Foundations of the Semantic Web: Ontology Engineering

Lecture 3

Common problems
Ontology Patterns
Re-representing properties and classes
Parts and Wholes

Alan Rector & colleagues

Special acknowledgement to Jeremy Rogers & Chris Wroe

1

Part 1: "Elephant Traps"

- 'Some' does not imply only
  'Only' does not imply some'
- Trivial satisfaction of universal restrictions
- Domain and Range Constraints
- What to do when it all turns red

2

someValuesFrom means "some"

- "some" means "at least 1"
- Dog_owner *equivalentClass*
  Person *and* hasPet *some* Dog
  - means:
    A Pet_owner is any person who has as a pet some (i.e. at least 1) dog
- Dog_owner *subclassOf*
  Person *and* hasPet *some* Dog
  - means
    All Pet_owners are people and have as a pet some (i.e. at least 1) dog.

3

allValuesFrom means "only"

- "only" means "no values except"
- First_class_lounge *equivalentClass*
  Lounge *and* hasOccupants *only* FirstClassPassengers
  - Means
    "A 'first class lounge' is any lounge where the occupants are *only* first class passengers"
    or
    "A first 'class lounge' is any lounge where there are *no* occupants *except* first class passengers"
- First_class_lounge *subclassOf*
  Lounge *and* hasOccupants *only* FirstClassPassengers
  - Means
    "All first class lounges have *only* occupants who are first class passengers"
    "All first class lounges have *no* occupants *except* first class passengers"
    "All first class lounges have *no* occupants who are *not* first class passengers"

4

"Some" does not mean "only"

- A "dog owner" might also own cats, and rats, and guinea pigs, and…
  - It is an open world, if we want a closed world we must add a closure restriction or axiom
- Dog_only_owner *equivalentClass*
  Person *and* hasPet *some* Dog and
  hasPet *only* Dog
- A "closure restriction" or "closure axiom"
  - The problem in making maguerita pizza a vegie pizza
  - Closure axioms use 'or' (disjunction)
  - dog_and_cat_only_owner *equivalentClass*
    hasPet *some* Dog and
    hasPet *some* Cat and
    hasPet *only* (Dog or Cat)

5

"Only" does not mean "some"

- There might be *nobody* in the first class lounge
  - That would still satisfy the definition
  - It would not violate the rules
- A pizza with *no* toppings satisfies the definition of a vegetarian pizza
  - Pizza & has_topping_ingredient *only* Vegetarian_topping
    - It has no toppings which are meat
      - It has no toppings which are not vegetables
        - » It has no toppings which aren't fish…
  - Analogous to the empty set is a subset of all sets
    - One reason for a surprising subsumption is that you have made it impossible for there to be any toppings
      - *"only* (ham cheese)"

6

Trivial Satisfiability

- A universal ('only') restriction with an unsatisfiable filler is "trivially satisfiable"
  - i.e. it can be satisfied by the case where there is no filler
    - If there is an existential or min-cardinality restriction, inferred or explicit, then the class will be unsatisfiable
      - Can cause surprising 'late' bugs

7

Part 2: Domain & Range Constraints

- Actually axioms
  - *Property P range(RangeClass*)
    means
    - owl:Thing
      restriction(P *only* RangeClass)
  - *Property P domain(DomainClass)*
    means
    - owl:Thing
      restriction(inverse(P) *only* DomainClass)

8

Range Restrictions: What happens if violated?

- Actually axioms
  - *Property eats range(LivingThing*)
    means
    - owl:Thing
      restriction(P *only* LivingThing)
  - Bird eats some Rock
    - All StoneEater eats some rocks
      - What does this imply about rocks?
        - » *Some* rocks are living things
        - » because only living things can be eaten
        - » What does this say about "all rocks"?

9

Domain restrictions: What happens if violated?

