



Elaborazione del Linguaggio Naturale 2007/2008

INTERAZIONE UOMO MACCHINA

Prof. Amedeo Cappelli

Studentessa: Alice Casti matr. 416470

Sommario

Elaborazione del linguaggio naturale nei robot

- *Interazione uomo robot*
- *Schemi di comunicazione*
- *Approccio cognitivo per la comprensione del linguaggio naturale*
- *Miglioramento dell'interazione uomo robot*

Utilizzo del linguaggio naturale nella domotica

- *Domotica*
- *Un'applicazione domotica: Genio*

ELABORAZIONE DEL LINGUAGGIO NATURALE NEI ROBOT

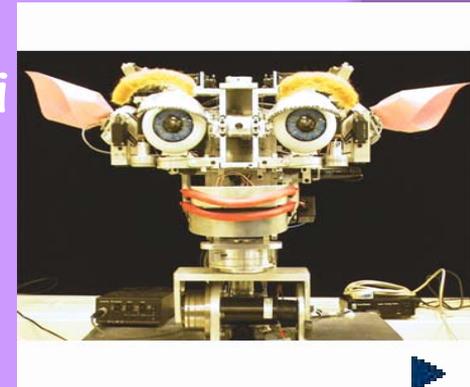


Interazione Uomo-Robot

- ⦿ *Prima del 2004 non si parlava di interazione uomo robot, in quanto i robot erano visti principalmente come delle macchine industriali programmabili che apportavano pochi cambiamenti nella HCI.*
- ⦿ *Differenze di percezione della HRI rispetto alla HCI:*
 1. *I robot autonomi sono modellati dalla mente delle persone con aspetto antropomorfo*
 2. *I robot spesso sono completamente mobili*
 3. *I robot autonomi prendono delle decisioni, apprendono qualcosa sia su loro stessi sia sul mondo che li circonda*
- ⦿ *L'ampliamento delle ricerche in campo di interazione uomo-robot permise, e permette tuttora, ad utenti inesperti di interagire con i robot in modo intuitivo e naturale.*

Interazione Uomo-Robot

- Nell'interazione uomo-robot si considerano maggiormente i **robot sociali**, che si dividono in 2 categorie:
 1. Basati sull'ispirazione biologica, ossia usano modelli di percezione e interazione umana per simulare un'intelligenza sociale simile a quella che si trova nelle creature viventi (Cog, Kismet)
 2. Progettati per una funzione, mostrano la loro competenza sociale solo in risposta al comportamento umano (Nursebot, CERO)



Interazione Uomo-Robot

- ⊙ È basata su:
 - Tipi di interfacce utilizzate
 - Modalità di dialogo
 - Interazioni sociali
- ⊙ *Soffre di problemi causati dal contesto fisico*
 - Il robot necessita di ottenere dati da più sensori per combinarli, per ottenere un numero maggiore di informazioni a riguardo dell'ambiente
 - Il robot necessita di combinare la rappresentazione di alto livello (come pianificazione e dialogo) con i dati dei sensori

Come comunicare?

- Verbali: Parlato
- Non verbali
 - Tracciamento dello sguardo
 - Gestì
 - Espressioni facciali
 - Cinesica
 - Prossemica
 - Aptica
 - Interfaccia Cervello-Computer



Modalità di dialogo

⊙ Teleoperazione:

- *L'utente impartisce comandi, o specifica un compito*
“Chiudi la porta”
“Girati di 90° a destra”
- *L'uomo monitora l'esecuzione*

⊙ Lavoro Cooperativo:

- *L'utente fornisce spiegazioni per correggere un errore*
- *L'utente fornisce al robot dettagli sull'ambiente in modo da tenere aggiornata la sua rappresentazione interna*
- *Il robot controlla l'esecuzione come parte di una squadra di lavoro*

SCHEMI DI COMUNICAZIONE

1) La comunicazione tramite il parlato

- Affinché le macchine possano comunicare con gli umani tramite il linguaggio naturale, necessitano di usare dei metodi di mappatura dei simboli con le corrispondenti informazioni sia nel contesto linguistico sia in quello non-linguistico.
- Il sistema di simboli può variare

Esempio:

Utente: “Metti i piatti nella credenza”



