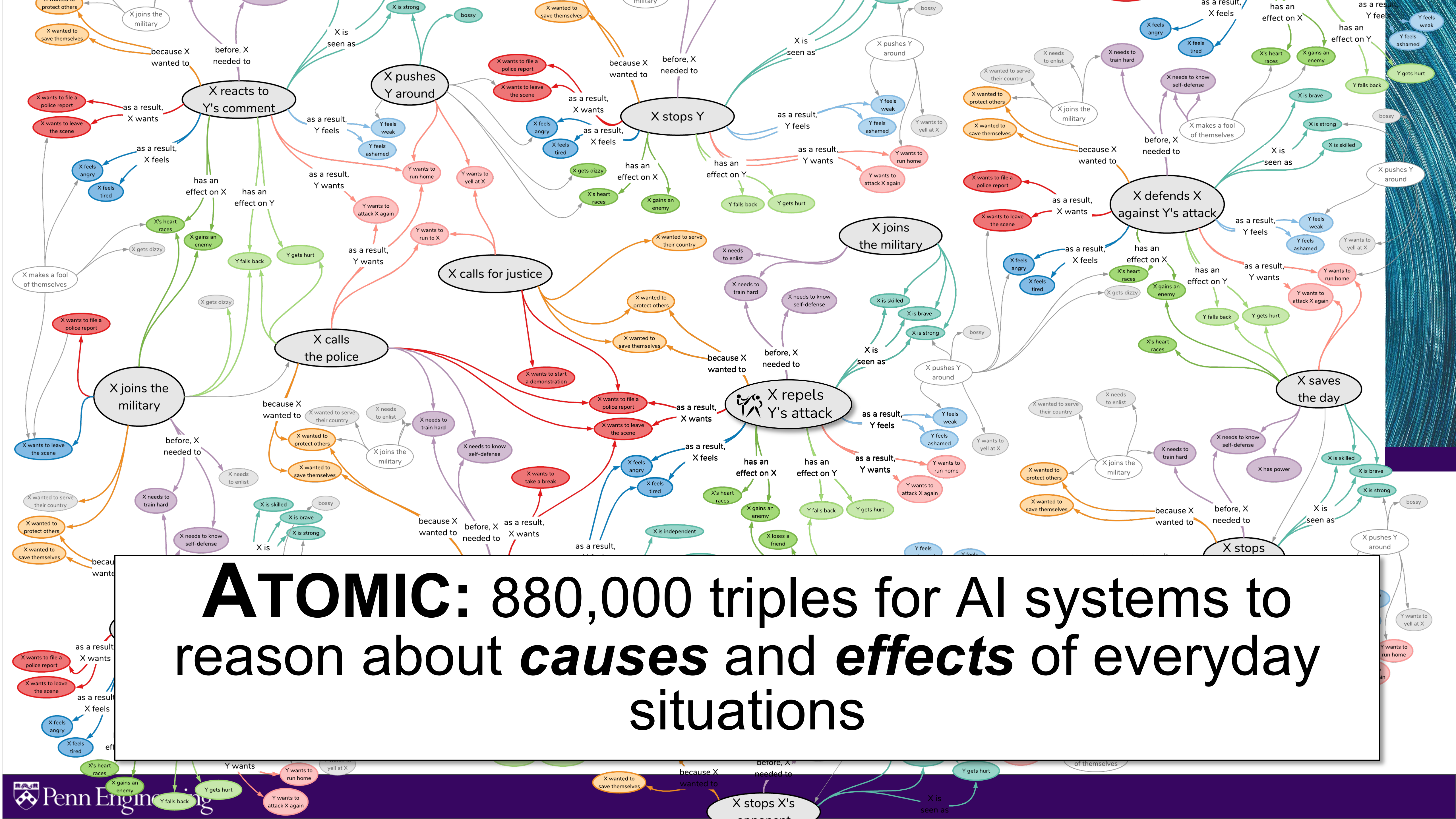
- Open Mind Commonsense assertions
- Wikipedia/Wiktionary semantic knowledge
- WordNet, Cyc ontological knowledge

http://conceptnet.io/

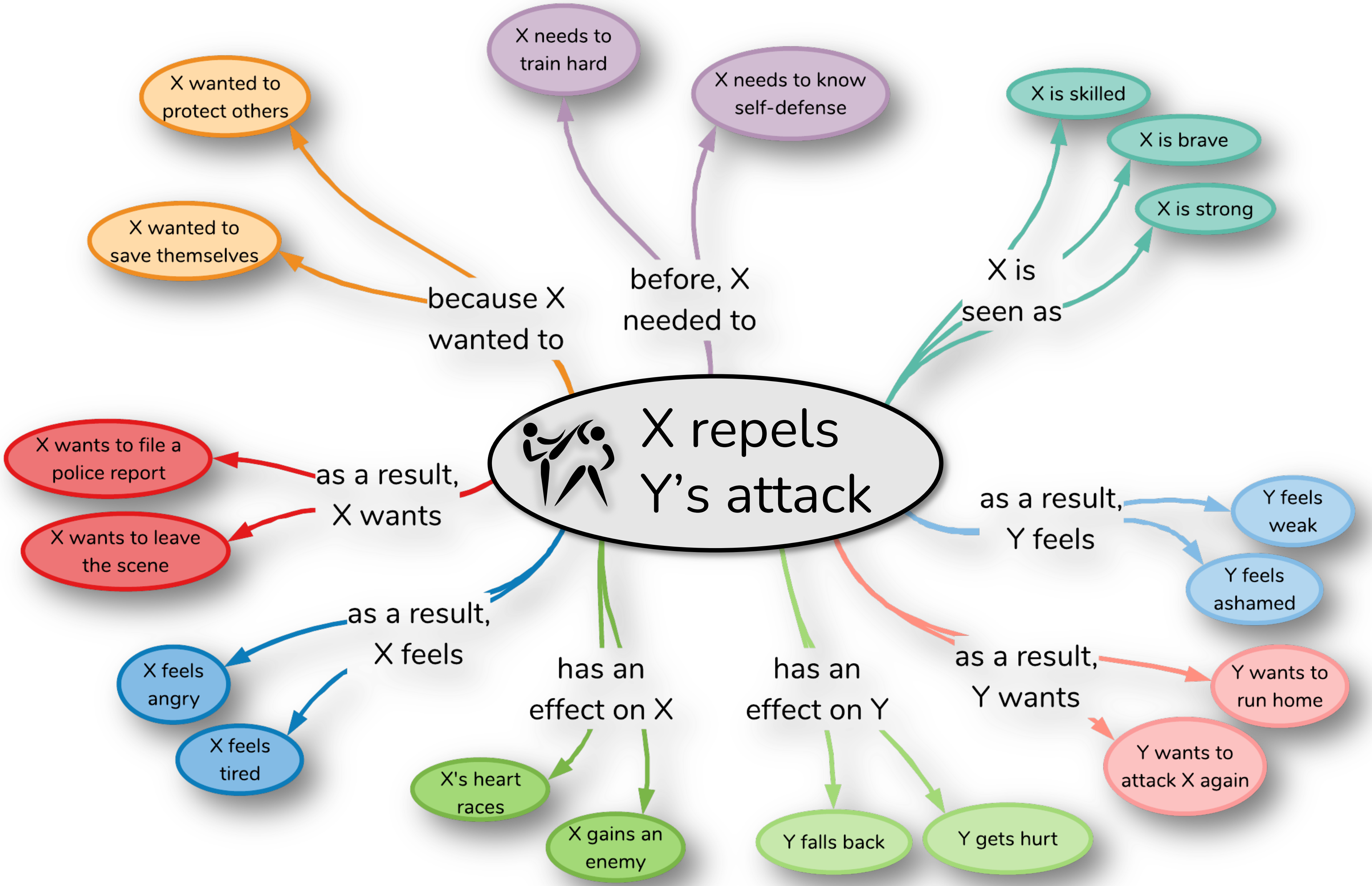
ATOMIC:

inferential knowledge in *natural language* form

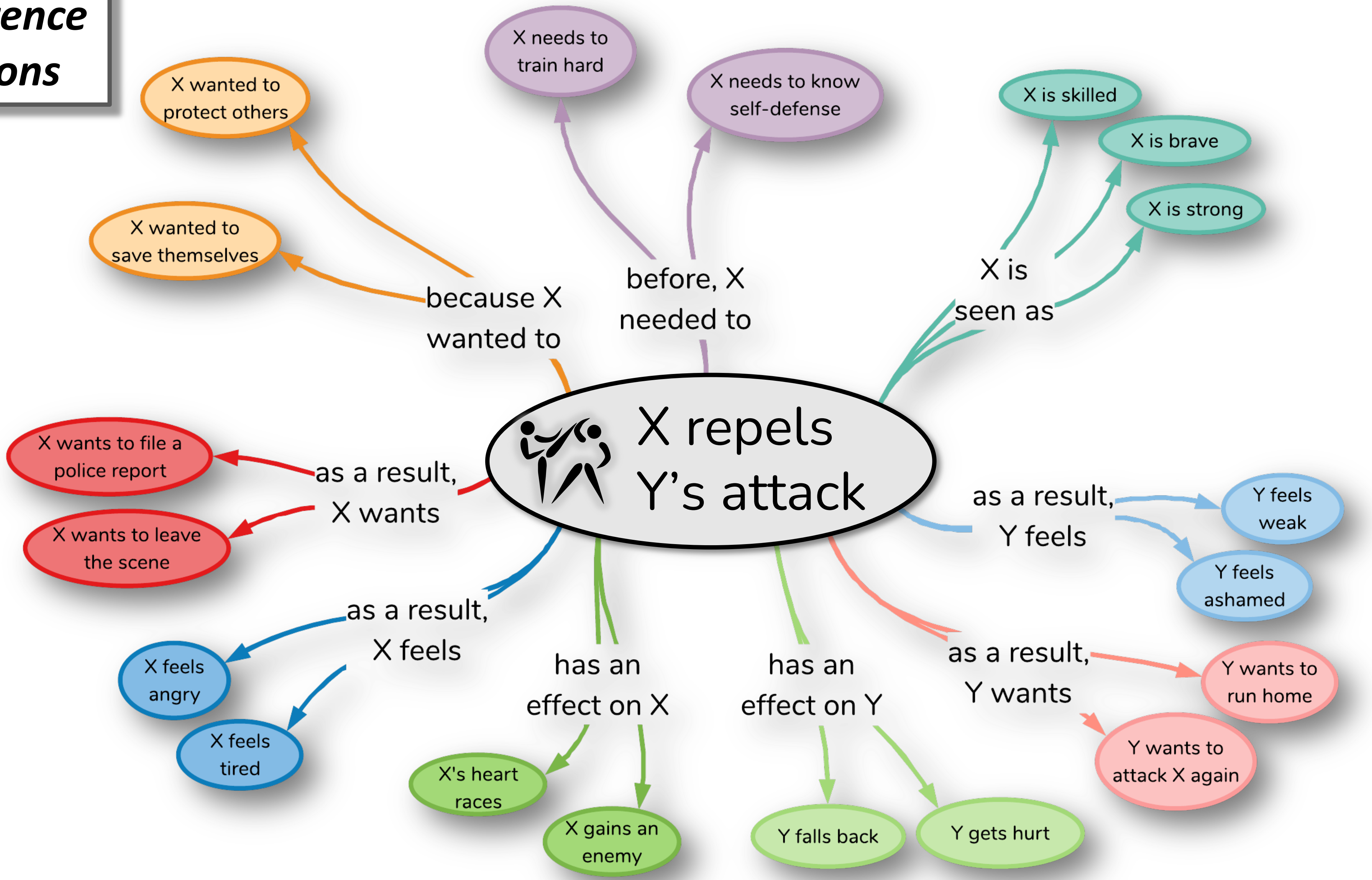
https://mosaickg.apps.allenai.org/kg_atomic

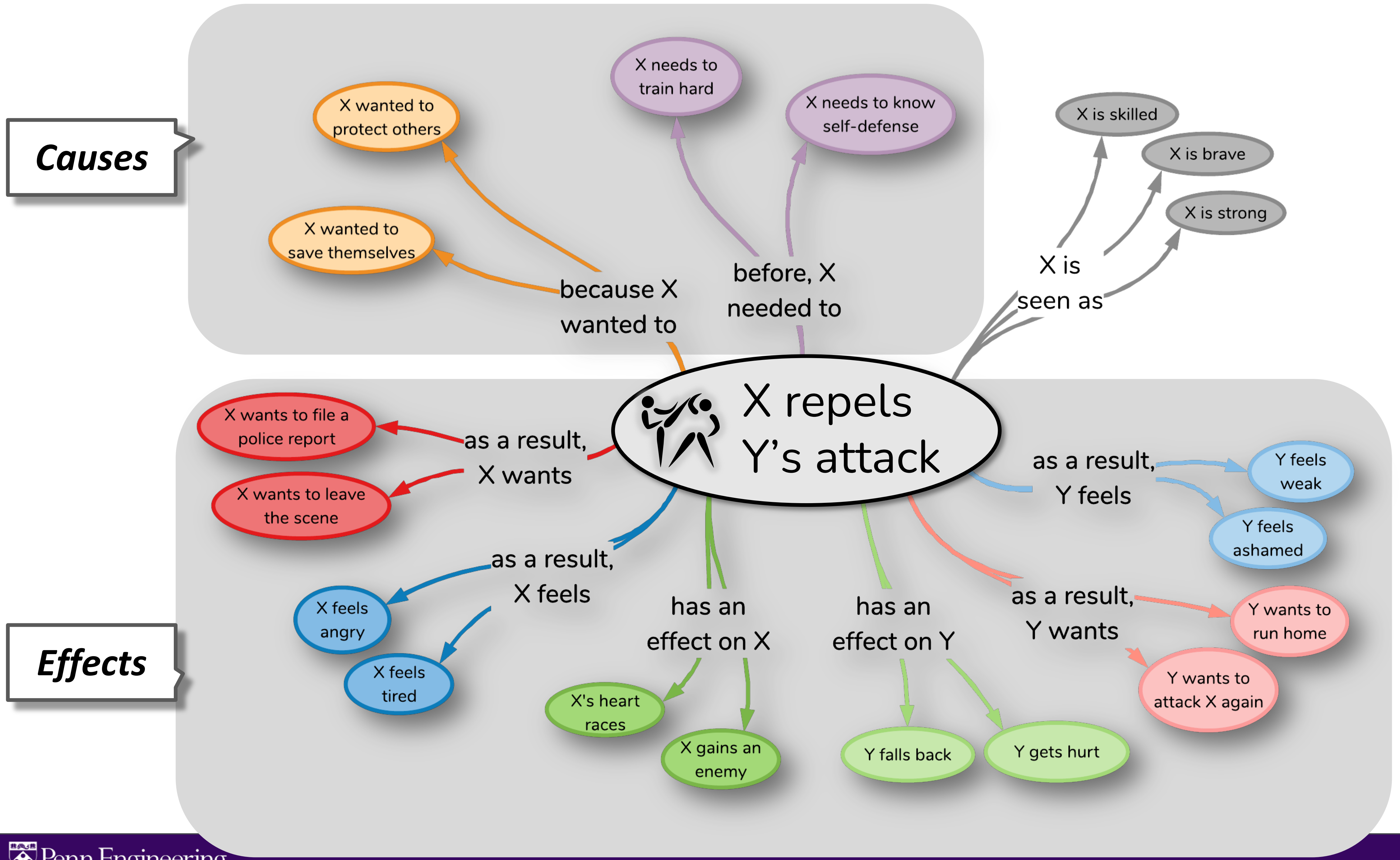


ATOMIC: 880,000 triples for AI systems to reason about ***causes*** and ***effects*** of everyday situations

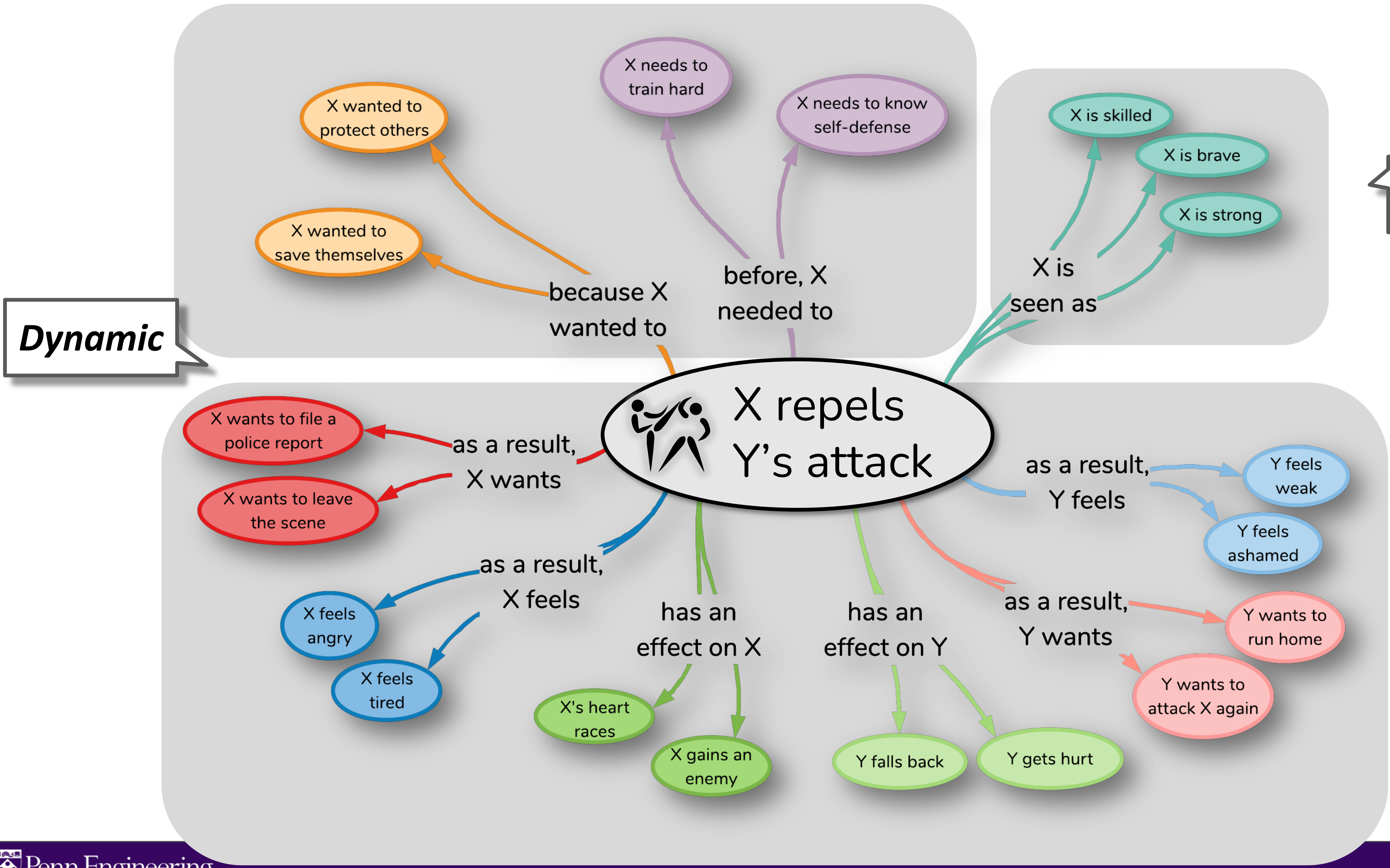

X repels Y's attack
because X wanted to
X wanted to protect others
X wanted to save themselves
before, X needed to
X needs to train hard
X needs to know self-defense
X is seen as
X is skilled
X is brave
X is strong
as a result, X wants
X wants to file a police report
X wants to leave the scene
as a result, 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

