

Moving toward Formalisation

Yizheng Zhao

zhaoyz@nju.edu.cn

Previously...

- We started the **knowledge acquisition** process...
 - to elicit tacit knowledge
 - in a variety of ways
 - about a set of terms (or concepts)
- But even there we could get **more explicit**
 - normalising terms (e.g., “symmetry or symmetric”?)
 - hierarchy (and other direct relations between terms)
 - categorizing terms (e.g., as modifiers or self-standing)
 - **constraining** and **defining** terms
- 2 important next steps
 1. getting even more explicit and precise
 - Refining our proto-representation
 2. getting actionable
 - Building a representation

Term extraction

- Highlight the **relevant, domain-dependent** terms in:
 - There are several sorts of domesticated animals, though by far the most are mammals (like us!). For example, our faithful pets, cats and dogs, are clearly domesticated (or we would not keep such dangerous carnivores in our homes), as is the delicious cow which is farmed in ever increasing numbers.

Step 1: Term extraction

- Highlight the **relevant, domain-dependent** terms in:
 - There are several sorts of **domesticated animals**, though by far the most are **mammals** (like **us!**). For example, our faithful **pets**, **cats** and **dogs**, are clearly **domesticated** (or we would not keep such **dangerous carnivores** in our **homes**), as is the **delicious*** yet **docile cow** which is **farmed** in ever **increasing numbers**.

Step 1: Term extraction

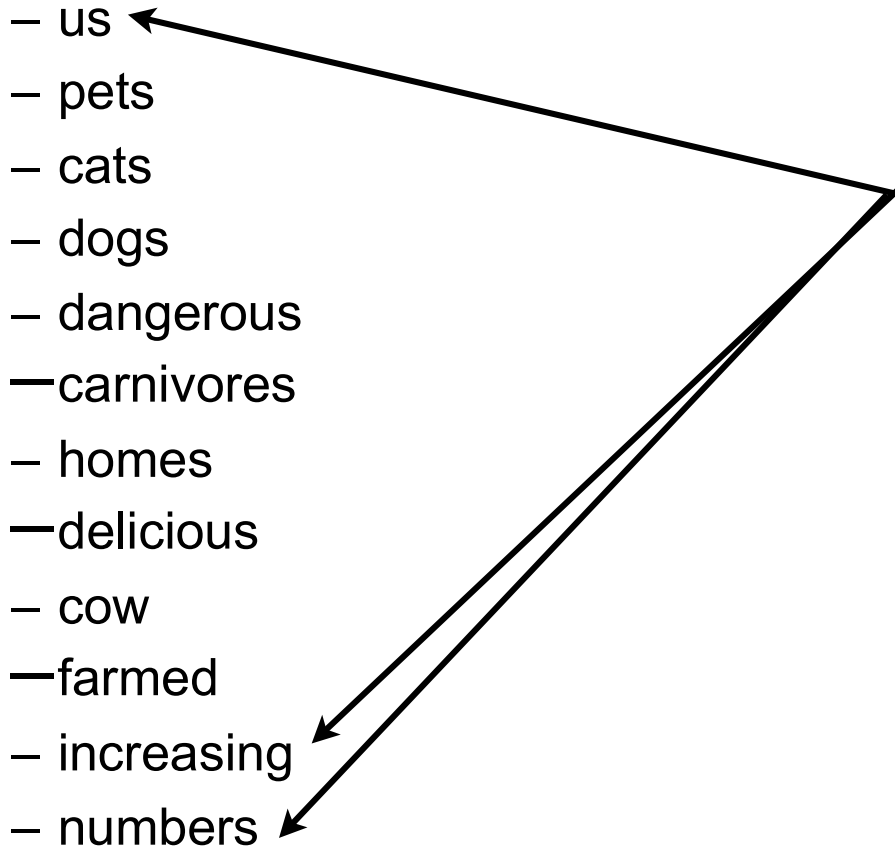
- We pull these out
 - domesticated
 - animals
 - mammals
 - us
 - pets
 - cats
 - dogs
 - dangerous
 - carnivores
 - homes
 - delicious
 - cow
 - farmed
 - increasing
 - numbers

Step 1: Term extraction

- We pull these out and **ponder**:

- domesticated
- animals
- mammals
- us
- pets
- cats
- dogs
- dangerous
- carnivores
- homes
- delicious
- cow
- farmed
- increasing
- numbers

These are quite odd
but in different ways



Step 1: Term extraction

- We pull these out and **ponder some more**:

