

Web semantico e logiche descrittive

M. Simi, 2014-2015
Cap 2 del "Description Logic Handbook"
Lezioni di U.Straccia

Categorie e oggetti

- Molti dei ragionamenti che si fanno sono sulle categorie piuttosto che sugli individui
- Se organizziamo la conoscenza in categorie (e sottocategorie) è sufficiente classificare un oggetto, tramite le proprietà percepite, per inferire le proprietà delle categorie a cui appartiene
- Ereditarietà come forma di inferenza

Ontologie di dominio

- Le idee delle reti semantiche e dei frame sono state raccolte in logiche specializzate
- Queste logiche sono alla base delle proposte per il *Web semantico*
- *Ontologia*: modello formale di un dominio di interesse (una concettualizzazione)
- Le relazioni di sottoclasse organizzano la conoscenza in tassonomie gerarchiche (come in botanica, biologia, nelle scienze librarie ...)

Il Web semantico

- La visione di Tim Berners-Lee (1998): da un Web "sintattico" per la comunicazione tra persone al Web "semantico", una grossa rete di informazioni collegate su scala globale comprensibili ai programmi (un database globale – *linked data*)
- Il protocollo di accesso ai dati è HTTP, la codifica è Unicode, il formato di scambio è XML.

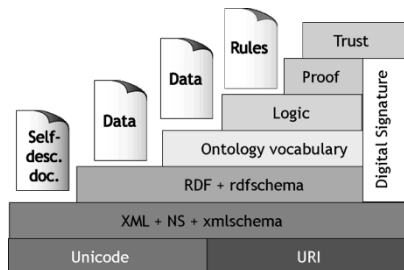
Le motivazioni del web semantico

- Il web di oggi, fatto di documenti, è adatto unicamente alla fruizione da parte delle persone
 - i motori di ricerca si basano su parole chiave e i risultati sono documenti, non dati utilizzabili dai programmi
 - Impossibile|difficile costruire dei programmi che fanno uso efficace dei contenuti web
- Il web semantico si propone di aiutare nella gestione della conoscenza presente nel web
- Un **web di dati** da condividere e riutilizzare tra applicazioni, aziende e comunità

Web semantico e gestione della conoscenza

- Ricerca dell'informazione
 - Ricerche meno sintattiche
- Estrazione dell'informazione
 - Adesso sono le persone che ricercano, interpretano e combinano i contenuti delle pagine web
- Manutenzione dell'informazione
 - Terminologie inconsistenti, informazioni obsolete
- Presentazione dell'informazione
 - Impossibile definire "viste" sui contenuti web

Il livelli del Web semantico



I livelli del web semantico

- Unicode e URI
- XML interoperabilità sintattica
- RDF (Resource Description Framework): per descrivere relazioni semantiche tra risorse (*soggetto, predicato, oggetto*)
- RDF schema (RDFS): definire classi di oggetti, relazioni tra classi, per vincolare domini e codomini delle relazioni; RDFS linguaggio per ontologie, poco espressivo

Web semantico e linguaggi per ontologie

- Linguaggi per l'aggiunta di un servizio inferenziale a RDF. Diverse proposte:
 - OIL gruppo europeo
 - DAML-ONT gruppo americano
 - DAML+OIL proposto come standard
 - OWL: Web Ontology Language, standard W3C.
- OWL evolve dalle logiche descrittive

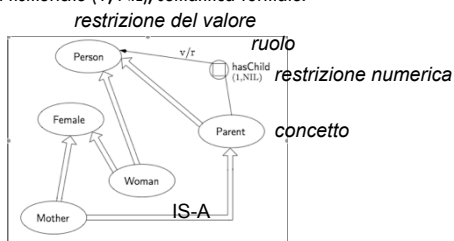
Logiche descrittive

Possono essere viste come:

- *Evoluzioni "logiche"* di linguaggi di KR "a rete", come i *frame* e le *reti semantiche*
- *Contrazioni* della logica del prim'ordine (FOL) per ottenere migliori proprietà computazionali

KL-One

- KL-One [Brachman-Scholze 1985] introduce le idee importanti delle DL: concetti e ruoli, restrizioni sui valori, restrizioni numeriche (1, NIL), semantica formale.



