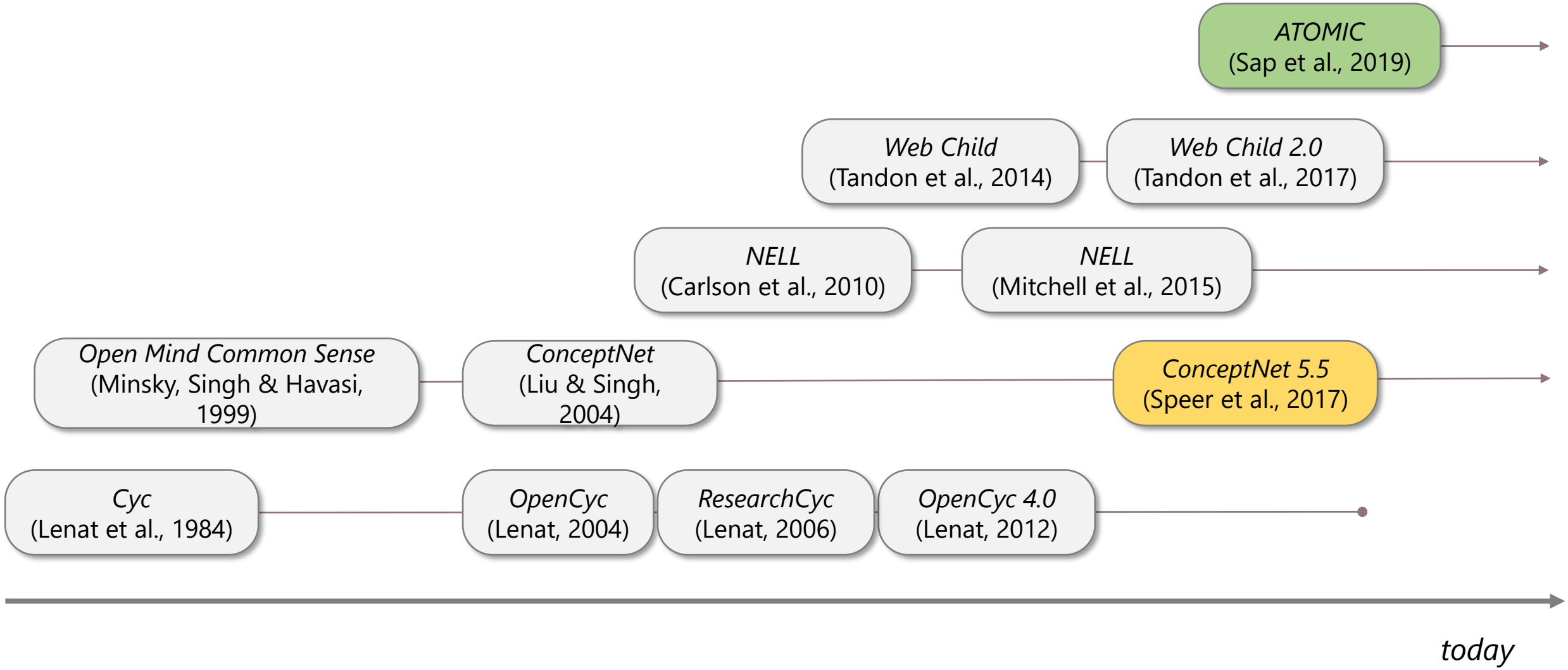
Knowledge Representations + Schemata

CIS-700 Interactive Fiction and Text Generation
Module 4 - 3/3/2022
Dr. Lara J. Martin