- domesticated

- animals

- mammals

- us

- pets

- cats

- dogs

- dangerous

- carnivores

- homes

- delicious

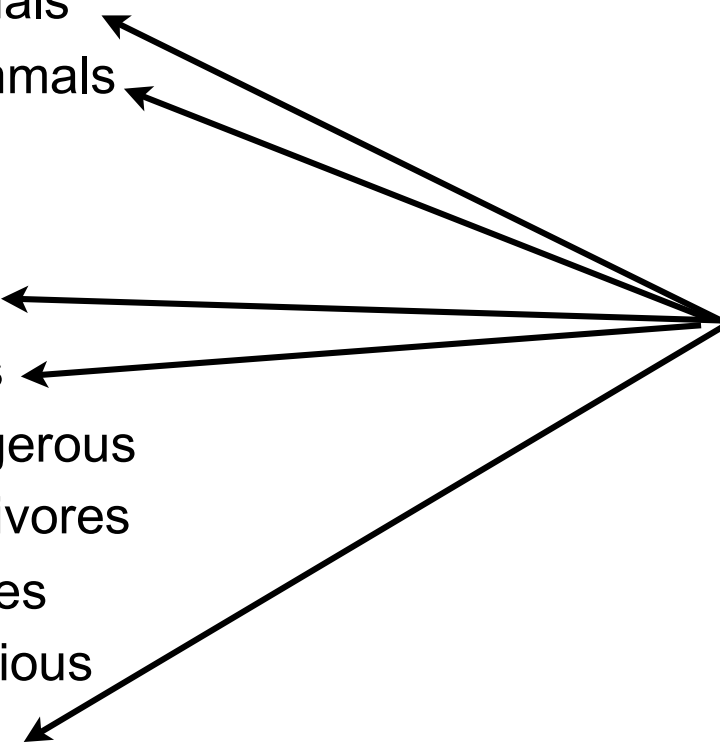
- cow

- farmed

- increasing

- numbers

These are similar
but have different levels
of generality, and
non-uniform spelling



Step 2: Grouping

- Base animal categories (noun-y terms)
 - animals
 - cats
 - dogs
 - mammals
 - cow
 - us
- Ways an animal can be (adjective-y terms)
 - domesticated
 - pets
 - dangerous
 - carnivores
 - delicious
 - farmed
- Stuff
 - homes
 - increasing
 - numbers

Step 2: Grouping

- Base animal categories (noun-y terms)
 - animals
 - cats
 - dogs
 - mammals
 - cow
 - us
- Ways an animal can be (adjective-y terms)
 - domesticated
 - pets
 - dangerous
 - carnivores
 - delicious
 - farmed
- Stuff
 - homes
 - increasing
 - numbers

Should we care about these?



A Key Slogan

to determine which terms to care about:

Representations are context sensitive & interest relative

- Context sensitive?
 - for which (kind of) application do we build KR?
- Interests?
 - Application needs
 - Teaching, categorising, data acquisition
 - Audience
 - Children, lay people, different disciplines, clinicians vs. researchers
- Establish context and relevant interests
 - Here: **context** is this class
 - Here: **interests** is to work up a reasonable example

Step 2: Grouping

- Base animal categories (noun-y terms)
 - animals
 - cats
 - dogs
 - mammals
 - cow
 - us
- Ways an animal can be (adjective-y terms)
 - domesticated
 - pets
 - dangerous
 - carnivores
 - delicious
 - farmed
- ~~Stuff~~
 - ~~– homes~~
 - ~~– increasing~~
 - ~~– numbers~~

Should we care about these?

No! (Why?)

Step 3: Normalise Terms

- Base animal categories (noun-y terms)

- animals
- cats
- dogs
- mammals
- cow
- us

Unify number & spelling

- Ways an animal can be (adjective-y terms)
 - domesticated
 - pets
 - dangerous
 - carnivores
 - delicious
 - farmed

Step 3: Normalise Terms

- Base animal categories (noun-y terms)
 - Animal
 - Cat
 - Dog
 - Mammal
 - Cow
 - **us** Give a good name
- Ways an animal can be (adjective-y terms)
 - domesticated
 - pets
 - dangerous
 - carnivores
 - delicious
 - farmed

Step 3: Normalise Terms

- Base animal categories (noun-y terms)
 - Animal
 - Cat
 - Dog
 - Mammal
 - Cow
 - Human
- Ways an animal can be (adjective-y terms)
 - domesticated
 - pets
 - dangerous
 - carnivores
 - delicious
 - farmed

Unify grammatical form & spelling

Step 3: Normalise Terms

- Base animal categories (noun-y terms)
 - Animal
 - Cat
 - Dog
 - Mammal
 - Cow
 - Human
- Ways an animal can be (adjective-y terms)
 - Domesticated
 - Pet
 - Dangerous
 - Carnivorous
 - Delicious
 - Farmed