Dimensione e forma piatti

Dimensione credenza

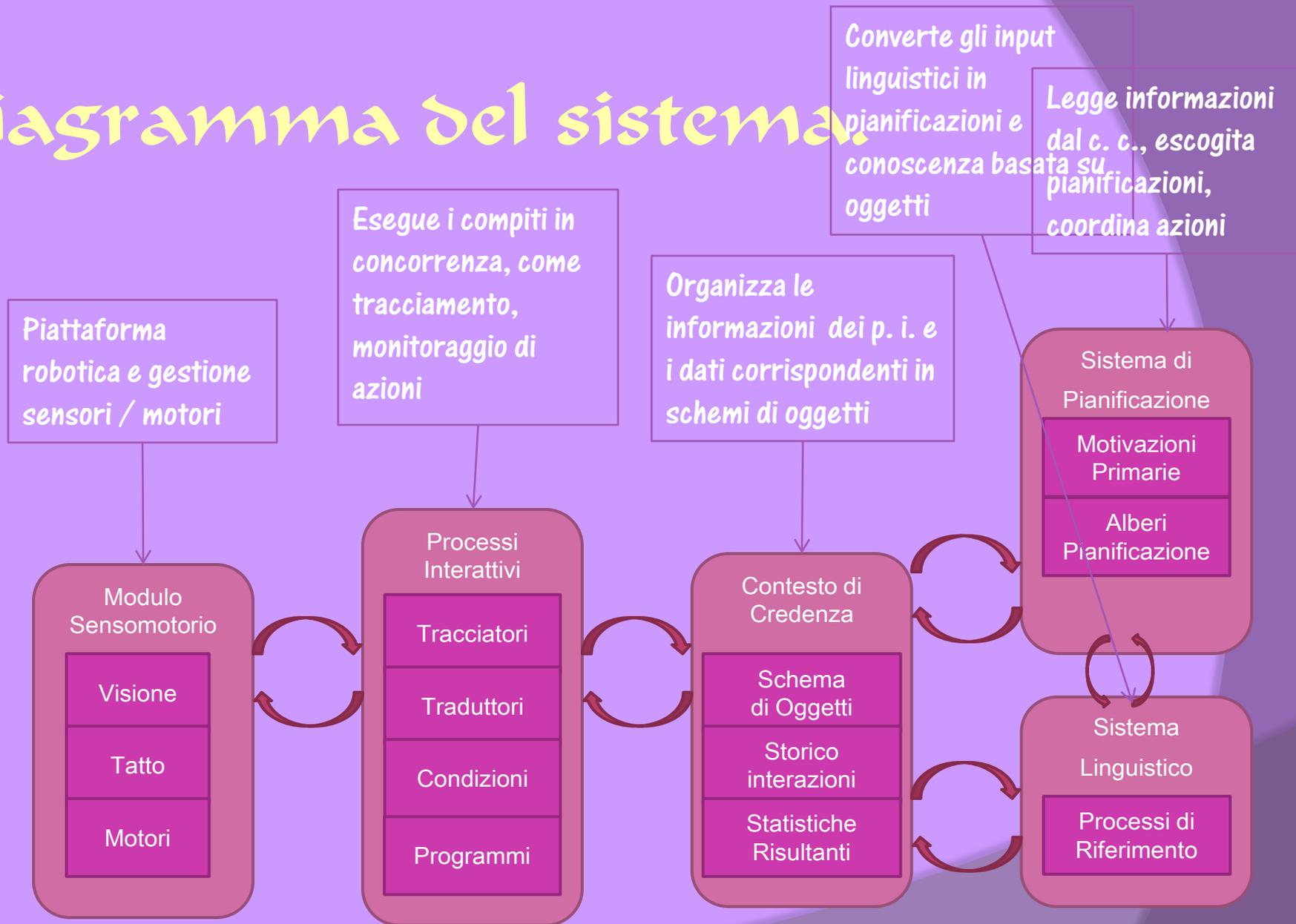
1) La comunicazione tramite il parlato

- ⊙ L'introduzione dell'utilizzo del linguaggio naturale in un robot non compromette la reattività dello stesso se il linguaggio è usato insieme ad altri processi interattivi, come sistemi di rilevamento delle collisioni e tracciamento visuale.
- ⊙ I dati dei processi interattivi rappresentano gli oggetti fisici di un ambiente in cui il robot opera. I dati sono organizzati in **schemi di oggetti**, per ottenere una più veloce integrazione tra dati provenienti da più sorgenti.

Gli schemi di oggetti

- ⦿ Gli schemi di oggetti sono utilizzati per coordinare gli aspetti discreti del linguaggio naturale e gli aspetti continui dati dall'ambiente dinamico sensomotorio.
- ⦿ Gli schemi di oggetti agiscono come delle entità discrete per gli scopi del linguaggio e della pianificazione. Organizzando i processi in modo continuo, è possibile ordinare i dati sensori in arrivo, in base alla rapidità di interazione necessaria, sia con il linguaggio sia con la pianificazione.
- ⦿ Gli schemi di oggetti forniscono un collegamento tra i riferimenti verbali discreti e le azioni motorie risultanti nel mondo continuo dei sensori.

Diagramma del sistema.