DL come formalizzazione di reti semantiche

- Verso gli anni 80 si ha una sterzata verso la logica delle reti semantiche
- Il processo consiste nel
 - riformulare i costrutti secondo i canoni della logica
 - eliminare i costrutti che non si prestano a tale riformulazione (*default* ed eccezioni)

Da KL-One alle logiche descrittive

- Logiche terminologiche
 - \mathcal{FL} - (Frame Language) [Brachman and Levesque, 1984] tradeoff tra espressività di un linguaggio di rappresentazione e la complessità del ragionamento
 - CLASSIC [Brachman 1991], limitato e completo
- Logiche descrittive
 - LOOM [MacGregor-Bates 1987], BACK [Nebel-vonLuck, 1988], espressivi e incompleti
 - KRIS [Baader, Hollunder, 1991], espressivi e completi
 - FaCT, DLP, Racer, sistemi ottimizzati per logiche espressive

Esempio

La seguente è una formula di una delle DL:

$(\text{and Paper } (\text{atmost } 2 \text{ hasAuthor})$
 $(\text{atleast } 2 \text{ hasAuthor})) [\text{paper3}]$

$\text{paper3: Paper } \sqcap (\leq 2 \text{ hasAuthor}) \sqcap (\geq 2 \text{ hasAuthor})$
 (sintassi alternativa)

equivalente a:

$\text{Paper}(\text{paper3}) \wedge$

$\exists x \text{ hasAuthor}(\text{paper3}, x) \wedge$

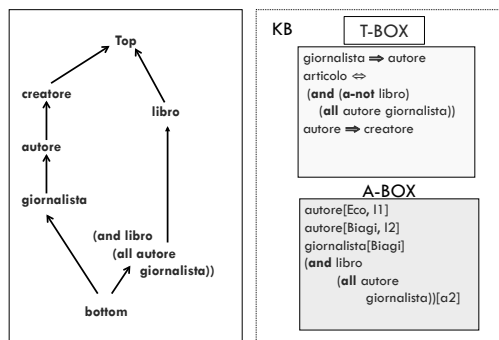
$\exists y \text{ hasAuthor}(\text{paper3}, y) \wedge x \neq y \wedge$

$\text{hasAuthor}(\text{paper3}, z) \Rightarrow (z = x) \vee (z = y)$

Concetti, ruoli, individui

- Ogni DL è caratterizzata da operatori per la costruzione di termini di due tipi:
 - Concetti, corrispondenti a relazioni unarie con operatori per la costruzione di concetti complessi: and (\sqcap), or (\sqcup), not (\neg), all (\forall), some (\exists), atleast ($\geq n$), atmost ($\leq n$), ...
 - Ruoli, corrispondenti a relazioni binarie ed eventualmente operatori per costruire ruoli complessi
- Individui: usati solo nelle asserzioni

Una KB basata su logica descrittiva



La logica \mathcal{AL} : la sintassi dei termini

$\langle \text{concetto} \rangle \rightarrow A$

- | \top (top, concetto universale)
- | \perp (bottom)
- | $\neg A$ (negazione atomica)
- | $C \sqcap D$ (intersezione)
- | $\forall R.C$ (restrizione di valore)
- | $\exists R.T$ (esistenziale debole)

$\langle \text{ruolo} \rangle \rightarrow R$

A, B concetti primitivi R ruolo primitivo

C, D concetti

Esempi

- Person \sqcap Female
- Person $\sqcap \neg$ Female
- Person $\sqcap \exists \text{ hasChild.T}$
- Person $\sqcap \forall \text{ hasChild.Female}$
- Person $\sqcap \forall \text{ hasChild.}\perp$

Semantica di \mathcal{AL}

$\Delta^{\mathcal{I}}$ dominio di interpretazione, un insieme di individui

\mathcal{I} funzione di interpretazione che assegna a:

- concetti atomici A : $A^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$
- ruoli atomici R : $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$
- nomi di individuo a : $a^{\mathcal{I}} \in \Delta^{\mathcal{I}}$

$\top^{\mathcal{I}} = \Delta^{\mathcal{I}}$

$\perp^{\mathcal{I}} = \emptyset$

$(\neg A)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus A^{\mathcal{I}}$ il complemento

$(C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}$

$(\forall R.C)^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid \forall b.(a, b) \in R^{\mathcal{I}} \rightarrow b \in C^{\mathcal{I}}\}$

$(\exists R.T)^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid \exists b.(a, b) \in R^{\mathcal{I}}\}$