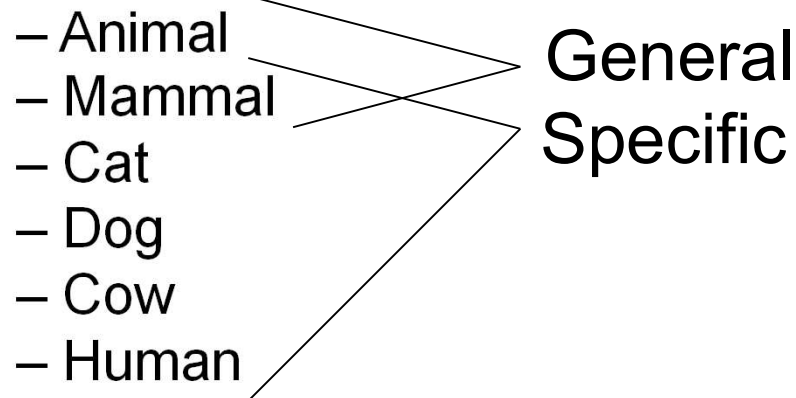
We have some background knowledge we can use to “round out” these terms

Step 3: Normalise Terms

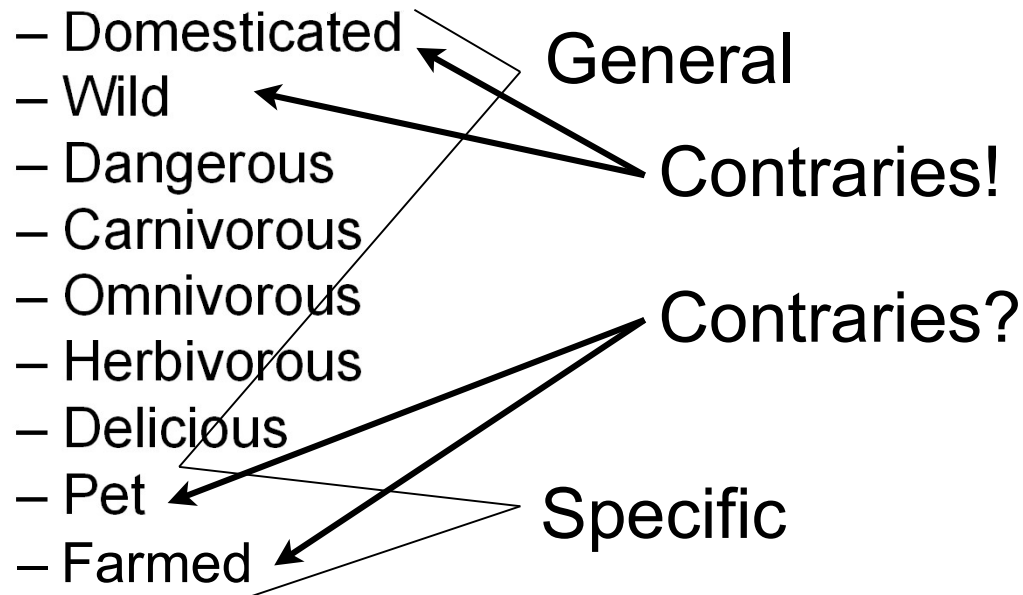
- Base animal categories (noun-y terms)
 - Animal
 - Cat
 - Dog
 - Mammal
 - Cow
 - Human
- Ways an animal can be (adjective-y terms)
 - Domesticated
 - Pet
 - Dangerous
 - Carnivorous
 - Omnivorous
 - Herbivorous
 - Delicious
 - Wild
 - Farmed

Step 4: Organise Terms

- Base animal categories (noun-y terms)



- Ways an animal can be (adjective-y terms)



Step 4: Organise Terms

- Base animal categories (noun-y terms)
 - General:
 - Animal
 - Mammal
 - Specific:
 - Cat
 - Dog
 - Cow
 - Human
- Ways an animal can be (adjective-y terms)
 - General:
 - Domesticated
 - Wild
 - Dangerous
 - Carnivorous
 - Omnivorous
 - Herbivorous
 - Delicious
 - Specific:
 - Pet
 - Farmed

Next:

What terms are definable?

Interlude: what is a definition?

- Mini-exercise:
- in the next 3 minutes,
agree with your neighbour on a **definition** for
 - pet
 - table (furniture)

Interlude: what is a definition?

- a statement that describes/fixes the meaning of a term
- can be
 - extensional: enumerate all elements a term describes e.g., good for “EU countries”
 - intensional: often using **genus–differentia** pattern
i.e., giving the next more general term (genus) plus
differentiating features for this term and its siblings
e.g., “An **endotherm** is an **organism** that maintains its body at
a metabolically favourable temperature.”