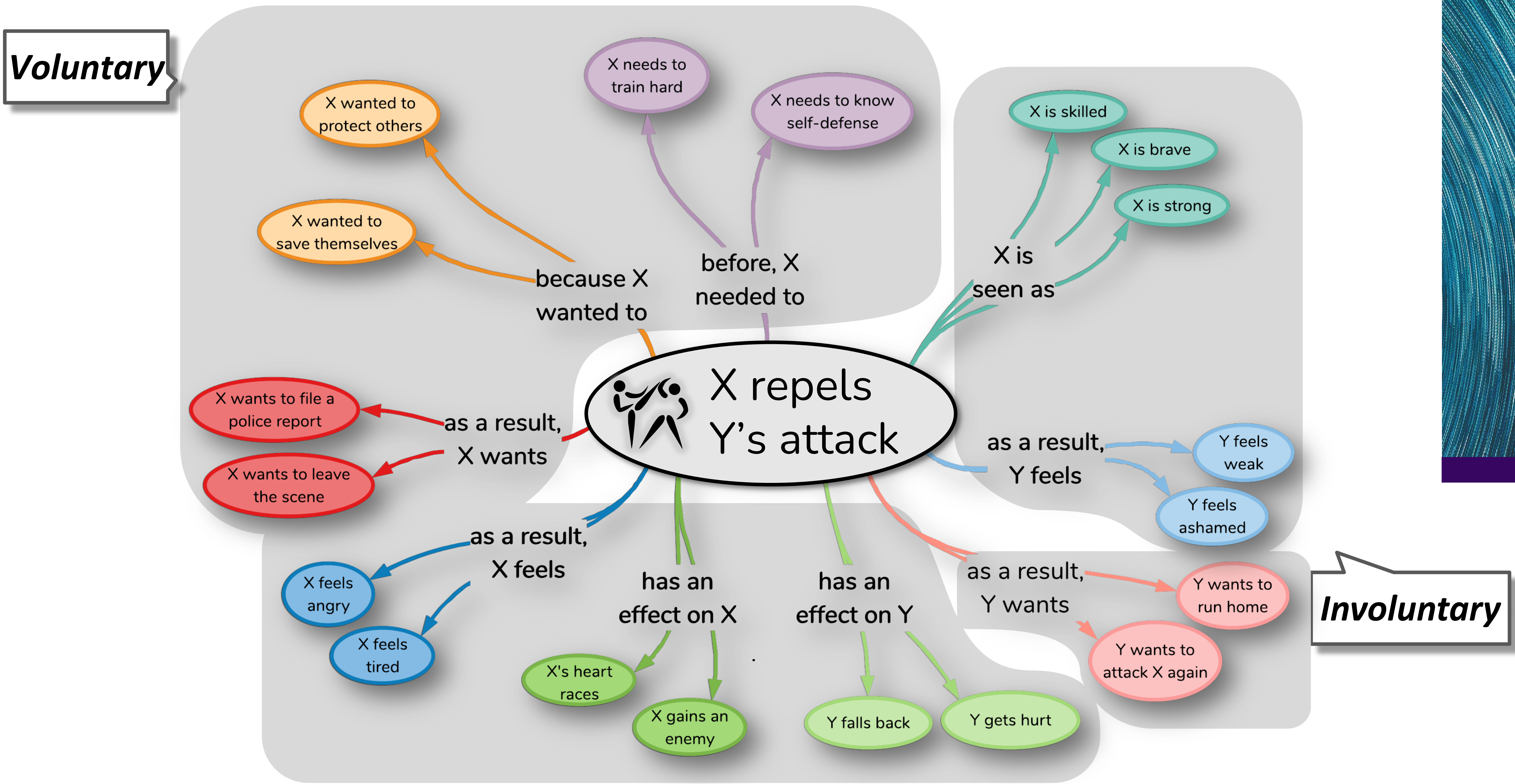


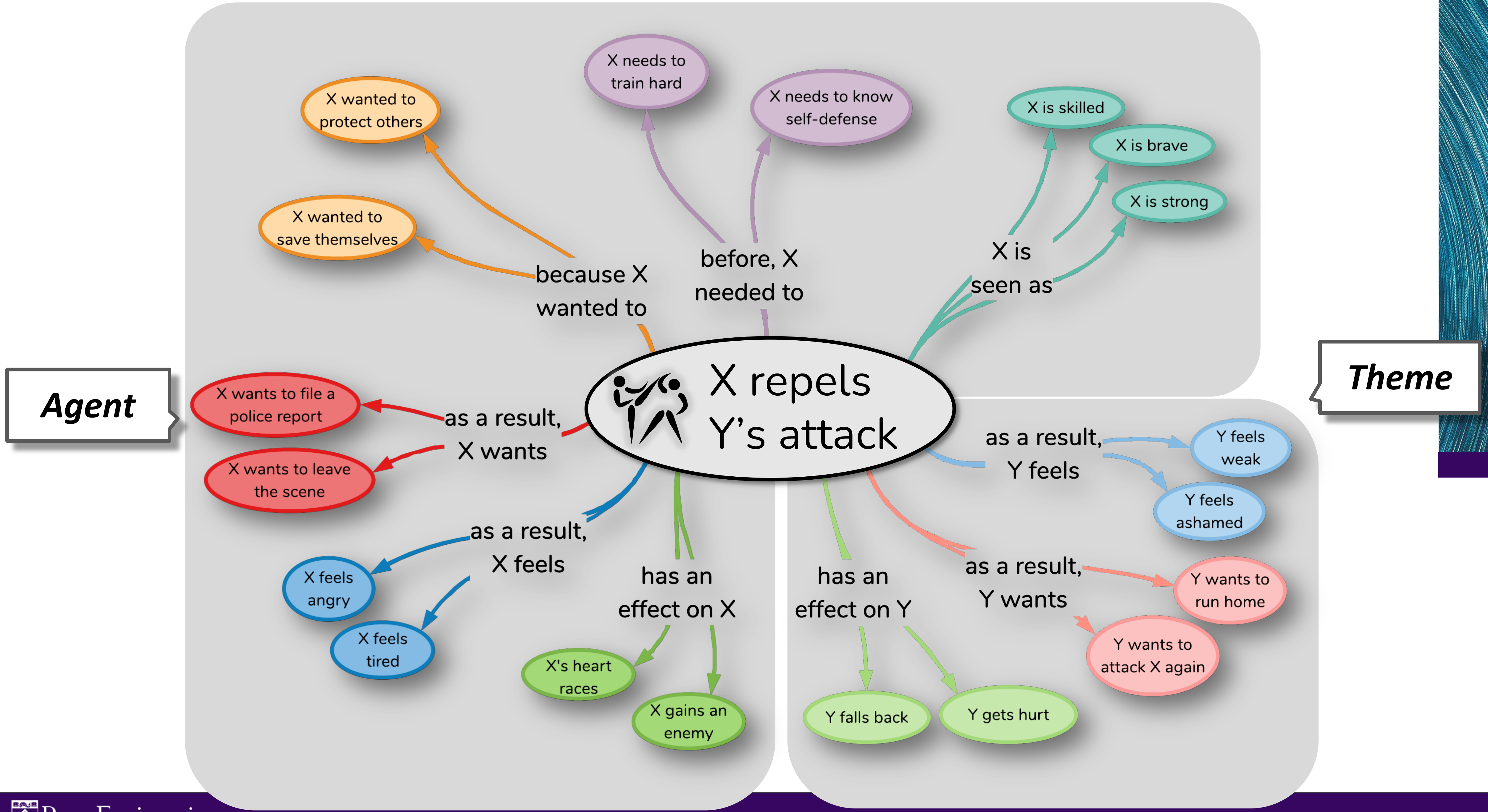


Causes
X wanted to protect others
X wanted to save themselves
because X wanted to
X needs to train hard
X needs to know self-defense
before, X needed to
X is skilled
X is brave
X is strong
X is seen as
X repels Y's attack
as a result, X wants
X wants to file a police report
X wants to leave the scene
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

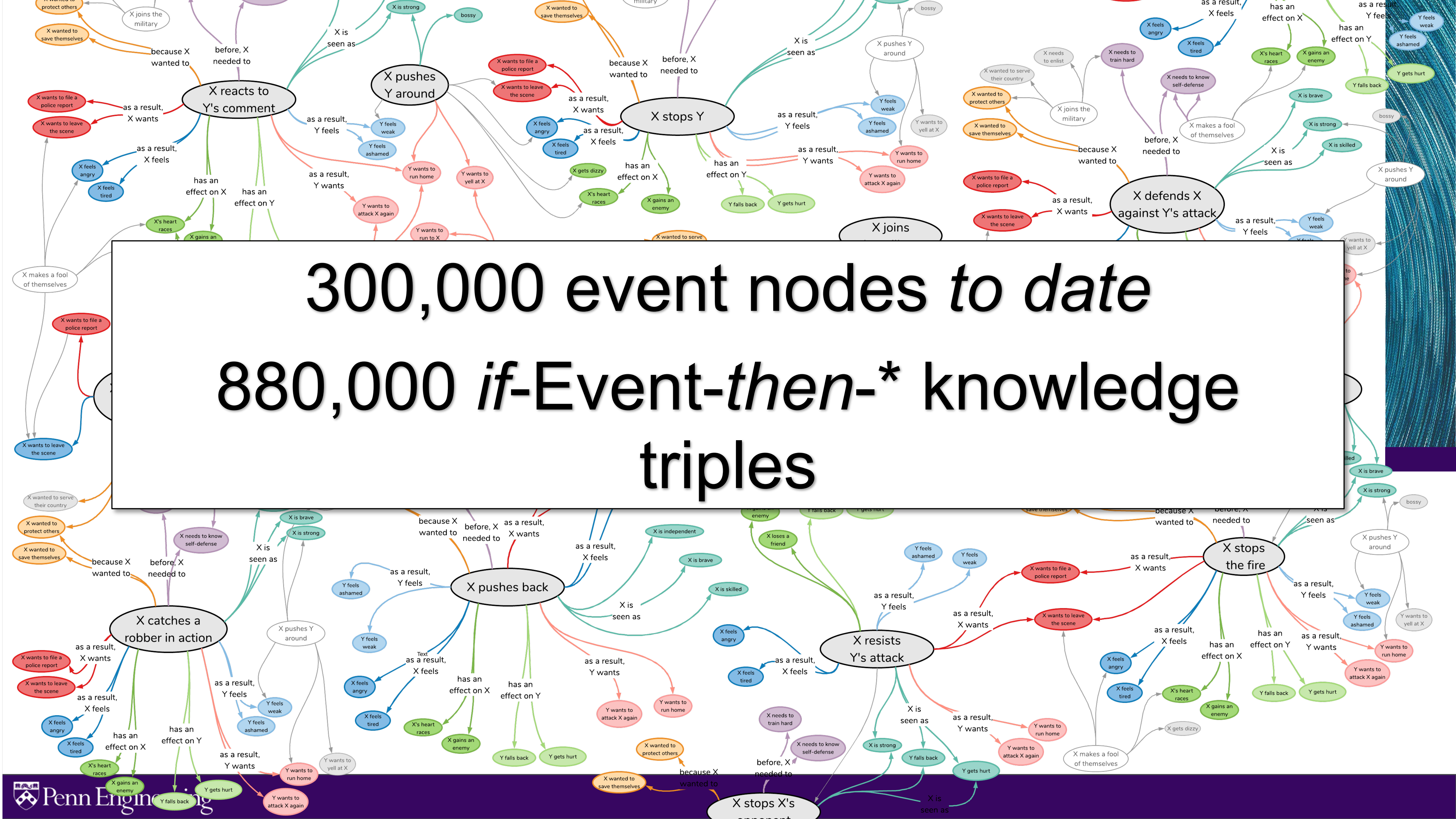


X repels Y's attack
Dynamic
Static
because X wanted to
X wanted to protect others
X wanted to save themselves
before, X needed to
X needs to train hard
X needs to know self-defense
X is seen as
X is skilled
X is brave
X is strong
as a result, X wants
X wants to file a police report
X wants to leave the scene
as a result, 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


Voluntary
X wanted to protect others
X needs to train hard
X needs to know self-defense
X is skilled
X is brave
X is strong
X wanted to save themselves
because X wanted to
before, X needed to
X is seen as
X repels Y's attack
X wants to file a police report
as a result, X wants
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
Involuntary



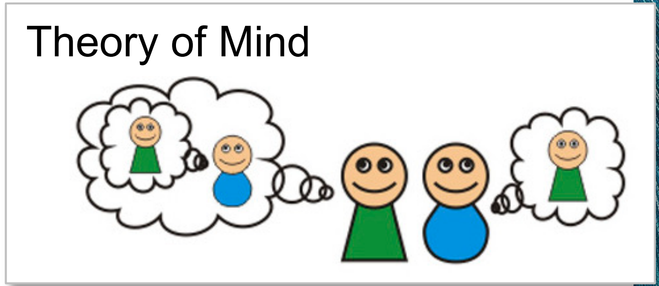
X repels Y's attack
Agent
Theme
because X wanted to
X wanted to protect others
X wanted to save themselves
before, X needed to
X needs to train hard
X needs to know self-defense
X is seen as
X is skilled
X is brave
X is strong
as a result, X wants
X wants to file a police report
X wants to leave the scene
as a result, 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



300,000 event nodes *to date*

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

ATOMIC: knowledge of *cause* and *effect*



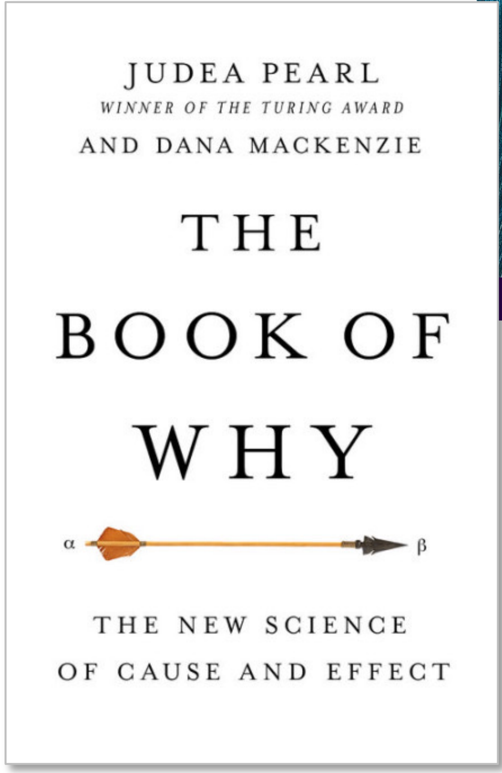
Humans have **theory of mind**, allowing us to

- make inferences about **people’s mental states**
- understand **likely events** that precede and follow
  (Moore, 2013)

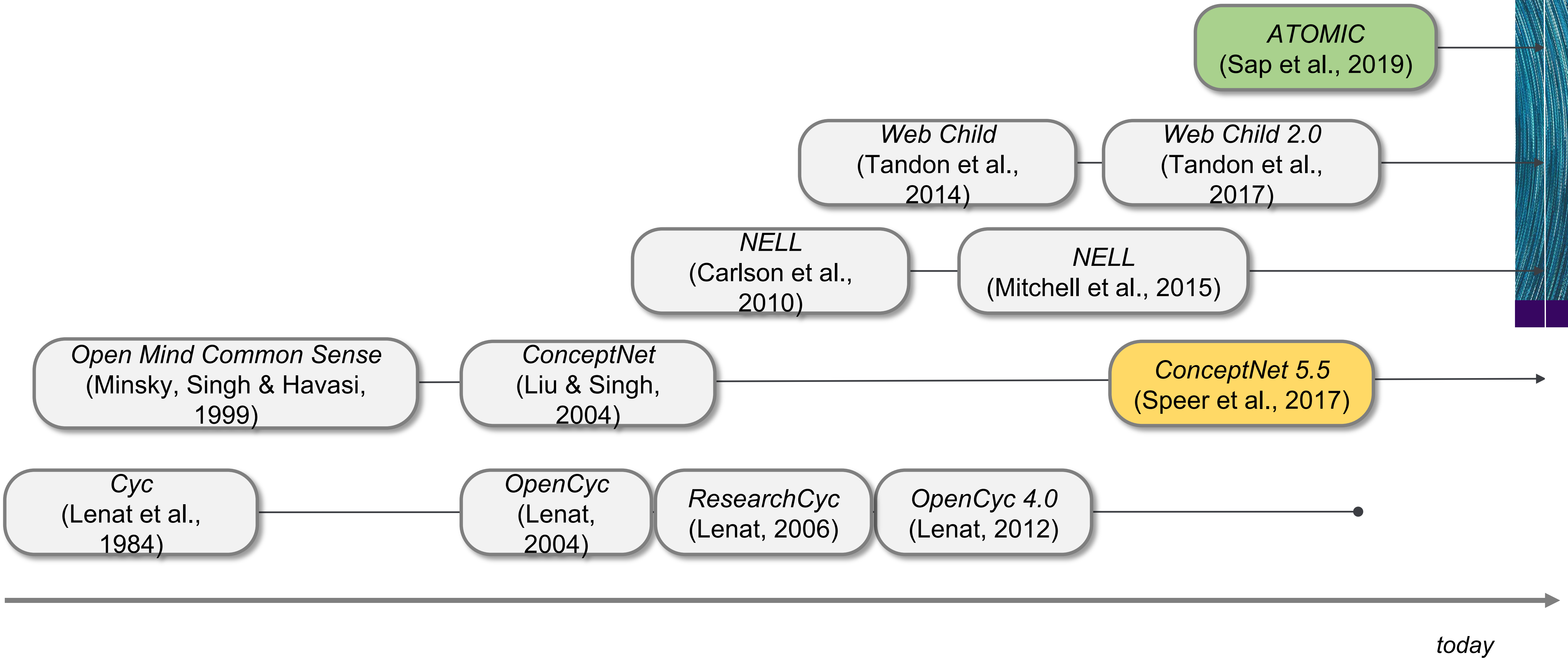
AI systems struggle with ***inferential*** reasoning

