Commonsense Reasoning

Lara J. Martin (she/they)

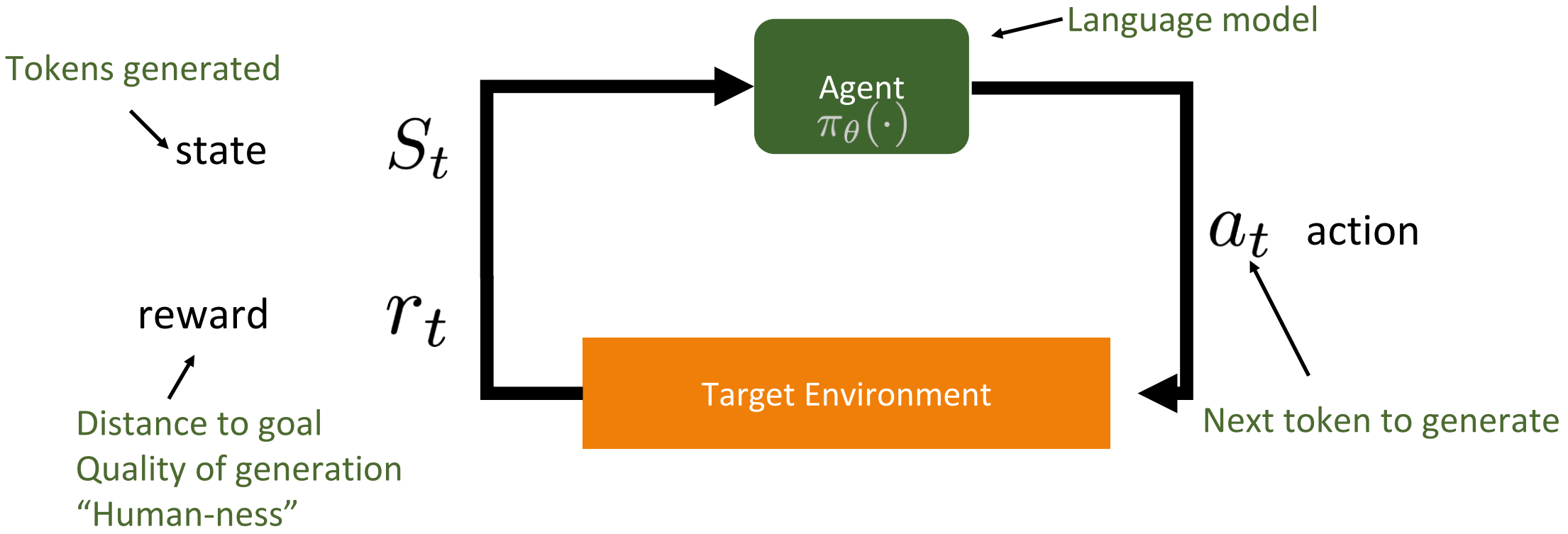
https://laramartin.net/interactive-fiction-class

Modified from slides by Chris Callison-Burch

Learning Objectives

- Consider the difference between human commonsense and machine reasoning
- Categorize the types of commonsense
- Find out about existing popular commonsense knowledge bases
- Distinguish when you would use one knowledge base over another

Review:
RL in the Context of Language Models…



$a_t \sim \pi_\theta(S_t)$: policy

Review: REINFORCE

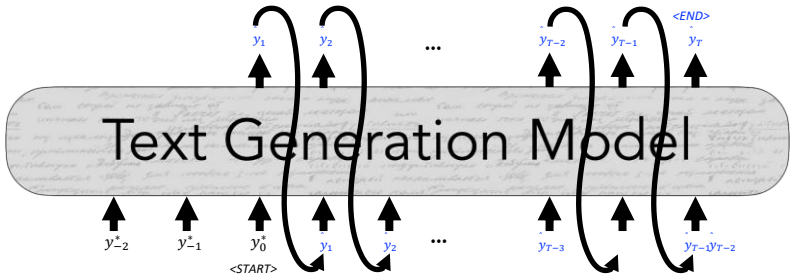
- Sample a sequence from your model, score the sequence, and use the score to train the model.

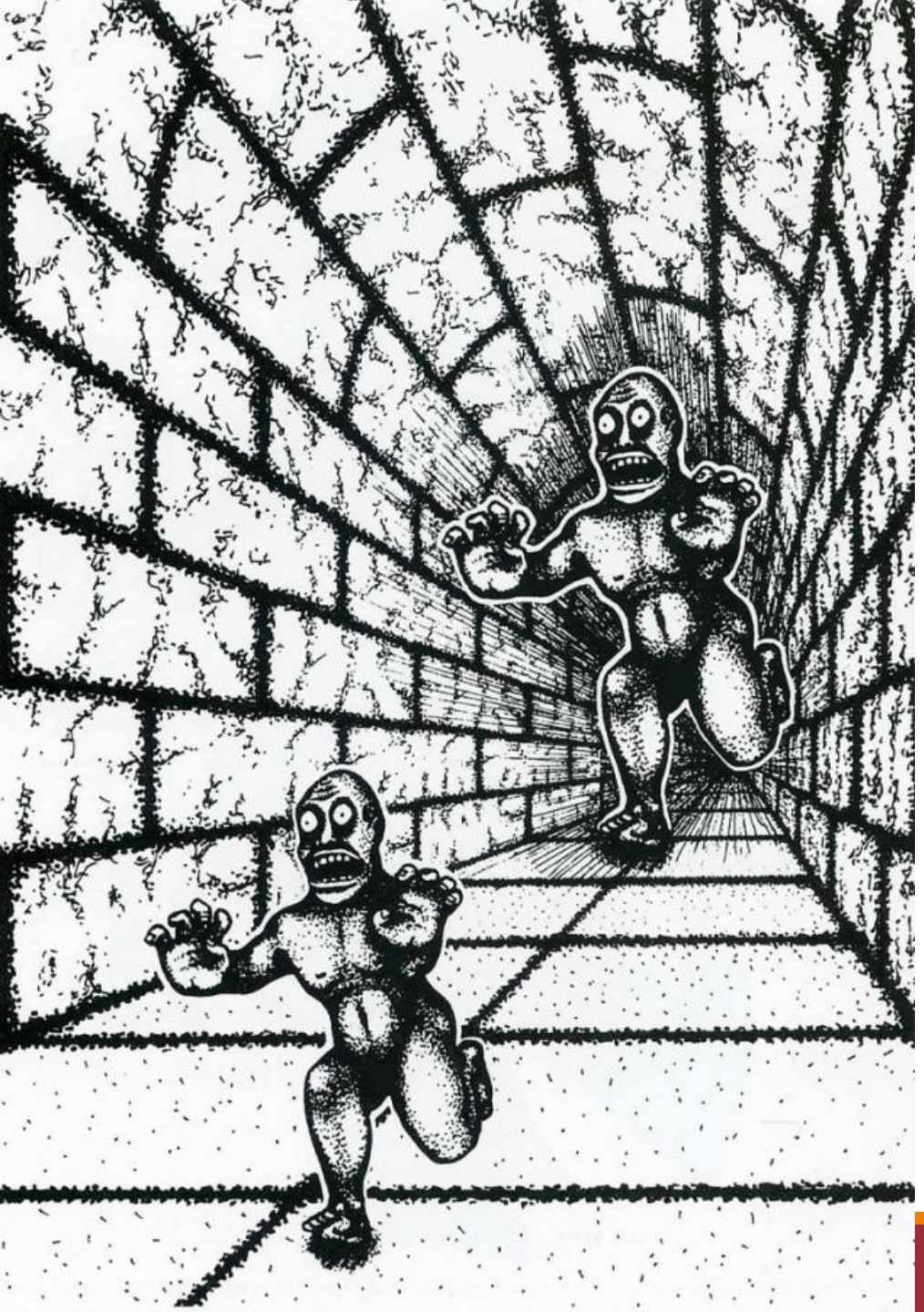
Next time, increase the probability of this sampled token in the same context.

$$L_{RL} = -\sum_{t=1}^{T} r(\hat{y}_t) \log P(y_t | \{y^*\}; \{\hat{y}\}_{<t})$$