Two consequences:

if Bob is an endotherm, then I know that...

if I find an organism that maintains its temperature..., then

Step 4: Organise Terms

- Base animal categories (noun-y terms)
 - General:
 - Animal
 - **Mammal**
 - Specific:
 - Cat
 - Dog
 - Cow
 - Human
- Ways an animal can be (adjective-y terms)
 - General:
 - Domesticated
 - Wild
 - Dangerous
 - **Carnivorous**
 - **Omnivorous**
 - **Herbivorous**
 - Delicious
 - Specific:
 - **Pet**
 - **Farmed**

Which terms are **easily** definable?

Which Terms are Definable?

- Base animal categories (noun-y terms)

- General:

- Animal = eats some Stuff
 - Mammal = has MammGlands

- Specific:

- Cat
 - Dog
 - Cow = eats only Grass
 - Human = Omnivore

- Ways an animal can be (adjective-y terms)

- General:

- Domesticated
 - Wild
 - Dangerous
 - Carnivorous = eats only Meat
 - Omnivorous = eats Meat & Plants
 - Herbivorous = eats only Plants
 - Delicious = tastes good

- Specific:

- Pet = lives with Humans
 - Farmed = is eaten/used

New Terms:
eats, lives, tastes...
= , only, &
Stuff
Plants, Meat,...

A first regimentation

- Base animal categories (noun-y terms)

- General:

- 1. Animal = eats some Stuff
 - 2. Mammal = has MammGlands

- Specific:

- Cat
 - Dog
 - 3. Cow = eats only Grass
 - 4. Human = Omnivore

- Ways an animal can be (adjective-y terms)

- General:

- Domesticated
 - Wild
 - Dangerous

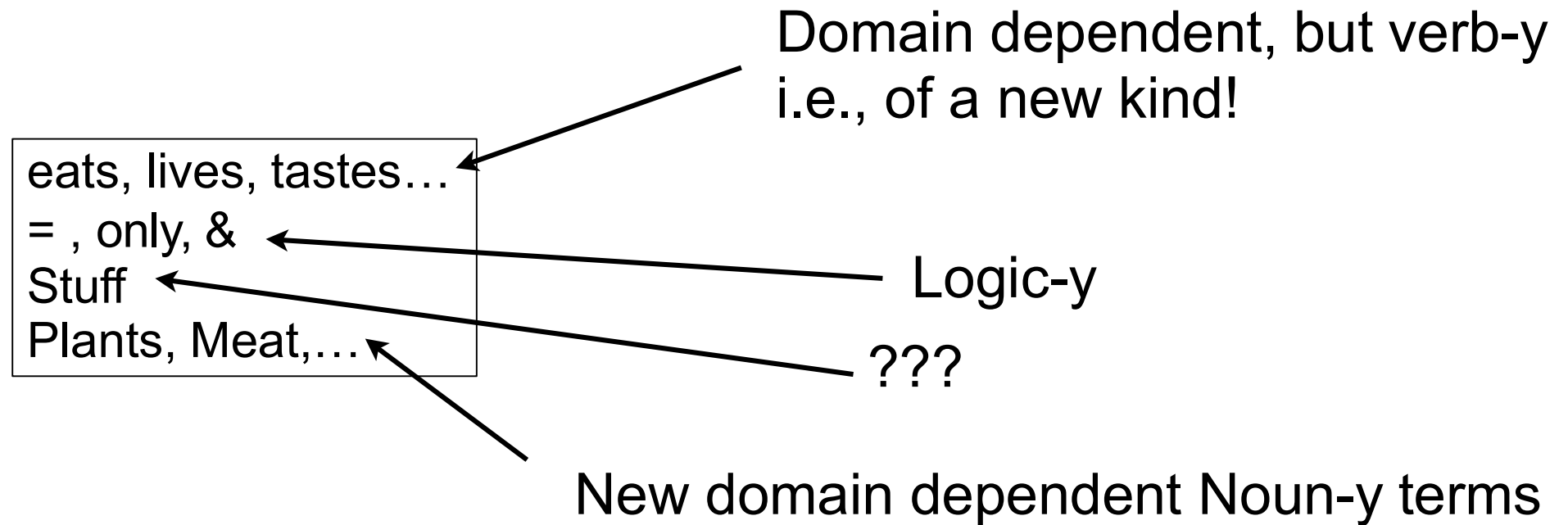
- 5. Carnivorous = eats only Meat
 - 6. Omnivorous = eats Meat & Plants
 - 7. Herbivorous = eats only Plants
 - 8. Delicious = tastes good

- Specific:

- 9. Pet = lives with Humans
 - 10. Farmed = is eaten/used

Which of these definitions
is really good?
I.e., is really a definition?

