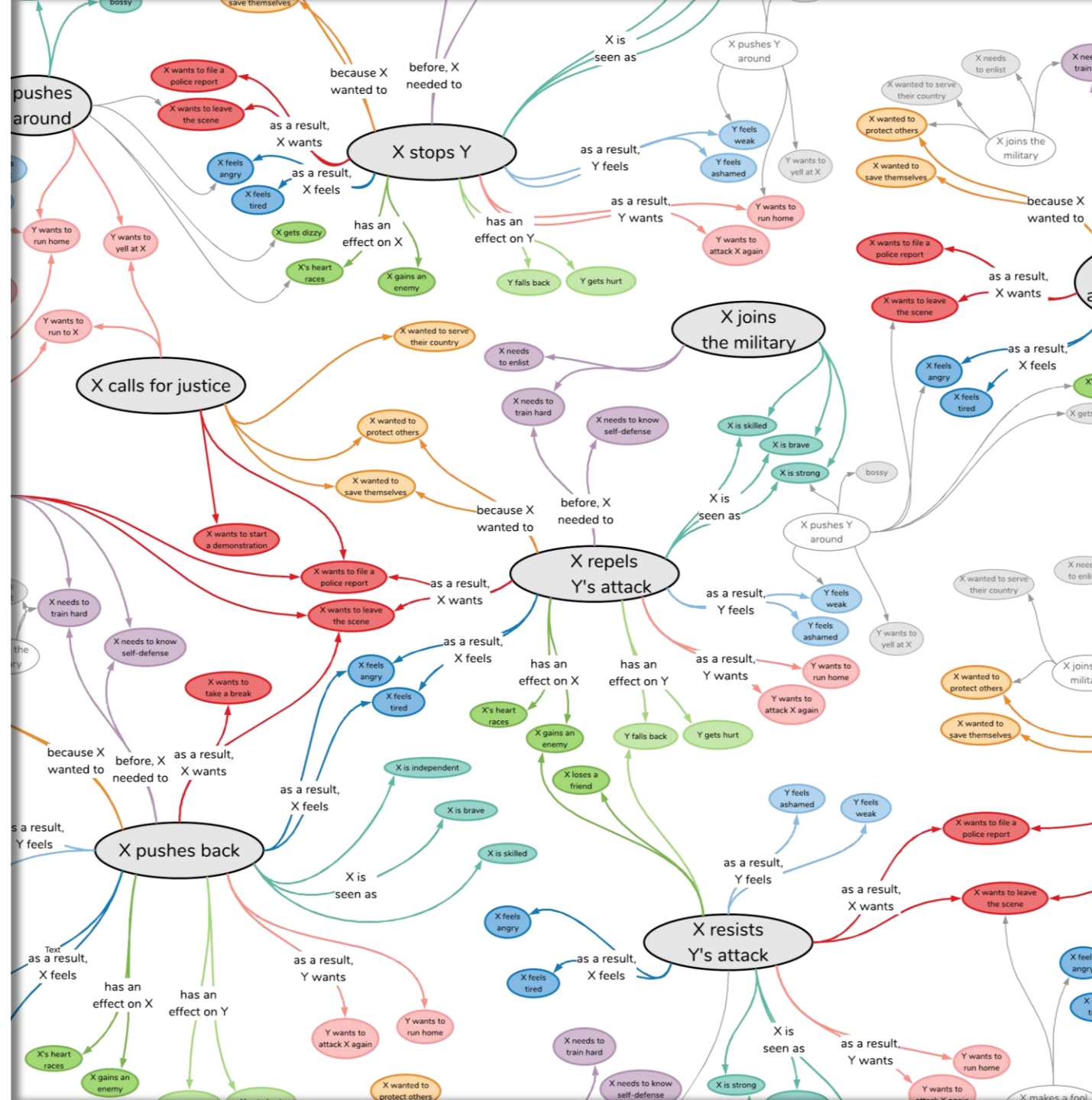


Commonsense resources



Grandma's glasses



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

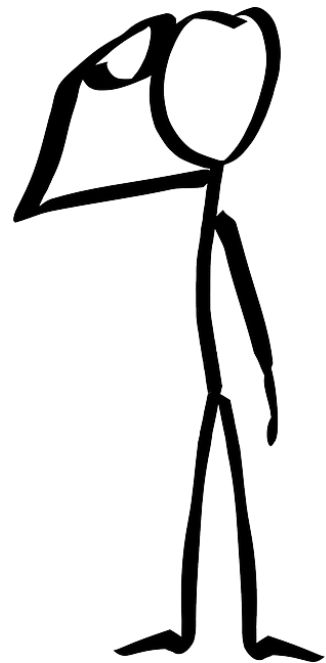
She couldn't pick them up, so she called Tom for help.

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

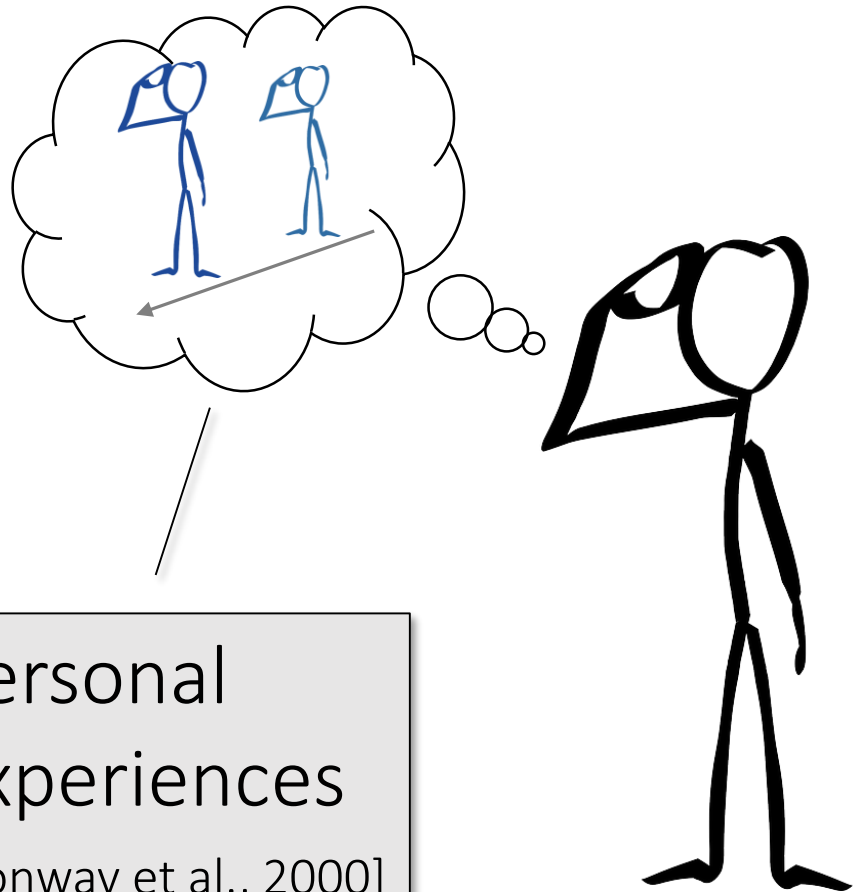
They realized that Tom broke her glasses by stepping on them.

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

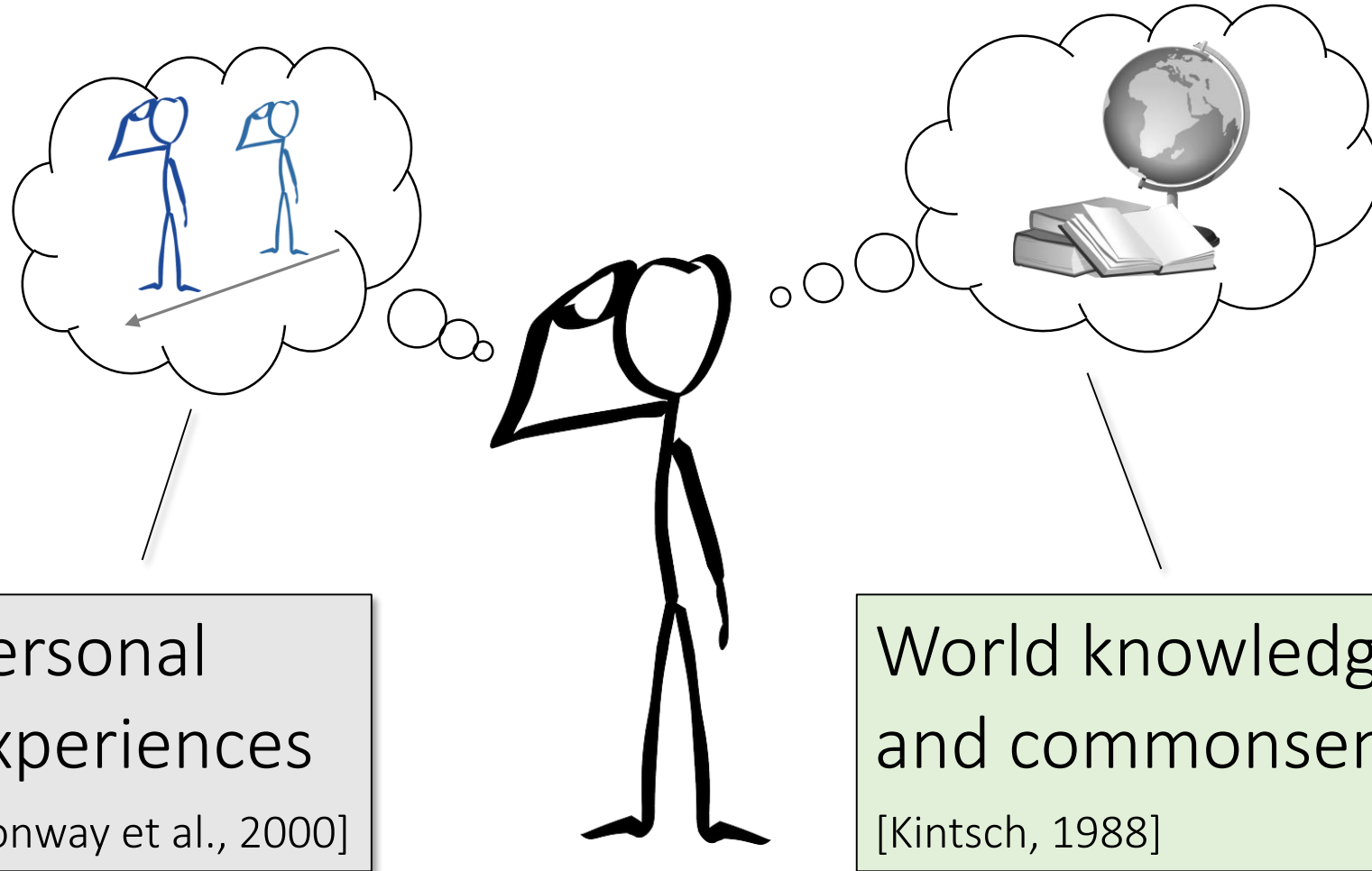
Humans reason about the world with
mental models [Graesser, 1994]



Humans reason about the world with **mental models** [Graesser, 1994]