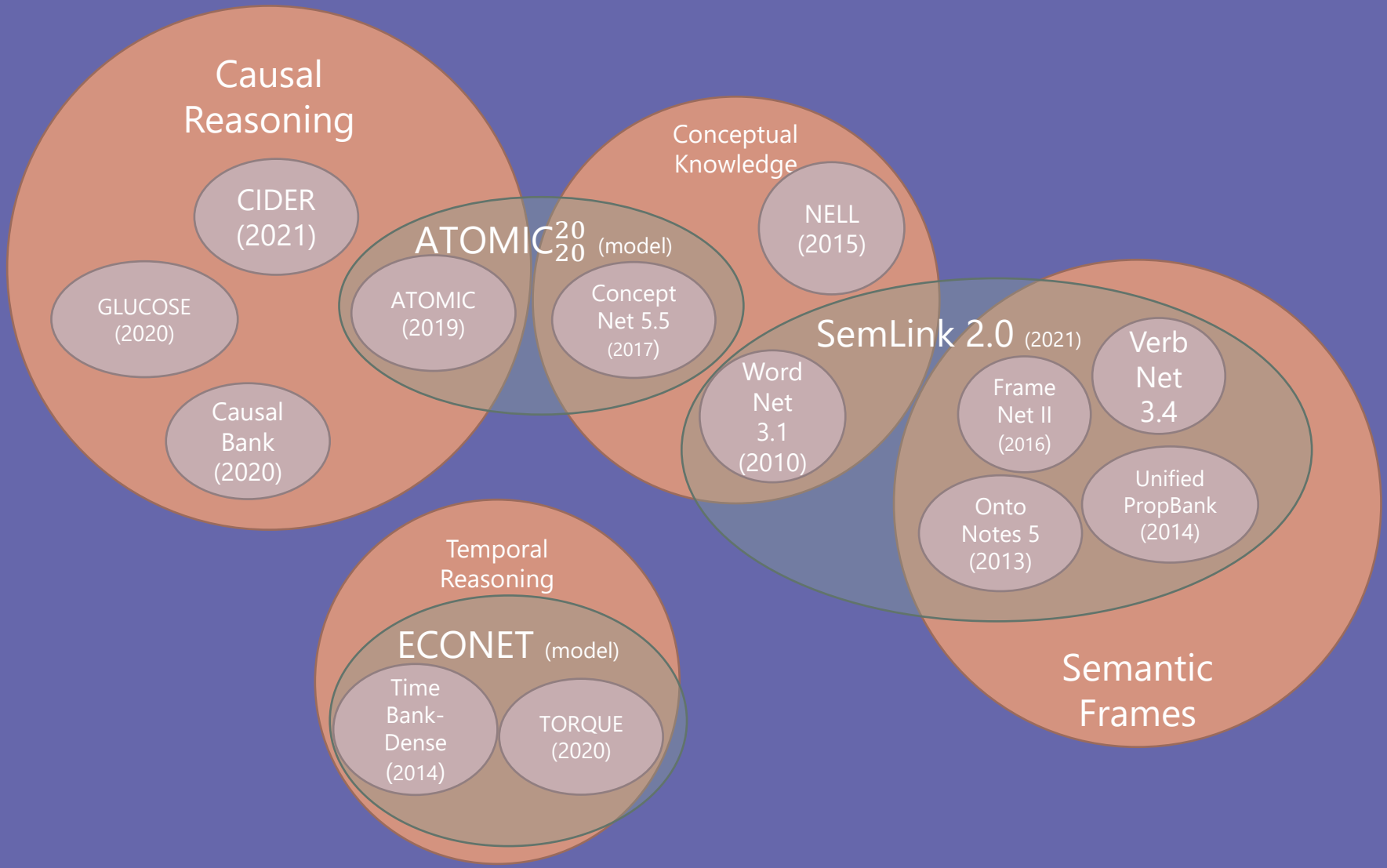
Learning Objectives

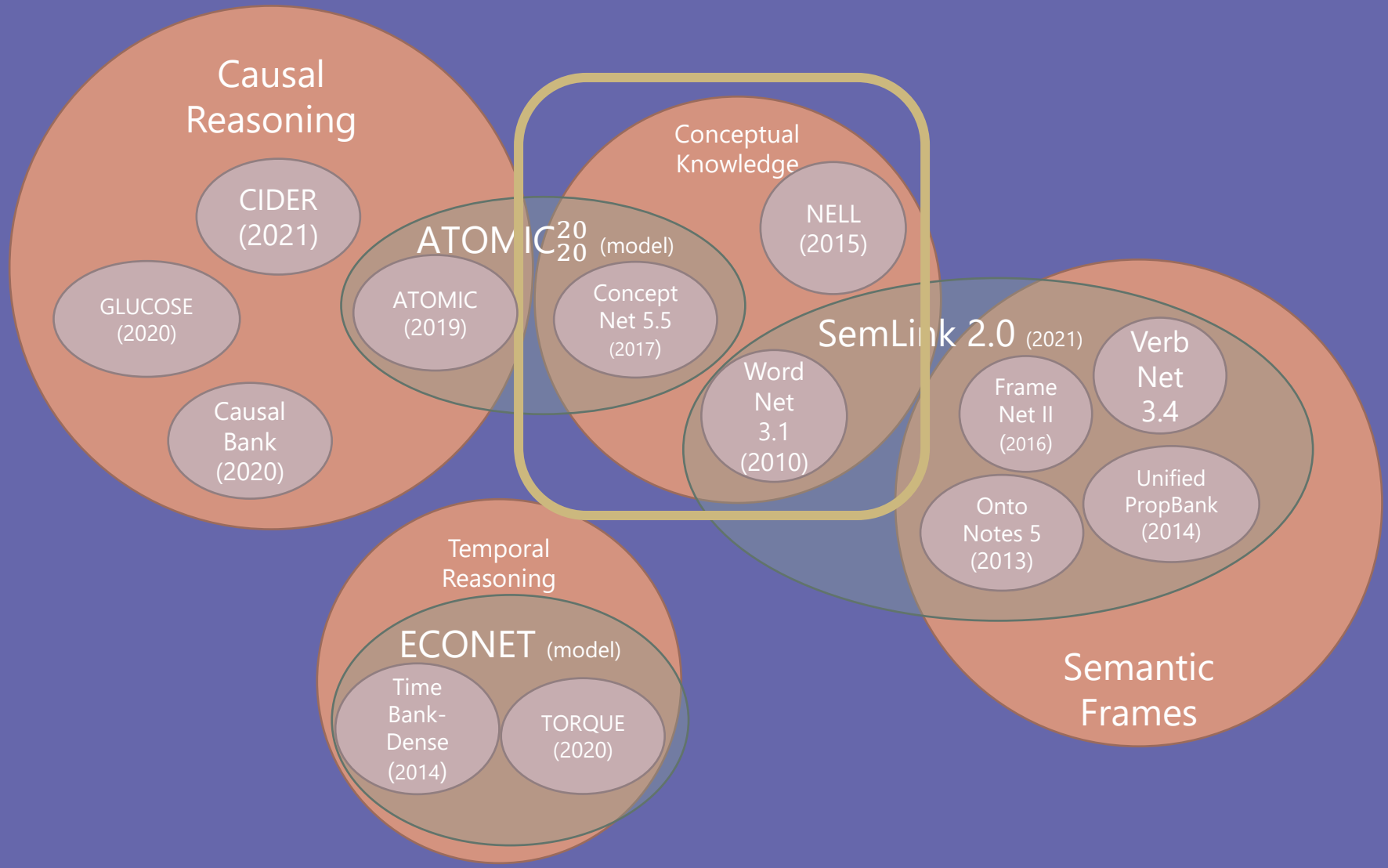
- Recognize what each knowledge representation can be used for and when to apply them
- Note where to find the knowledge databases, how to access them, and call their APIs
- Manipulate data from knowledge representations to form schemas

History of Knowledge Bases



From slides for the ACL 2020 Commonsense Tutorial by Yejin Choi, Vered Shwartz, Maarten Sap, Antoine Bosselut, and Dan Roth


Causal Reasoning
CIDER (2021)
GLUCOSE (2020)
Causal Bank (2020)
ATOMIC$^{20}_{20}$ (model)
ATOMIC (2019)
Conceptual Knowledge
Concept Net 5.5 (2017)
NELL (2015)
SemLink 2.0 (2021)
Word Net 3.1 (2010)
Frame Net II (2016)
Verb Net 3.4
Onto Notes 5 (2013)
Unified PropBank (2014)
Semantic Frames
Temporal Reasoning
ECONET (model)
Time Bank-Dense (2014)
TORQUE (2020)


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

Conceptual Knowledge

Concepts

A mental representation/reference for something grounded in the real world

- Can be explained via a set of features (e.g. a CAT is a carnivore with fur, sharp claws, long tail, …)
- Can be abstract (e.g. UNICORN)
- Can be composed to create more complex concepts (e.g. black cat = BLACK + CAT)

Lexical concepts

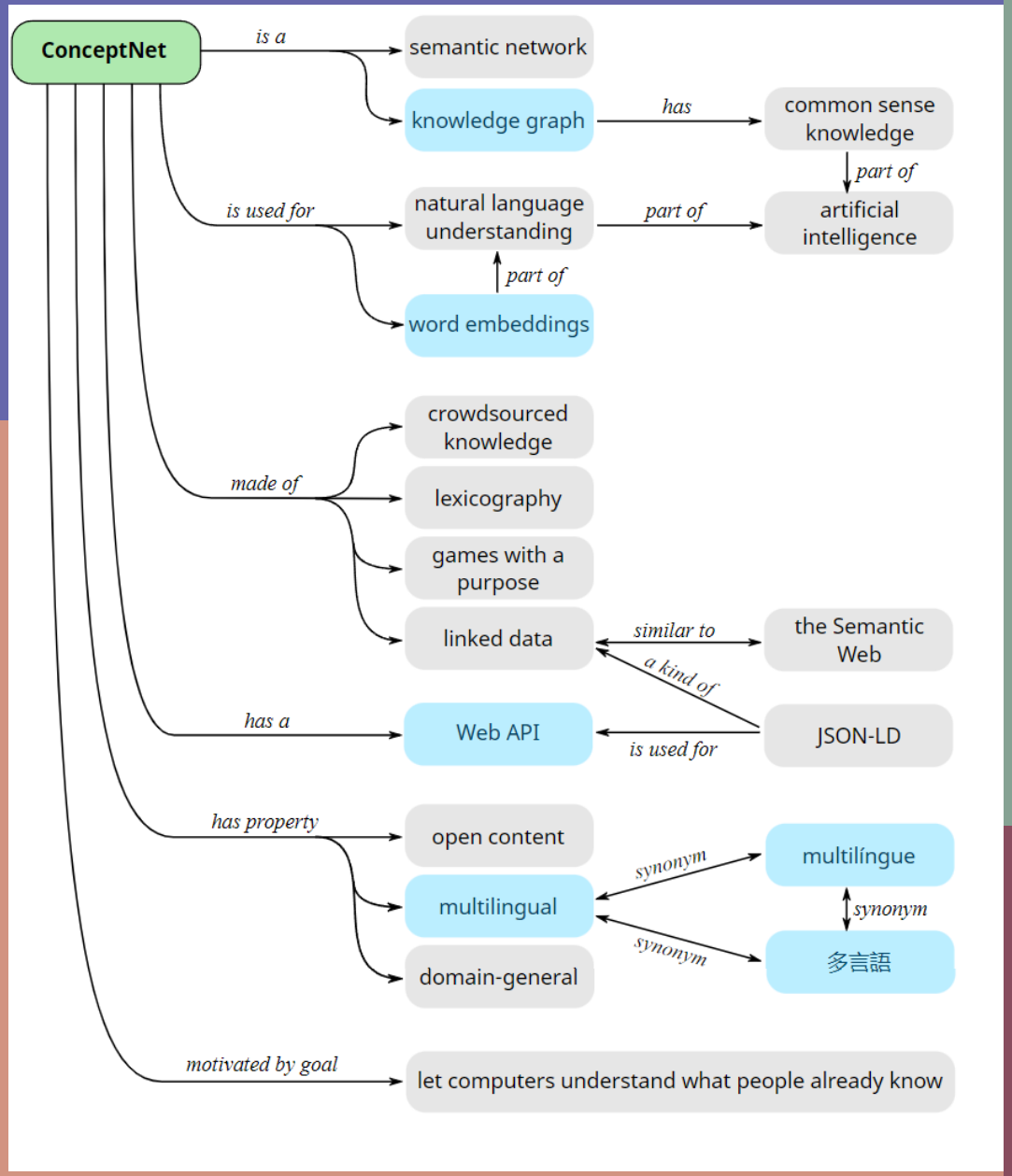
- Words to refer to ideas

Margolis, E. E., & Laurence, S. E. (1999). *Concepts: Core Readings*. The MIT Press.

ConceptNet 5.5

https://conceptnet.io/

Data Source: crowdsourced + other sources like Wikis, OpenCyc
(Core) Languages: English, French, Italian, German, Spanish, Russian, Portuguese, Japanese, Dutch, Chinese
Use: python package, AWS, raw data



R. Speer, J. Chin, and C. Havasi, "ConceptNet 5.5: An Open Multilingual Graph of General Knowledge," in *AAAI Conference on Artificial Intelligence (AAAI)*, 2017, pp. 4444–4451.

NELL

http://rtw.ml.cmu.edu/rtw/

Data Source: web crawling
Languages: English
Use: raw data

Recently-Learned Facts twitter Refresh

instance	iteration	date learned	confidence
the_net_tv_show is a TV show	1111	06-jul-2018	99.9
philippians_2_15_16 is an ethnic group	1111	06-jul-2018	97.3
preservation_meeting is a perception action	1111	06-jul-2018	95.4
media___methods_magazine is a magazine	1111	06-jul-2018	99.8
jane_williams is a U.S. politician	1111	06-jul-2018	100.0
polo is a sport taught in the country __america	1116	12-sep-2018	99.2
burlington_international is an attraction that will be fall in city burlington	1116	12-sep-2018	93.8
andrew_jacobs is a journalist that writes for the publication times	1112	24-jul-2018	98.4
air_france has acquired netherlands	1111	06-jul-2018	100.0
la_voz is a newspaper in the city santa_barbara	1116	12-sep-2018	100.0

