

*Progetto ROBOCARE: sistema multi-agenti  
con componenti fisse e robotiche mobili intelligenti*

Settore "Piattaforme ITC abilitanti complesse ad oggetti distribuiti"  
MIUR legge 449/97 per l'anno 2000

Robo  
Care

## L'interazione Uomo-Robot

*Amedeo Cappelli, Emiliano Giovannetti*

Laboratorio KDD  
Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo", CNR, Pisa

{Amedeo.Cappelli,Emiliano.Giovannetti}@isti.cnr.it

## The ROBOCARE Technical Reports

---

RC-TR-1103-1

L'interazione Uomo-Robot  
Amedeo Cappelli, Emiliano Giovannetti

Robocare Technical Report N. 1

November 2003

Corresponding Author:

Amedeo Cappelli

Laboratorio KDD

Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo", CNR, Pisa

Amedeo.Cappelli@isti.cnr.it

"RoboCare: A Multi-Agent System with Intelligent Fixed and Mobile Robotic Components"  
A project funded by the Italian Ministry for Education, University and Research (MIUR) under  
Law 449/97 (Funds 2000).

<http://robocare.ip.rm.cnr.it>



© 2003

ISTC-CNR

Consiglio Nazionale delle Ricerche

Istituto di Scienze e Tecnologie della Cognizione

Viale Marx, 15 - 00137 ROMA



# L'Interazione Uomo-Robot\*

**Amedeo Cappelli, Emiliano Giovannetti**

Laboratorio KDD

Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo", CNR, Pisa

[Amedeo.Cappelli@isti.cnr.it](mailto:Amedeo.Cappelli@isti.cnr.it)

[Emiliano.Giovannetti@isti.cnr.it](mailto:Emiliano.Giovannetti@isti.cnr.it)

## Abstract

L'Interazione Uomo-Robot (HRI, *Human-Robot Interaction*) è un'area di ricerca multidisciplinare in costante sviluppo ricca di spunti per ricerche avanzate e trasferimenti di tecnologia. Essa gioca un ruolo fondamentale nella realizzazione di robot che operano in ambienti aperti e cooperano con gli esseri umani. Compiti di questo tipo richiedono lo sviluppo di tecniche che permettano ad utenti inesperti di usare i loro robot in modo semplice e sicuro utilizzando interfacce intuitive e naturali. In questo lavoro, dopo un'introduzione riguardante le questioni fondamentali dell'HRI, saranno indagate le diverse classificazioni di robot e una particolare tassonomia dell'Interazione Uomo-Robot, dopodiché saranno presentate le diverse modalità di interazione possibili tra un robot e un essere umano corredate da una serie di applicazioni di interfacce uomo-robot avanzate con particolare riguardo per i robot di assistenza agli anziani e alle persone diversamente abili.

## 1 Introduzione

L'impiego dei robot su vasta scala industriale ha portato ad un sostanziale aumento della produttività ed ha permesso l'abbassamento dei costi di produzione. Parallelamente ai progressi della tecnologia è stato possibile rendere i robot sempre più indipendenti dall'operatore umano e in grado di muoversi con maggiore autonomia all'interno di un ambiente di lavoro.

La maggior parte dei robot industriali lavorano in un ambiente controllato e conosciuto e svolgono azioni ripetitive che fanno parte di compiti pre-programmati. L'esempio tipico è quello della catena di montaggio, dove i robot eseguono ripetutamente la stessa operazione.

Questo approccio è adatto in tutti quegli ambienti "sicuri" in cui non si presentano situazioni "impreviste". Nel caso in cui un robot si trovi ad agire in un ambiente complesso e dinamico è necessario che sia dotato di sistemi di percezione esterna in modo da "vedere" ciò che lo circonda e rispondere in maniera adeguata e autonomamente alle diverse situazioni, anche quelle impreviste.

L'autonomia di questo tipo di robot "evoluti" deriva dalla capacità di elaborare le informazioni raccolte dai sensori e di pianificare la sequenza di azioni da intraprendere. Alcuni esempi di robot di questo tipo sono quelli impiegati nell'esplorazione spaziale, capaci di riconoscere ed evitare ostacoli in ambienti impervi e sconosciuti a grandissime distanze da qualsiasi operatore.

Lo studio di sistemi robotici in grado di percepire l'ambiente in modo complesso e capaci di interagire con esso è uno dei temi più attuali nei principali laboratori di robotica nel mondo, specie nel contesto dell'interazione uomo-robot e nello sviluppo di interfacce naturali.

---

\* Questa ricerca è parzialmente finanziata dal MIUR (Ministero Italiano per l'Educazione, l'Università e la Ricerca) nell'ambito del Progetto RoboCare (Sistema multiagenti con componenti fisse e robotiche mobili intelligenti).

Utilizzare un robot in un ambiente naturale e dinamico abitato da esseri umani impone precisi requisiti riguardanti la percezione sensoriale, la mobilità e la destrezza, nonché la capacità di pianificare compiti, prendere decisioni ed effettuare ragionamenti.

Ad ogni modo, la tecnologia attuale, allo stato dell'arte, non è ancora in grado di soddisfare tutti questi requisiti. Un limite allo sviluppo di questo tipo di robot "sociali" deriva dalla mancanza di interfacce appropriate che permettano una interazione e una comunicazione naturali, intuitive e versatili, in altre parole, *human-friendly*. Interfacce di questo tipo sono considerate essenziali per programmare ed istruire il robot in modo efficiente.

Questo rapporto si propone di fornire una panoramica sulle questioni e le applicazioni concernenti l'Interazione Uomo-Robot e, specificatamente, vuol costituire un punto di riferimento per lo studio di interfacce per robot mobili sviluppate nell'ambito del progetto "RoboCare" [Cesta et al., 2003], coordinato dall'Istituto di Scienze e Tecnologie della Cognizione del CNR di Roma. Tale progetto mira allo sviluppo di sistemi distribuiti nei quali agenti software e robotici collaborano per fornire servizi in ambienti nei quali gli esseri umani hanno necessità di assistenza, come all'interno di case di cura.

Lo studio e la realizzazione di sistemi complessi in grado di eseguire compiti di questo tipo richiede la sinergia di un gran numero di discipline, che il progetto RoboCare effettivamente coinvolge, come la comunicazione, la rappresentazione della conoscenza, l'interazione uomo-robot, l'apprendimento, il ragionamento simbolico collettivo e individuale.

## 2 L'evoluzione della Robotica: dalle Industrie alla Società

Il termine "robot", coniato nel 1921 dal romanziere e commediografo ceco Karel Čapek nella sua opera "Rossum's universal robots", deriva dal vocabolo ceco "robota" che significa "lavoro forzato". Nell'opera di Čapek il robot protagonista, contrariamente al suo status di macchina umanoide priva di sensibilità, si dimostra all'altezza delle facoltà umane giungendo ad asservire e a distruggere il proprio creatore. I primi veri robot (ad uso industriale) sono apparsi trent'anni dopo.

Ciò che ha spinto l'uomo a progettare e costruire i robot è la necessità di avere a disposizione dei mezzi che possano sostituirlo nelle attività più onerose o rischiose. Tra le prime applicazioni "avanzate" spiccano i "telemanipolatori", bracci meccanici controllati da un operatore umano e utilizzati per maneggiare sostanze pericolose, come materiali radioattivi all'interno di centrali nucleari. Più avanti, grazie allo sviluppo dell'elettronica e dell'informatica, hanno visto la luce bracci meccanici completamente motorizzati e programmabili, il più famoso dei quali, il PUMA (*Programmable Universal Manipulator for Assembly*, manipolatore universale programmabile per assemblaggio), realizzato grazie ad un finanziamento della General Motors da Victor Scheinman, un ricercatore del MIT<sup>1</sup> (Massachusetts Institute of Technology), che ha segnato, per molti, l'inizio dell'era dei robot.

### 2.1 Tappe Fondamentali della Storia dei Robot

#### 1478

Un disegno di Leonardo da Vinci, secondo l'interpretazione di un ingegnere meccanico statunitense di nome Mark Elling Rosheim, rappresenterebbe il progetto di un robot programmabile. Il disegno in questione (Figura 1) è contenuto nel Codice Atlantico conservato a Milano nella Biblioteca Ambrosiana. Rosheim sostiene che quella che finora è stata considerata l'antenata dell'automobile era in realtà una

---

<sup>1</sup> Ai termini sottolineati corrispondono indirizzi internet riportati in appendice.

piattaforma programmabile, che doveva sorreggere degli automi da trasportare poi in una determinata località.

## 1920

La parola “robot” appare per la prima volta nell’opera teatrale di Karel Capek;

## 1942

Attraverso i racconti ed i romanzi di fantascienza di Isaac Asimov (Figura 2) il termine “robotica” diviene popolare. A lui si deve l’idea delle “tre leggi della robotica” (descritte per la prima volta nella storia “Runaround”) e del “cervello positronico”. Asimov, nei suoi racconti, prevede un mondo in cui i robot sono parte integrante della società e la robotica costituisce una potente industria.

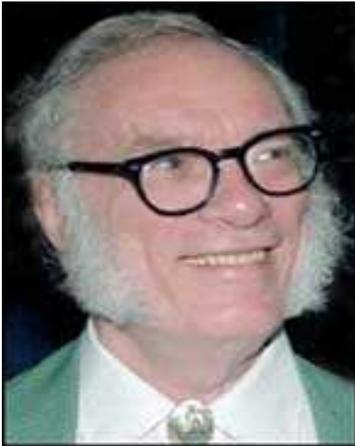


Figura 2: Isaac Asimov.

## 1974

Il braccio manipolatore comandato dal computer conosciuto come “Stanford Arm” (Figura 3), realizzato da Victor Scheinman, è utilizzato per assemblare automaticamente la pompa idraulica di una Ford Modello T per mezzo di sensori ottici e di contatto.

