

Ontologies and World Modeling

Lara J. Martin (she/they)

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

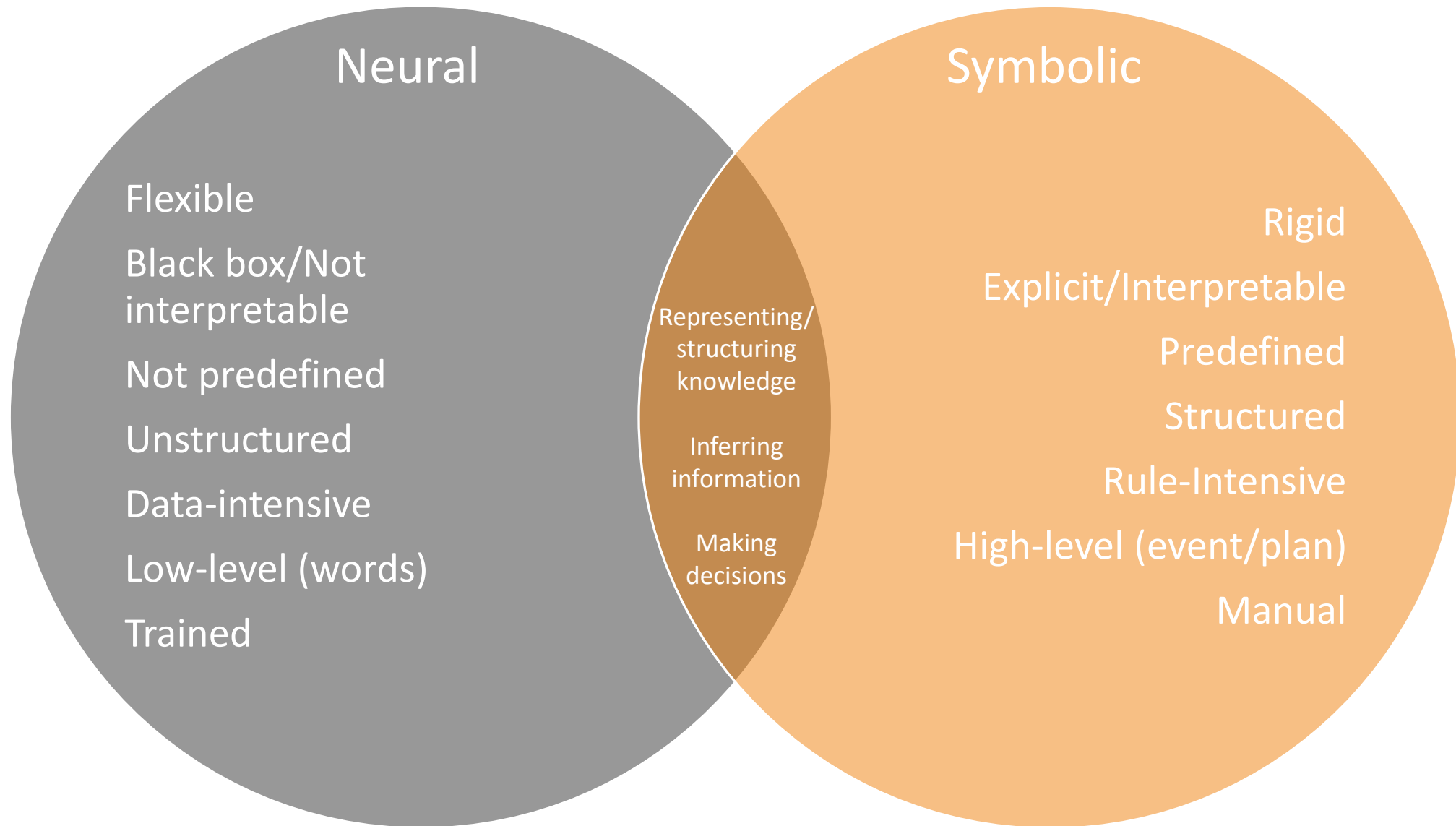
Modified from slides from Susan Brown

Learning Objectives

Define an ontology and its components

Compare scripts, knowledge graphs, and ontologies

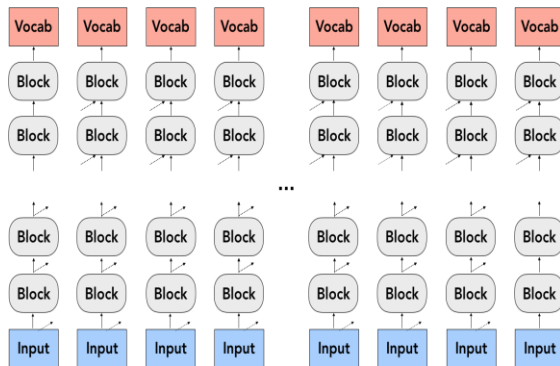
Evaluate how world modeling has evolved through the years



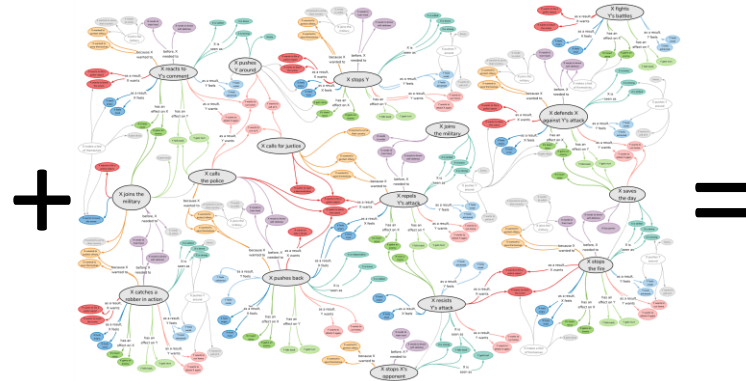
Review: Ways of combining neural and symbolic methods

Commonsense Transformers

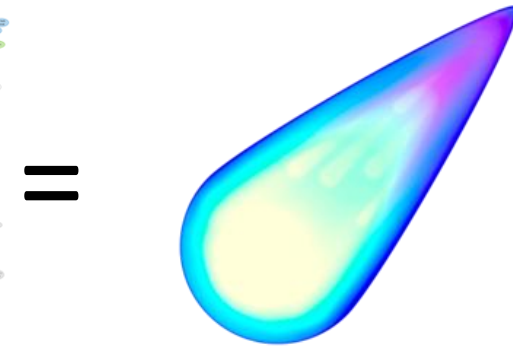
- Language models implicitly represent knowledge
- Finetune them on knowledge graphs to learn structure of knowledge
- Resulting knowledge model generalizes structure to other concepts