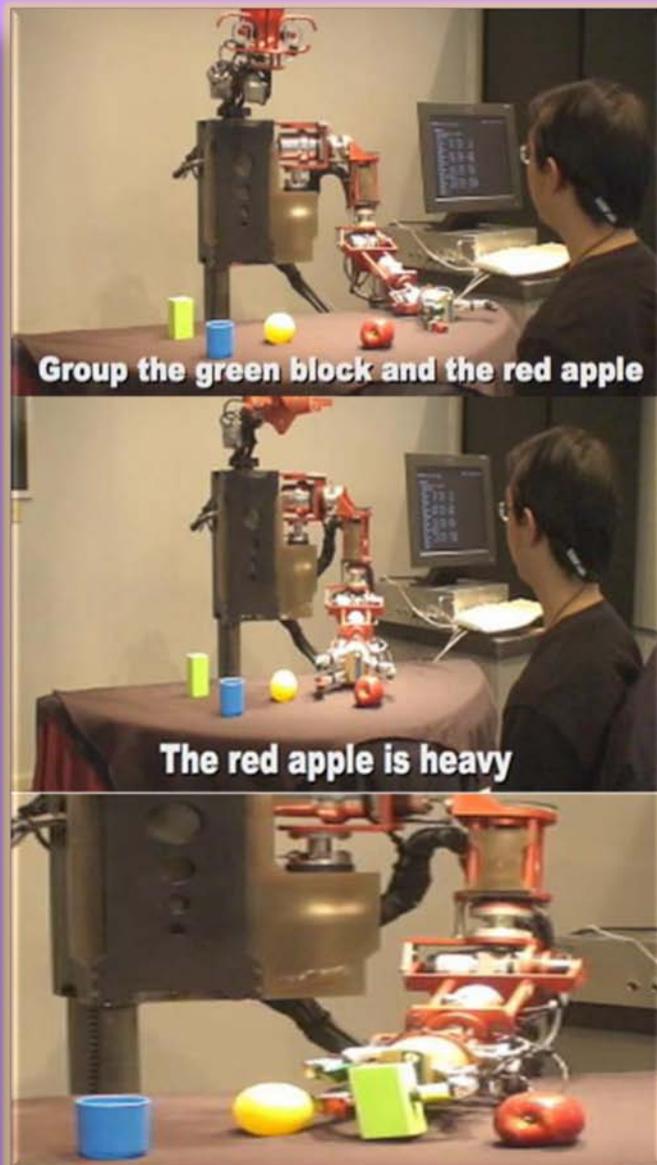


Sistema linguistico

- Una metodologia per categorizzare gli oggetti è basata sugli attributi che definiscono l'oggetto stesso (colore, peso, forma).
- Il sistema linguistico gestisce tipi di input linguistico quali:
 - Input Descrittivo, presenta maggiori informazioni di un dato oggetto, che andranno a far parte degli attributi
Esempio “La palla rossa è pesante”
 - Input Direttivo definisce una richiesta da parte di un essere umano
Esempio “Solleva la palla rossa”
 - Input Correttivo indica la modifica di una precedente richiesta da parte di un essere umano
Esempio “No... la palla verde”

Sistema linguistico

- ◎ Il sistema gestisce l'input del parlato a diversi livelli.
 1. Il riconoscitore del parlato converte il flusso audio in un reticolo delle probabili parole, di cui fa il parsing tramite un parser Earley per instanziare azioni e pianificazioni del sistema.
 2. Il processo di riferimento si occupa di individuare le corrispondenze tra gli schemi di oggetti e token nominali e, cercare nel contesto di credenza, gli schemi di oggetti con attributi coincidenti
 3. Il processo di riferimento collega lo schema di oggetti identificato con il predicato verbale a cui si riferiva l'oggetto.



Esempio

L'interlocutore umano chiede di mettere assieme mela rossa e blocco verde.

Il robot può agire spostando o l'uno o l'altro, sceglie di spostare la mela.

L'interlocutore umano aggiunge "La mela è pesante" quindi il robot modifica la sua precedente scelta.

2) La comunicazione HRI integrata con il sistema aptico

- ⊙ Usata soprattutto nell'ambito di lavoro collaborativo tra uomini e robot.
- ⊙ Quando due agenti lavorano insieme alla manipolazione di oggetti è importante che essi abbiano un modo per identificare l'entità da usare tra quelle presenti nel dominio disponibile: **generazione di espressioni di riferimento - GRE** .
- ⊙ Rappresentano tipologie di dialogo multimodali.

2) La comunicazione HRI integrata con il sistema aptico

- ⊙ *Le espressioni di riferimento definiscono dei modelli di accessibilità*
 - *Accessibilità bassa: entità appena introdotte nel discorso*
 - *Maggiore accessibilità: entità che sono centro dell'attenzione causano un'elaborazione delle corrispondenti espressioni di riferimento*
- ⊙ *Possono includere anche riferimenti del contesto visuale e azioni di riferimento aptico.*

Esempi di riferimenti aptico

- *Monta questo (→ indicando le ali dell'aereo) alla fusoliera*
- *Ora ci serve questo (→ afferrando un cubo verde)*

Come funziona?

ALGORITMO INCREMENTALE

[Dale e Reiter 1995]

Produce un'espressione di riferimento per l'oggetto obiettivo, selezionando un insieme di attributi che lo definiscono.

Progressivamente sceglie come attributo dell'oggetto quello che almeno un oggetto del dominio non possiede.

Ripete la procedura finché non ha definito l'oggetto obiettivo della ricerca.

Esempio:

1. Grande, rosso, pesce a righe

2. Piccolo, verde, pesce a righe

3. Grande, rosso, insetto a righe

4. Minuscolo, rosso, insetto a pallini

5. Piccolo, verde, pesce a pallini

{Tipo, Colore, Dimensione, Motivo}

3) Interfaccia Cervello-Computer

- ⊙ Rappresenta un tentativo di fornire una nuova modalità di comunicazione accrescitiva per pazienti umani che soffrono di gravi problemi neuromuscolari.
- ⊙ L'obiettivo è dare a questi utenti, che possono essere completamente paralizzati, una comunicazione di base e un controllo delle loro capacità così che loro possano esprimere i loro desideri o addirittura possano comandare programmi di elaborazione testi, bracci robotici.
- ⊙ Queste interfacce determinano le intenzioni dell'utente tramite i segnali elettrici del cervello registrati sullo scalpo (EEG) oppure da elettrodi impiantati chirurgicamente sulla superficie corticale o nel cervello.
- ⊙ I segnali vengono trasformati in comandi da far eseguire al computer o ad altri dispositivi.

