

Laurea Specialistica in Informatica a.a. 2006-2007

Interazione Uomo-Macchina II:

Interfacce Intelligenti

Nicole Novielli e Fiorella de Rosis

novielli_at_di.uniba.it

Introduzione

Prima parte: Formalizzazione e Ragionamento

- 1.1. Ragionamento logico:
 - Formalizzazione
 - Risoluzione
- 1.2. Ragionamento incerto
 - Reti Causali Probabilistiche
 - Reti dinamiche
 - Apprendimento di Reti

Seconda parte: Modelli di Utente

- 2.1. Modelli logici
- 2.2. Modelli con incertezza

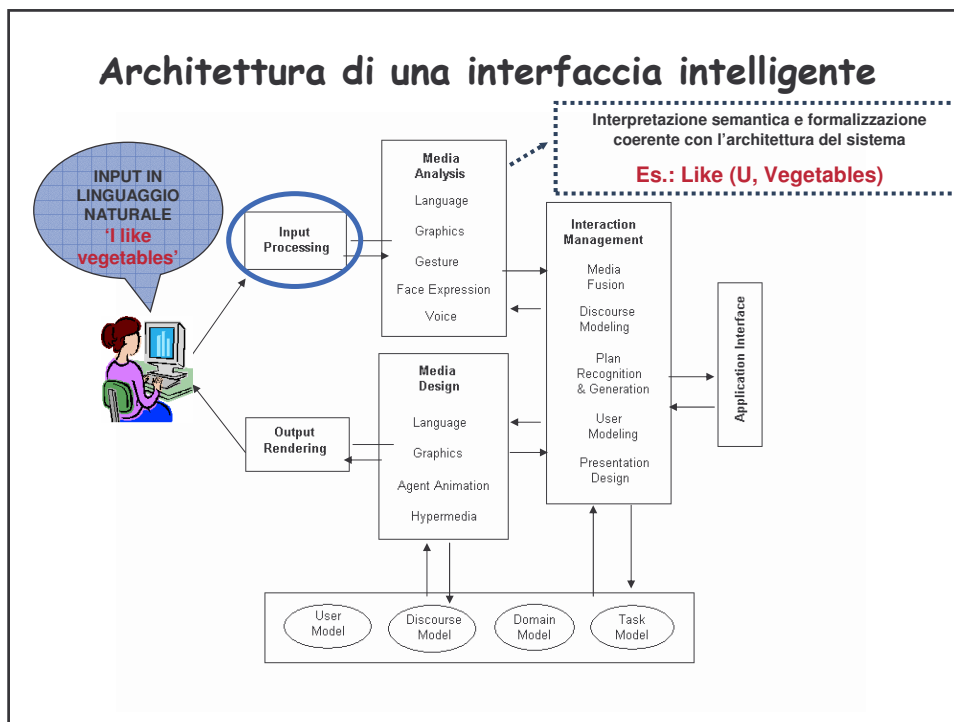
Terza parte: Interazione in linguaggio naturale

- 3.1. Generazione di messaggi
 - Introduzione
 - Teorie
 - Metodi

3.2. *Comprensione di messaggi*

Quarta parte: Simulazione di dialoghi

Architettura di una interfaccia intelligente



Semantica

L'**obiettivo** del riconoscimento in un sistema di dialogo è quello di **formalizzare le frasi in linguaggio naturale tramite un linguaggio ad hoc** (ad es. linguaggio logico), allo scopo di 'ragionare' su di esse.

The dog ate.

Ate(dog)

A man saw a cat with a telescope

Saw(man, CatWithTelescope) gatto dotato di telescopio ☺

Oppure: Saw(man, Cat, UsingTelescope) usando un telescopio
nelle due interpretazioni della frase che abbiamo visto.

Jack downloaded the computer game

Downloaded(Jack, ComputerGame)

**Attenzione:
per comprendere questa
esercitazione è necessario fare
riferimento alla teoria, illustrata
nella lezione 'NLU parte
seconda'**

Esercitazione 2b

**Riconoscimento di atti comunicativi tramite LSA applicata a move
utente in linguaggio naturale**

Un esempio: AutoTutor (Graesser et Al.)¹

SCENARIO: Sistema completamente automatizzato per la gestione di dialoghi di tutoring con studenti

Gli argomenti che il sistema è in grado di gestire riguardano *hardware, sistemi operativi e Internet*, nell'ambito di un corso introduttivo

L'interazione si basa sulla metafora dell'agente conversazionale: un tutor virtuale, personificato da una *talking head*, dialoga con gli studenti proponendo loro quesiti su argomenti random e valutando la correttezza nelle risposte ricevute