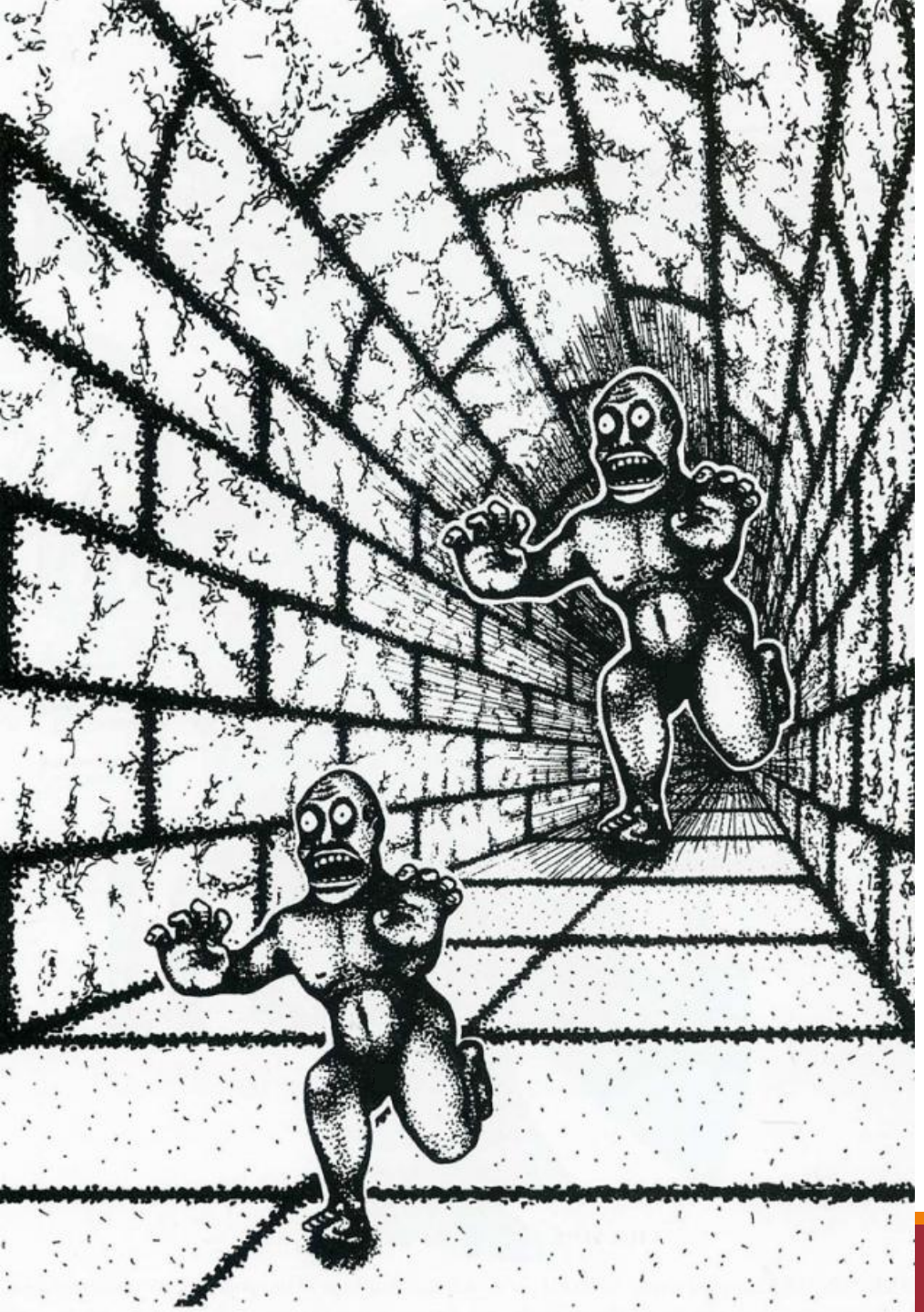
... but increase it more if I get a higher reward from the reward function.

- $r(\cdot)$: Your reward model
- y^* :Input sequence given to the model
- $\hat{y}$:The sequence sampled from the model given y^*





What does this picture show?



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”

ChatGPT passes exams from law and business schools

By Samantha Murphy Kelly, CNN Business

4 minute read · Updated 1:35 PM EST, Thu January 26, 2023

CADE METZ BUSINESS MAY 22, 2017 3:12 PM

AlphaGo Is Back to Battle Mere Humans—and It's Smarter Than Ever

s in China to take on the world's top-ranked
e there for every move.

JULY 12, 2022 | 6 MIN READ

Google Engineer Claims AI Chatbot Is Sentient: Why That Matters

Is it possible for an artificial intelligence to be sentient?

BY LEONARDO DE COSMO

AI in real-life usage: Can't win an argument with your partner? Get ChatGPT to do it for you

Girlfriend goes viral for using ChatGPT to make her arguments when couple disagrees, apparently telling her partner that they lack 'emotional bandwidth.'

Darren Allan
Tech Reporter

Published Oct 17, 2024 9:05 AM CDT
3 minutes read time

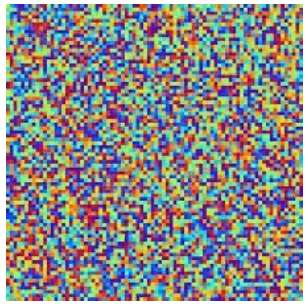
Newsweek SUBSCRIBE FOR $1 Login

Tech | Joe Rogan | The Joe Rogan Experience | Kamala Harris | Donald Trump

Who Is Joe Rogan Voting For? We Asked ChatGPT

Published Oct 15, 2024 at 6:38 AM EDT

Updated Oct 15, 2024 at 6:07 PM EDT


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.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. Developmental Science, 10(1), 89-96. doi: 10.1111/j.1467-7687.2007.00569.x

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

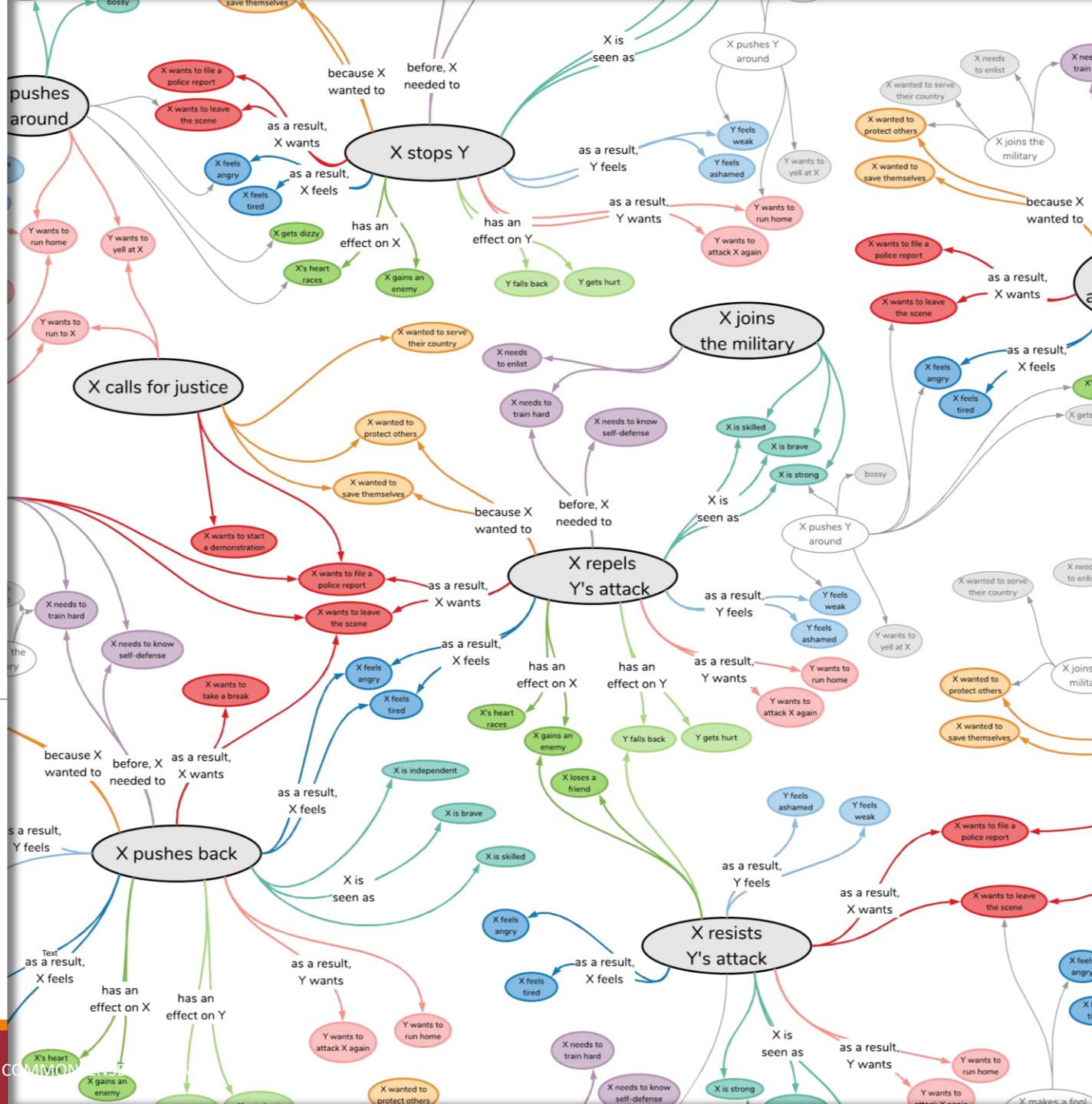
Isaac Asimov's "Three Laws of Robotics"

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

- Isaac Asimov, 1942 short story "Runaround"

https://cdn.britannica.com/82/195182-050-97684526/Isaac-Asimov-1979.jpg

Commonsense resources



Grandma’s glasses

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

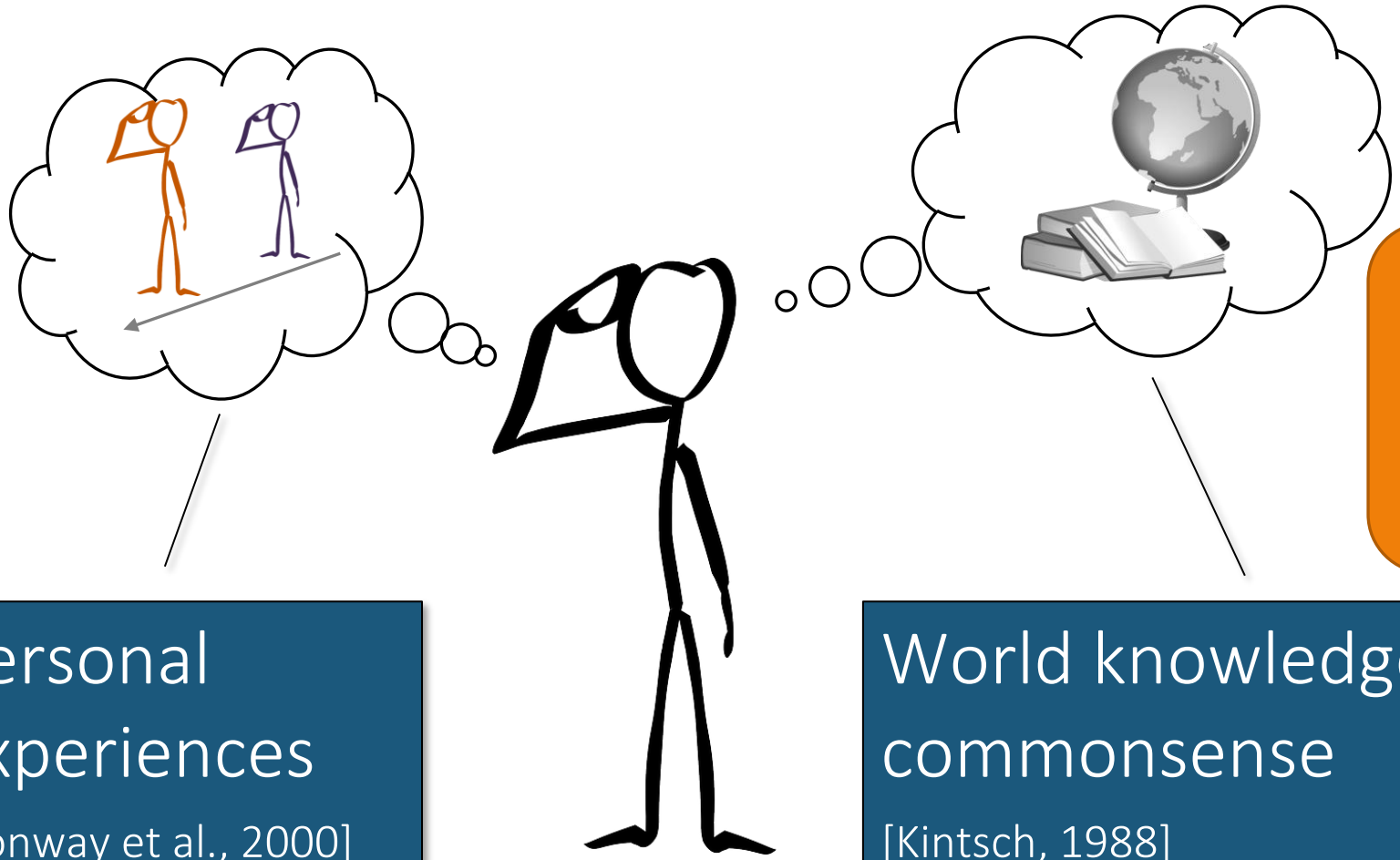
She couldn’t pick them up, so she called Tom for help.