Pre-trained
Language Model

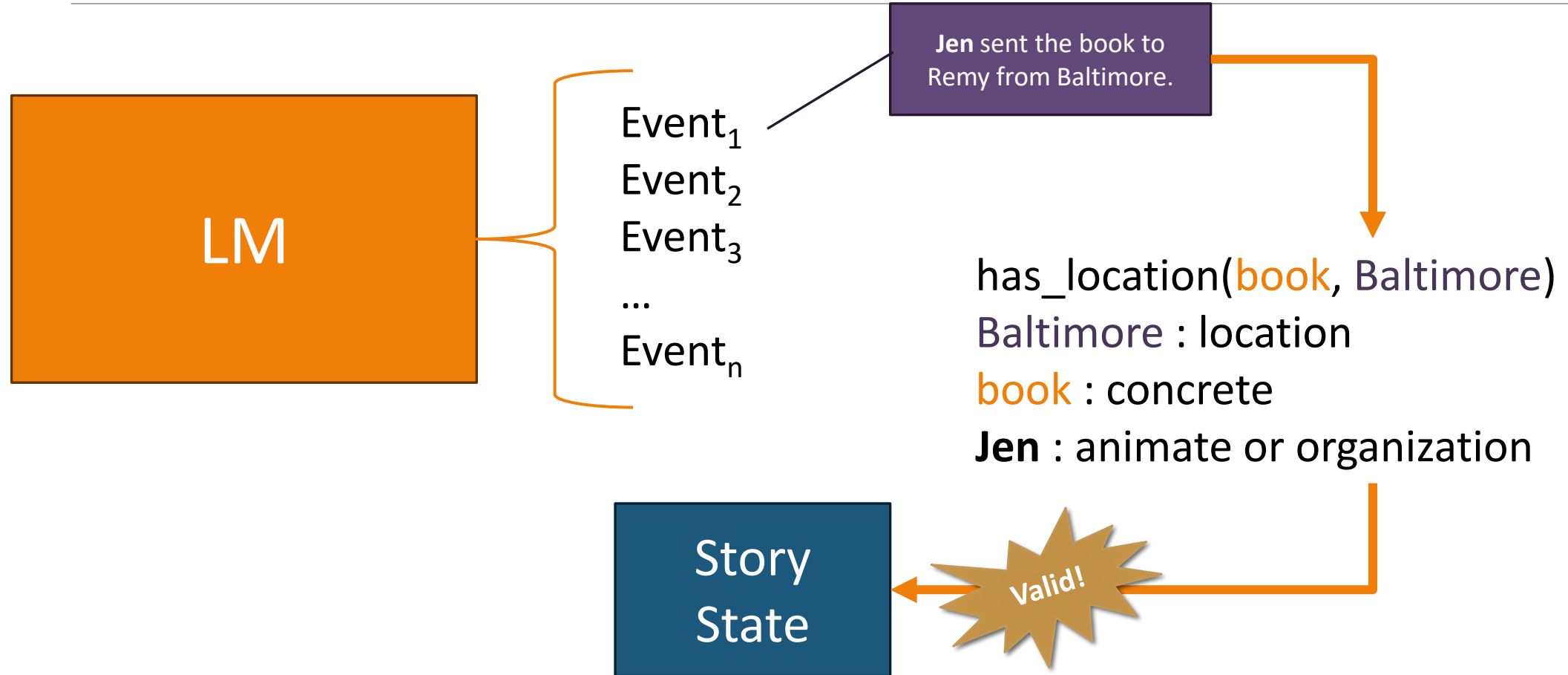


Seed Knowledge
Graph Training



COMET

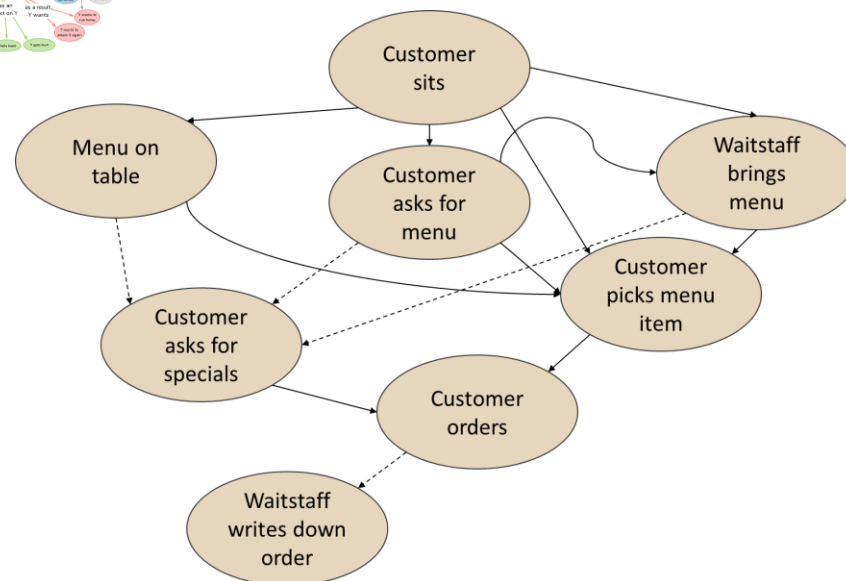
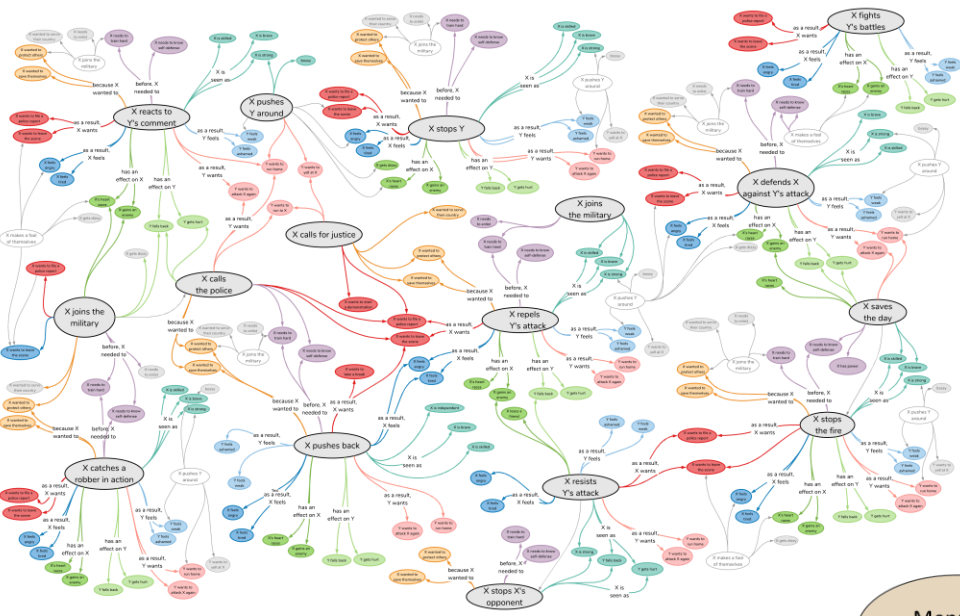
How does a neural network fit in here?



A quick question

What is a *script*?

What's the difference between a script and a knowledge graph?



Ontologies

Semantic representations and predicate logic

Franco likes Frasca.

First order logic:

$$\exists e \text{Liking}(e) \wedge \text{Liker}(e, \text{Franco}) \wedge \text{Liked}(e, \text{Frasca})$$

VerbNet:

The lion tamer jumped the lion through the hoop.

has_location(e1, Theme, Initial_Location)

do(e2, Agent)

motion(e3, Theme, Trajectory)

has_location(e4, Theme, Destination)

cause(e2, e3)

Semantics

Let's start with the basics of what we might want to say about some world.

- There are entities in this world.
- We'd like to assert properties of these entities.
- And we'd like to assert relations among them.

Let's call a scheme that can capture these things *a model*

And let's claim that we can use basic *set theory* to represent such models.

We can do this with *an ontology*.

From vocabulary to ontology

Vocabulary

- Fixed set of terms

Taxonomy

- Fixed set of terms with subset relations between terms

Ontology

- Fixed set of terms with structured relationships between terms, generalization, specialization of terms

Logic-based ontology

- Ontology that is written in a formal language that is underpinned by a logic, giving it a precisely specified semantics, and computable relationships between terms

What is an ontology

Describes a domain

- concepts
- properties and attributes of those concepts
- constraints on properties and attributes
- individuals

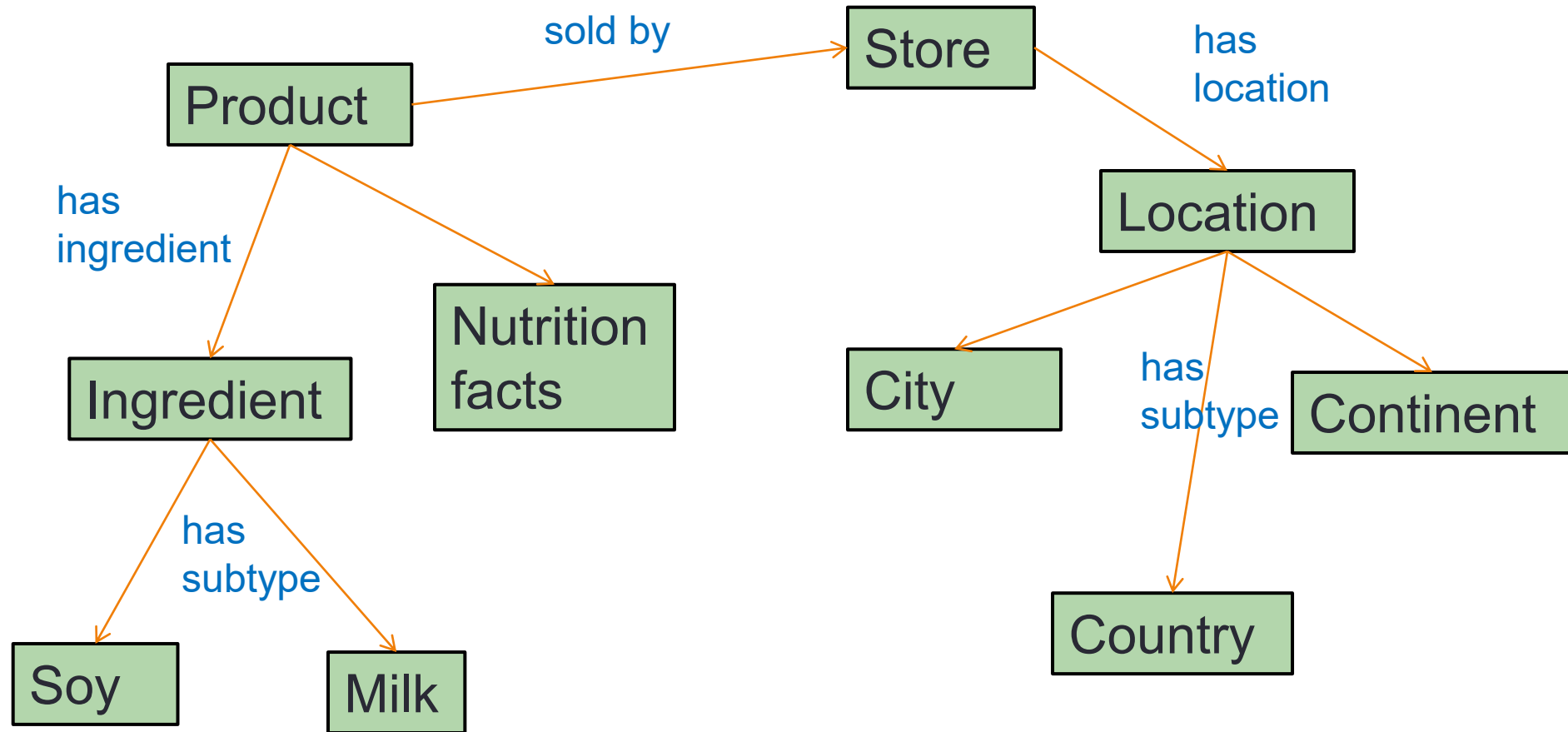
Defines

- a common vocabulary
- a shared understanding

Can be used with reasoning agents

- to infer new facts from existing definitions

Imagine a mind map for the domain



Ontology basics (using OWL)

Axioms

Basic **statements** in an ontology.
An ontology is a set of axioms

Entities

Used to refer to basic **things in the domain** of interest.