OBIETTIVO: il tutor virtuale deve

- simulare un atteggiamento collaborativo, nella costruzione della risposta alla domanda,
- fornire un feedback immediato in caso di risposte (parzialmente) corrette o meno

¹ Graesser et Al.: Using Latent Semantic Analysis to Evaluate the Contributions of Students in AutoTutor

AutoTutor - come avviene l'interazione

- a. AutoTutor esegue la propria *move* (mossa di dialogo), pronunciandola grazie ad uno speech synthesizer ed adottando un'intonazione ed un'espressione del volto appropriate

Le domande di AutoTutor sono poste in modo da richiedere una risposta ragionata piuttosto che in forma schematica (vengono evitate ad es. domande del tipo 'fill-in the blank')

- b. Parte un sotto-dialogo in cui, iterativamente si ripetono i seguenti passi, sino a che la soluzione alla risposta non viene raggiunta in modo completo:
1. L'utente digita le proprie risposte, anche parziali, utilizzando una tastiera
 2. AutoTutor valuta la correttezza della risposta dello studente, fornisce il feedback (positivo o negativo) associato, produce suggerimenti in caso di risposte completamente o parzialmente errate

AutoTutor - tipi di move eseguibili dal sistema

- (1) Positive immediate feedback. "That's right" "Yeah"
- (2) Neutral immediate feedback. "Okay" "Uh-huh"
- (3) Negative immediate feedback. "Not quite" "No"
- (4) Pumping for more information. "Uh-huh" "What else"
- (5) Prompting for specific information. "The primary memories of the CPU are ROM and _____"
- (6) Hinting. "The hard disk can be used for storage" or "What about the hard disk?"
- (7) Elaborating. "CD ROM is another storage medium. "
- (8) Splicing in correct content after a student error. This is a correction.
- (9) Questioning. "So once again, what is the function of a CPU?"
- (10) Summarizing. "So to recap," <succinct recap of answer to question>

AutoTutor

Definizione di un range di possibili risposte per ogni argomento

Ho bisogno di una base di conoscenza cui fare riferimento, formalizzata in qualche modo...

Es.: ad ogni domanda associo un set di risposte parziali positive (A_i)...

QUESTION: Why do computers need operating systems?

GOOD-ANSWER-1: The operating system helps load application programs.

GOOD-ANSWER-2: The operating system coordinates communications between the software and the peripherals.

GOOD-ANSWER-3: The operating system allows communication between the user and the hardware.

GOOD-ANSWER-4: The operating system helps the computer hardware run efficiently.

... e di risposte parziali negative (B_i)

La risposta completa ideale consiste in un set di N risposte parziali corrette

$\{A_1, A_2, \dots, A_n\}$

Come formalizzo questa conoscenza in modo che sia confrontabile con le risposte degli studenti?

AutoTutor - LSA per rappresentare la conoscenza

Corpus di apprendimento: 2 libri + 30 articoli relativi agli argomenti trattati

Lessico di circa 10.000 parole inglesi, con informazioni sulla POS

(ad es.: 'program' può essere un nome, un verbo o un aggettivo)

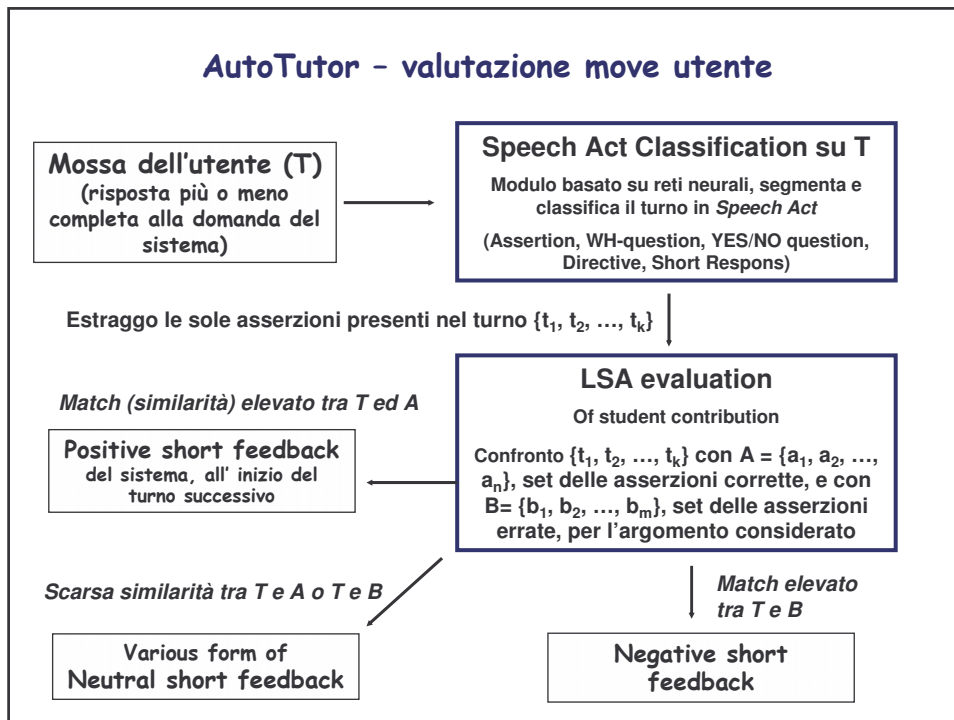
La conoscenza è formalizzata utilizzando una matrice documenti x termini in cui l'unità di testo (**documento**) coincide con un **paragrafo** di testo.