- Actually axioms
  - *Property eats domain(LivingThing)*
    means
    - owl:Thing
      restriction(inverse(eats) *only* LivingThing)
    - "Only living things eat anything"
  - StoneEater eats some Stone
    - All StoneEaters eat some Stone
      - Therefore All StoneEaters are living things
        - » If StoneEaters are not already classified as living things, the classifier will reclassify ('coerce') them
        - » If StoneEaters is disjoint from LivingThing it will be found disjoint

10

Example of Coercion by Domain violation

- has_topping: *domain*(Pizza) *range*(Pizza_topping)

  class Ice_cream_cone
  has_topping *some* Ice_cream

- If Ice_cream_cone and Pizza are *not* disjoint:
  - Ice_cream_cone is classified as a kind of Pizza
    …but: Ice_cream is *not* classified as a kind of Pizza_topping
    - Have shown that:
      all Ice_cream_cones are a kinds of Pizzas,
      but only that:
      some Ice_cream is a kind of Pizza_topping
      - » *Only domain constraints can cause reclassification*

11

Reminder Subsumption means necessary implication

- *"B is a kind of A"*
  means
  "All Bs are As"

  - *"Ice_cream_cone is a kind of Pizza"*
    means
    "All ice_cream_cones are pizzas"
    - From "Some Bs are As" we can deduce very little of interest in DL terms
      - » "some ice_creams are pizza_toppings"
        says nothing about "all ice creams"

12

Summary:Domain & Range Constraints Non-Obvious Consequences

- Range constraint violations – *unsatisfiable or ignored*
  - If filler and RangeClass are disjoint: *unsatisfiable*
  - *Otherwise nothing happens!*
- Domain constraint violations – *unsatisfiable or coerced*
  - If subject and DomainClass are disjoint: *unsatisfiable*
  - *Otherwise, subject reclassified (coerced) to kind of DomainClass!*
- *Furthermore cannot be fully checked before classification*
  - *although tools can issue warnings.*

13

Part 3: What to do when "Its all turned red"

Don't Panic!

- Unsatisfiability propagates – so trace it to its source
  - Any class with an unsatisfiable filler in a *someValuesFor* (existential) restriction is unsatisfiable
  - Any *subclass* of an unsatisfiable class is unsatisfiable
  - Therefore errors propagate, trace them back to their source
- Only a few possible sources
  - Violation of disjoint axioms
  - Unsatisfiable expressions in some restrictions
    - Confusion of "and" and "or"
  - Violation of a universal (only) constraint (including range and domain constraints)
    - Unsatisfiable domain or range constraints
- Tools coming RSN

14

Part 4 – Patterns: n-ary relations

15

Saying something about a restriction

- Not just
  - that an animal is dangerous,
  - but why
  - And how dangerous
  - And how to avoid
- But can say nothing about properties
  - except special thing
    - Super and subproperties
    - Functional, transitive, symmetric

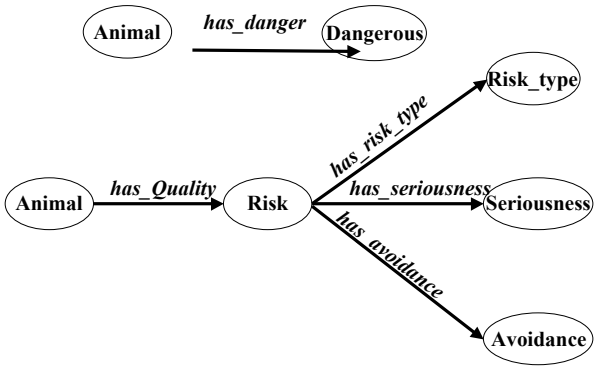
16

Re-representing properties as classes

- To say something about a property it must be re-represented as a class
  - property:has_danger → Class: Risk
    - plus property: Thing has_quality Risk
    - plus properties: Risk has_reason
      has_risk_type
      has_avoidance_measure
  - Sometimes called "reification"
    - But "reification" is used differently in different communities