What about these new terms?



Let's try to formalise!

untitled-ontology-4 (http://www.semanticweb.org/sattler/ontologies/2016/3/untitled-ontology-4) : [/Use...

< > untitled-ontology-4 Search...

Active Ontology x Entities x Classes x Individuals by class x DL Query x

Class hierarchy (inferred)
Class hierarchy

Class hierarchy: Cow

Annotations Usage

Annotations: Cow

Annotations +

`rdfs:comment` Definable

`rdfs:comment` eats only plants

`rdfs:comment` Self-Standing

Description: Cow

Equivalent To +

SubClass Of +

`Mammal`

owl:Thing

Animal

Mammal

Human

Dog

Cow

Cat

Carnivorous

Domesticated

Herbivorous

Meat

Omnivorous

Pet

Plant

Wild

Asserted

To use the reasoner click Reasoner > Start reasoner ☒ Show Inferences

Underlying OWL Language

Class: Cow

Annotations:

rdfs:comment "eats only Plants",

rdfs:comment "Definable",

rdfs:comment "SelfStanding"

SubClassOf:

Mammal

OWL has many syntaxes;
this is one of them called **Manchester Syntax**

Recall the regimentation

- Base animal categories (noun-y terms)

- General:

- 1. Animal = eats some Stuff
 - 2. Mammal = has MammGlands

- Specific:

- Cat
 - Dog
 - 3. Cow = eats only Grass
 - 4. Human = Omnivore

- Ways an animal can be (adjective-y terms)

- General:

- Domesticated
 - Wild
 - Dangerous

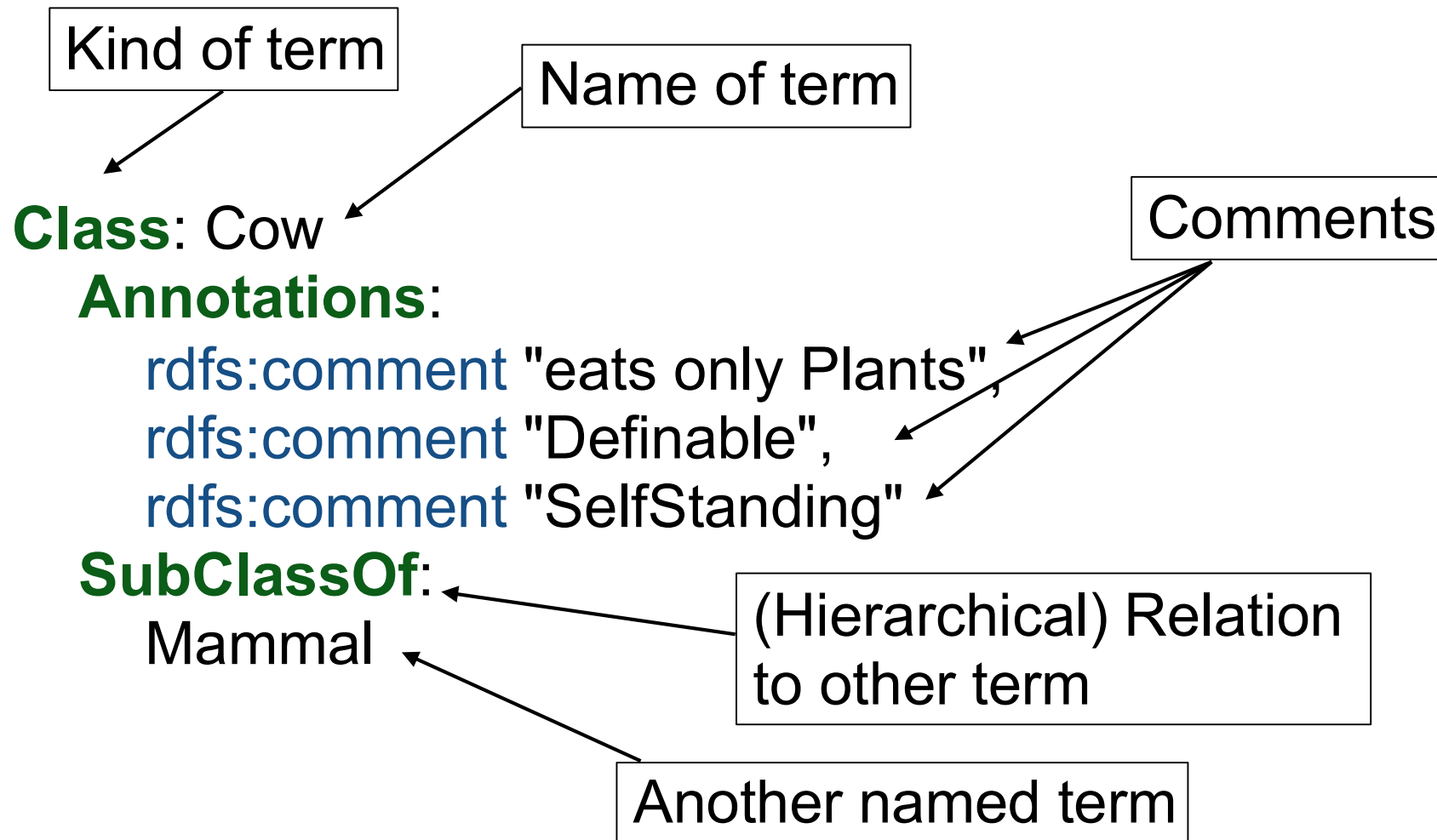
- 5. Carnivorous = eats only Meat
 - 6. Omnivorous = eats Meat & Plants
 - 7. Herbivorous = eats only Plants
 - 8. Delicious = tastes good

