Semantics

- Introduction to main challenges
- Semantic representation
  - Logic formalism
  - Frame based and ontologies
  - Lexical semantics
- Semantic interpretation

Introduction to main challenges 0

- Semantics is the study of the meaning of linguistic sentences
- It involves
  - **Semantics Representation.** Formal representions of meaning

    Usually based on one of the following:

    - **First predicate calculus**
    - **Semantic Networks**
    - **Frame-based and ontologies**
  - **Semantic interpretation.** Theories and algorithms for mapping sentences and its formal representations –

    Usually based on semantics compositional: an object is obtained from the semantic interpretation of its components.

Introduction to main challenges 0

- Semantic representation and interpretation must be based in a formal theory, not in an "ad-hoc" process. This theory must support:
  - ambiguity
  - complex phenomena: negation, quantification, inferences, etc.
  - an interface mechanism between sintax and semantic must be defined
- The semantic interpretation of an object is obtained from the semantic interpretation of its components

Introduction to main challenges 1

SEMANTIC AMBIGUITY

- More than one semantic interpretation is possible for a given sentence

- *Peter gave a cake to the children*

Resolving ambiguous input

- Multiple alternative linguistic structures can be built
  - *I* ***made her duck***
  - **Ambiguities in the sentence**

Resolving ambiguous input

- Multiple alternative linguistic structures can be built
  - *I **made her duck***
    - *I cooked waterfowl for her*
    - *I cooked waterfowl belonging to her*
    - *I created the (plaster?) duck she owns*
    - *I caused her to quickly lowed her head or body*
    - *I waved my magic wand and turned her into undifferentiated waterfowl*
  - **Ambiguities in the sentence**
    - ***Duck*** can be noun(waterfowl) or a verb (go down)

      -> syntactic and semantic ambiguity
    - ***Her*** can be a dative pronoun or a possessive pronoun -> syntactic ambiguity
    - ***Make*** can be create or cook -> lexical ambiguity

Introduction to main challenges 2

- Different forms of inferences

  The red car ->The car is not completely red but only the external part

  The same adjective could have several meaning depending on the object

Examples obtained from recipes for cooking

One coffee spoon of sugar -> the quantity of sugar that corresponds to that in a coffee spoon

- Different levels of granularity and abstraction

  350 gr. of beans, two pieces of fruit, plenty of oil

Introduction to main challenges

Objects not quantified

- Mass

  3 Kg de rice

- Not formal metrics

  A cup of rice

- Not specific quantities

  A little bit of salt, some sugar

Introduction to main challenges

- Enumeration

Three potatoes, salt and pepper

- Global reference

Fresh fruit, garlic

- Quantification

A tea spoon of sugar

- Disjunction

One big potato or two small ones

- Not exhaustive lists

***Apples, bananas, oranges, etc*...**

Meaning Representation based on logic[1]

input:

¿Who organizes the party?

logical form:

```
(question
   (referent (X))
      (∃ X instance (X, persona)
      (el1 (Y instance(Y, party))
            (∃ Z  instance(Z, organizes)
                  present(Z)
                  value_prop(Z, agent, X)
                  value_prop(Z, patient,Y)))))
```

Meaning Representation based on logic$_2$

- This form includes four different types of knowledge:
  - Logical: The form
  - Conceptual: person, party, organize, agent,patient
  - Speech act: question
  - Pragmatics: searching for an answer, the value of X (instance of person)
- The semantic formalism must support these different types of knowledge

Meaning Representation based on $logic_3$

- A finite set of **functions** with arguments
- A finite set of **predicates** (functions that return a boolean value) with arguments
- A finite set of **constants** and **variables**
- A finite set of logical **connectors**
- A finite set of **quantifiers,** that will be applied over the predicates

Meaning Representation based on logic$_4$

- Three types of objects:
  - Boolean
    - True or false
  - Entities
    - Classes and their elements
    - Specifications of space and time
  - Functions or predicates

A cat eats a fish

$$(\exists\ X{:}cat\ \ (\exists\ Y{:}fish\ \ eats\ (X, Y)))$$

Meaning Representation based on logic$_5$

- Quantification
- Negation
- Conjuntion
- Disjuntion

The Semantic Networks 1

- Labeled directed graphs
  - nodes ==> concepts (classes or types) / objects (instances)
  - edges ==> binary relations (binary predicates) between concepts