Il progetto ADAPT

COMPRENSIONE DEL LINGUAGGIO NATURALE

Cosa è ADAPT... Active Perception

for Thought

- È un progetto nato dalla collaborazione di 3 gruppi universitari di ricerca, mettendo insieme conoscenze di linguistica, robotica ed intelligenza artificiale.
- Ha lo scopo di costruire un'architettura robotica di tipo cognitivo che possa percepire e controllare qualcosa in tempo reale. Deve possedere le seguenti qualità:
 - essere in grado di lavorare in ambienti non strutturati;
 - abilità di pianificazione e risoluzione dei problemi;
 - essere in grado di imparare dall'esperienza;
 - essere dotato di metodi per la comprensione ed uso del linguaggio naturale, generazione del parlato.

Architettura di ADAPT

⊙ È composta da 2 componenti:

- ▶ Linguaggio chiamato **Robot Schema**, *RS*, usato per rappresentare il mondo come una rete di schemi concorrenti. Il linguaggio *RS* è un linguaggio con un modello di elaborazione concorrente basato sulla semantica di reti di automi di porta. Un automa di porta è un automa a stati finiti dotato di un insieme di porte di comunicazione sincrone.

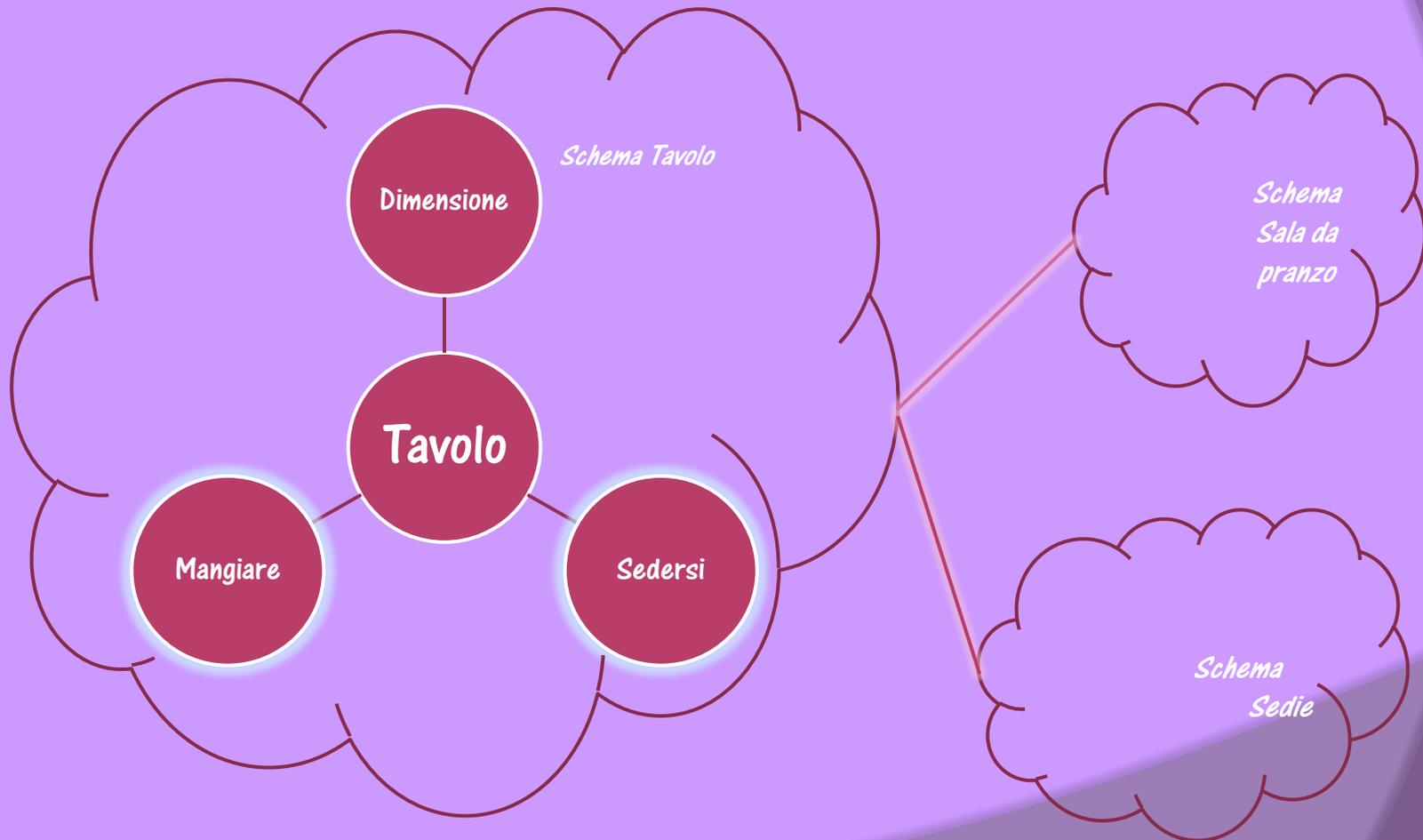
Costruisce una rete di schemi senso-motori per modellare la dinamica del robot e dell'ambiente circostante. Gli schemi sono una rete organizzata di parole, fatti, percezioni su degli aspetti del mondo.