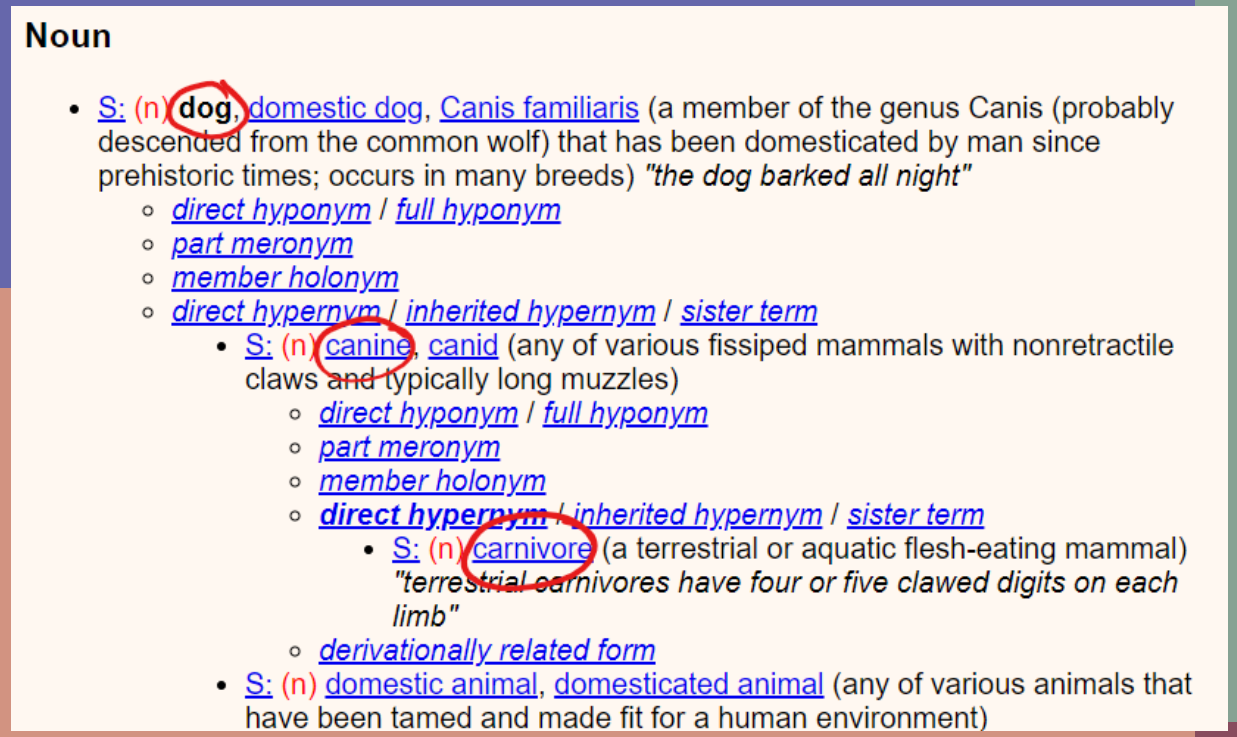
Mitchell, Tom, and Edward Fredkin. "Never-ending language learning." *2014 IEEE International Conference on Big Data (Big Data)*. 2014.

WordNet 3.1

https://wordnet.princeton.edu/

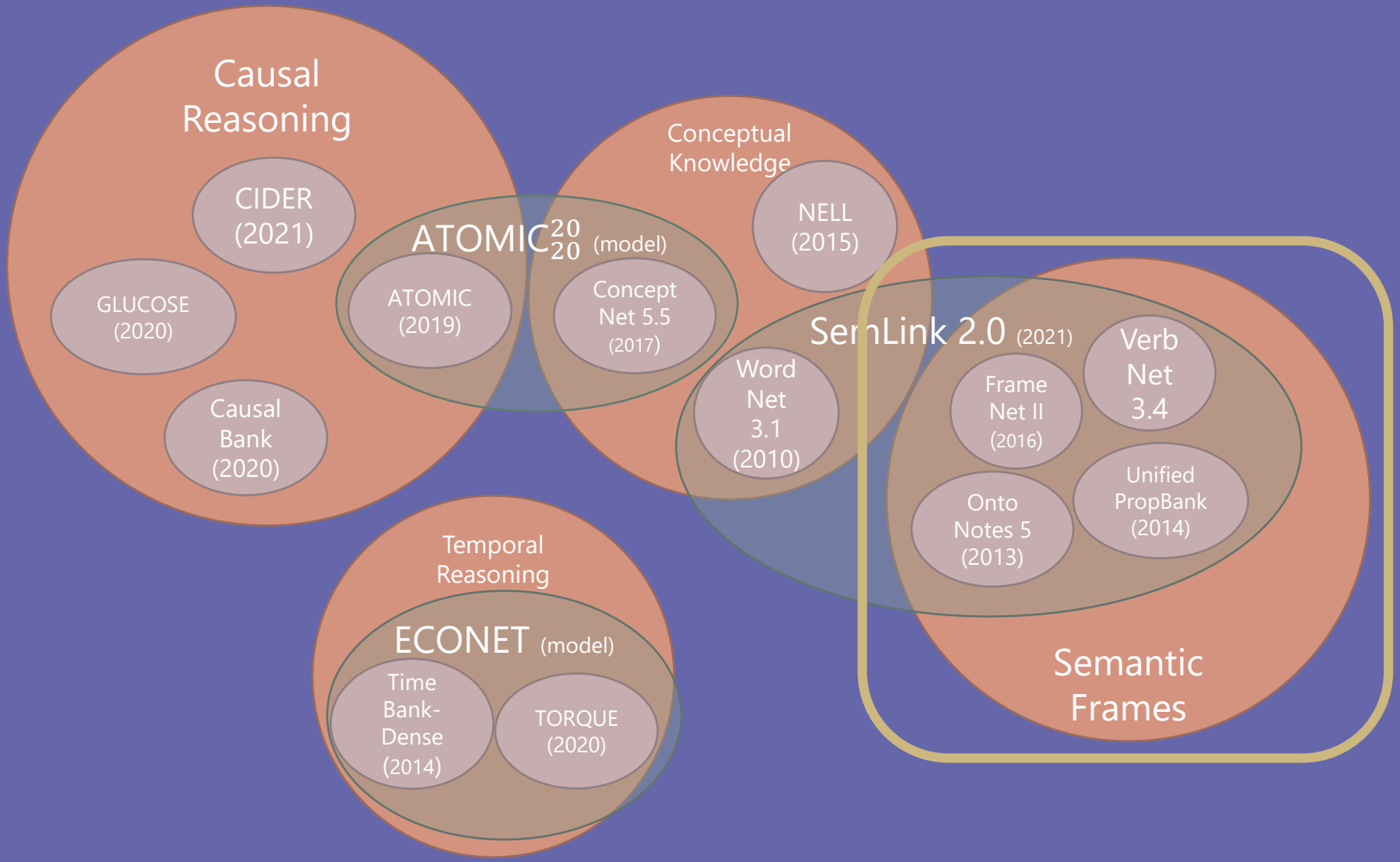
Use: hierarchical dictionary of “cognitive synonyms”
Data Source: hand-crafted
Languages: English, but other have made similar efforts: http://globalwordnet.org/
Use: nltk, raw data

Demo: http://wordnetweb.princeton.edu/perl/webwn



Princeton University "About WordNet." WordNet. Princeton University. 2010.

Semantic Frames & Propositions


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

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

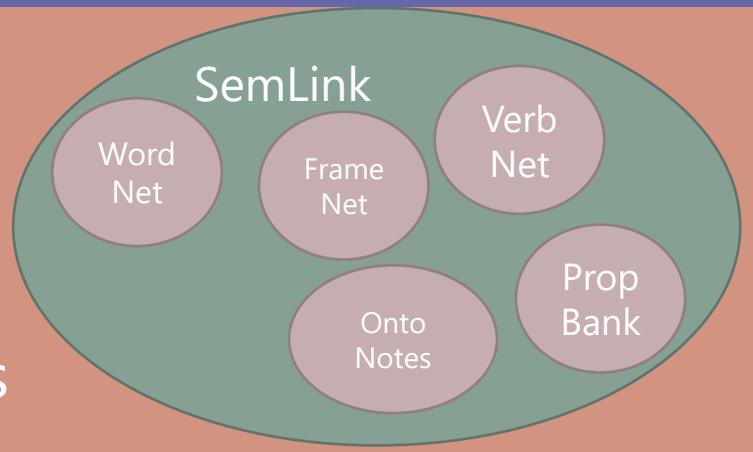
Josef Ruppenhofer, Michael Ellsworth, Miriam R. L Petruck, Christopher R. Johnson, Collin F. Baker, & Jan Scheffczyk. *FrameNet II: Extended Theory and Practice* (Revised November 1, 2016.)
Fillmore, Charles J. (1982). "Frame semantics". In The Linguistic Society of Korea, eds. *Linguistics in the Morning Calm*. Seoul: Hanshin. 111-37.

