

Aristotele e l'informatica

Maurizio Lenzerini

DIPARTIMENTO DI INFORMATICA
E SISTEMISTICA ANTONIO RUBERTI



SAPIENZA
UNIVERSITÀ DI ROMA

Lectio Magistralis, Aula Magna, 16 dicembre 2010

Complessità dei sistemi informativi

Frammento di archivio elettronico di un istituto bancario:

CUC	TS_START	TS_END	ID_GRUP	FLAG_CP	FLAG_CF	FATTURATO	FLAG_FATT	
124589	30-lug-2004	1-gen-9999	92736	S	N	195000,00	N	
140904	15-mag-2001	15-giu-2005	35060	N	N	230600,00	N	
124589	5-mag-2001	30-lug-2004	92736	N	S	195000,00	S	
-452901	13-mag-2001	27-lug-2004	92770	S	N	392000,00	N	
129008	10-mag-2001	1-gen-9999	62010	N	S	247000,00	S	
-472900	10-mag-2001	1-gen-9999	62010	S	N	0 00	N	
130976	7-mag-2001	9-lug-2003	75680					

Complessità dei sistemi informativi

*valore negativo indica “cliente fittizio”,
ovvero una holding di persone fisiche*

CUC	TS_START	TS_END	ID_GRUP	FLAG_CP	FLAG_CF	FATTURATO	FLAG_FATT	
124589	30-lug-2004	1-gen-9999	92736	S	N	195000,00	N	
140904	15-mag-2001	15-giu-2005	35060	N	N	230600,00	N	
124589	5-mag-2001	30-lug-2004	92736	N	S	195000,00	S	
-452901	13-mag-2001	27-lug-2004	92770	S	N	392000,00	N	
129008	10-mag-2001	1-gen-9999	62010	N	S	247000,00	S	
-472900	10-mag-2001	1-gen-9999	62010	S	N	0 00	N	
130976	7-mag-2001	9-lug-2003	75680					