Successivamente Scheinman perfeziona il braccio e crea il PUMA. Il primo braccio meccanico programmabile, più rudimentale del PUMA e capace di un numero limitato di operazioni, era già stato realizzato dall’inventore americano George Devol jr. nel 1954.

## 1983

I primi robot mobili sono presentati al pubblico statunitense come strumenti adatti al servizio domestico e all’educazione.

## 1995

Circa 700,000 robot sono operativi nell’industria mondiale. Oltre 500,000 sono usati in Giappone, circa 120,000 nell’Europa occidentale e 60,000 negli Stati Uniti.

## 2.2 I Robot Industriali

I robot industriali possono essere classificati in funzione delle operazioni che sono in grado di eseguire:

**Prima generazione: (anni '60).** Sono macchine programmabili senza alcuna possibilità di controllo delle modalità reali di esecuzione e senza comunicazione con l’ambiente esterno. In genere sono semplici bracci meccanici privi di meccanismi di servocontrollo (e per questo molto rumorosi a causa

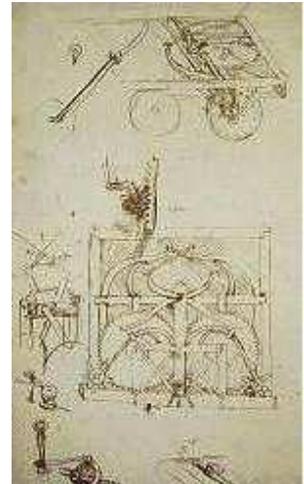


Figura 1: Leonardo, Codice Atlantico, foglio 812.

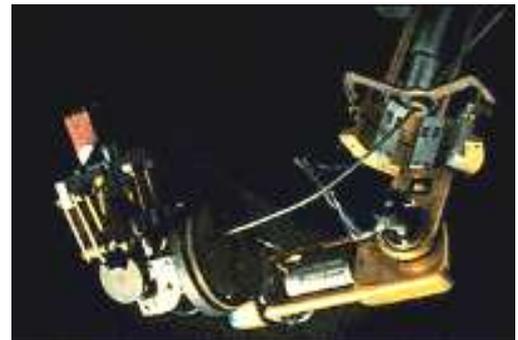


Figura 3: Lo “Stanford Arm”.

degli urti dei bracci contro i fermi meccanici) e a funzionamento pneumatico. L'uso di questi robot era prevalentemente limitato al carico/scarico di macchine e alla movimentazione di materiali.

La prima società a produrre robot fu, nel 1961, la statunitense Unimation. In Italia, il primo robot realizzato risale al 1969 ad opera della D.E.A. e si trattava di una macchina per effettuare la saldatura a punti.

**Seconda generazione: (anni '70).** Sono macchine programmabili a comportamento autoadattivo con elementari possibilità di riconoscimento dell'ambiente esterno per mezzo di sensori tattili e visivi. I robot di seconda generazione sono dotati di servocontrollo e possono essere programmati per spostamenti punto a punto o lungo un percorso continuo. Questo tipo di sistemi dispone di software dedicati per applicazioni specifiche e può risultare difficile utilizzare lo stesso robot per svolgere un compito diverso da quello per cui è stato programmato. Anche per quanto riguarda le avarie, i robot di seconda generazione dispongono di limitate capacità diagnostiche e si limitano ad accendere spie luminose in caso di guasti lasciando molto spesso all'operatore il compito di risalire all'origine del malfunzionamento. Questi robot venivano utilizzati per compiti più complessi, come il controllo dei centri di lavoro o la saldatura delle carrozzerie.

Verso la metà degli anni '70 alla robotica industriale si rivolgono le grandi case automobilistiche ed altre grandi aziende, come ACMA (Renault), COMAU (Fiat), Volkswagen ecc. Nel 1973 la società italiana "Olivetti" costruì, per uso interno, un sistema di montaggio a più bracci, nel quale vennero utilizzati per la prima volta dei sensori tattili. Era il prototipo del modello che sarebbe stato commercializzato dal 1976 con il nome di SIGMA.

**Terza generazione: (anni '80 - oggi).** Sono macchine autoprogrammabili interagenti con l'ambiente esterno e con l'operatore in modo complesso (attraverso la visione, la voce, ecc.) e in grado di autoistruirsi per mezzo di tecniche di intelligenza artificiale per l'esecuzione di un compito assegnato. I robot della terza generazione sono apparecchiature ad alta tecnologia che operano sotto servocontrollo. Attualmente, i robot di questo tipo sono impiegati per eseguire compiti "intelligenti", come la saldatura ad arco adattiva, nel corso delle quali il robot utilizza la visione o la percezione "attraverso l'arco" per localizzare il giunto di saldatura ed ottenere informazioni di guida del movimento. Questa generazione di robot si è rapidamente evoluta al punto da riuscire a svolgere operazioni altamente sofisticate come le ispezioni tattili, le lavorazioni a mano libera e le operazioni di assemblaggio.

**Quarta generazione: (?).** Probabilmente, i robot della quarta generazione saranno delle apparecchiature costruite con materiali in fibra o con composti di grafite, saranno più leggeri e più robusti e utilizzeranno sofisticati meccanismi di percezione sensoriale per una migliore coordinazione occhio-mano di cui avranno bisogno per svolgere le operazioni più complesse.

### 2.3 I Robot di Servizio

I cosiddetti "robot di servizio", o "robot assistenti", sono ancora lungi dal raggiungere il potenziale economico dei propri cugini robotici ad uso industriale. Utilizzare un robot in un ambiente naturale non modificato abitato da esseri umani richiede precisi requisiti concernenti la percezione sensoriale, la mobilità e la destrezza nonché la capacità di pianificazione dei compiti, di prendere decisioni ed effettuare ragionamenti.

Per interagire con un robot di servizio destinato ad operare in un ambiente sociale, quindi con una pluralità di soggetti, è indispensabile l'uso di un codice di comunicazione naturale: è impensabile utilizzare una tastiera o un mouse per comunicare, ad esempio, con un robot mobile dedicato alla pulizia

della casa e per esprimere naturalmente tutte le funzionalità che un robot domestico dovrebbe avere e che, secondo Khan [1998], sono:

- prendere e trasportare;
- preparare un pasto;
- pulire la casa;
- monitorare segnali di vita;
- assistere nella deambulazione;
- gestire l'ambiente;
- comunicare tramite la voce;
- intraprendere azioni di emergenza (fuoco, intrusione, ecc.)

Nel 1989 Joseph F. Engelberger, nel suo libro "Robotics in Service", sosteneva che nel 1995 i robot di tipo diverso da quello industriale avrebbero superato di gran lunga quelli impiegati in ambito manifatturiero.

Gli anni Novanta non sono stati testimoni di questa evoluzione dei robot di servizio. La previsione di Engelberger è risultata fallace per quanto concerne il settore della costruzione e vendita dei robot; invece si è dimostrata sostanzialmente vera per quanto concerne le attività di ricerca e di laboratorio.

### **2.3.1 Aspetti Sociali e Collaborativi.**

Dal momento che questi "robot di servizio" sono destinati a divenire parte della vita degli esseri umani, si rendono necessari studi appositi volti all'indagine degli aspetti sociali e collaborativi legati all'interazione uomo-robot. Da un lato, è possibile analizzare i modelli e le abilità comunicative e sociali umane ed utilizzarle come punto di partenza, dall'altro, per contro, sono indispensabili studi empirici di interazione uomo-robot in situazioni reali.

Presso l'IPLab di Stoccolma, ad esempio, sono state affrontate e studiate le relazioni sociali e di collaborazione all'interno di un ufficio tra impiegati "diversamente abili" ed il robot mobile autonomo CERO [Severinson-Eklundh et al., 2003] realizzato presso il laboratorio stesso (cfr. 4.7).

Alcuni studi sull'interazione sociale tra robot ed esseri umani sono stati svolti presso l'Università di Reading e l'Università di Edimburgo [Dautenhahn et al., 1999]. La ricerca è stata basata sull'assunto che per studiare lo sviluppo cognitivo dei robot è necessario considerare i "robot nella società", utilizzando un approccio che considera le interazioni sociali come fondamentali per la progettazione di funzionalità cognitive sempre più ricche.

L'HRP-2P è un prototipo di robot umanoide realizzato nell'ambito dello Humanoid Robotics Project giapponese. Il robot è alto 154 centimetri, pesa 58 Kg ed è dotato di 30 gradi di libertà. L'HRP-2P è utilizzato per esperimenti di deambulazione su terreni impervi, per studi sulla capacità di rialzarsi in piedi in seguito ad una caduta e, in generale, per tutte le situazioni pertinenti alla collaborazione uomo-robot.

All'Università di Stanford, al "Center for Work Technology and Organization" (WTO), è in corso uno studio sul campo con un robot mobile autonomo chiamato HELPMATE (Figura 4), che opera in un ambiente ospedaliero. Progettato dalla Pyxis Corporation, HELPMATE funziona a batterie e funge da corriere che risponde a richieste programmate trasportando materiali tra differenti impiegati e luoghi diversi all'interno dell'ospedale. Lo studio di Stanford è di natura prevalentemente etnografica: i

ricercatori raccolgono dati qualitativi osservando le interazioni tra gli impiegati ed il robot e intervistando gli impiegati circa le loro esperienze. Gli esperimenti condotti mirano allo studio di come l'aspetto fisico e l'atteggiamento psicologico del robot influenzino l'attitudine delle persone verso il robot stesso e motivino il desiderio di collaborare con esso. Inoltre, lo studio prevede una comparazione del comportamento degli impiegati prima e dopo l'introduzione del robot.

