UNIVERSITY OF Southampton
School of Electronics and Computer Science

Knowledge Representation, Ontologies and Logic

Dr Nicholas Gibbins
32/4037
nmg@ecs.soton.ac.uk

Knowledge Representation

- Knowledge representation is a central concern of the Semantic Web
- Knowledge must be organised for later use
- A good knowledge representation ‘naturally’ represents the problem domain
- An unintelligible knowledge representation is wrong
- Most AI systems (and therefore SW systems) consist of
  - Knowledge Base
  - Inference Mechanism (Inference Engine)

Knowledge Representation

- Knowledge Base (KB)
  - Forms the system's intelligence source
  - Inference mechanism uses contents of KB to reason and draw conclusions
- Inference mechanism
  - Set of procedures that are used to examine the knowledge base to answer questions, solve problems or make decisions within the domain

Knowledge Representation

- Major knowledge representation schemes:
  - Logic
  - Production rules
  - Semantic Networks
  - Frames
- Semantic Web combines aspects of all of these schemes
  - We will concentrate on the logical aspects

UNIVERSITY OF Southampton

School of Electronics and Computer Science

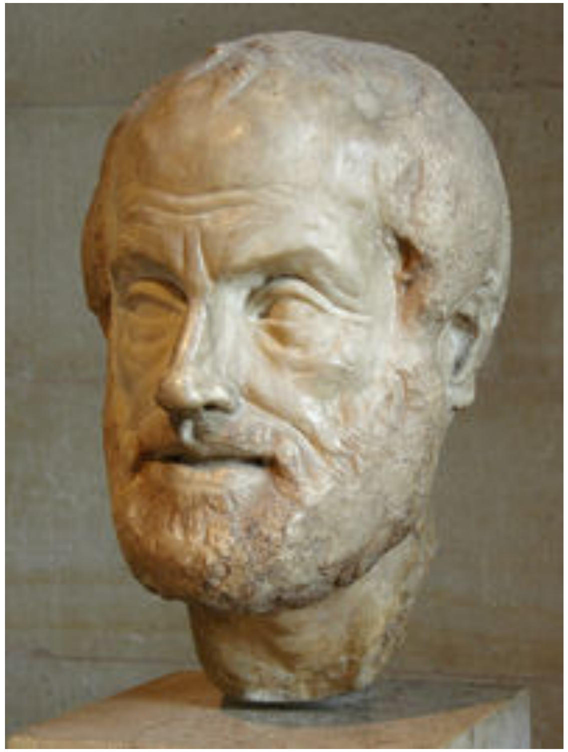
Ontologies

Defining the ‘O’ word

UNIVERSITY OF Southampton
School of Electronics
and Computer Science

Ontology, *n.*

1. a. *Philos*. The science or study of being; that branch of metaphysics concerned with the nature or essence of being or existence.



Oxford English Dictionary, 2004

The Celestial Empire of Benevolent Knowledge

UNIVERSITY OF Southampton
School of Electronics and Computer Science

On those remote pages it is written that animals are divided into:

a. those that belong to the Emperor
b. embalmed ones
c. those that are trained
d. suckling pigs
e. mermaids
f. fabulous ones
g. stray dogs
h. those that are included in this classification
i. those that tremble as if they were mad
j. innumerable ones
k. those drawn with a very fine camel's hair brush
l. others
m. those that have just broken a flower vase
n. those that resemble flies from a distance

J-L Borges, The Analytical Language of John Wilkins, 1952

Defining the ‘O’ word

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- An ontology is a **specification** of a **conceptualisation**
- **Specification**: A formal description
- **Conceptualisation**: The objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships that hold among them
- Referred to in the philosophical literature as **Formal Ontology**

T. R. Gruber. A translation approach to portable ontologies. Knowledge Acquisition, 5(2):199-220, 1993

Ontology in Computer Science

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- Ontologies as engineered artifacts:
  - constituted by a specific vocabulary used to describe a certain reality, plus
  - a set of explicit assumptions regarding the intended meaning of the vocabulary