17

Re-representing the property has_danger as the class Risk



18

Lions are dangerous

- All lions pose a deadly risk of physical attack that can be avoided by physical separation
- All lions have the quality risk that is
  - of type some physical attack
  - of seriousness some deadly
  - has avoidance means some physical separation

19

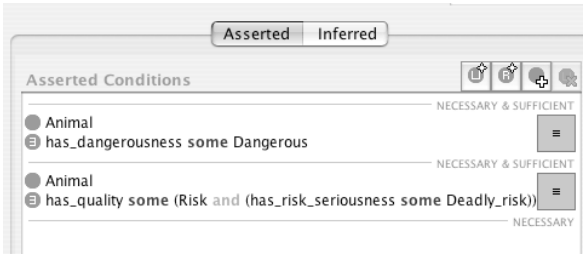
Can add a second definition of Dangerous Animal

- A dangerous animal is any animal that has the quality Risk that is Deadly
  - or
- Dangerous_animal =
  - Animal
    has_quality *some*
    (Risk AND has_seriousness *some* Deadly)
  - [NB: "that" paraphrases as "AND"]

20

In the tool

- Dangerous_animal =
  - Animal
    has_quality *some*
    (Risk AND has_seriousness *some* Deadly)



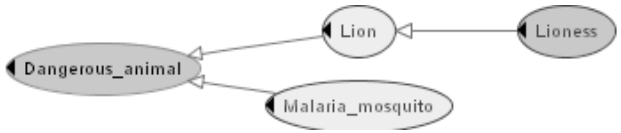
This says that

- Any animal that is Dangerous

  is also

  An animal that has the quality Risk
  with the seriousness Deadly

Anopheles Mosquitos now count as dangerous



- Because they have a deadly risk of carrying disease

Multiple definitions are dangerous

- Better to use one way or the other
  - Otherwise keeping the two ways consistent is difficult
  - … but ontologies often evolve so that simple Properties are re-represented as Qualities
    - Then throw away the simple property

Often have to re-analyse

- What do we mean by "Dangerous"
  - How serious the danger?
  - How probable the danger?
  - Whether from individuals (Lions) or the presence or many (Mosquitos)?
- Moves to serious questions of "ontology"
  - The information we really want to convey
    - Often a sign that we have gone to far
      - So we will stop

25

Part 5 – More Patterns: Part-whole relations

26

Part-whole relations

One method: NOT a SWBP draft

- How to represent part-whole relations in OWL is a commonly asked question
- SWBP will put out a draft.
- This is one approach that will be proposed
  - It has been used in teaching
  - It has no official standing

27

Part Whole relations

- OWL has no special constructs
  - But provides the building blocks
- Transitive relations
  - Finger is_part_of Hand
    Hand is_part_of Arm
    Arm is_part_of Body
    → therefore →
  - Finger is_part_of Body

28

Implementation Pattern
Transitive properties with non-transitive "direct" subproperties

- Transitive properties should have non-transitive children
  - isPartOf : transitive
    isPartOfDirectly : non-transitive
- Split which is used in "partial" descriptions (subclass axioms) and complete definitions (equivalentClasses axioms)
  - "Partial Definitions" (subclass axioms) use non-transitive version
  - "Complete Definitions" equivalentClasses axioms use transitive version
- Benefits
  - Allows more restrictions in domain/range constraints and cardinality
    - Allows the hierarchy along that axis to be traced one step at a time
    - Allow a good approximation of pure trees
      - Make the nontransitive subproperty functional
        - » Transitive properties can (almost) never be functional (by definition, a transitive property has more than one value in any non-trivial system)
    - Constraints on transitive properties easily lead to unsatisfiability

29

Many kinds of part-whole relations

- Physical parts
  - hand-arm
- Geographic regions
  - Hiroshima - Japan
- Functional parts
  - cpu – computer
- See Winston & Odell
  Artale
  Rosse