Complessità dei sistemi informativi

S indica che il cliente è capogruppo del gruppo a cui appartiene

S indica che il cliente è capofila del gruppo a cui appartiene

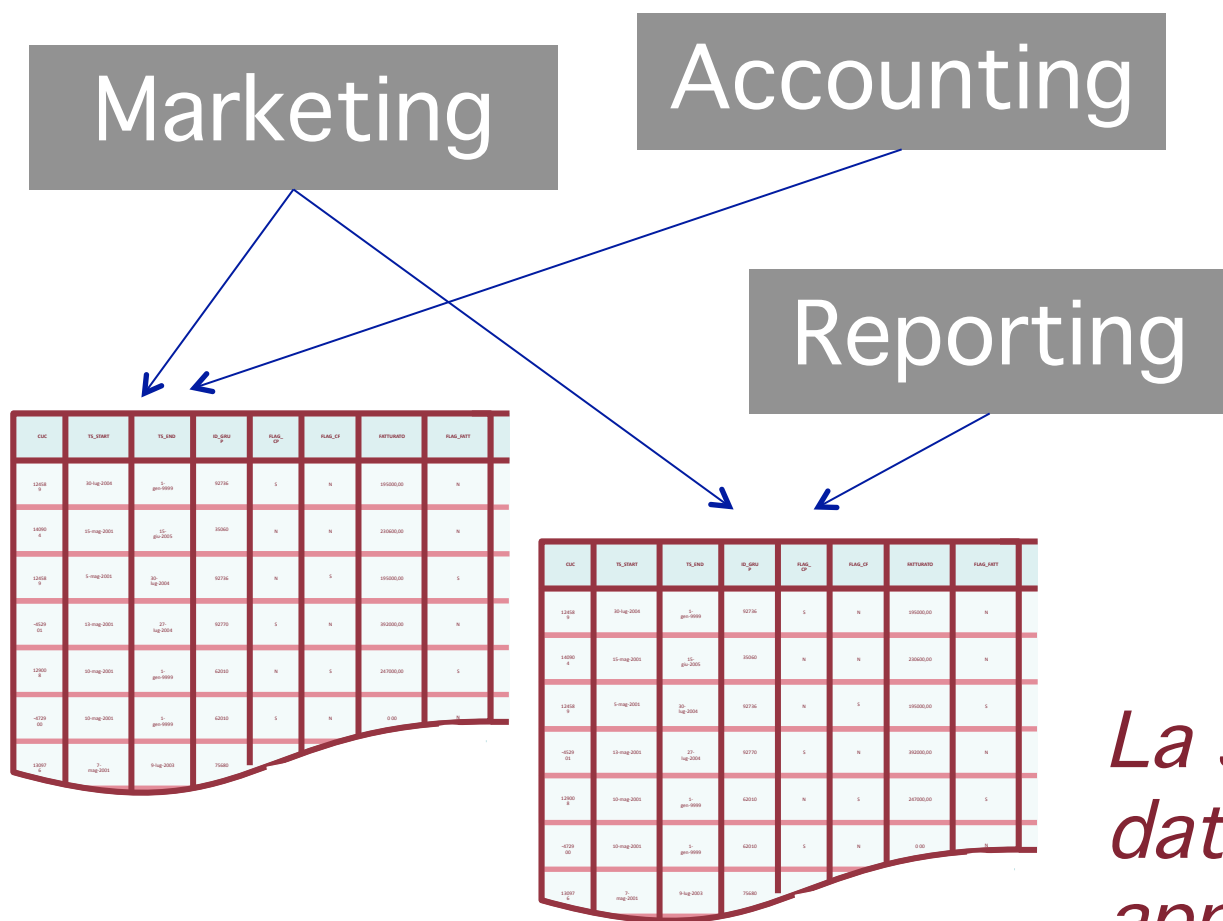
CUC	TS_START	TS_END	ID_GRUP	FLAG_CP	FLAG_CF	FATTURATO	FLAG_FATT	
124589	30-lug-2004	1-gen-9999	92736	S	N	195000,00	N	
140904	15-mag-2001	15-giu-2005	35060	N	N	230600,00	N	
124589	5-mag-2001	30-lug-2004	92736	N	S	195000,00	S	
-452901	13-mag-2001	27-lug-2004	92770	S	N	392000,00	N	
129008	10-mag-2001	1-gen-9999	62010	N	S	247000,00	S	
-472900	10-mag-2001	1-gen-9999	62010	S	N	0 00	N	
130976	7-mag-2001	9-lug-2003	75680					

Complessità dei sistemi informativi

*N indica che il campo
FATTURATO non è valido*

CUC	TS_START	TS_END	ID_GRUP	FLAG_CP	FLAG_CF	FATTURATO	FLAG_FATT	
124589	30-lug-2004	1-gen-9999	92736	S	N	195000,00	N	
140904	15-mag-2001	15-giu-2005	35060	N	N	230600,00	N	
124589	5-mag-2001	30-lug-2004	92736	N	S	195000,00	S	
-452901	13-mag-2001	27-lug-2004	92770	S	N	392000,00	N	
129008	10-mag-2001	1-gen-9999	62010	N	S	247000,00	S	
-472900	10-mag-2001	1-gen-9999	62010	S	N	0 00	N	
130976	7-mag-2001	9-lug-2003	75680					

Complessità dei sistemi informativi



La semantica dei dati è nascosta nelle applicazioni



Aristotele

Soluzione: modellare il sistema

Due contributi di Aristotele (384 a.c.)