Humans reason about the world with **mental models** [Graesser, 1994]



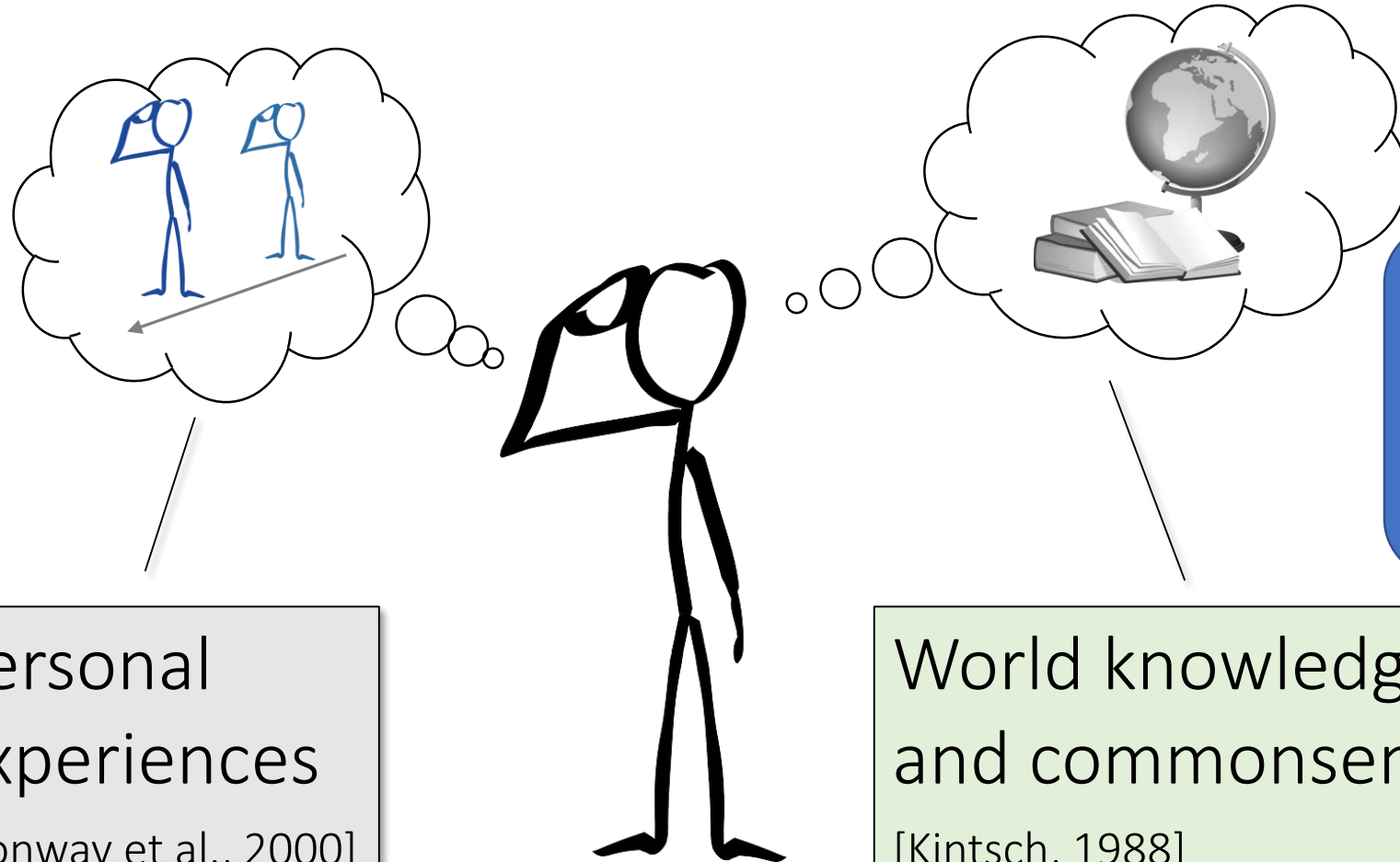
Personal
experiences

[Conway et al., 2000]

World knowledge
and commonsense

[Kintsch, 1988]

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.

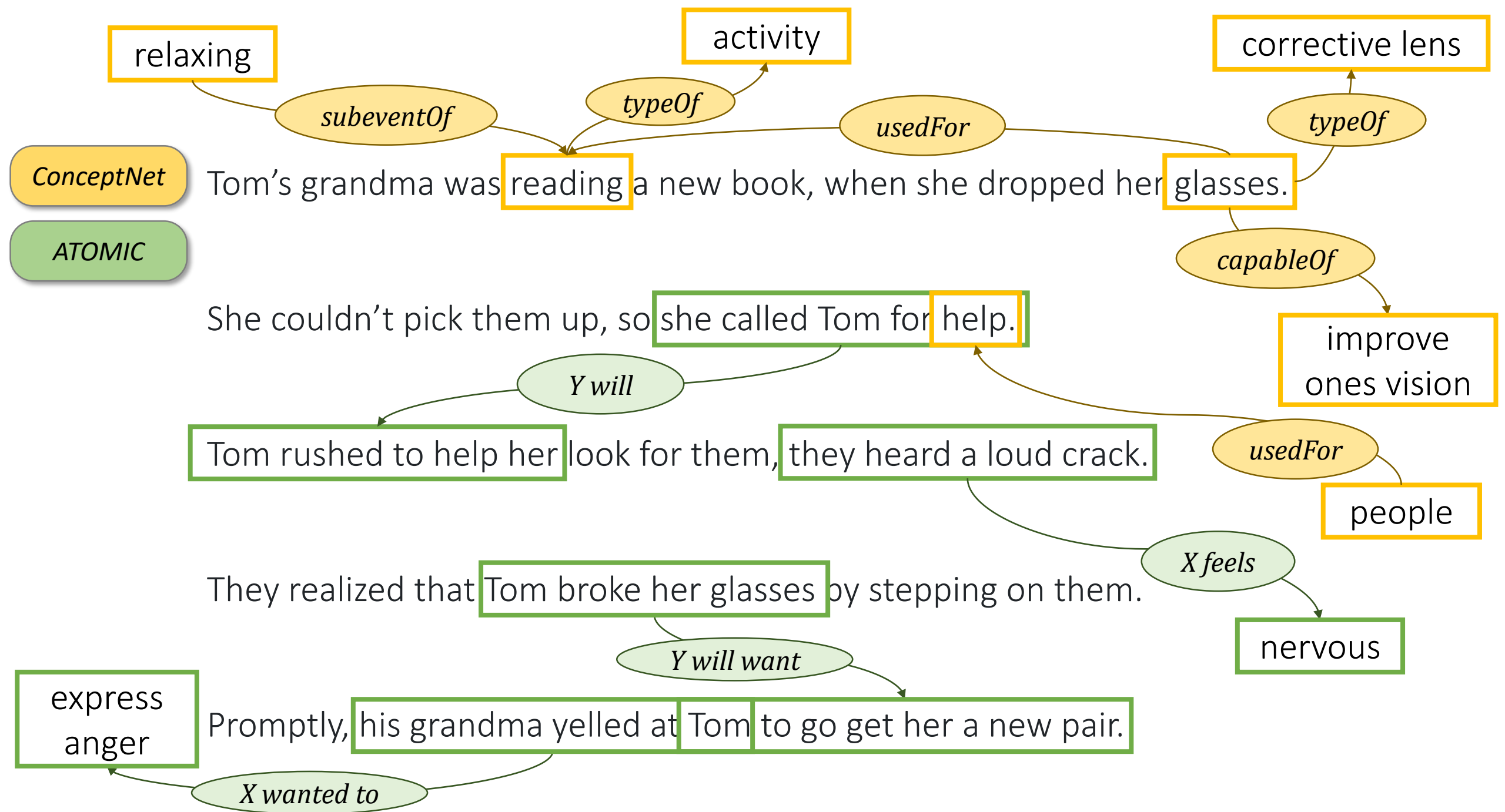
Y will

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.



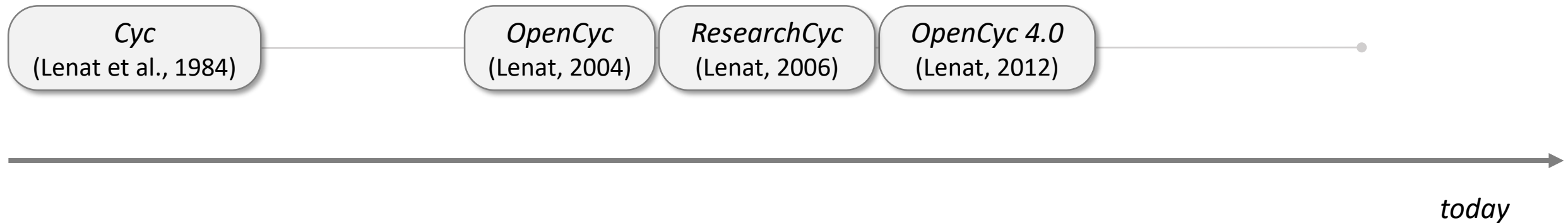
Overview of existing resources

Cyc
(Lenat et al., 1984)

today



Overview of existing resources



Overview of existing resources

Open Mind Common Sense
(Minsky, Singh & Havasi, 1999)

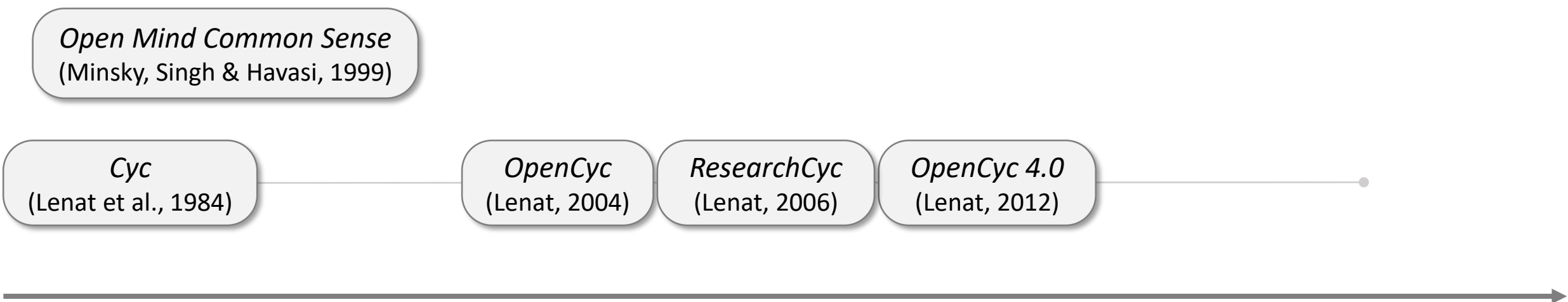
Cyc
(Lenat et al., 1984)

OpenCyc
(Lenat, 2004)

ResearchCyc
(Lenat, 2006)