Class Expressions

Combinations of entities that form more **complex descriptions out of simpler ones.**

Axioms specify the relationships between entities and class expressions

OWL Axioms

Some examples...

Cat **SubClassOf** Animal

SubClassOf
Cats are Animals

Cat **DisjointWith** Dog

DisjointClasses
Cats are not Dogs

Tibbs **Type** Cat

ClassAssertion
Tibbs is a Cat

Betty hasPet Tibbs

PropertyAssertion
Betty has Tibbs as a pet

hasPet **Domain** Person

Domain
Anything that has a pet is
Person

Class expressions

Some examples...

Cat **or** Dog

The class of individuals that instances of Cat or Dog (or both!)

Person **and** PetOwner

The class of individuals that are both instances of Person and PetOwner

hasPet **some** Cat

The class of individuals that have at least one hasPet relationship to an individual that is an instance of Cat

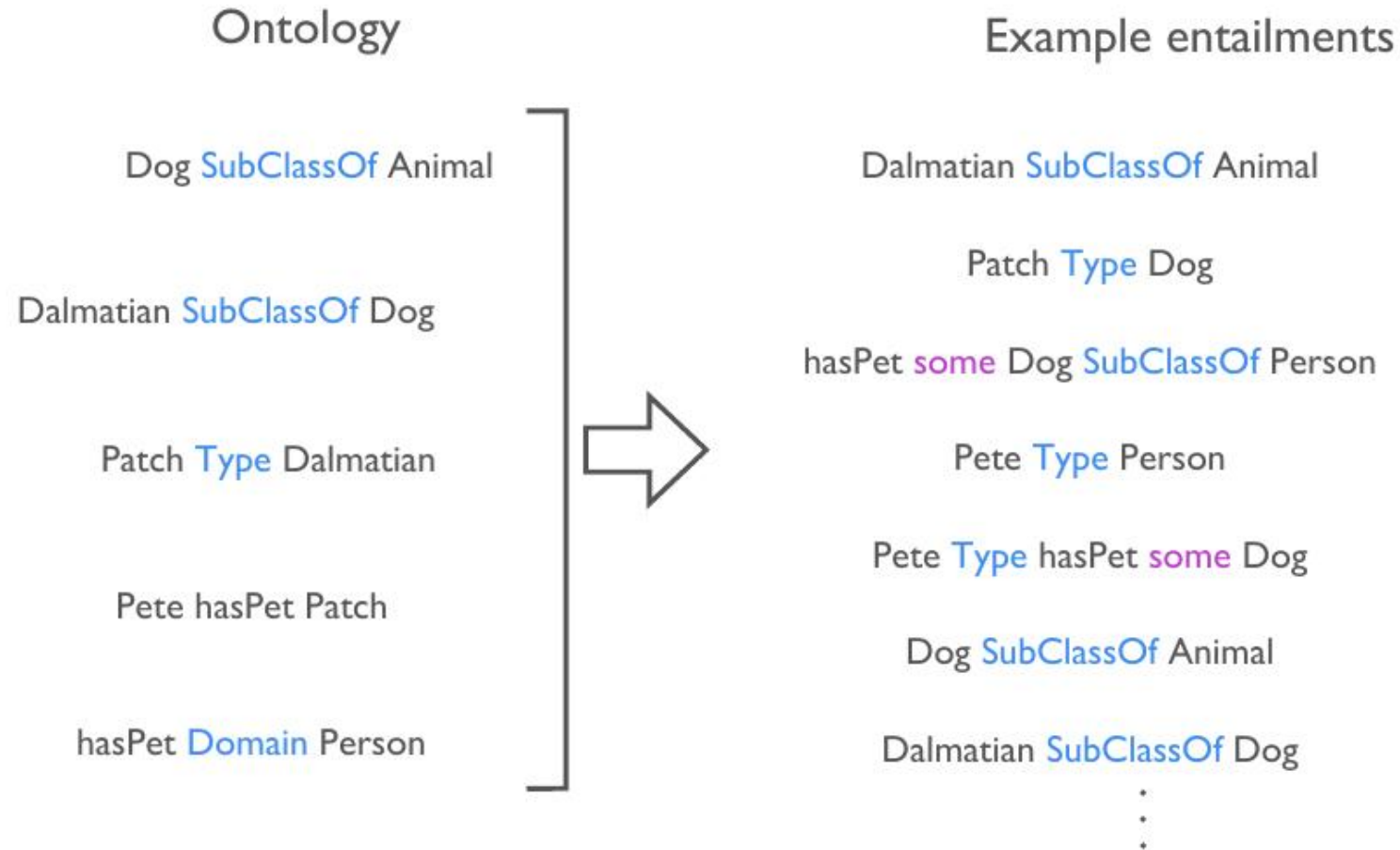
Person **and** hasPet **some** Cat

The class of individuals that are both instances of Person and hasPet some Cat

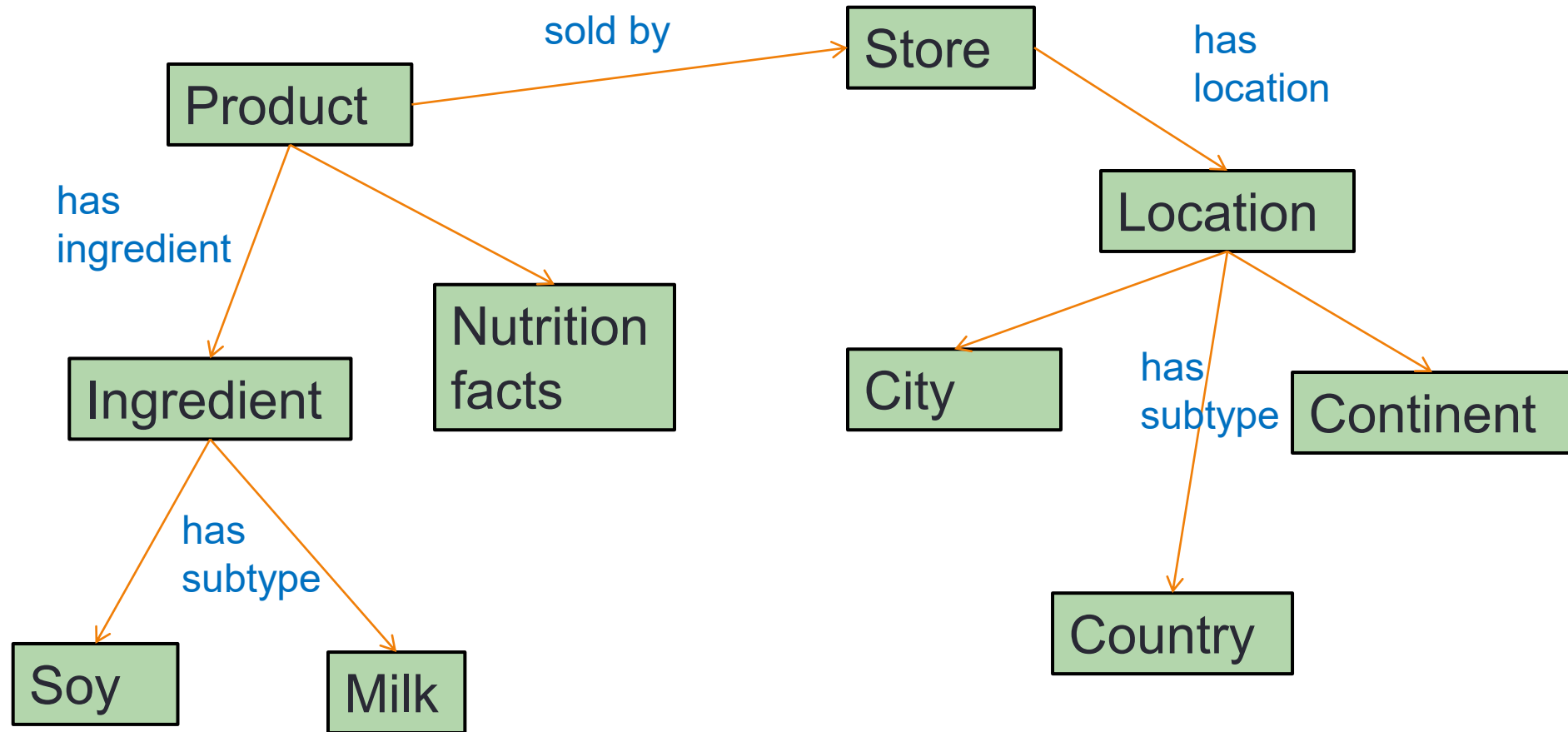
Person **and not** (hasPet **some** (Cat **or** Dog))