- 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)



Overview of existing resources



Existing knowledge bases

ATOMIC
(Sap et al., 2019)

NELL
(Mitchell et al., 2015)

ConceptNet 5.5
(Speer et al., 2017)

OpenCyc 4.0
(Lenat, 2012)

Existing knowledge bases

Represented in **symbolic logic**
(e.g., LISP-style logic)

Represented in **natural language**
(how humans *talk* and *think*)

NELL
(Mitchell et al., 2015)

OpenCyc 4.0
(Lenat, 2012)

ConceptNet 5.5
(Speer et al., 2017)

ATOMIC
(Sap et al., 2019)

```
(#$implies
  (#$and
    (#$isa ?OBJ ?SUBSET)
    (#$genls ?SUBSET ?SUPERSET))
  (#$isa ?OBJ ?SUPERSET))
```

Existing knowledge bases

Represented in **symbolic logic**
(e.g., LISP-style logic)

Represented in **natural language**
(how humans *talk* and *think*)

NELL
(Mitchell et al., 2015)

OpenCyc 4.0
(Lenat, 2012)

ConceptNet 5.5
(Speer et al., 2017)

Knowledge of "***what***"
(taxonomic: `A isA B`)

Knowledge of "***why***" and "***how***"
(inferential: *causes* and *effects*)

ATOMIC
(Sap et al., 2019)

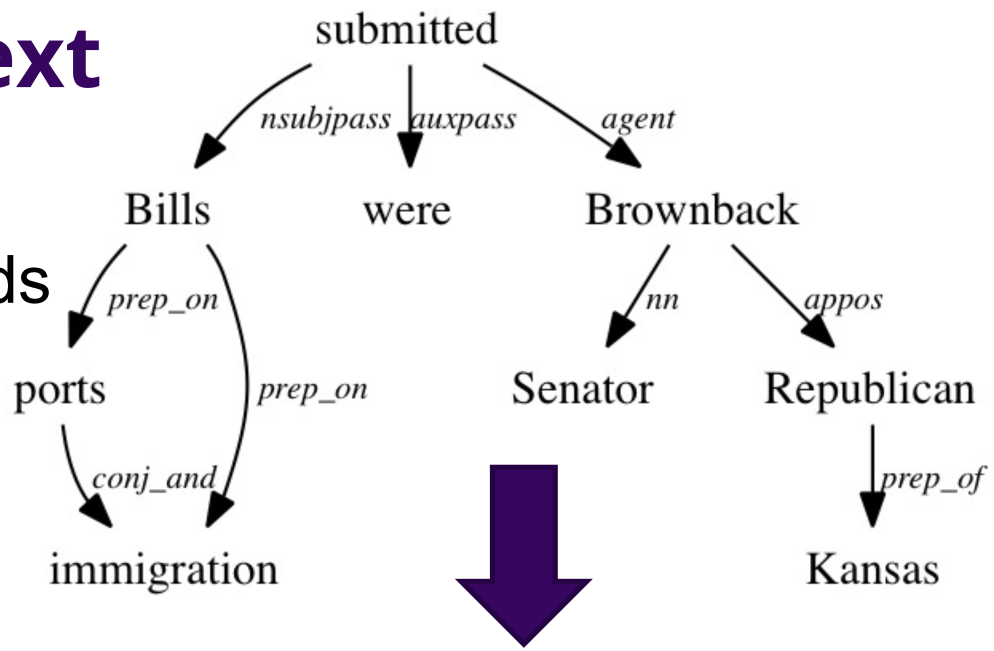
Q: How do you gather commonsense knowledge at scale?

A: It depends on the type of knowledge

Extracting commonsense from text

Based on information extraction (IE) methods

1. Read and parse text
2. Create candidate rules
3. Filter rules based on quality metric



`isA(senator,Brownback)`
`location(Kansas,Brownback)`
~~isA(senator,Kansas~~)
...

Advantage:
can extract knowledge automatically

Example system:
Never Ending Language Learner (*NELL;* Carlson et al., 2010)

Some commonsense cannot be extracted

Text is subject to **reporting bias** (Gordon & Van Durme, 2013)

- Noteworthy events
  - Murdering 4x more common than exhaling

found when extracting commonsense knowledge on four large corpora using Knext (Gordon & Van Durme, 2013)

Commonsense is not often written

- ***Grice's maxim of quantity***

When communicating, people try to be as informative as they possibly can, and give as much information as is needed, and no more.

Eliciting commonsense from humans

Experts create knowledge base

Advantages:

- Quality guaranteed
- Can use complex representations (e.g., CycL, LISP)

Drawbacks:

- Time cost
- Training users

OpenCyc 4.0 (Lenat, 2012)

WordNet (Miller et al., 1990)

Non-experts write knowledge in natural language phrases

Natural language

- Accessible to non-experts
- Different phrasings allow for more nuanced knowledge

Fast and scalable collection

- Crowdsourcing
- Games with a purpose

ATOMIC (Sap et al., 2019)

ConceptNet 5.5 (Speer et al., 2017)

Knowledge bases and mitigating biases

PersonX clutches a gun → because X wanted to

- to be safe
- to protect himself
- to protect themselves
- to defend themselves
- to defend himself

ATOMIC (Sap et al., 2019)

Jaquain clutches a gun → because X wanted to

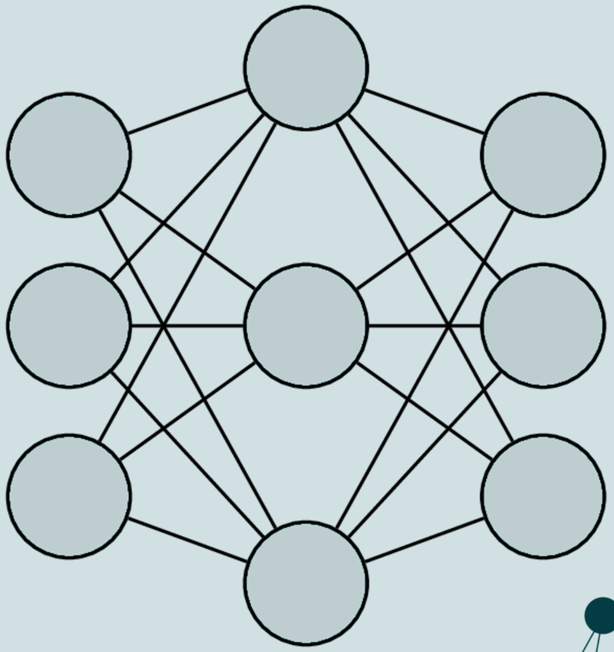
- to kill someone
- none
- to protect himself
- to be safe
- to protect themselves

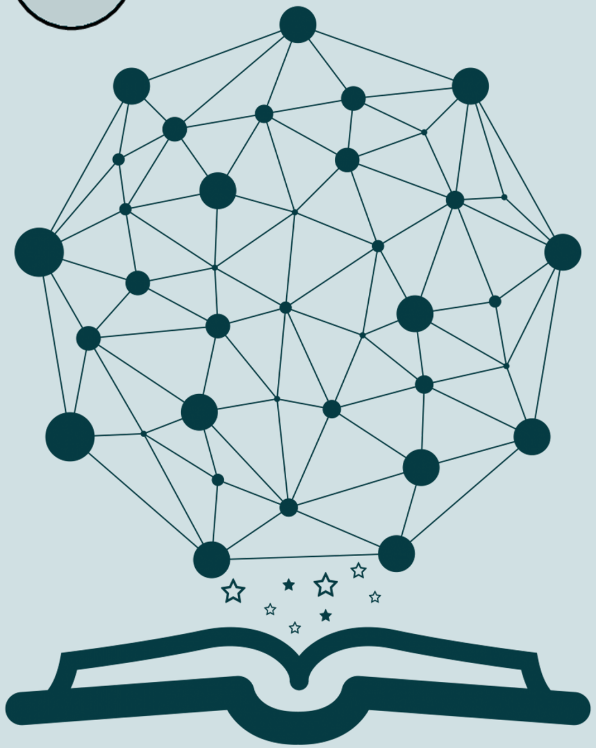
Karen clutches a gun → because X wanted to

- to be safe
- to protect himself
- to shoot
- to get the gun
- none

COMET (Bosselut et al., 2019): ATOMIC + OpenAI GPT

Neural and Symbolic Models of Commonsense Reasoning





Katrina had the financial means to afford a new car while Monica did not, since _____ had a high paying job.

WINOGRANDE: An Adversarial Winograd Schema Challenge at Scale.
Keisuke Sakaguchi, Ronan Le Bras, Chandra Bhagavatula, and Yejin Choi. AAAI 2020.

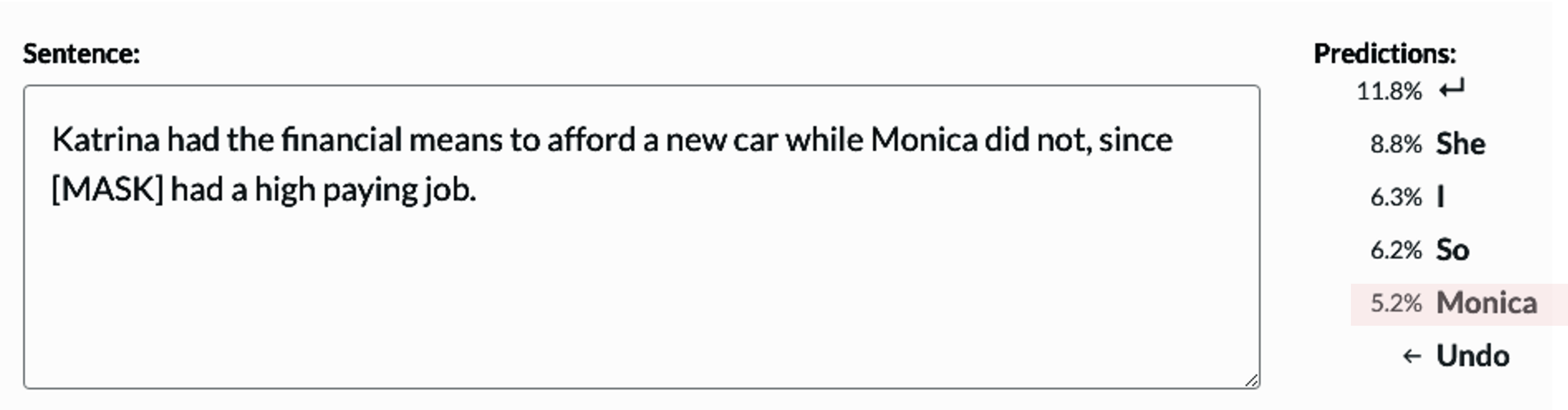
Neural Architecture

[CLS] Katrina had the financial means to afford a new car while Monica did not, since [SEP] Katrina had a high paying job.

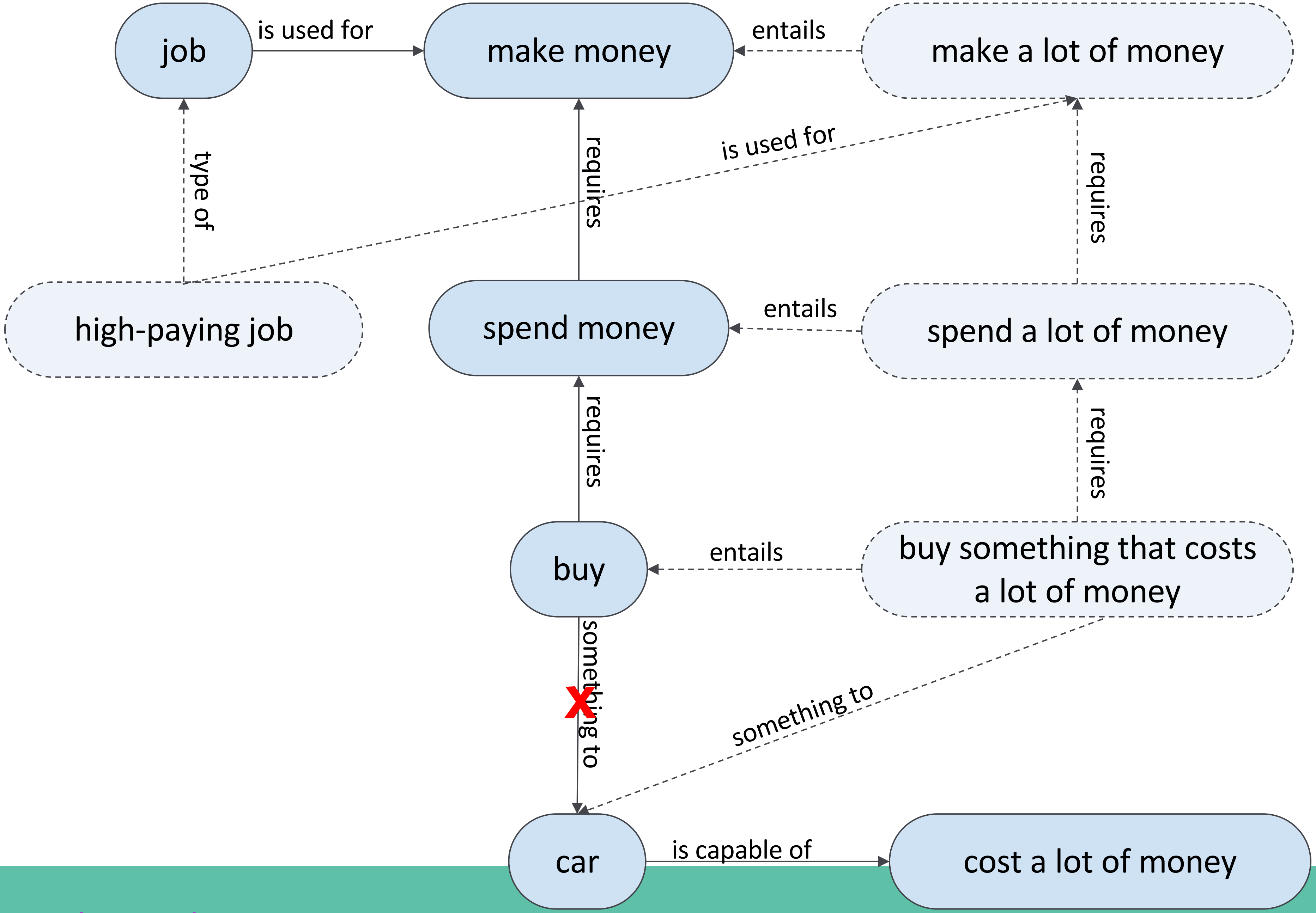
[CLS] Katrina had the financial means to afford a new car while Monica did not, since [SEP] Monica had a high paying job.

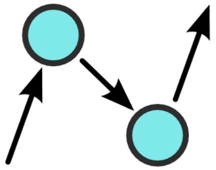
0.51
0.49

Masked Language Models



https://demo.allennlp.org/masked-lm

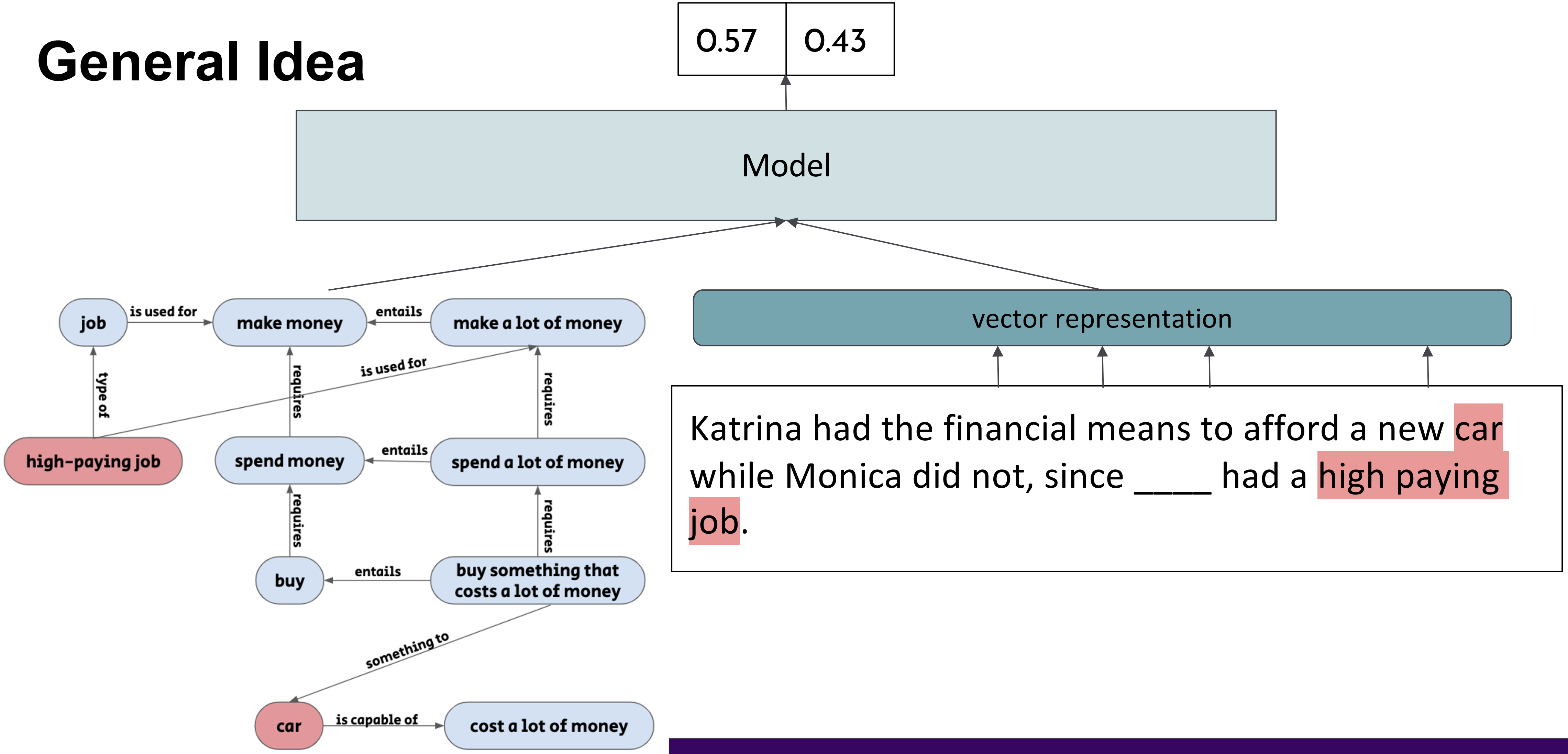




http://conceptnet5.media.mit.edu/

Incorporating External Knowledge into Neural Models

General Idea

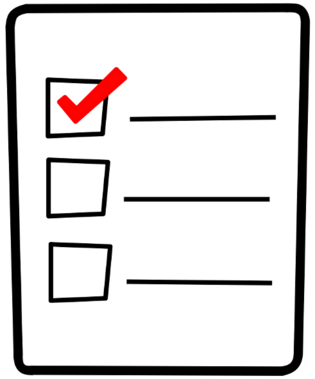


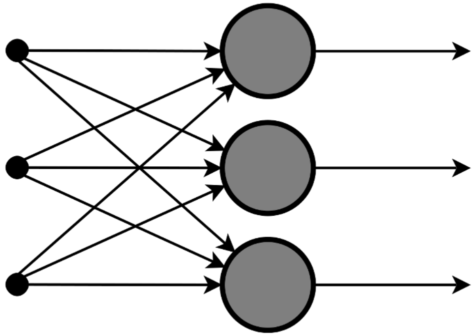
Incorporating External Knowledge into Neural Models