30

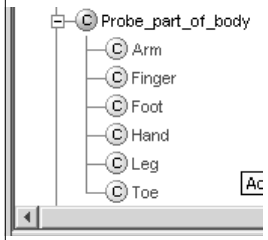
Simple version

- One property *is_part_of*
  - transitive
    - Finger is_part_of *some* Hand
      Hand is_part_of *some* Arm
      Arm is_part_of some Body

31

Get a simple list

- Probe_part_of_body =
  Domain_category
  is_part_of some Body



- Logically correct
  - But may not be what we want to see

32

Injuries, Faults, Diseases, Etc.

- A hand is not a *kind of* a body
  - … but an injury to a hand is a kind of injury to a body
- A motor is not a *kind of* automobile
  - … but a fault in the motor is a kind of fault in the automobile
- And people often expect to see partonomy hierarchies

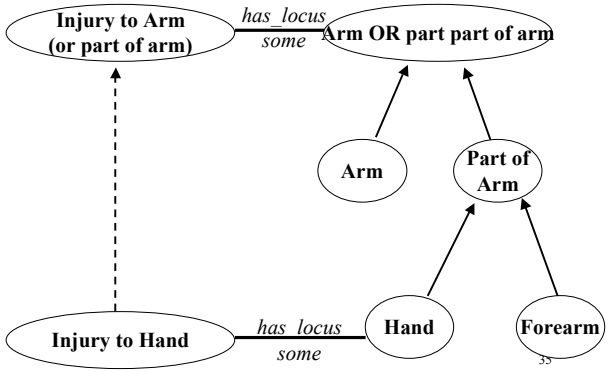
33

Being more precise: "Adapted SEP Triples"

- Body ('as a whole')
  - Body
- The Body's parts
  - is_part_of *some* Body
- The Body and it's parts
  - Body OR is_part_of *some* Body
- Repeat for all parts
  - Use 'Clone class' or
  - NB: 'JOT' Python plugin is good for this

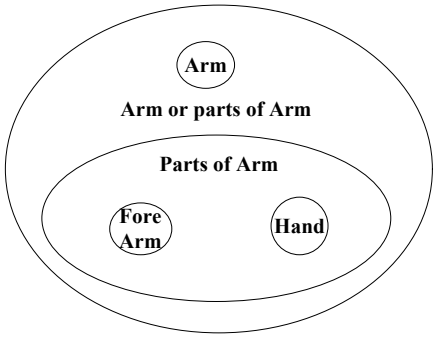
34

Adapted SEP triples: UML like view



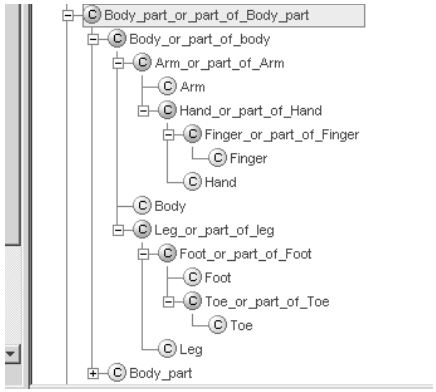
35

Adapted SEP triples: Venn style view



36

Resulting classification: Ugly to look at, but correct



Using part-whole relations: Defining injuries or faults

- Injury_to_Hand = Injury has_locus *some* Hand_or_part_of_hand
- Injury_to_Arm = Injury has_locus *some* Arm_or_part_of_Arm
- Injury_to_Body = Injury has_locus *some* Body_or_part_of_Body

- The expected hierarchy from point of view of anatomy

Parts & Wholes in More Detail

Parts & Wholes, containment, connection and adjacency – common sense merology

- Standard lexical semantic versions motivated by history
  Many philosophical versions motivated by topology
  - This version motivated primarily by anatomy and engineering
- Classic knowledge representation work is
  - Odell, J. J. (1994). "Six different kinds of composition." Journal of Object Oriented Programming **5**(8): 10-15.
    - A short readable summary
      - Not complete nor completely up to date
  - Winston, M., R. Chaffin, et al. (1987). "A taxonomy of part-whole relations." Cognitive Science **11**: 417-444.
- Merology – the study of parts and wholes
  - A quick glance at Google…