The class of individuals that are instances of Person but not instances of the class of individuals that have at least one hasPet relationship to and individual that is an instance of the class Cat or Dog

Entailment



Imagine a mind map for the domain



Defining classes

A class is a concept in the domain

- a class of products
- a class of ingredients
- a class of dairy products

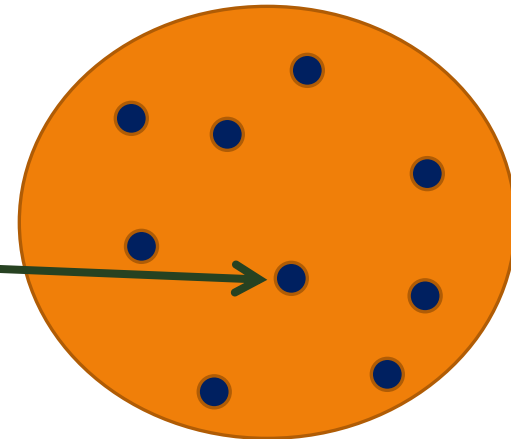
A class is a set of elements with similar properties

Instances of classes

- a box of cereal that you are buying

*box of cereal
you just bought*

class of cereal products

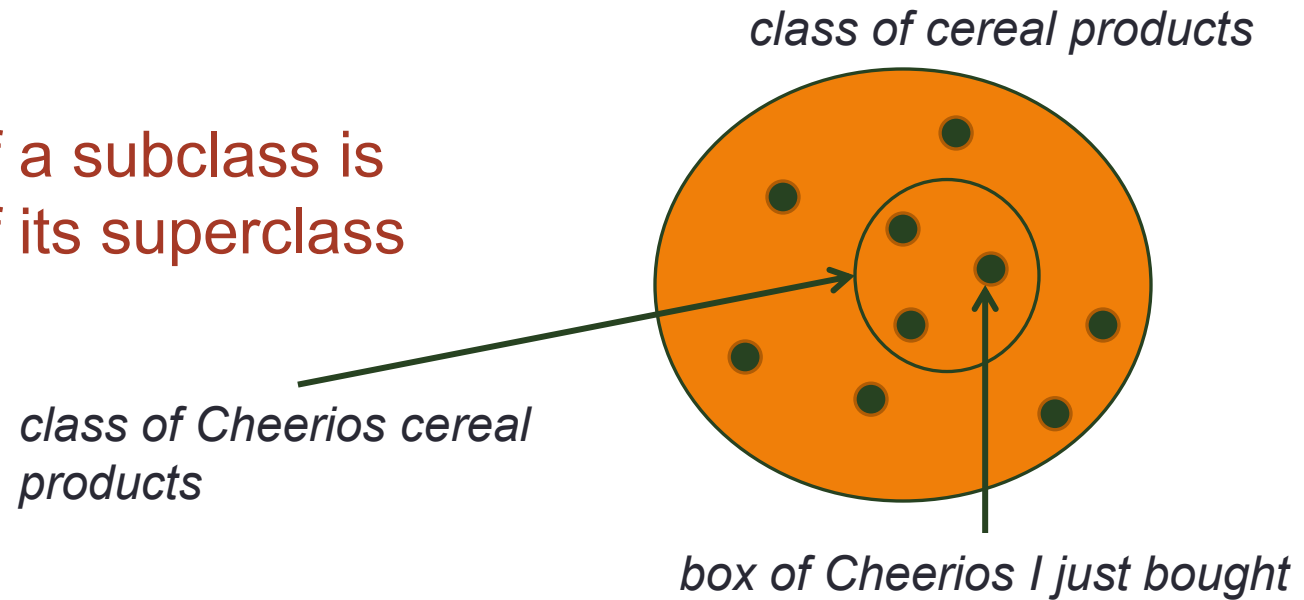


Class inheritance

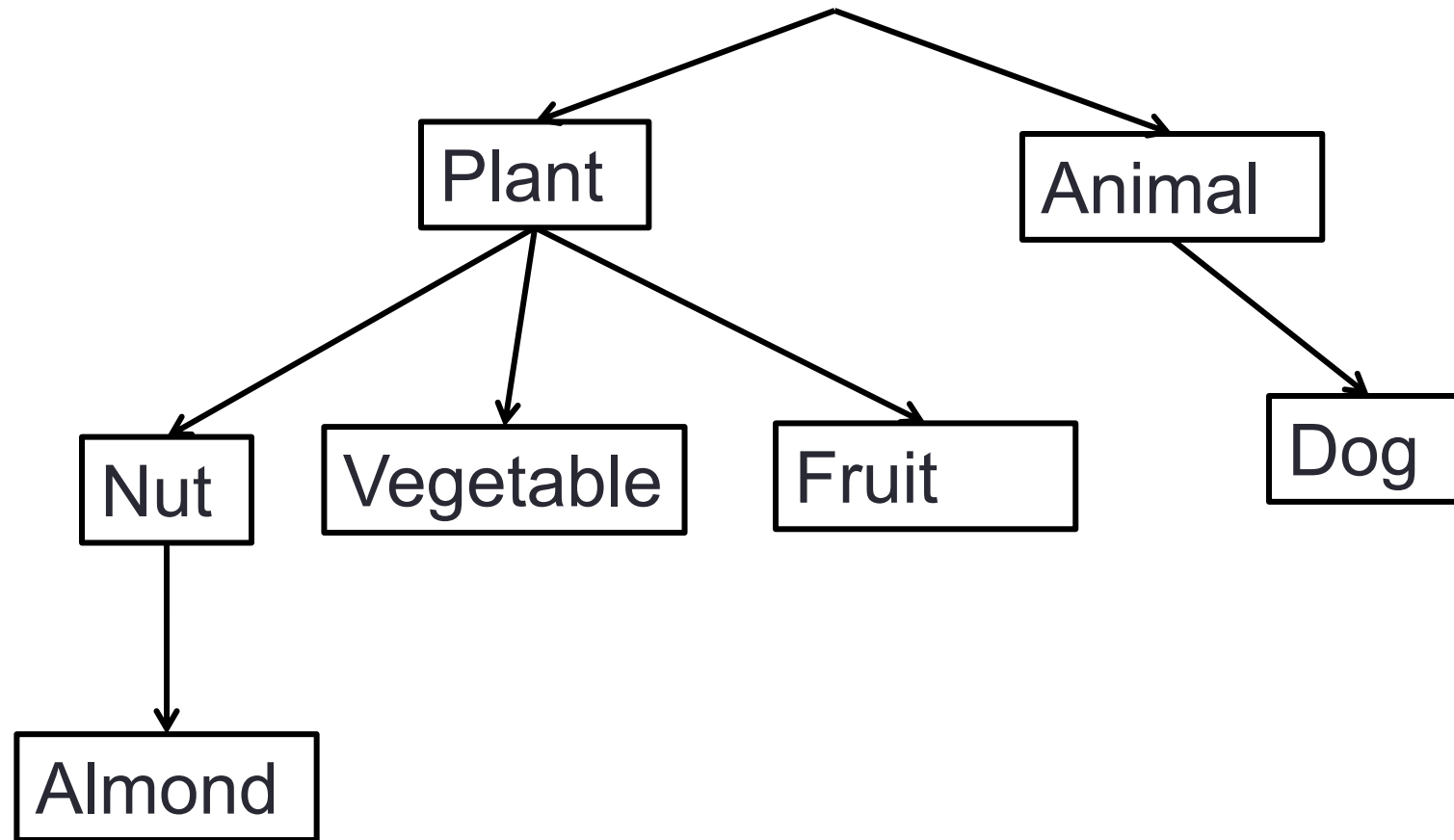
Classes usually constitute a taxonomic hierarchy (a subclass-superclass hierarchy)

an IS-A hierarchy:

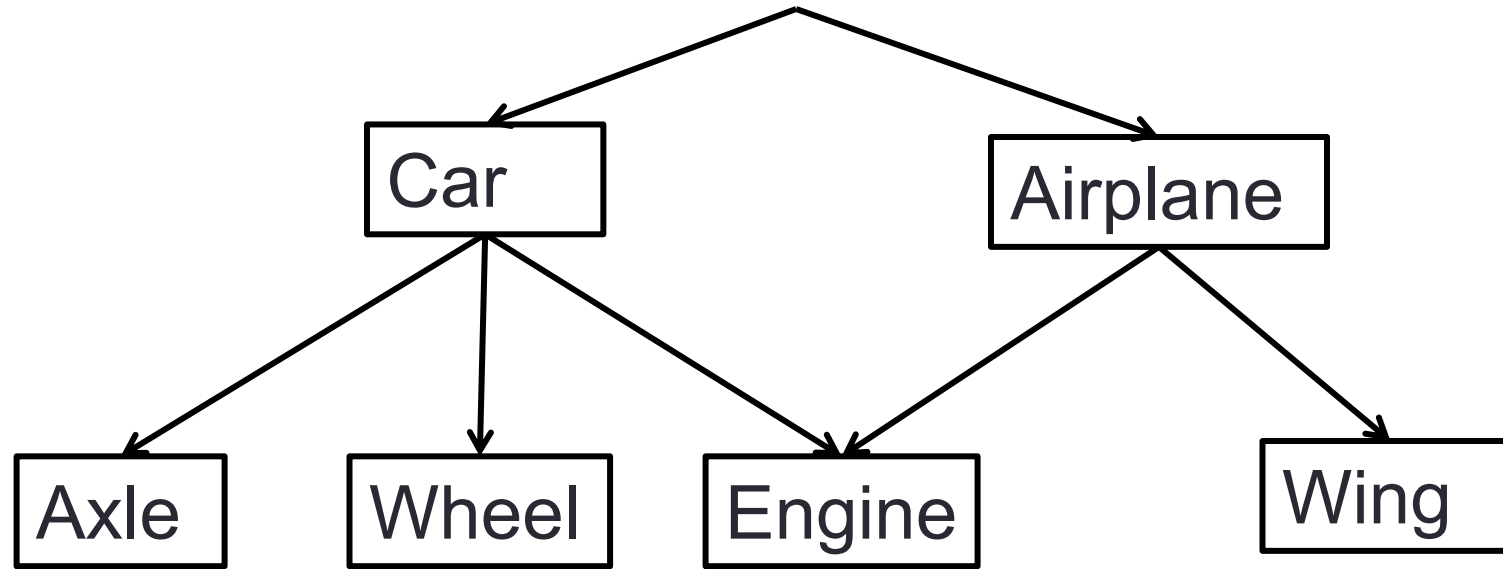
an instance of a subclass is
an instance of its superclass



If you think of a class as a set, a subclass is a subset



**Subclass-Superclass
relations?**



**Subclass-Superclass
relations?**

Defining properties

Products *have a price*



price

Products *are produced by* a manufacturer



produced by

Products *have an expiration date*



has expiration
date

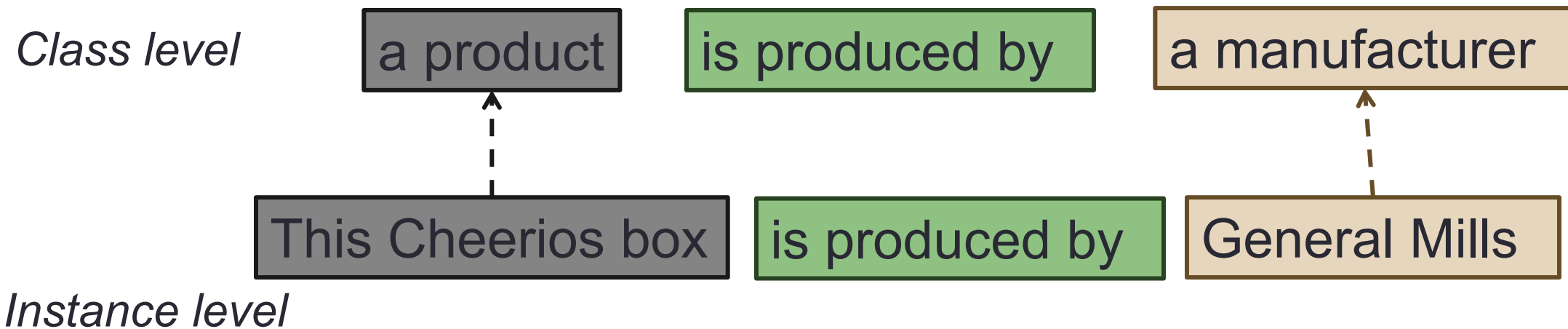
Products *have ingredients*



has ingredient

Properties describe instances

Properties associated with a class describe the **attributes** and **relationships of the instances** of the class



Individuals

Individuals are the last level of the ontology; they cannot be further specified

They represent a materialization of the descriptions at the class level

This is the level at which the actual data is put in

The data depends on the application

- grocery app?
- tracking terrorist organizations?

Ontologies for NLP

Move from words to concepts/entities underlying the words

A conceptual ontology with links to lexical items

Bio-NLP

Event extraction and participant tracking

Events in ontologies

Events difficult to model in an ontology

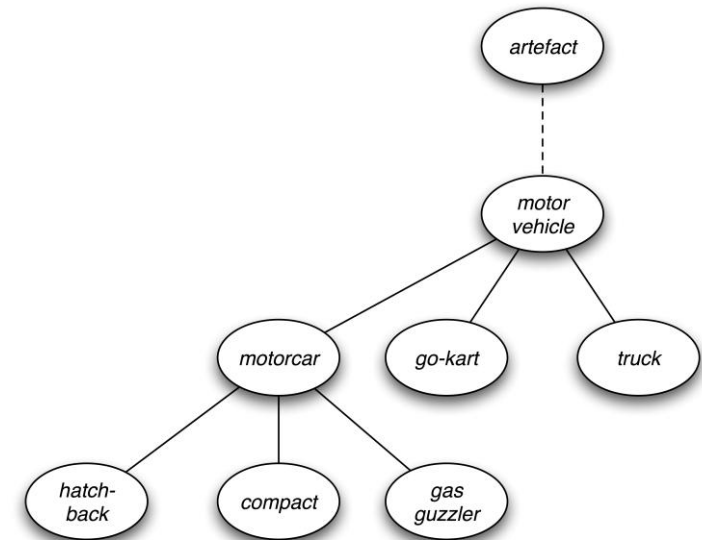
- is-a relations tricky to determine (killing, crime, murder, death)
- where does an event start and end? (surgical event)

Usually represented as **relations** between entities

- relations can't have links to lexical items
- relations can't have individuals (you might want to make lexical items the individuals or instances in annotation)

Most ontologies have **shallow** models of events

- WordNet
- SUMO (Suggested Upper Merged Ontology)



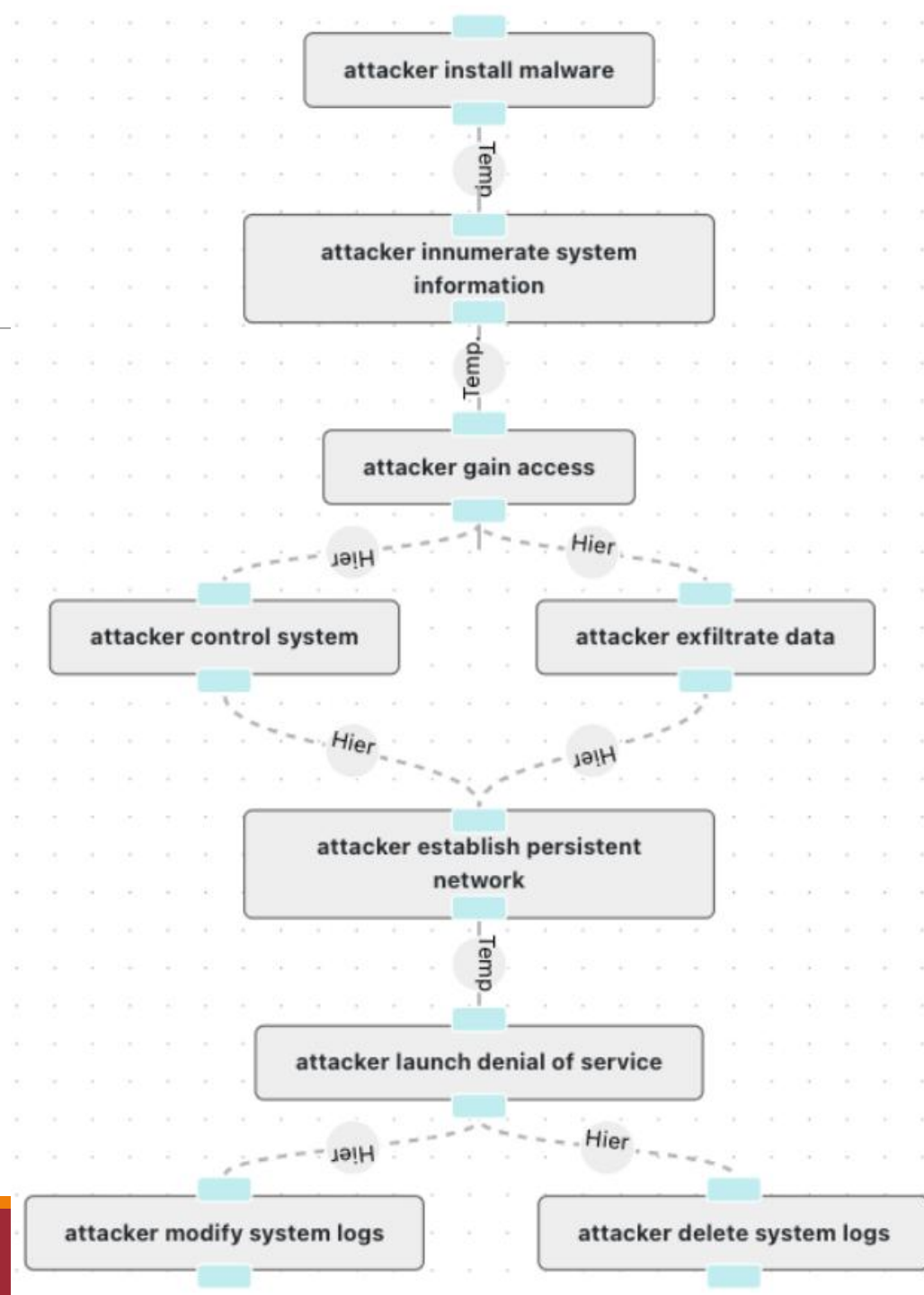
<https://docs.huihoo.com/nltk/0.9.5/en/ch02.html>