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.

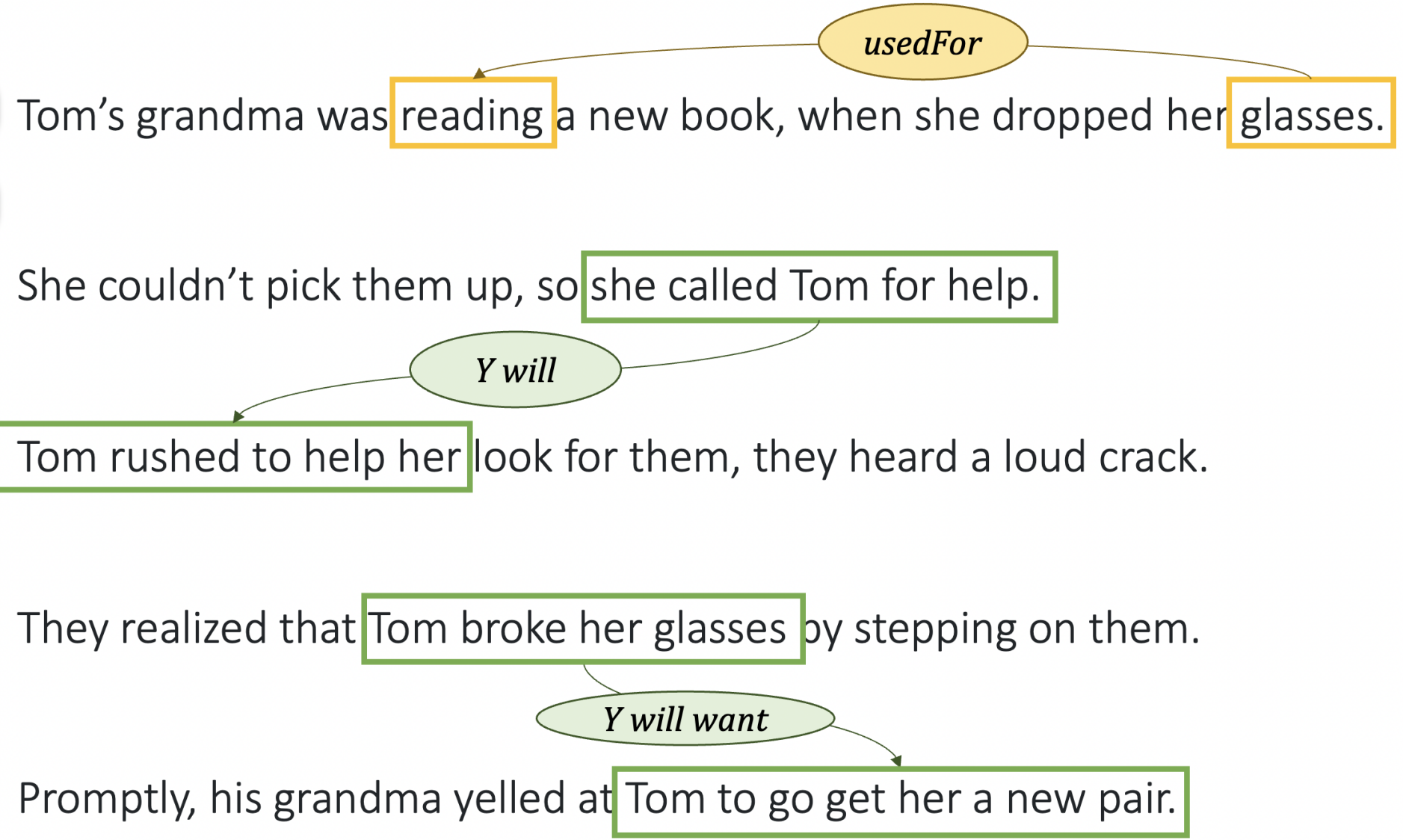
Tom rushed to help her look for them, they heard a loud crack.

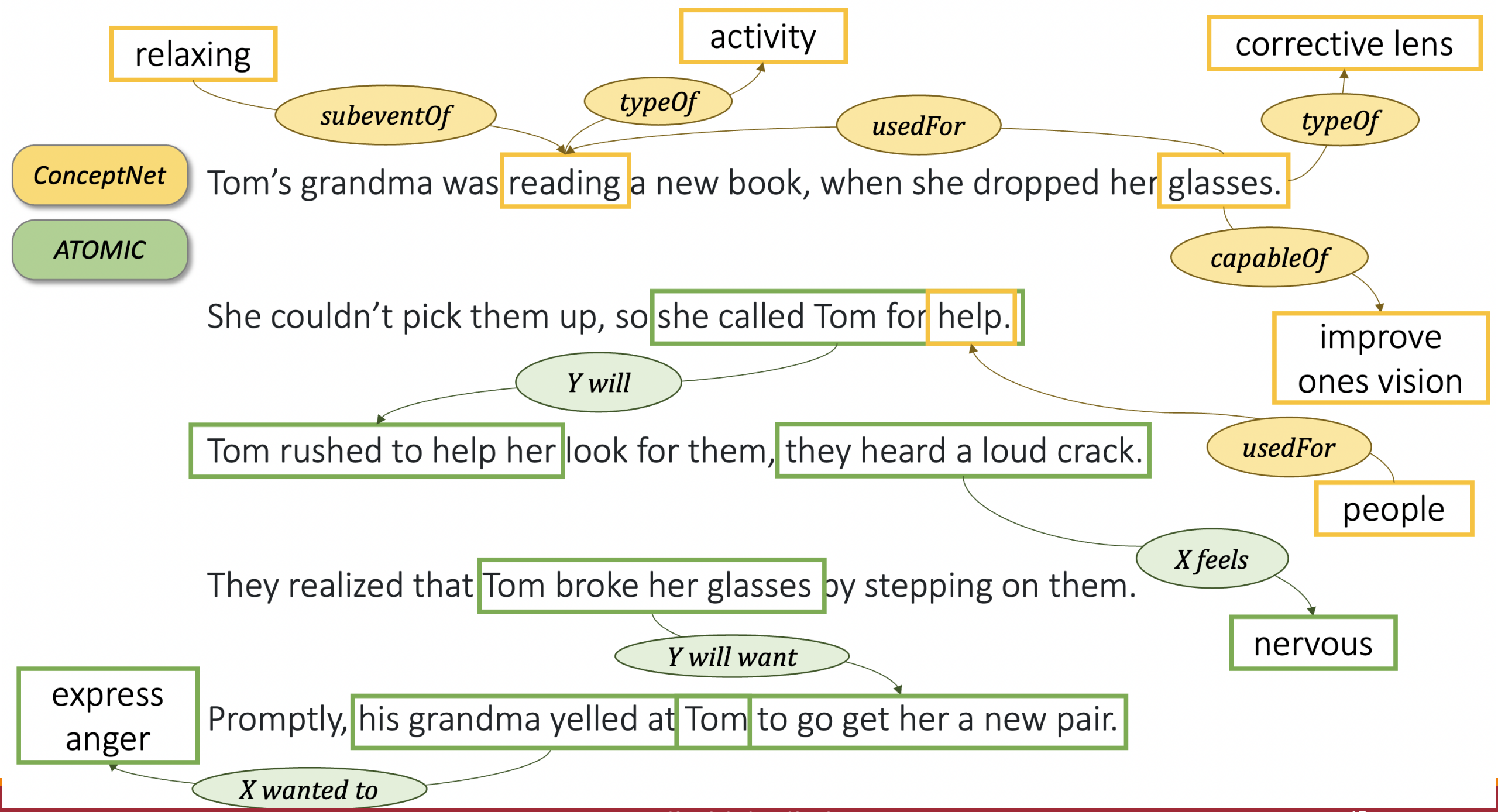
They realized that Tom broke her glasses by stepping on them.