- Global organization of the knowledge base
- Inference rules (basically, inheritance)

The Semantic Networks 2

- **Advantages**
  - Visibility
  - Associative representation. Efficient access
  - Appropriate for knowledge searching and inference
  - Representation of both general and specific knowledge
  - Supporting complex matching processes

The Semantic Networks 2

- **Disadvantages**
  - Representation of relations of arity higher than two is difficult (unary and binary relations are easily represented)
  - Representation of logic operations such as negation, implication and disjuntion is difficult
  - Representation of quantification is difficult

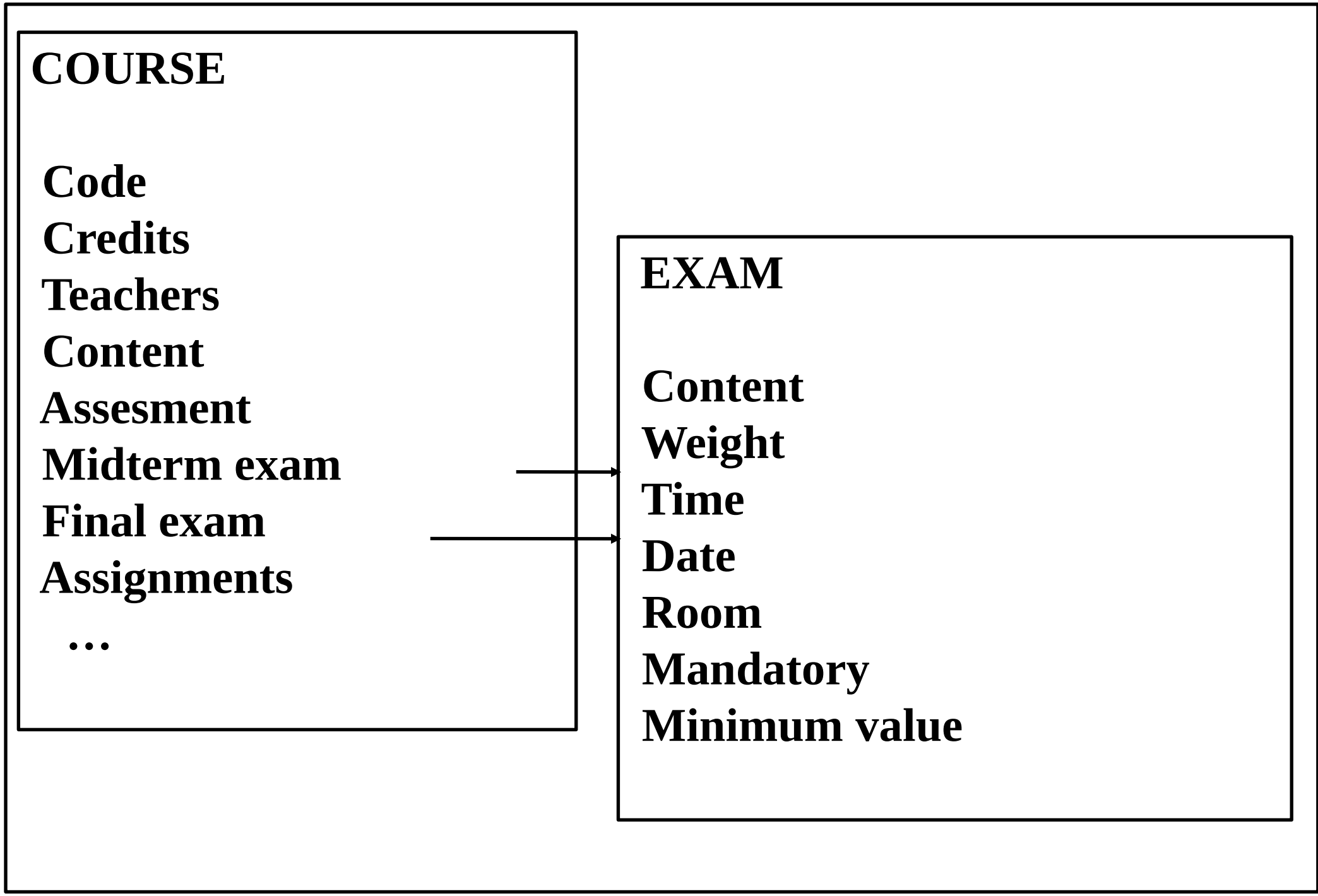
Frames 1

- A flexible way of representing concepts
- Representing not standard objects
- Concepts described by properties (or attributes)