SemLink/Unified Verb Index 2.0

https://github.com/cu-clear/semlink

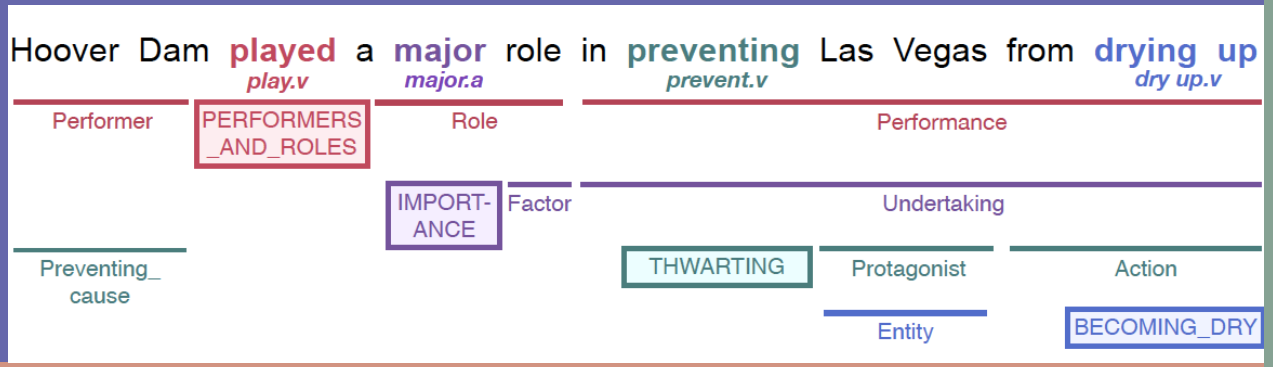
Combines 4 systems:
VerbNet, PropBank, FrameNet, WordNet and OntoNotes

Use: above link



Kevin Stowe, Jenette Preciado, Kathryn Conger, Susan Windisch Brown, Ghazaleh Kazeminejad, James Gung, and Martha Palmer. 2021. SemLink 2.0: Chasing Lexical Resources. In *Proceedings of the 14th International Conference on Computational Semantics (IWCS)*, pages 222–227, Groningen, The Netherlands (online). Association for Computational Linguistics.

FrameNet II

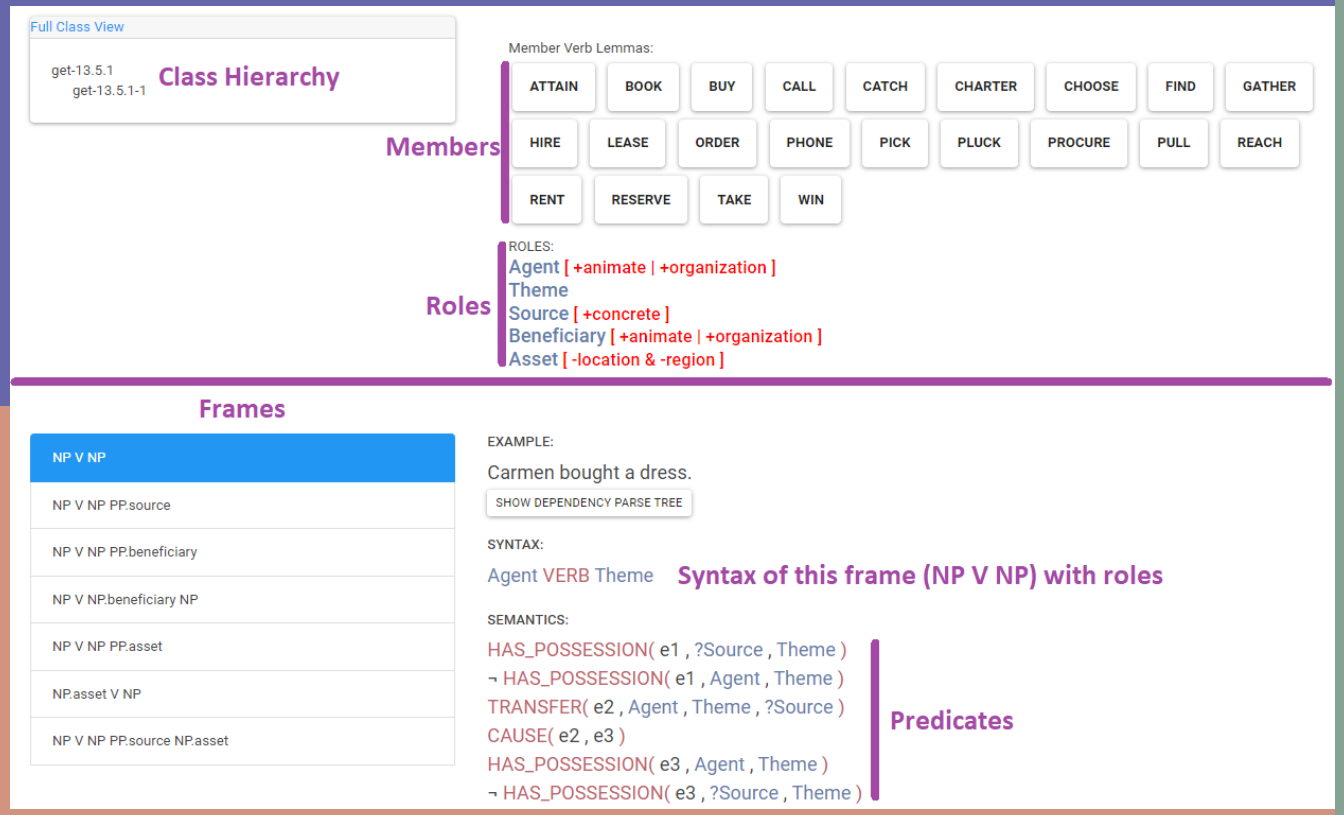


https://framenet.icsi.berkeley.edu/fndrupal/

Data Source: British National Corpus, US newswire, American National Corpus; annotated
Languages: English, global initiative: https://www.globalframenet.org/
Use: Open-SESAME; Raw data needs to be requested

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.)
Picture from Open-SESAME (Swabha Swayamdipta, Sam Thomson, Chris Dyer, & Noah A. Smith. "Frame-Semantic Parsing with Softmax-Margin Segmental RNNs and a Syntactic Scaffold" on arXiv.

VerbNet v3.4



https://verbs.colorado.edu/verbnet/

Verb classes based on Beth Levin (1993)

Data Source: hand-crafted

Languages: English

Use: raw data, my code (will be provided in upcoming homework), semparse

Demo: https://uvi.colorado.edu/uvi_search

K. Kipper Schuler, "VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon," University of Pennsylvania, 2005.

Levin, B. (1993) "English Verb Classes and Alternations: A Preliminary Investigation", University of Chicago Press, Chicago, IL.

Unified* PropBank

http://propbank.github.io/

Proposition→ true/false statement
Data Source: hand-crafted; added to PennTreebank
Languages: English, Hindi, Chinese, Arabic, Finnish, Portuguese, Basque, Turkish (Plus a way to map English to different languages)
Use: raw data

*semantic propositions regardless of part of speech (e.g. create & creation)

Event relation: Offer

25. **Predicate**: *offer*-verb
Roleset id: offer.01 transaction
Roles: Arg0: entity offering
Arg1: commodity
Arg2: price
Arg3: benefactive or entity offered to
Example: *He offered to buy the house.*

26. **Predicate**: *offer*-noun
Roleset id: offer.01 transaction
Roles: Arg0: entity offering
Arg1: commodity
Arg2: price
Arg3: benefactive or entity offered to
Example: *His offer to buy the house...*
He made an offer to buy the house.

27. **UNIFIED ROLESET**
Predicate aliases: *offer*-verb, *offer*-noun
Roleset id: offer.01 transaction
Roles: Arg0: entity offering
Arg1: commodity
Arg2: price
Arg3: benefactive or entity offered to
Example: *He offered to buy the house.*
His offer to buy the house..
He made an offer to buy the house.

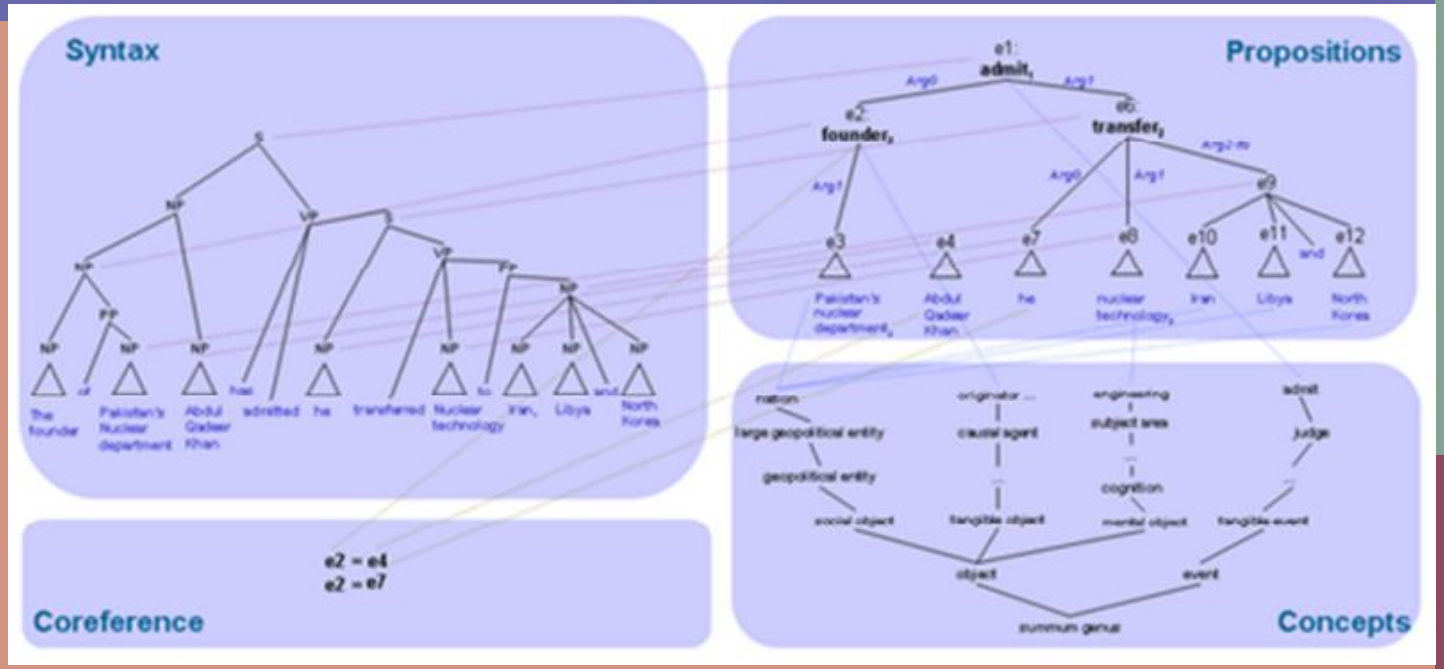
```
(o / offer-01
  :ARG0 (h2 / he)
  :ARG1 (b2 / buy-01
          :ARG0 h2
          :ARG1 (h3 / house)))
```