Architettura di ADAPT(2)

- ▶ Ad ogni schema è associato un automa di porta, il quale definisce la semantica formale per il corrispondente schema. Le varie componenti degli automi sono così organizzate:
 - Stati -> semantica per relazioni e predicati (aggettivi, avverbi)
 - Transizioni -> semantica per i verbi

**Gli automi definiscono le basi
per l'uso del linguaggio naturale**

Esempio di schema



Architettura di ADAPT(3)

- ▶ Architettura cognitiva chiamata **SOAR**, usata per modellare le percezioni come delle attività di problem solving (risoluzione dei problemi).

Quest'architettura organizza il processo di risoluzione dei problemi in una gerarchia di sotto-obiettivi e usa un metodo di aggregazione in blocchi per generalizzare e velocizzare quel processo.

Fornisce un modello cognitivo integrato con un intero intervallo di capacità cognitive, includendo percezione, pianificazione deliberativa, emozioni, reazioni e linguaggio naturale.

La mancanza di parallelismo e meccanismi temporali sono assicurati dall'integrazione di SOAR con RS.

Architettura di ADAPT(4)

- ◎ Altre componenti:
 - Modello del mondo multimediale rappresentato tramite la piattaforma Ogre3D. Per mezzo di Ogre è possibile includere gli aspetti della dinamica e della grafica del mondo di ADAPT.
SOAR contiene la parte simbolica del mondo, che consiste di annotazioni che descrivono le entità, le relazioni e le attività di Ogre.
 - Il server del robot che si interfaccia direttamente con l'hardware sottostante.
 - SoarRobot è la parte che si occupa dell'interazione tra tutte le componenti del sistema e l'utente.
- ◎ Il modello del mondo di ADAPT è connesso esclusivamente a SOAR, quindi i collegamenti con l'esterno sono direttamente gestiti ed elaborati da SOAR.
- ◎ Solo gli oggetti dell'ambiente, importanti per il compito attuale, sono elaborati al fine della visualizzazione.

A cosa serve il modello del mondo

- ⊙ Sistema di visione costituito da 2 componenti:
 - Uno lavora ad alto livello di dettaglio, producendo mappe basilari di ineguaglianza, flussi immagine di grana grossa, indipendenti dal compito eseguito
 - L'altro lavora a basso livello di dettaglio, eseguendo compiti come riconoscimento degli oggetti, filtri sui dati, dipendentemente dal compito eseguito.
- ⊙ Semantica del linguaggio fornisce un'integrazione tra la grammatica e il significato del linguaggio simbolico usando le primitive spaziali per dare un significato a molte azioni e relazioni. [*Grammatica Cognitiva di Langacker*]

Gestione del linguaggio naturale

[Green & Lehman]

- L'interazione uomo-robot è realizzata tramite un'interfaccia del discorso basata sul dialogo.
- NL-Soar implementa un modello basato sulle ricette del discorso per la comprensione e generazione del dialogo.

PAROLE DEL
TESTO

Accesso
Lessicale

Modello
Sintattico

Modello
Semantico

Modello del
discorso

COMPRENSIONE
DEL TESTO

- Apprende le ricette del discorso, che sono la generalizzazione delle pianificazioni del discorso di un agente, effettuando la pianificazione del dialogo.

Componenti per NPL

- **Automati a stati finiti (FSA)**

Limitano la capacità di adattamento alle situazioni del mondo

- **Architettura di credenze-desiderio-intenzioni (BDI)**

Modellano gli obiettivi e le credenze dei partecipanti del discorso così come sono percepite dal robot

- **Dialoghi basati sul compito/struttura**

Eliminano il determinismo dei FSA ma non riescono ad adattarsi agli ambienti altamente dinamici

- **Costrutti di riconoscimento dei piani**

- **Motore di movimento del dialogo (DME)**

Seleziona una nuova azione a seconda del contesto dell'enunciato, gli obiettivi e l'evolversi del modello della conversazione

Comprensione attraverso la generazione

- ⦿ *Costruisce una spiegazione per l'enunciato di input, cercando di generare un enunciato simile a quello di partenza.*
- ⦿ *SOAR genera una risposta adeguata al discorso, cercando tra le possibili combinazioni di obiettivi e aspettative della persona con cui il robot parla.*

VANTAGGI:

Comprensione più profonda del mondo rispetto a quello che i modelli statistici possono fornire

SVANTAGGI:

Approccio lento con una ricerca lunga, che potrebbe fallire

Esempi

Approccio Verbale

- ⊙ “Abbassa la fiamma”
- ⊙ KB: Fiamma Alta \Rightarrow Cibo cuoce più velocemente
- ⊙ Obiettivo: Cibo cotto lentamente