Recipe

Task

Story ending,
Machine Comprehension
Social common sense
NLI



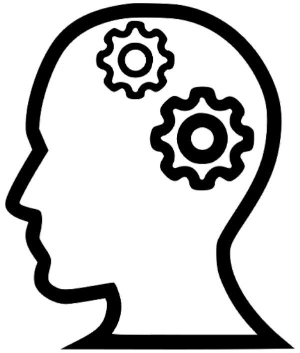


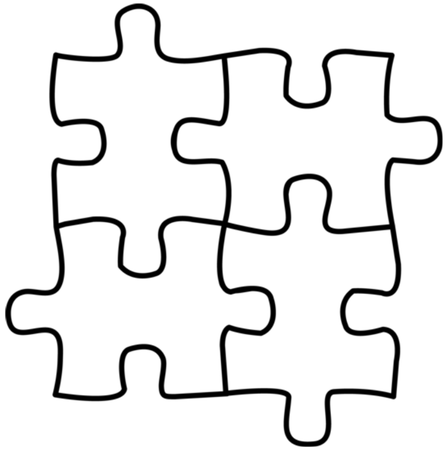
Neural Component

Pre/post pre-trained
language models

Knowledge Source

Knowledge bases,
extracted from text,
hand-crafted rules





Combination Method

Attention, pruning,
word embeddings,
multi-task learning

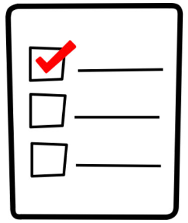
Story Ending Task (RocStories)

Agatha had always wanted pet birds.
So one day she purchased two pet finches.
Soon she couldn't stand their constant noise.
And even worse was their constant mess.

↓

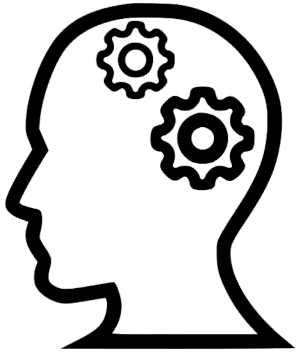
Agatha decided to buy two more. (Wrong)
Agatha decided to return them. (Right)

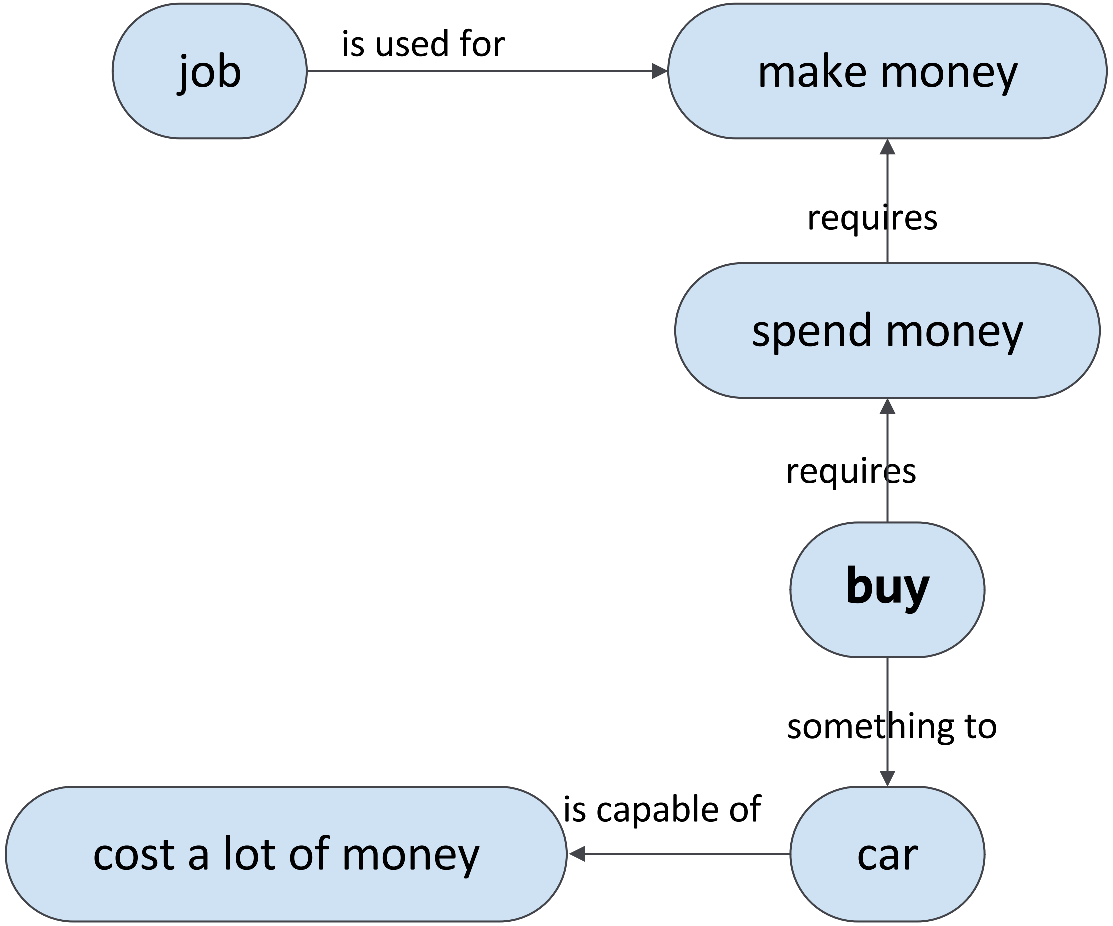
Task



A Corpus and Cloze Evaluation for Deeper Understanding of Commonsense Stories. *Nasrin Mostafazadeh, Nathanael Chambers, Xiaodong He, Devi Parikh, Dhruv Batra, Lucy Vanderwende, Pushmeet Kohli, and James Allen.* NAACL 2016.

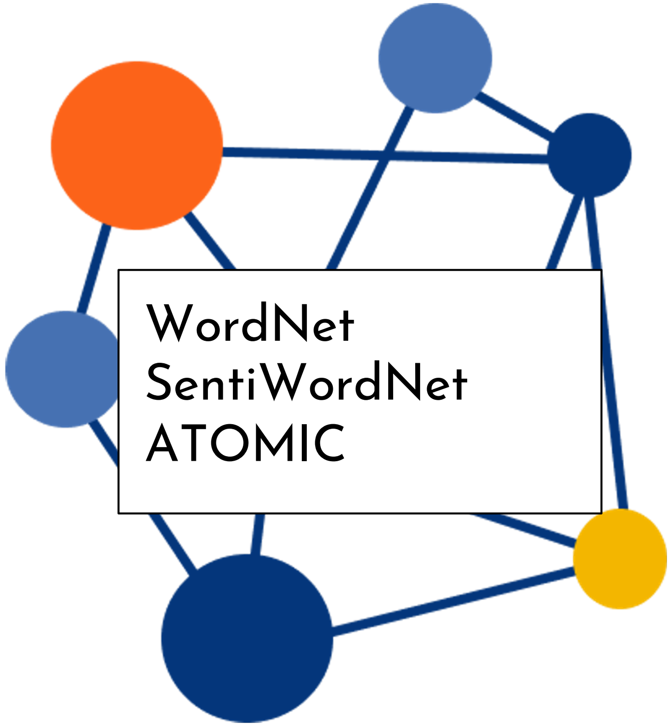
ConceptNet





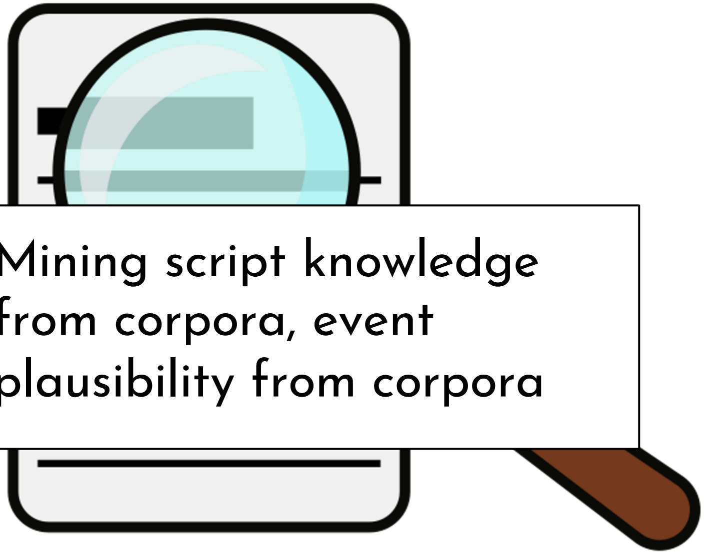
Conceptnet 5.5: An open multilingual graph of general knowledge. *Robyn Speer, Joshua Chin, and Catherine Havasi.* AAAI 2017.

Other Knowledge Sources



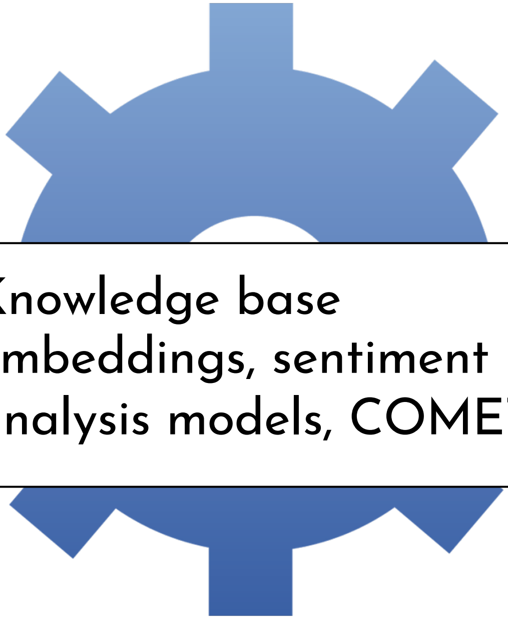
WordNet
SentiWordNet
ATOMIC

Mining from Text



Mining script knowledge from corpora, event plausibility from corpora

Knowledge Bases



Knowledge base embeddings, sentiment analysis models, COMET

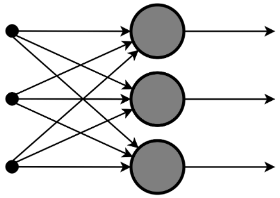
Tools

Neural Component

[CLS] Katrina had the financial means to afford a new car while Monica did not, since [SEP] Katrina had a high paying job.

[CLS] Katrina had the financial means to afford a new car while Monica did not, since [SEP] Monica had a high paying job.

0.51
0.49

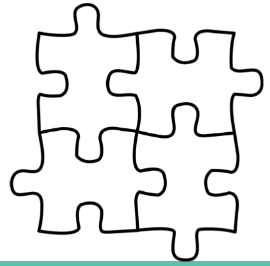


Combination Method

1. Incorporate into scoring function

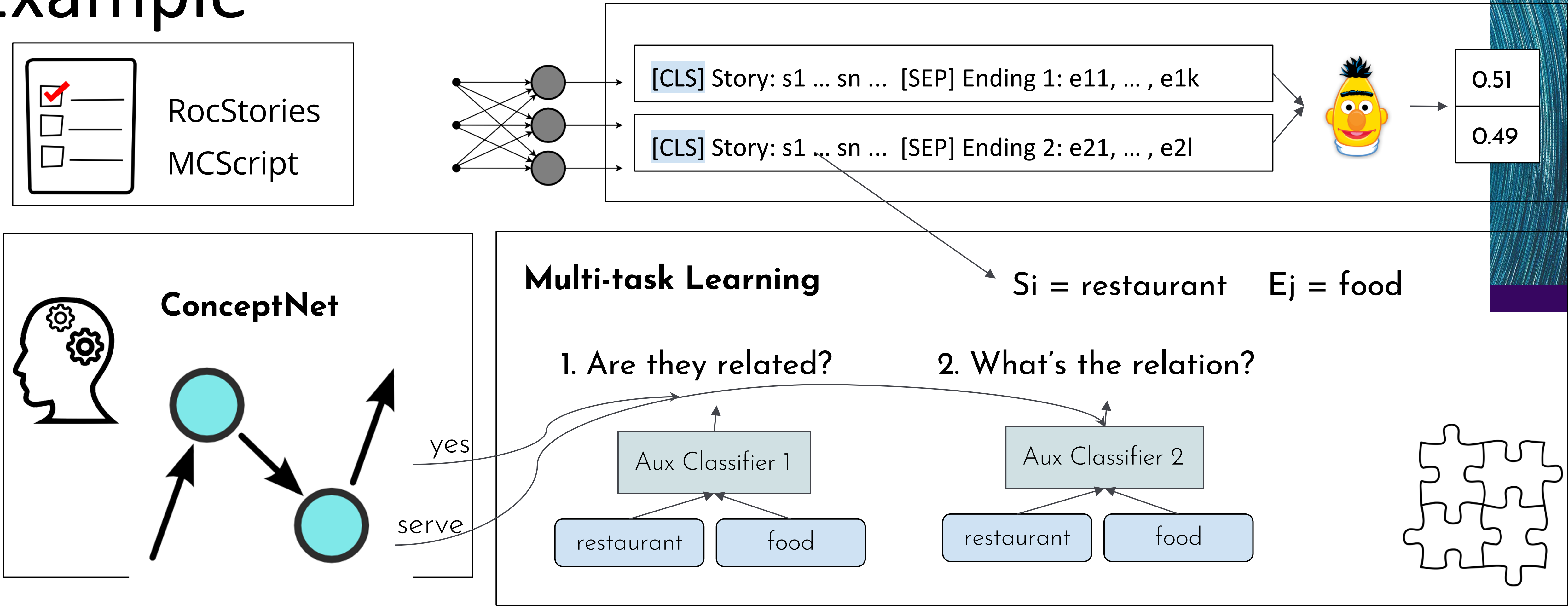
2. Symbolic → vector representation
   - (+attention)

3. Multi-task learning



Incorporating External Knowledge into Neural Models

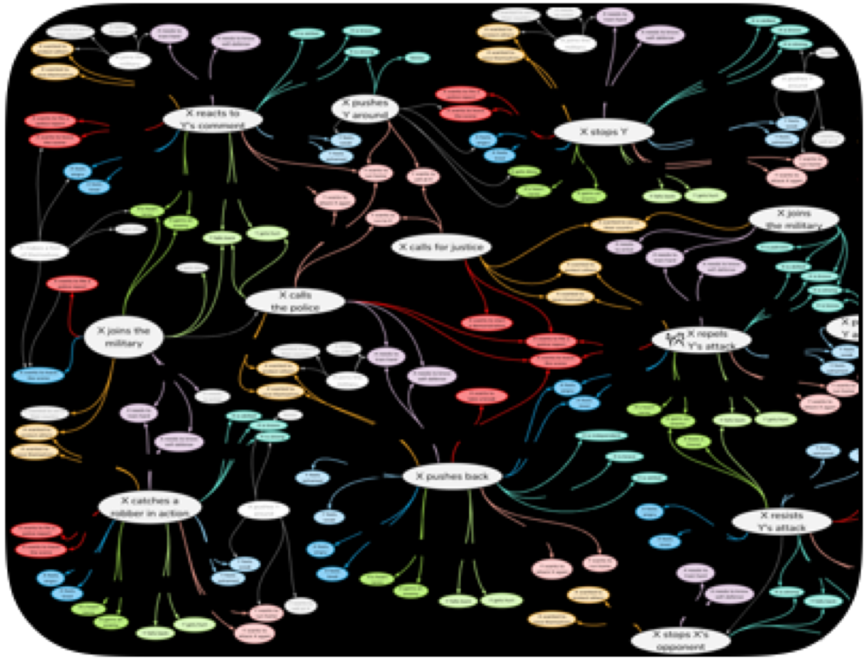
Example



Incorporating Commonsense Reading Comprehension with Multi-task Learning. *Jiangnan Xia, Chen Wu, and Ming Yan.* CIKM 2019.