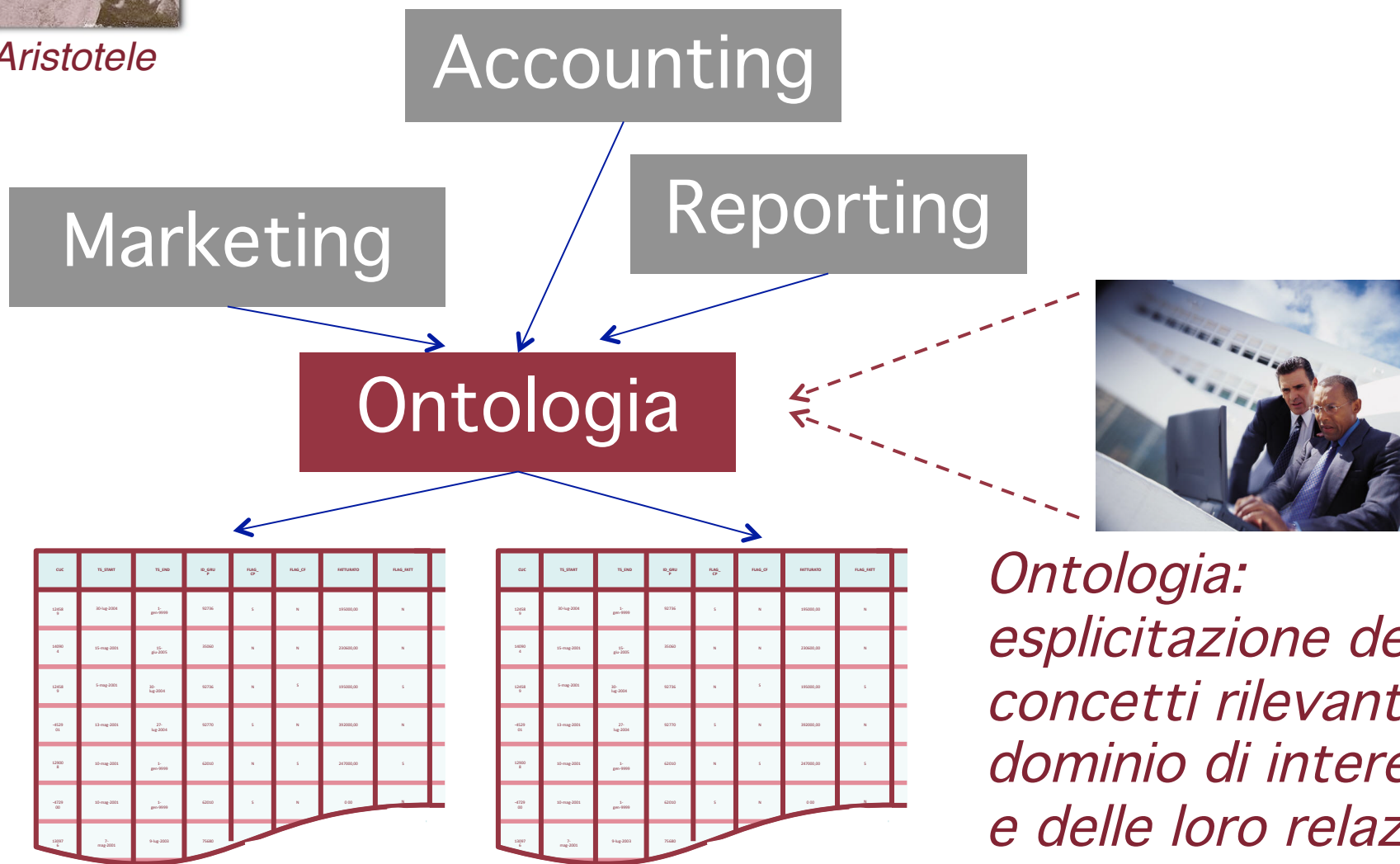
- Ontologia: analisi degli enti e delle loro proprietà