Parts & wholes: Some examples

- The leg is part of the chair
- The left side of the body is part of the body
- The liver cells are part of the liver
- The ignition of part of the electrical system of the car
- The goose is part of the flock
- Manchester is part of England
- Computer science is part of the University

41

Five families of relations

- Partonomic
  - Parts and wholes
    - The lid is part of the box
  - Constitution
    - The box is made of cardboard
  - Membership?
    - The box is part of the shipment
- Nonpartonomic
  - Containment
    - The gift is contained in the box
  - Connection/branching/Adjacency
    - The box is connected to the container by a strap

42

Some tests

- True kinds of *part-of* are transitive and
  A fault to the part is a fault in the whole
  - The finger nail is part of the finger is part of the hand is part of the upper extremity is part of the body
    - Injury to the fingernail is injury to the body
  - The tail-light is part of the electrical system is part of the car
    - A fault in the tail light is a fault in the car
- *Membership* is not transitive
  - The foot of the goose is part of the goose but not part of the flock of geese
    - Damage to the foot of the goose is not damage ot the flock of geese
- *Containment* is transitive but things contained are not necessarily parts
  - A fault (e.g. souring) to the milk *contained in* the bottle is not damage to the bottle
- Some kinds of part-whole relation are questionably transitive
  - Is the cell that is part of the finger a part of the body?
    - Is damage to the cell that is part of the finger damage to the body?
      - Not necessarily, since the cells in my body die and regrow constantly

43

Structural parts

- The leg is *a component of* of the table
  - Discrete
  - connected,
  - clear boundary,
  - specifically named
  - may be differently constituted
  - Can have metal legs on a wooden table or vice versa
- The left side is *a subdivision of* the table
  - 'Side', 'Lobe', 'segment', 'region',…
    - Arbitrary, similarly constituted,
    - components typically fall into one or another subdivision;
    - defined in relation to something else;
    - sensible to talk about what fraction it is: half the table, a third of the table, etc.

44

Propagates_via / transitive_across

- Components of subdivisions are components of the whole, *but*
  subdivisions of components are not subdivisions of the whole
  - A the left side of the steering wheel of the car is not a subdivision of the left side of the car (at least not in the UK)
- No consistent name for this relation between properties
  - We shall call it *propagates_via* or *transitive_across*
    - Also known as "*right identities*"
  - Not supported in most DLs or OWL directly
    - Although an extension to FaCT to support it exists
    - Heavily used in medical ontologies (GRAIL and SNOMED-CT)

45

No simple solution: Here's one of several nasty kluges

- Component_of_table is defined as a component of table or any subdivision of table
  - Must do it for each concept
    - A Schema rather than an axiom
      - No way to say "same as"
      - No variables in OWL
        - » or most DLs
- SCHEMA:
  Components_of_X ≡
  isComponentOf some
  (X or (some isSubDivisionOf X))
  - Tedious to do
    - Schemas to be built into new tools

46

- x p1 y and y p2 z THEN x p1 z
- x hasLocus Hand and Hand isPartOf Arm-->

X has_locus ARM.

hasLocus o isPartOf --> hasLocus

Is_component_of o is_subdivision_of --> is_component_of

'o' == "composed with"

All xyzß. p1(x,y) & p2(y,z) --> p1(x,z)

p1 o p2 --> p1

47

Functional parts

- Structural parts form a contiguous whole
  - May or may not contribute to function
    e.g. decorative parts, vestiges such as the human appendix, "spandrels"[1], accidental lumps and bumps
- The remote control is part of the projection system
  - May or may not be physically connected to it
    - Part of a common function
- Biology examples:
  - The endocrine system
    - The glands are not connected, but form part of a functioning system communicating via hormones and transmitters
    - The blood-forming system
      - Bone marrow in various places, the spleen, etc.

