Ontologies

Lara J. Martin (she/they)

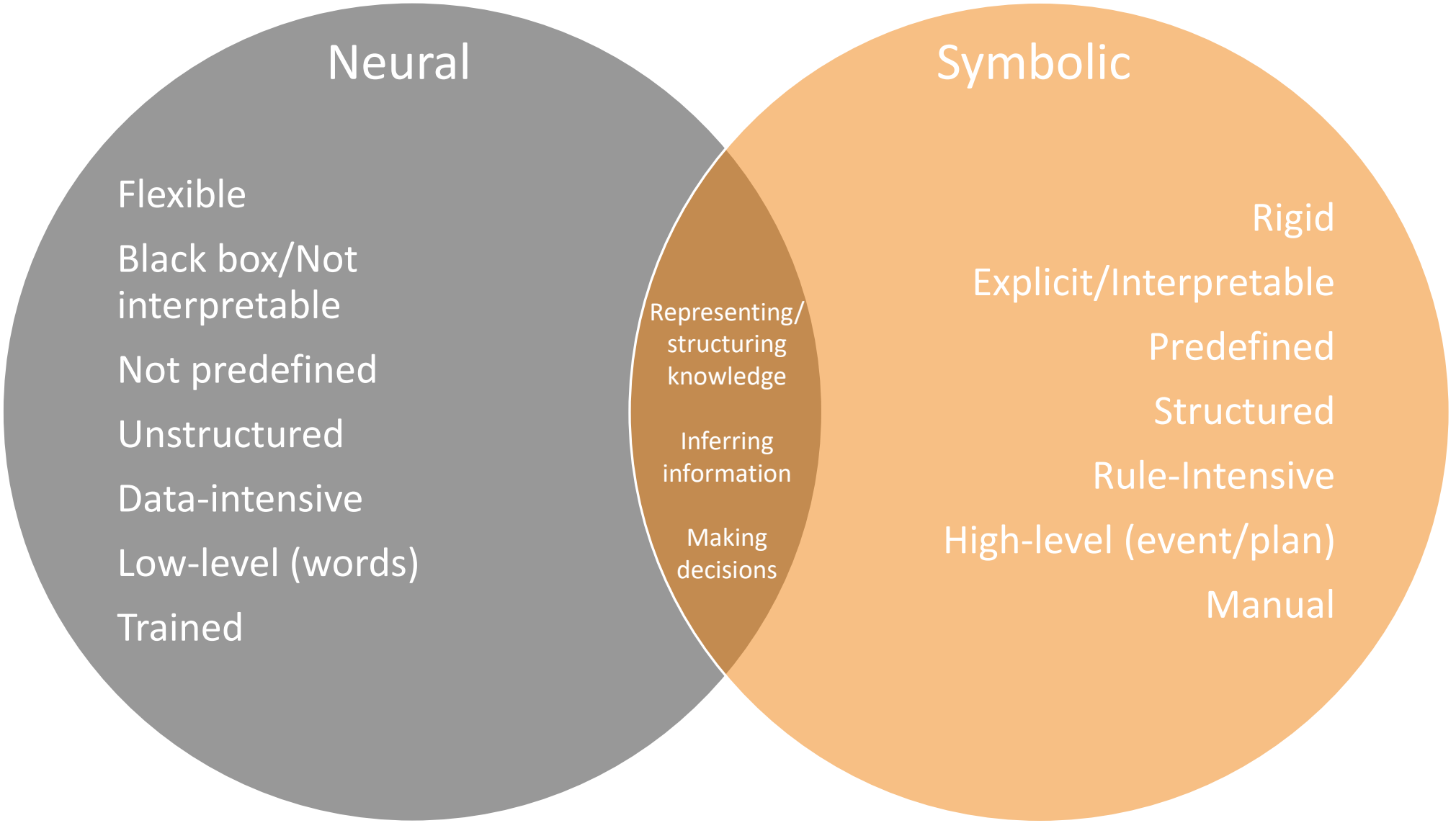
https://laramartin.net/interactive-fiction-class

Modified from slides from Susan Brown

Learning Objectives

Define an ontology and its components

Compare scripts, knowledge graphs, and ontologies


Neural
Flexible
Black box/Not interpretable
Not predefined
Unstructured
Data-intensive
Low-level (words)
Trained
Representing/ structuring knowledge
Inferring information
Making decisions
Symbolic
Rigid
Explicit/Interpretable
Predefined
Structured
Rule-Intensive
High-level (event/plan)
Manual

Review: Ways of combining neural and symbolic methods

During training

- Such as in reinforcement learning or retrieval-augmented generation (RAG)

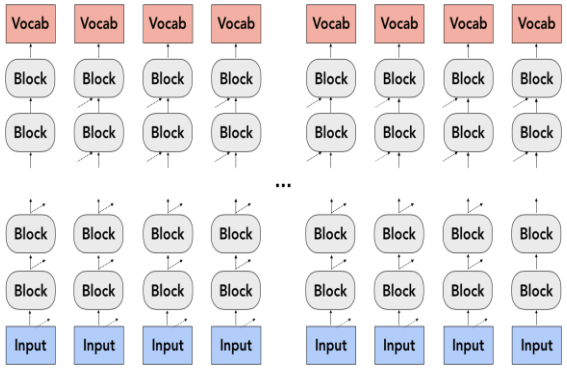
After training

- Like a symbolic “wrapper” – helps validate what the NN is doing

Others??