La Carnegie Mellon University e la Stanford University, nell'ambito del "Project on People and Robots", si occupano di robot di servizio e assistenza, con particolare riguardo a tre aree di ricerca: il design dei robot, l'interazione uomo-robot e come funzionano i robot all'interno di un gruppo di lavoro. Come dominio iniziale della ricerca sono state selezionate comunità per anziani e ospedali, dove i robot di servizio possono svolgere compiti utili di assistenza e monitoraggio.

Per quanto riguarda il gruppo incaricato di eseguire gli studi di design, sono stati generati metodi quantitativi e qualitativi per valutare i punti di incontro tra gli attributi del prodotto e le esigenze maturate dall'esperienza delle persone. Il lavoro di questa unità mira ad una comprensione dettagliata di come i robot possono essere progettati, dal punto di vista estetico (Figura 5) e sociale, per andare incontro alle necessità e ai desideri delle persone [DiSalvo et al., 2002].

L'unità di ricerca del progetto che si occupa di Interazione Uomo-Robot ha come obiettivo la comprensione scientifica di come le persone interagiscono con i robot autonomi sociali: in particolare cerca di capire in che modo le persone si creano un modello mentale di robot come macchina dagli attributi simil-umani. In base a queste ricerche l'unità intende creare principi per la progettazione di un



**Figura 4:** HELPMATE.



**Figura 5:** Studi sull'aspetto della testa (CMU).

assistente robotico che riguardino il dialogo, i processi di interazione e i meccanismi di feedback che vadano incontro alle necessità sociali e pratiche e che incoraggino la cooperazione tra le persone e il robot. Tra gli studi previsti figurano:

- lo sviluppo di misure di antropomorfismo;
- l'analisi dei modelli mentali delle persone circa i robot e le macchine basate sui computer;
- la comparazione degli effetti di diversi progetti robotici sulla cooperazione delle persone con i robot;
- lo sviluppo e il test di interfacce per robot interattivi.

La terza unità di ricerca del progetto si interessa dei robot nel contesto, ovvero di come i robot sociali operano come membri di un team di lavoro. L'obiettivo è quello di capire meglio come le persone rispondano ai robot mobili in ambienti lavorativi e come l'uso di tali robot possa influenzare il lavoro delle persone.

**2.3.2 I Robot Umanoidi Autonomi.** A differenza dei robot di servizio autonomi, capaci di svolgere un numero limitato di compiti specifici senza la supervisione umana, i robot umanoidi combinano abilità avanzate di manipolazione e processi cognitivi simil-umani, in una forma antropomorfa in modo da essere in grado di operare in ambienti non modificati adatti agli esseri umani ed utilizzare, perciò, gli strumenti e le apparecchiature che usa normalmente l'uomo. I robot di questo tipo sono progettati per ricevere istruzioni circa i compiti da eseguire per mezzo di "interfacce multi-modali" (cfr. 4.7) che combinano l'uso del parlato, dei gesti, delle espressioni facciali, ecc.

I robot umanoidi sono stati per molti anni appannaggio esclusivo della fantascienza, soltanto recentemente, grazie all'evoluzione tecnologica da un lato (potenza di elaborazione, meccatronica, percezione robotica, ecc.) e agli studi cognitivi dall'altro (linguistica, psicologia, ecc.), è stato possibile sviluppare strumenti di intelligenza artificiale integrati in sistemi operazionali autonomi. A differenza dei robot umanoidi della fantascienza, però, i sistemi di oggi non sono creature super-intelligenti e super-veloci, anche se molti progressi sono stati compiuti e non è raro stupirsi assistendo alla dimostrazione di un nuovo robot umanoide in azione.

Presso il JSK Laboratory, dell'Università di Tokyo, sono stati progettati e realizzati diversi robot mobili umanoidi, l'ultimo dei quali si chiama H7, utilizzati soprattutto come piattaforme sperimentali per lo studio della deambulazione.

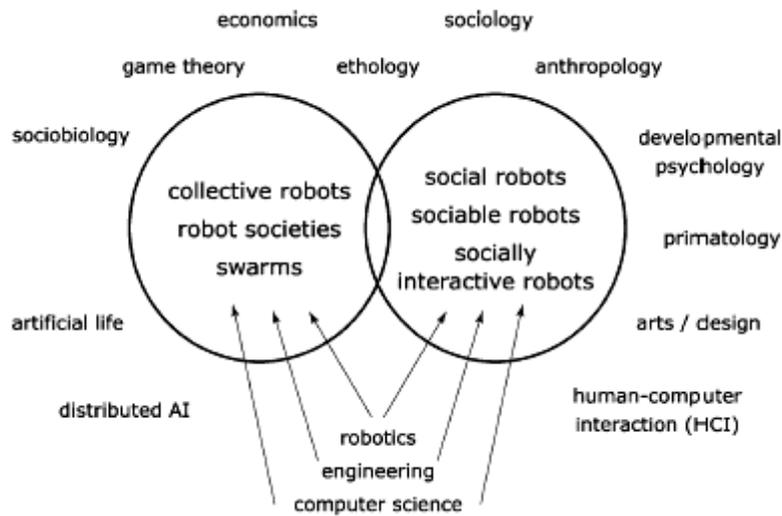
E' targato Honda il robot umanoide intelligente chiamato ASIMO (Figura 6). Il robot è in grado di interpretare le posture e i gesti degli esseri umani e di interagire con essi in diversi modi: può salutare le persone che gli si avvicinano, seguirle, spostarsi nella direzione indicata, riconoscere le facce delle persone e chiamarle per nome. Oltre alle sopracitate capacità di interazione ASIMO è capace di riconoscere l'ambiente in cui si trova, seguire oggetti in movimento, camminare, salire e scendere le scale e perfino restare in equilibrio su di una gamba.



**Figura 6:** ASIMO.

## 2.4 Robot Collettivi vs. Robot Individuali

Prima di passare al prossimo capitolo, dedicato all'Interazione Uomo-Robot, è opportuno analizzare in dettaglio le varie tipologie che è possibile individuare parlando di robot sociali distinguendo tra robot sociali "collettivi" e robot sociali "individuali" [Fong et al., 2003] intendendo i primi come facenti parte di un gruppo di robot ("group social collective robots") la cui funzione è prevalentemente "di squadra" mentre i secondi come agenti individuali ("individual social robots") più "individualisti" e orientati all'interazione con l'uomo. Nella Figura 7 i robot sociali individuali sono indicati semplicemente come "robot sociali" in contrapposizione ai robot "collettivi". Intorno ai cerchi rappresentanti le varie tipologie di robot sono riportati i campi di studio maggiormente coinvolti nello studio della relativa categoria di robot. L'intersezione dei due cerchi rappresenta quei robot individuali in cui l'individualità gioca un ruolo minore e, analogamente, quei robot collettivi dotati di maggiore autonomia individuale.



**Figura 7:** Robot sociali “collettivi” (a sinistra) e robot sociali “individuali” (a destra).

## 2.5 Gli Artefatti Attivi

Una particolare categoria di sistemi robotici intelligenti, destinati probabilmente ad entrare a far parte della società umana nei prossimi anni, è costituita dai cosiddetti “artefatti attivi”.

L’obiettivo che si pongono al MIT, nell’ambito del consorzio “Things that Think”, è quello di instillare intelligenza all’interno di semplici oggetti di uso quotidiano, come porte e sedie. Questi artefatti attivi sono robot destinati ad operare nell’ambiente umano progettati per essere molto semplici da utilizzare. Per facilitare l’interazione tra gli esseri umani e questi robot sono utilizzate modalità di interazione “non-verbali”, la comunicazione corporale e le cosiddette “interfacce tangibili” [Ullmer et al., 1997; Ishii et al., 1997].

Tra gli artefatti attivi già realizzati figura una sedia robotica [Terdada et al., 2002] capace di muoversi verso l’operatore umano non appena quest’ultimo manifesti l’intenzione di sedersi tramite un movimento della mano: non appena il gesto viene riconosciuto, attraverso l’uso di una telecamera, la sedia si muove autonomamente fino a raggiungere il punto dove l’utente l’ha richiamata.

## 3 L’Interazione Uomo-Robot

Due enti statunitensi, il National Science Foundation (NSF) e il Department of Energy (DOE), hanno definito l’interazione uomo-robot di importanza strategica [Murphy et al, 2001] e hanno sottolineato quanto la ricerca in questo campo sia solo agli inizi. L’importanza dell’HRI è motivata soprattutto dai recenti sviluppi tecnologici che hanno permesso ai robot di uscire dalle mura delle fabbriche ed entrare tra le mura domestiche. Per fare questo, però, sono necessari meccanismi di interazione avanzati per permettere a persone senza particolari esperienze o conoscenze di interagire facilmente con i robot.

Le questioni legate alle interfacce e all'interazione con l'uomo sono da molto tempo oggetto della ricerca nell'ambito della robotica. Tipicamente, le persone che lavorano con i robot hanno mansioni di supervisionamento e/o teleoperazione: i primi studi volti al miglioramento delle interfacce sono stati motivati per facilitare questo tipo di interazione. Fino a poco tempo fa, comunque, l'attenzione della comunità robotica è stata principalmente "robot-centrica" con maggiore enfasi sulla sfida tecnologica per ottenere controllo e mobilità intelligenti. Soltanto di recente, i progressi scientifici e tecnologici hanno permesso di fare predizioni di questo tipo: "entro una decade, robot che rispondono al telefono, aprono la posta, consegnano documenti a dipartimenti differenti, preparano il caffè, puliscono e passano l'aspirapolvere potrebbero occupare ogni ufficio." [Lorek, 2001].

Per fare in modo che i robot possano svolgere questo tipo di compiti è opinione comune che necessitino di un'intelligenza di tipo "umano" e che siano capaci di interagire con gli esseri umani (e talvolta tra di loro) nel modo in cui gli uomini comunicano tra loro. Questo approccio alla robotica, di tipo "uomo-centrico", enfatizza lo studio degli esseri umani come modelli per i robot. Ad ogni modo, in conseguenza del miglioramento delle prestazioni fisiche dei robot, si sta concretizzando la possibilità di utilizzarli sia nei luoghi comuni, come uffici, fabbriche, case e ospedali, sia in ambienti più tecnici come stazioni spaziali, pianeti, miniere, fondali oceanici, ecc.