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

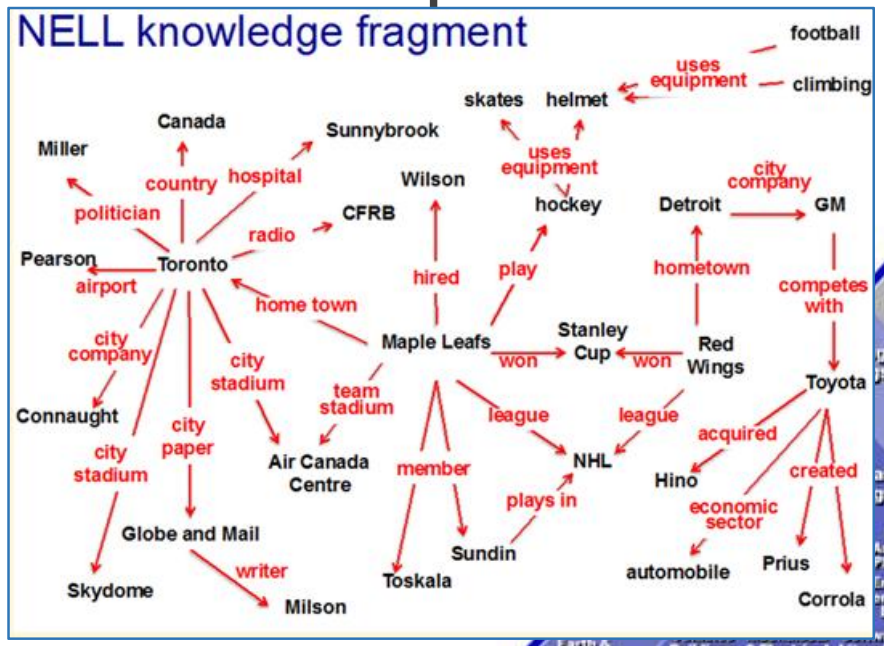
ConceptNet

ATOMIC




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

Some important resources



ATOMIC
(Sap et al., 2019)

ild 2.0
al., 2017)

Open Mind Comm
(Minsky, Singh & Ha

Net 5.5
al., 2017)

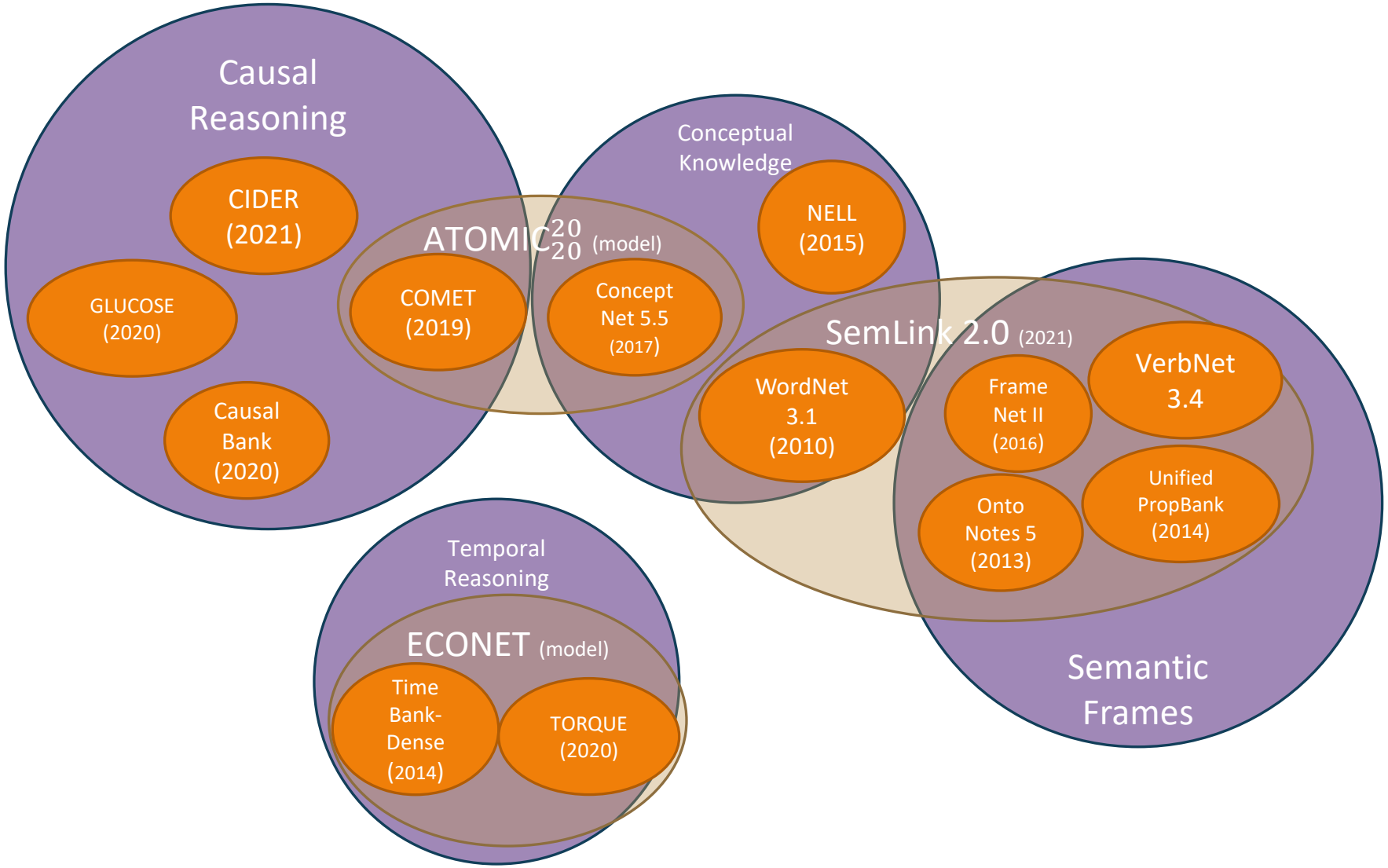
Cyc
(Lenat et al., 1984)

OpenCyc
(Lenat, 2004)

ResearchCyc
(Lenat, 2006)

OpenCyc 4.0
(Lenat, 2012)

today


Causal Reasoning
CIDER (2021)
GLUCOSE (2020)
Causal Bank (2020)
ATOMIC2020 (model)
COMET (2019)
Conceptual Knowledge
Concept Net 5.5 (2017)
NELL (2015)
SemLink 2.0 (2021)
WordNet 3.1 (2010)
Frame Net II (2016)
VerbNet 3.4
Onto Notes 5 (2013)
Unified PropBank (2014)
Semantic Frames
Temporal Reasoning
ECONET (model)
Time Bank-Dense (2014)
TORQUE (2020)

How do you create a commonsense resource?

Desirable properties for a commonsense resource

COVERAGE

Large scale

Diverse knowledge types

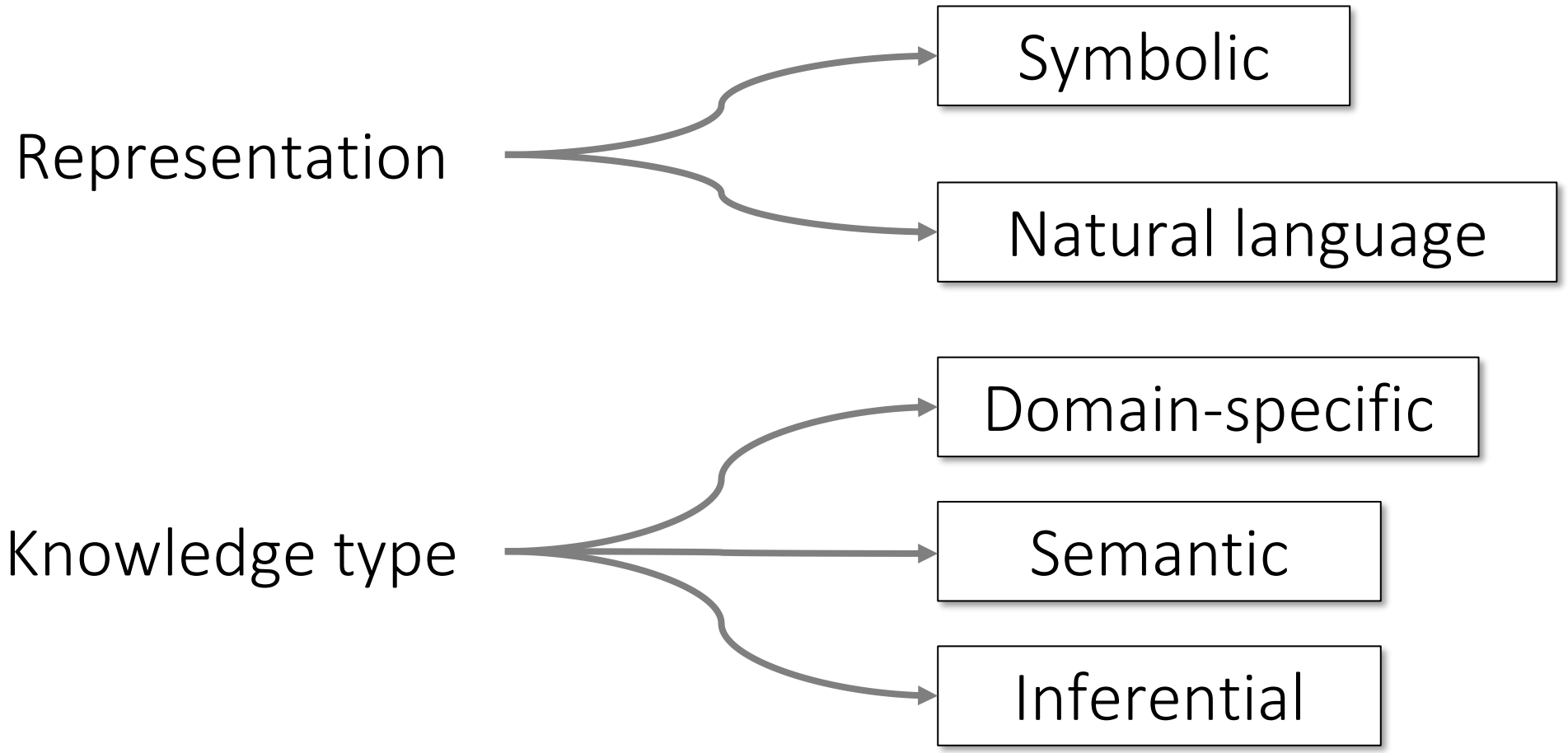
USEFUL

High quality knowledge

Usable in downstream tasks

Multiple resources tackle different knowledge types

Creating a commonsense resource



What is a semantic frame?