Definizioni

- Un'interpretazione \mathcal{I} è un *modello* di un concetto C sse $C^{\mathcal{I}} \neq \emptyset$ (insieme vuoto). Lo stesso per i ruoli.
- $C \equiv D$ (equivalenti) sse $C^{\mathcal{I}} = D^{\mathcal{I}}$ per ogni \mathcal{I}
- Esempio:
 - $\forall \text{haFiglio.Femmina} \sqcap \forall \text{haFiglio.Studente} =$
 - $\forall \text{haFiglio.Femmina} \sqcap \text{Studente}$

Esempio 1

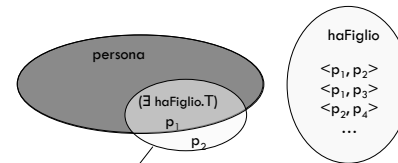
Articolo $\sqcap \exists \text{haAutore.T} \sqcap$

$\forall \text{haAutore.Giornalista}$

“I insieme degli articoli che hanno almeno un autore, e i cui autori sono tutti giornalisti”

Esempio 2

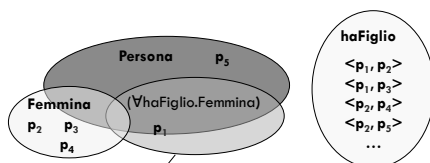
Persona $\sqcap \exists \text{haFiglio.T}$



$\{a \in \Delta^{\mathcal{I}} \mid \exists b (a, b) \in \text{haFiglio}^{\mathcal{I}}\}$

Esempio 3

Persona $\sqcap \forall \text{haFiglio.Femmina}$



$\{a \in \Delta^{\mathcal{I}} \mid \forall b (a, b) \in \text{haFiglio}^{\mathcal{I}} \Rightarrow b \in \text{Femmina}^{\mathcal{I}}\}$

Logiche più espressive

\mathcal{U} : unione, $(C \sqcup D)^{\mathcal{I}} = (C^{\mathcal{I}} \cup D^{\mathcal{I}})$

\mathcal{E} : esistenziale pieno

$(\exists R.C)^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid \exists b.(a, b) \in R^{\mathcal{I}} \wedge b \in C^{\mathcal{I}}\}$

\mathcal{N} : restrizioni numeriche

$(\geq n R)^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid |\{b \mid (a, b) \in R^{\mathcal{I}}\}| \geq n\}$ (atleast)

$(\leq n R)^{\mathcal{I}} = \{a \in \Delta^{\mathcal{I}} \mid |\{b \mid (a, b) \in R^{\mathcal{I}}\}| \leq n\}$ (atmost)

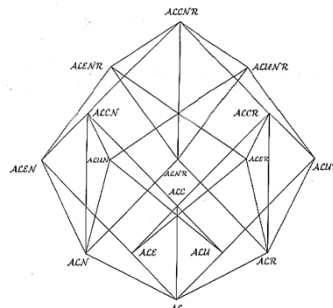
n , numero intero $|\cdot|$ cardinalità dell'insieme

C : complemento pieno, $(\neg C)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$

La famiglia \mathcal{AL}

- Le diverse DL sono ottenute aggiungendo altri costruttori di termini ad \mathcal{AL}
 $\mathcal{AL}[U][E][M][C]$
- Non tutte distinte
 - $\mathcal{AL}UE = \mathcal{AL}C$ dato che $(C \sqcup D) \equiv \neg(\neg C \sqcap \neg D)$
 e $\exists R.C \equiv \neg \forall R. \neg C$
 - $\mathcal{AL}CN = \mathcal{AL}UEN$

Il reticolo della famiglia \mathcal{AL}



© Paolo Buongarzone & Rossella

Il linguaggio della T-BOX

- Assiomi terminologici \mathcal{T}
 - $C \sqsubseteq D$ inclusione di concetti, $C^I \subseteq D^I$
 - $R \sqsubseteq S$ inclusione di ruoli, $R^I \subseteq S^I$
 - $C \equiv D$ uguaglianza di concetti, $C^I \equiv D^I$
 - $R \equiv S$ uguaglianza di ruoli, $R^I \equiv S^I$
- I soddisfa \mathcal{T} sse soddisfa ogni elemento in \mathcal{T}

Terminologia (T-BOX)

- Definizioni:** uguaglianze che introducono un simbolo sulla sinistra
 - $Mother \equiv Woman \sqcap hasChild.Person$
- Terminologia:** i simboli compaiono sulla sinistra non più di una volta.
- Simboli primitivi:** compaiono solo a destra
- Simboli definiti:** compaiono anche sulla sinistra
- Assumiamo \mathcal{T} acicliche.