Approccio Non-Verbale

- ⊙ “Il robot osserva l’umano a cui cammina affianco: l’uomo guardando avanti rallenta perché passa un’auto”
- ⊙ Ricostruisce un modello del mondo con uomo e veicolo
- ⊙ Simula la situazione opposta al comportamento tenuto

MIGLIORAMENTO DELL'INTERAZIONE UOMO-ROBOT

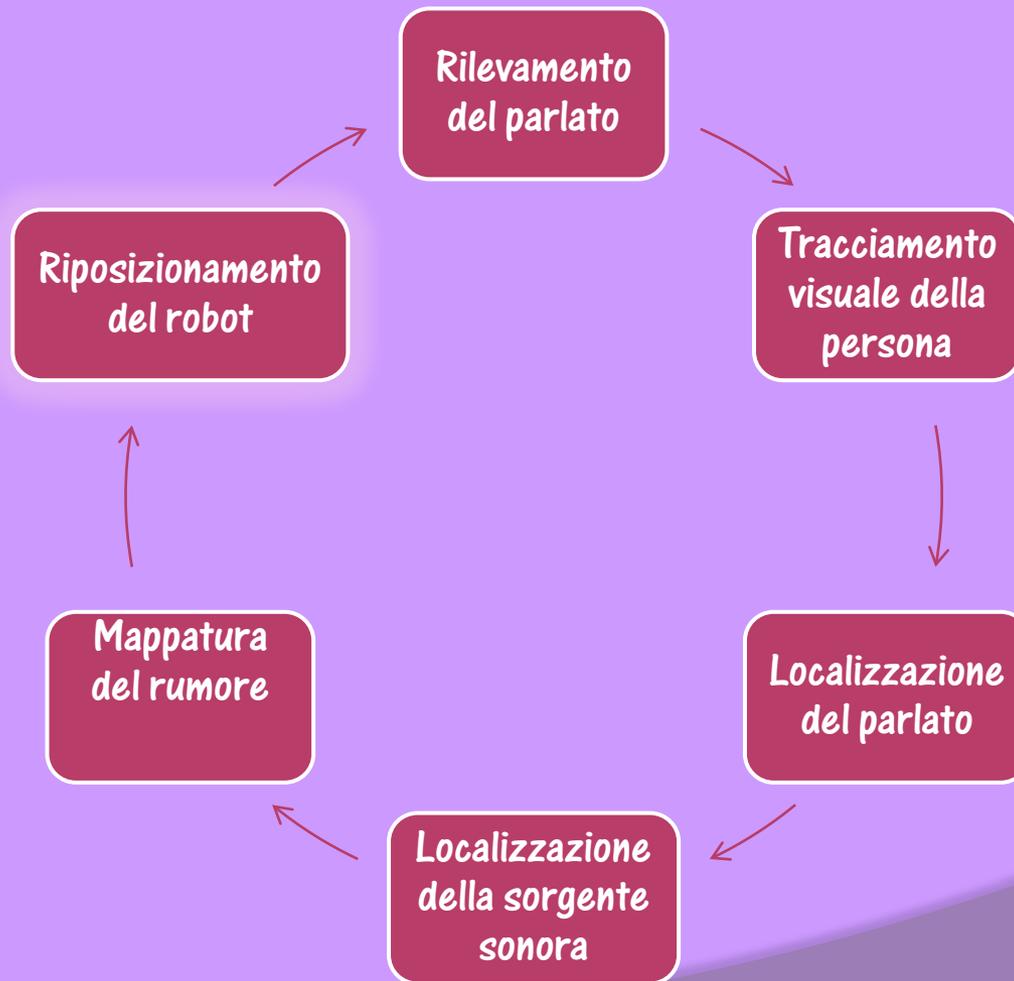
Perché parlare di miglioramento?

- Rumore
- Deterioramento degli enuncianti
- Scarsa comprensibilità

Cosa fanno in questo caso i robot?

1. *Adattano dinamicamente il volume del parlato a seconda della distanza dell'ascoltatore e del livello di rumore percepito*
2. *Inseriscono delle pause per evitare le eccessive interferenze date dalle sorgenti di rumore dell'ambiente*
3. *Si ruotano in modo da essere rivolti verso l'utente*
4. *Si riposizionano quando l'interferenza da rumore diventa troppo estesa*