"people understand the meaning of words largely by virtue of the frames which they evoke"

- Understanding words in context
- Based on recurring experiences

Josef Ruppenhofer, Michael Ellsworth, Miriam R. L Petruck, Christopher R. Johnson, Collin F. Baker, & Jan Scheffczyk. *FrameNet II: Extended Theory and Practice* (Revised November 1, 2016.)

Fillmore, Charles J. (1982). "Frame semantics". In The Linguistic Society of Korea, eds. *Linguistics in the Morning Calm*. Seoul: Hanshin. 111-37.

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 →

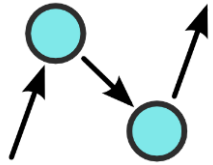
Things used for reading

- 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:

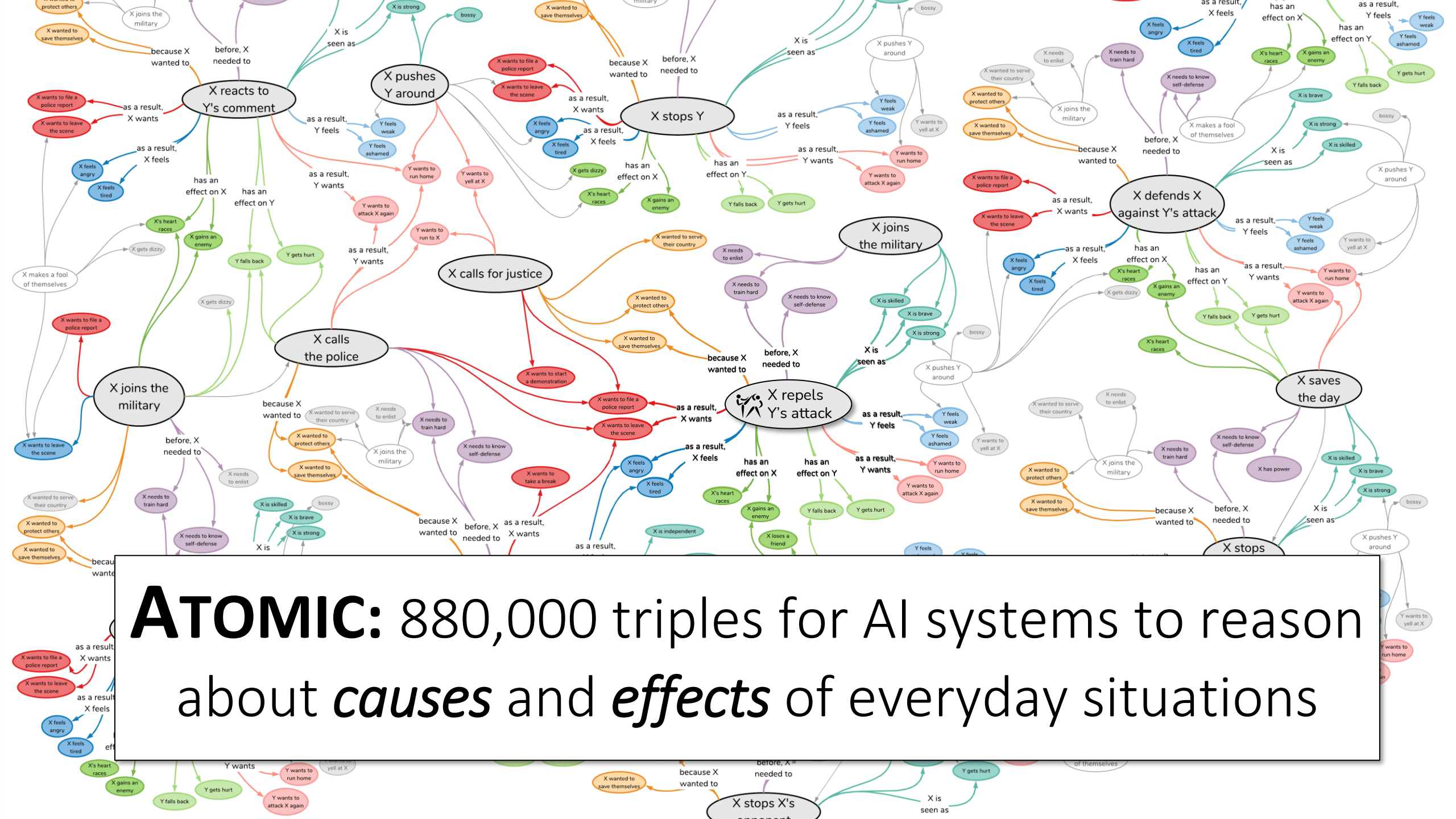
- Open Mind Commonsense assertions
- Wikipedia/Wiktionary semantic knowledge
- WordNet, Cyc ontological knowledge

http://conceptnet.io/

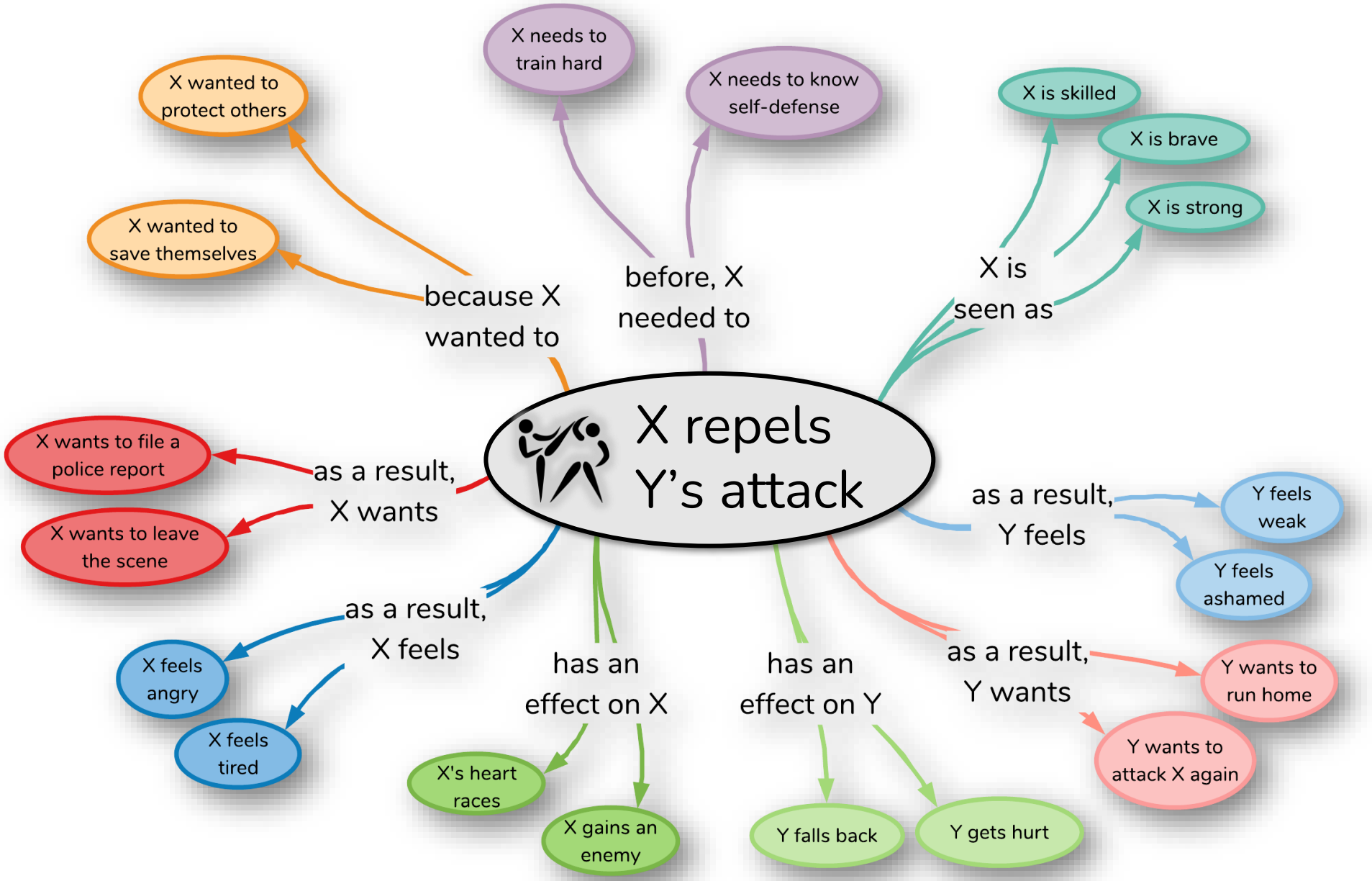
ATOMIC:

inferential knowledge in *natural language* form

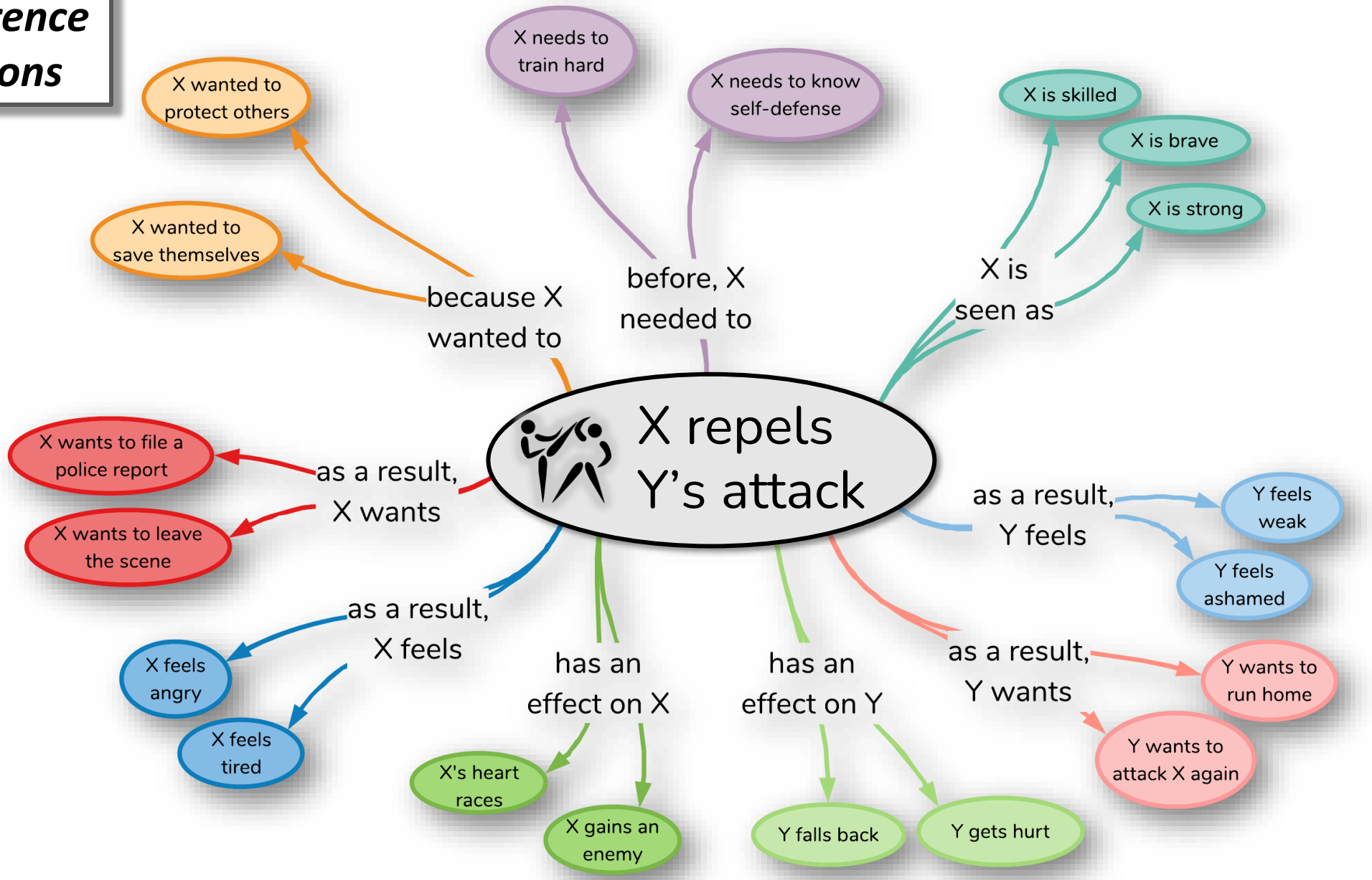
https://github.com/allenai/comet-atomic-2020

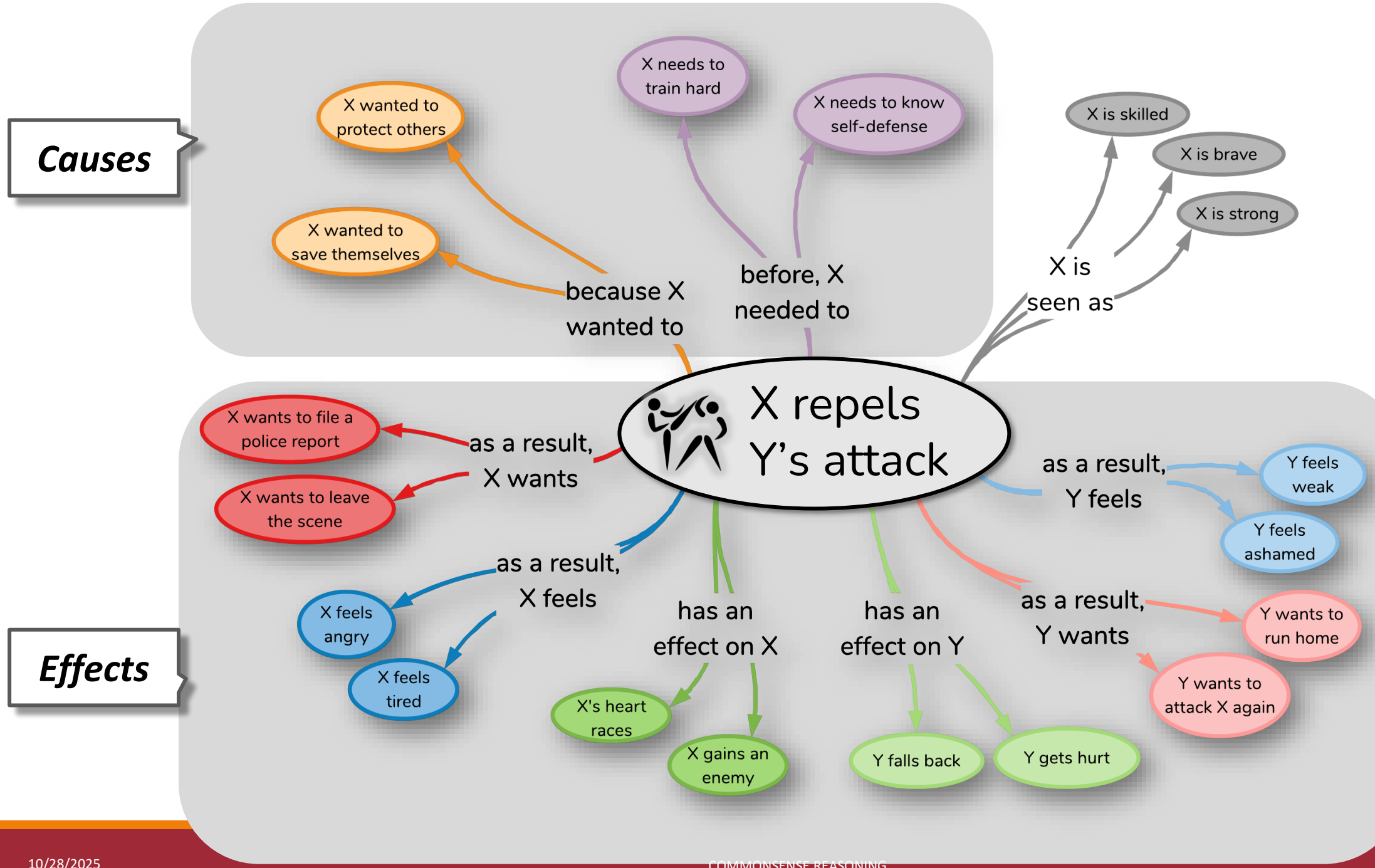


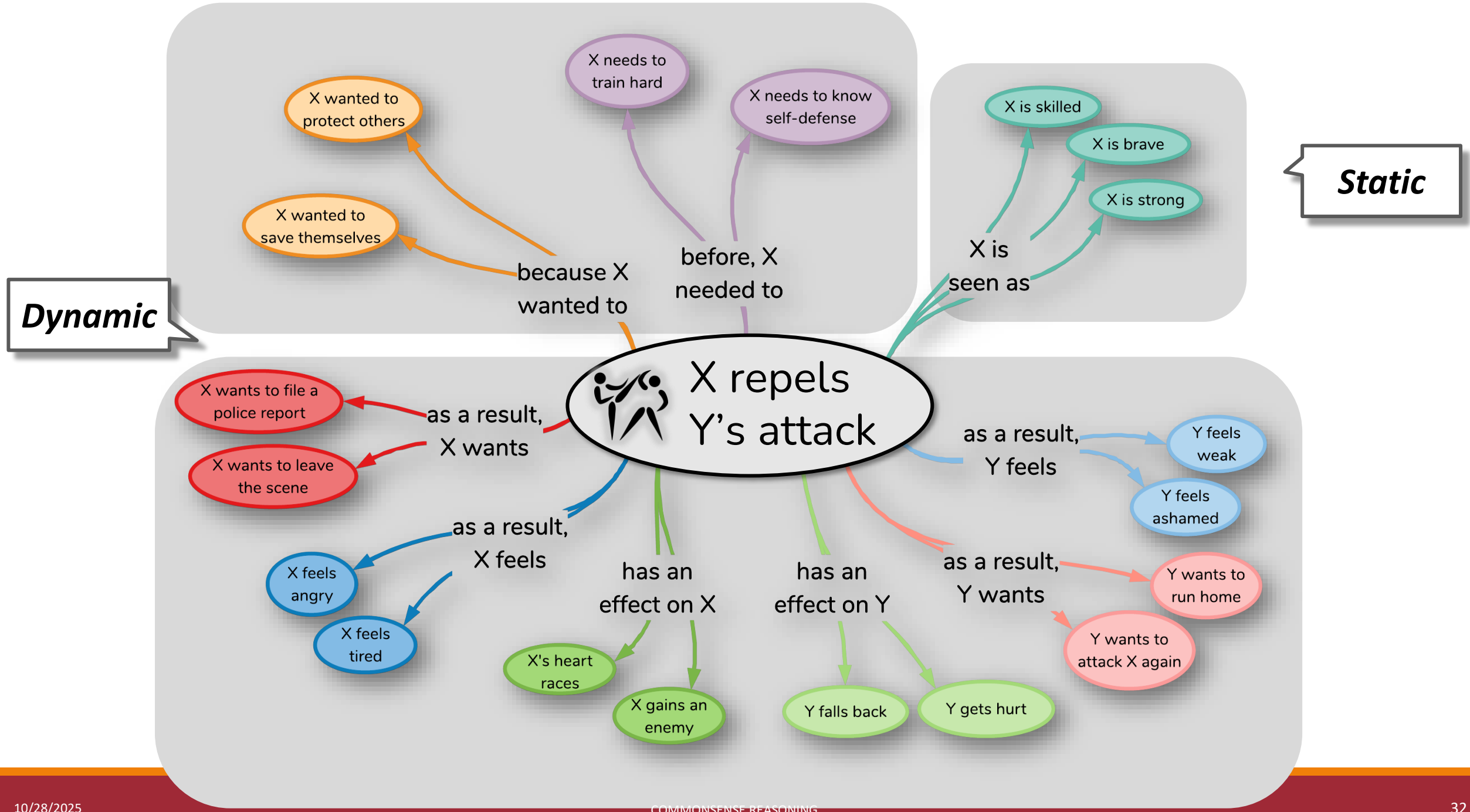
ATOMIC: 880,000 triples for AI systems to reason about ***causes*** and ***effects*** of everyday situations