Martha Palmer, Dan Gildea, Paul Kingsbury, The Proposition Bank: A Corpus Annotated with Semantic Roles *Computational Linguistics Journal*, *31:1*, 2005.
Claire Bonial, Julia Bonn, Kathryn Conger, Jena Hwang and Martha Palmer (2014) PropBank: Semantics of New Predicate Types. *The 9th edition of the Language Resources and Evaluation Conference*. Reykjavik, Iceland.

OntoNotes 5.0

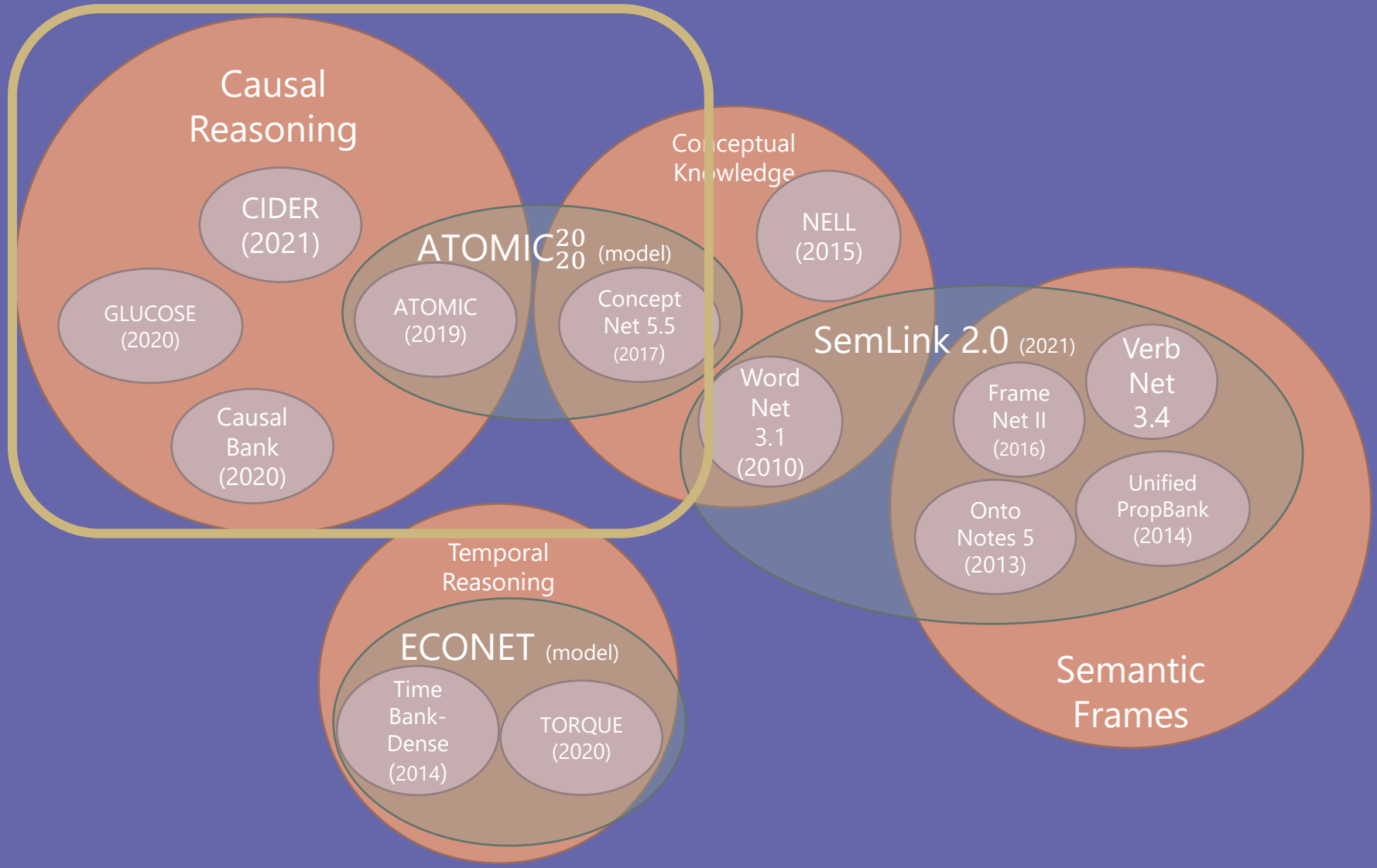
https://catalog.ldc.upenn.edu/LDC2013T19

Data Source: news, telephone conversations, blogs, talk shows, etc.
Languages: English, Chinese, Arabic
Use: raw data (same link)



S. S. Pradhan, E. Hovy, M. Marcus, M. Palmer, L. Ramshaw and R. Weischedel, "OntoNotes: A Unified Relational Semantic Representation," International Conference on Semantic Computing (ICSC 2007), 2007, pp. 517-526, doi: 10.1109/ICSC.2007.83.

Causal Orderings


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

RECAP: CAUSAL VS PROBABILISTIC ORDERINGS

Causal

Occur because of one another

Example:

I pour dog food in my dog's bowl.

My dog eats dog food.

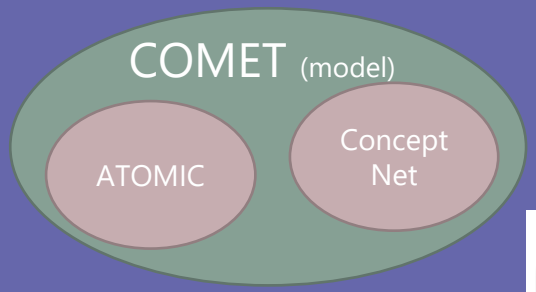
Probabilistic

Occur frequently together (not necessarily because they had to)

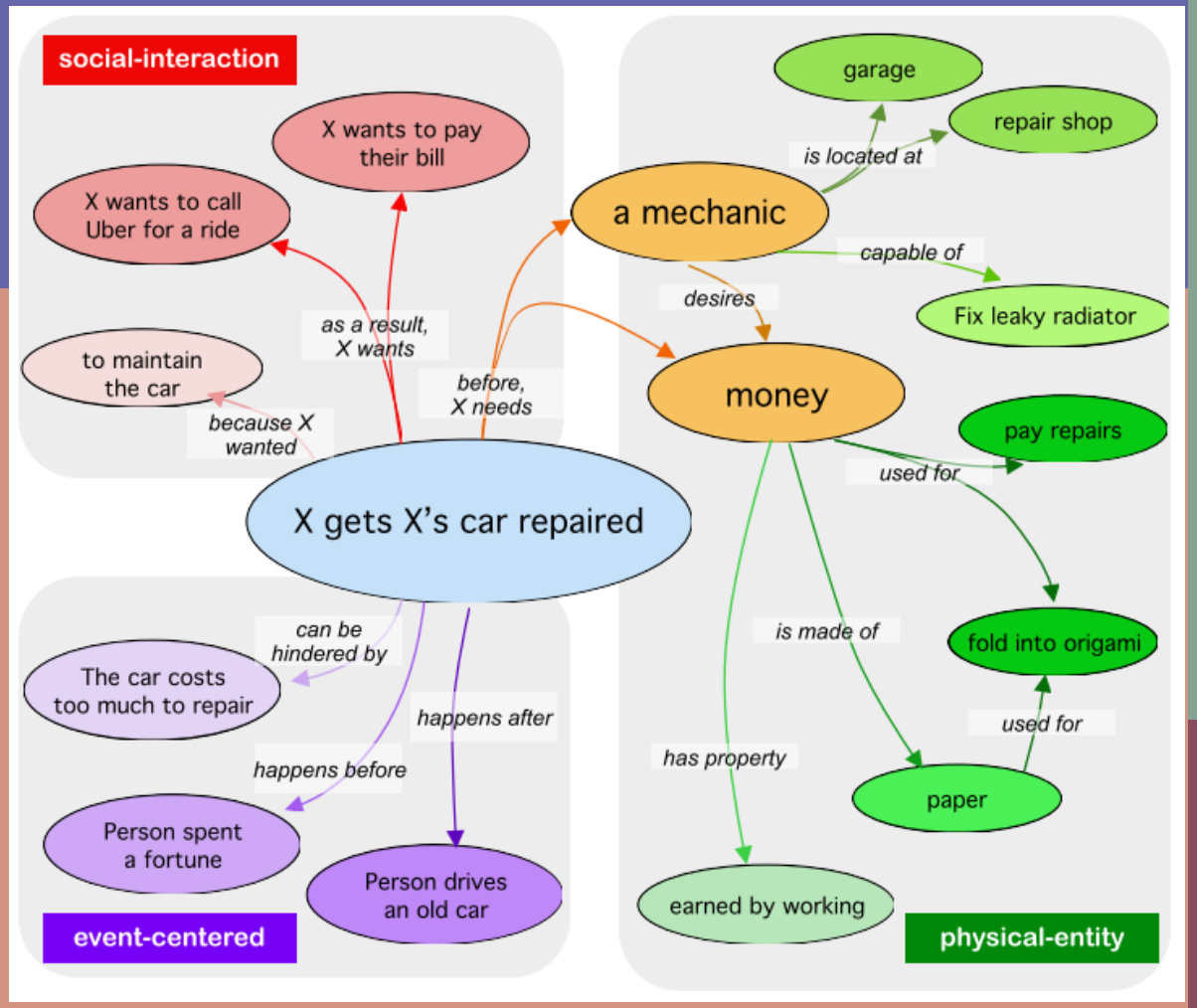
Example:

I pour dog food in my dog's bowl.

I pet my dog.



(COMET)-ATOMIC$^{20}_{20}$



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

GPT-3 COMET With ConceptNet + ATOMIC

Data Source: crowdsourcing

Languages: English

Use: above link

Demo: https://mosaickg.apps.allenai.org/kg_atomic2020

J. D. Hwang *et al.*, "(COMET-)ATOMIC2020: On Symbolic and Neural Commonsense Knowledge Graphs," 2021.

GLUCOSE

https://github.com/ElementalCognition/glucose

Causal relations within ROCStories

Data Source: crowdsourcing
Languages: English
Use: above link

Dimension	**Semi-structured Specific Statement and Inference Rule**: antecedent *connective* consequent
1: Event that directly causes or enables X	A car (subject) turned (verb) in front of (preposition) him (object) *Causes/Enables* Gage (subject) turned (verb) his bike (object)
	Sth$_A$ (subject) turns (verb) in front of (preposition) Sth$_B$ (that is Someone$_A$'s vehicle) (object) *Causes/Enables* Someone$_A$ (subject) turns (verb) Sth$_B$ (object1) away from (preposition) Sth$_A$ (object2)
2: Emotion or basic human drive that motivates X	Gage (subject) wants (verb) safety (object) *Causes/Enables* Gage (subject) turned (verb) his bike (object)
	Someone$_A$ (subject) wants (verb) safety (object) *Causes/Enables* Someone$_A$ (subject) moves (verb) away from (preposition) Something$_A$ (that is dangerous) (object)
3: Location state that enables X	Gage (subject) was (verb) close to (preposition) a car (object) *Enables* Gage (subject) turned (verb) his bike (object1) away from (preposition) the car (object2)
	Someone$_A$ (subject) is (verb) close to (preposition) Something$_A$ (object) *Enables* Someone$_A$ (subject) moves (verb) away from (preposition) Something$_A$ (object)
4: Possession state that enables X	Gage (subject) possesses (verb) a bike (object) *Enables* Gage (subject) turned (verb) his bike (object)
	Someone$_A$ (subject) possesses (verb) Something$_A$ (object) *Enables* Someone$_A$ (subject) moves (verb) Something$_A$ (object)
5: Other attributes enabling X: N/A (the dimension is not applicable for this example)	
6: Event that X directly causes or enables	Gage (subject) turned (verb) his bike (object) *Causes/Enables* He (subject) fell (verb) off his bike (object)
	Someone$_A$ (subject) turns (verb) Sth$_B$ (that is Someone$_A$'s vehicle) (object) *Causes/Enables* Someone$_A$ (subject) falls (verb) off Sth$_B$ (object)
7: An emotion that is caused by X: N/A	
8: A change in location that X results in	Gage (subject) turned (verb) his bike (object1) away from (preposition) the car (object2) *Results in* Gage (subject) was (verb) further (object1) from (preposition) the car (object2)
	Someone$_A$ (subject) moves (verb) away from (preposition) Something$_A$ (object) *Results in* Someone$_A$ (subject) is (verb) further from (preposition) Something$_A$ (object)

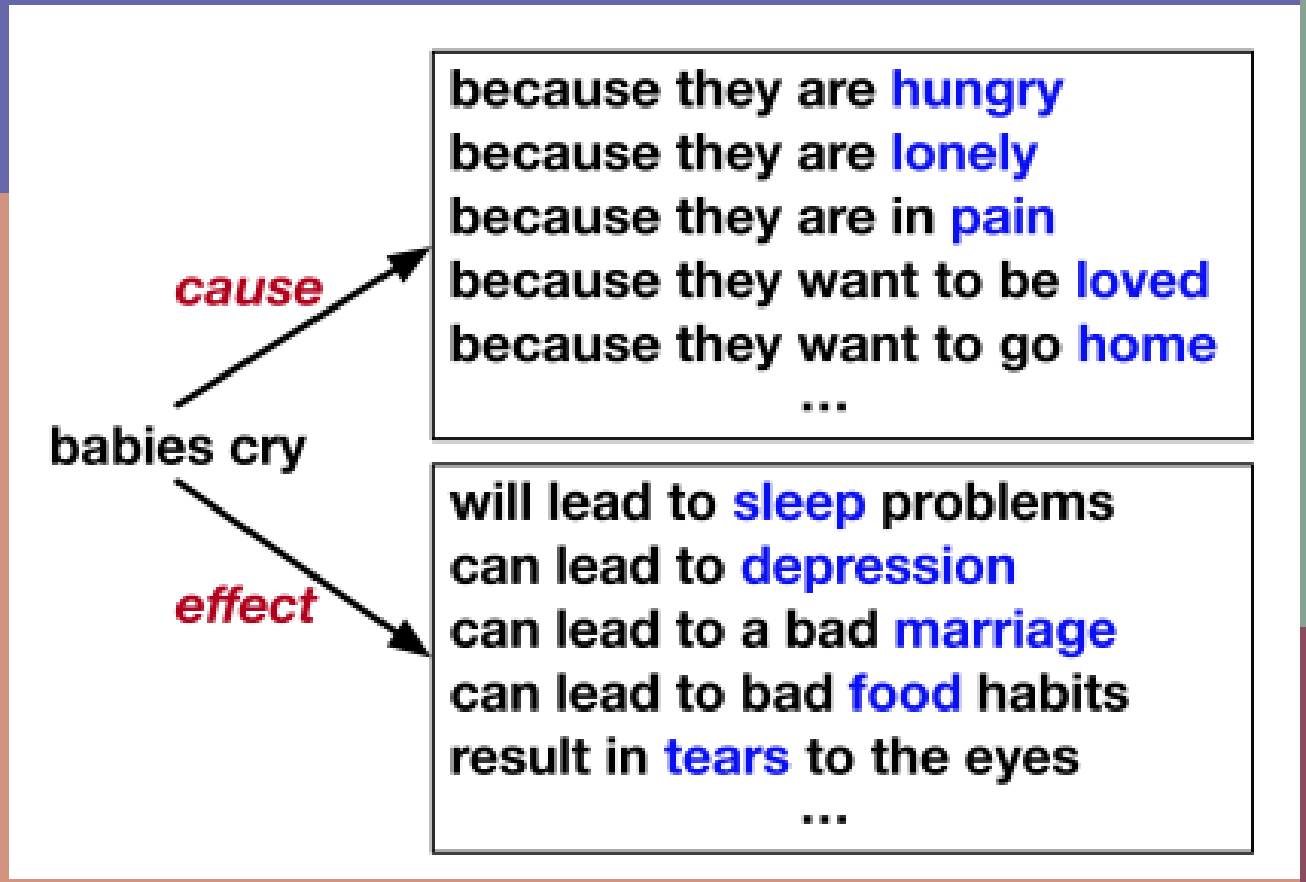
N. Mostafazadeh *et al.*, "GLUCOSE: GeneraLized and COntextualized Story Explanations," in *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2020, vol. 4569–4586.