“Delle cose dette, ciascuna significa o *ousía* o una quantità o una qualità o in relazione a qualcosa o un luogo o un tempo. Ed è *ousía* ad esempio «uomo», «cavallo»; una quantità per esempio «due cùbiti»; una qualità per esempio «bianco»; in relazione a qualcosa per esempio «maggiore»; un luogo ad esempio «piazza»; in un tempo ad esempio «ieri».”
- Logica: studio di come argomenti corretti seguono necessariamente da premesse corrette (si veda lo studio sul sillogismo e quello sulla quantificazione)



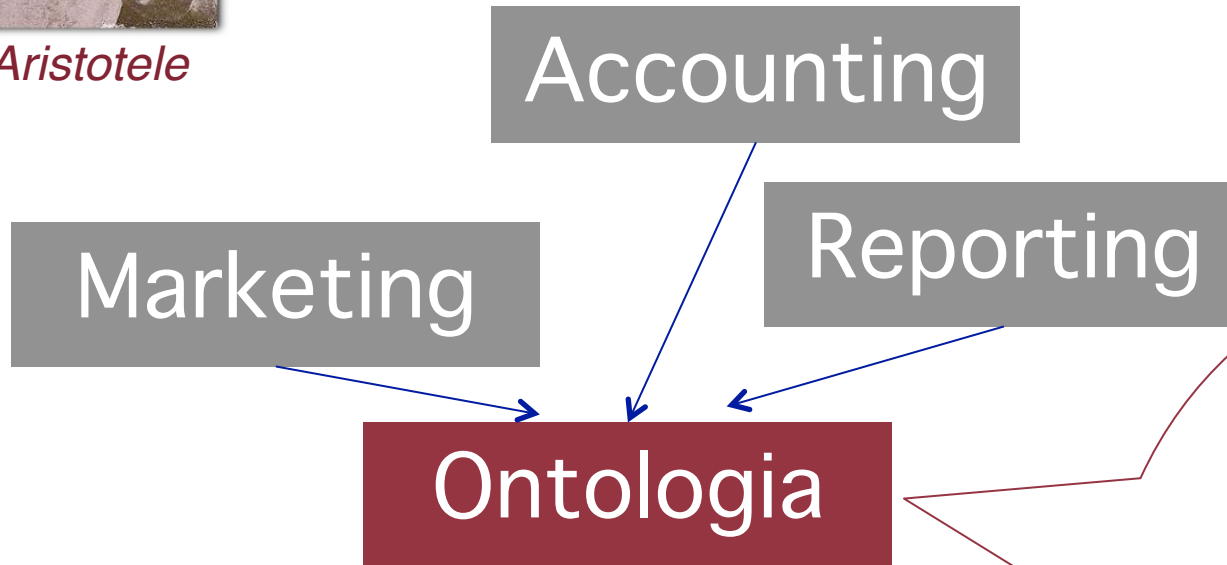
Aristotele

Soluzione: ontologia del dominio





Soluzione: ontologia del dominio



- Un cliente fittizio è una holding di persone fisiche
- Responsabile di un gruppo è la capofila se la capogruppo è sotto monitoraggio
- Un cliente estero non ancora censito è sotto monitoraggio

DEC	TS_DENET	TS_FMS	SL_FMS P	PLAS EP	PLAS_CP	NETSUMED	PLAS_NET
12018 S	30-mag 2004	1- pm-2004	92796	S	N	195000,00	N
14000 S	15-mag 2005	15- pm-2005	20040	N	N	230000,00	N
12018 S	5-mag 2005	20- kg-2004	92796	N	S	195000,00	S
-0129 SD	13-mag 2005	23- kg-2004	92796	S	N	203000,00	N
12000 S	10-mag 2005	5- pm-2005	42000	N	S	247000,00	S
-0129 SD	10-mag 2005	5- pm-2005	42000	S	N	0-00	S
13007 S	11- mag-2004	9-kg-2005	75000				