- Shared understanding
- Facilitate communication
  - Establish a joint terminology for a community of interest
  - Normative models…
- Inter-operability: sharing and reuse

Ontology Structure

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- Ontologies typically have two distinct components:
- Names for important concepts in the domain
  - Elephant is a concept whose members are a kind of animal
  - Herbivore is a concept whose members are exactly those animals who eat only plants or parts of plants
  - Adult_Elephant is a concept whose members are exactly those elephants whose age is greater than 20 years
- Background knowledge/constraints on the domain
  - Adult_Elephants weigh at least 2,000 kg
  - All Elephants are either African_Elephants or Indian_Elephants
  - No individual can be both a Herbivore and a Carnivore

Informal Usage

UNIVERSITY OF Southampton
School of Electronics
and Computer Science

- Informally, ‘ontology’ may also be used to describe a number of other types of conceptual specification:
  - Controlled vocabulary
  - Taxonomy
  - Thesaurus

- Study of ontology is not limited to computer scientists and philosophers
- Rich tradition of knowledge representation and ontology in library and information science…
- …but they talk about classification and metadata instead of ontologies

Controlled Vocabularies

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- An explicitly enumerated list of terms, each with an unambiguous, non-redundant definition
- No structure exists between terms - a controlled vocabulary is a flat list
- Examples:
  - Library of Congress Subject Headings (LCSH)
  - Medical Subject Headings (MeSH)

Taxonomies

- A collection of controlled vocabulary terms organised into a hierarchical structure
- Each term is in one or more parent-child relationships
- May be several different types of parent-child relationship:
  - Type-instance
  - Genus-species
  - Part-whole (referred to as meronymy)

Taxonomy Examples

- Library classification schemes
  - Library of Congress
  - Dewey Decimal
  - UDC
- Linnean Classification
  - Kingdom, Phylum, Class, Order, Family, Genus, Species, Subspecies
- MeSH Tree Structures

Taxonomy Examples

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- Dewey Decimal
  - 500s - Natural Sciences and Mathematics
  - 530s - Physics
  - 537 - Electricity and Electronics
- Library of Congress
  - Q - Science
  - QA - Mathematics
  - QA71-90 - Instruments and machines
  - QA75-76.95 - Calculating machines
  - QA75.5-76.95 - Electronic computers and computer science
  - QA76-76.765 - Computer software

Polyhierachical Taxonomies

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- Also known as faceted taxonomies
- Define several orthogonal hierarchies
- Objects may be classified under multiple hierarchies
- Example: Universal Decimal Classification
  - Facets for language, relation to other subjects
  - 004.8 - artificial intelligence
  - 616 - clinical medicine
  - 004.8=20 - artificial intelligence in English
  - 004.8:616 - artificial intelligence and clinical medicine
  - 004.8:616=20 - AI and clinical medicine in English

Thesauri

- A thesaurus is a taxonomy with additional relations showing lateral connections
  - Related Term (RT)
  - See Also
- Parent-child relation usually described in terms of Broader Terms (BT) and Narrower Terms (NT)
- Thesauri also typically contain scope notes which define the meaning of a term

Thesaurus Example

Apples

Scope notes: The fruit of any member of the species *Malus pumila*

Broader term: Foodstuffs

Related terms: Cooking Ingredients
Taxable Foodstuffs
Horticulture

Narrower terms: Granny Smiths

See also: Apple Trees

Use: For Apple computers use Personal Computers (Apple)

Ontology

- An ontology further specialises types of relationships (particularly related term)
- A ontology typically includes:
  - Class definitions and hierarchy
  - Relation definitions and hierarchy
- An ontology may also include the following:
  - Constraints
  - Axioms
  - Rule-based knowledge

Summary

Controlled Vocabulary + Hierarchy = Taxonomy

Taxonomy + lateral relations = Thesaurus

Thesaurus + typed relations
+ constraints
+ rules
+ axioms = Ontology

UNIVERSITY OF Southampton

School of Electronics and Computer Science

A Description Logic Primer

Description Logics

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- Family of knowledge representation formalisms
- Decidable subset of first order predicate logic (FOPL)
  - Unary predicates denote class membership

$$Person(x)$$

  - Binary predicates denote relations (roles) between instances

$$hasChild(x, y)$$

- Model-theoretic formal semantics
- Underlying formalism for OWL

Defining ontologies with Description Logics

- Describe classes (concepts) in terms of their necessary and sufficient attributes