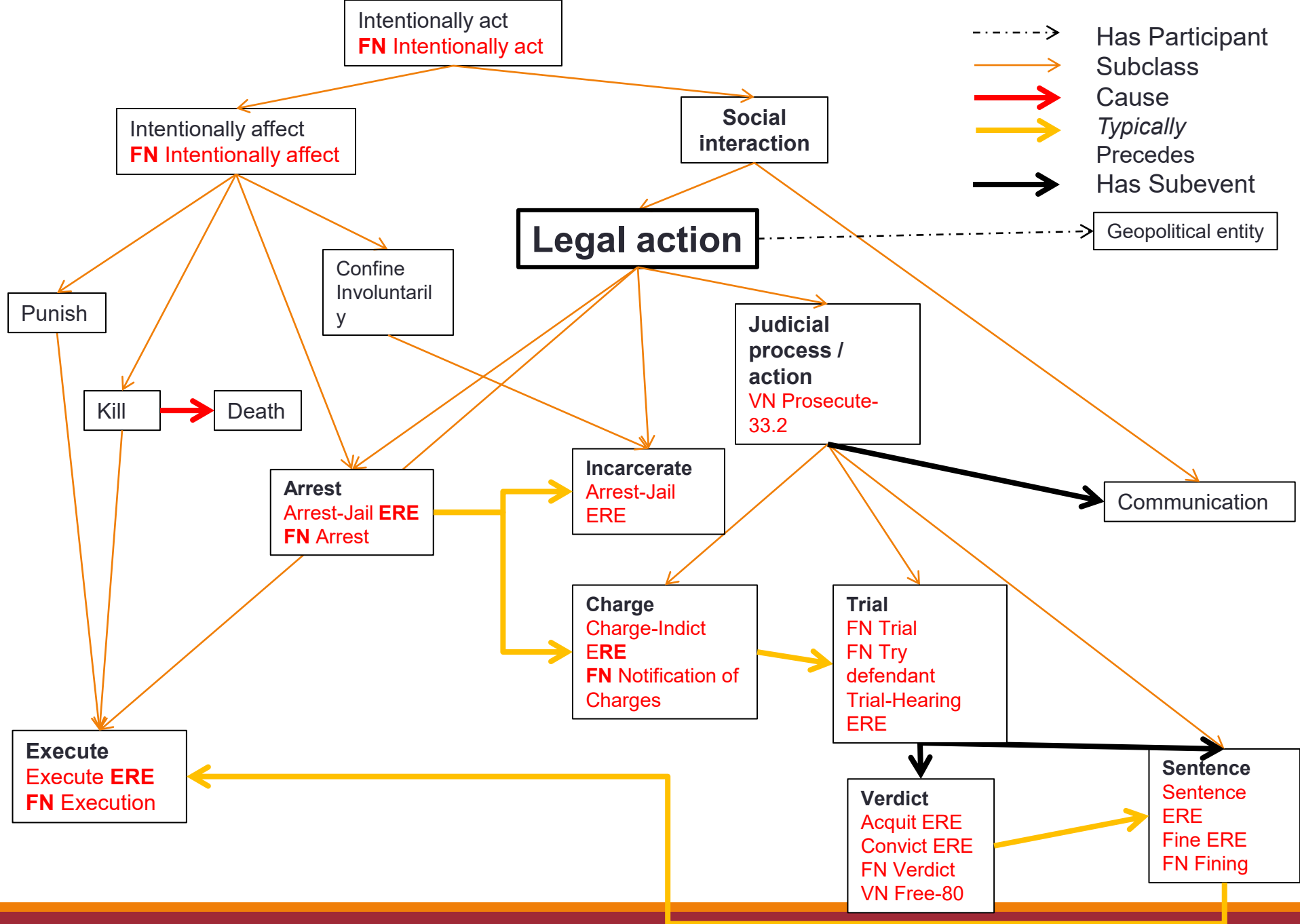
Event Schemas

Structured representation for tracking how events unfold

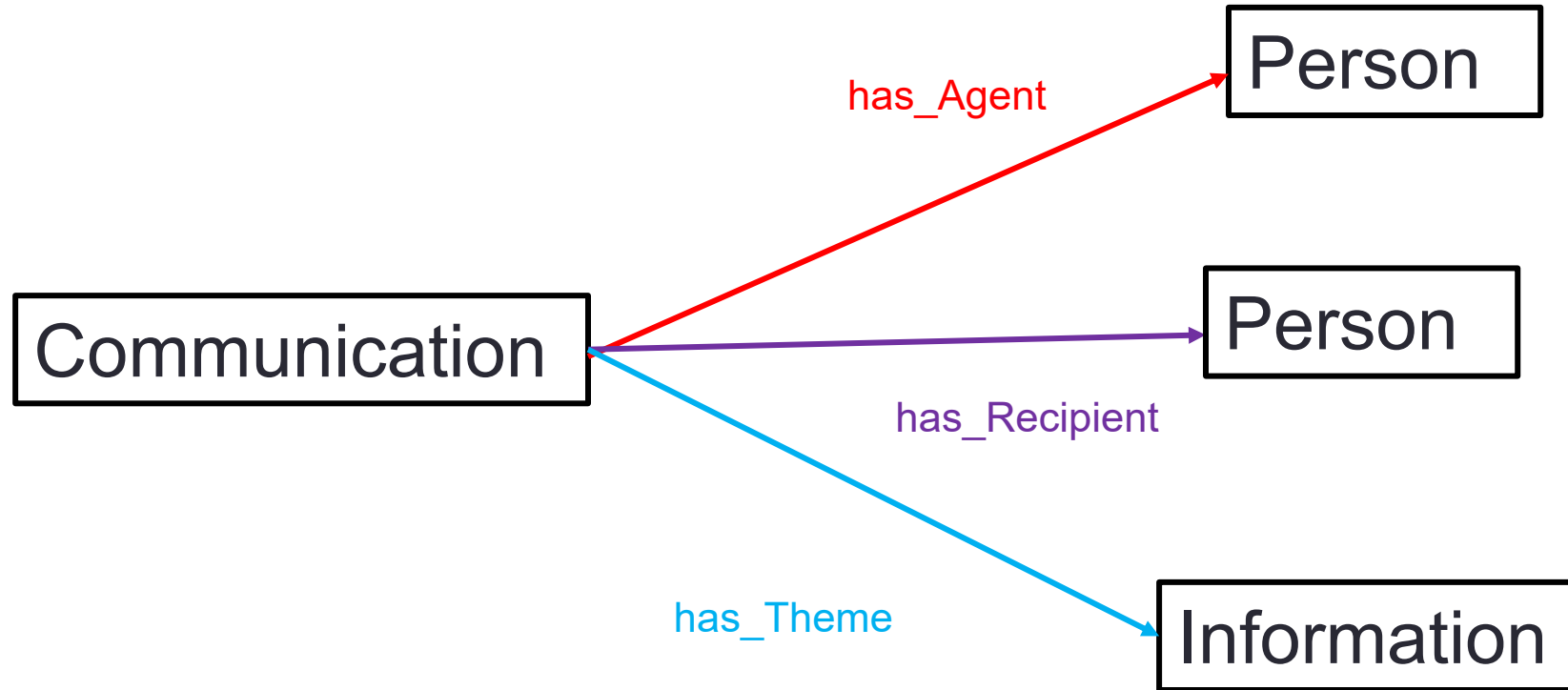
Similar to scripts but can be more complicated

- E.g., contain hierarchical nodes





Event-Object Relations



Creation

Label	ArtifactExistence.Creation
Description	The act of creation or invention in which an entirely novel and unique physical or informational entity (or event) is formed for the first time from raw materials or components, either intentionally or through a causative event
Slot Role	Slot Argument Constraints
Creator	per, org, gpe, sid, event
Thing created	abs, fac, com, veh, wea, pth, inf, event?
Components/Materials	com, nat
Place	fac, loc, gpe
Temporal	
Start and End	(times specific to event)
Duration	1 minute through multiple years

Wear

Label	Wear
Description	Bearing or having clothing or other objects on the person

Slot Role	Slot Argument Constraints
Wearer	per
Thing worn	com
Body_Location	bod
Place	fac, loc, gpe

Temporal	
Start and End	(times specific to event)
Duration	1 minute through multiple years