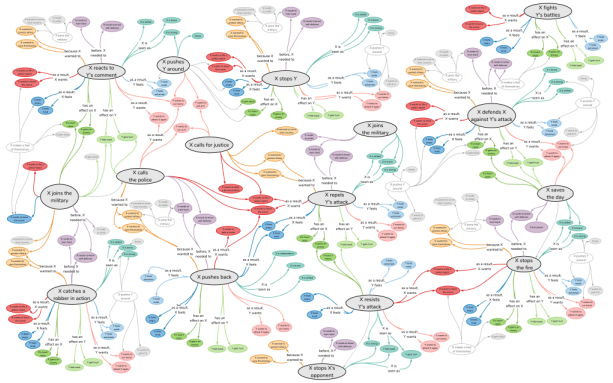
Commonsense Transformers

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



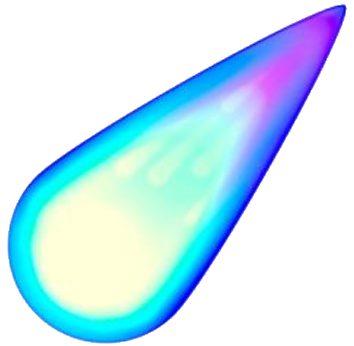
Pre-trained
Language Model

+



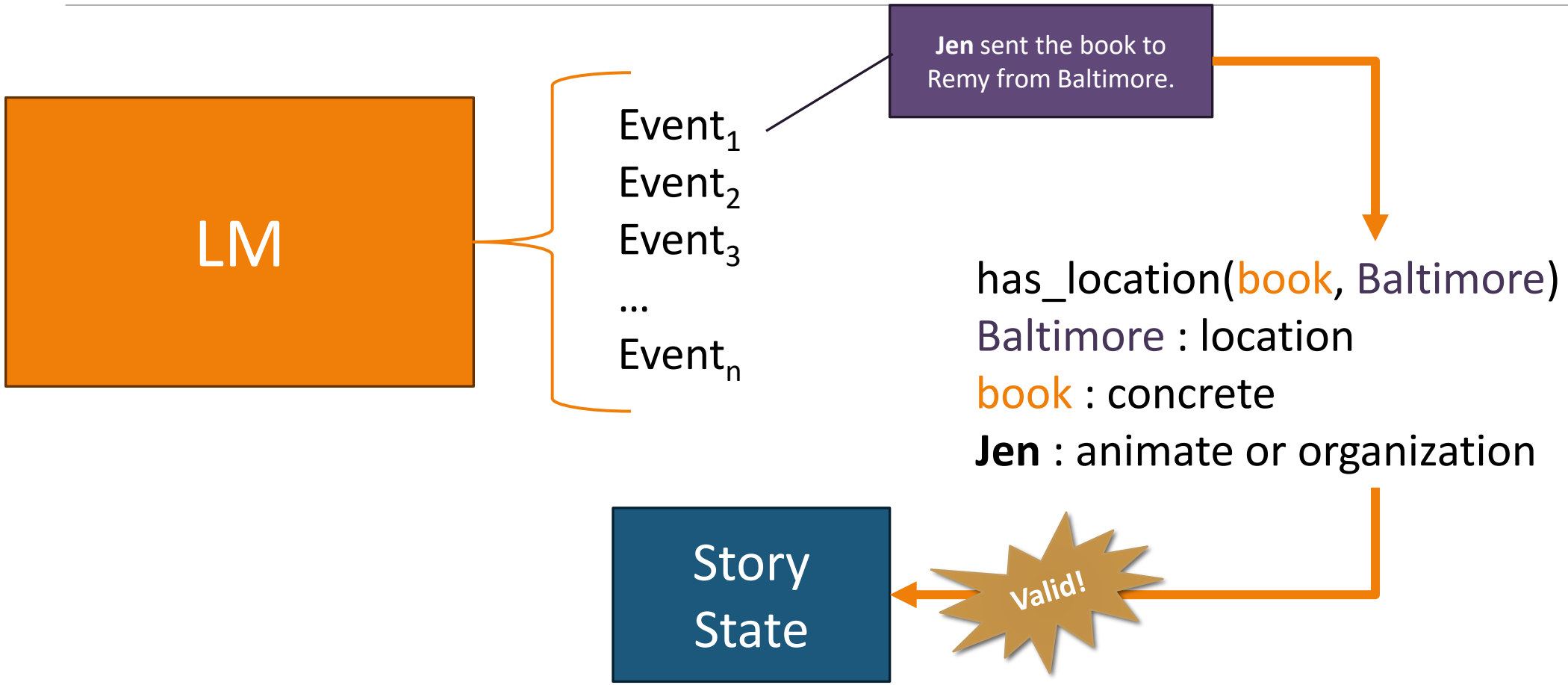
Seed Knowledge
Graph Training

=



COMET

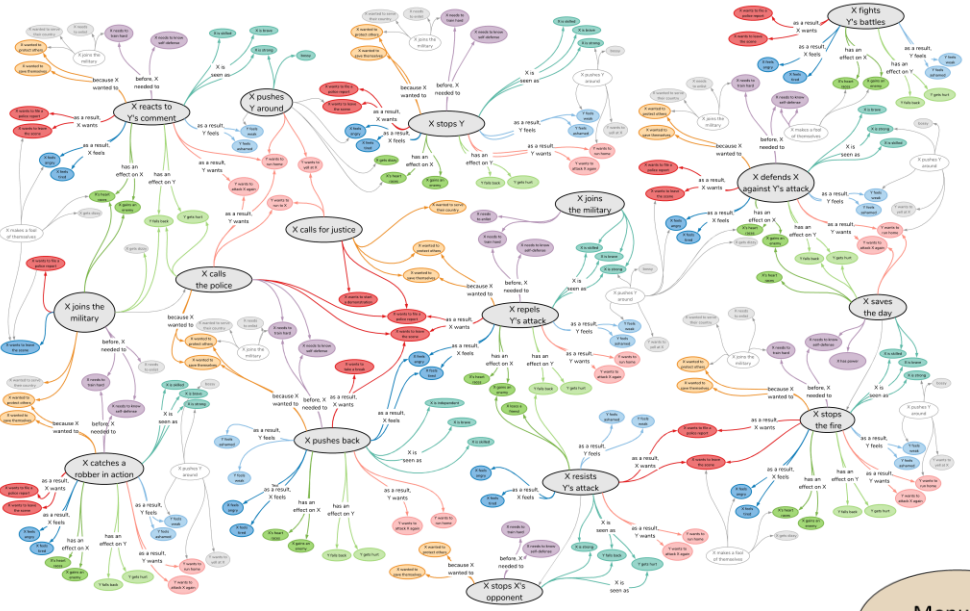
How does a neural network fit in here?