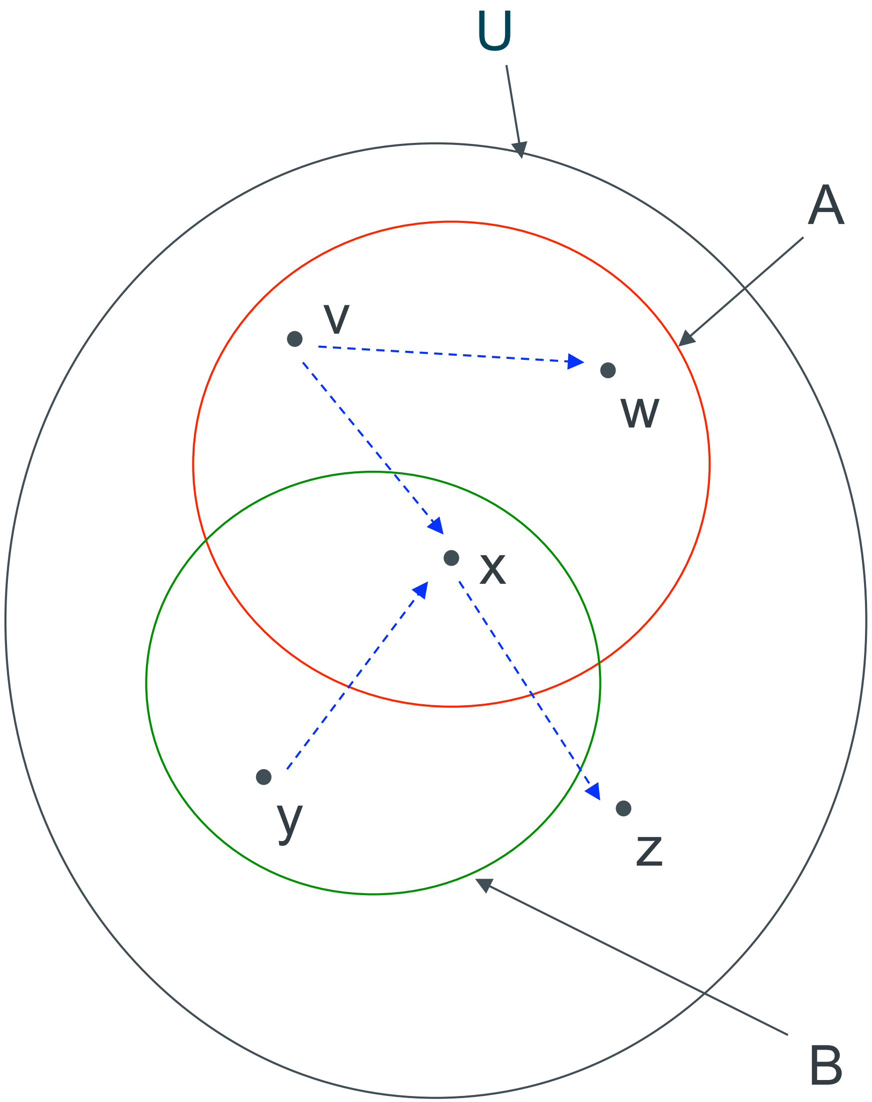
- Consider an attribute A of a class C:
- A is a necessary attribute of C
  - If an object is an instance of C, then it has A
- A is a sufficient attribute of C
  - If an object has A, then it is an instance of C

Description Logic Reasoning

- Designed for three reasoning tasks:
  - Satisfaction
    “Can this class have any instances?"
  - Subsumption
    "Is every instance of class A necessarily an instance of class B?"
  - Classification
    "What classes is this object an instance of?"