- Specific:

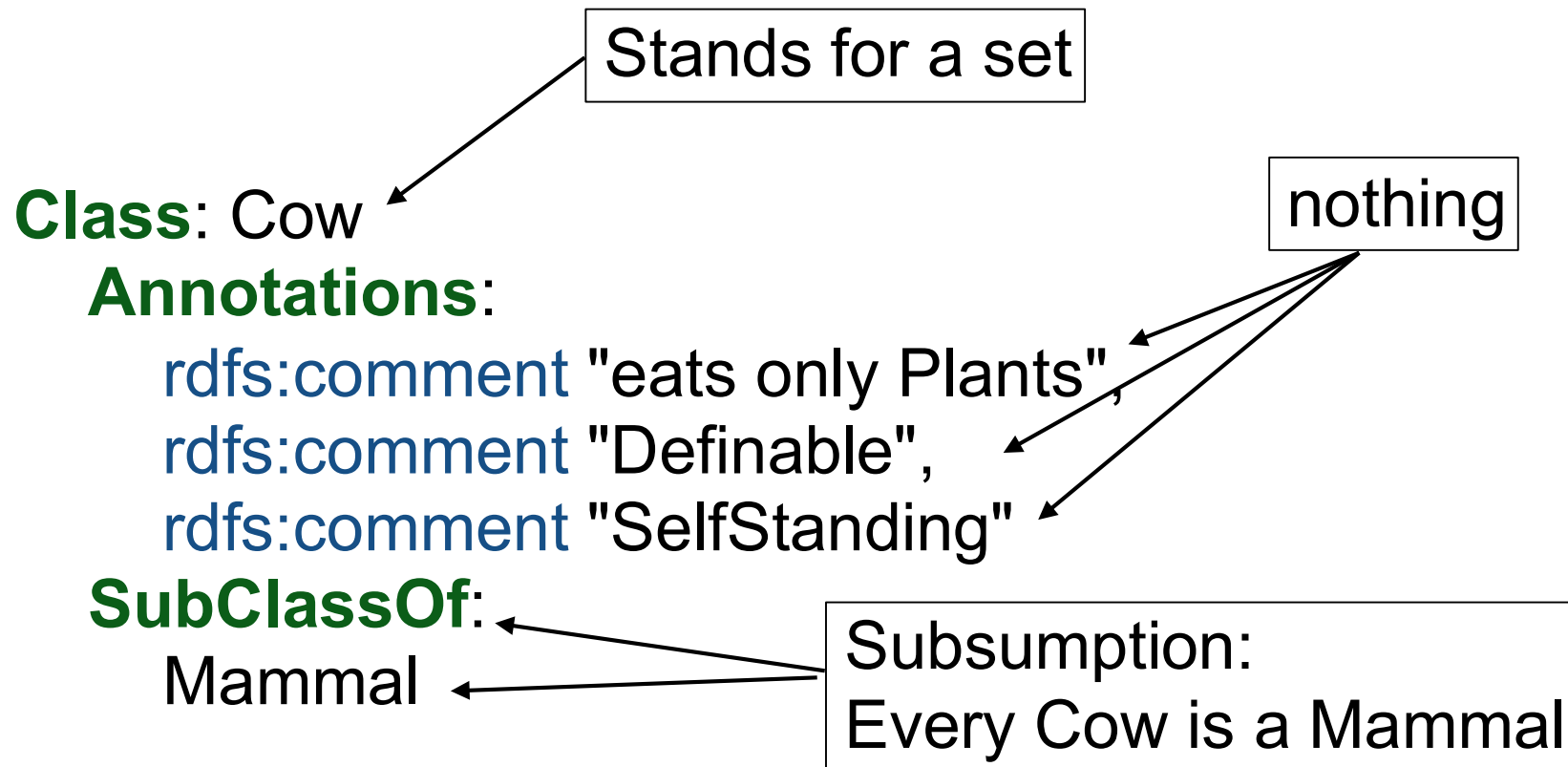
- 9. Pet = lives with Humans
 - 10. Farmed = is eaten/used

Which of these definitions
is really good?
I.e., is really a definition?