[1] See Stephen J Gould

48

If something is both a structural and functional part…

- Must put in both restrictions explicitly
  - Can create a common child property but this gets complicated with the different kinds of structural parts
  - Better to put syntactic sugar in tools
    - But syntactic sugar has not arrived, so for this course you have to do it by hand!
      - Coming Real Soon Now (RSN)

49

So far we have

- isPartOf
  - isStructuralPartOf
    - isSubdivisionOf
    - isComponentOf
  - isFunctionalPartOf
- Many other varieties
  - Layers, surfaces, …
- Many other constraints, e.g.
  - Dimensions must match
    - 3-D things can only be structural parts of 3-D things
  - boundaries have one less dimension than the things they bound
    - surfaces bound volumes, lines bound areas
  - layers of subdivisions are subdivisions of layers of the whole
    - the skin of the finger is a subdivision of the skin of the upper hand
- Can add isSubprocessOf –
  - similar to isComponentOf

50

What about containment

- X **is_contained_in** Y **is_structural_partOf** Z → X **is_contained_in** Z
- Rigorous version needs analogous schema to subdivision
  - **Contained_in**_X ≡ **Contained_in** *someValuesFor* (X *or* (*someValuesFor* **is_structural_part_of** X))
- Weak approximation
  - make **contained_in** a super-property of **is_structural_part**
    - Not right – implies all structural parts are contained in the whole
      - » A "kluge"

51

Class, Instances & Properties

- Naming Conventions - useful hints
  - Classes
    - Nouns, singular, begin with upper case
    - Begin with an upper case letter. Subsequent words in lower case, e.g. "Person", "Professor_of_computer_science", etc.
  - Individuals
    - Nouns, singular, begin with lower case and usually have a designator or article, e.g.
      - "person_1", "aPerson", "anApple"," apple_"1, etc.
    - Names - all words in upper case, e.g. "Manchester_University"
  - Properties
    - Verbs, all lower case
    - Where convenient of the form "has_X" with inverse "is_X_of", e.g. "has_part" / "is_part_of" ; "has_module" / "is_module_of"

52

"I am *an* individual"

- My naming convention should be
  - Hence **Alan_Rector** in the examples or **"aPerson"**
- So are
  - This year's version of CS646 –
    - Hence CS646_2003 in the examples
  - You
  - The University of Manchester
  - This room, its furniture, etc.
  - Your thoughts, understanding, …
  - This lecture, the lab following it, …
  - …

53

"Person" is a class

- My naming convention should be
  - **"Person"**
- So are
  - **CS646** (the class of all CS646 modules in all years)
  - **Professor, Student, …**
  - **University**
  - **Room, Desk, Internal_room, Difficult_module, …**
  - **Idea, Understanding, Thought_about_description_logic**
  - **Lecture, Lab_associated_with_lecture, …**
  - **…**

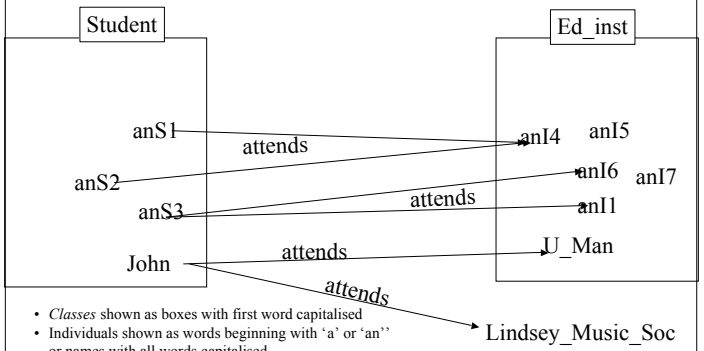
54

Lecturing on a course is a property

- The naming convention should be
  - lectures_on / is_lecturer_for_module
- So are
  - has_annual_version/is_annual_version_of
  - attends / is_attended_by
  - has_thought / is_thought_of
  - has_lecture / is_lecture_of