Ogni paragrafo corrisponde ad una risposta parziale positiva per un dato argomento

Gli elementi della matrice corrispondono ai valori di frequenza dei termini nei documenti

Tramite SVD si ottiene una rappresentazione a 200 dimensioni

AutoTutor - valutazione move utente



AutoTutor - valutazione LSA

Per la valutazione dell'accuratezza del riconoscimento tramite LSA è stato raccolto un testing corpus di 192 risposte di studenti a domande riguardanti gli argomenti che AutoTutor è in grado di trattare

Agli studenti non è stato chiesto di interagire col sistema ma solo di rispondere in modo approfondito e ragionato, in linguaggio naturale

Quattro 'rater' indipendenti (2 mediamente esperti, 2 esperti) hanno annotato il corpus, giudicando, per ogni risposta studente, la percentuale di *asserzioni* in accordo con la risposta completa ideale

(es.: se su 4 asserzioni ritrovate in una risposta dell'utente, per 3 si verifica 'matching' con uno degli aspetti della risposta completa, allora viene assegnato uno score di .75 come valutazione di correttezza della risposta)

Lo score LSA viene calcolato allo stesso modo (sulle asserzioni)

Correlazione tra LSA e annotazione

Analisi di correlazione tra le due variabili (score annotatori e score LSA) (.49)

non statisticamente significativa la differenza con .51, valore di correlazione tra l'annotazione dei due annotatori mediamente esperti

decisamente inferiore al coefficiente di correlazione tra l'annotazione di due esperti (.78)

Sembra quindi che il riconoscimento tramite LSA produca una valutazione della correttezza in linea con le performance di esperti di livello intermedio piuttosto che elevato

Tuttavia, la maggior parte dei tutor 'reali' ha un livello di abilità intermedio, quindi tale risultato è da intendersi come positivo (LSA simula in modo eccellente questa categoria di tutor)

AutoTutor - ricapitolando

Obiettivo: assistere il processo di apprendimento tramite un simulatore di dialogo completamente automatizzato

Per fare ciò, è necessario *adattare* lo stile (short feedback all'inizio del turno) e il contenuto (suggerimenti in caso di risposte sbagliate) del dialogo al livello di comprensione mostrato dagli studenti

LSA per la rappresentazione della conoscenza nel dominio e la valutazione delle move utente in input

Fasi cruciali dello studio (da fare in modo ragionato ed in accordo con l'obiettivo finale):

- definizione del lessico (10.000 parole con info su POS)
- definizione dell'unità di testo (paragrafo)
- definizione della dimensionalità cui ridurre la matrice tramite SVD (200)
- validazione del modello (in questo caso tramite analisi di correlazione)

A questo punto il modello è pronto per essere utilizzato (classificazione di nuovo input)

Cosa fare per l'esercitazione - scenario

SCENARIO: Simulatore di dialoghi persuasivi con un ECA, nel campo dell'*healthy eating*

Vogliamo che il simulatore sia in grado di adattare il proprio stile di interazione, i contenuti delle proprie mosse e la strategia persuasiva all'attitudine mostrata dall'utente nel corso del dialogo (reazioni ai suggerimenti dell'ECA)

Dopo una fase iniziale di raccolta di informazioni sull'utente, il sistema ritiene di dover convincere l'utente a consumare più vegetali. Effettua quindi il seguente tentativo di persuasione:

S: *'Diverse ricerche a livello internazionale hanno dimostrato l'importanza di una **dieta ricca di vegetali e di frutta**. Idealmente, bisognerebbe mangiare almeno cinque porzioni di frutta e verdura al giorno. La porzione consigliata risulta essere di circa 80 grammi: per esempio una mela, due o tre cucchiari di verdure congelate o un bicchiere di succo di frutta. In generale, si consiglia un consumo quotidiano di una porzione di verdura cotta e una di verdura cruda, e di due o tre porzioni di frutta fresca'*

S: *'Diverse ricerche a livello internazionale hanno dimostrato l'importanza di una **dieta ricca di vegetali e di frutta**. Idealmente, bisognerebbe mangiare almeno cinque porzioni di frutta e verdura al giorno. La porzione consigliata risulta essere di circa 80 grammi: per esempio una mela, due o tre cucchiari di verdure congelate o un bicchiere di succo di frutta. In generale, si consiglia un consumo quotidiano di una porzione di verdura cotta e una di verdura cruda, e di due o tre porzioni di frutta fresca'*



Gli utenti possono **reagire in maniera differente** al suggerimento:

U: 'E come lo sai?'	<i>U chiede all'ECA di giustificare la propria asserzione</i>	ASKJUSTIFY
U: 'lo mangio un sacco di frutta'	<i>U informa l'ECA circa una propria abitudine</i>	INFORM
U: 'Hai torto, non sono d'accordo'	<i>U dichiara di essere in disaccordo con il sistema</i>	DISCONFIRM
U: 'Un bel peccato di gola è meglio di una dieta bilanciata!'	<i>U rifiuta il suggerimento</i>	REJECT
U: 'Così tante? Ho sentito tesi contrastanti al proposito'	<i>U si dichiara non persuaso</i>	CHALLENGE

Riconoscimento dell'atto comunicativo

Obiettivo dell'esercitazione: applicare l'LSA per il riconoscimento dei differenti atti comunicativi che possiamo individuare nelle reazioni dell'utente ai tentativi di persuasione da parte del sistema

Comm. Act	Purpose	Target	Examples
UNCERTAIN	R nods without expressing any clear opinion	All	mmm
ASKIF	R ask the truth value of a fact	All	Do you think my diet is correct?
ASKINFO	R asks for more information about some topic	All	How could I substitute fruits?
ASKJUSTIFY	R asks the system to justify its statement		And how do you know it?
INFORM	R provides some evidence about his/her attitudes or behaviour	Likes Implies CanDo Evidence	I eat meat, fish, vegetables, lots of fruits...
CONFIRM	R declares to agree with the evidence provided by the system	Likes Implies CanDo Evidence	Right, I agree
DISCONFIRM	R declares to disagree with the evidence provided by the system	Likes Implies CanDo Evidence	No, you're wrong. I don't agree
IREBUTTAL	R presents an exception that falsifies the system argument	Likes Implies CanDo	I love unbridled life, with light aversion towards healthy food.
OBJECT	R argues about the truth value of a premise of the suggestion	Likes Implies CanDo	Are you joking? So you mean I have to bring a fruit bag with me, at work?
ACCEPT	R declares to agree with the received suggestion	Sh-Do	Understood! So I should try to do it?
COMMIT	R commits him/herself to apply the received suggestion	Sh-Do	Ok, I will do it
CHALLENGE	R declares to not be persuaded by the suggestion	Sh-Do	So many portions of fruits? I've heard contrary theories on this topic
REJECT	R refuses the suggestion	Sh-Do	But... a sin of gluttony is better than any healthy and balanced diet!
S-REBUTTAL	R presents an exception that falsifies the suggestion	Sh-Do	I don't want to avoid sweets at all

L'esercitazione 2b prevede

- annotazione di un corpus (in italiano) degli atti comunicativi individuabili nelle reazioni degli utenti ai suggerimenti dell'ECA
Le modalità dell'annotazione e il linguaggio di markup da utilizzare saranno descritti in dettaglio ai gruppi coinvolti in questa esercitazione (2b)
- rappresentazione del corpus annotato tramite LSA previa opportuna definizione del dizionario dei termini
Ad es.: definizione di categorie linguistiche basate sulla semantica di singoli termini o sequenze di parole (vedi esempio self introduction slide. n. 16)
- sviluppo (in Java, C o altro linguaggio a scelta) dei moduli necessari per
 - Costruire la matrice termini x documenti a partire dal corpus annotato e dal dizionario dei termini definito
 - Classificare in funzione dell'atto comunicativo i turni utente in input
- Valutazione della bontà del riconoscimento (accuratezza) su un testing corpus