CausalBank

https://nlp.jhu.edu/causalbank/

Large, graph-based cause & effect

Data Source: Common Crawl Corpus
Languages: English
Use: raw data, COD3S

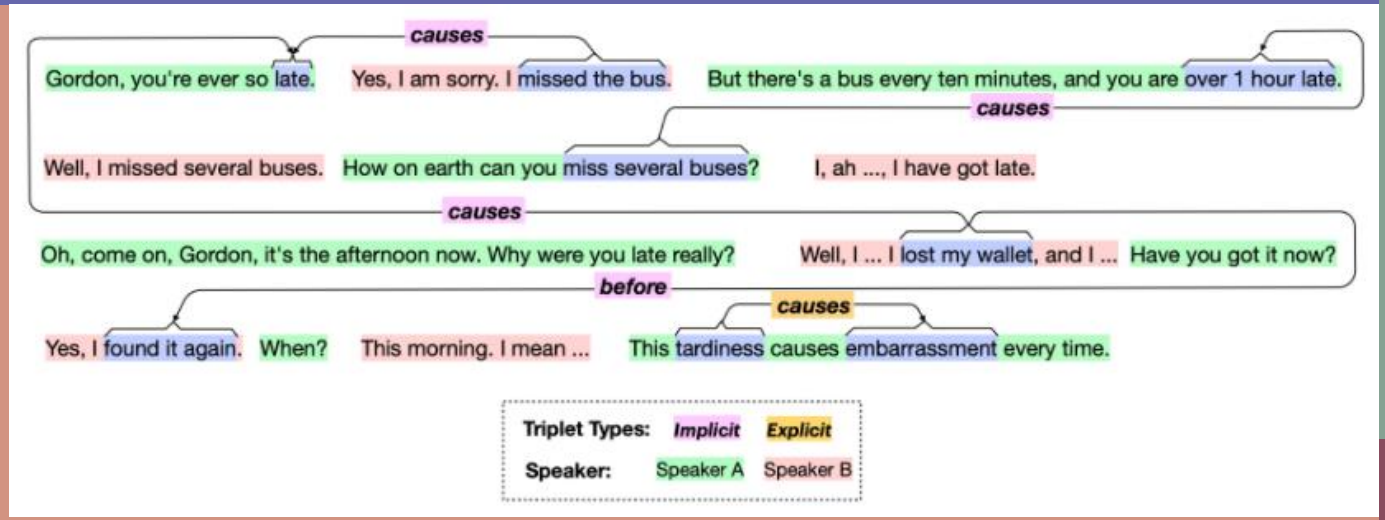


Z. Li, X. Ding, T. Liu, J. E. Hu, and B. Van Durme, "Guided Generation of Cause and Effect," in *International Joint Conference on Artificial Intelligence (IJCAI)*, 2020, pp. 3629–3636, doi: 10.24963/ijcai.2020/502.

CIDER

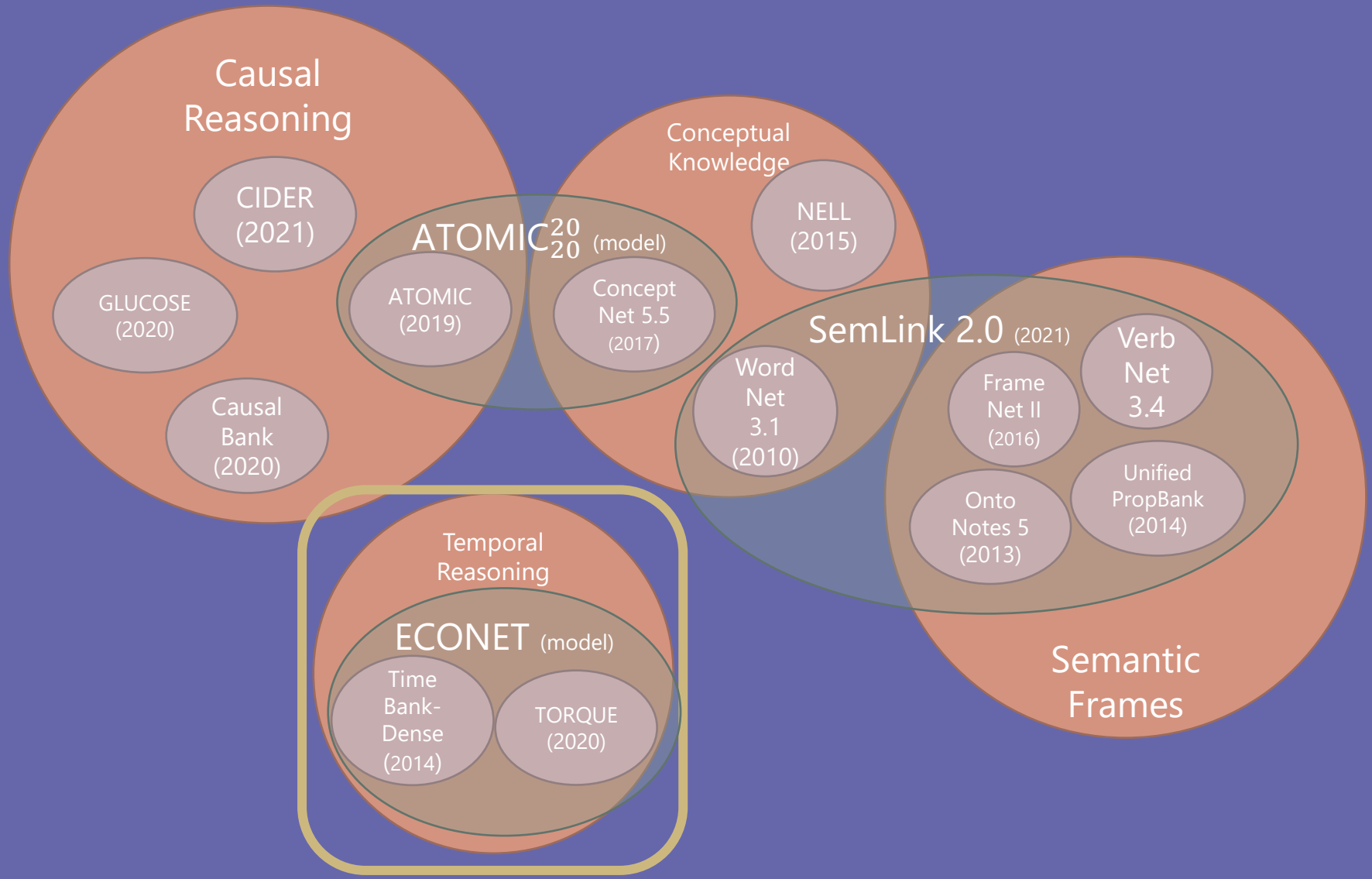
https://cider-task.github.io/cider/

Data Source: annotated dyadic (2-person) dialogues
Languages: English
Use: repo



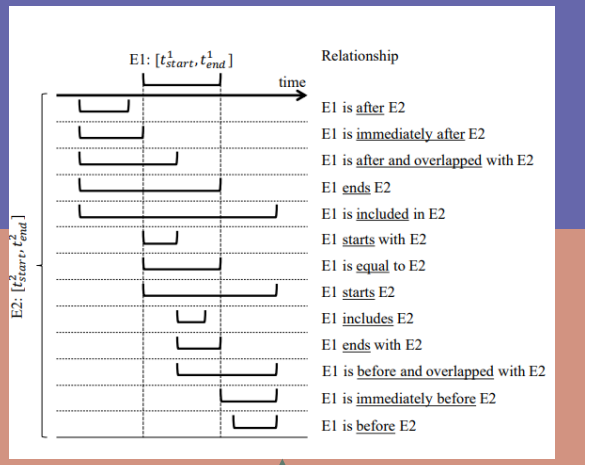
D. Ghosal, P. Hong, S. Shen, N. Majumder, R. Mihalcea, and S. Poria, "CIDER: Commonsense Inference for Dialogue Explanation and Reasoning," 2021.