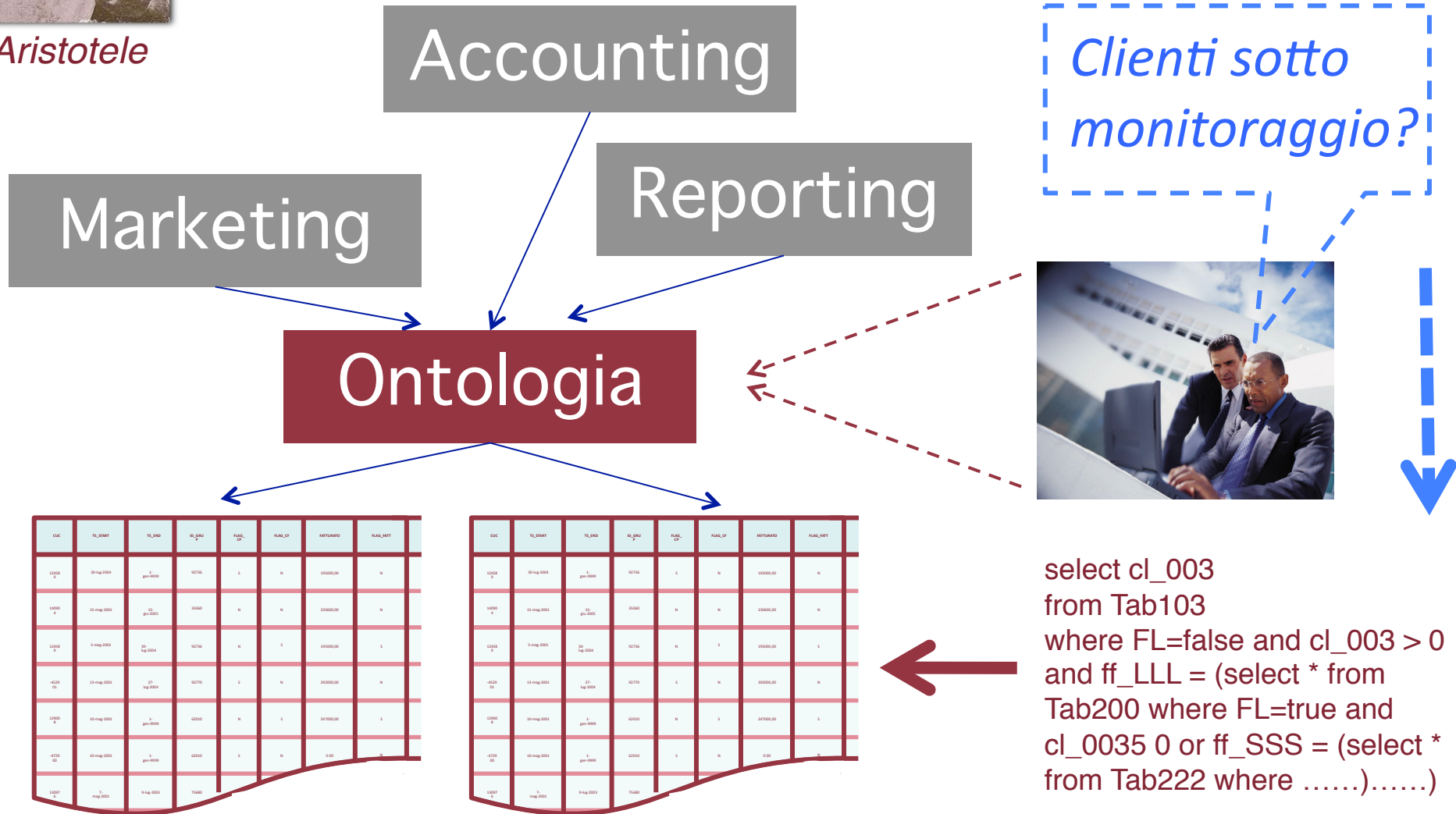
QXZ	TLS_EST1	TLS_EST2	SL_EST %	PLAS_CP	PLAS_CD	NOTUSED	PLAS_NET
1245 S	80 mgp 2004	5 gm 2000	52734	S	N	193000.00	N
14000 S	80 mgp 2004	10 gm 2000	50260	N	N	230000.00	N
1245 S	5 mgp 2003	20 gm 2000	52734	N	S	193000.00	S
4529 S1	13 mgp 2001	20 gm 2000	52730	S	N	350000.00	N
12600 S	80 mgp 2004	5 gm 2000	42000	N	S	247000.00	S
4529 S2	80 mgp 2004	5 gm 2000	42000	S	N	0.00	S
13007 S	5 mgp 2001	0 mgp 2003	71480				