Adattamento del robot al rumore



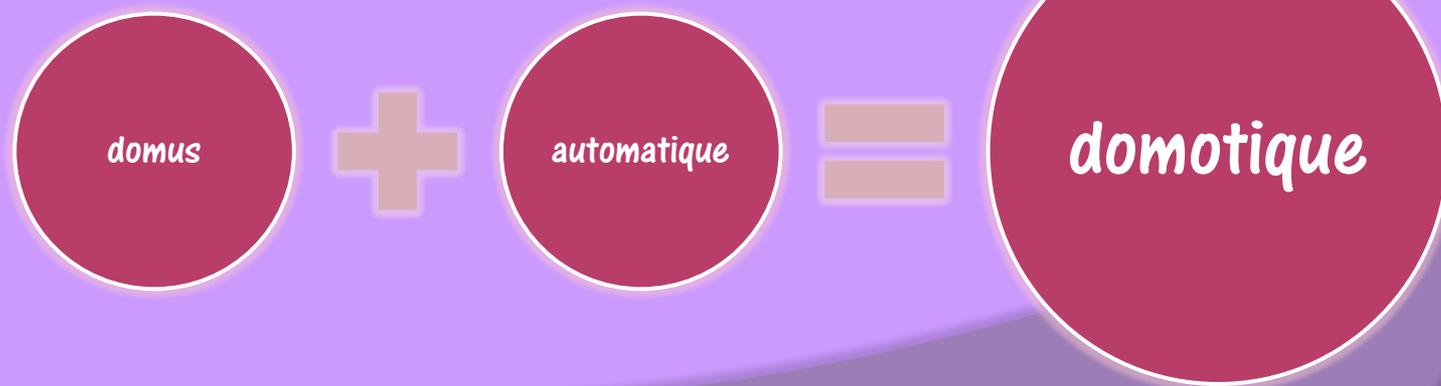
UTILIZZO DEL LINGUAGGIO NATURALE NELLA DOMOTICA



Definizione

Domotica: “è la disciplina che si occupa di studiare le tecnologie atte a migliorare la qualità della vita nella casa (e più in generale degli ambienti antropizzati) grazie all’automazione e al controllo di processi ripetitivi oppure poco performanti e all’integrazione dei sistemi”

[<http://it.wikipedia.org/wiki/Domotica>]



Caratteristiche della Domotica

- ◎ **Intelligenza ambientale:** ambiente dotato di uno svariato numero di tecnologie, tra loro complementari, non invasive. Queste tecnologie sono in grado di offrire servizi agli utenti o possedere delle caratteristiche auspicabili, **minimizzando le abilità necessarie per l'interazione con il sistema.**
- ◎ Un'intelligenza ambientale deve essere in grado di
 - Riconoscere gli utenti e i loro stati ed agire di conseguenza
 - Avere un comportamento predittivo basato sull'ambiente e sui suoi abitanti
 - Produrre servizi in tempo reale
 - Permettere di accedere ai molti servizi indipendentemente da dove si trova l'utente (ubiquità)
 - Collegare in modo naturale l'utente ai mezzi di interfacce multimodali basate sulla voce, leggendo i movimenti e i gesti, generando immagini.

Caratteristiche dell'Intelligenza ambientale

- Invisibile: deve essere incorporata negli oggetti
- Mobile: deve essere in grado di essere trasportata in giro
- Consapevolezza del contesto: deve esaminare l'ambiente locale per recuperare informazioni utili
- Comportamento di previsione: deve agire a seconda dell'utilità senza richieste esplicite dell'utente
- Comunicazione naturale: deve comunicare con i potenziali utenti tramite voce e gesti
- Adattabile: deve essere in grado di reagire a tutti i tipi di situazioni eccezionali, in modo flessibile, senza interrompere il servizio

Domotica come supporto per anziani e diversamente abili

Obiettivo: prolungare la vita autonoma delle persone assistite nelle loro case, riducendo al minimo l'intervento delle cure intensive.