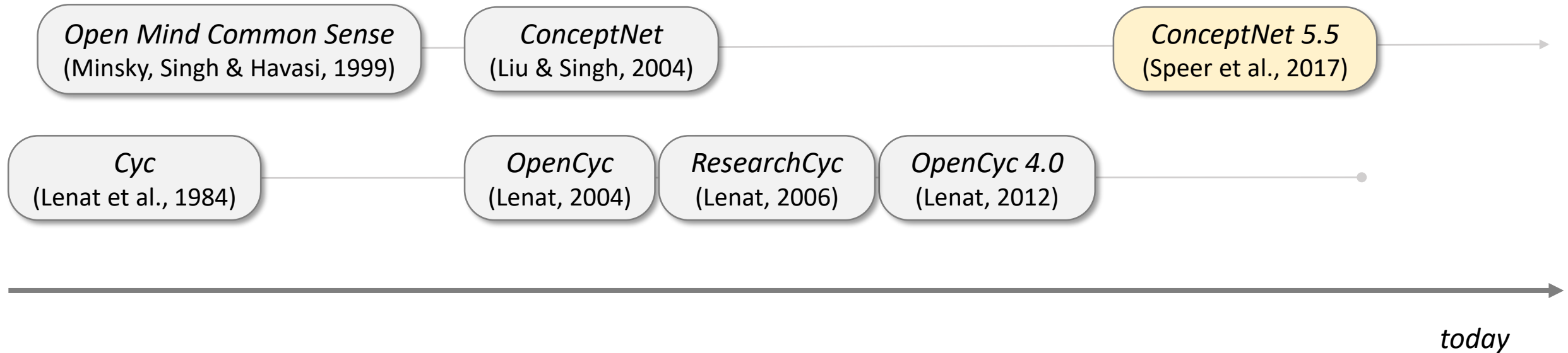
OpenCyc 4.0
(Lenat, 2012)

today

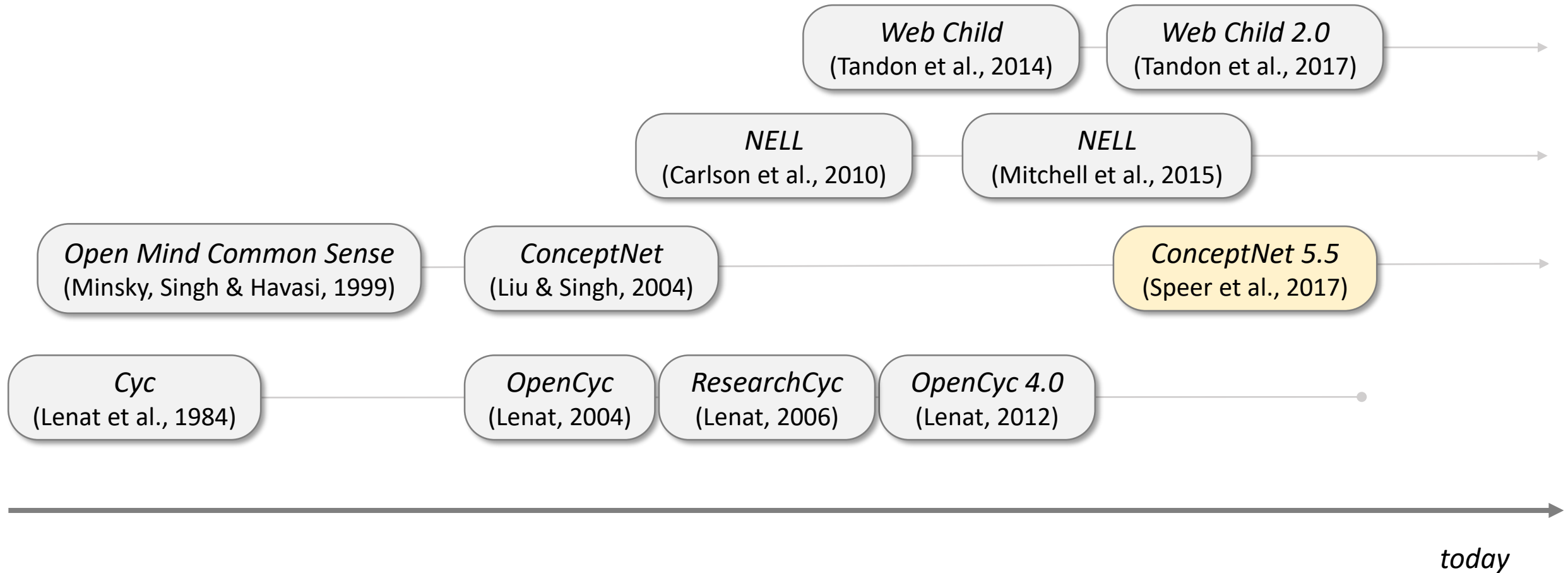


The diagram illustrates the evolution of Cyc knowledge bases over time. A horizontal timeline arrow at the bottom points to the right, labeled 'today'. Above the arrow, five rounded rectangular boxes are connected by a horizontal line. From left to right, the boxes contain: 'Cyc (Lenat et al., 1984)', 'OpenCyc (Lenat, 2004)', 'ResearchCyc (Lenat, 2006)', 'OpenCyc 4.0 (Lenat, 2012)', and a final box containing 'Open Mind Common Sense (Minsky, Singh & Havasi, 1999)'. The 'Open Mind Common Sense' box is positioned higher than the others and is not connected to the main timeline line.

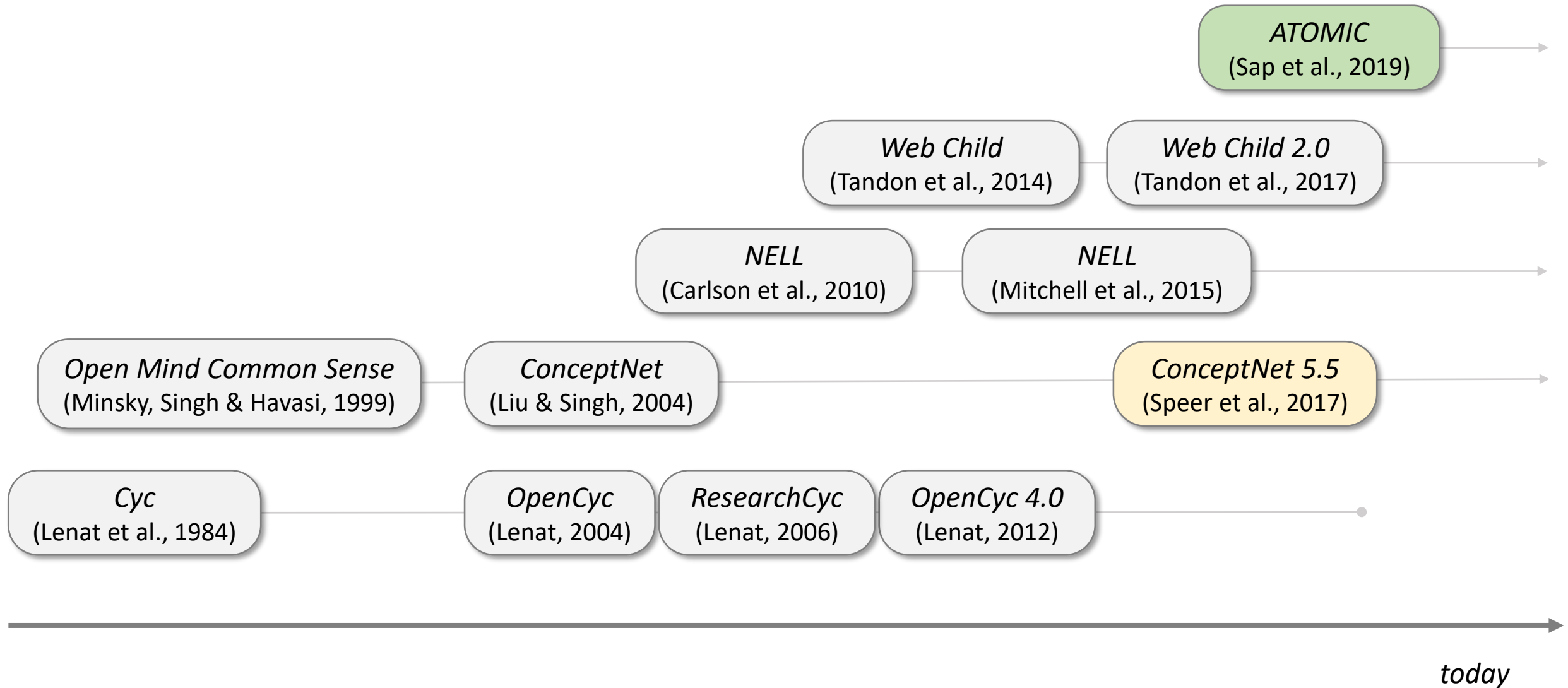
Overview of existing resources



Overview of existing resources



Overview of existing resources



How do you create a commonsense resource?