Aristotele

Soluzione: ontologia del dominio





Leibniz

Come si esprime l'ontologia?

L'**idea meravigliosa** di Leibniz (1646):

- creare un alfabeto di simboli che denotino concetti e relazioni tra concetti, e sulla base di questo alfabeto costruire un *linguaggio formale*
- definire un *calcolo simbolico* (*calculus ratiocinator*) per stabilire quali enunciati del linguaggio sono validi, e, più in generale, per individuare le relazioni logiche che si instaurano tra gli enunciati

Logica e [mate/infor]matica

- **Matematica → Logica**

L'algebra viene utilizzata come mezzo per esprimere le leggi del ragionamento (opera di Boole, 1815)

- **Informatica → Logica**

L'informatica può fornire strumenti algoritmici per il ragionamento automatico

- **Logica → Matematica**

La logica viene utilizzata per formalizzare tutta la matematica (programma di Hilbert, 1862)

- **Logica → Informatica**

La logica fornisce gli strumenti per modellare concettualmente un dominio di interesse



La mia ricerca

- **Anni '80**

L'ontologia come prodotto intermedio nella progettazione del sistema informativo

- **Anni '90**

La logica per esprimere l'ontologia: linguaggi logici per rappresentare concetti e loro relazioni e sistemi per effettuare deduzioni

- **Anni '00**

Algoritmi per il ragionamento automatico su ontologie espresse in logica e traduzione di servizi in termini di accesso ai dati



Gödel

Potenzialità e limiti dei sistemi logici

- **Tesi di dottorato di Gödel (1906)**

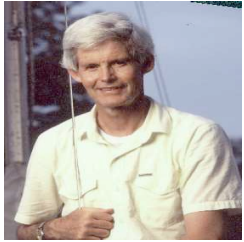
Sistema deduttivo “corretto e completo” per la logica matematica

- **Teorema di incompletezza di Gödel**

In ogni sistema coerente T sufficientemente espressivo da catturare l'aritmetica, esiste una formula valida che non è dimostrabile usando T

- **Teorema di Church (1903) - Turing (1912)**

Non esiste alcun procedimento meccanico che, avendo in input una formula della logica del primo ordine, sia in grado di stabilire se essa sia un teorema oppure no



Cook

Limiti pratici del calcolo

	$n = 10$	$n = 20$	$n = 50$	$n = 100$	$n = 1000$
n	$10\mu s$	$20\mu s$	$50\mu s$	$100\mu s$	$1\ ms$
$n \cdot \log_2 n$	$33,2\mu s$	$86,4\mu s$	$0,28ms$	$0,6ms$	$9,9ms$
n^2	$0,1ms$	$0,4ms$	$2,5ms$	$10ms$	$1s$
n^3	$1ms$	$8ms$	$125ms$	$1s$	$16,6mn$
2^n	$1ms$	$1s$	$35,7a$	$\approx 10^{14}c$	∞

Complessità computazionale degli algoritmi e tempo di esecuzione dei programmi

Nota: il migliore algoritmo per verificare se una formula Booleana è valida richiede tempo esponenziale rispetto alla dimensione della formula