Sanitize

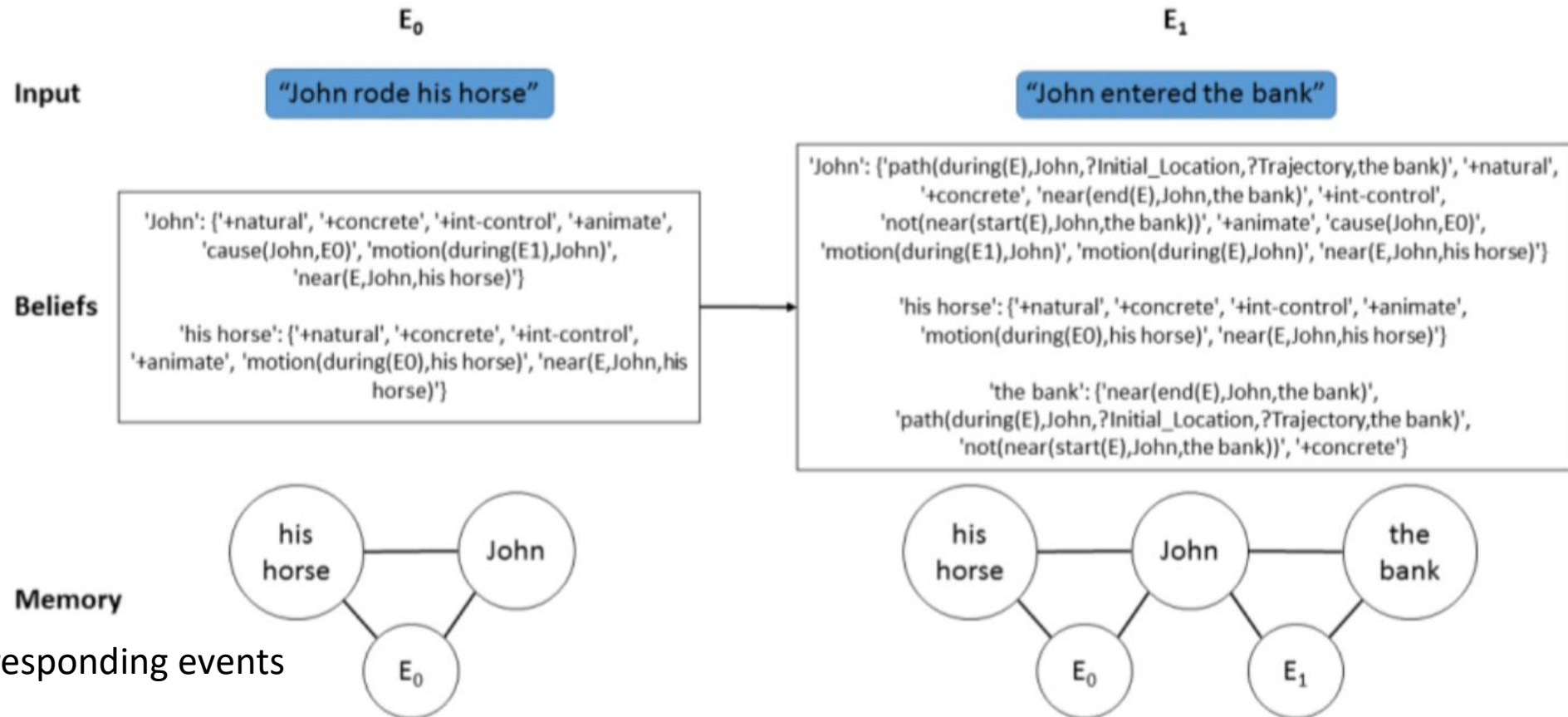
Label	Sanitize
Description	Rendering pathogens harmless through methods including use of heat, antiseptics and antibacterial agents
Slot Role	Slot Argument Constraints
Agent/Sanitizer	per, org, gpe, sid
Sanitized object	fac, com, veh, wea
Sanitizing substance	com, nat
Pathogen	pth
Place	fac, loc, gpe
Temporal	
Start and End	(times specific to event)
Duration	1 minute through multiple years

What is the problem with this approach?

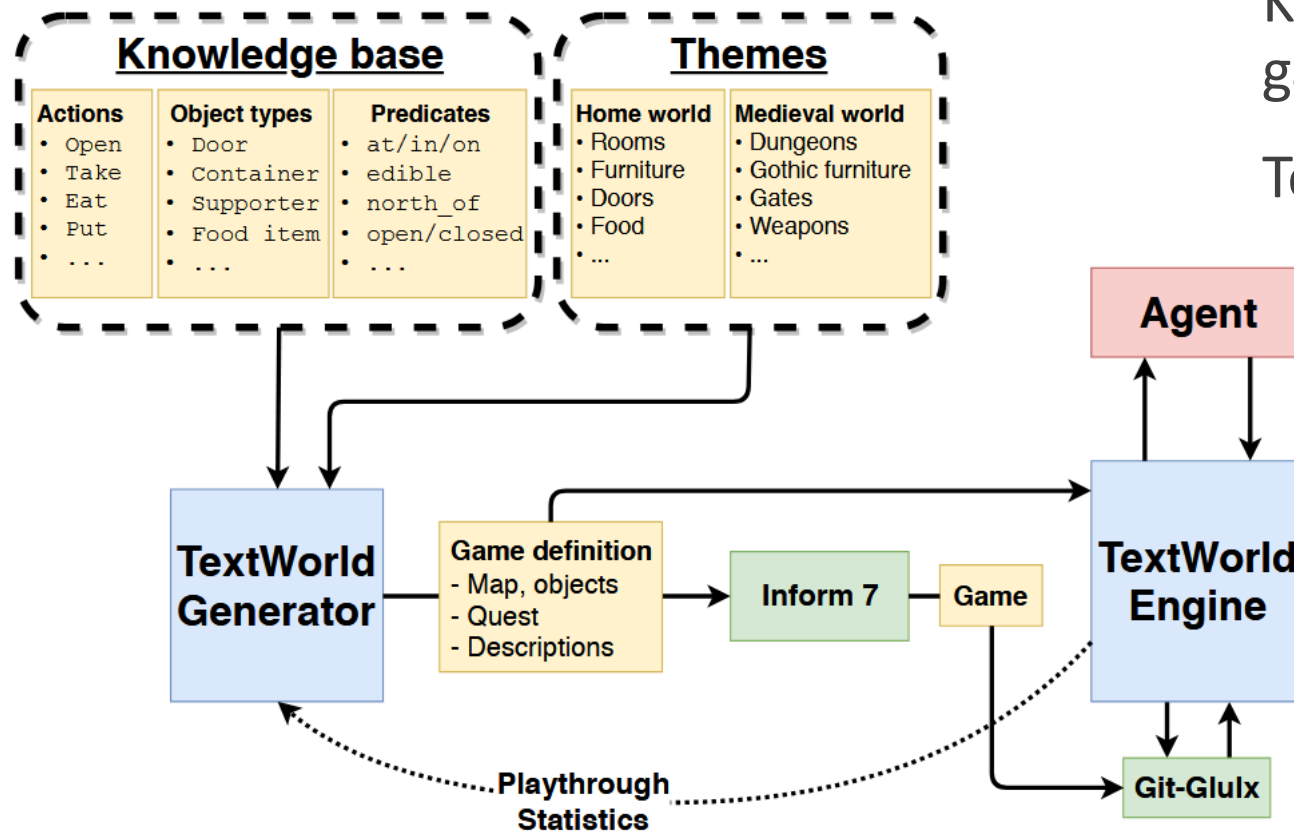
What might happen when using an ontology like this in an actual application?