Classes as sets

UNIVERSITY OF Southampton
School of Electronics and Computer Science



Concept Constructors

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- Boolean class constructors $\neg C, C \sqcup D, C \sqcap D$
- Restrictions on role successors $\forall R.C, \exists R.C$
- Number restrictions (cardinality constraints) on roles $\leq n\ R, \geq n\ R, = n\ R$
- Nominals (singleton concepts) $\{x\}$
- Universal class, top $\top$
- Contradiction, bottom $\bot$

Role Constructors

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- Concrete domains (datatypes)
- Inverse roles R^-
- Transitive roles R^+
- Role composition $R \circ S$

OWL and Description Logics

- Not every description logic supports the constructors on the previous page
- More constructors = more expressive = higher complexity
- OWL DL is equivalent to the logic $\mathcal{SHOIN}(\mathbf{D})$
    - Atomic concepts and roles
    - Boolean operators
    - Universal, existential restrictions
    - Role hierarchies
    - Nominals
    - Inverse and transitive roles
    - Number restrictions
    - (but not role composition)

Boolean Class Operations

UNIVERSITY OF Southampton
School of Electronics and Computer Science

$\text{Child} \sqcap \text{Happy}$

- The class of things which are both children and happy

$\text{Rich} \sqcup \text{Famous}$

- The class of things which are rich or famous (or both)

$\neg\text{Happy}$

- The class of things which are not happy

Universal Restriction

$\forall \mathrm{hasPet.Cat}$

- The class of things all of whose pets are cats
  - Or, which only have pets that are cats
  - Note: includes those things which have no pets

Existential Restriction

UNIVERSITY OF Southampton
School of Electronics and Computer Science

$\exists \mathrm{hasPet.Cat}$

- The class of things which have some pet that is a cat
- Note: must have at least one pet

Cardinality Restrictions

University of Southampton
School of Electronics and Computer Science

- The class of things with more than one country of origin

$\geq 2 \text{ originCountry}$

- The class of quadrapeds

$= 4 \text{ hasLeg}$

Translating Description Logic to Predicate logic

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- Every concept C is translated into a predicate logic formula $\phi_C(x)$
- Boolean class constructors

$$\phi_{\neg C}(x) = \neg\phi_C(x)$$

$$\phi_{C \sqcap D}(x) = \phi_C(x) \wedge \phi_D(x)$$

$$\phi_{C \sqcup D}(x) = \phi_C(x) \vee \phi_D(x)$$

- Restrictions

$$\phi_{\exists R.C}(y) = \exists x.R(y,x) \wedge \phi_C(x)$$

$$\phi_{\forall R.C}(y) = \forall x.R(y,x) \Rightarrow \phi_C(x)$$

Knowledge Bases

UNIVERSITY OF Southampton
School of Electronics
and Computer Science

- A description logic knowledge base (KB) has two parts:
- TBox: terminology
  - A set of axioms describing the structure of the domain (i.e., a conceptual schema)
  - Concepts, roles
- ABox: assertions
  - A set of axioms describing a concrete situation (data)
  - Instances

TBox Axioms

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- Concept inclusion

  $C \sqsubseteq D$ (C is a subclass of D)

- Concept equivalence

  $C \equiv D$ (C is equivalent to D)

- Role inclusion

  $R \sqsubseteq S$ (R is a subproperty of S)

- Role equivalence

  $R \equiv S$ (R is equivalent to S)

- Role transitivity

  $R^+ \sqsubseteq R$ (R composed with itself is a subproperty of R)

Translating Description Logic to Predicate logic

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- Concept inclusion $C \sqsubseteq D$

  $\forall x.\phi_C(x) \Rightarrow \phi_D(x)$

- Concept equivalence $C \equiv D$

  $\forall x.\phi_C(x) \Leftrightarrow \phi_D(x)$

ABox Axioms

University of Southampton
School of Electronics and Computer Science

- Concept instantiation
  - x:D
  - x is of type D
- Role instantiation
  - <x,y>:R
  - x has R of y