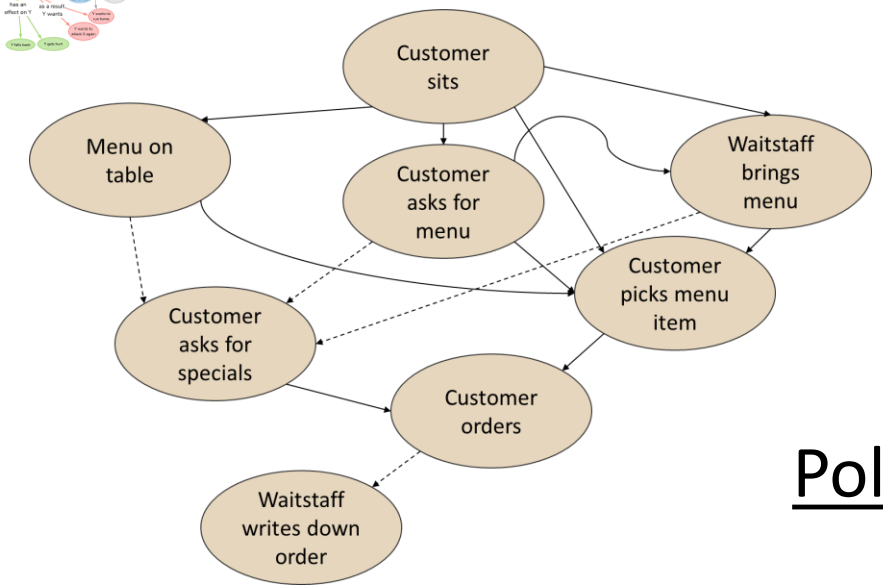


A quick question

What is a *script*?

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





PollEv.com/laramartin527

Ontologies

Semantic representations and predicate logic

Franco likes Frasca.

First order logic:

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

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

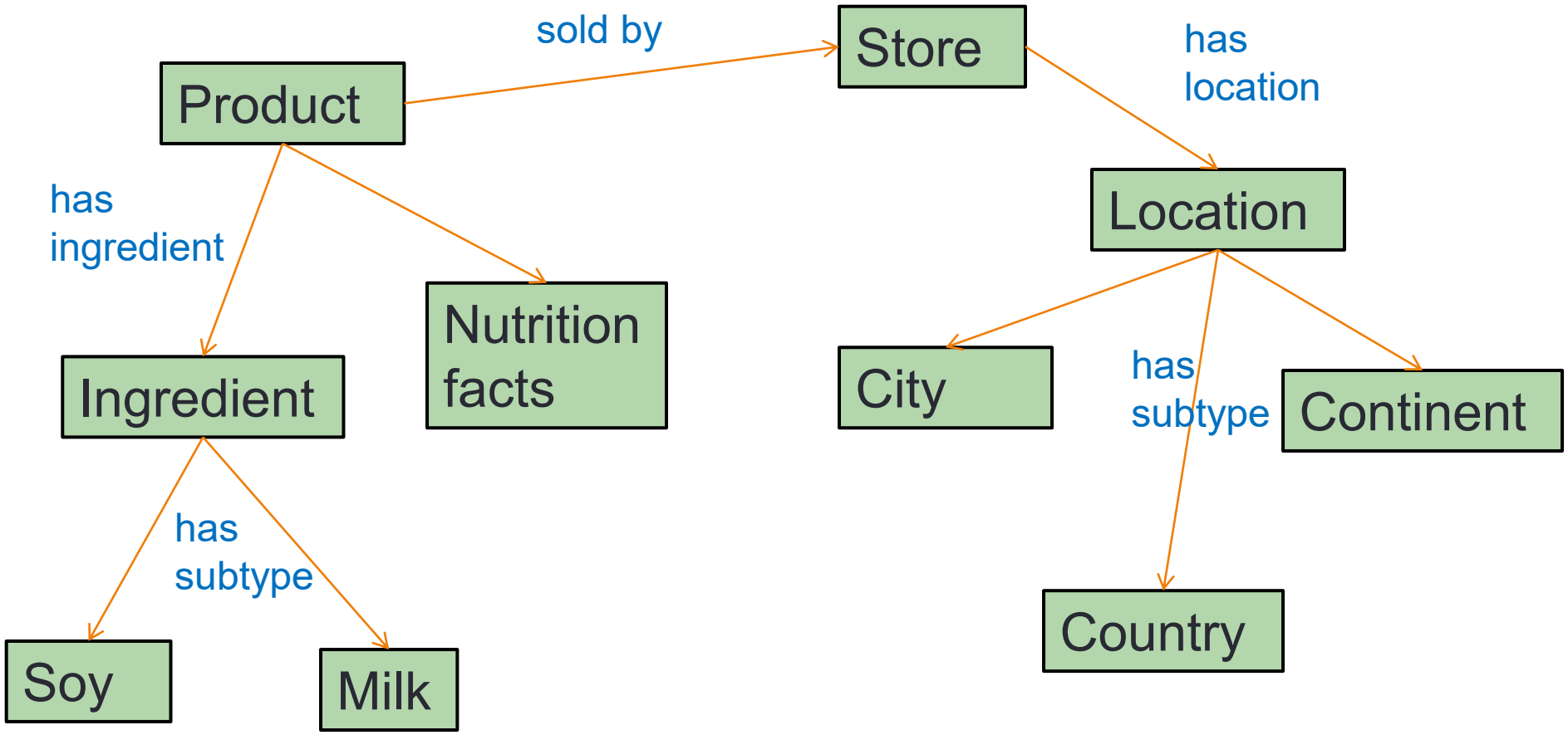
Defines

- 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

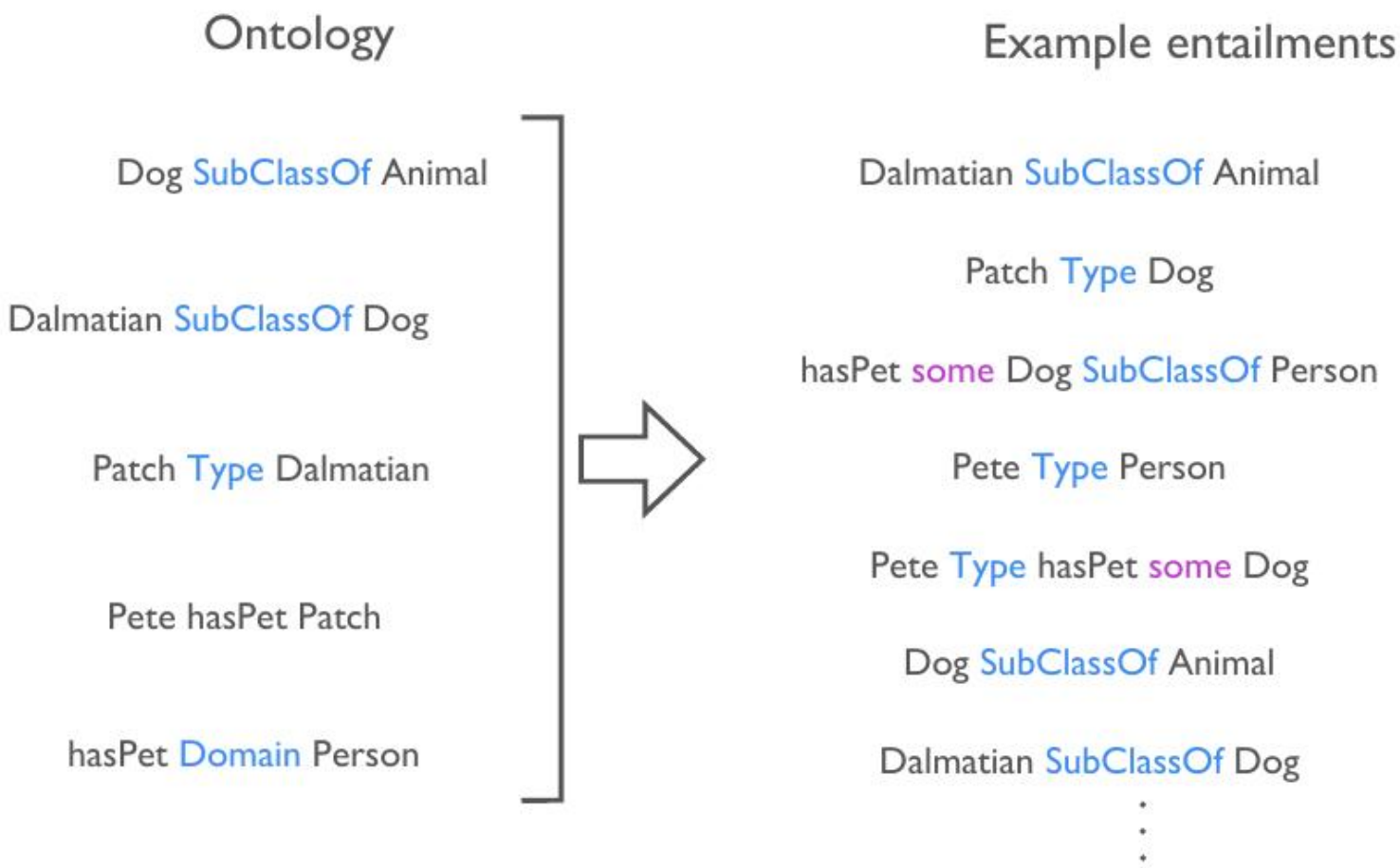
Person and hasPet some Cat

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


Ontology
Dog SubClassOf Animal
Dalmatian SubClassOf Dog
Patch Type Dalmatian
Pete hasPet Patch
hasPet Domain Person
Example entailments
Dalmatian SubClassOf Animal
Patch Type Dog
hasPet some Dog SubClassOf Person
Pete Type Person
Pete Type hasPet some Dog
Dog SubClassOf Animal
Dalmatian SubClassOf Dog