Desiderata for a good commonsense resource

Coverage

- Large scale
- Diverse knowledge types

Useful

- High quality knowledge
- Usable in downstream tasks

Desiderata for a good commonsense resource

Coverage

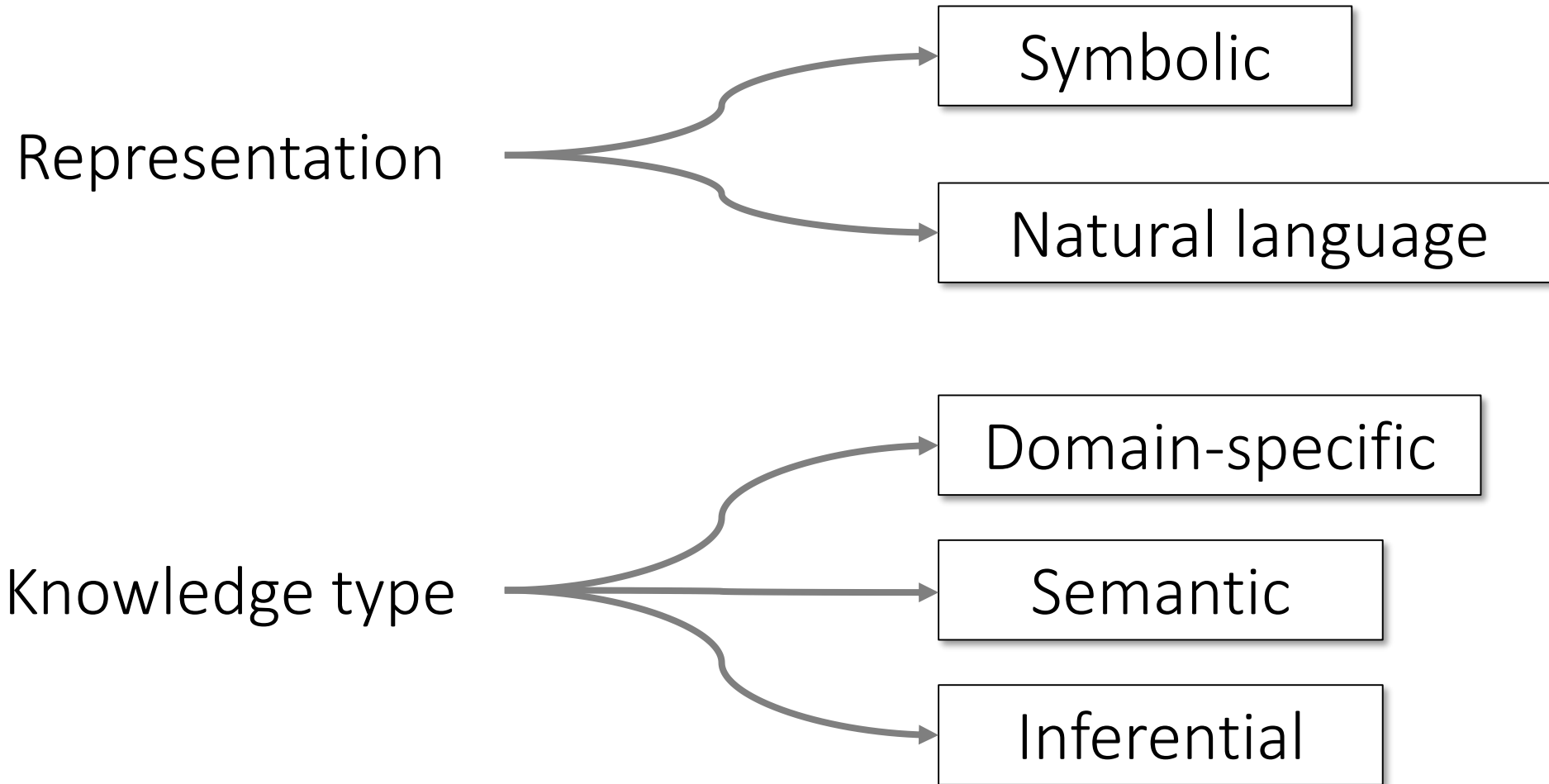
- Large scale
- Diverse knowledge types

Useful

- High quality knowledge
- Usable in downstream tasks

Multiple resources tackle different
knowledge types

Creating a commonsense resource



CONCEPTNET:

semantic knowledge in natural language form

en reading

An English term in ConceptNet 5.8

reading is a subevent
of...

en you learn →

en turning a page →

en learning →

en **reading**

An English term in ConceptNet 5.8

Related terms

en book →

en books →

en book →

reading is a subevent
of...

en you learn →

en turning a page →

en learning →

en **reading**

An English term in ConceptNet 5.8

Related terms

en book →

en books →

en book →

Effects of reading

en learning →

en ideas →

en a headache →

reading is a subevent
of...

en you learn →

en turning a page →

en learning →

en **reading**

An English term in ConceptNet 5.8

Related terms

- en** book →
- en** books →
- en** book →

Effects of reading

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

reading is a subevent of...

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

en **reading**

An English term in ConceptNet 5.8

reading is a type of...

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

Related terms

en book →

en books →

en book →

Effects of reading

en learning →

en ideas →

en a headache →

reading is a subevent of...

en you learn →

en turning a page →

en learning →

en **reading**

An English term in ConceptNet 5.8

reading is a type of...

en an activity →

en a good way to learn →

en one way of learning →

en one way to learn →

Types of reading

en browse (n, communication) →

en bumf (n, communication) →

en clock time (n, time) →

en miles per hour (n, time) →

Related terms

en book →

en books →

en book →

Effects of reading

en learning →

en ideas →

en a headache →

reading is a subevent of...

en you learn →

en turning a page →

en learning →

en **reading**

An English term in ConceptNet 5.8

reading is a type of...

en an activity →

en a good way to learn →

en one way of learning →

en one way to learn →

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) →

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 →

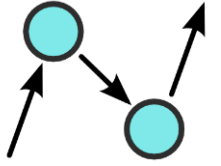
Things used for reading

- 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?

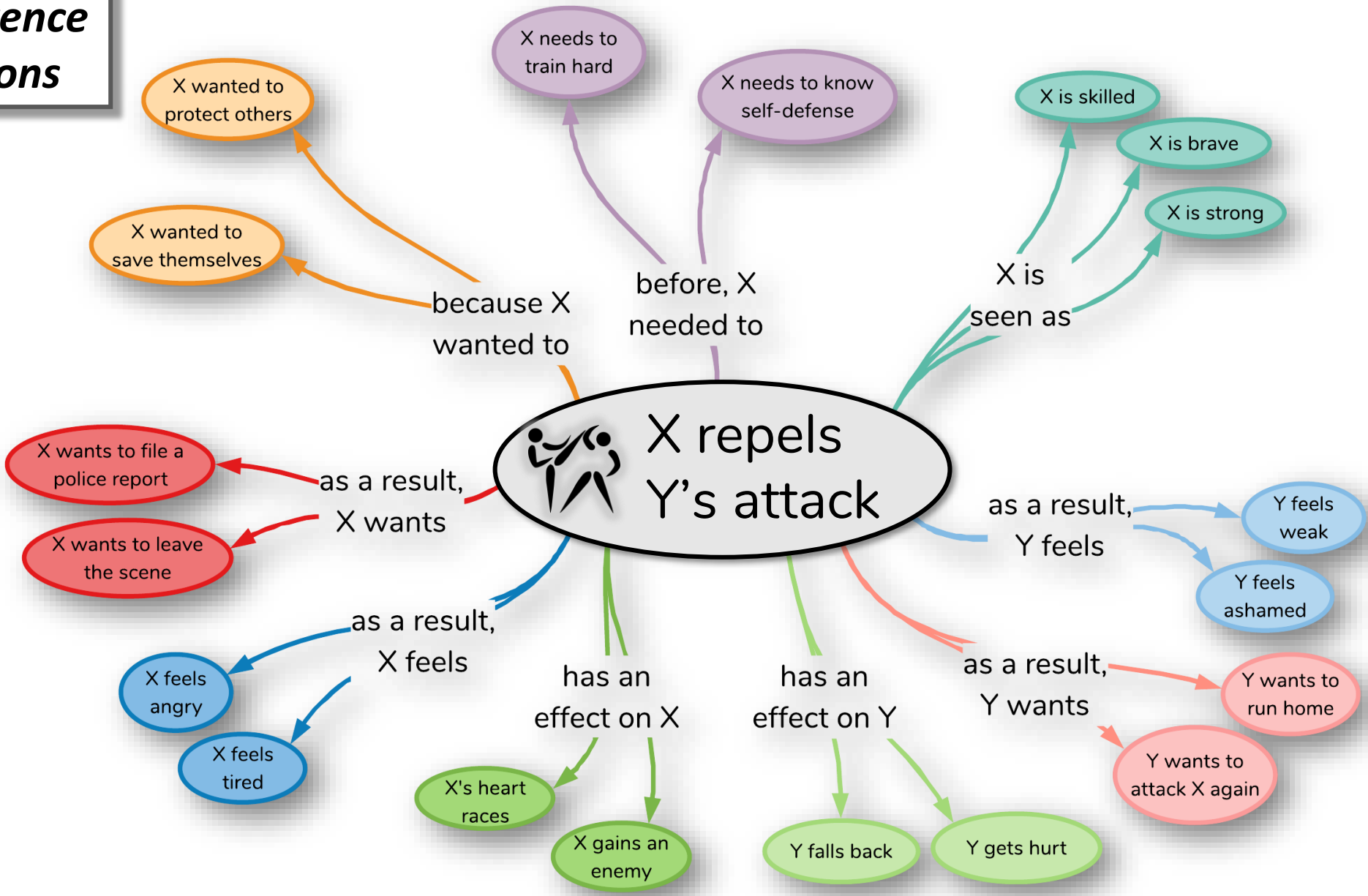


- General commonsense knowledge
- 21 million edges and over 8 million nodes (as of 2017)
 - Over 85 languages
 - In English: over 1.5 million nodes
- Knowledge covered:
 - Open Mind Commonsense assertions
 - Wikipedia/Wiktionary semantic knowledge
 - WordNet, Cyc ontological knowledge

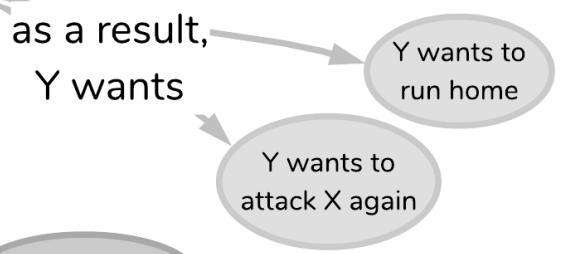
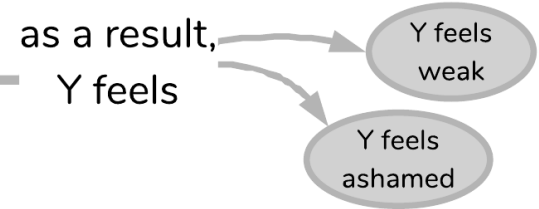
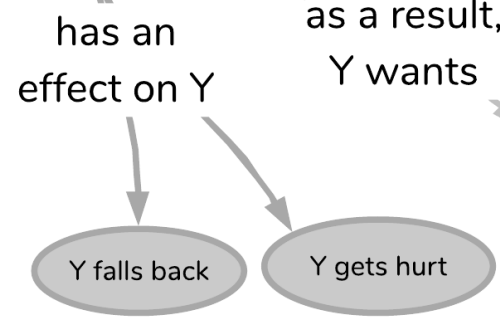
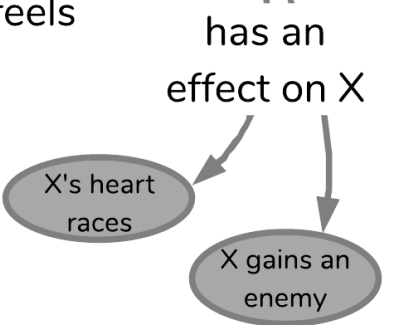
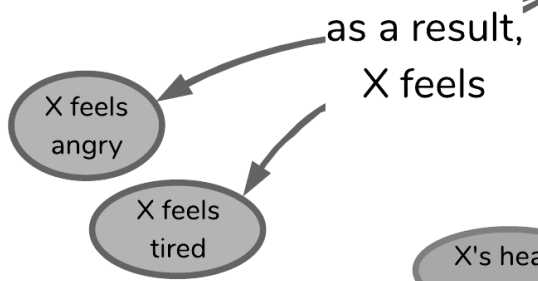
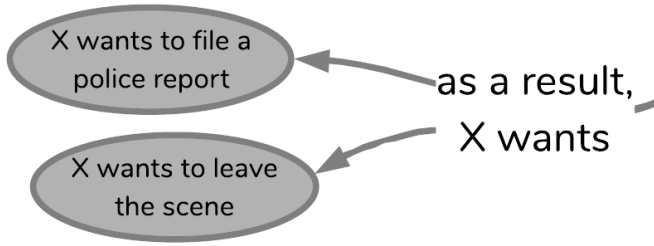
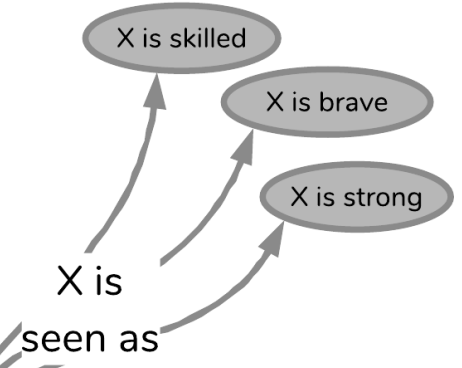
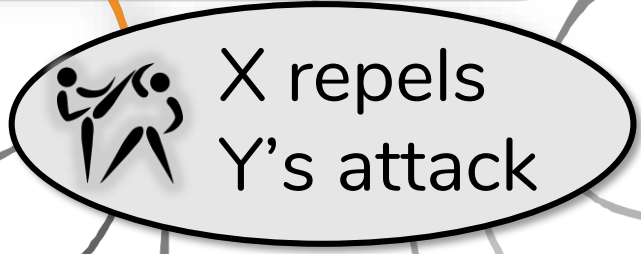
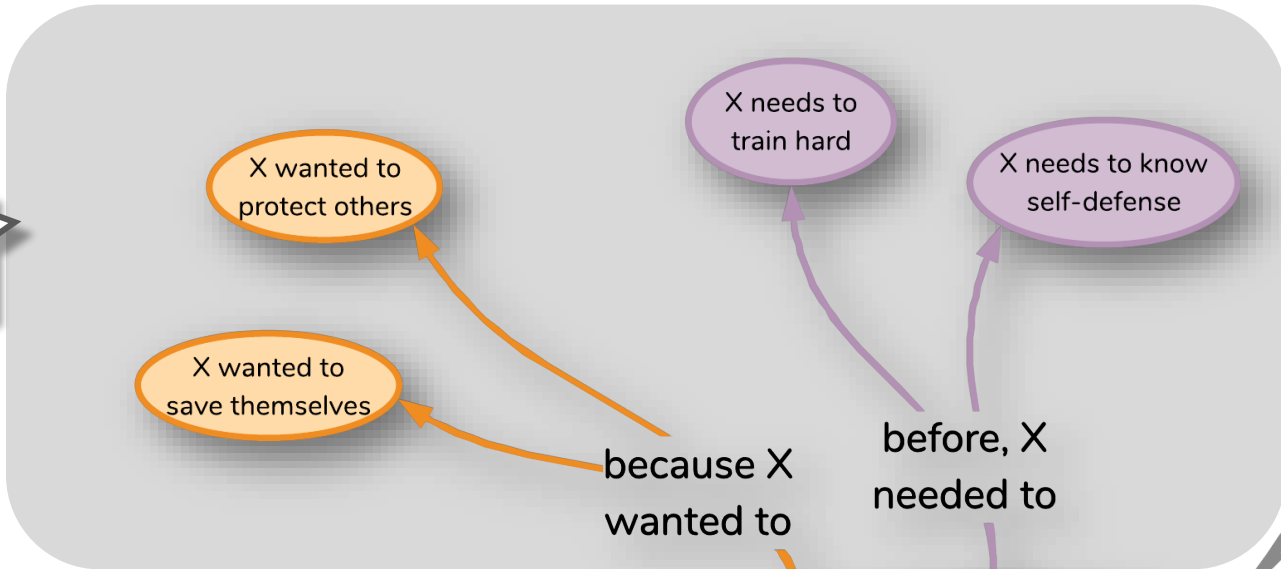
ATOMIC:

inferential knowledge in natural language form

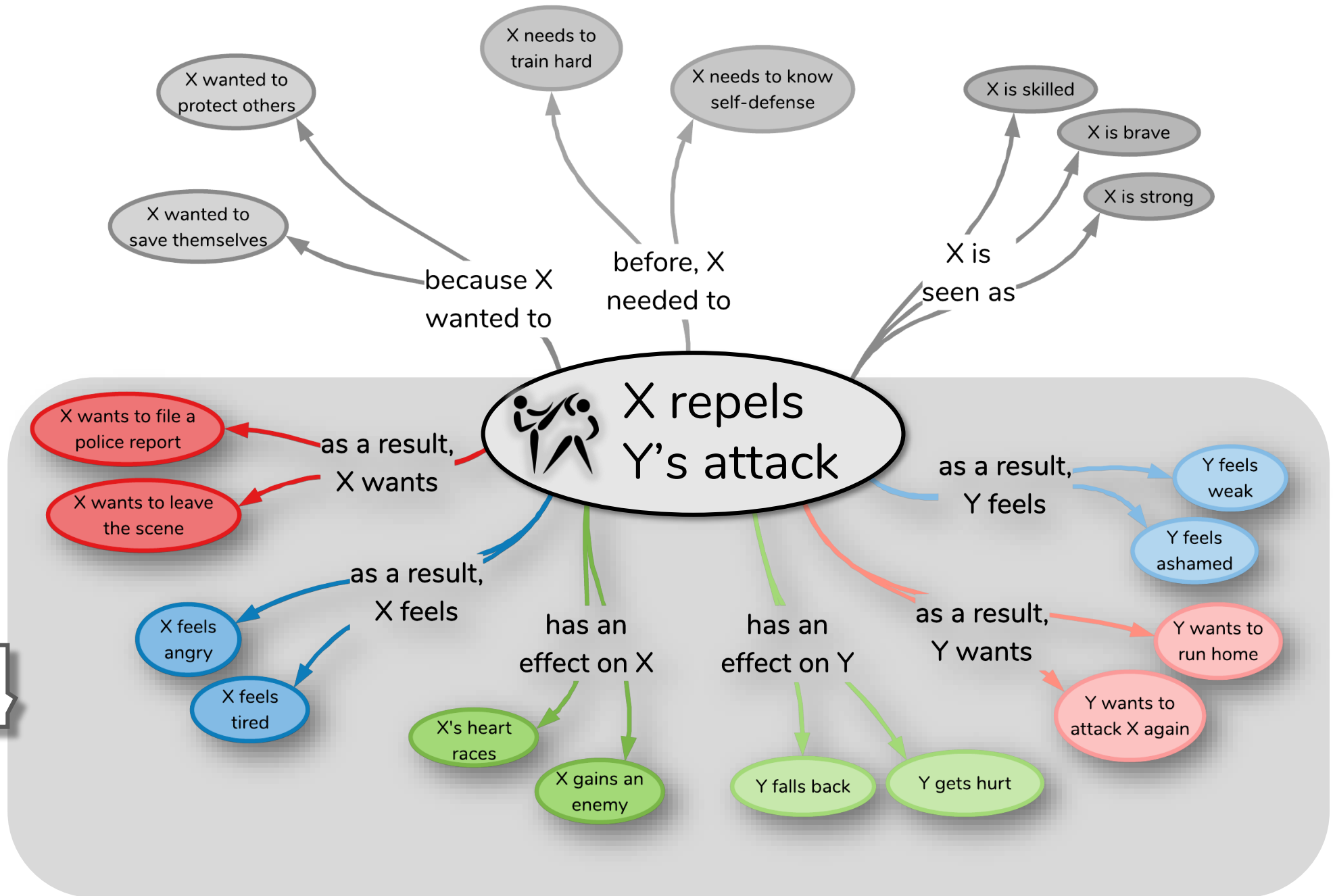
nine inference dimensions



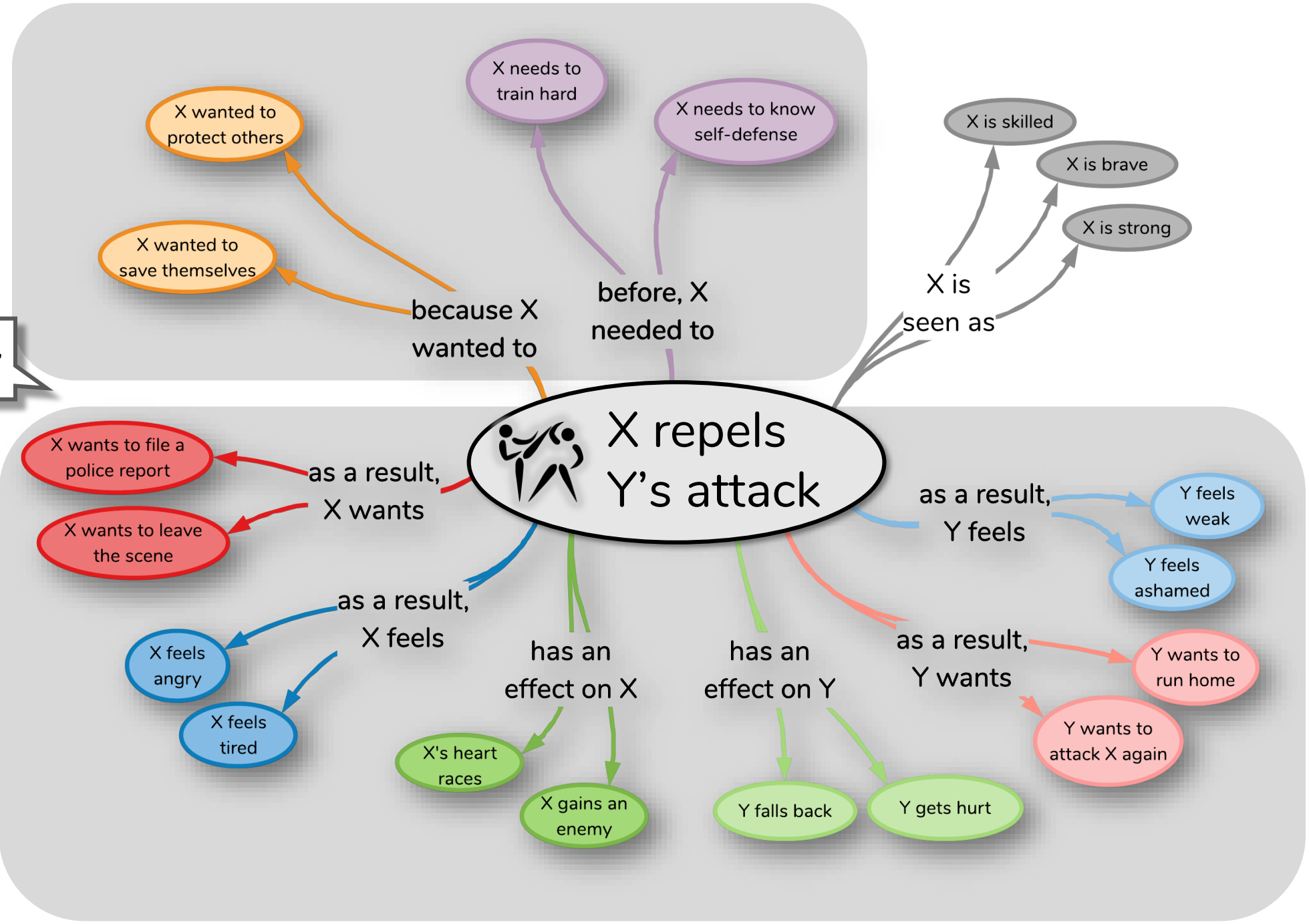
Causes

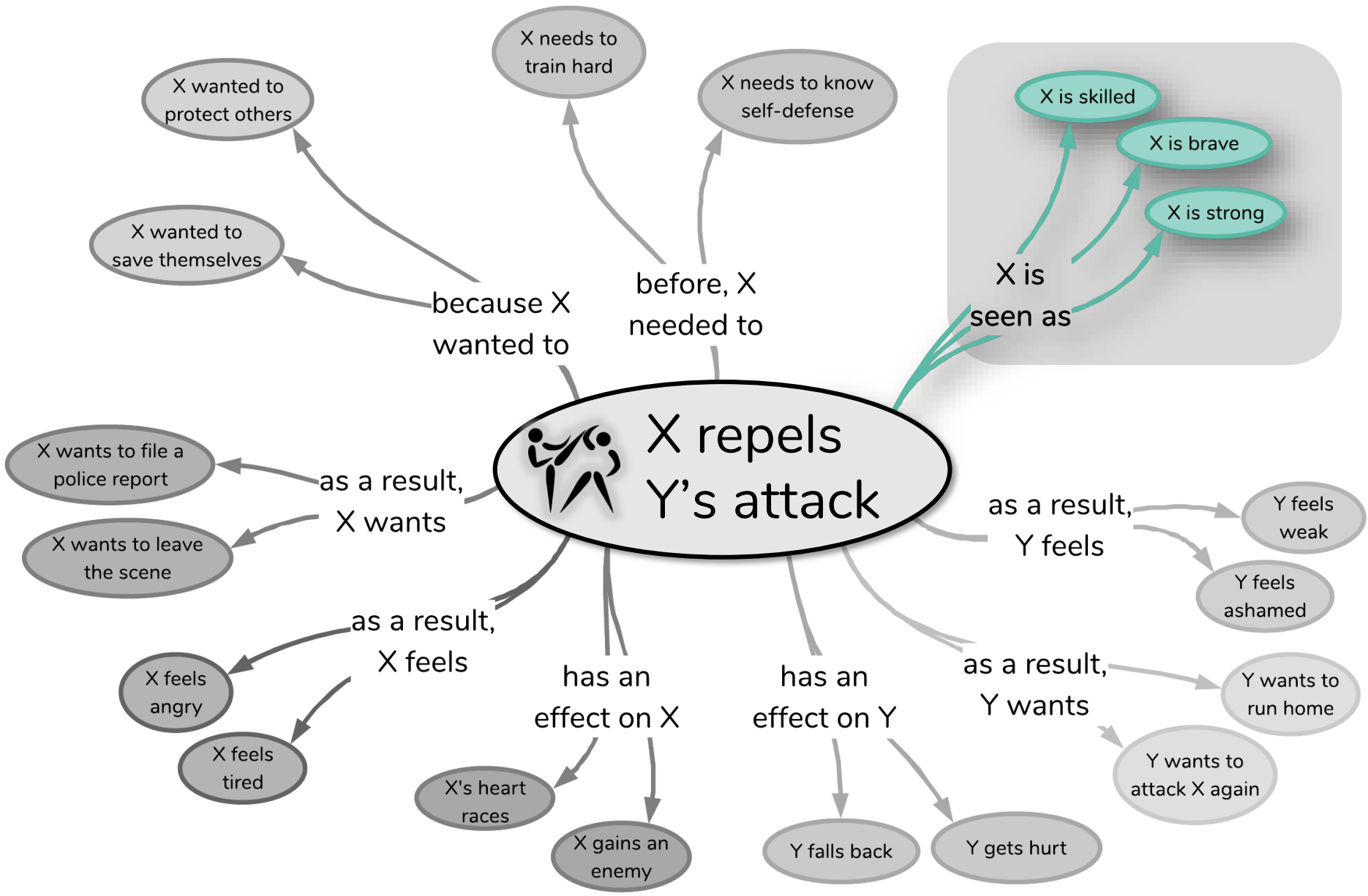


Effects



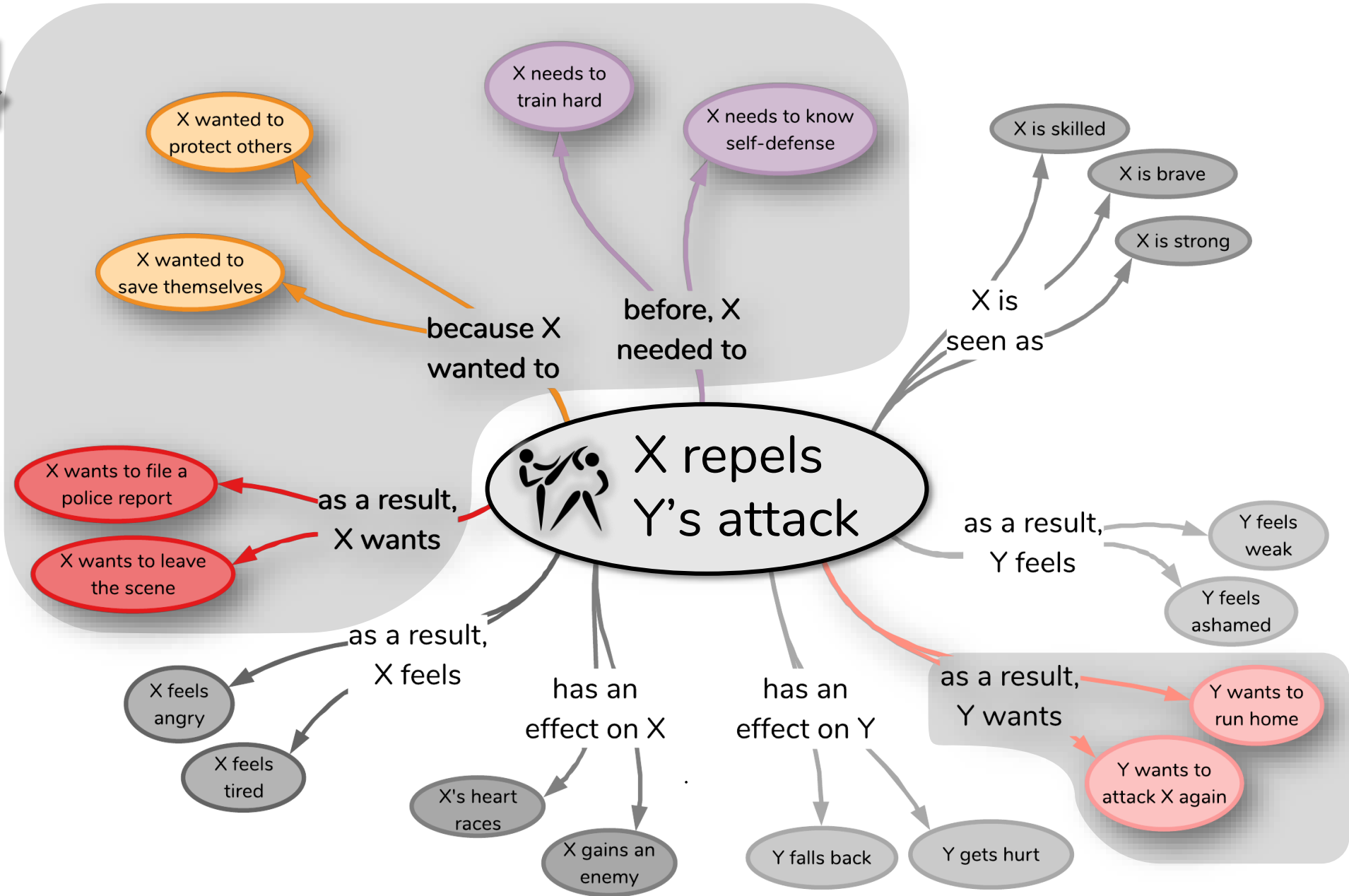
Dynamic

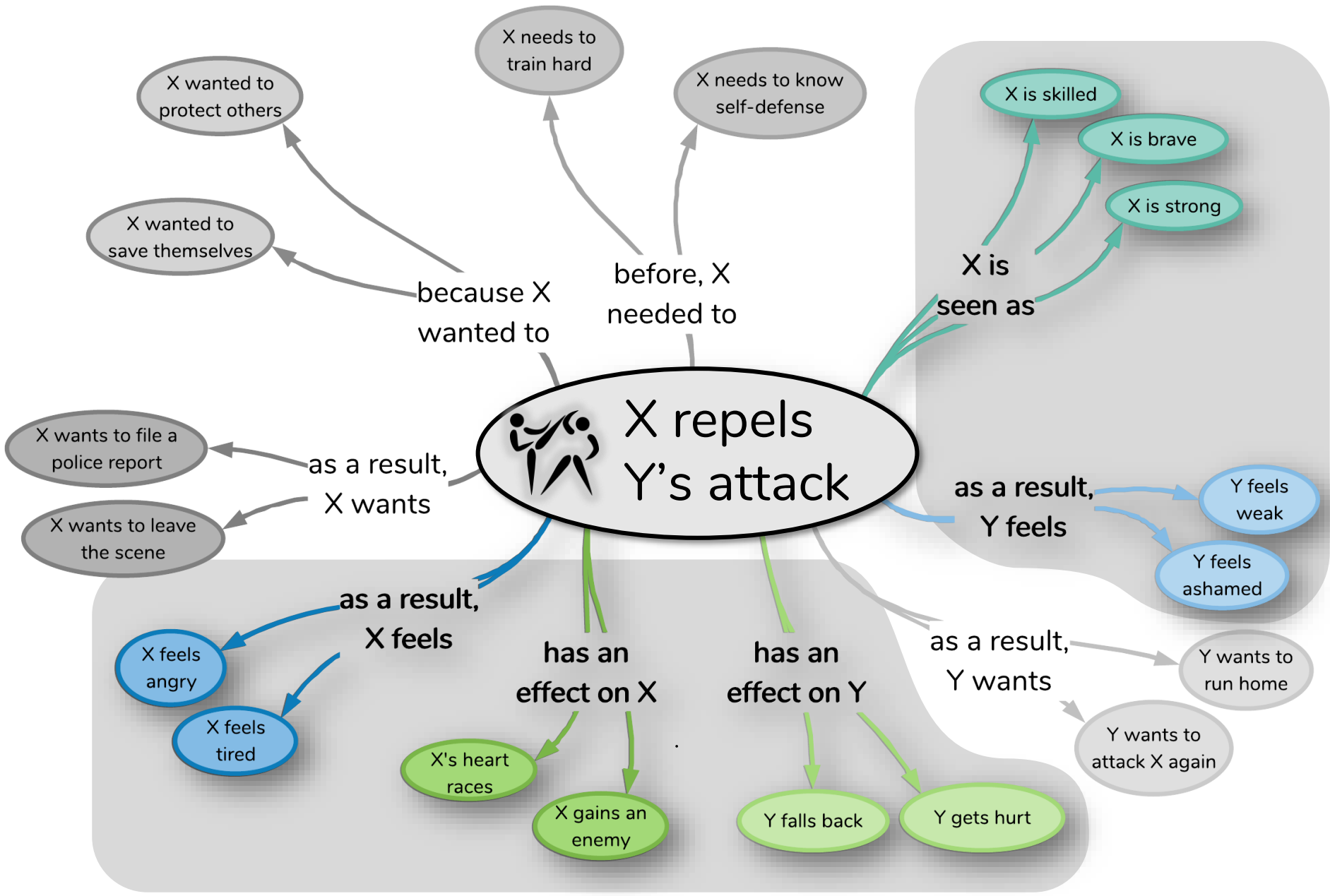




Static

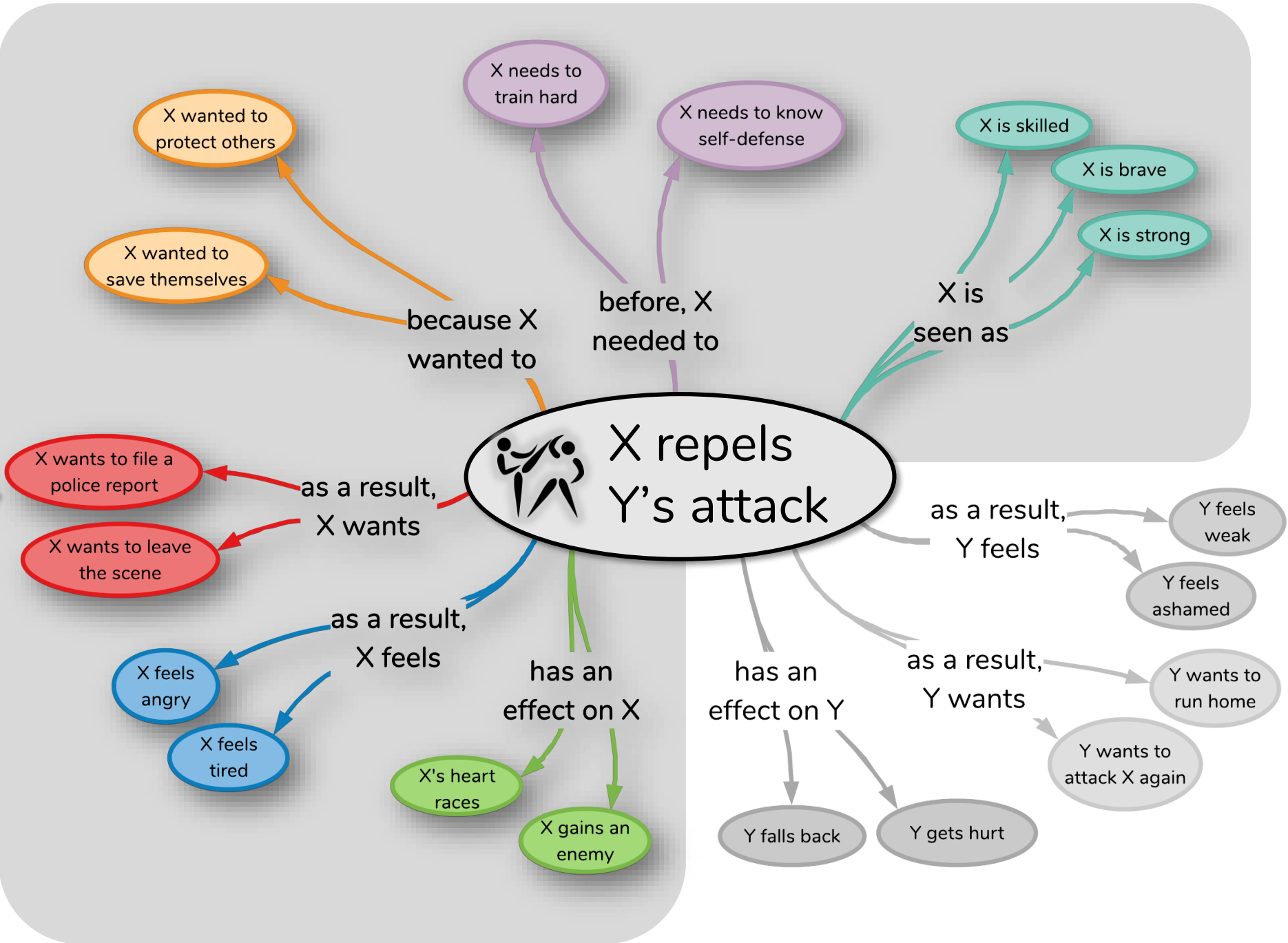
Voluntary

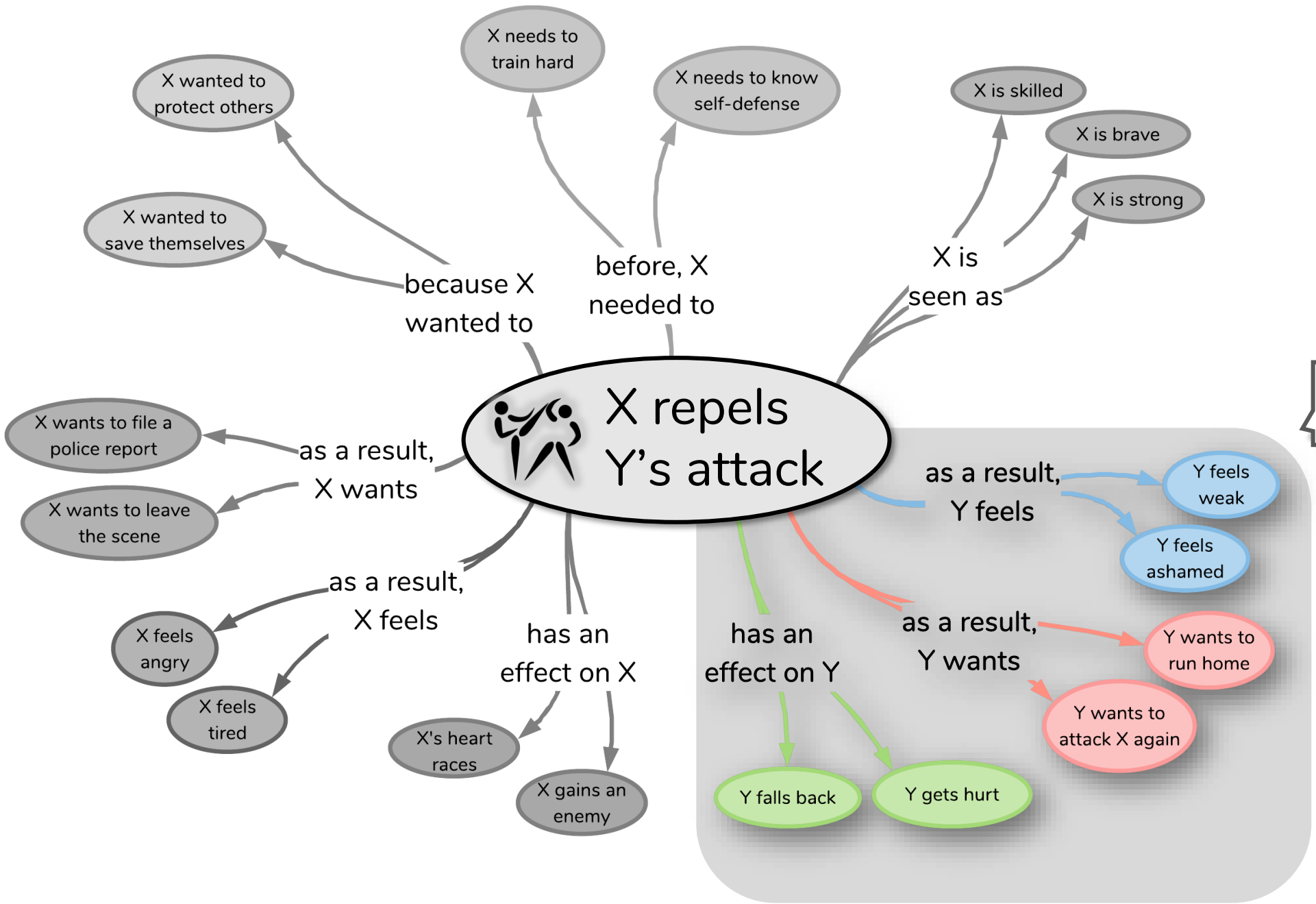




Involuntary

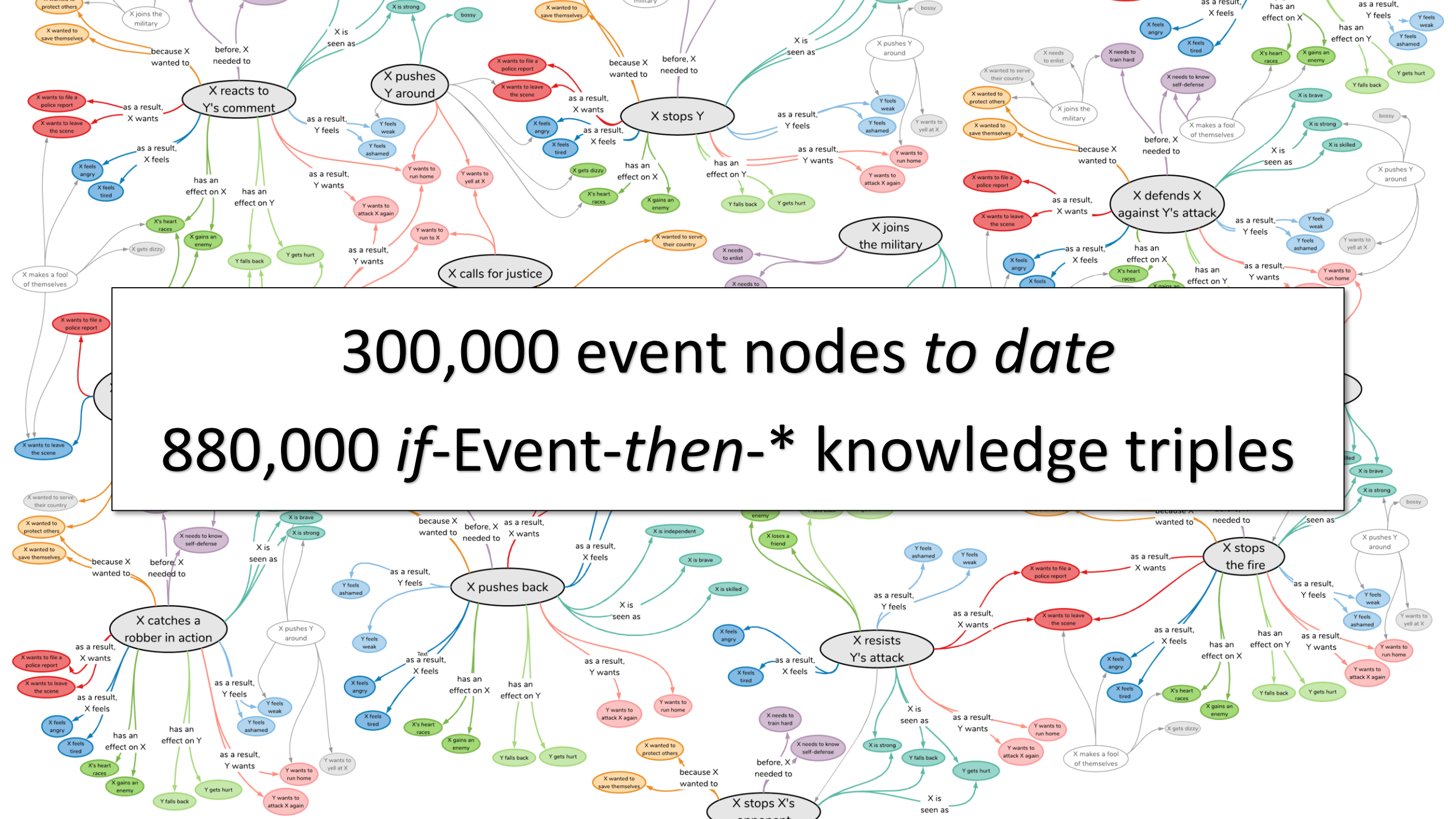
Agent





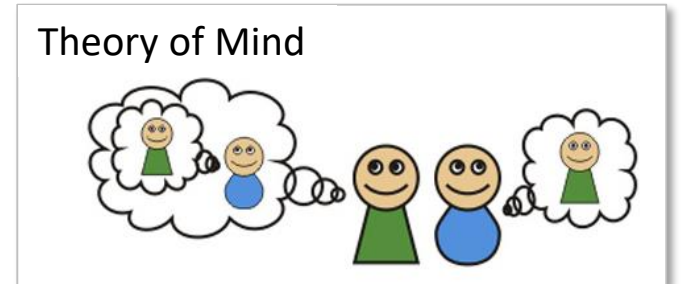
Theme

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



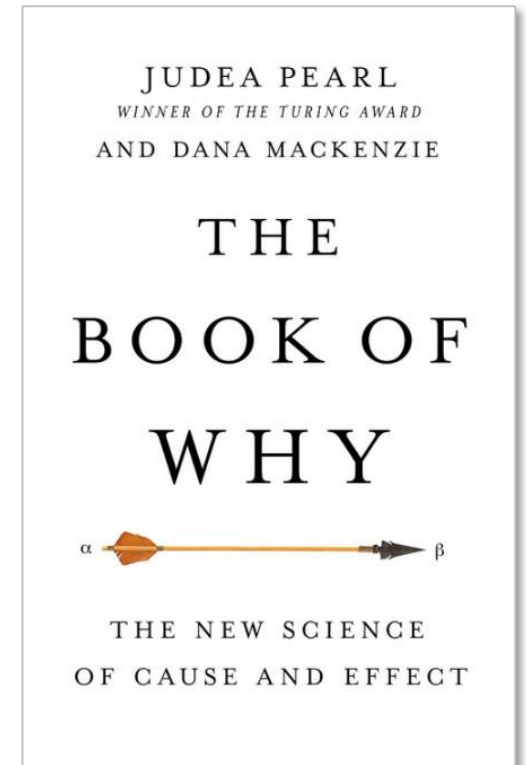
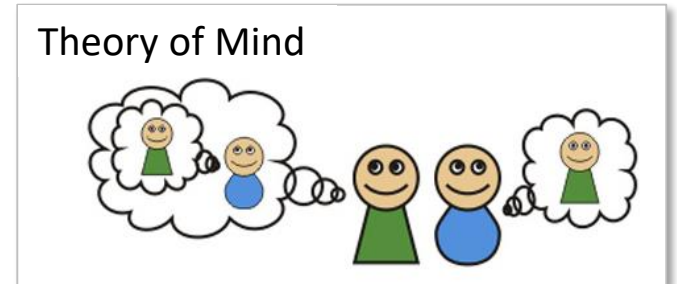
ATOMIC: knowledge of *cause* and *effect*

- Humans have **theory of mind**, allowing us to
 - make inferences about **people's mental states**
 - understand **likely events** that precede and follow (Moore, 2013)

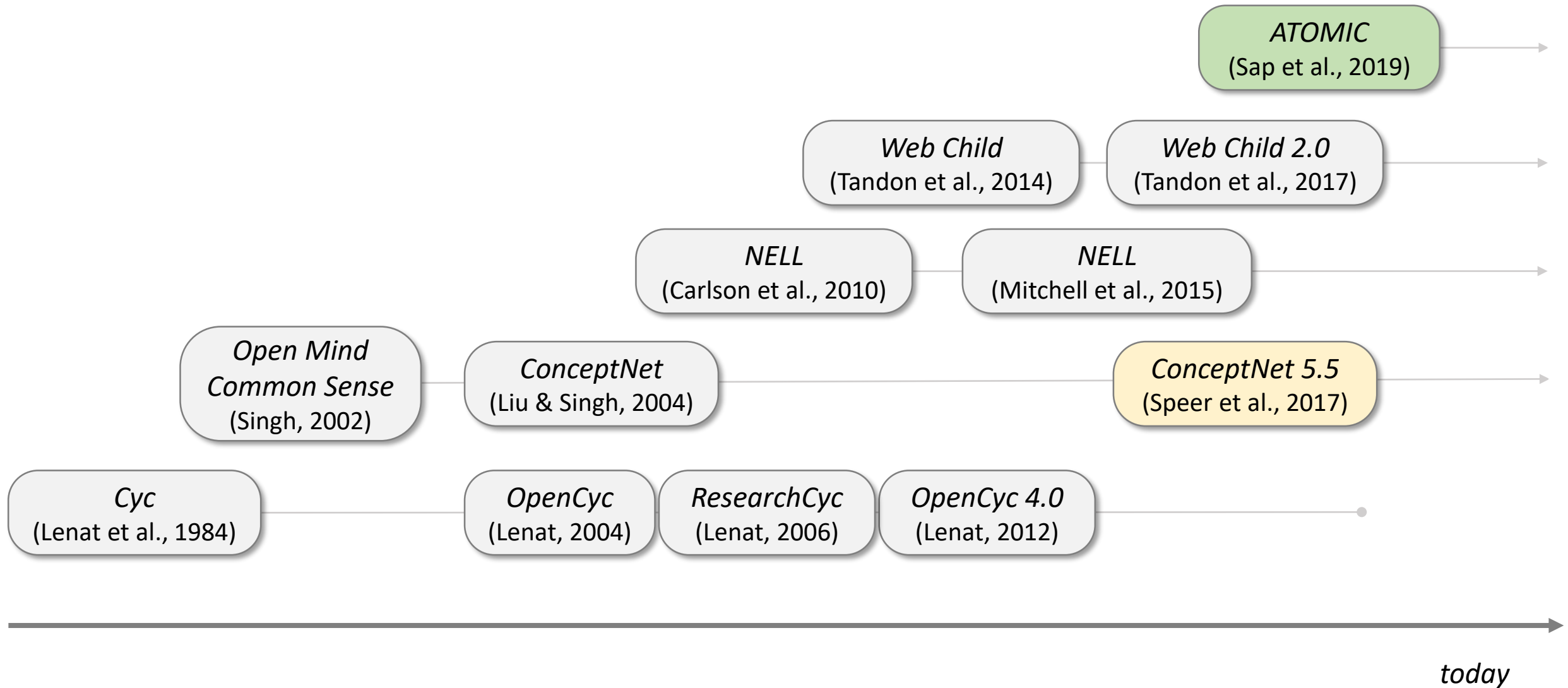