Axiom Exercises

- Every person is either living or dead
- Every successful man has a beautiful wife
- No elephants can fly
- A curry is an Indian stew with a spicy ingredient
- All Englishmen are mad

$$\text{Person} \sqsubseteq \text{Living} \sqcup \text{Dead}$$

$$\text{Man} \sqcap \text{Successful} \sqsubseteq \exists \text{hasWife}.\text{Beautiful}$$

$$\text{Person} \sqcap \text{Male} \sqcap \text{Successful} \sqsubseteq$$
$$\exists \text{hasSpouse}.(\text{Beautiful} \sqcap \neg\text{Male} \sqcap \text{Beautiful})$$

$$\text{Elephant} \sqcap \text{FlyingThing} \equiv \bot$$

$$\text{Curry} \equiv \text{Stew} \sqcap$$
$$\exists \text{hasOrigin}.\{\text{India}\} \sqcap$$
$$\exists \text{hasIngredient}.\text{Spicy}$$

$$\exists \text{bornIn}.\{\text{England}\} \sqcap \text{Male} \sqsubseteq \text{Mad}$$

Tips for Creating Class Expressions

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- Don't Panic!
- No single 'correct' answer - different modelling choices possible
- Break sentence down into pieces
  - e.g. "successful man", "spicy ingredient" etc
- Look for indicators of axiom type:
  - "Every X is Y" - inclusion axiom
  - "X is Y" - equivalence axiom
- Remember that $\forall$R.C is satisfied by instances which have no value for R

Description Logic Semantics

UNIVERSITY OF Southampton
School of Electronics and Computer Science

- Δ is the domain (non-empty set of individuals)

- Interpretation function $\cdot^{\mathcal{I}}$ (or ext()) maps:
  - Concept expressions to their extensions (set of instances of that concept, subset of Δ)
  - Roles to subset of $\Delta \times \Delta$
  - Individuals to elements of Δ

- Examples:
  - $C^{\mathcal{I}}$ is the set of instances of C
  - $(C \sqcup D)^{\mathcal{I}}$ is the set of instances of either C or D