Defining classes

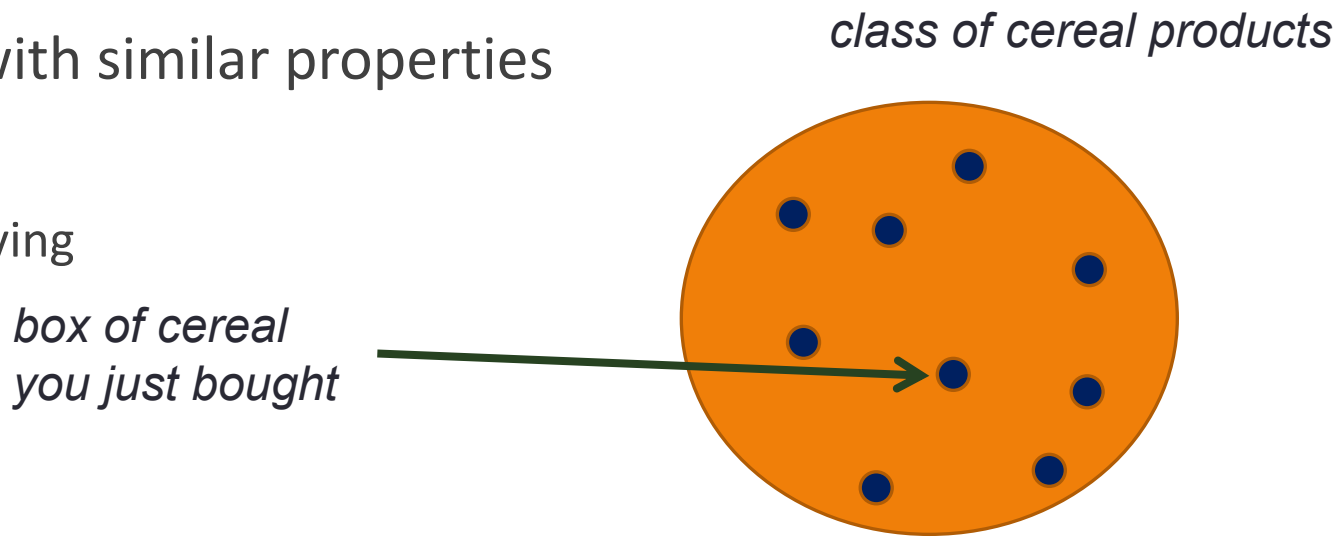
A class is a concept in the domain

- 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

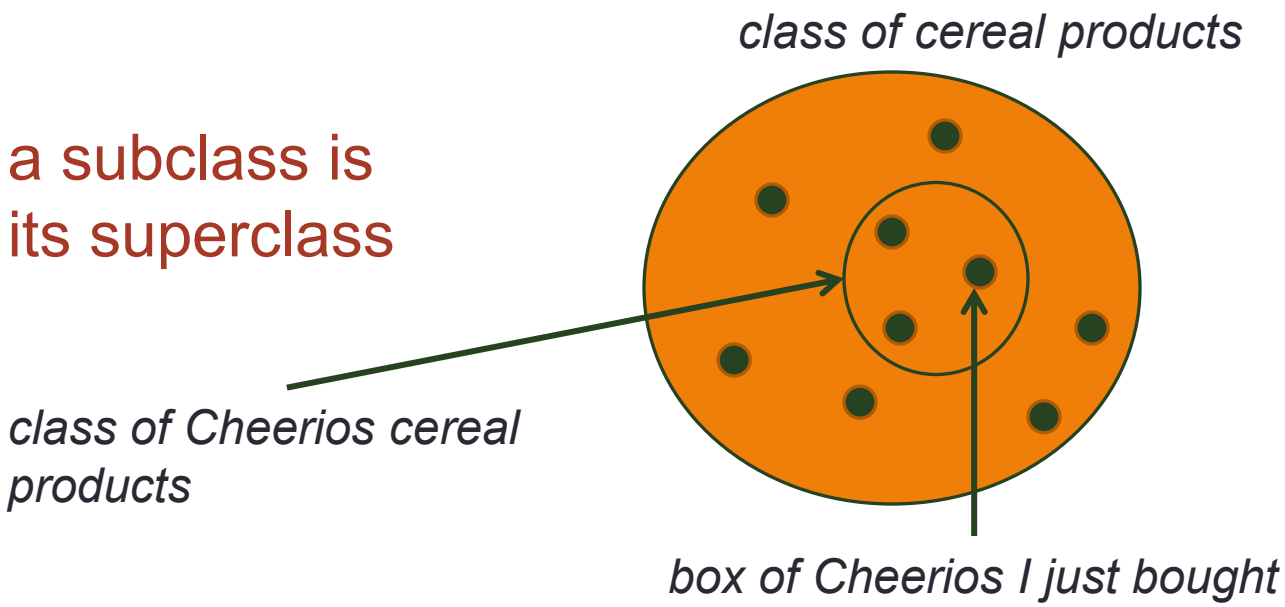


Class inheritance