Cat
eats: fish
Domestic: yes

Meal
Who_eats: cat
What_eats: fish

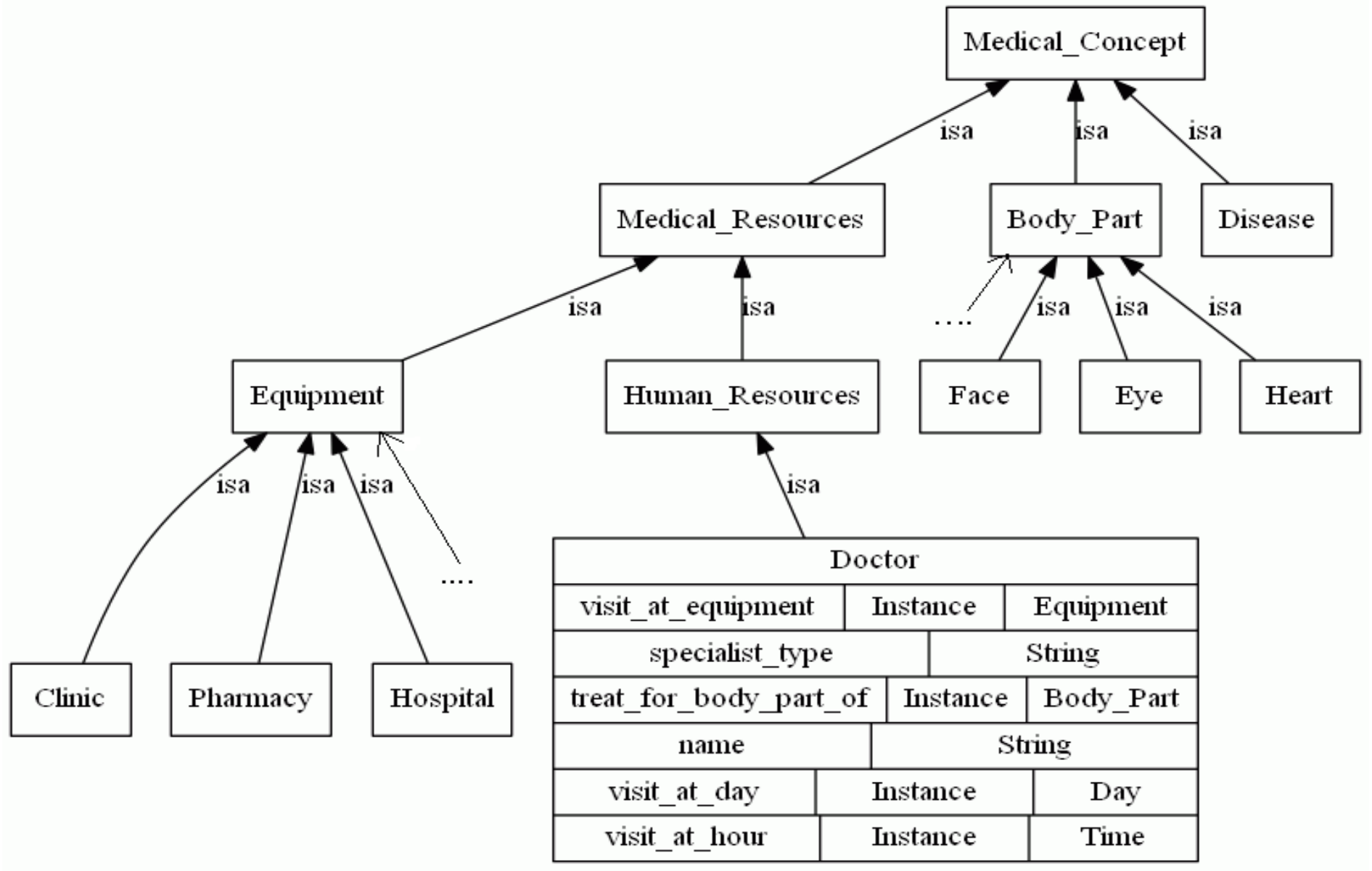
Frames 2


COURSE
Code
Credits
Teachers
Content
Assesment
Midterm exam
Final exam
Assignments
…
EXAM
Content
Weight
Time
Date
Room
Mandatory
Minimum value