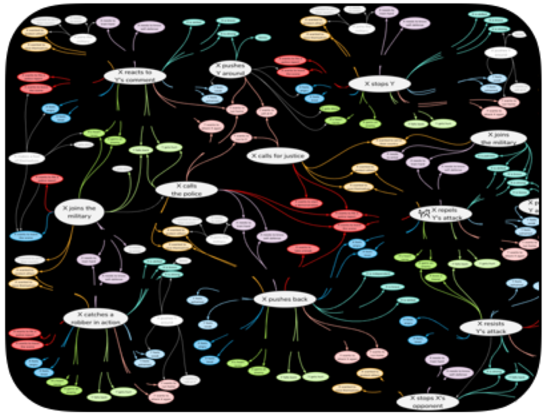
Limitations of Knowledge Graphs

- Insufficient Coverage
- Not 100% Accurate
- Limited expressivity



Limitations of Knowledge Graphs

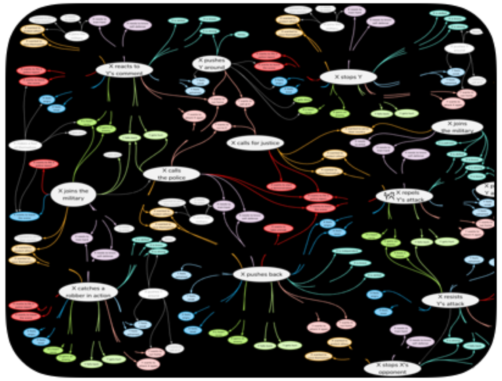
Kai knew that things were getting out of control and managed to keep his temper in check



Limitations of Knowledge Graphs

- Situations rarely found **as-is** in commonsense knowledge graphs

ATOMIC



(Sap et al., 2019)

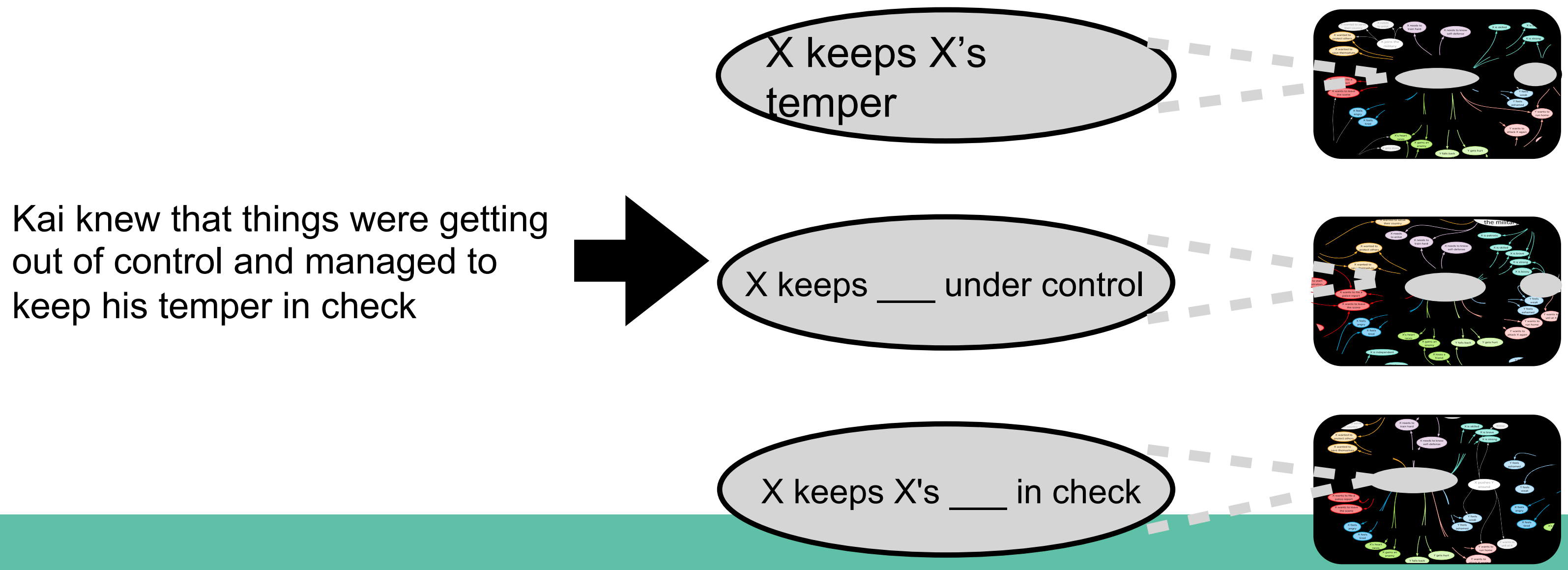
(X goes to the mall,
Effect on X, buys clothes)

(X goes the mall,
Perception of X, rich)

(X gives Y some money,
Reaction of Y, grateful)

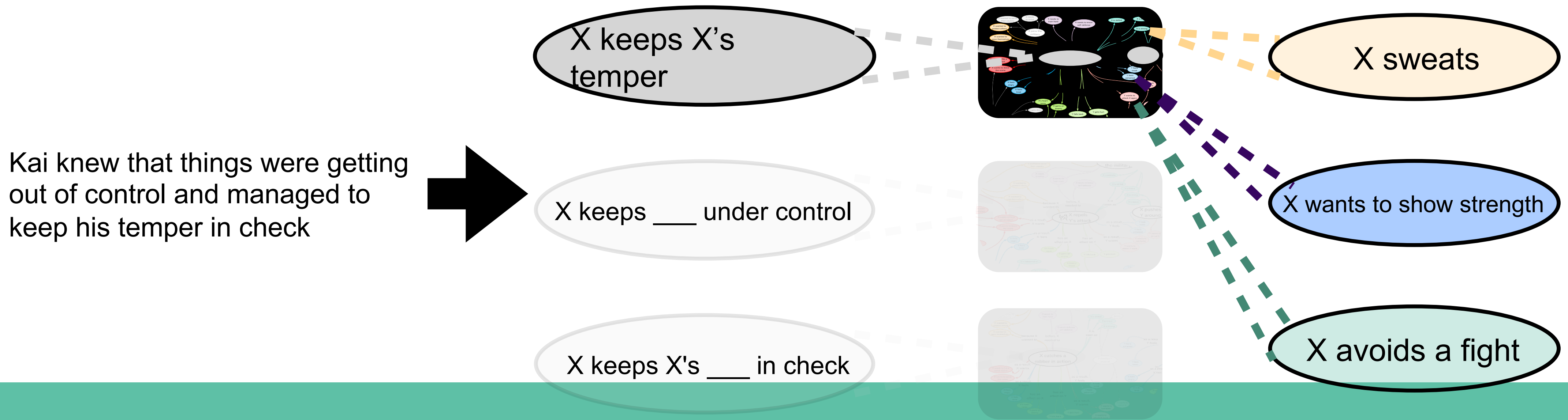
Limitations of Knowledge Graphs

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



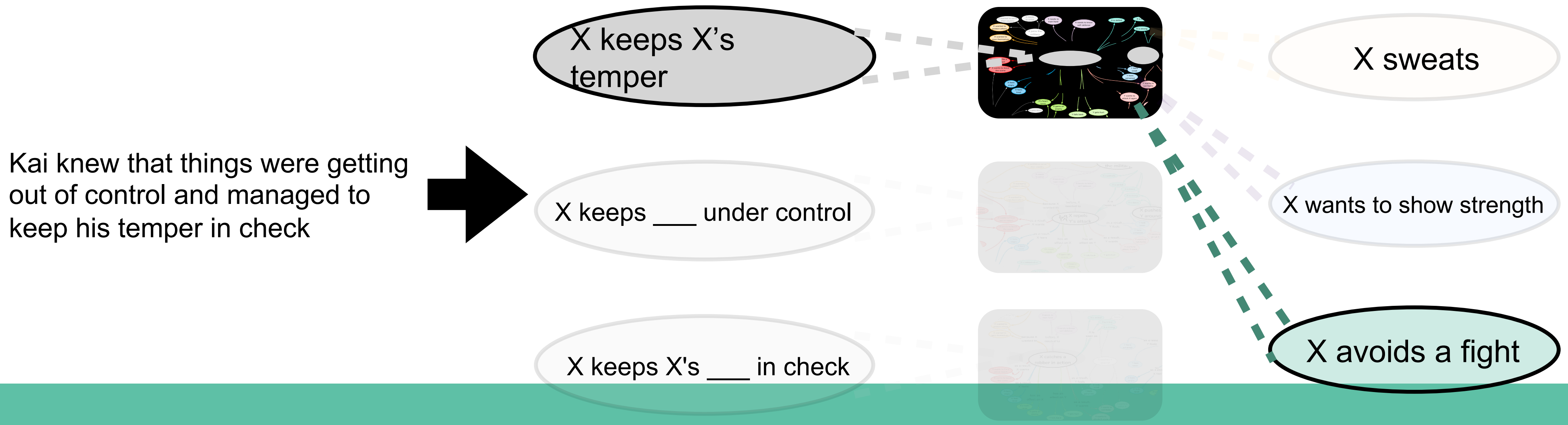
Limitations of Knowledge Graphs

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



Limitations of Knowledge Graphs

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



Challenge

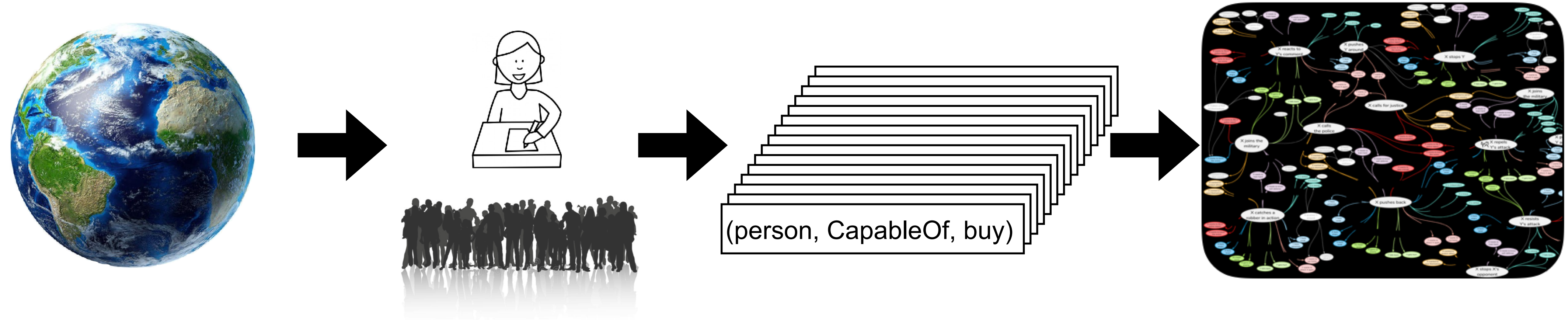
How do we provide machines with
large-scale commonsense knowledge?

Constructing Knowledge Graphs

Observe world

Write commonsense knowledge facts

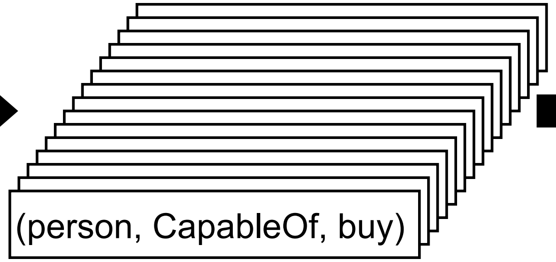
Store facts in knowledge graph



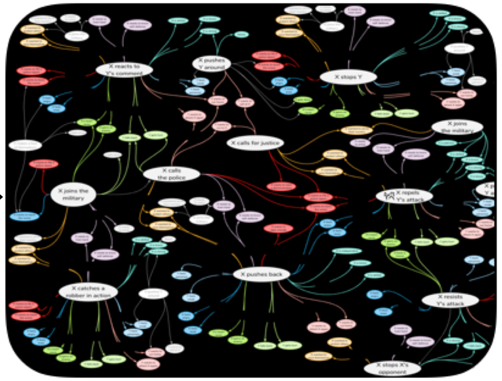
Constructing Symbolic Knowledge Graphs

Observe world

Write commonsense knowledge facts



Store facts in knowledge graph



(Miller, 1995)

(Singh et al., 2002)

(Lenat, 1995)

(Sap et al., 2019)

Challenges of Prior Approaches

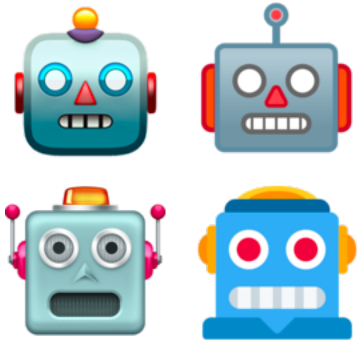
- Commonsense knowledge is **immeasurably vast**, making it **impossible to manually enumerate**

Constructing Knowledge Graphs Automatically

Gather Textual Corpus

John went to the grocery **store** to buy some **steaks**. He was going to prepare **dinner** for his **daughter**’s **birthday**. She was turning 5 and would be starting elementary **school** soon.

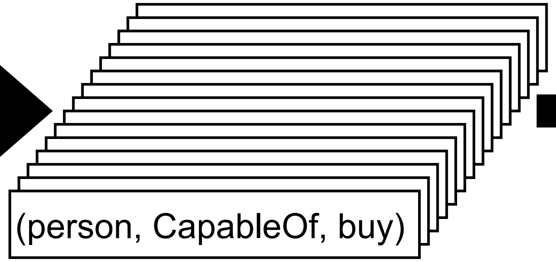
Automatically extract knowledge



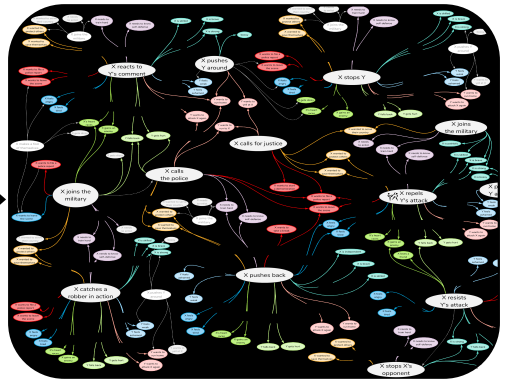
(Schubert, 2002)

(Banko et al., 2007)

(Zhang et al., 2020)



Store in knowledge graph



ConceptNet

An open, multilingual knowledge graph

(Speer et al., 2017)

Webchild

(Tandon et al., 2019)

Encyclopedic vs. Commonsense Knowledge

Encyclopedic Knowledge	Commonsense Knowledge
Explicitly written in text	Often assumed Grice's Maxim of Quantity
Ontological Mentions	
Deviations rarely written	

Encyclopedic vs. Commonsense Knowledge

Encyclopedic Knowledge	Commonsense Knowledge
Explicitly written in text	Often assumed Grice's Maxim of Quantity
Ontological Mentions	Complex Mentions e.g., Causal If-Then Knowledge
Deviations rarely written	

Encyclopedic vs. Commonsense Knowledge

Encyclopedic Knowledge	Commonsense Knowledge
Explicitly written in text	Often assumed Grice's Maxim of Quantity
Ontological Mentions	Complex Mentions e.g., Causal If-Then Knowledge
Deviations rarely written	Reporting Bias murders 4x more common than breating

Challenges of Prior Approaches

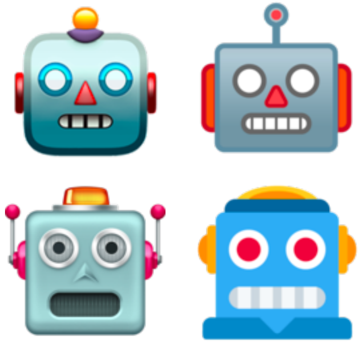
- Commonsense knowledge is **immeasurably vast**, making it **impossible to manually enumerate**
- Commonsense knowledge is often implicit, and often **can’t be directly extracted from text**

Constructing Knowledge Graphs Automatically

Gather Textual Corpus

John went to the grocery **store** to buy some **steaks**. He was going to prepare **dinner** for his **daughter**'s **birthday**. She was turning 5 and would be starting elementary **school** soon.

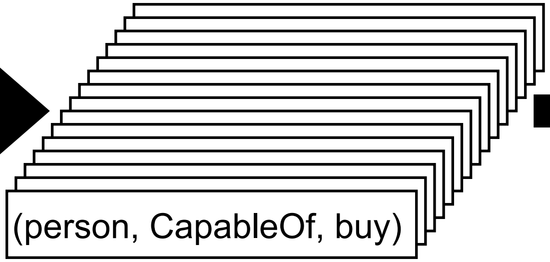
Automatically extract knowledge



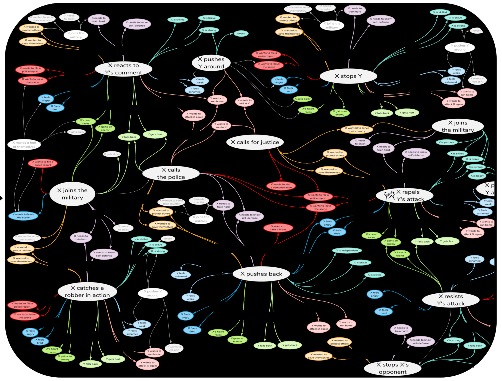
(Schubert, 2002)

(Banko et al., 2007)

(Zhang et al., 2020)



Store in knowledge graph



ConceptNet

An open, multilingual knowledge graph

(Speer et al., 2017)

Webchild

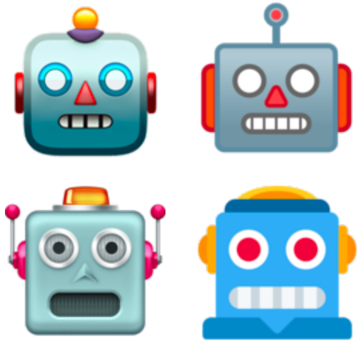
(Tandon et al., 2019)

Knowledge Base Completion

Gather training set of knowledge tuples

(person, CapableOf, buy)

Learn relationships among entities

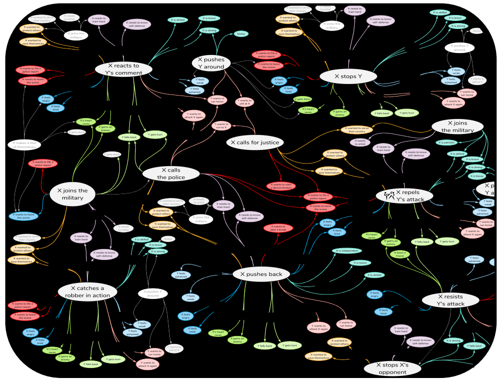


(Socher et al., 2013)
(Bordes et al., 2013)
(Riedel et al., 2013)
(Toutanova et al., 2015)
(Yang et al., 2015)
(Trouillon et al., 2016)
(Nguyen et al., 2016)
(Dettmers et al., 2018)

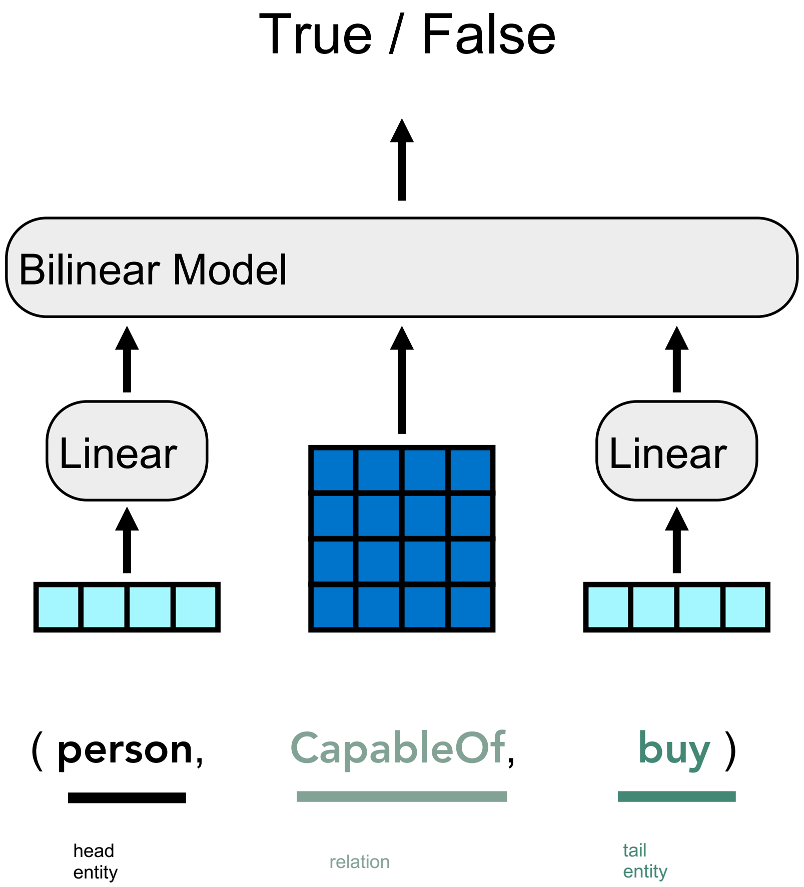
Predict new relationships

(person, CapableOf, ?)

