



### 3. I sistemi per la rappresentazione delle conoscenze



- Come abbiamo visto, il nostro obiettivo principale è rendere certe conoscenze disponibili ai sistemi informatici; il problema, quindi, è come rappresentare le conoscenze in un formato che sia *Computer readable*
  - nel senso che un computer possa “leggere” le conoscenze e soprattutto utilizzarle per eseguire certi compiti d’interesse applicativo.
- Dato che questo corso non ha una prospettiva informatica, non ci piaceremo all’interno della problematica delle applicazioni al web, ma concentreremo la nostra attenzione sulle rappresentazioni *simboliche* e *dichiarative*.
- Il paradigma di riferimento per le rappresentazioni dichiarative è la *logica simbolica* (o *formale*), con particolare riguardo alla *logica dei predicati del primo ordine* (*first order logic*, FOL).



I-3

2

## Introduzione (2)

- In FOL si assume che tutte le rappresentazioni riguardino un insieme non vuoto (e per il resto arbitrario) di individui, detto *universo* o *dominio*.
- Di questi individui possiamo rappresentare *proprietà* oppure *relazioni* che li leghino fra loro;
  - da questo punto di vista, un fatto è semplicemente il sussistere di una proprietà di un determinato individuo
    - “Barbara è bionda”
  - oppure il sussistere di una relazione fra più individui
    - “Alberto è più alto di Barbara”
    - “Alberto ha dato il suo cellulare a Barbara”



I-3

3

## Introduzione (3)

- Ogni metodo per la rappresentazione delle conoscenze si basa su due componenti principali:
  - un *linguaggio per la rappresentazione di conoscenze*
  - un insieme di *procedure di ragionamento*.



I-3

4

## Linguaggi di rappresentazione

- Un linguaggio per la rappresentazione di conoscenze è un linguaggio formale, con sintassi testuale o grafica, le cui *espressioni* sono utilizzate per rappresentare elementi di conoscenza.
  - Come esempio vogliamo rappresentare il significato del termine "madre" come "donna con almeno un figlio".
  - Per utilizzare FOL parafrasiamo la definizione di "madre" come:

( $x$  è una madre) se e solo se ( $x$  è una donna ed esiste almeno un  $y$  tale che  $x$  è genitore di  $y$ ).



I-3

5

## Esempio di rappresentazione in FOL

- Possiamo esprimere questa definizione usando una notazione testuale, come ad esempio il linguaggio di FOL solitamente presentato nei libri di logica

$$\forall x (\text{MADRE}(x) \leftrightarrow \text{DONNA}(x) \wedge \exists y \text{GenDi}(x,y))$$

- dove:
  - $\forall x \dots$  per ogni  $x$ , ... (*il quantificatore universale*)
  - $\exists y \dots$  esiste un  $y$  tale che ... (*il quantificatore esistenziale*)
  - $\dots \leftrightarrow \dots$  se e solo se ... (*il connettivo bicondizionale*)
  - $\dots \wedge \dots$  e ... (*il connettivo di congiunzione*)
  - $\text{MADRE}(x)$ ,  $\text{DONNA}(x)$  due *predicati monoargomentali*, che esprimono proprietà di  $x$
  - $\text{GenDi}(x,y)$  un *predicato biargomentale*, che esprime una relazione binaria fra  $x$  e  $y$

[Per i simboli di FOL si è usato il font Symbol]



I-3

6

## Ragionamento automatico

- Nel contesto in cui ci stiamo muovendo, per “ragionamento” s’intende il *ragionamento deduttivo* o *deduzione*.
- Una deduzione è un processo che fa passare
  - da alcune espressioni (dette *premesse* o *ipotesi*)
  - a un’espressione (detta *conclusione* o *tesi*),
  - in modo tale da *conservare l’eventuale verità delle premesse*
- in altre parole, se le premesse sono vere, lo sarà anche la conclusione.



I-3

7

## Esempio di “ragionamento”

- Ad esempio, dati come premesse
  - la definizione di “madre”,
  - il fatto che Laura è una donna,
  - il fatto che Laura è genitore di Marco,
- si può dedurre come conclusione che
  - Laura è una madre.