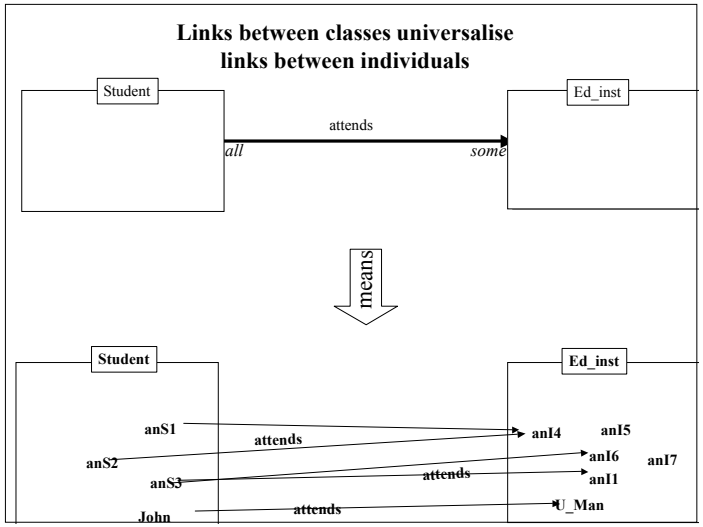
55

Individuals, Properties & Classes



- *Classes* shown as boxes with first word capitalised
- Individuals shown as words beginning with 'a' or 'an'' or names with all words capitalised
- Properties shown as labels for lines lines

56



Individuals in Ontologies

- **Simple test 1:**
  "Can it have kinds" – if so, it is a class
  - "Kinds of dog" makes sense
  - "Kinds of person" makes sense
  - "Kinds of Alan Rector" does not make sense
  - "Kinds of Module" makes sense
  - "Kinds of CS646_2003" does not make sense
  - "Kinds of jacket" makes sense
  - "Kinds of the 'jacket I am wearing'" does not make sense

Individuals in Ontologies (cont)

- Simple test 2:
  If you say something about it,
  if you have made a new concept, then it is a class
  if you have just stated a fact about it, it is is an individual.
  - "Big dog" is a new class of dog
    - "Rover is big" just says something about Rover
      - Which would allow us to infer that Rover is a member of the class of "Big Dogs"
  - "Men with beards" is a new class
    - "Alan rector has a beard" is a fact about Alan Rector
      - Which would allow us to infer that he is a member of the class of "Men with Beards"

Clues in English

- Articles + singular indicate individual
  - 'the book there on the shelf' – an individual
  - 'a book' – an unspecified individual
- Proper nouns (almost always) indicate individuals
  - Alan Rector, Ian Horrocks, Cross Street, Manchester, England, …
- Plurals usually indicate classes
  - 'the books' – probably a class
    - Although possibly an individual aggregation
  - And perversely the English convention is to name classes in the singular

More clues in English

- a '…that…' clause and usually indicates a class
  - "The Modules that are available for ACS"
    - Perversely by convention Classes are given names in the singular in English
      - "Module that…"
- a '…which…" clause depends on local usage
  - Some English stylebooks would have 'which' clauses used only for individuals, others say there is no real difference between 'that' and 'which'
    - "MS Word usually asks for 'that' with plurals (classes) and 'which' with singulars
- No perfect guide, must take case by case.