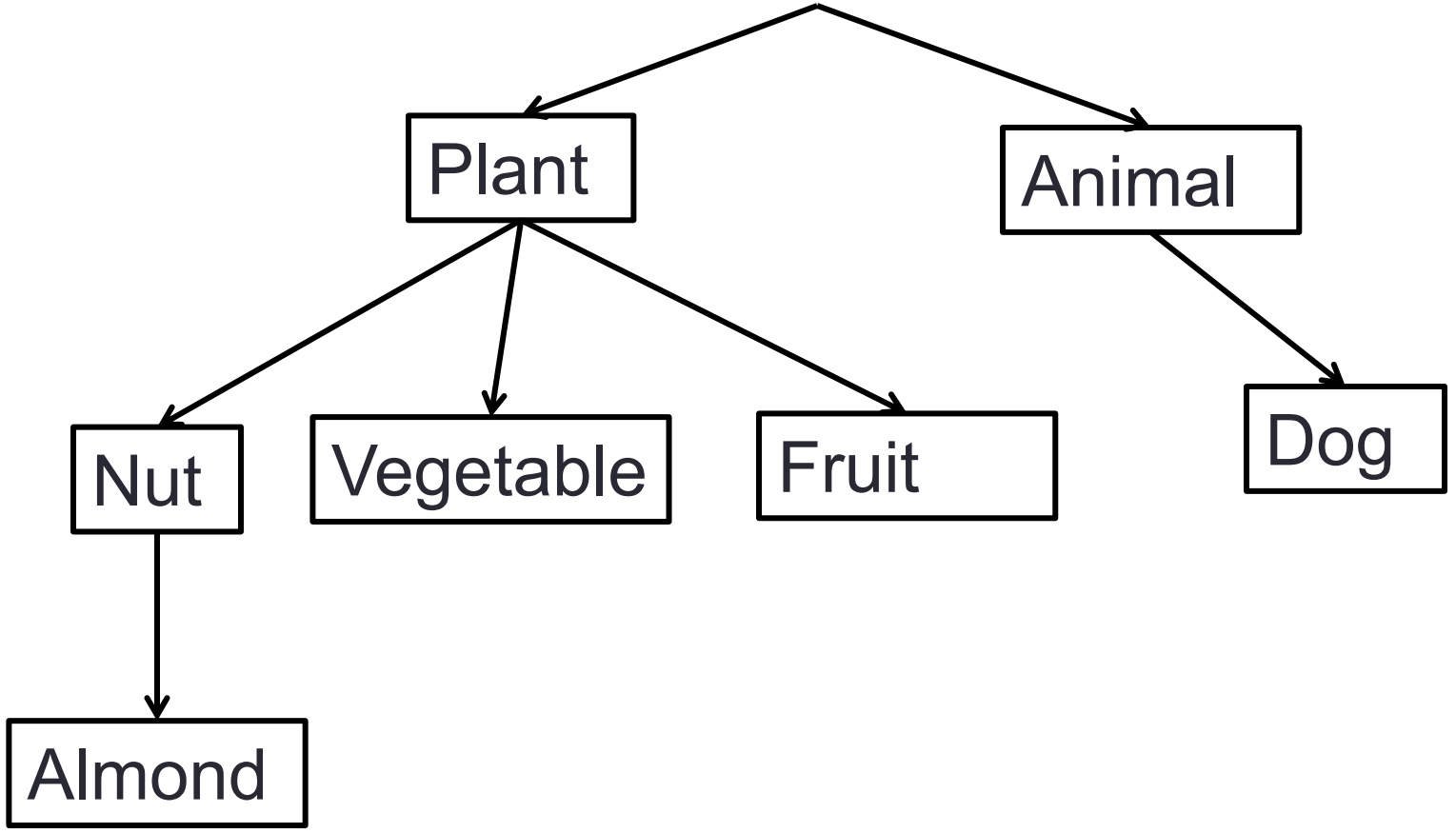
Classes usually constitute a taxonomic hierarchy (a subclass-superclass hierarchy)

an IS-A hierarchy:

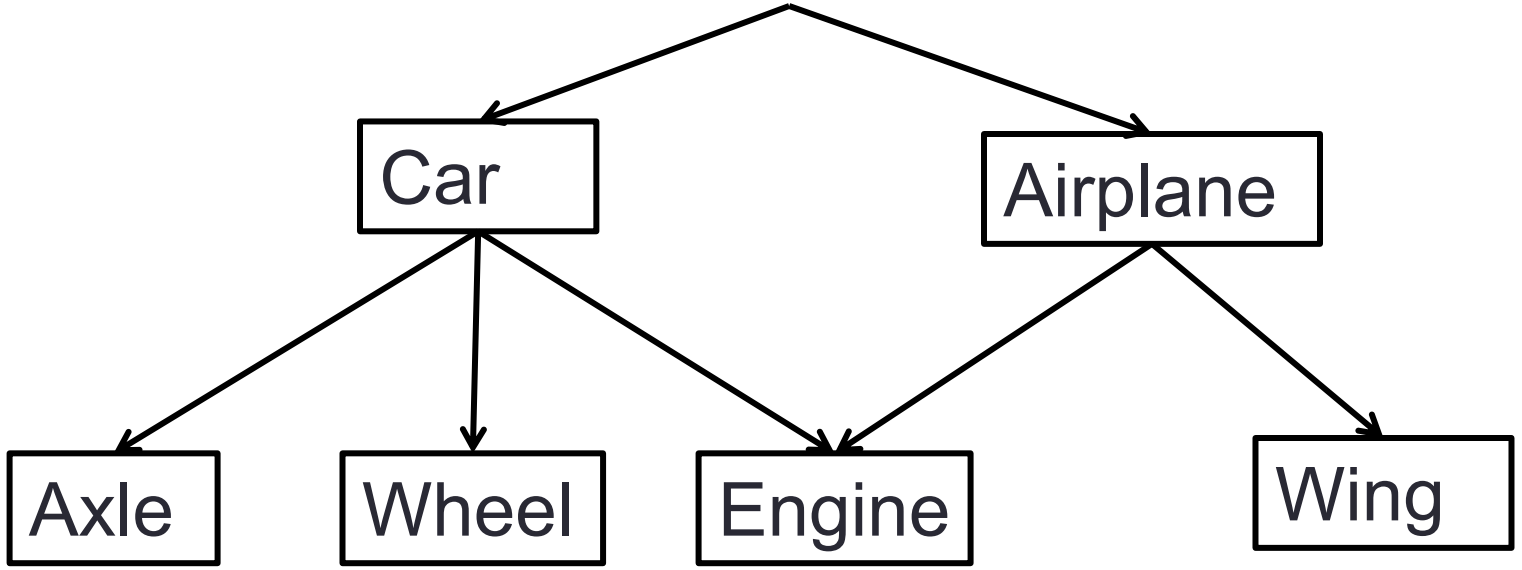
an instance of a subclass is
an instance of its superclass