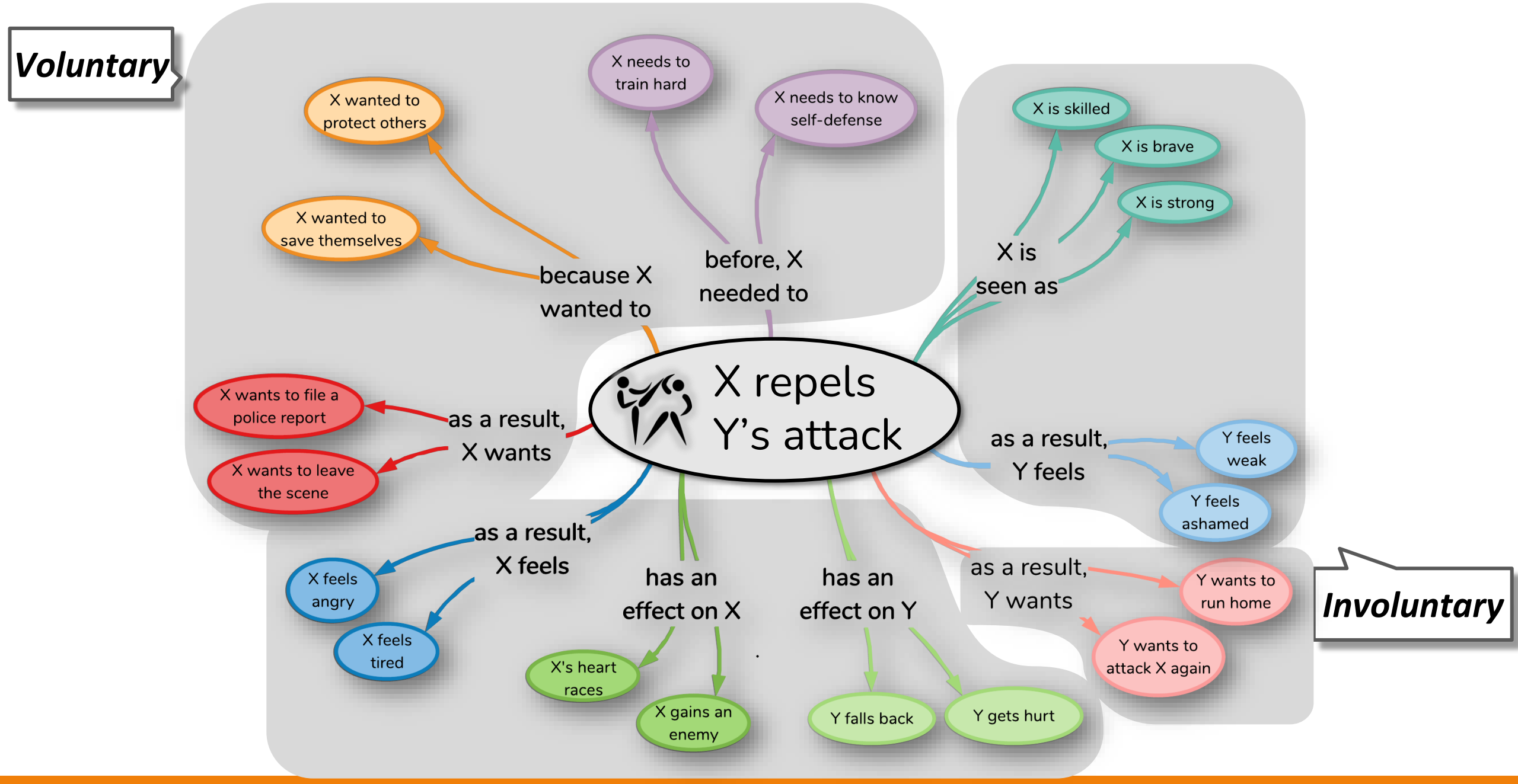

X repels Y's attack
because X wanted to
X wanted to protect others
X wanted to save themselves
before, X needed to
X needs to train hard
X needs to know self-defense
X is seen as
X is skilled
X is brave
X is strong
as a result, X wants
X wants to file a police report
X wants to leave the scene
as a result,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