61

Leaf nodes are not Individuals

- Leaf node
  - Depends on ontology – may be very detailed, e.g.
    - Golden_retriever_bitch_from_karmella_kennels_from_2003_litter
      - Individual in that class "Halo"
- Even if there is only one possible individual, a leaf node is not an individual
  - Transferable_skills_course_for_first_year_PhD_students_in_CS_department
    - There might be other courses besides CS700
      - Its not impossible, just untrue
- Only individuals if there *could never be* kinds
  - CS646_2003
    - There can never be a "kind" of this year's course

62

Keeping the Ontology Re-usable

- If we make leaf nodes individuals, we close off any extension to more granular kinds
  - Make the ontology specific to our immediate needs
  - Make extensions require radical surgery

63

Comparison with "Instances" in databases, frames, and OO programming

- "Individuals" in ontologies are slightly different than in OO programming or data bases
- Test for individual
  - Ontologies – could it sensibly have kinds
  - Databases – is it going to be stored in a field in the database
  - OO programming – is it going to be an operational object in the program
  - RDF(S) – still some confusion
    - Anything can be an individual

64

"Tangle at the Top"

- Many OO environments require that everything be an instance of something.
  - If everything must be an instance of something, then we have an infinite regress
    - Most systems stop it by having something be an instance of itself
      - Protégé, Smalltalk, and Java Class
      - RDF(S), OWL-Full: rdf:resource
- Being an instance of yourself violates the semantics of OWL-DL
  - In OWL-DL, classes are not instances of anything
    - They are interpreted as the intensions of sets of individuals\
      - (In OWL-Full Classes may be instances; also in RDF(S))
  - Without very careful limitations, will make any logic inconsistent
    - It took ≥ 30 years from Russell's letter to Cantor to Zermelo-Frankel's consistent axiomatization of set theory, and another 10 years to von Neuman-Bernays-Goedel's axiomatization, which is usually used today.

More vocabulary "Intensions" & "Extensions"

- "Intension"
  - The meaning of something
    The definition of a class
    - "The lecturer the application part of this module"
    - "The evening star"
- "Extension"
  - The things which satisfy the meaning –
    the members of the class
    - Alan Rector
    - The planet Venus

Extensional equality vs Intensional Equivalence

- Two sets are equal if their extensions are equal
  - In a particular model
    - The extensions of "The evening star" and "The morning star" are equal
- Two intensions are equivalent if if their extensions *must* be equal –
  - i.e. if their being unequal would be a contradiction in *any* model satisfying the same axioms
    - "Three sided polygon" is equivalent to "Three angled polygon" given the axioms of geometry

'T-Box' and 'A-Box'

- 'T-Box' (Terminology Box)
  - Definitions and restrictions on classes
- 'A-Box'(Assertions box)
  - Descriptions and assertions of individuals
- DLs (& OWL DL) work best for T-Box
  - Large general A-Boxes are intractable
    - A change anywhere can propagate anywhere else
- Individuals in defining classes, e.g. "Lecturers on CS646" or "John's shirts"
  - Often best implement as 'pseudo-individuals'

Nominals - oneOf

- Abstract syntax:
  Individuals should be able to be imported into class restrictions via *oneOf*
  - Staff_for_CS646_2003 ≡
    restriction teaches some *oneOf* {CS_646_2003}
- Manchester syntax assumes nominals:
  - Staff_for_CS646_2003 ≡
    restriction teaches *value* CS_646_200

69

Pseudo-Individuals to simulate Nominals Simulating Individuals as Leaf Nodes

- For use in nominals, it often works better in current technology to simulate individuals as leaf nodes
  (A Very Nasty Kluge - bug often the best engineering with the tools available.)
  - Follow the naming convention, and use a suffix such as "_ind" or "_inst"
  - Mark them in the comment field. Perhaps create a special annotation property.
    - pseudo-individual:true
    - Or make them all a sub of 'Pseudo individual'
- …but beware: You it is incomplete - not all inferences will be made
  - Particularly concerning inverses

70

Individuals in Protégé

- On the Individuals Tab
  - A form is automatically generated for with a field for every property for which the class is explicitly in the domain.
- NB we will do very little with individuals in this course

71