Description Logic Semantics

UNIVERSITY OF Southampton
School of Electronics and Computer Science

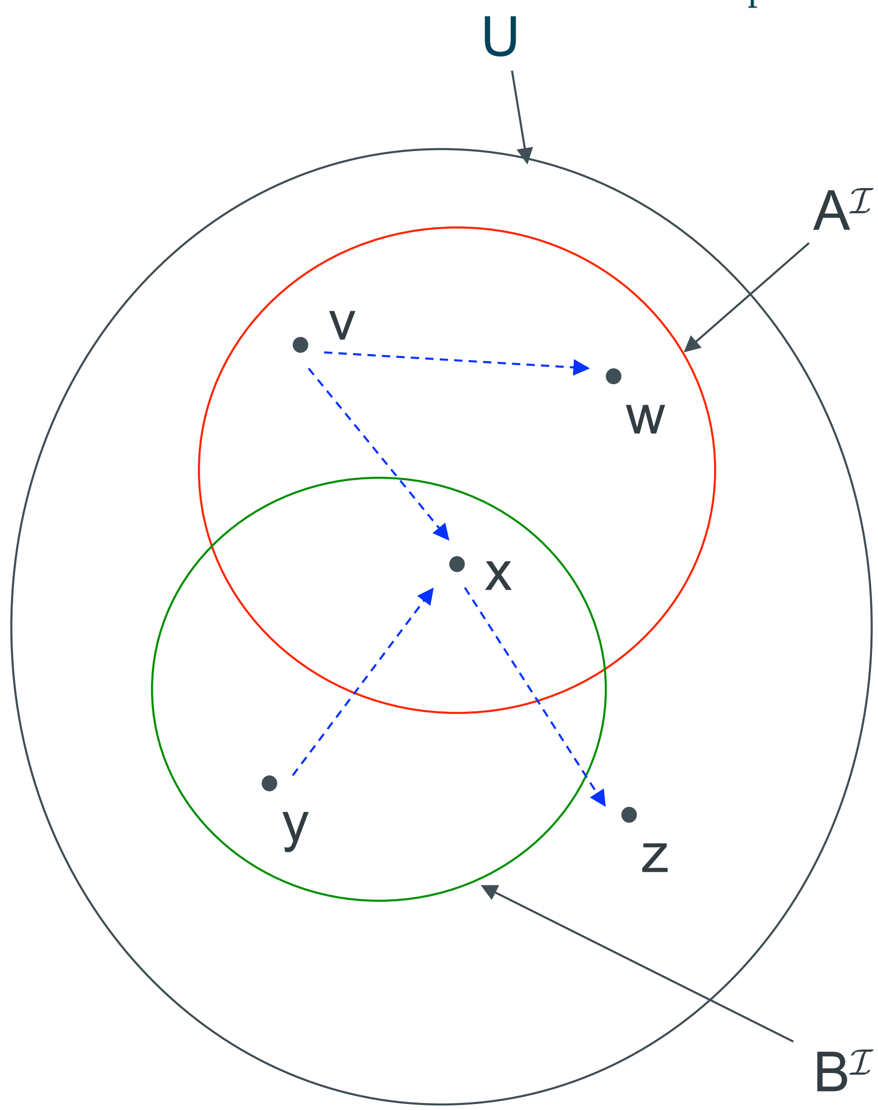
Syntax	Semantics	Notes
$(\mathrm{C} \sqcap D)^{\mathcal{I}}$	$\mathrm{C}^{\mathcal{I}} \cap D^{\mathcal{I}}$	Conjunction
$(\mathrm{C} \sqcup D)^{\mathcal{I}}$	$\mathrm{C}^{\mathcal{I}} \cup D^{\mathcal{I}}$	Disjunction
$(\neg C)^{\mathcal{I}}$	$\Delta \setminus C^{\mathcal{I}}$	Complement
$(\exists R.C)^{\mathcal{I}}$	$\{\mathrm{x} \mid \exists y.\langle x, y\rangle \in R^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\}$	Existential
$(\forall R.C)^{\mathcal{I}}$	$\{\mathrm{x} \mid \forall y.\langle x, y\rangle \in R^{\mathcal{I}} \Rightarrow y \in C^{\mathcal{I}}\}$	Universal
$\geq n\ R$	$\{\mathrm{x} \mid \#\{y \mid \langle x, y\rangle \in R^{\mathcal{I}}\} \geq n\}$	Min cardinality
$\leq n\ R$	$\{\mathrm{x} \mid \#\{y \mid \langle x, y\rangle \in R^{\mathcal{I}}\} \leq n\}$	Max cardinality
$= \mathrm{n\ R}$	$\{\mathrm{x} \mid \#\{y \mid \langle x, y\rangle \in R^{\mathcal{I}}\} = n\}$	Cardinality
$(\bot)^{\mathcal{I}}$	$\emptyset$	Bottom
$(\top)^{\mathcal{I}}$	Δ	Top

Interpretation Example

UNIVERSITY OF Southampton
School of Electronics and Computer Science

$Y = \{v, w, x, y, z\}$
$A^{\mathcal{I}} = \{v, w, x\}$
$B^{\mathcal{I}} = \{x, y\}$
$R^{\mathcal{I}} = \{(v, w), (v, x), (y, x), (x, z)\}$