Store in knowledge graph



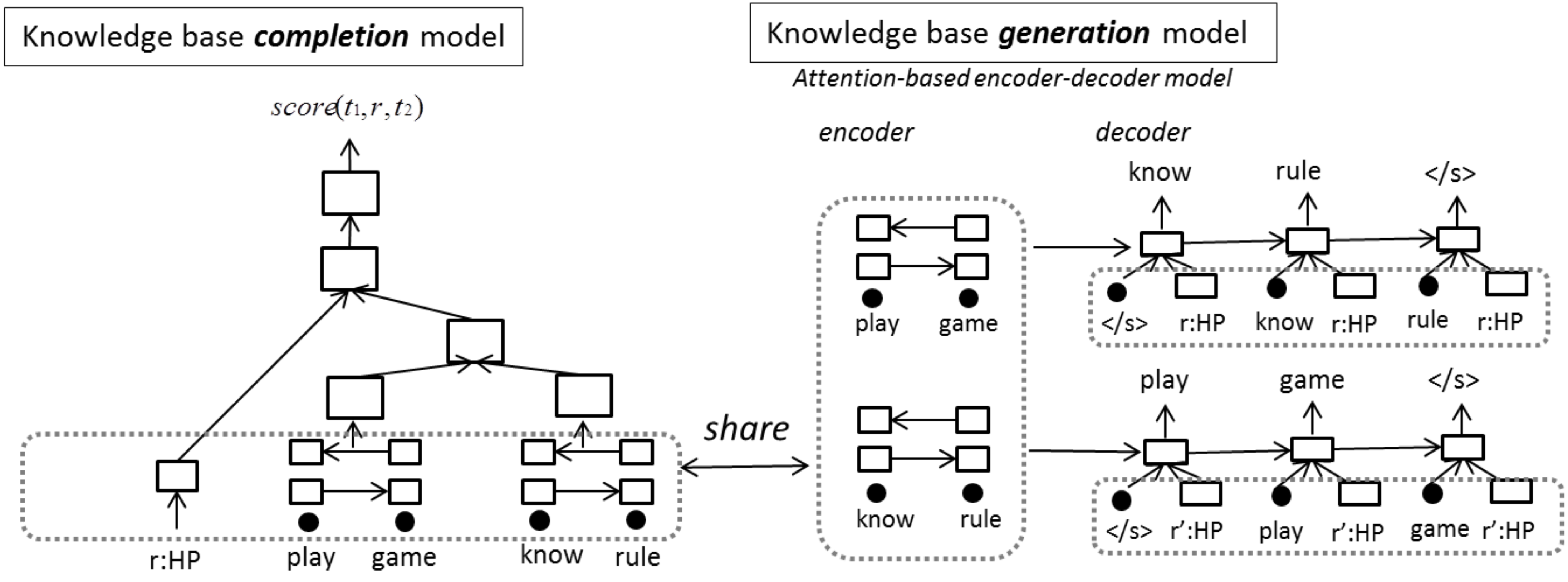
Commonsense Knowledge Base Completion



Only high confidence predictions are validated

Low Novelty

Commonsense Knowledge Base Completion and Generation!



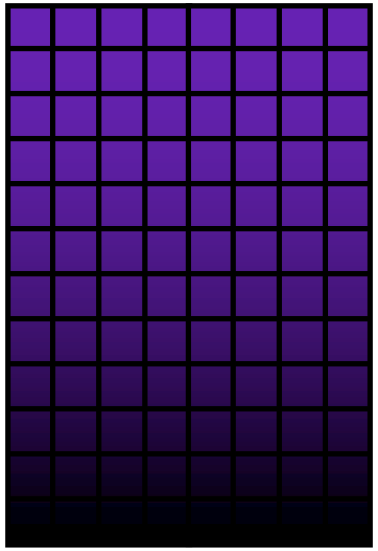
Saito et al., 2018

Challenges of Prior Approaches

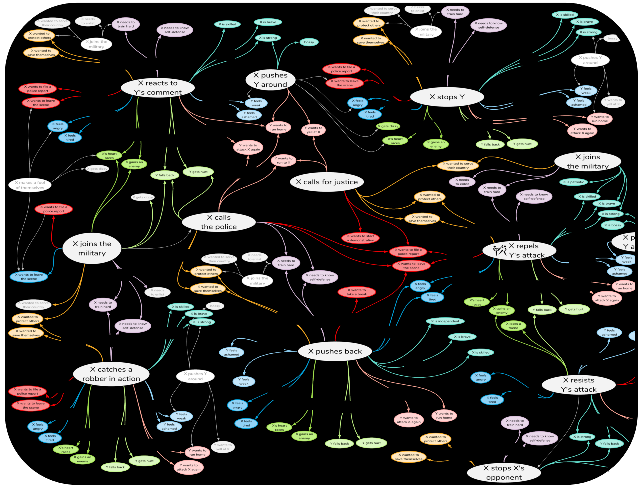
- Commonsense knowledge is **immeasurably vast**, making it **impossible to manually enumerate**
- Commonsense knowledge is often implicit, and often **can't be directly extracted from text**
- Commonsense knowledge resources are quite sparse, making them **difficult to extend by only learning from examples**

Solution Outline

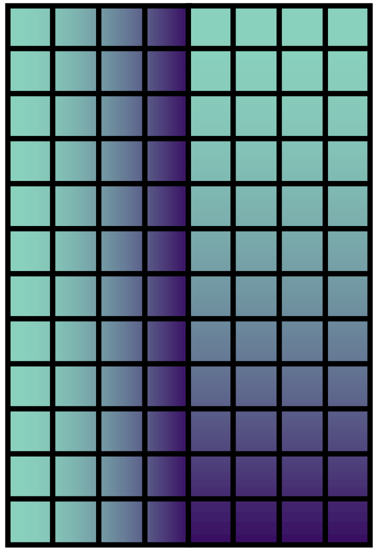
- Leverage manually curated commonsense knowledge resources
- Learn from the examples to induce new relationships
- Scale up using language resources



Learn word embeddings from language corpus



Retrofit word embeddings on semantic resource



Learn knowledge-aware embeddings

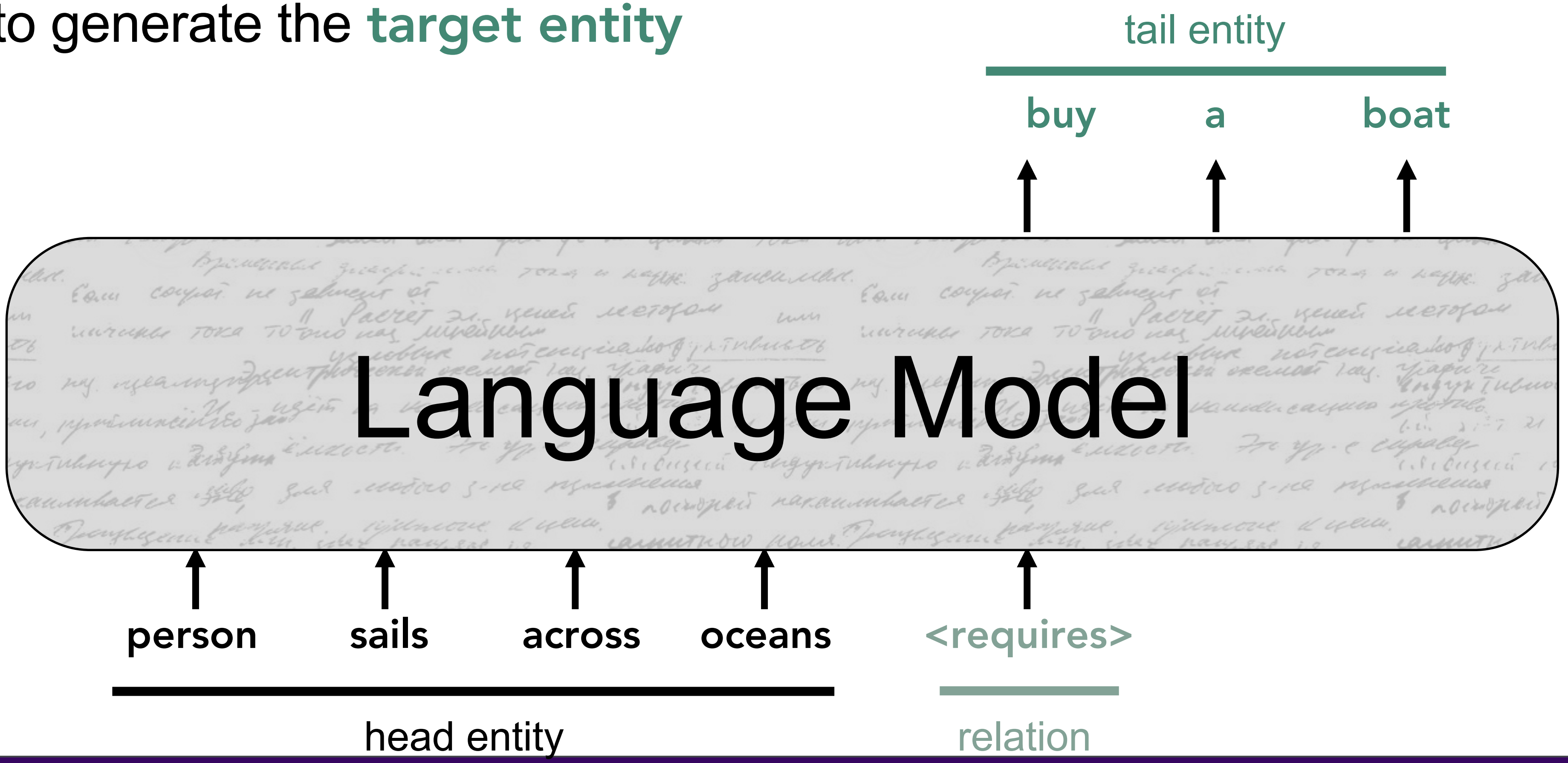
Faruqui et al., 2015, Speer et al., 2017

Structure of Knowledge Tuple

Learning Structure of Knowledge

Given a **seed entity** and a **relation**,
learn to generate the **target entity**

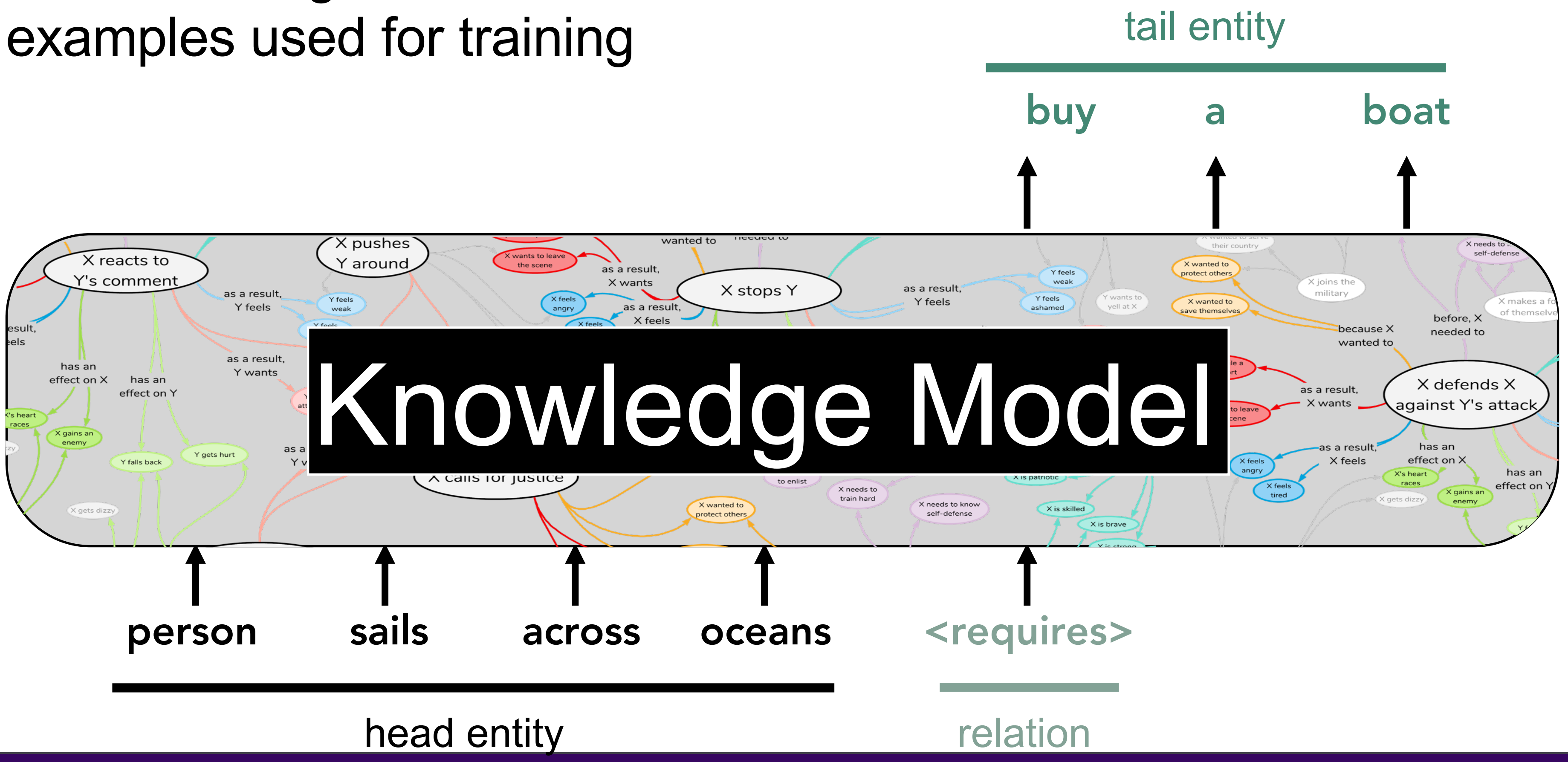
$$\mathcal{L} = -\sum \log P(\textit{target words} | \textit{seed words, relation})$$

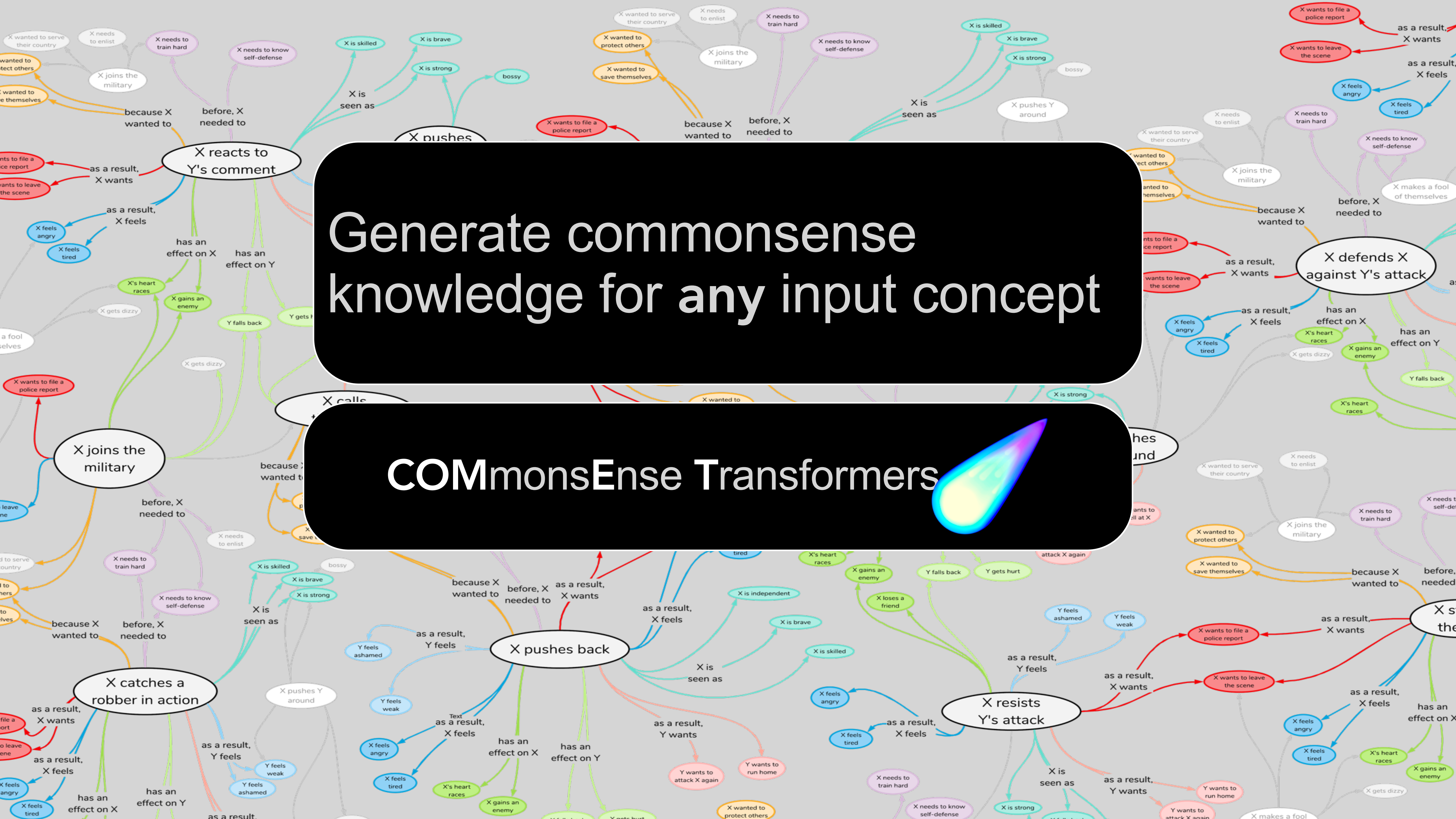


(Bosselut et al., 2019)