I-3

8

- In FOL la deduzione può essere rappresentata in vari modi, ad esempio tramite la prova seguente
  1.  $\forall x (MADRE(x) \leftrightarrow DONNA(x) \wedge \exists y GenDi(x,y))$  premessa
  2.  $DONNA(laura)$  premessa
  3.  $GenDi(laura,marco)$  premessa
  4.  $MADRE(laura) \leftrightarrow DONNA(laura) \wedge \exists y GenDi(laura,y)$  da 1 per eliminazione di  $\forall$
  5.  $\exists y GenDi(laura,y)$  da 3 per introduzione di  $\exists$
  6.  $DONNA(laura) \wedge \exists y GenDi(laura,y)$  da 2 e 3 per introduzione di  $\wedge$
  7.  $MADRE(laura)$  da 4 e 6 per eliminazione di  $\leftrightarrow$
- il metodo di prova utilizzato è noto come “calcolo della deduzione naturale”; l’ultima riga della prova rappresenta la conclusione



I-3

9

## Espressività e “decidibilità”

- Utilizzando FOL si possono esprimere conoscenze molto articolate e, almeno in linea di principio, eseguire in modo automatico ragionamenti complessi.
- C’è però un problema:
  - in FOL la procedura di deduzione (detta anche *procedura di prova o calcolo*) non è una procedura di decisione, ma soltanto di *semidecisione*.
- Ciò significa che:
  - quando la conclusione è deducibile dalle premesse, la procedura termina in un numero finito di passi producendo una prova;
  - quando invece la conclusione non è deducibile dalle premesse, la procedura può non terminare.



I-3

10

## Espressività e "decidibilità" (2)

- In parole povere, un sistema informatico può "andare in ciclo" se tenta di dedurre da un insieme di premesse una conclusione che in effetti non è deducibile dalle premesse.
- L'indecidibilità di FOL dipende dalla notevole espressività del linguaggio:
  - più un linguaggio di rappresentazione è espressivo,
  - più sono problematiche le procedure di ragionamento basate su di esso.



I-3

11

## Sottolinguaggi di FOL

- Molte ricerche nel campo dei linguaggi di rappresentazione delle conoscenze hanno l'obiettivo di identificare un sottolinguaggio di FOL tale che:
  - il linguaggio sia comunque abbastanza espressivo per le applicazioni;
  - la deduzione si basi su una procedura di decisione
    - che quindi termini in ogni caso dopo un numero finito di passi, sia quando la conclusione è deducibile dalle premesse, sia quando non lo è;
  - la procedura di decisione abbia complessità computazionale accettabile
    - ovvero richieda una quantità accettabile di risorse di calcolo.



I-3

12

## Le Logiche Descrittive

- I sistemi di questo tipo hanno preso il nome di *logiche descrittive* (*description logic*, DL). Le DL utilizzano una sintassi semplificata rispetto a FOL.
- Ad esempio, le tre premesse
  1.  $\forall x (\text{MADRE}(x) \leftrightarrow \text{DONNA}(x) \wedge \exists y \text{GenDi}(x,y))$
  2.  $\text{DONNA}(\text{laura})$
  3.  $\text{GenDi}(\text{marco}, \text{laura})$
- in logica descrittiva verrebbero rappresentate come:
  1.  $\text{MADRE} \equiv \text{DONNA} \exists \text{GenDi}$
  2.  $\text{DONNA}(\text{laura})$
  3.  $\text{GenDi}(\text{laura}, \text{marco})$

[Per i simboli delle DL si è utilizzato il font Lucida Math Symbol]



I-3

13

## Le rappresentazioni

- In svariate aree dell'informatica si utilizzano sistemi di rappresentazione per scopi ed usi differenti.
- Sono ad esempio sistemi di rappresentazione le basi di dati, i diagrammi UML, i documenti XML.
- Al di là delle ovvie differenze, tutti i sistemi di rappresentazione condividono alcune caratteristiche comuni, che cercheremo qui di analizzare.
- Il primo concetto da analizzare è la distinzione fra *realtà* e *rappresentazione*.
- Ogni rappresentazione, infatti, è necessariamente la rappresentazione di qualcosa, e questo "qualcosa" è un frammento della realtà interessante per qualche applicazione.