Frames 3

- A flexible way of representing concepts
- Representing not standard objects
- Concepts described by properties (or attributes)
- Classes and instances
- Two basic standard relationships between classes and instances
  - ***isa*** and ***instance***
- Inheritance of properties
- New relationships between objects can be defined
- Facetts describing attributes

Looking for Medical Specialists



Frames and ontolgies

- Ontologies are based on frame formalism
- Ontologies present a more formal representation of concepts and relations than frames
- Ontologies support inference and more complex reasoning than frames
- They can differ in
  - Type, granularity and domain

Lexical Semantics$_1$

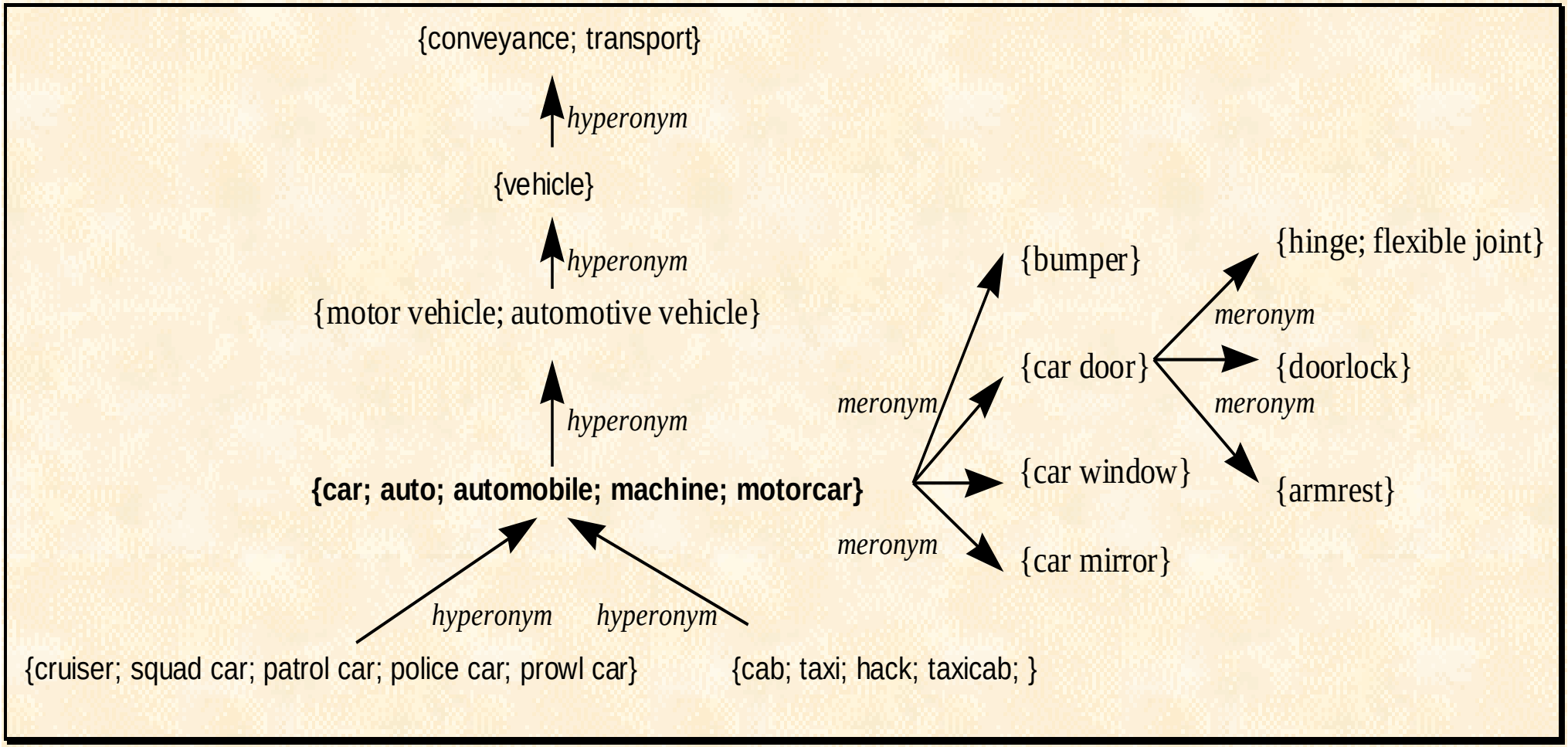
- Semantic dictionaries versus Ontologies
- Examples
  - WordNet
  - EuroWordnet
  - UMLS
  - Verbnet