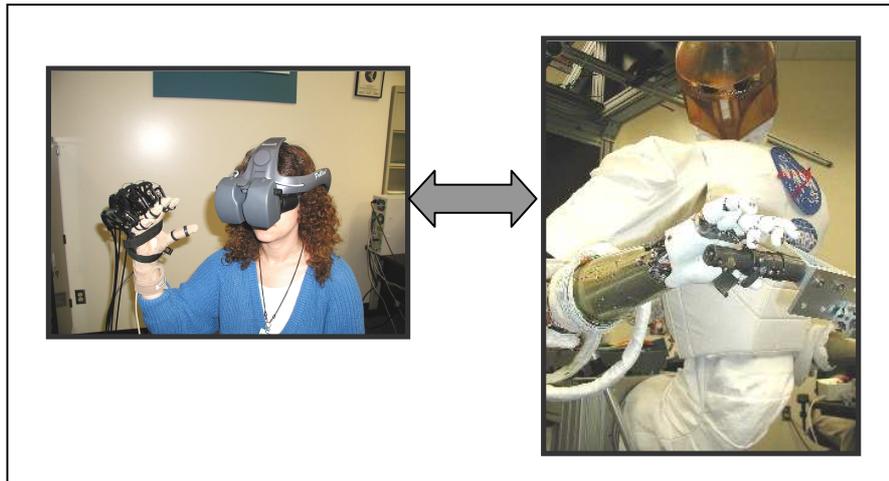
Prima che robot intelligenti siano sviluppati e integrati nella nostra società, però, è necessario studiare attentamente la natura delle relazioni uomo-robot e l'impatto che queste relazioni possono avere nel futuro della società umana. Una buona strategia per fare questo è attingere alla grande esperienza già maturata nell'ambito della comunità dell'Interazione Uomo-Computer (HCI, *Human-Computer Interaction*), dove sono state studiate le direzioni dello sviluppo tecnologico e il suo impatto sugli esseri umani.

### 3.1 Una Tassonomia dell'Interazione Uomo-Robot

Per meglio comprendere la natura dell'Interazione Uomo-Robot riportiamo di seguito una particolare tassonomia [Yanco et al., 2002] in cui è possibile inquadrare le possibili situazioni interattive. Tra le altre ricerche svolte in questo senso citiamo [Agah, 2001] in cui è presentata una tassonomia dell'interazione tra uomo e sistemi intelligenti e [Dudek et al., 2002] e [Balch, 2002] in cui sono studiate tassonomie concernenti l'interazione con robot multipli.

**Livello di autonomia / livello di intervento.** La prima categoria della tassonomia concerne il livello di intervento necessario per controllare un robot. Ad un estremo (massima autonomia) troviamo i robot completamente autonomi: sistemi di questo tipo, ad esempio, possono essere i robot utilizzati per mansioni di trasporto di oggetti all'interno di un ufficio [Severinson-Eklundh et al., 2003] oppure robot utilizzati per visite guidate come "Polly" che fa visitare il settimo piano dell'Artificial Intelligence Laboratory del MIT [Horswill, 1995]. All'estremo opposto si pongono i robot totalmente privi di autonomia che richiedono il controllo costante e diretto da parte di un operatore umano. ROBONAUT (Figura 8), della NASA, fornisce un esempio di sistema di "telepresenza" utilizzabile sia da terra che a bordo di una nave spaziale [Ambrose et al., 2000].

**Rapporto numero di persone/numero di robot.** Questa categoria tassonomica non misura l'interazione tra l'uomo e il robot ma indica semplicemente il numero di persone coinvolte nell'interazione diviso il numero di robot.

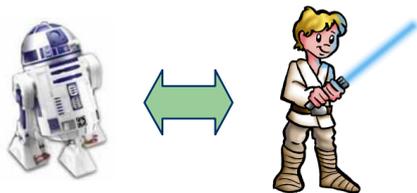


**Figura 8:** Telepresenza col sistema ROBONAUT.

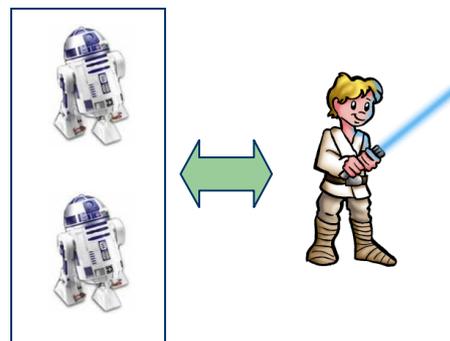
**Livello di interazione condivisa tra team.** Per valutare la tipologia di interazione tra uomo e robot è necessario distinguere diversi casi in base alla composizione del team di operatori umani e del team di robot.

Presentando le varie possibilità, cominciamo con l'interazione più semplice, ovvero quella costituita da un operatore umano e un singolo robot (Figura 9a). In questo caso l'umano impartisce comandi al robot il quale restituisce le informazioni raccolte dai sensori all'operatore. Un uomo che interagisce con la propria sedia a rotelle robotica costituisce un esempio di interazione di questo tipo.

Nel caso in cui un uomo interagisca con un team di robot (un gruppo di robot coordinato) si ha la situazione riportata in Figura 9b: l'operatore assegna un compito ad un insieme di robot i quali decidono tra loro come portarlo a termine e come suddividersi i compiti. Come esempio di interazione di questo tipo si può pensare ad un operatore che comandi ad un gruppo di robot di "pulire la casa": i robot si divideranno i compiti decidendo chi si occuperà dei pavimenti, chi dei vetri e così via.



**Figura 9a:** Un robot – un operatore.



**Figura 9b:** Team di robot – un operatore.

Diverso è il caso in cui un singolo operatore umano impartisca compiti differenti a robot differenti (Figura 9c): i singoli robot non necessariamente sono tenuti a sapere come si comporteranno gli altri robot e ognuno agirà autonomamente. Per portare un esempio si pensi ad una situazione analoga alla precedente in cui però l'operatore impartisca ai robot compiti specifici oppure assegni ad ogni robot una stanza diversa di cui occuparsi.

In modo complementare, possiamo considerare situazioni in cui operatori multipli comandino un solo robot. In Figura 9d è riportata la situazione interattiva in cui un team di operatori si accorda per impartire un singolo comando ad un robot. Due ospiti di una casa di cura che decidono insieme di chiedere ad un robot di portare loro un mazzo di carte costituisce un esempio di interazione di questo tipo.

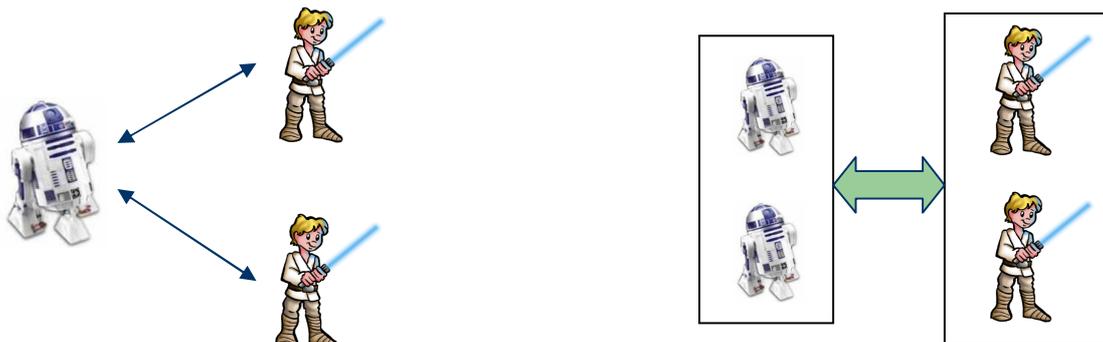


**Figura 9c:** Robot multipli – un operatore.

**Figura 9d:** Un robot – team di operatori.

Nella Figura 9e gli operatori umani agiscono indipendentemente e impartiscono comandi differenti allo stesso robot che dovrà necessariamente stabilire quale compito eseguire per primo e risolvere eventuali situazioni di conflitto. Un’interazione di questo genere può verificarsi, ad esempio, se due persone diverse comandano allo stesso robot di portare loro un certo oggetto: il robot dovrà decidere chi servire per primo e quindi portare a termine i due compiti nell’ordine stabilito.

Gli ultimi tre casi di interazione da analizzare coinvolgono operatori umani multipli che interagiscono con robot multipli. In Figura 9f un gruppo di operatori umani si coordina per impartire un ordine ad un gruppo di robot i quali, a loro volta, si organizzeranno per portare a termine il compito stabilito. Un’interazione di questo tipo si può avere, ad esempio, se due infermieri all’interno di un ospedale si accordano per comandare ad un gruppo di robot di fare le pulizie in una determinata area del reparto.



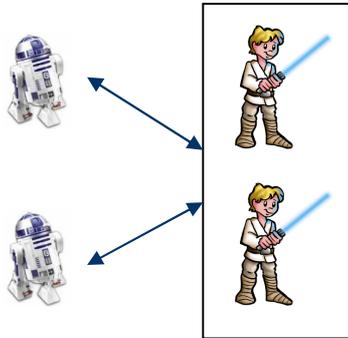
**Figura 9e:** Un robot – operatori multipli.