Una terminologia aciclica

Woman	\equiv	Person \sqcap Female
Man	\equiv	Person $\sqcap \neg$ Woman
Mother	\equiv	Woman $\sqcap \exists hasChild.Person$
Father	\equiv	Man $\sqcap \exists hasChild.Person$
Parent	\equiv	Father \sqcup Mother
Grandmother	\equiv	Mother $\sqcap \exists hasChild.Parent$
MotherWithManyChildren	\equiv	Mother $\sqcap \geq 3 hasChild$
MotherWithoutDaughter	\equiv	Mother $\sqcap \forall hasChild. \neg$ Woman
Wife	\equiv	Woman $\sqcap \exists hasHusband.Man$

Espansione di \mathcal{T}

- Se una terminologia è aciclica può essere espansa sostituendo ai simboli definiti le loro definizioni.
- Il processo converge e l'espansione, \mathcal{T}^e , è unica.
- Proprietà di \mathcal{T}^e :
 - in \mathcal{T}^e ogni uguaglianza ha la forma $C \equiv D^e$ dove D^e contiene solo simboli primitivi
 - \mathcal{T}^e contiene gli stessi simboli primitivi e definiti di \mathcal{T}
 - \mathcal{T}^e è equivalente a \mathcal{T}

Esempio: espansione

Woman \equiv Person \sqcap Female
 Man \equiv Person $\sqcap \neg(\text{Person} \sqcap \text{Female})$
 Mother \equiv (Person \sqcap Female) $\sqcap \exists \text{hasChild.Person}$
 Father \equiv (Person $\sqcap \neg(\text{Person} \sqcap \text{Female})) \sqcap \exists \text{hasChild.Person}$
 Parent \equiv ((Person $\sqcap \neg(\text{Person} \sqcap \text{Female})) \sqcap \exists \text{hasChild.Person}) \sqcup ((\text{Person} \sqcap \text{Female}) \sqcap \exists \text{hasChild.Person})$
 Grandmother \equiv ((Person \sqcap Female) $\sqcap \exists \text{hasChild.Person}) \sqcap \exists \text{hasChild}(((\text{Person} \sqcap \neg(\text{Person} \sqcap \text{Female})) \sqcap \exists \text{hasChild.Person}) \sqcup ((\text{Person} \sqcap \text{Female}) \sqcap \exists \text{hasChild.Person}))$
 MotherWithManyChildren \equiv ((Person \sqcap Female) $\sqcap \exists \text{hasChild.Person}) \sqcap \geq 3 \text{hasChild}$
 MotherWithoutDaughter \equiv ((Person \sqcap Female) $\sqcap \exists \text{hasChild.Person}) \sqcap \forall \text{hasChild}.\neg(\text{Person} \sqcap \text{Female})$
 Wife \equiv (Person \sqcap Female) $\sqcap \exists \text{hasHusband}.\neg(\text{Person} \sqcap \neg(\text{Person} \sqcap \text{Female}))$

Assiomi di inclusione

- **Specializzazione:** un assioma di inclusione la cui parte sinistra è atomica.

Woman \sqsubseteq Person

- **Normalizzazione:** Una terminologia *generalizzata* [con assiomi di inclusione], se aciclica, può essere trasformata in una terminologia equivalente con solo definizioni:

$A \sqsubseteq C \rightarrow A \equiv A' \sqcap C$

con A' nuovo simbolo primitivo

Proprietà della normalizzazione

- Sia \mathcal{T} una terminologia generalizzata e \mathcal{T}' una sua normalizzazione.
 1. Ogni modello di \mathcal{T}' è un modello di \mathcal{T} .
 2. Per ogni modello \mathcal{I} di \mathcal{T} c'è un modello \mathcal{I}' di \mathcal{T}' che ha lo stesso dominio di \mathcal{I} e coincide con \mathcal{I} sui concetti e ruoli atomici in \mathcal{T} .
- Le inclusioni non aggiungono potere espressivo nel caso di terminologie acicliche