World Modeling with Symbolic Systems

State Tracking Using VerbNet



TextWorld



Knowledge-guided building of text games

Text generation using a CFG

StoRM

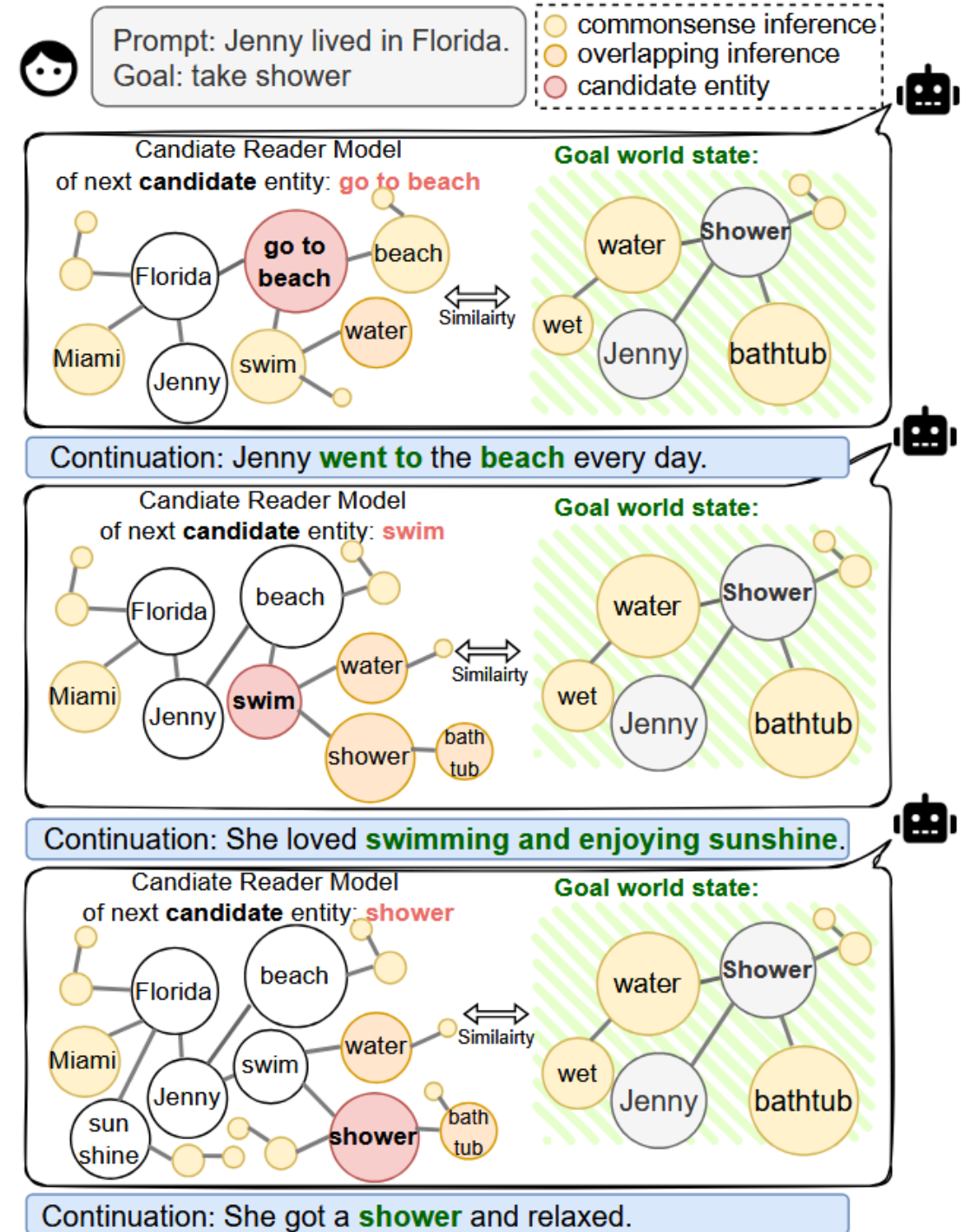
Reader model and goal world states compared

Tuples extracted from sentences are turned into knowledge graphs using semantic role labeling (i.e., filling in roles of frames)

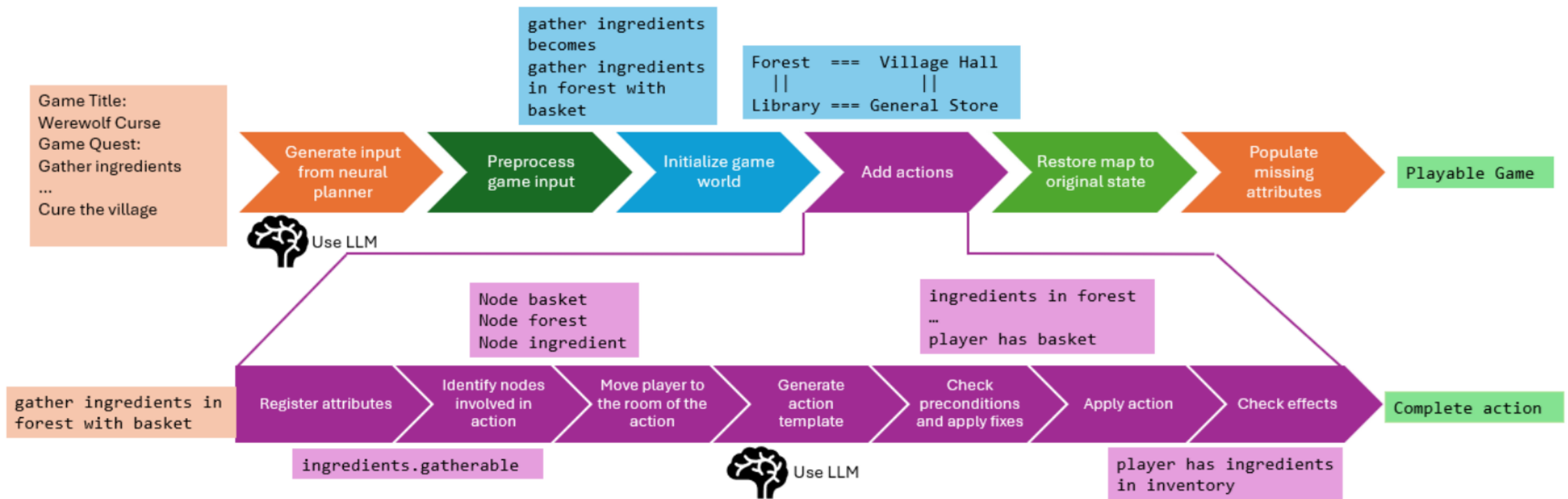
KG expanded with COMET2020

Templates were filled in with RoBERTa (infilling)

- [subject] <mask> <mask> swim <mask>



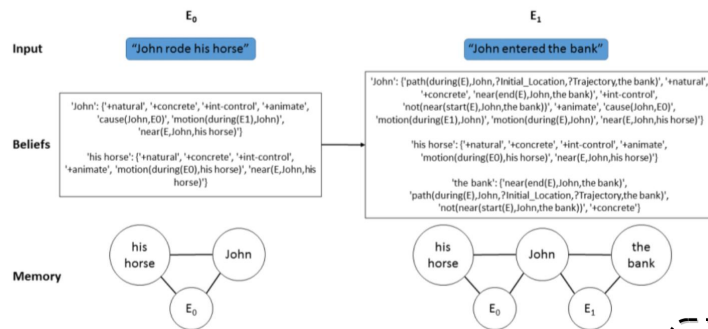
Story2Game



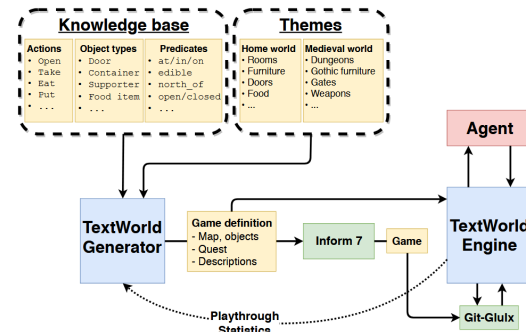
Think-Pair-Share

How has the way we track world states or modeled worlds changed over the years?

What has stayed the same?

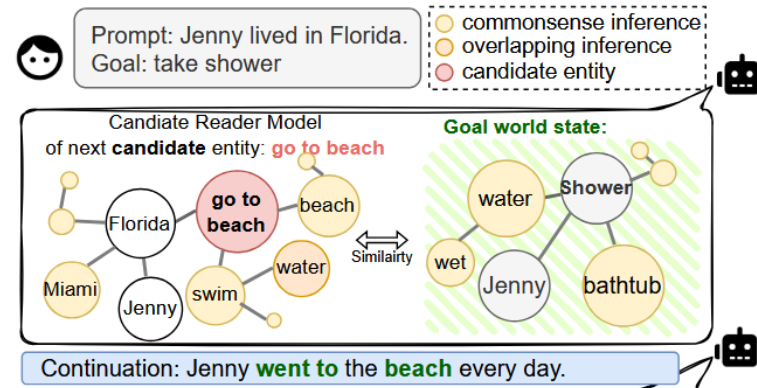


2016

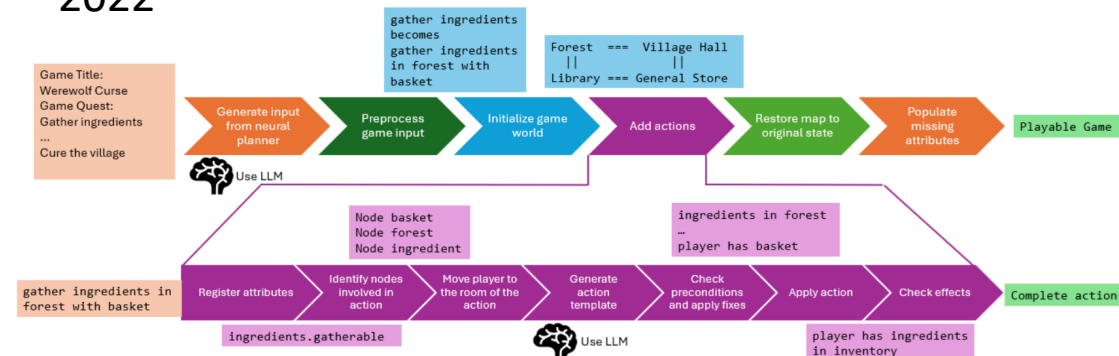


2018

ONTOLOGIES AND WORLD MODELING



2022



2025