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

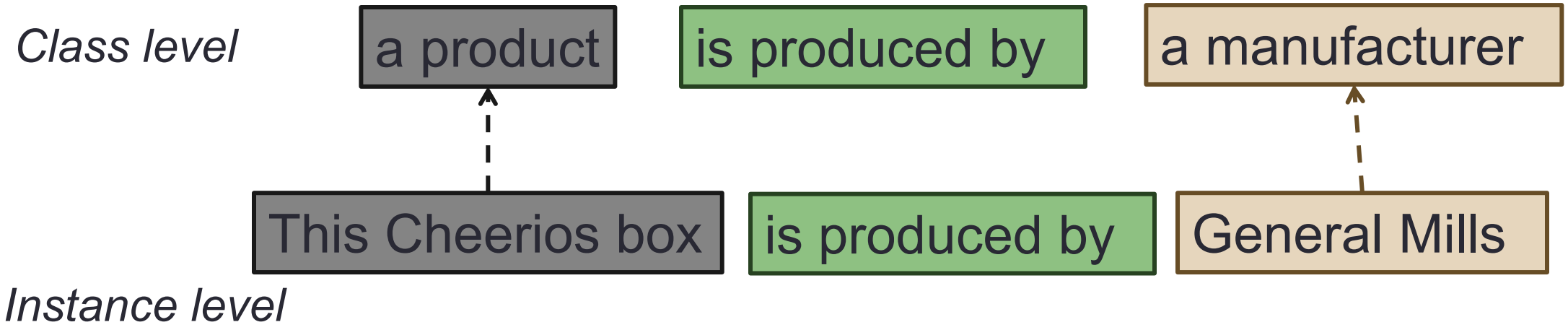
Products ***are produced by*** a manufacturer → produced by

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

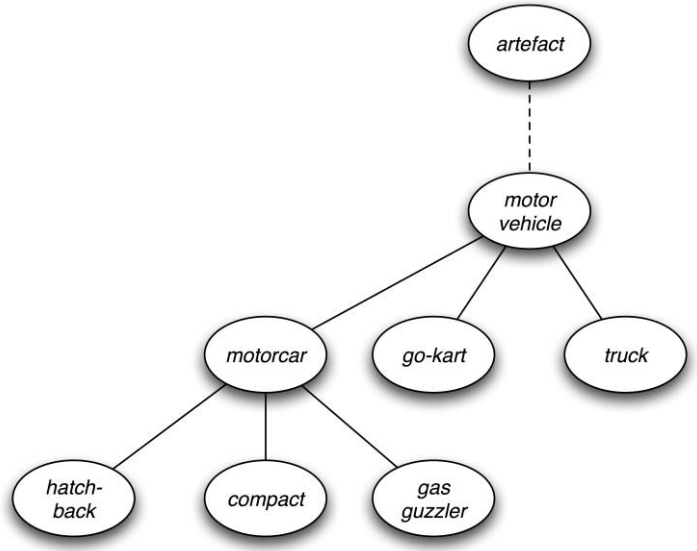
- is-a relations tricky to determine (killing, crime, murder, death)
- where does an event start and end? (surgical event)

Usually represented as relations between entities

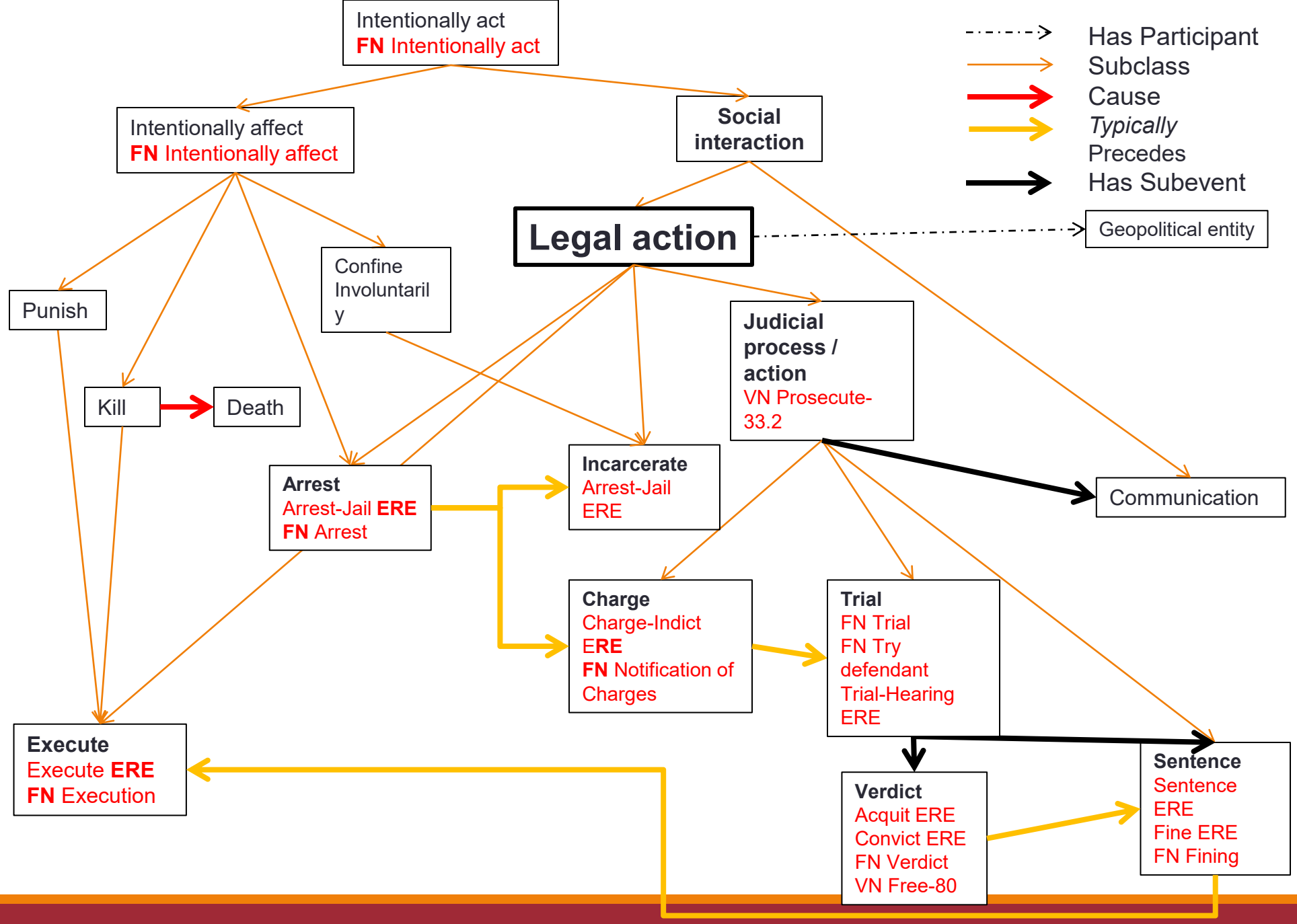
- relations can't have links to lexical items
- relations can't have individuals (you might want to make lexical items the individuals or instances in annotation)