**Figura 9f:** Team di robot – team di operatori.

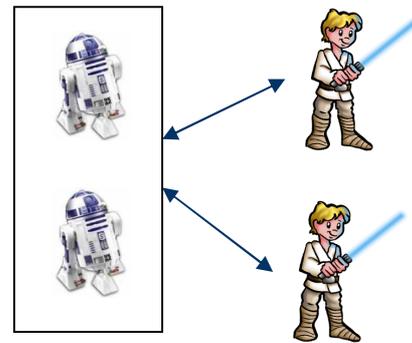
Nel caso in cui un team di operatori impartisca ordini diversi a robot diversi si ha la situazione riportata in Figura 9g. In questo caso gli umani si accordano sui comandi da impartire ai singoli robot che eseguiranno i compiti assegnati autonomamente. Per fare un esempio analogo al precedente, è sufficiente immaginare una situazione in cui due infermieri si mettono d’accordo per impartire ordini

separatamente ad un robot “pulitore” per occuparsi di una stanza e ad un robot medico affinché controlli lo stato di salute di un certo paziente.

L’ultima situazione è quella in cui un insieme non coordinato di persone impartisce ordini ad un gruppo coordinato di robot (Figura 9h). I robot dovranno stabilire delle priorità, risolvere situazioni conflittuali e spartirsi i compiti prima di cominciare ad eseguirli. Un gruppo di robot pulitori che riceve ordini da persone distinte cade in questa categoria.



**Figura 9g:** Robot multipli – team di operatori.



**Figura 9h:** Team di robot – operatori multipli.

**Supporto alle decisioni per gli operatori.** Questa categoria tassonomica concerne il tipo di informazione disponibile all’operatore per il supporto alle decisioni. La categoria può essere suddivisa in tre sottocategorie:

- *informazione dei sensori disponibile:* indica l’ubicazione e la tipologia dei sensori disponibili sulla piattaforma robotica;
- *informazione dei sensori fornita:* costituisce il sottoinsieme di sensori indicati nella precedente categoria utilizzati effettivamente per il supporto alle decisioni;
- *tipo di fusione dei sensori:* specifica le eventuali fusioni di informazione sensoriale utilizzata per il supporto alle decisioni (ad es: *sonar* e *ladar* per costruire una mappa da visualizzare per l’operatore);
- *pre-processing:* verifica lo stato iniziale dei sensori necessario per il supporto alle decisioni.

La misura dell’uso dell’informazione proveniente dai sensori nell’interfaccia utente permette di determinare come la quantità di supporto alle decisioni influenzi la prestazione dell’operatore: sarà più semplice controllare un robot che fornisca un supporto alle decisioni nella forma di una mappa piuttosto che una lista delle letture grezze provenienti dai sensori sonar.

**Criticità.** La criticità misura quanto può essere dannosa l’esecuzione errata di un compito. L’interazione con una sedia a rotelle robotica, ad esempio, avrà un alto valore di criticità: se la sedia fallisce nel rilevare una rampa di scale in discesa può far precipitare il suo operatore causando gravi lesioni. All’estremo opposto, un cagnolino robotico giocattolo che non funziona bene non provoca alcun danno.

**Tassonomia spazio/temporale.** Questa tassonomia [Ellis et al., 1991] suddivide l’Interazione Uomo-Robot in quattro categorie in base allo spazio che separa l’operatore dal robot e alla sincronicità o meno dell’interazione (Figura 10):

		Tempo	
		stesso	diverso
Spazio	stesso	Sedia a rotelle robotica	Robot industriali
	diverso	Ricerca e soccorso urbano	Mars Rover

**Figura 10:** Categoria tassonomica spazio/temporale con esempi.

Robot utilizzati per l'esplorazione spaziale, come il Mars Rover, sono classificati come "asincroni" e "non-collocati", dal momento che sono largamente autonomi e lavorano a milioni di chilometri dalla Terra. All'estremo opposto, una sedia a rotelle robotica costituirà un sistema "sincrono" e "collocato": risponderà istantaneamente ai comandi del disabile che la utilizza.

**Composizione del team di robot.** L'ultima categoria presentata riguarda la composizione del team di robot che può essere composto da robot dello stesso tipo o di tipo diverso. Team omogenei potranno interagire più facilmente per mezzo di una singola interfaccia, mentre per team eterogenei sarà più difficile presentare all'operatore un'informazione coerente utile per il supporto alle decisioni.

#### 4 Le Modalità di Interazione

La comunicazione tra un robot ed un essere umano può avvenire sia per mezzo dei dispositivi "classici" dell'Interazione Uomo-Computer (tastiera, mouse, touchpad, monitor, schermi sensibili al tocco, ecc.) oppure utilizzando modalità interattive più "naturali" (*human-friendly*), tipiche dell'interazione uomo-uomo, in grado di coinvolgere tutti i sensi umani e i canali di comunicazione, come il parlato, la visione, la gestualità e il tatto.

Dal momento che siamo particolarmente interessati all'interazione tra l'uomo e i robot sociali, dedicheremo questo capitolo all'indagine di questa seconda categoria di modalità interattive più "orientate all'uomo". In generale, è possibile distinguere sei categorie principali di modalità di interazione "naturali", anche se in alcuni casi, come evidenzieremo più avanti, i confini tra le diverse modalità appaiono sfumati:

1. parlato;
2. gesti;
3. espressioni facciali;
4. tracciamento dello sguardo;

5. prossemica e cinesica;
6. aptica.

Come risulta dai dati raccolti attraverso un sondaggio d'opinione [Khan, 1998] la maggior parte delle persone, interrogate circa i propri desideri in termini di interazione uomo-robot, predilige nettamente l'uso del parlato in combinazione ad altre modalità. Per questo motivo si rende necessario, da parte della comunità che si occupa di Interazione Uomo-Robot, procedere con ricerche di usabilità al fine di investigare fino a che punto la tecnologia disponibile oggi sia in grado di soddisfare le esigenze degli utenti e quanto le metodologie sviluppate in ambito di Interfacce in Linguaggio Naturale (NLI, *Natural Language Interfaces*) e multimodalità siano capaci di fornire efficienti mezzi di interazione con i robot.

Per valutare le caratteristiche comunicative di cui dovranno essere dotati i robot di prossima generazione, è particolarmente utile, specie da un punto di vista progettuale "uomo-centrico", analizzare le esperienze comunicative degli esseri umani e utilizzare questa conoscenza come punto di riferimento per la realizzazione delle interfacce uomo-robot.

La gamma di sistemi di comunicazione e interazione che le persone utilizzano include interfacce "faccia-a-faccia", interfacce mediate "uomo-a-uomo" e interfacce "uomo-macchina".

Nella comunicazione "faccia-a-faccia" la persone utilizzano il linguaggio (parlato), i gesti e gli sguardi per scambiare concetti, attitudini e opinioni. Normalmente, la comunicazione umana è ricca di fenomeni come ellissi, atti parlati indiretti e riferimenti ad oggetti ed azioni [Donellan, 1996; Milde et al., 1997]. Un'altra caratteristica implicita nella comunicazione uomo-uomo, l'ambiguità, è considerata una delle più importanti fonti di difficoltà nell'ambito dell'Elaborazione del Linguaggio Naturale. Le ambiguità incorporate nella conversazione uomo-a-uomo devono perciò essere accuratamente studiate e tenute in considerazione nella progettazione di interfacce uomo-robot [Grice, 1975].

Per quanto riguarda le Interfacce in Linguaggio Naturale sono già stati effettuati diversi studi [Ogden et al, 1997] e sviluppati diversi sistemi di dialogo telefonici [Bernsen et al, 1997] in grado di mettere in comunicazione un essere umano con una macchina. D'altra parte, l'integrazione di queste interfacce all'interno di robot richiede nuove strategie di dialogo, differenti sia dai sistemi telefonici sia dai sistemi NLI basati su workstation.

Per comprendere le difficoltà relative alla comunicazione uomo-robot si consideri la seguente situazione: un robot mobile ed un utente si trovano fisicamente nella stessa stanza e al robot viene chiesto di "andare a sinistra". L'esecuzione corretta dell'azione implica due direzioni differenti in relazione alla posizione del robot rispetto all'utente. In altre parole, il robot deve individuare l'ambiguità del termine "sinistra" il cui significato è pragmaticamente influenzato dalla posizione relativa dei due interlocutori. Un problema di questo tipo può essere risolto automaticamente oppure instaurando un dialogo appropriato con l'utente, possibilmente mediante una "interfaccia multimodale" (*multi-modal interface*) che, in situazioni di questo tipo, può aiutare sensibilmente nell'esecuzione del compito.

Le interfacce multimodali sono considerate particolarmente vantaggiose grazie alle loro proprietà di alta ridondanza, maggiore percettibilità, migliore accuratezza e i possibili effetti sinergici delle differenti modalità di comunicazione individuali.

Nella maggior parte dei sistemi odierni, per ovviare a questo tipo di problemi, i comandi vengono impartiti direttamente tramite manipolazione del dispositivo di input, per esempio tramite tastiere, manopole e pulsanti. Il processo di interazione viene controllato e gestito attraverso una rappresentazione grafica su uno schermo che può essere parte integrante del sistema robotico.

L'obiettivo che si desidera ottenere è spostare il campo di interazione dalla superficie di uno schermo allo spazio reale di una stanza [Bolt, 1980] che l'utente e il robot possono condividere. Le sfide

che i ricercatori si pongono per la realizzazione di interfacce uomo-robot dotate di una adeguata combinazione di modalità di interazione e comunicazione sono molteplici.

Il primo passo da compiere è stabilire alcuni “principi guida” per la progettazione di sistemi interattivi, come i robot mobili, che permettano di minimizzare la complessità dei dispositivi di input (guanti, microfoni montati sulla testa, sistemi di inseguimento dell’occhio, ecc.) utilizzati fino ad oggi nella ricerca in ambito di interazione multi-modale.

Analogamente, si rendono necessarie linee guida per quanto concerne la sicurezza, l’autorità di comando e la subordinatazza dei robot: un punto di partenza potrebbe essere costituito dalla famose “Tre Leggi della Robotica” di Isaac Asimov [Asimov, 1995].