Temporal Orderings


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

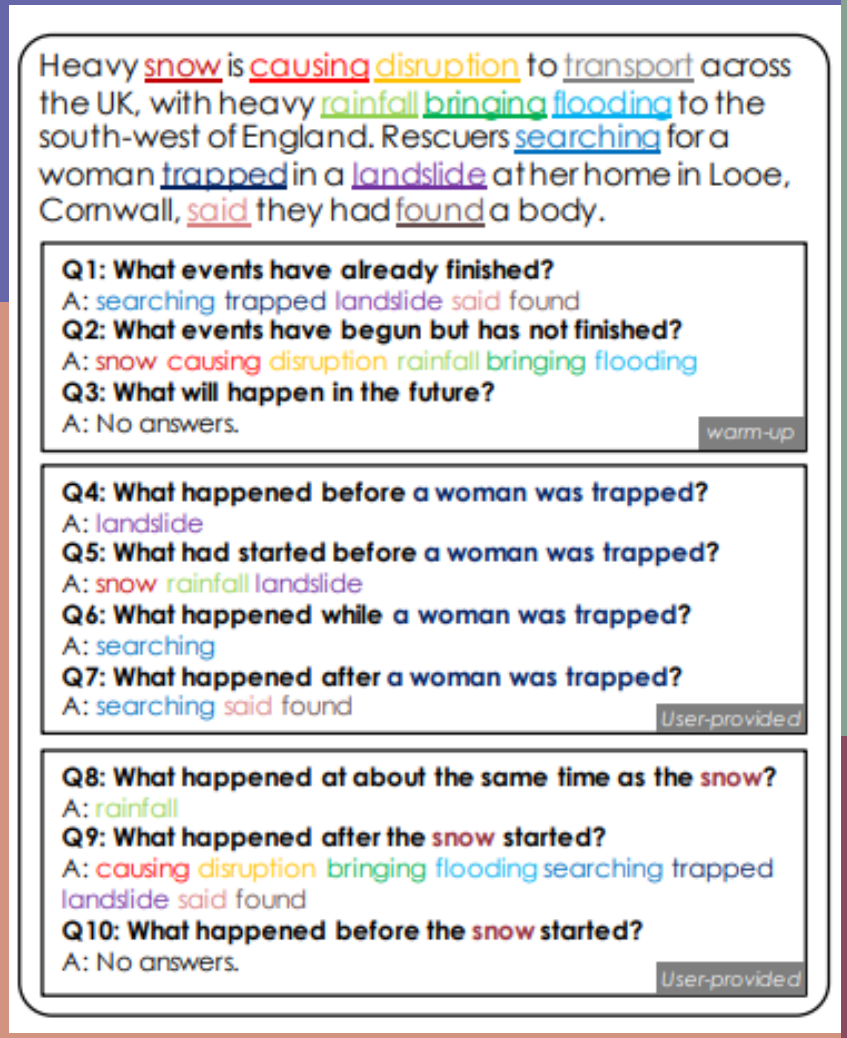
TORQUE

https://allenai.org/data/torque



Contains information about time span

Data Source: crowdsourcing
Languages: English
Use: raw data, ECONET

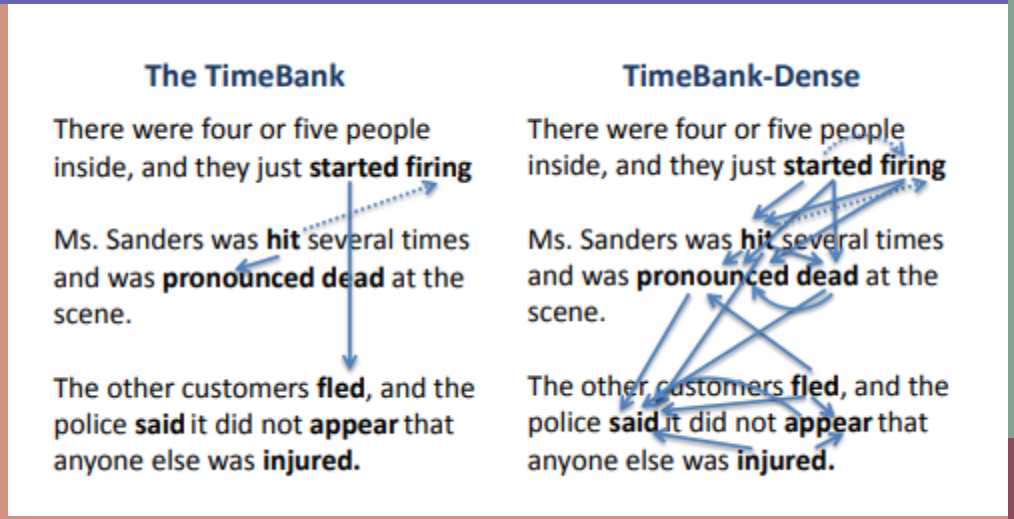


Qiang Ning, Hao Wu, Rujun Han, Nanyun Peng, Matt Gardner, and Dan Roth. 2020. TORQUE: A Reading Comprehension Dataset of Temporal Ordering Questions. *In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1158–1172, Online. Association for Computational Linguistics.

TimeBank-Dense

https://www.usna.edu/Users/cs/nchamber/caevo/

Data Source: re-annotated TimeBank (news articles annotated)
Languages: English
Use: CAEVO, ECONET



Nathanael Chambers, Taylor Cassidy, Bill McDowell, and Steven Bethard. 2014. Dense Event Ordering with a Multi-Pass Architecture. *Transactions of the Association for Computational Linguistics*, 2:273–284.

Schemas

Schema

"a representation of a plan or theory in the form of an outline or model." (from www.lexico.com/)

- Holds a set of facts/information (extracted from input text using knowledge representation)
- Can be used to capture the state of a fictional world and be updated when the fictional world changes (can be changed over time)

Using VerbNet

Jen sent the book to Remy from Atlanta.

ROLES Agent Theme Destination Initial_Location

has_location(e1, book, Atlanta)
do(e2, **Jen**)
cause(e2, e3)
motion(e3, book)
!has_location(e3, book, Atlanta)
has_location(e4, book, Remy)

PREDICATES

Initial_Location : location

Theme : concrete

Agent : animate or organization

SELECTIONAL RESTRICTIONS

Pre-Conditions and Effects

Jen sent the book to Remy from Atlanta.

Pre-Conditions

has_location(e1, book, Atlanta)

do(e2, **Jen**)

cause(e2, e3)

motion(e3, book)

!has_location(e3, book, Atlanta)

has_location(e4, book, Remy)

Effects

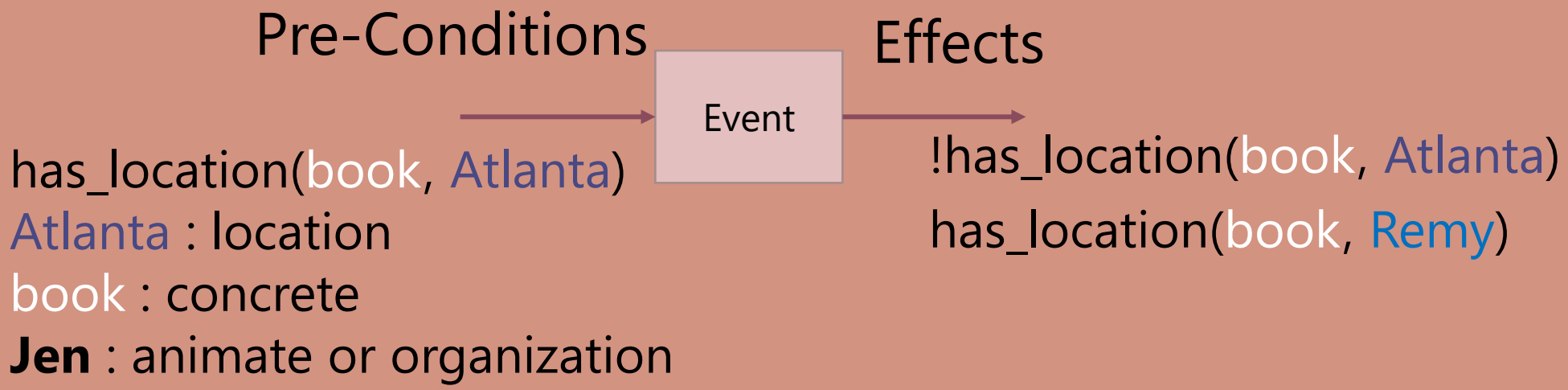
~~Atlanta : location~~

~~book : concrete~~

~~**Jen** : animate or organization~~

Pre-Conditions and Effects

Jen sent the book to Remy from Atlanta.



Resulting State Representation

Jen sent the book to Remy from Atlanta.

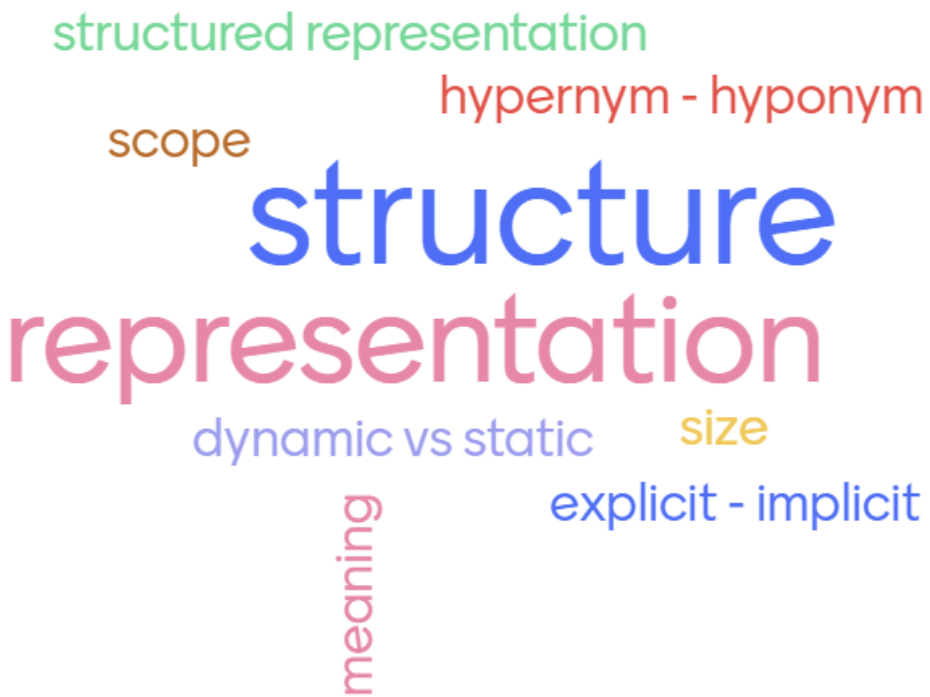
Atlanta : location
book : concrete
Jen : animate or organization
!has_location(book, Atlanta)
has_location(book, Remy)

What's the difference between a schema and a concept?

Go to **www.menti.com** and use the code **3167 2634**

What's the difference between a schema and a concept?

Mentimeter



17

In-Class Activity

https://interactive-fiction-class.org/in_class_activities/schemas/schemas.html

Please fill out this mid-semester survey

https://forms.gle/bQXZz3y8xzrU7wJ68

Have a good spring break!