Lexical Semantics$_2$
Example: WordNet

- WordNet
  - University de Princeton (Fellbaum,1998)
  - Lexicalized concepts (words)
    - synsets: set of synonyms
  - Includes nouns, verbs, adjetives and adverbs
  - Related by semantic relationships
    - Sinonima
    - antonyms
    - hyperonym-hyponym
    - implication
    - cause
    - ...
  - Wn1.5, Wn1.6,Wn 1.7: 120.000 words, 100.000 synsets
  - Wn2.0, Extended WordNet

http://www.cogsci.princeton.edu/~wn/

Fragmento de WN1.5

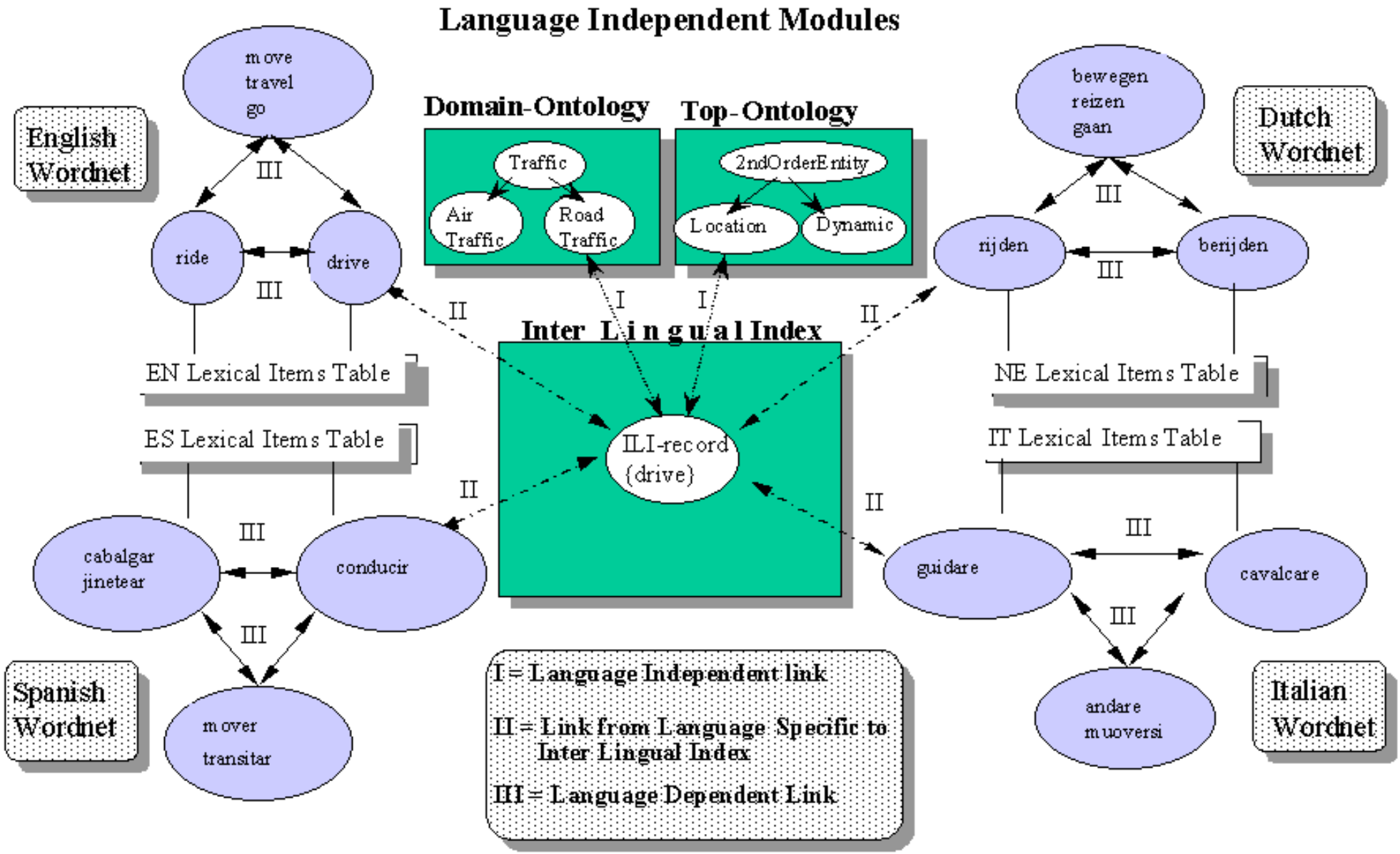


Lexical Semantics$_3$
EuroWordNet

- Project LE-2 4003 Telematics Application Programme of the European Community
- Semantic networks in different languages (Integrated)
  - English Universidad de Sheffield
  - Dutch Univ. de Amsterdam
  - Italian I.L.C. de Pisa
  - Spanish UB, UPC, U.N.E.D
- Covers basically nouns and verbs (50.000 meanings for each language)
- Rich in semantic relationships
  - inter/intra lingual, inter/intra category
- EWN2
  - German, Czech, Estonian, French
- Extensions to Catalan, Galician and Basque
- Improvements

http://www.hum.uva.nl/~ewn/
http://www.lsi.upc.es/~nlp/

Architecture of the EuroWordNet Data Structure



Lexical Semantics$_4$
Example: UMLS

- UMLS (Unified Medical Language System)
  - National Library of Medecine, USA Department of Health and Human Services
  - Resources set
    - Metathesaurus
      - 330.000 concepts, 735.000 terms
    - Semantic Network
      - Set of predefined semantic categories (135 semantic types, 51 relationships)