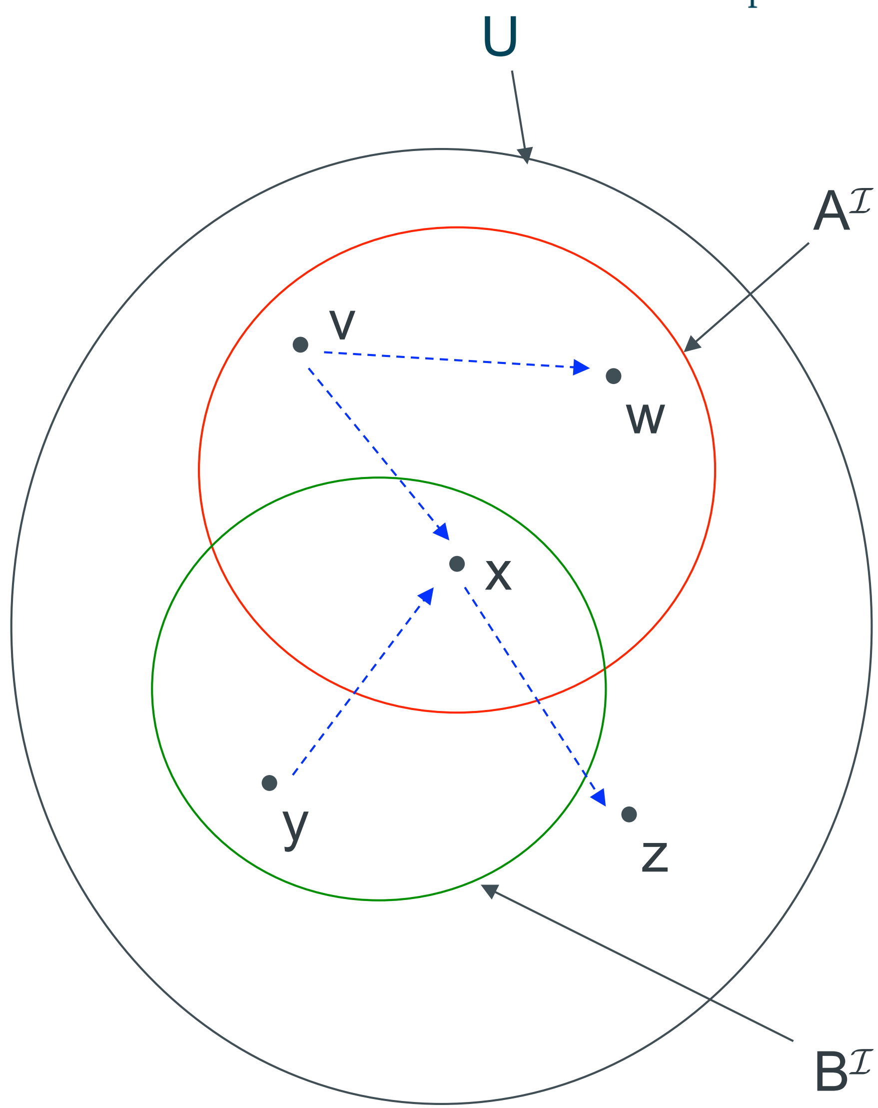
$(\neg B)^{\mathcal{I}} =$
$(A \sqcup B)^{\mathcal{I}} =$
$(\neg A \sqcap B)^{\mathcal{I}} =$
$(\exists R.B)^{\mathcal{I}} =$
$(\forall R.B)^{\mathcal{I}} =$
$(\exists R.(\exists R.A))^{\mathcal{I}} =$
$(\exists R.\neg(A \sqcap B))^{\mathcal{I}} =$



Answers

$Y = \{v, w, x, y, z\}$
$A^{\mathcal{I}} = \{v, w, x\}$
$B^{\mathcal{I}} = \{x, y\}$
$R^{\mathcal{I}} = \{(v, w), (v, x), (y, x), (x, z)\}$

$(\neg B)^{\mathcal{I}} = \{v, w, z\}$
$(A \sqcup B)^{\mathcal{I}} = \{v, w, x, y\}$
$(\neg A \sqcap B)^{\mathcal{I}} = \{y\}$
$(\exists R.B)^{\mathcal{I}} = \{v, y\}$
$(\forall R.B)^{\mathcal{I}} = \{y, w, z\}$
$(\exists R.(\exists R.A))^{\mathcal{I}} = \{\}$
$(\exists R.\neg(A \sqcap B))^{\mathcal{I}} = \{v, x\}$



DL Reasoning Revisited

- Satisfaction (can this class have any instances)
  - C is satisfiable w.r.t. K iff
    there exists some model I of K, $C^I \neq \bot$
- Subsumption
  - C subsumes D w.r.t. K iff
    for every model I of K, $C^I \supseteq D^I$
- Equivalence (mutual subsumption)
  - C is equivalent to D w.r.t. K iff
    for every model I of K, $C^I = D^I$
- Classification
  - x is an instance of C w.r.t. K iff
    for every model I of K, $x^I \in C^I$

- (where K is a knowledge base, I is an interpretation of K)