Contributi del mio gruppo

Entra in gioco l'ingegneria

- Progettazione di linguaggi logici che offrano un compromesso ottimale tra
 - potere espressivo nella descrizione di ontologie
 - complessità computazionale degli algoritmi
- Realizzazione di un sistema per la rappresentazione della conoscenza che consente di
 - definire ontologie,
 - collegare, mediante la nozione di mapping, l'ontologia alle risorse computazionali,
 - rispondere, mediante le tecniche di deduzione automatica, alle interrogazioni in modo efficiente



Conclusioni

- L'informatica non è una disciplina di mero contenuto tecnologico
- Una parte rilevante dell'informatica fonda le sue radici nella logica, e dal secolo scorso gioca un ruolo importante nello sviluppo della filosofia e della matematica
- *“Computer Science is the continuation of logics by other means”* (G. Gottlob, 1956)



La logica nella nostra società

Dalla prima pagina di uno dei più importanti quotidiani nazionali:

“C'è ancora chi crede che il rigore nei conti pubblici contraddica lo sviluppo dell'economia, ma anni di storia italiana con spesa facile e stagnazione economica ci insegnano esattamente il contrario.”

Si tratta di un noto errore logico: la **fallacia del conseguente**. Per confutare

rigore *implica* non sviluppo

si afferma che possono convivere

non rigore (spesa facile) e non sviluppo (stagnazione economica)
quando, al contrario, l'unico modo sarebbe affermare che possono convivere

rigore e sviluppo



Proposta per la scuola: due “I” ed una “L”

Le tre “I”	→	Due “I” ed una “L”
Impresa	→	+ Italiano
Inglese	→	Inglese
Informatica	→	Logica



I miei studenti ed il mio gruppo

Francesco
Donini



Marco
Cadoli



Andrea
Schaerf



Zoran Majkic



Andrea
Cali



Daniela Berardi



ex-studenti



Giuseppe De
Giacomo



Diego
Calvanese



Domenico
Lembo



Antonella Poggi

il gruppo



Riccardo
Rosati

Marco
Ruzzi



studenti attuali



Fabio Savo



Floriana
Di Pinto

All the people we used to know they're an illusion to me now... Some are mathematicians some are carpenter's wives. Don't know how it all got started, I don't know what they're doin' with their lives. But me, I'm still on the road headin' for another joint. We always did feel the same, we just saw it from a different point of view. Tangled up in blue. **Bob Dylan**



Marco Cadoli (5/12/1965 – 21/11/2006)



Marco entrò nella mia stanza dell'ufficio in via Buonarroti poco più di ventidue anni fa per chiedermi la tesi, con la sua educazione, la sua figura elegante, il suo sguardo limpido, e la sua intelligenza prorompente...

Marco ha fornito contributi di altissimo livello scientifico nelle basi di dati e nell'intelligenza artificiale, coniugando l'amore per la logica con la passione per l'informatica e l'ingegneria

ed è a lui che dedico questa lezione.

Alcuni lavori rilevanti

Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini: *What can Knowledge Representation do for Semi-Structured Data?* AAAI/IAAI 1998: 205-210

Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini: *Representing and Reasoning on XML Documents: A Description Logic Approach*. J. Log. Comput. 9(3): 295-318 (1999)

Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini, Moshe Y. Vardi: *View-Based Query Processing and Constraint Satisfaction*. LICS 2000: 361-371

Diego Calvanese, Giuseppe De Giacomo, Maurizio Lenzerini, Daniele Nardi: *Reasoning in Expressive Description Logics*. Handbook of Automated Reasoning 2001: 1581-1634

Maurizio Lenzerini: *Data Integration: A Theoretical Perspective*. PODS 2002: 233-246

Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Riccardo Rosati: *Tractable Reasoning and Efficient Query Answering in Description Logics: The DL-Lite Family*. J. Autom. Reasoning 39(3): 385-429 (2007)

Diego Calvanese, Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Riccardo Rosati: Inconsistency tolerance in P2P data integration: An epistemic logic approach. Inf. Syst. 33(4-5): 360-384 (2008)