Un’altra questione fondamentale che deve essere studiata e discussa riguarda la scelta della modalità di interazione che si rende desiderabile con un robot di un certo tipo. Per esempio, consideriamo un dispositivo robotico avanzato equivalente ad un agente intelligente autonomo nel quale l’utente ripone la propria fiducia: come dovrebbe essere progettato il sistema di interazione e comunicazione? Se il robot agisce come un agente sociale forse dovrebbe avere un’interfaccia illustrativa, oppure dovrebbe avere la possibilità di interagire non solo con l’utente primario ma poter lasciare messaggi vocali alla segreteria di una terza persona oppure mandare delle e-mail a terzi.

#### 4.1 Il Parlato

L’ interazione con un robot per mezzo della voce, cioè poter dare istruzioni o ricevere risposte tramite il parlato, costituisce uno degli obiettivi fondamentali nello sviluppo delle interfacce uomo-robot. Le interfacce basate sul parlato, fino a poco tempo fa, sono rimaste confinate tra le mura dei laboratori di ricerca, ma, non appena i primi robot hanno mosso i primi passi nel mondo reale, si sono presentate nuove opportunità di utilizzo.

Man mano che i robot si fanno più complessi e capaci di eseguire compiti sempre più sofisticati, il linguaggio naturale appare un’alternativa più che desiderabile alla selezione di un comando per mezzo di una tastiera o alla visualizzazione di menu su di uno schermo. In ogni caso non è detto che il parlato venga considerato il mezzo di comunicazione ideale in ogni circostanza: in molti casi sono di gran lunga preferibili dispositivi di interazione “vecchio stile”, come nei casi di tele-operazione (utilizzando dei joystick) oppure quando sia necessario specificare al robot un certo obiettivo da raggiungere su una mappa (selezionando con un mouse su uno schermo) ed in tutti gli altri casi in cui sono coinvolti strumenti di uso comune, come tagliaerba, aspirapolvere, ecc., per i quali si predilige, almeno fino ad oggi, l’uso di pulsanti e piccoli display.

E’ possibile individuare due categorie di situazioni tipiche, non necessariamente disgiunte, nelle quali può essere utilizzata con successo un’interfaccia vocale (*speech interface*):

- L’utente ha le mani o gli occhi impegnati;
- L’uso di dispositivi di input convenzionali è considerato inopportuno e quindi sconsigliato.

Tipiche situazioni che cadono nella seconda categoria si hanno nell’interazione con robot mobili di servizio, soprattutto in ambienti domestici, dove il robot è libero di muoversi e mal si presta a ricevere comandi e a fornire feedback attraverso dispositivi classici. Nel caso in cui il robot sia utilizzato come strumento di supporto da persone con necessità particolari sia hanno situazioni del primo tipo.



Figura 11a: KAMRO.

**4.1.1 La progettazione di interfacce basate sul parlato.** Il primo passo da compiere nella progettazione di un'interfaccia basata sul parlato concerne l'Elaborazione del Linguaggio Naturale (NLP, *Natural Language Processing*). Per poter instaurare una comunicazione bidirezionale si rendono necessarie tecniche di Comprensione del Linguaggio Naturale (NLU, *Natural Language Understanding*) e di Generazione di Linguaggio Naturale (NLG, *Natural Language Generation*). In particolare, il sistema deve essere dotato della capacità di comprendere comandi vocali per mezzo di tecniche di Riconoscimento del Parlato (*Speech Recognition*) per tradurre la frase parlata nella relativa rappresentazione testuale interna. Analogamente, sarà necessario fare uso della Generazione di Parlato (*Speech Synthesis*) per tradurre in voce le frasi che il robot dovrà rivolgere all'utente. Per maggiori informazioni riguardo alle tecnologie del parlato si veda [SpeechLinks](#).

Infine, una volta che il sistema robotico sia in grado di comprendere e produrre linguaggio naturale, si rende opportuno affrontare altre questioni. Oltre a tutte le difficoltà che si incontrano normalmente nella realizzazione di una componente di comprensione del linguaggio naturale, infatti, la vera sfida, per molti ricercatori, è la capacità di mantenere traccia del contesto corrente nel quale il robot viene usato. Inoltre, ad un robot di servizio che operi in un luogo pubblico, si richiede di interagire tramite un'interfaccia basata sul parlato con una molteplicità di utenti a ciascuno dei quali è permesso, per quanto possibile, di utilizzare modalità comunicative individuali.

**4.1.2 Stato dell'arte del parlato.** KANTRA è un'interfaccia basata sul parlato sviluppata presso l'Università di Karlsruhe e l'Università di Saarland, applicata ad un robot mobile dotato di due braccia chiamato KAMRO (Figura 11a) [Leangle et al., 1995]. L'approccio che è stato scelto è basato sul dialogo e affronta la questione dell'Interazione Uomo-Robot presentando quattro situazioni principali:

- specificazione del task;
- monitoraggio dell'esecuzione;
- spiegazione del recupero da errore;
- aggiornamento e descrizione della rappresentazione dell'ambiente.

Nella Figura 11b è riportata l'architettura del sistema KANTRA. Possono essere distinti tre moduli principali per l'accesso alla conoscenza necessaria all'interpretazione e alla generazione di testo:

- **Modulo di analisi:** l'input in linguaggio naturale deve essere tradotto da un parser in una rappresentazione semantica codificata in un linguaggio di rappresentazione della conoscenza;
- **Modulo di valutazione:** le frasi sono interpretate in relazione alla conoscenza del mondo interna del sistema intelligente; questo componente costituisce l'interfaccia vera e propria tra l'accesso in linguaggio naturale e il sistema autonomo. Dal sistema applicativo proviene l'informazione di feedback destinata al sistema di dialogo in contatto con l'utente;
- **Modulo di generazione:** l'informazione proveniente dal modulo di valutazione deve essere tradotta in frasi in linguaggio naturale in relazione al contesto situazionale.

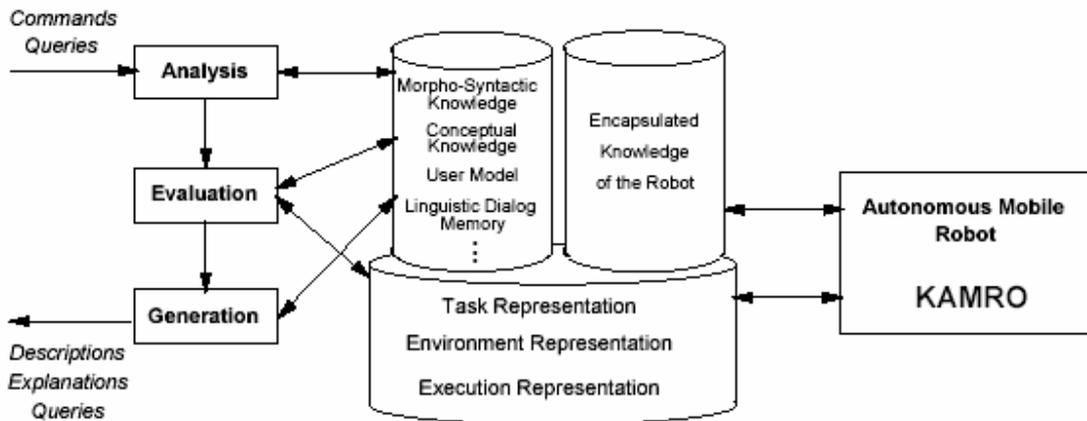
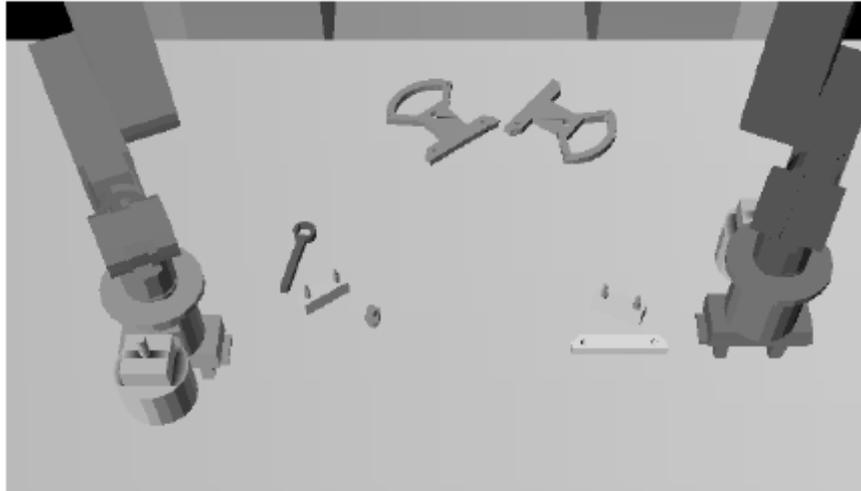


Figura 11b: Architettura di KANTRA.

Il modello dell'ambiente deve essere disponibile sia al robot KAMRO che all'interfaccia KANTRA: questo modello, per essere consistente con il modo reale circostante, necessita di aggiornamenti continui, soprattutto in seguito all'esecuzione di un comando. Per interpretare correttamente una frase pronunciata dall'operatore umano è necessario poter identificare gli oggetti menzionati nella frase: è possibile usare espressioni spaziali in modo da descrivere la locazione di un oggetto e poterlo quindi identificare [Herskovits, 1986]. KAMRO è stato progettato per ricevere istruzioni costituite da frasi corte: il sistema è in grado, in una certa misura, di completare l'informazione non specificata da solo ed eliminare alcune componenti di incertezza tramite i moduli dedicati al dialogo.