Our mini-formalisation



Meaning? Semantics?



Benefits of this formalisation?

Class: Cow

Annotations:

`rdfs:comment` "eats only Plants",

`rdfs:comment` "Definable",

`rdfs:comment` "SelfStanding"

SubClassOf:

Mammal

- Gives some structure to our set of terms:
 - a hierarchy that we can browse
 - we can retrieve classes
 - we can search for comments



Side note: A “Computer View”

Class: Blah

Annotations:

`rdfs:comment` “b123 623 7y3”,

`rdfs:comment` "mch345",

`rdfs:comment` "lkjherhjhhhh"

SubClassOf:

Foo

Better Annotations

Class: Cow

Annotations:

`rdfs:comment "eats only Plants",`

`isDefinable True`

`hasGrammaticalType SelfStanding`

Use good
annotation
properties



SubClassOf:

Mammal

For less string-hackery and
easier data-entry

A Better Definition

Class: Cow

Annotations:

isDefinable True

hasGrammaticalType SelfStanding

EquivalentTo:

eats **only** Plant

SubClassOf:

Mammal

...exact meaning/semantic later!

We Need a Syntax!

- A simple grammar for **descriptions** (aka class expressions)
- Examples
 - Animal that eats only Animal
 - eats some (not Animal)
 - not (eats only Animal and some Animal)

description ::= **conjunction** 'or' **conjunction** { 'or' **conjunction** }
 | **conjunction**

conjunction ::= **classIRI** 'that' ['not'] **restriction**
 { 'and' ['not'] **restriction** }
 | **primary** 'and' **primary** { 'and' **primary** }
 | **primary**

primary ::= ['not'] (**restriction** | **atomicClass**)

restriction ::= **Property** 'some' **primary**
 | **Property** 'only' **primary**

atomicClass ::= [A-Z][a-zA-Z]* (*in camel case*)

Property ::= [a-z][a-zA-Z]* (*in camel case*)

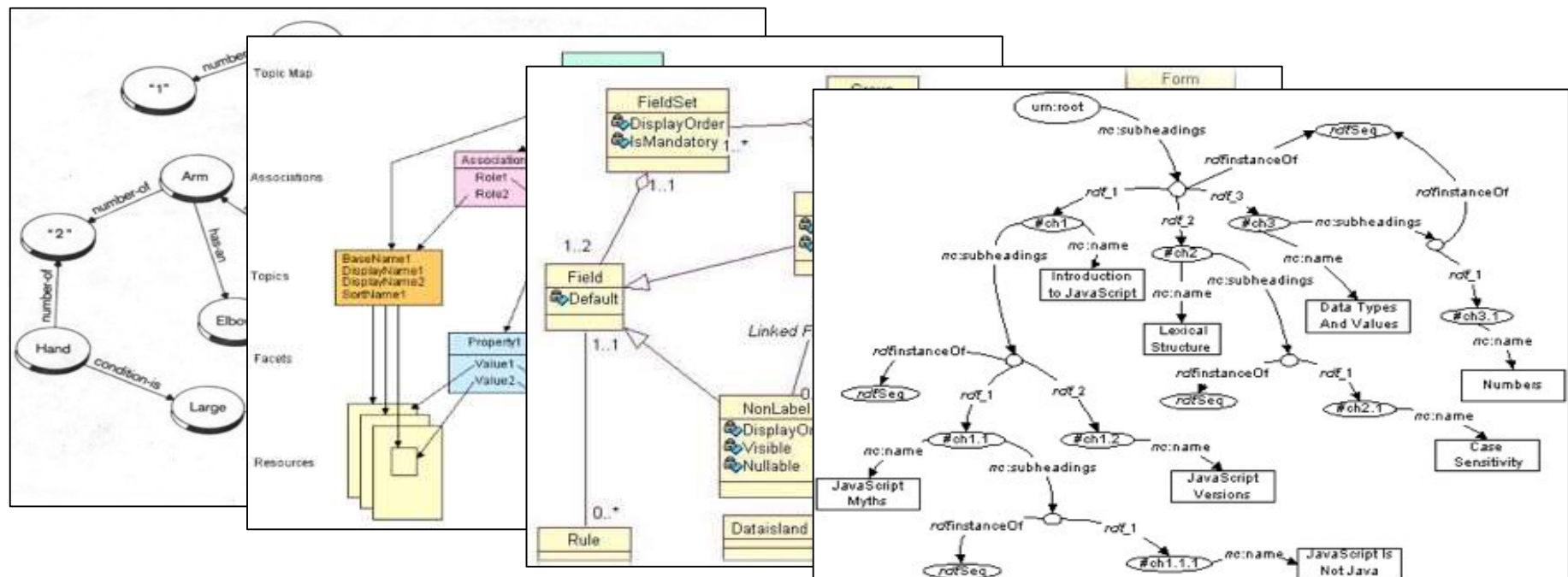
We Need More Syntax!