    - Links to the source vocabularies
      - 30 sources (multilingual)
    - Specialized lexicon with morpho-syntactic information

Lexical Semantics$_5$
VerbNet

- Computational Verbal lexicon
- Associations between syntax and semantics
  - Syntactic frames (subcategorization patterns) and selection restrictions.
  - Lexical semantic information– predicate/argument structure
  - Semantic components represented as predicates

    Links to WordNet synsets
- Entries based on precise description of Levin classes
- Temporal properties represented in an explicity form
  - during(E), end(E), result(E)

Lexical Semantics$_7$
VerbNet

Hit Class

<<MEMBERS>>	[<hit1>, <kick1>,<slap1>…]
<<THEMATIC ROLES>>	Agent(A), Patient(P), Instrument(I)
<<SELECT RESTRICTIONS>>	Agent [+animate], Patient [+concrete], Instrument [+concrete,-animate]

<<FRAMES and PREDICATES>>

Basic Transitive	A V P	Manner (during(E), directedmotion,A)^ Manner (end(E), forceful,A)^ Contact(end(E),A,P)
Transitive with Instrument	AVP with I	Manner (during (E)directedmotion,I)^ Manner (end (E),forceful,I)^ Contact (end(E),I,P)
Conative	AV at P	Manner (during (E), directedmotion, A)
With/against alternation	A V I against/on P	Manner(during (E), directedmotion, I)^ Manner(end(E), forceful, I)^ Contact (end(E), I, P)

Lexical Semantics$_8$
Levin classes (3100 verbs)

- Intermediate level between syntax and semantics
- 47 top level classes, 193 second and third level

- Based on pairs of syntactic patterns.