A fronte di determinati comandi impartiti dall'utente, inoltre, è possibile che si presentino situazioni di errore: un oggetto, ad esempio, può essere appoggiato soltanto se prima è stato afferrato dal manipolatore. Informazioni di questo tipo, sebbene intese dall'utente, spesso non sono esplicitamente inserite nella frase che costituisce il comando. Un altro problema può verificarsi nel caso in cui un operatore impartisca un numero di comandi del tipo "prendi" superiore al numero di manipolatori di cui è dotato il robot: se tra di essi non è inserito alcun comando del tipo "metti" il sistema non sarà capace di portare a termine il compito. Situazioni di questo tipo sono risolte dal modulo di pianificazione, responsabile della corretta esecuzione dei compiti.

Per capire come funziona l'interfaccia KANTRA supponiamo che il robot sia posto di fronte alla situazione rappresentata in Figura 11c: per analizzare una frase del tipo "Take the spacer between the shaft and the lever!" ("Prendi il distanziatore tra l'albero e la leva!") è necessario mettere in relazione il verbo con una operazione del robot e individuare gli oggetti menzionati nella frase. L'input è elaborato attraverso un parser basato su una grammatica di unificazione che produce la struttura di superficie della frase che, a sua volta, viene elaborata dalla componente di valutazione. In definitiva, il verbo "Take" viene messo in relazione con l'operazione del robot implicita "pick", mentre l'oggetto diretto della frase con l'oggetto con il quale il robot deve eseguire l'operazione. Nel caso in cui siano presenti diversi oggetti dello stesso tipo sul piano di lavoro, ad esempio, quattro "spacer" come nella situazione iniziale riportata in Figura 11c, le relazioni spaziali ("between the shaft and the lever") fornite con la frase devono essere interpretate rispetto all'informazione visiva e geometrica.



**Figura 11c:** KANTRA: situazione iniziale.

Uno degli “*spacer*” è l’oggetto da localizzare, la relazione è “*between*” mentre “*shaft*” e “*lever*” costituiscono gli oggetti di riferimento per la relazione spaziale. Per ogni “*spacer*” vengono create le cosiddette “proposizioni spaziali” che consistono della relazione, l’oggetto da localizzare e gli oggetti di riferimento, ad esempio: (*between spacer:1 shaft:1 lever:1*). Se anche gli oggetti di riferimento sono ambigui devono essere create proposizioni spaziali per ogni combinazione possibile, dopodiché, per ciascuna delle proposizioni, viene calcolato un “valore di applicabilità” per trovare quale dei quattro “*spacers*” è localizzato nella posizione più “tipica” rispetto a questa relazione spaziale. Nell’esempio riportato è “*spacer:2*”, quindi è necessario trasferire il comando “*pick spacer:2*” a KAMRO dove sarà interpretato dal modulo di esecuzione dei piani. Nella Figura 11d, il sistema ha eseguito l’operazione ed ha afferrato lo “*spacer:2*”.



**Figura 11d:** Il sistema ha afferrato uno “*spacer*”.

All’Università di Edimburgo, il robot mobile Godot [Theobalt et al., 2002] è stato usato come banco di prova per un’interfaccia tra un sistema robotico di navigazione a basso livello e un sistema simbolico basato sul dialogo.

Nell’ambito del progetto “Instruction Based Learning for Mobile Robots” del Robotic Intelligence Laboratory dell’Università di Plymouth, è stato realizzato un robot mobile in grado di ricevere istruzioni vocali su come spostarsi da un posto ad un altro all’interno di una città di miniatura [Bugmann, 2003]. Il sistema è in grado di ricevere istruzioni del tipo “*prendi la prima a sinistra*”, “*prosegui dritto finché...*”, “*se la strada è bloccata prendi quest’altra*”, ecc.

Il sistema Kairai è il risultato di un progetto di ricerca congiunto tra l'Università di New York ed il Tokyo Institute of Technology. Il sistema incorpora un certo numero di robot software 3-D con i quali è possibile conversare. Accetta comandi vocali, li interpreta ed esegue i relativi compiti in uno spazio virtuale [Tanaka et al., 2002]. Nello spazio virtuale possono trovarsi diversi robot software contemporaneamente: si consideri, ad esempio, la situazione rappresentata in Figura 12 dove sono presenti un cavallo, un pollo, un uomo di neve ed un cameraman, l'ultimo dei quali è invisibile ma manipola la sua telecamera per inquadrare lo spazio virtuale corrente. Oltre agli agenti robotici vi sono, per terra, due sfere rosse e due sfere blu. Per mezzo della voce dell'operatore, Kairai accetta un comando alla volta: il seguente esempio mostra come il sistema sia in grado di gestire situazioni anaforiche, ellittiche e vaghe.



**Figura 12:** Il sistema Kairai in azione.

**OPERATORE:** “Horse, push the sphere located in the left to the front of Chicken”

*Kairai decide quale delle sfere è quella specificata nel comando e dove il cavallo deve spostarla in funzione dello stato corrente dello spazio virtuale. In base all'interpretazione del comando il sistema fa eseguire al cavallo l'azione di spinta. Supponiamo che il colore della sfera sia blu.*

**OPERATORE:** “Push the red sphere, too.”

*Kairai decide qual è la sfera rossa che il cavallo deve spingere e lascia che l'agente robotico porti a termine il compito. L'espressione “the red sphere” è un esempio di espressione deittica.*

**OPERATORE:** “Chicken, push it, too.”

*Kairai risolve l'ambiguità anaforica di “push it” utilizzando il contesto, ovvero i comandi precedenti. In questo caso “push it” fa riferimento alla sfera rossa, che il cavallo ha spinto. Il pollo esegue l'azione.*

**OPERATORE:** “Further.”

*Sebbene non sia presente alcun soggetto, oggetto o verbo, Kairai “aumenta” queste parole ellittiche considerando il contesto accumulatosi attraverso il dialogo: il sistema fa spostare al pollo la sfera rossa un po' più avanti. Per mezzo della visualizzazione Kairai determina anche quanto lontano il pollo sposti la sfera. Questo è un classico problema di “vaghezza” del linguaggio naturale.*

Il progetto Hygeiorobot si è concluso nel 2001 ed ha coinvolto il National Centre for Scientific Research “Demokritos” di Atene e l'Università di Piraeus nella realizzazione di metodi e strumenti per il controllo e la navigazione di un robot mobile di servizio per l'assistenza ospedaliera. [Spiliotopoulos et al., 2001]. In questo contesto si sono rivelati particolarmente adatti sistemi capaci di comunicare attraverso la voce e in grado di gestire dialoghi, dal momento che il robot non trasportava una tastiera o altri dispositivi di interazione comuni e doveva essere utilizzato da persone con poca o nessuna esperienza informatica.



**Figura 13:** AIBO.

esplorare ambienti sconosciuti, eseguire compiti di trasporto e manipolazione e, soprattutto, di interagire e comunicare, anche con operatori inesperti, per mezzo del linguaggio naturale. Il robot è utilizzato come piattaforma di sperimentazione per l'uso in ambienti abitati da esseri umani: per questo motivo è stato dotato di un aspetto umanoide con altezza e peso analoghi a quelli di un essere umano.

Riportiamo di seguito un esempio di dialogo tra HERMES ed un operatore, che chiede al robot di prendere un bicchiere dalle sue mani ed appoggiarlo su un tavolo:

**OPERATORE:** "Hello, Hermes!"

**HERMES:** "Hello! What can I do for you?"

**OPERATORE:** "Take over glass."

*(Il robot distende il braccio e allarga la pinza che funziona come mano.)*

**HERMES:** "Please, hand over glass!"

*(l'operatore appoggia il bicchiere nella "mano" del robot che stringe delicatamente la pinza afferrando saldamente il bicchiere, quindi avvicina a sé il bicchiere e lo osserva.)*

**HERMES:** "What else can I do for you?"

**OPERATORE:** "Place it onto the table!"

*(A questo punto il robot descrive la serie di task che ritiene opportuno eseguire per portare a termine il compito assegnatogli e, infine, chiede conferma all'operatore...)*

**HERMES:** "I will search for the table! Then I will approach the table! Then I will place the glass onto the table! Do you want me to do these tasks?"

Tra i robot disponibili sul mercato spicca AIBO (Figura 13), un robot a quattro zampe realizzato dalla Sony in cui il parlato è parte integrante dell'interfaccia utente. AIBO è un robot autonomo casalingo da intrattenimento che i produttori ritengono dotato di intelligenza artificiale. Il suo comportamento simula quello di un cane nella sua abilità di camminare e giocare con funzioni integrate per simulare emozioni, istinti, apprendimento e crescita. La Sony non considera AIBO come il sostituto di un cane, ma un sistema studiato per favorire l'interazione uomo-robot.

Il robot umanoide HERMES (Figura 14) è stato creato presso la Bundeswehr University di Monaco di Baviera per dimostrare che cosa può essere fatto con i componenti di nuova generazione e i metodi di controllo avanzati [Bischoff et al., 2003]. HERMES (*Humanoid Experimental Robot for Mobile Manipulation and Exploration Services*) è in grado di



**Figura 14:** HERMES.

OPERATORE: “Yes, please!”

*(Il robot analizza l'ambiente nel quale deve operare, cerca il tavolo sul quale appoggiare il bicchiere ma ne trova due: si trova in una situazione di ambiguità che risolve interrogando l'operatore per chiedere di specificare il tavolo...)*

**HERMES:** “**I am searching for the table! I found two tables! Shall I approach the left or right table?**”

OPERATORE: “The left table!”

**HERMES:** “**I am driving towards the left table!**”

*(il robot calcola il percorso da seguire per raggiungere il tavolo di sinistra e si avvicina ad esso.)*

**HERMES:** “**I am placing the glass onto the table!**”

*(HERMES appoggia il bicchiere sul tavolo.)*

**HERMES:** “**What else can I do for you?**”

OPERATORE: “Nothing, thank you!”