Learning Structure of Knowledge

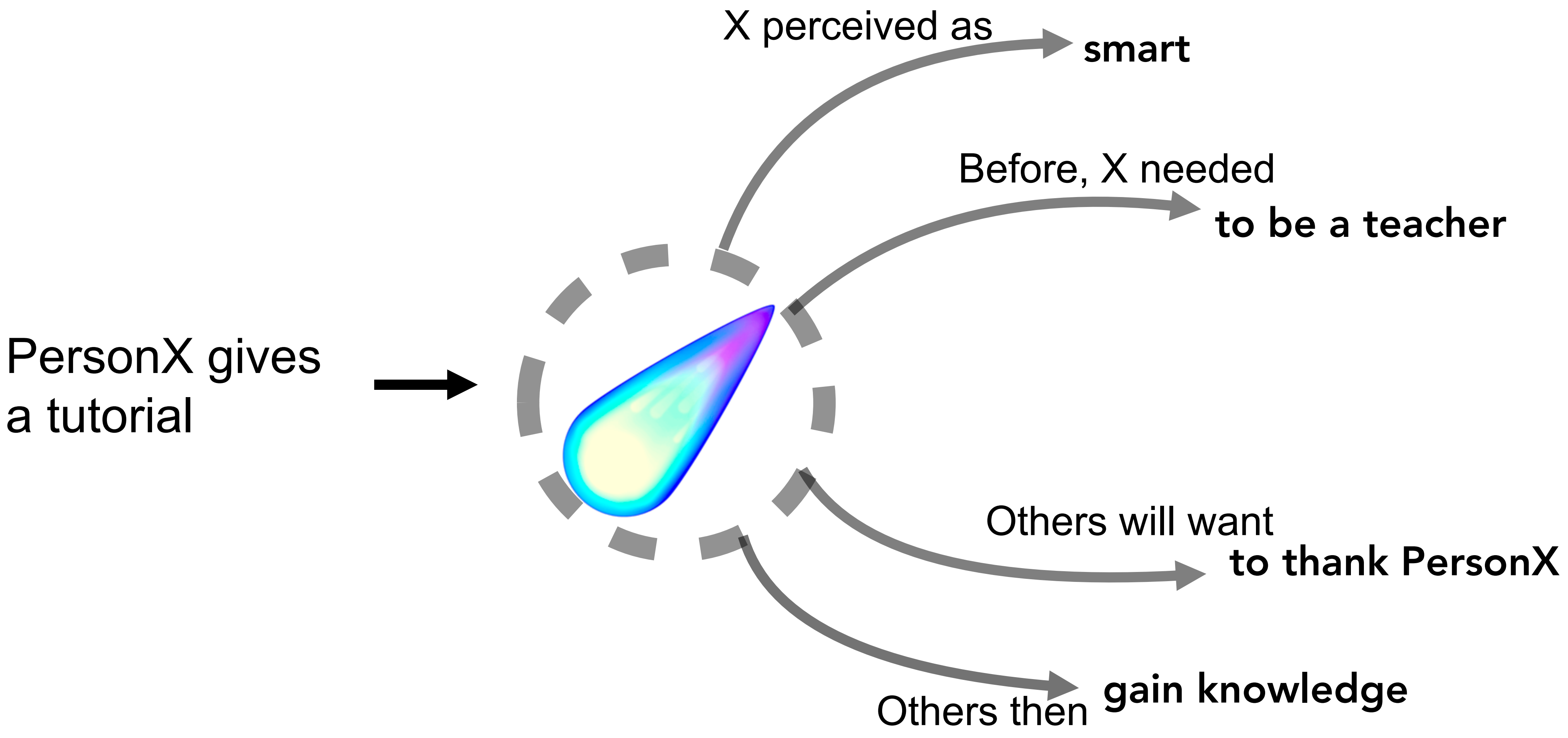
Language Model ⟶ **Knowledge Model:** generates knowledge of the structure of the examples used for training



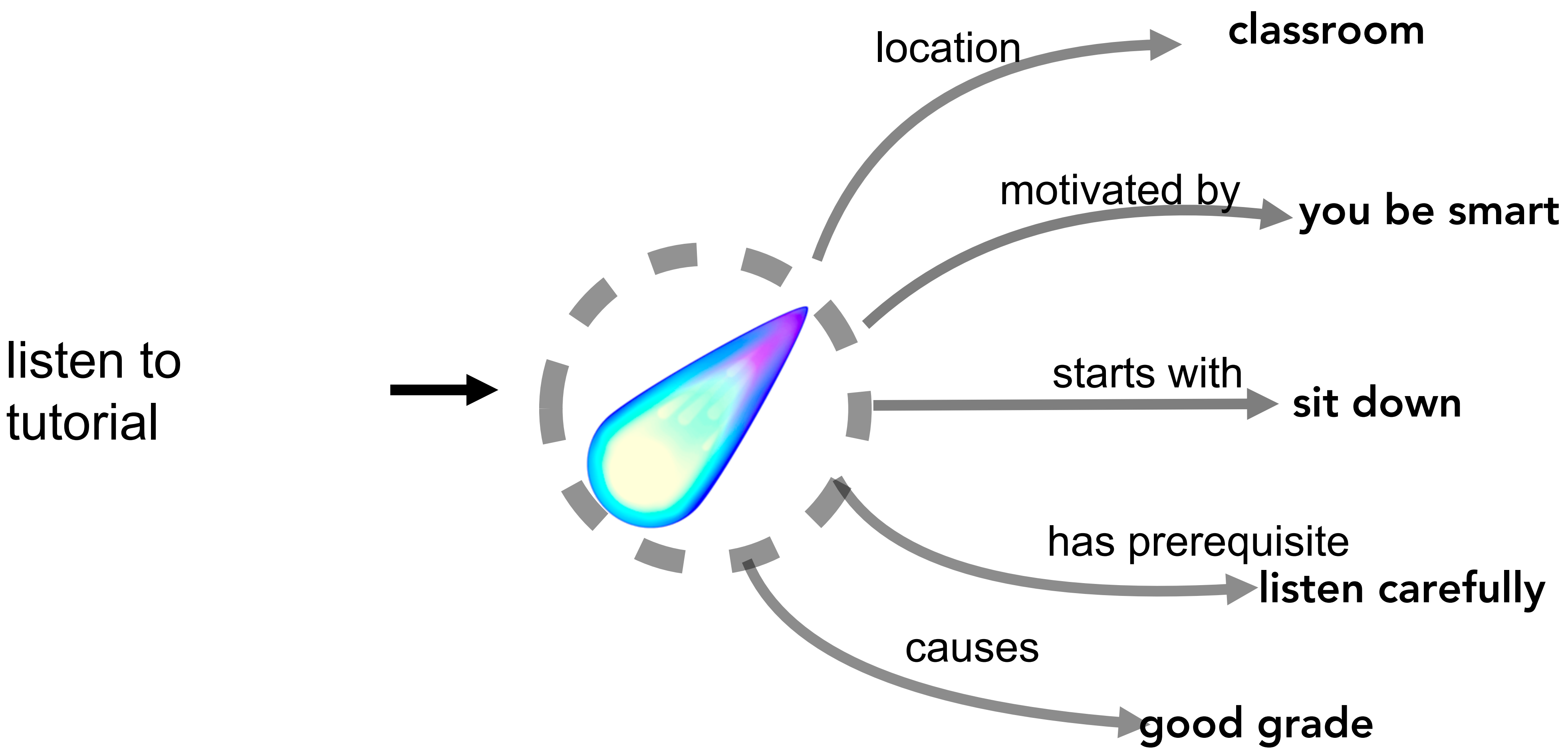


Generate commonsense knowledge for any input concept
COMmonsEnse Transformers
X reacts to Y's comment
X joins the military
X catches a robber in action
X pushes back
X resists Y's attack
X defends X against Y's attack

COMET - ATOMIC



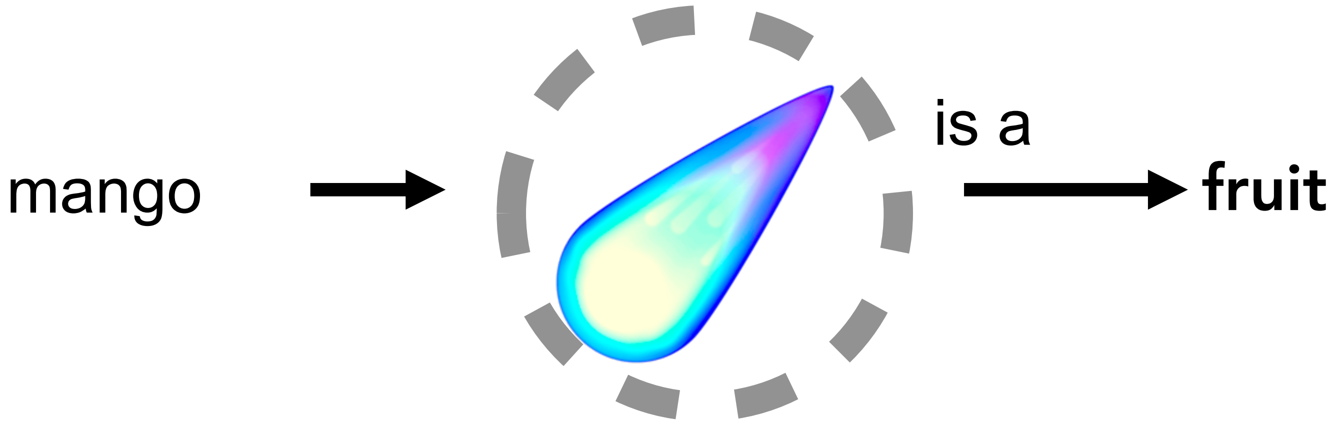
COMET - ConceptNet



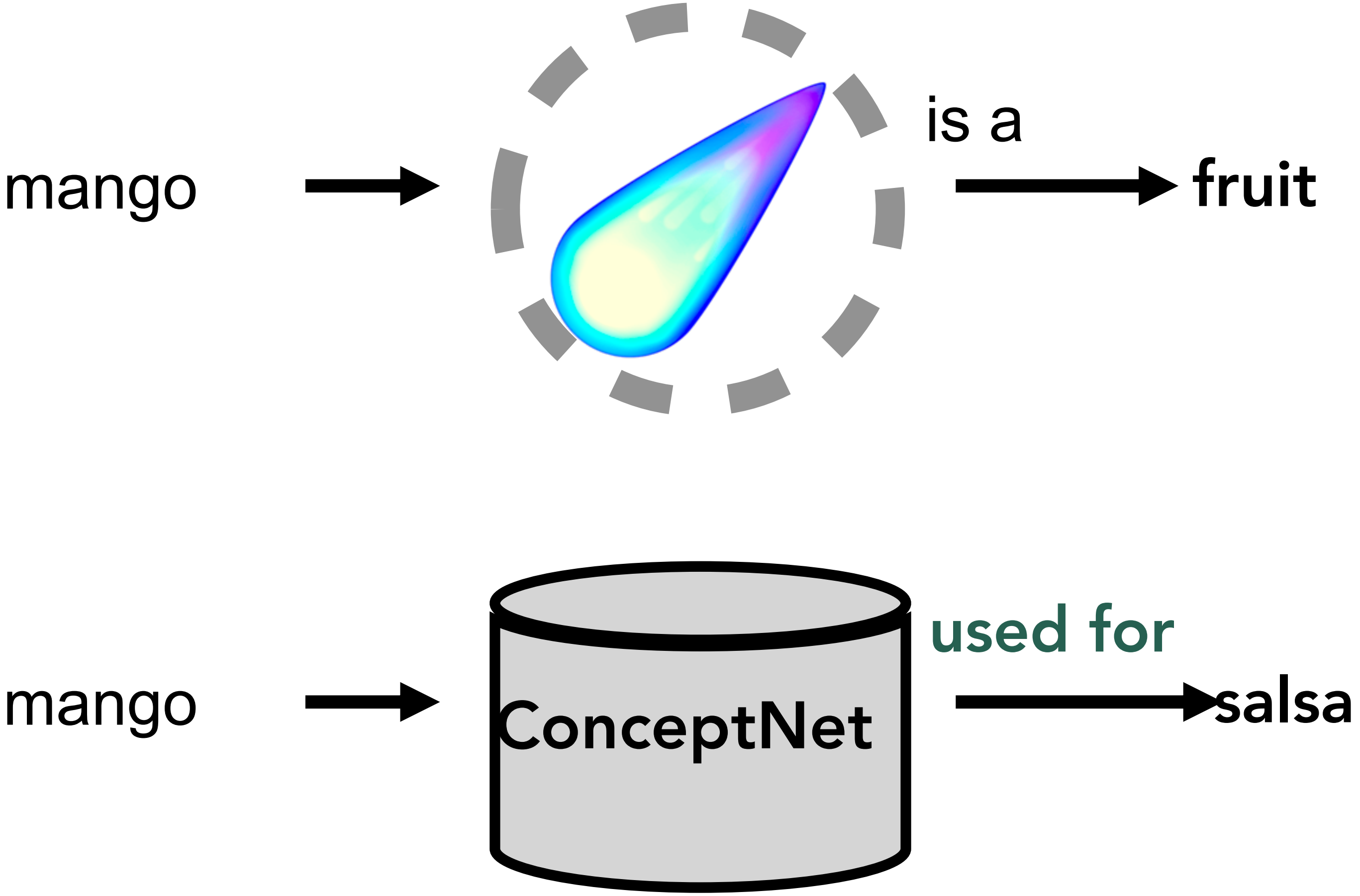
Question

Why does this work?