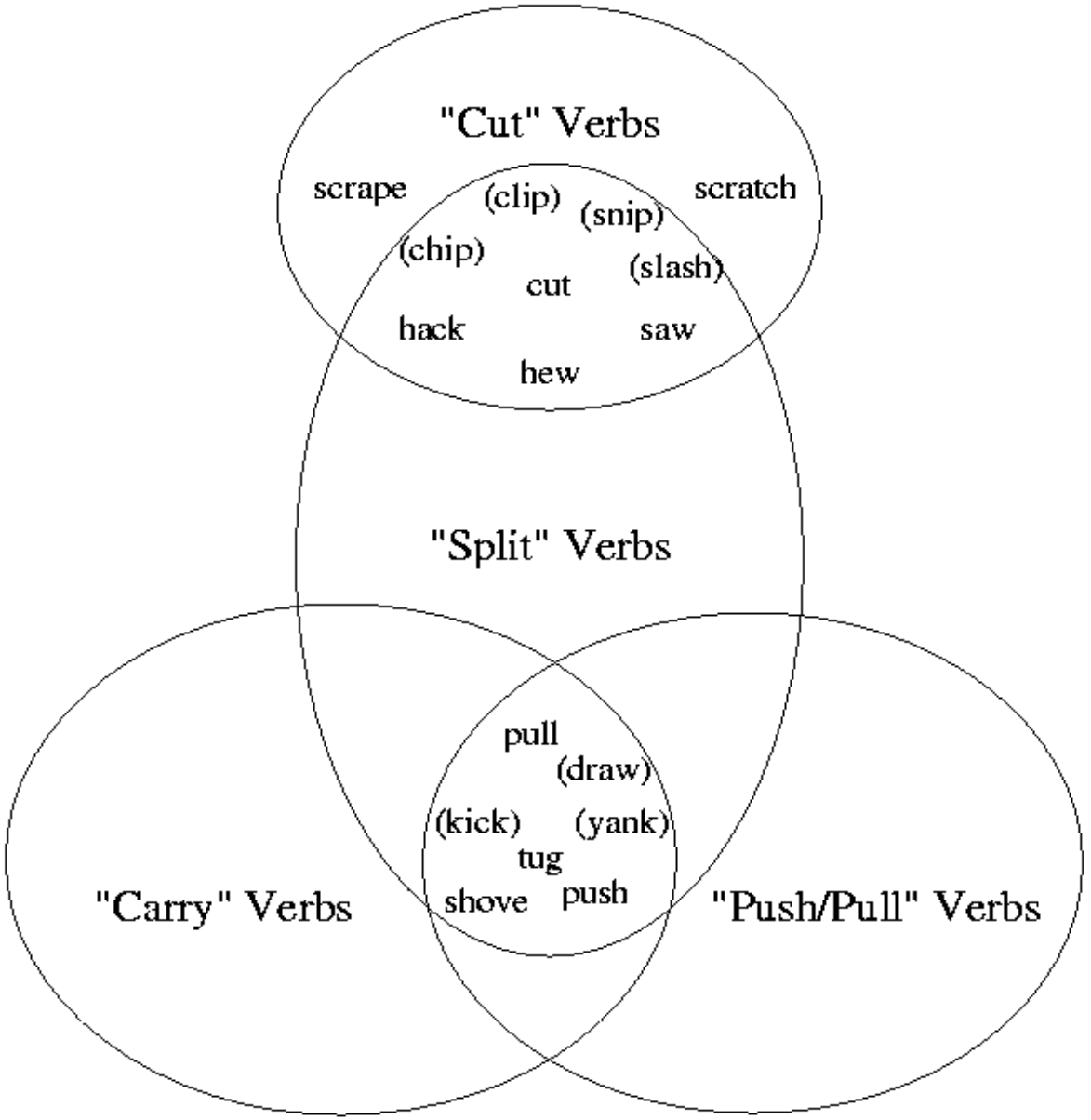
  John broke the jar. / Jars break easily. / The jar broke.
  *John cut the bread. / Bread cuts easily. / *The bread cut.*
  *John hit the wall. / *Walls hit easily. / *The wall hit.*

- Reflect implicit semantic component

  contact, directed motion,
  exertion of force, change of state

- Synonymy, syntactic patterns (subcategorization patterns)

Lexical Semantics$_9$
Intersective Levin classes



Lexical Semantics$_{10}$
Regular Sense Extensions

John pushed the chair. **+force, +contact**

John pushed the chairs apart. **+ch-state**

John pushed the chairs across the room. **+ch-loc**

John pushed at the chair. **-ch-loc**

The train whistled into the station. **+ch-loc**

The truck roared past the weigh station. **+ch-loc**