- A simple grammar for **axioms** (aka propositions, statements)
- Examples
 - Class: CarnivorousAnimal EquivalentTo: Animal that eats only Animal
 - Class: Cow SubClassOf: eats some (not Animal)
 - Class: ConfusedCow SubClassOf:
not (eats only Animal and some Animal)
- What does it all mean!?
- Coming in 10 minutes...

```
classFrame ::= 'Class:' atomicClass  
    { 'Annotations:'    annotation { ',' annotation }  
    | 'SubClassOf:'     description { ',' annotation }  
    | 'EquivalentTo:'   description { ',' annotation } }
```

Which Syntax?

- OWL is textual - would a graphical be better?
- In OWL, an ontology is a (web) **document** that we can
 - parse, import, syntax check and
 - draw graphs for!



Exploring Benefits

- E.g., Omnivorous
 - Annotations:
comment "Carnivorous
and
Herbivorous"
 - has no meaning
 - so let's be explicit:
 - add definition in class
description
 - run reasoner
 - check inferred class
hierarchy
- ➡our definition was wrong!

This screenshot shows a software interface with two main panels. The left panel, titled 'Class hierarchy (inferred)', displays a tree structure of classes: owl:Thing, Animal, Mammal, Cat, Cow, Dog, Human, Carnivorous, Domesticated, Herbivorous, Meat, Omnivorous (highlighted in blue), and Pet. The right panel has tabs for 'Annotations' and 'Usage'. The 'Annotations' tab is active, showing an annotation for 'Omnivorous' with the comment 'Both Carnivorous and Herbivorous'. Below this, a 'Description: Omnivorous' section is visible.

This screenshot shows the same interface after running a reasoner. The left panel's class hierarchy now includes 'Omnivorous' as a subclass of 'Herbivorous'. The right panel's 'Annotations' tab remains the same, but the 'Description: Omnivorous' section now shows an 'Equivalent To' relationship with 'Carnivorous and Herbivorous', indicating that the reasoner has inferred that the original comment was true.

Exploring Benefits II

- E.g., Cows
 - Annotations:
comment “Animal that eats only Plants”
 - has no meaning
 - so let’s be explicit:
 - add definition in class description
 - run reasoner
 - check inferred class hierarchy
- ➔ our class hierarchy is improved: Cows are indeed herbivores!

The image displays two screenshots of a class hierarchy editor, likely from a software tool like Protégé.

Top Screenshot: The class hierarchy on the left shows a tree structure starting with `owl:Thing`. Under `owl:Thing` is `Animal`, which has subclasses `Carnivorous`, `Domesticated`, and `Herbivorous`. `Herbivorous` has subclasses `Cow`, `Meat`, `Omnivorous`, `Pet`, and `Plant`. The `Herbivorous` class is selected and highlighted in blue. On the right, the **Description: Herbivorous** panel shows the class description as `Equivalent To` `+ eats only Plant`, where `eats` is in pink and `only` is in purple.

Bottom Screenshot: The class hierarchy on the left is similar, but `Cow` is now selected and highlighted in blue. The `Herbivorous` class is no longer highlighted. On the right, the **Description: Cow** panel shows the class description as `Equivalent To` `+ Mammal and eats only Plant`, where `Mammal` is in blue, `and` is in teal, `eats` is in pink, and `only` is in purple. Below this, the `Herbivorous` class is listed as a subclass.

First Benefits!

- Links for “free”
 - Tools make **implicit** links **explicit**
 - We don’t have to encode every link ourselves
 - Different modality
 - Instead of is-a/subsumption relations...focus on **meanings**
 - ...we can think local rather than global
- Verification
 - Definitions have **consequences**
 - Wrong links
 - Detectable problems
 - Links so wrong they are never right