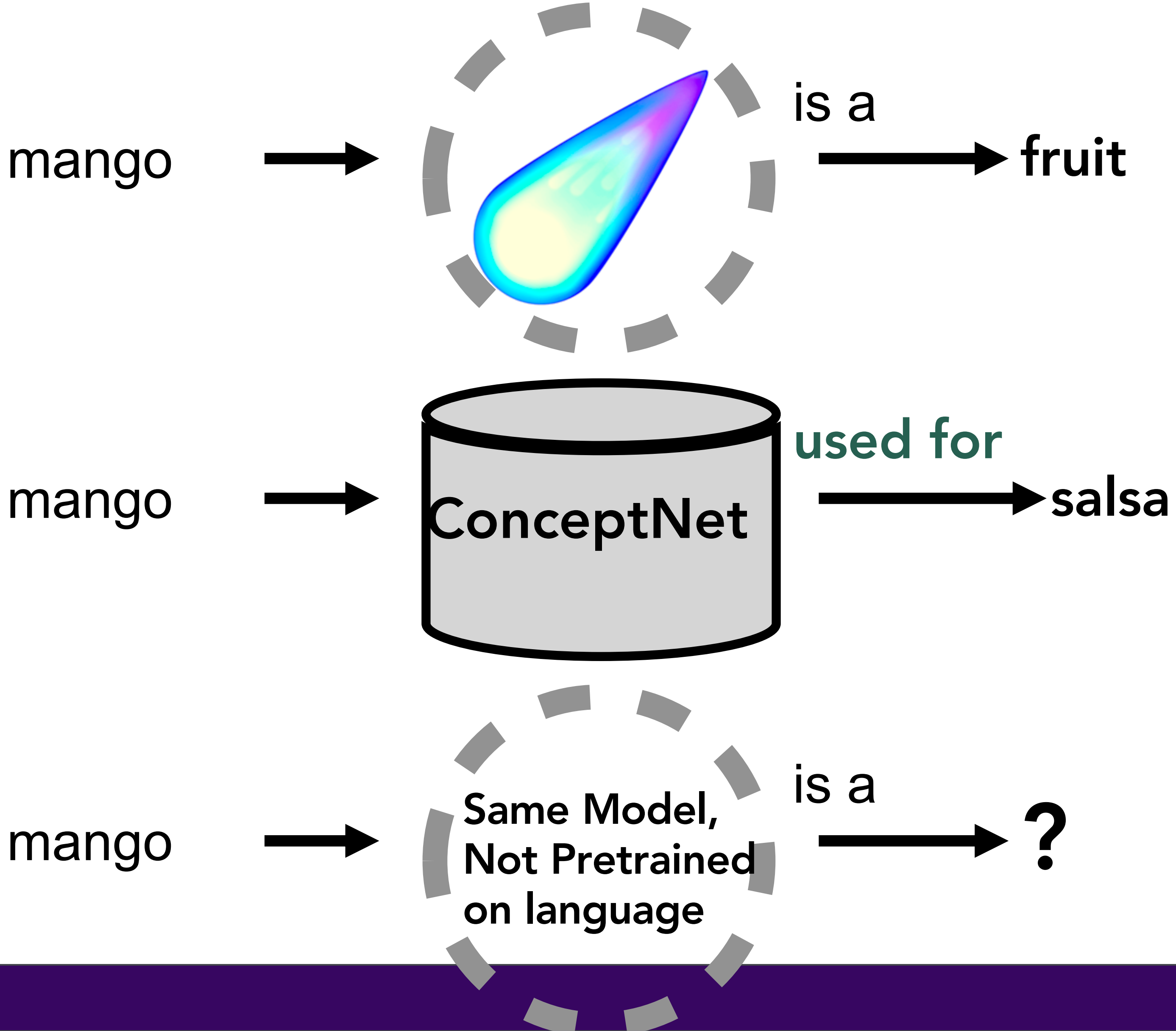
Transfer Learning from Language



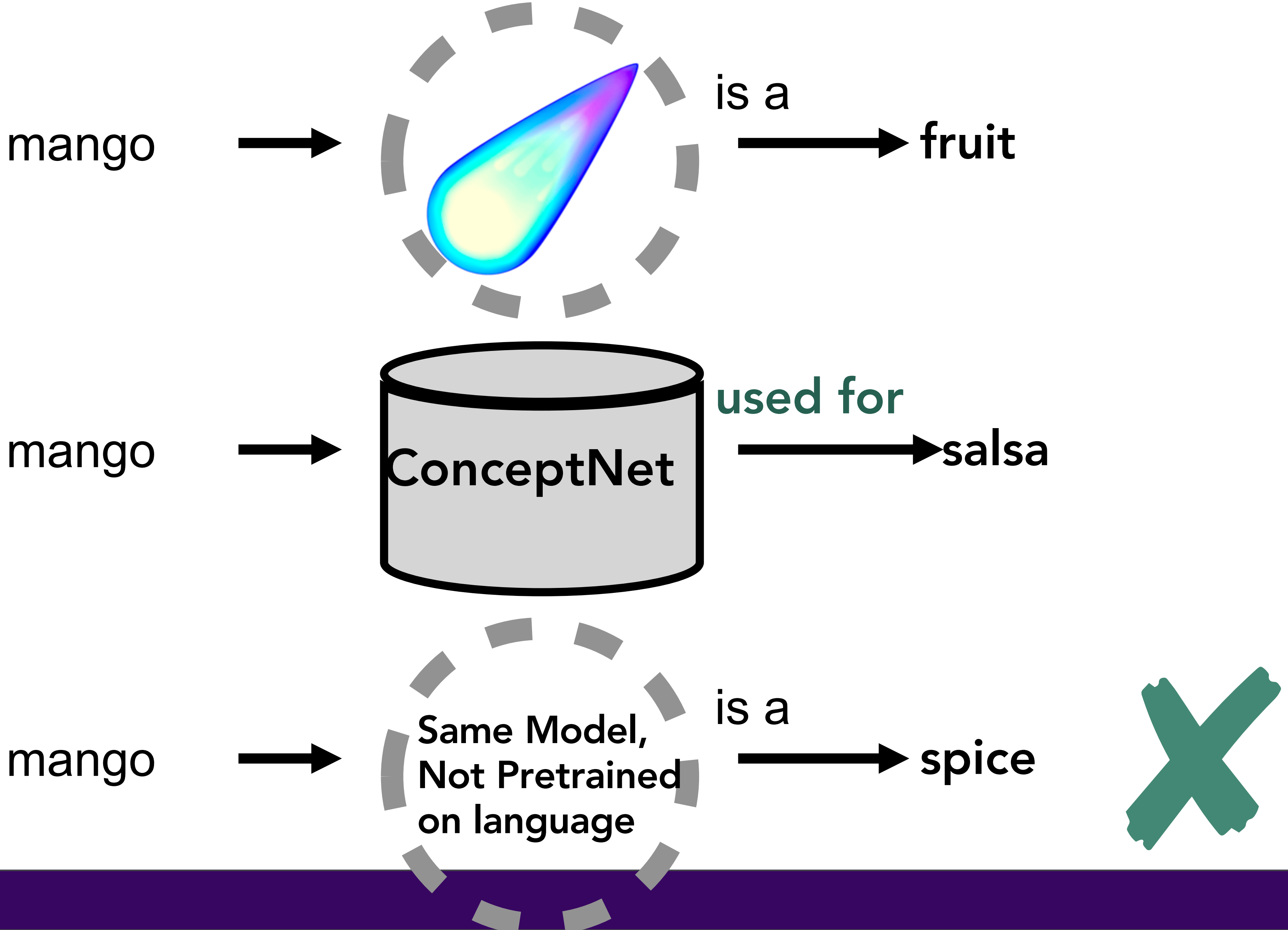
Transfer Learning from Language



Transfer Learning from Language



Transfer Learning from Language



Question

Can’t a language model do the same thing?

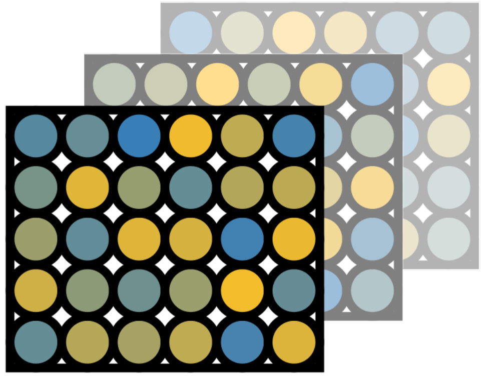
Unsupervised Commonsense Probing

(Dante, <born_in>, ?)

map relation to one or more natural language sentences

"Dante was born in [MASK]."

LM

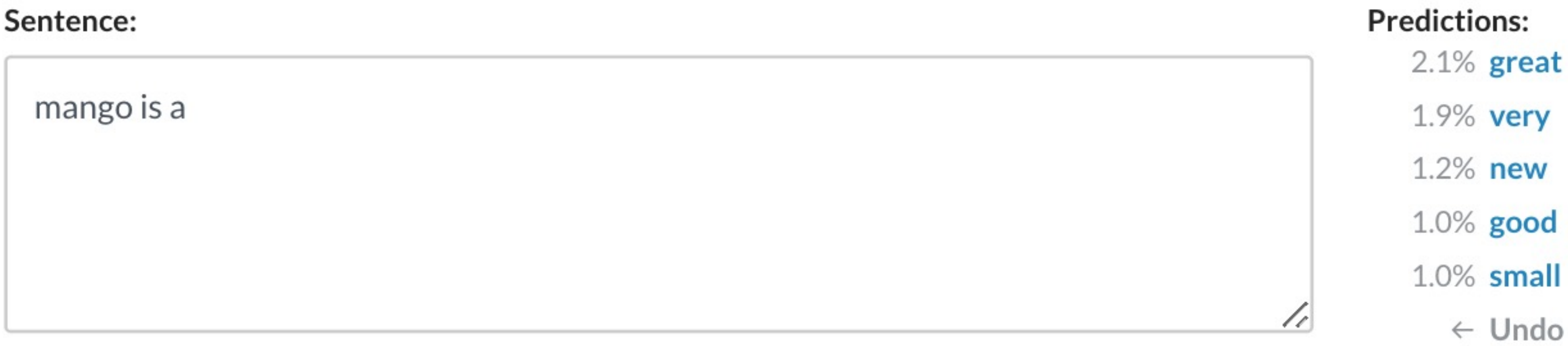


Neural LM Memory Access → Florence

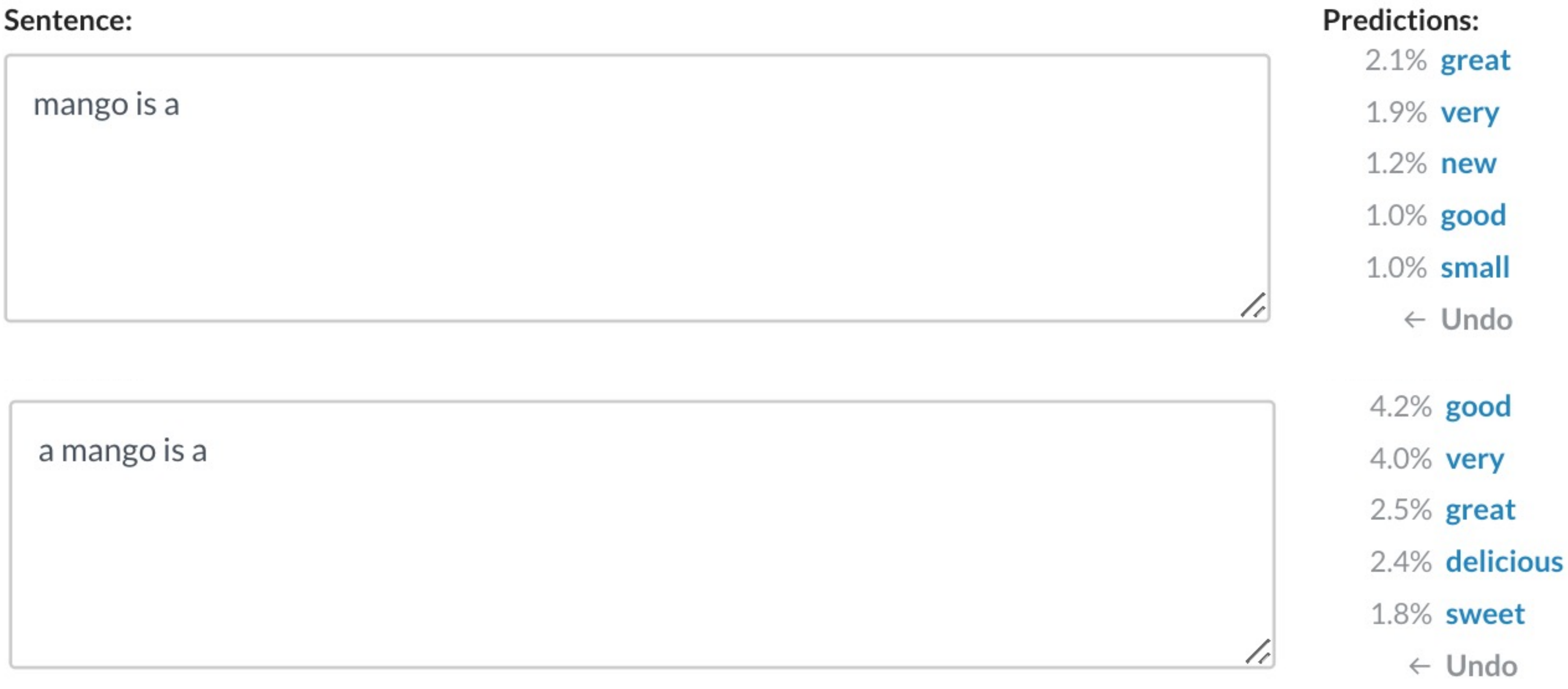
e.g. ELMo/BERT

Petroni et al., 2019; Feldman et al., 2019

Do Language Models know this?



Do Language Models know this?



https://demo.allennlp.org/next-token-lm

Do Language Models know this?

Sentence:

mango is a

Predictions:

2.1%	great
1.9%	very
1.2%	new
1.0%	good
1.0%	small

← Undo

a mango is a

4.2%	good
4.0%	very
2.5%	great
2.4%	delicious
1.8%	sweet

← Undo

Sentence:

A mango is a

Predictions:

4.2%	fruit
3.5%	very
2.5%	sweet
2.2%	good
1.5%	delicious

← Undo

https://demo.allennlp.org/next-token-lm

Do Masked Language Models know this?

Sentence:

mango is a [MASK]

Mask 1 Predictions:
69.7% .
9.3% ;
1.7% !
0.8% vegetable
0.7% ?

Sentence:

mango is a [MASK] .

Mask 1 Predictions:
7.6% staple
7.6% vegetable
4.6% plant
3.5% tree
3.5% fruit

Sentence:

A mango is a [MASK] .

Mask 1 Predictions:
16.0% banana
12.1% fruit
5.9% plant
5.5% vegetable
2.5% candy

https://demo.allennlp.org/masked-lm

Sensitivity to cues

Candidate Sentence S_i	$\log p(S_i)$
"musician can playing musical instrument"	-5.7
"musician can be play musical instrument"	-4.9
"musician often play musical instrument"	-5.5
"a musician can play a musical instrument"	$-\mathbf{2.9}$

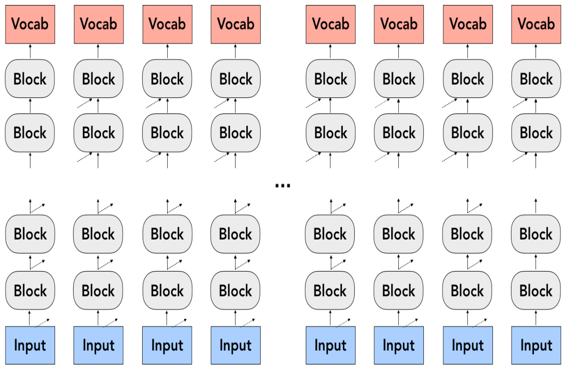
Feldman et al., 2019

Prompt	**Model Predictions**
A ____ has fur.	dog, cat, fox, ...
A ____ has fur, is big, and has claws.	cat, **bear**, lion, ...
A ____ has fur, is big, has claws, has teeth, is an animal, eats, is brown, and lives in woods.	**bear**, wolf, cat, ...

Weir et al., 2020

Commonsense Transformers

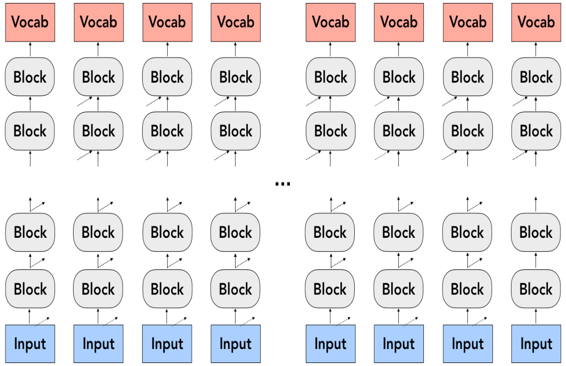
- Language models **implicitly represent knowledge**



Pre-trained
Language Model

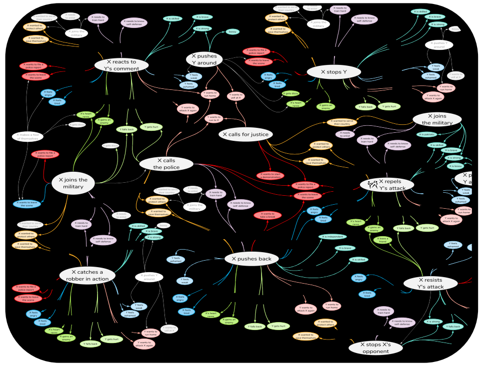
Commonsense Transformers

- Language models **implicitly represent knowledge**
- Re-train them on knowledge graphs to **learn structure of knowledge**



Pre-trained
Language Model

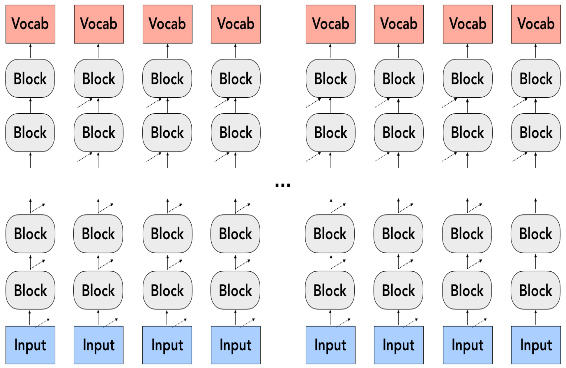
+



Seed Knowledge
Graph Training

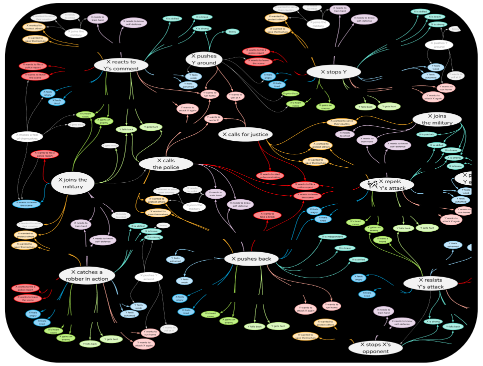
Commonsense Transformers

- Language models **implicitly represent knowledge**
- Re-train them on knowledge graphs to **learn structure of knowledge**
- Resulting knowledge model **generalizes structure** to other concepts



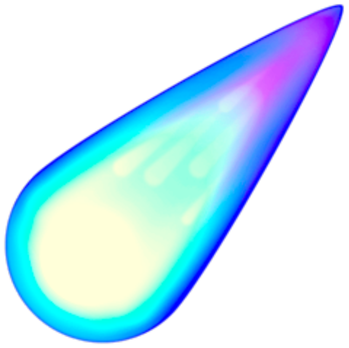
Pre-trained
Language Model

+



Seed Knowledge
Graph Training

=



COMET

Question

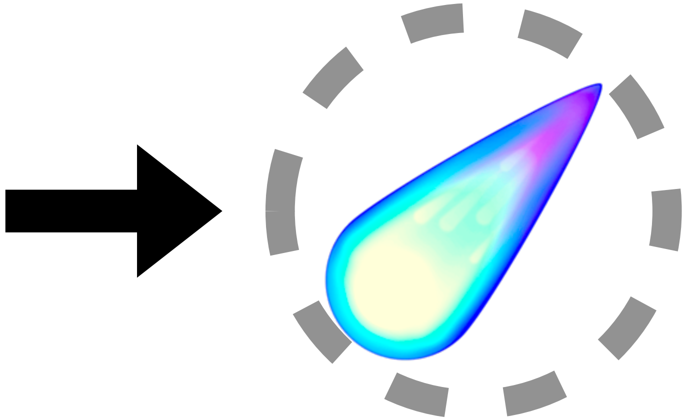
What are the implications of this knowledge representation?

Commonsense Knowledge for any Situation

transformer-style architecture — input format is natural language

- event can be fully parsed

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

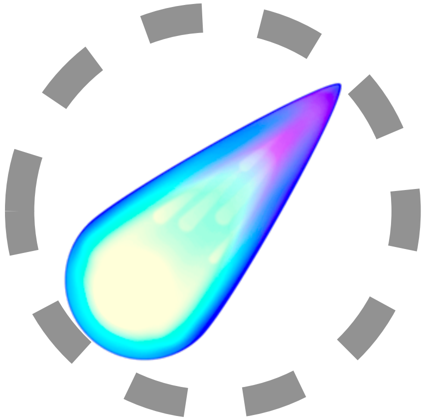


Commonsense Knowledge for any Situation

transformer-style architecture — input format is natural language

- event can be fully parsed
- knowledge generated **dynamically** from neural knowledge model

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