Most ontologies have shallow models of events

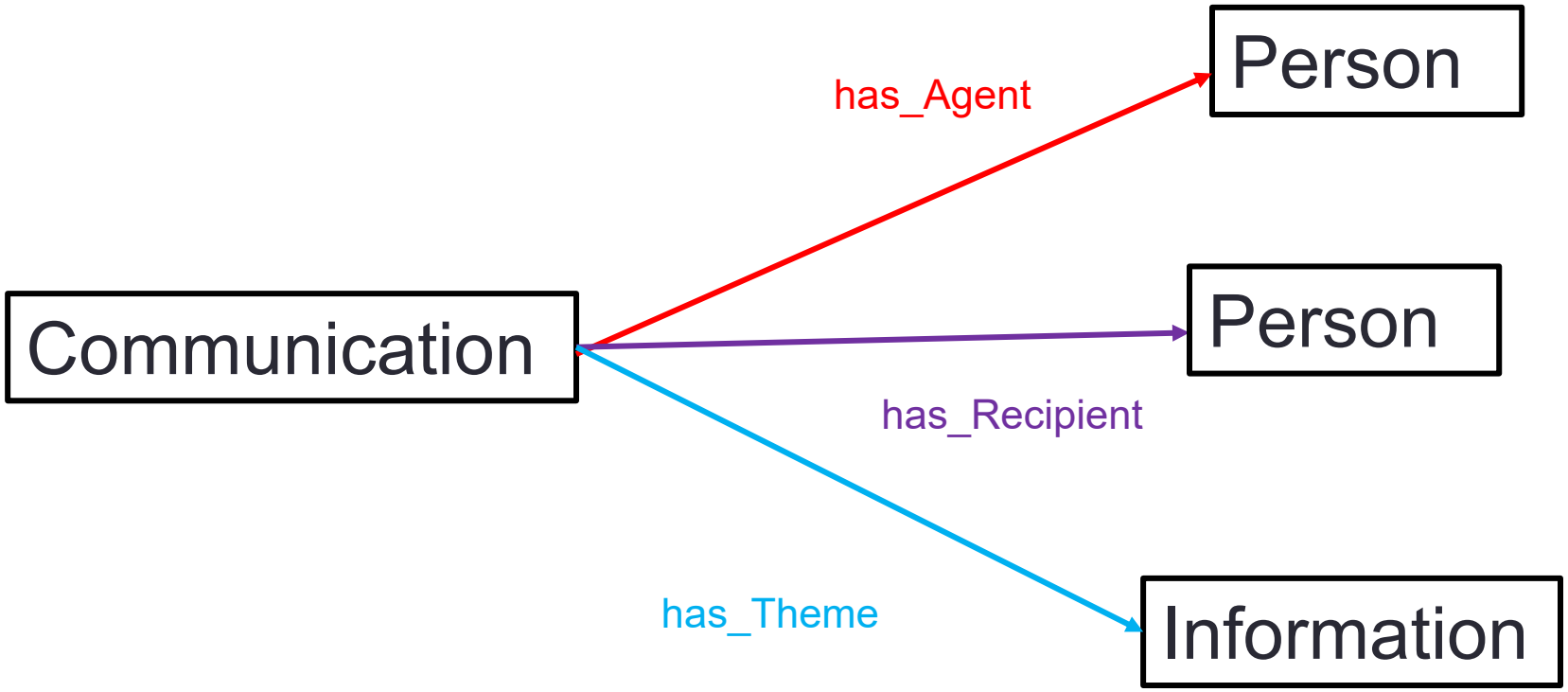
- WordNet
- SUMO (Suggested Upper Merged Ontology)



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


Intentionally act
FN Intentionally act
Intentionally affect
FN Intentionally affect
Social interaction
Has Participant
Subclass
Cause
Typically Precedes
Has Subevent
Legal action
Geopolitical entity
Punish
Confine Involuntaril y
Kill
Death
Judicial process / action
VN Prosecute-33.2
Arrest
Arrest-Jail ERE
FN Arrest
Incarcerate
Arrest-Jail
ERE
Communication
Charge
Charge-Indict
ERE
FN Notification of Charges
Trial
FN Trial
FN Try defendant
Trial-Hearing
ERE
Execute
Execute ERE
FN Execution
Verdict
Acquit ERE
Convict ERE
FN Verdict
VN Free-80
Sentence
Sentence
ERE
Fine ERE
FN Fining

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	person, organization, geopolitical entity, different sides of a conflict, event
Thing created	abstract, facility, tangible product, vehicle, weapon, pathogen, information, event?
Components/Materials	tangible product, natural materials
Place	facility, location, geopolitical entity

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	person
Thing worn	tangible product
Body_Location	body part
Place	facility, location, geopolitical entity

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	person, organization, geopolitical entity, different sides of a conflict
Sanitized object	facility, tangible product, vehicle, weapon
Sanitizing substance	tangible product, natural materials
Pathogen	pathogen
Place	facility, location, geopolitical entity
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?

- The system might have problems staying at the appropriate level
- Are all entities, events, properties captured? And if it's too big, can it be processed? (scalable)
- How do events interact?