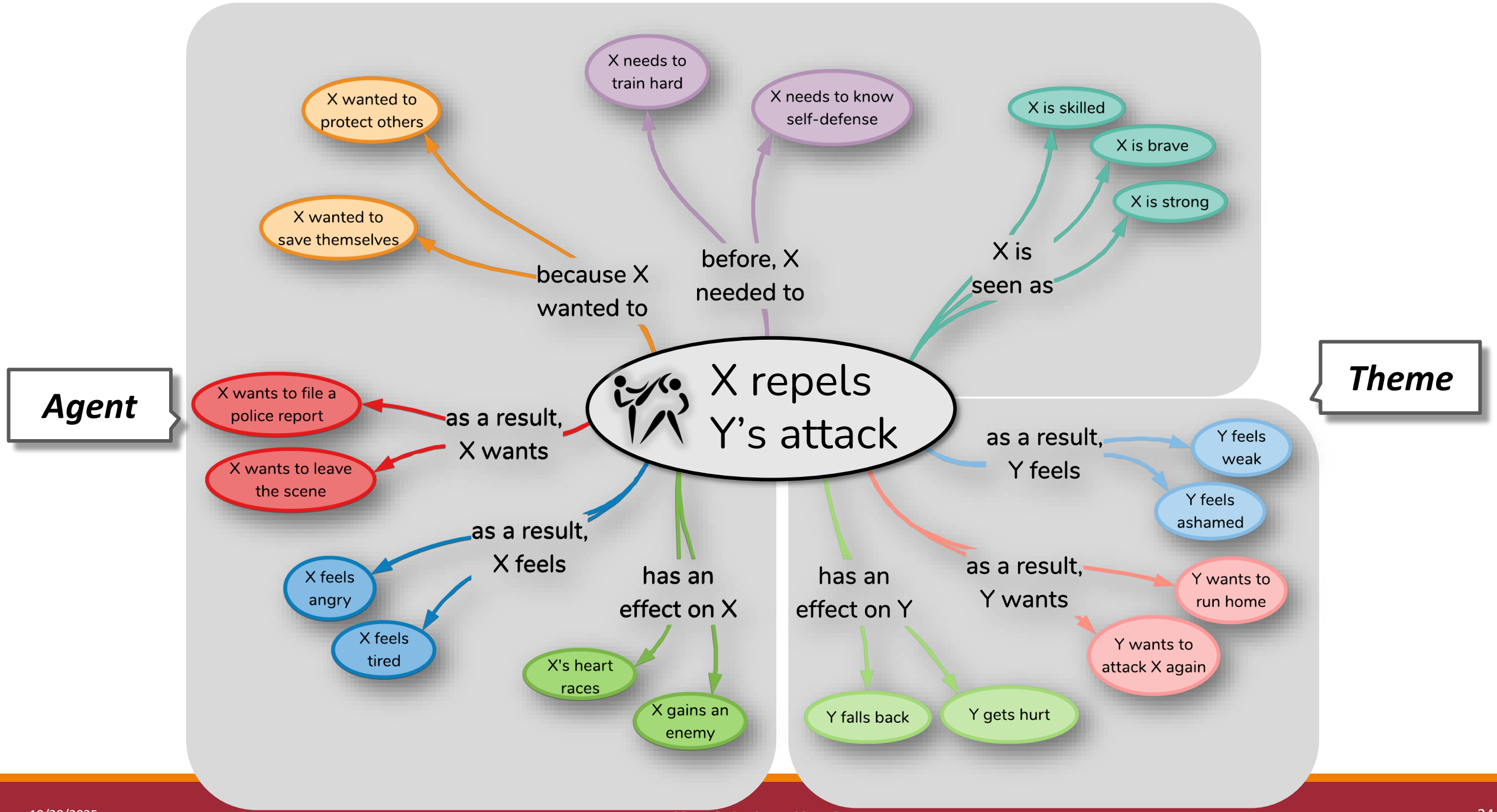
nine inference dimensions

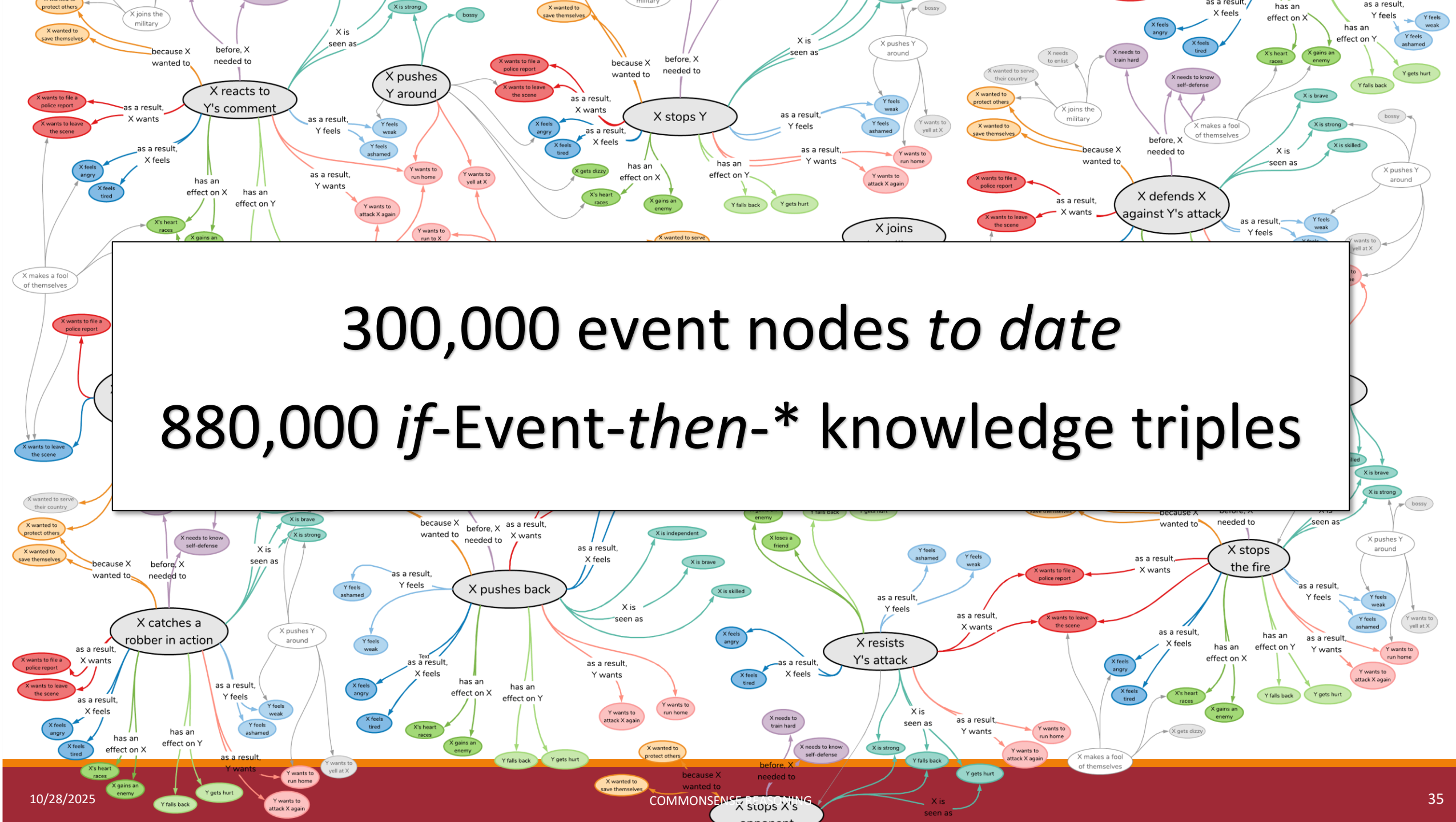



Causes
X wanted to protect others
X needs to train hard
X needs to know self-defense
X wanted to save themselves
because X wanted to
before, X needed to
X is skilled
X is brave
X is strong
X is seen as
X repels Y's attack
X wants to file a police report
X wants to leave the scene
as a result, X wants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


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


Agent
Theme
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, X feels
X feels angry
X feels tired
has an effect on X
X's heart races
X gains an enemy
as a result, Y feels
Y feels weak
Y feels ashamed
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


300,000 event nodes *to date*
880,000 *if*-Event-*then*-* knowledge triples

ATOMIC: knowledge of *cause* and *effect*

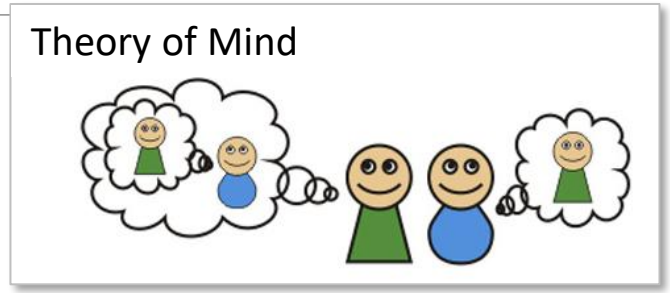
Humans have **theory of mind**, allowing us to

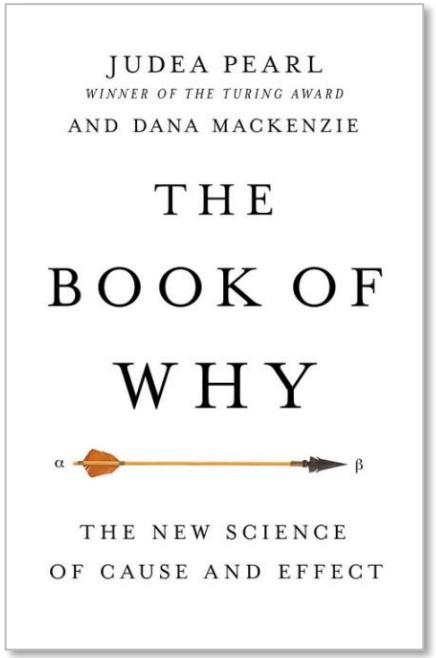
- make inferences about **people's mental states**
- understand **likely events** that precede and follow (Moore, 2013)

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

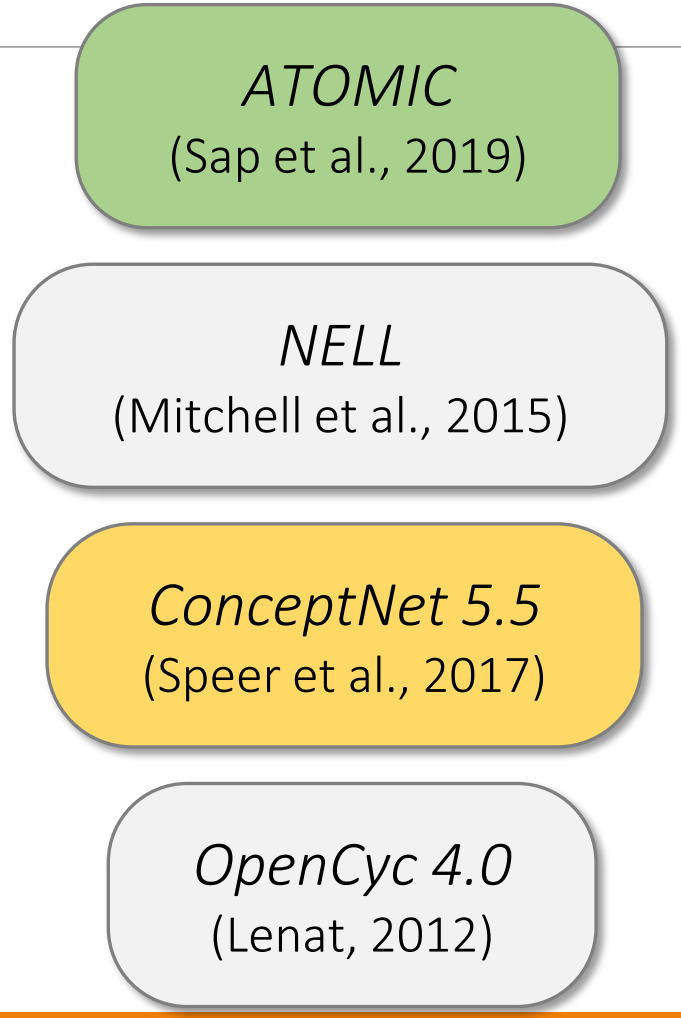
- only find **complex correlational patterns** in data
- **limited to the domain** they are trained on

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





Ways of categorizing existing knowledge bases



Ways of categorizing 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))
```

Ways of categorizing 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)

When would you want to use a taxonomic knowledge base?

How about an inferential?

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

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

ATOMIC
(Sap et al., 2019)

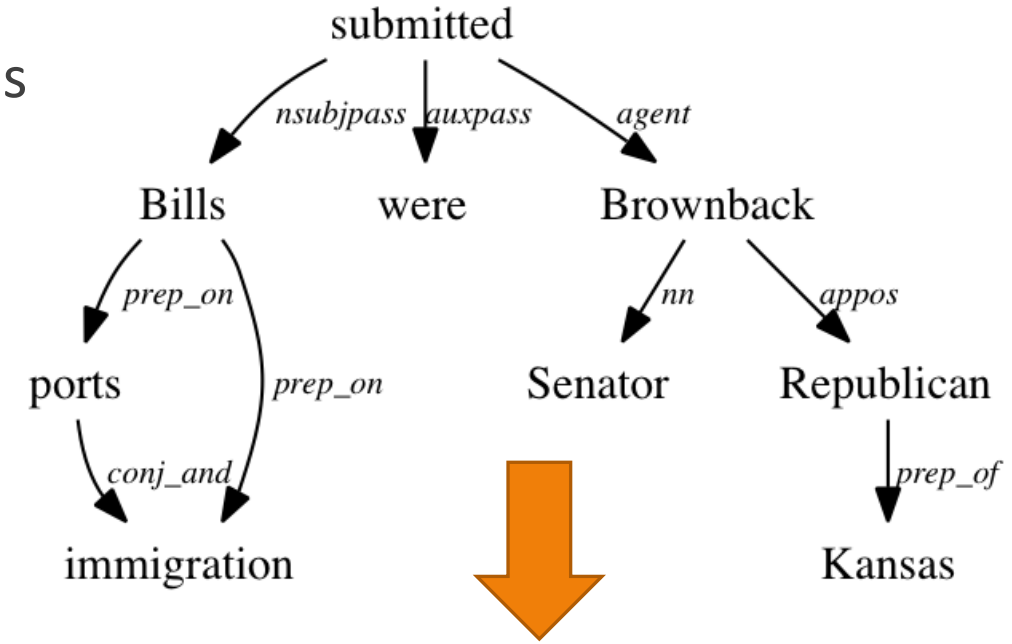
Q: How do you gather commonsense knowledge at scale?

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

Commonsense is not often written

- Grice's maxim of quantity

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

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

NON-EXPERTS WRITE KNOWLEDGE IN NATURAL LANGUAGE PHRASES

What is an “expert” in this case?

OpenCyc 4.0 (Lenat, 2012)

WordNet (Miller et al., 1990)

ATOMIC (Sap et al., 2019)

ConceptNet 5.5 (Speer et al., 2017)

Knowledge Check

What are some advantages of asking experts to create a commonsense knowledge base?