I-3

14

## Tre livelli di una rappresentazione

- Livello 1: modello concreto.
  - Il modello concreto di un frammento di realtà è la rappresentazione dei fatti che sussistono nel frammento di realtà (conoscenze fattuali sul frammento di realtà);
    - ad esempio, "Emanuele Della Valle è un docente di una facoltà di Scienze MM.FF.NN. dell'Università dell'Insubria".



I-3

15

## Tre livelli di una rappresentazione (2)

- Livello 2: *modello concettuale*.
  - Il modello concettuale di un frammento di realtà è una rappresentazione dei concetti utilizzati per formulare il modello concreto (conoscenze terminologiche e nomologiche);
    - ad esempio, per esprimere il fatto dell'esempio precedente è necessario definire i concetti di persona (dotato di nome e cognome), di docente, di facoltà e così via.





I-3

16

## Tre livelli di una rappresentazione (3)

- Livello 3: *metamodello*.
  - Il metamodello di un sistema di rappresentazione è la specifica degli strumenti formali utilizzabili per definire il modello concettuale e il modello concreto;
    - ad esempio: il linguaggio di FOL, oppure il formalismo degli schemi E-R, oppure il formalismo del diagramma delle classi UML, e così via.



I-3

17

## Tre livelli di una rappresentazione (4)

- Livello 1: *modello concreto*.
  - Il modello concreto di un frammento di realtà è la rappresentazione dei fatti che sussistono nel frammento di realtà (conoscenze fattuali sul frammento di realtà)
- Livello 2: *modello concettuale*.
  - Il modello concettuale di un frammento di realtà è una rappresentazione dei concetti utilizzati per formulare il modello concreto (conoscenze terminologiche e nomologiche);
- Livello 3: *metamodello*.
  - Il metamodello di un sistema di rappresentazione è la specifica degli strumenti formali utilizzabili per definire il modello concettuale e il modello concreto;



I-3

18

## Rappresentazioni a confronto

- Livello 1: *modello concreto*:
  - conoscenze fattuali sul frammento di realtà
- Livello 2: *modello concettuale*:
  - conoscenze terminologiche e nomologiche
- Livello 3: *metamodello*:
  - gli strumenti formali utilizzabili per definire il modello concettuale e il modello concreto;

	<i>basi di dati</i>	<i>specifiche SW</i>	<i>logica</i>	<i>RDF</i>	<i>DL</i>
meta modello	il formalismo E-R	il formalismo UML	la logica predicativa del primo ordine	il formalismo delle triple RDF	Una particolare DL
Modello concettuale	un modello concettuale E-R	un diagramma delle classi	una teoria del primo ordine	un documento RDFS	una TBox
modello concreto	un'istanza di base di dati	le istanze delle classi	il "diagramma" di un modello	un documento RDF	un'ABox



I-3

19

## Diversi usi dei Modelli Concettuali

- In che cosa, allora, un KBS si distingue, poniamo, da una rappresentazione nel campo delle basi dei dati?
- La differenza principale sta in questo:
  - nei campi delle basi di dati i modelli concettuali (E-R) rappresentano un passo importante nello sviluppo di un sistema, ma non fanno parte del prodotto finale;
  - nei KBS, invece, i modelli concettuali sono parte integrante del prodotto finale, in quanto sono disponibili e vengono utilizzati a run time dai sistemi software.



I-3

20

## Componenti di un KBS

- una base di conoscenze (KB), a sua volta spesso costituita da due componenti distinte:
  - una *terminological box (TBox)*,
    - che rende disponibile e accessibile la rappresentazione formale del modello concettuale di un frammento di realtà;
  - un'*assertion box (ABox)*,
    - che rende disponibile e accessibile la rappresentazione formale del modello concreto di un frammento di realtà;



I-3

21

## Componenti di un KBS (2)

- un'*interfaccia di accesso*, che consente l'accesso alle conoscenze contenute nella TBox e nella ABox da parte di applicazioni software, e tipicamente di:
  - un insieme di *servizi di ragionamento (reasoning services)*,
    - ovvero un'applicazione software in grado di dedurre nuove conoscenze dalle conoscenze contenute nella TBox e nella ABox;
  - un'*interfaccia di editing*,
    - ovvero un'applicazione software che consente la gestione dei contenuti della TBox e della ABox da parte di un operatore umano;
  - un insieme di *applicazioni software*
    - che dipendono dalla specifica destinazione del KBS.