Il linguaggio delle asserzioni: A-BOX

- Una A-BOX è un insieme di asserzioni di due tipi:
 - $a:C$, asserzioni su concetti, $a^{\mathcal{I}} \in C^{\mathcal{I}}$
 - $(b, c):R$, asserzioni su ruoli, $(b^{\mathcal{I}}, c^{\mathcal{I}}) \in R^{\mathcal{I}}$
 - $a, b, c, d \dots$ sono meta-simboli per individui
 - \mathcal{I} fornisce anche una interpretazione per i simboli di individuo

Esempio di A-BOX

Mary: Mother	Peter: Father
(Mary, Peter): hasChild	(Peter, Harry): hasChild
(Mary, Paul): hasChild	

- Assunzione di *mondo aperto* (OWA): non si assume di specificare tutto
- Assunzione di *nome unico* (UNA): simboli diversi, individui diversi

DL come frammenti del FOL

- È possibile convertire asserzioni delle logiche descrittive in formule FOL
- La traduzione avviene attraverso la definizione di una funzione di traduzione $t(C, x)$ che restituisce una formula del FOL con x libera

$$t(C, x) \mapsto C(x)$$

Traduzione da DL a FOL

$\tau(C \sqsubseteq D) \mapsto \forall x. t(C, x) \Rightarrow t(D, x)$
 $\tau(a:C) \mapsto t(C, a)$
 $\tau((a, b):R) \mapsto R(a, b)$

$\tau(\top, x) \mapsto true$
 $\tau(\perp, x) \mapsto false$
 $\tau(A, x) \mapsto A(x) \quad A \text{ atomico}$
 $\tau(C \sqcap D, x) \mapsto t(C, x) \wedge t(D, x)$
 $\tau(C \sqcup D, x) \mapsto t(C, x) \vee t(D, x)$

Traduzione da DL a FOL (cont.)

$\tau(\neg C, x) \mapsto \neg t(C, x)$
 $\tau(\exists R.C, x) \mapsto \exists y. R(x, y) \wedge t(C, y)$
 $\tau(\forall R.C, x) \mapsto \forall y. R(x, y) \Rightarrow t(C, y)$

Esempi di traduzione

$\tau(\text{HappyFather} \sqsubseteq \text{Man} \sqcap \exists \text{hasChild} . \text{Female}) =$
 $\forall x. t(\text{HappyFather}, x) \Rightarrow t(\text{Man} \sqcap \exists \text{hasChild} . \text{Female}, x) =$
 $\forall x. \text{HappyFather}(x) \Rightarrow t(\text{Man}, x) \wedge t(\exists \text{hasChild} . \text{Female}, x) =$
 $\forall x. \text{HappyFather}(x) \Rightarrow \text{Man}(x) \wedge t(\exists \text{hasChild} . \text{Female}, x) =$
 $\forall x. \text{HappyFather}(x) \Rightarrow \text{Man}(x) \wedge \exists y. \text{hasChild}(x, y) \wedge \text{Female}(y)$
 $\tau(a:\text{Man} \sqcap \exists \text{hasChild} . \text{Female}) =$
 $\text{Man}(a) \wedge (\exists y. \text{hasChild}(a, y) \wedge \text{Female}(y))$

Sintassi alternativa Lisp-like

\top	\rightarrow *top*	$(\leq n R.C)$	\rightarrow (at-most n R C)
\perp	\rightarrow *bottom*	$\{\bar{a}\}$	\rightarrow (one-of a)
$\neg C$	\rightarrow (not C)	R^{-}	\rightarrow (inv R)
$C \sqcap D$	\rightarrow (and C D)	$C \sqsubseteq D$	\rightarrow (implies C D)
$C \sqcup D$	\rightarrow (or C D)	$A = C$	\rightarrow (define-concept A C)
$\exists R.C$	\rightarrow (some R C)	$R \sqsubseteq P$	\rightarrow (implies-role R P)
$\forall R.C$	\rightarrow (all R C)	$fun(f)$	\rightarrow (functional f)
$(\geq n R)$	\rightarrow (at-least n R)	$trans(R)$	\rightarrow (transitive R)
$(\leq n R)$	\rightarrow (at-most n R)	aC	\rightarrow (instance a C)
$(\geq n R.C)$	\rightarrow (at-least n R C)	$(a, b):R$	\rightarrow (related a b R)