Possono essere classificati in 2 modi diversi

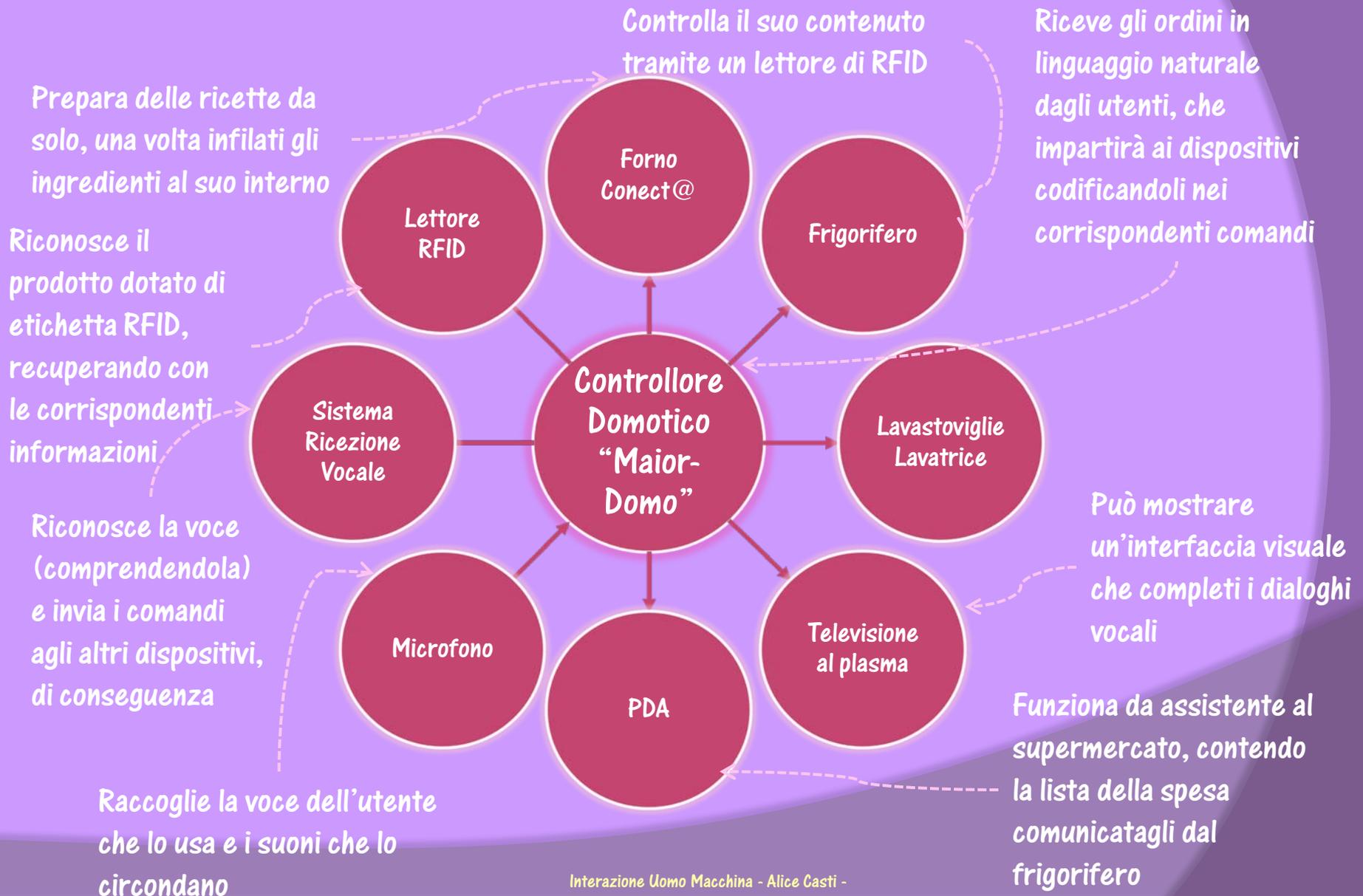
⊙ Per i servizi offerti

- Trattamento dell'emergenza (es. cadute, attacchi di cuore)
- Miglioramento dell'autonomia (es. assistenza medica, assistenza nel mangiare, assistenza nei viaggi)
- Comodità (es. ritrovamento oggetti, servizi d'orientamento)

⊙ Per luogo di assistenza

- Dentro casa
- All'aperto

Un'applicazione domotica: Genio



Capacità di Genio

- ⊙ Riconosce la voce, capendo la frase pronunciata e elabora la richiesta
- ⊙ Instaura un dialogo con l'utente, in linguaggio naturale
- ⊙ Spegne/accende dispositivi
- ⊙ Controlla la posta elettronica
- ⊙ Svolge attività di intrattenimento

Esempio:

Utente: "Maior-Domo"

MD: "Dimmi"

Utente: "Lava i miei abiti quando rientro da lavoro"

MD: "Quando tornerai da lavoro?"

Utente: "Alle 8 di sera"

MD: "A quale programma li lavo?"

Utente: "Delicato, 30° C"

MD: "Lavatrice programmata"

Bibliografia

1. A. Cappelli, E. Giovannetti, “L’interazione uomo robot” Robocare Technical Report 2003
2. M. Heerink, B. Krose, B. Wielinga, V. Evers “Enjoyment, Intention to User And Actual Use of a Conversational Robot by Elderly People”, ACM 2008
3. B. Robins, K. Dautenhahn, R. Boekhorst, C. Nehaniv, “Behaviour Delay and Robot Expressiveness in Child-Robot Interactions: A User Study on Interaction Kinesics”, ACM 2008
4. K. Sugiura, N. Iwahashi, “Learning Objects-Manipulation Verbs for Human-Robot Communication”, MISI November 2007
5. K. Hsiao, S. Vasoughi, S. Tellex, “Object Schemas for Responsive Language Use”, ACM 2008
6. H. Jacobsson, N. Hawes, G. Kruijff, J. Wyatt, “Crossmodal Content Binding in Information-Processing Architectures”, ACM 2008
7. D. Benjamin, D. Lonsdale, D. Lyons, “A Cognitive Robotics Approach to Comprehending Human Language and Behaviours”, ACM 2007
8. J. Nehmer, A. Karshmer, “Living Assistance System – An Ambient Intelligence Approach”, ICSE 2006 - ACM

Bibliografia (2)

9. S. Kiesler, P. Hinds “Human-Robot Interaction”, Human-Computer Interaction 2004
10. A. Steinfeld, T. Fong, D. Kaber “Common metrics for Human-Robot Interaction” ACM 2006
11. M. E. Foster, E. Gurman Bard, M. Guhe, R. Hill, J. Oberlander, A. Knoll, “The Roles of Haptic-Ostensive Referring Expressions in Cooperative, Task-based Human-Robot Dialogue” ACM 2008
12. J. R. Wolpaw, “Brain-Computer Interfaces (BCIs) for Communication and Control” ACM 2007
13. A. Nijholt, B. Allison, M. Moore Jackson, D. Tan, J. Del R. Millán, B. Graitmann, “Brain-Computer Interfaces for HCI and Games”, CHI 2008
14. E. Martinson, D. Brock, “Improving Human-Robot Interaction through Adaptation to the Auditory Scene”, HRI 2007 – ACM
15. A. Gárate, N. Herrasti, A. López, “GENIO: An Ambient Intelligence application in home automation and entertainment environment”, EUSAI conference 2005