ATOMIC: knowledge of *cause* and *effect*

- 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)

NELL
(Mitchell et al., 2015)

OpenCyc 4.0
(Lenat, 2012)

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

ConceptNet 5.5
(Speer et al., 2017)

ATOMIC
(Sap et al., 2019)

Existing knowledge bases

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

NELL

(Mitchell et al., 2015)

OpenCyc 4.0

(Lenat, 2012)

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

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

ConceptNet 5.5

(Speer et al., 2017)

ATOMIC

(Sap et al., 2019)

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

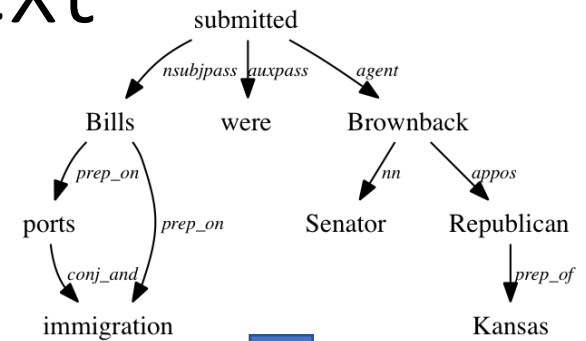
1. Read and parse text
2. Create candidate rules
3. Filter rules based on quality metric

Advantage:
can extract knowledge automatically

Example system:

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

... more on this later with temporal commonsense



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

Some commonsense cannot be extracted

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

- Idioms & figurative usage
“Black sheep problem”
- 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)

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)

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)

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

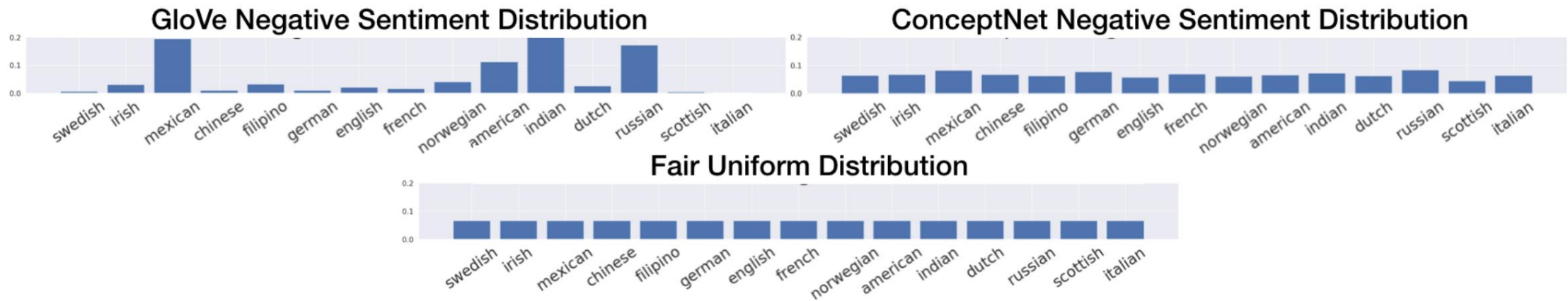
- Natural language
 - Accessible to non-experts
 - Different phrasings allow for more nuanced knowledge
- Fast and scalable collection
 - Crowdsourcing
 - Games with a purpose

ATOMIC
(Sap et al., 2019)

ConceptNet 5.5
(Speer et al., 2017)

Knowledge bases and mitigating biases

- Different data collection methods suffer from social biases differently
- ConceptNet word embeddings have less demographic biases than GloVe embeddings [Sweeney & Najafian, 2019]



Knowledge bases and mitigating biases

PersonX clutches a gun

ATOMIC (Sap et al., 2019)

because X
wanted to

to be safe

to protect himself

to protect themselves

to defend themselves

to defend himself

Knowledge bases and mitigating biases

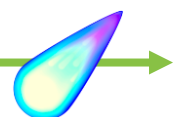
PersonX clutches a gun

ATOMIC (Sap et al., 2019)

because X
wanted to

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

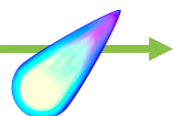
Jaquain clutches a gun



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

What's next with commonsense resources?

- Use them with models in downstream tasks
 - Reading comprehension, QA tasks, etc.
- Create inference or reasoning engines
 - Knowledge base construction, multi-hop reasoning, etc.



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

Thanks! Questions?

people

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