Presso i Mitsubishi Electric Research Laboratories (MERL), le interfacce basate sul parlato vengono studiate da lungo tempo e le ricerche procedono su due binari paralleli: dispositivi incentrati sul parlato (*Speech-Centric Devices*) da un lato e interfacce basate sulla conversazione dall'altro. Per quanto riguarda i dispositivi incentrati sul parlato, al MERL sono interessati alla creazione di nuovi strumenti progettati a partire dal concetto di interfaccia basata sul parlato, a differenza di ciò che avviene nella maggior parte delle applicazioni in cui, molto spesso, si utilizzano comandi vocali come semplice alternativa alla pressione di tasti di selezione.

L'altro campo di ricerca riguarda le interfacce basate sulla conversazione. Come è noto, l'accuratezza dei migliori sistemi di riconoscimento del parlato non consentono di usare la voce come interfaccia primaria per compiti complessi. D'altra parte sono proprio questi compiti che possono trarre il massimo vantaggio dalle interfacce basate sul parlato. Al MERL sono in corso quindi studi che utilizzano i principi della teoria del discorso collaborativo umano per la costruzione di una struttura di base conversazionale sopra la quale poter implementare interfacce basate sul parlato.

Le ricerche più recenti mirano alla realizzazione di interfacce in grado di stabilire una conversazione tra una persona e un robot riguardo ad un determinato compito da raggiungere [Sidner et al., 2003] e coinvolgono, oltre allo studio della competenza linguistica e della conoscenza del mondo necessarie per una corretta esecuzione del compito, anche quello dei comportamenti che manifestino le intenzioni del robot sull'esecuzione e la finalizzazione dell'interazione (*engagement behaviors*).

## 4.2 I Gestii

Il riconoscimento dei gesti umani è un'area di ricerca in continuo sviluppo, soprattutto nell'ambito dell'interazione uomo-computer e uomo-robot ottenuta per mezzo di interfacce multimodali.

Diversi studi si sono interessati al ruolo dei gesti nell'interazione uomo-robot [Breazeal, 2001; Kanda et al., 2002]. Molti ricercatori si sono interessati dell'aspetto collaborativo e del dialogo tra uomo e robot [Fong et al., 2001] e sono stati effettuati diversi studi con agenti bidimensionali in grado di riprodurre gesti durante una conversazione [Cassell et al., 2000; Johnson et al., 2000] pur non incorporando nel sistema anche una fase di riconoscimento.

**4.2.1 Il Riconoscimento di Gest.** Per quanto riguarda le tecniche utilizzate per il riconoscimento dei gesti sono stati adottati diversi approcci: per mezzo di analisi di sequenze video [Cedras et al, 1995], mediante tecniche di riconoscimento in tempo reale [Bobick et al, 1996] in cui una “super-posizione” di immagini diverse è stata utilizzata per l'estrazione di caratteristiche e attraverso approcci che utilizzano Modelli di Markov Nascosti (HMM, *Hidden Markov Models*), come in [Starnier et al, 1995] in cui si riconosce il linguaggio dei gesti americano attraverso l'uso di guanti colorati. Altri studi hanno fatto uso di HMM per il riconoscimento di gesti in tempo reale: in [Rigoll et al, 1998] è presentato un sistema capace di riconoscere e distinguere diversi tipi di gesti con le mani, come il saluto e l'indicazione, e gesti della testa. Gli HMM del sistema sono stati addestrati per mezzo di un database di 24 gesti isolati eseguiti da 14 persone diverse.

Nei laboratori del Centre for Autonomous Systems (CAS) si occupano principalmente di interpretazione dei gesti. Il loro obiettivo è lo sviluppo di un sistema di comprensione dei gesti per robot di servizio domestici in grado di integrare le informazioni gestuali con il parlato. Per altri studi sull'interpretazione dei gesti da parte di robot si veda [Zelek et al., 2002].

**4.2.2 La Produzione di Gest.** Sempre presso il MERL sono studiati principi di interazione uomo-robot (Figura 15) in cui è possibile integrare riconoscimento e produzione di gesti in fase di conversazione e collaborazione [Sidner et al., 2002]. A questo scopo è stato realizzato un pinguino robot in grado di interagire con i visitatori del laboratorio: durante l'interazione il robot parla, gesticola ed è in grado di seguire con lo sguardo il proprio interlocutore. Inoltre, il pinguino robot è capace di interpretare i gesti del visitatore e determinare il suo grado di interesse all'interazione.



**Figura 15:** Pinguino robot al MERL.

### 4.3 Le Espressioni Facciali

La faccia di un essere umano può essere considerata una sorta di “finestra” affacciata sui meccanismi che governano le emozioni e la vita sociale [Gutta et al, 1996]. Per un uomo non è difficile riconoscere una faccia, anche in presenza di notevoli cambiamenti di aspetto derivati da diverse condizioni di visibilità, espressioni, età, acconciature diverse, ecc. Una macchina in grado di riconoscere un volto si presta ad innumerevoli applicazioni, come l'identificazione di criminali, il ritrovamento di bambini scomparsi, la verifica di carte di credito, il recupero di video-documenti e altro ancora.

Nell'ambito dell'Interazione Uomo-Robot, la capacità di riconoscere e produrre espressioni facciali permette al robot di allargare le proprie capacità comunicative, interpretando, da un lato, le emozioni che si dipingono sul volto del proprio interlocutore e, dall'altro, tradurre i propri intenti comunicativi in espressioni modellandole sulla propria faccia robotica.

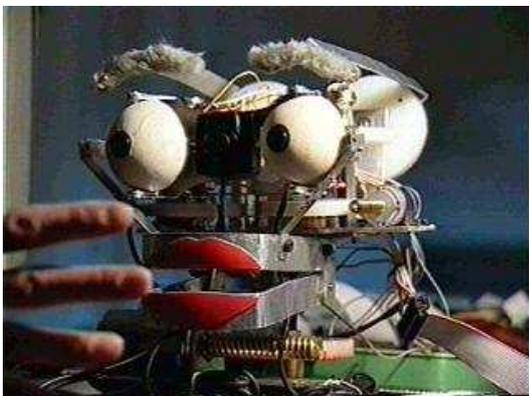
A rigore, questa modalità di interazione costituisce un sottoinsieme comunicativo della “cinesica” (cfr. 4.5), ovvero della disciplina che si occupa, in generale, dei movimenti del corpo, mimica facciale inclusa. Dal momento che le espressioni facciali giocano un ruolo particolarmente importante e costituiscono una sfida complessa di per sé, sono state poste come modalità interattiva distinta.

**4.3.1 Il Riconoscimento delle Espressioni Facciali.** Le tecniche utilizzate per il riconoscimento delle espressioni facciali sono molto simili a quelle utilizzate per il riconoscimento dei volti, dal momento che

entrambe le discipline richiedono l'individuazione di particolari caratteristiche facciali, o “*facial features*”, sul volto da analizzare. Alcuni studi, inoltre, hanno utilizzato tecniche per il riconoscimento delle caratteristiche facciali per il tracciamento dello sguardo (cfr. 4.4) nell'ambito dell'Interazione Uomo-Computer [Grant, 1998]. Infatti, una volta che le caratteristiche facciali sono state individuate, è possibile conoscere la direzione della testa e, successivamente, seguire la direzione dello sguardo.

Una tecnica che viene spesso utilizzata consiste nel fornire al robot un *template* deformabile da applicare alla regione catturata dall'immagine video dove presumibilmente è contenuta la faccia: questa “maschera” viene quindi utilizzata e modellata opportunamente per individuare e tracciare le caratteristiche facciali chiave come gli occhi, il naso e la bocca. Una volta ottenuto questo risultato, il sistema deve seguire i movimenti della testa, seguire i tratti somatici individuati e rilevare quei cambiamenti che contraddistinguono un'espressione facciale.

Esperimenti di riconoscimento di volti sono stati effettuati presso l'Artificial Intelligence Laboratory del MIT utilizzando la piattaforma robotica Kismet (cfr. 4.5) e affrontando il problema dell'identificazione della persona integrando l'informazione visiva con tecniche di riconoscimento della voce [Aryananda, 2001].



**Figura 16:** Aryan.

Hossein Mobahi, dell'Università di Tehran, ha progettato e realizzato Aryan (Figura 16), una faccia robotica in grado di riconoscere volti, interpretare le espressioni facciali e di produrle per esprimere emozioni [Mobahi, 2003]. Aryan costituisce la prima tappa per la realizzazione di un robot umanoide interattivo completo: per il momento Mobahi ha realizzato, utilizzando principalmente strumenti e componenti artigianali, la faccia robotica, le componenti elettroniche ed il software per la visione artificiale. Per il momento il robot è in grado di individuare autonomamente il volto di una persona, le mani e le caratteristiche facciali e di seguirne i movimenti tramite il suo sistema di visione attiva.

**4.3.2 La produzione di espressioni facciali.** K-Bot (Figura 17) è una testa robotica messa a punto da David Hanson presso l'“Institute for Interactive Arts and Engineering” dell'Università del Texas. Appartenente alla categoria dei cosiddetti “robot sociali”, K-bot può modulare ben 28 espressioni facciali, come gioia, rabbia e disprezzo, in risposta alle espressioni facciali dipinte sul viso dell'interlocutore umano, che il robot è in grado di individuare e interpretare con i propri “occhi”. Di fronte a una faccia allegra e soddisfatta, ad esempio, anche il viso di K-bot si illumina in un sorriso di contentezza: le informazioni visive raccolte dalle telecamere sono trasformate in comandi per i 24 attuatori, deputati al controllo degli occhi, dell'inclinazione della testa e del movimento della pelle artificiale che la riveste.



**Figura 17:** K-Bot.



**Figura 18:** Leonardo (Photo Copyright Sam Ogden.)

Leonardo (Figura 18) è un piccolo robot frutto della collaborazione del “MIT Media Laboratory”, che si è occupato della parte scientifica, e lo Stan Winston Studio di Los Angeles, che ne ha curato l’aspetto esteriore. Oltre ad essere stato utilizzato per numerosi studi, come la produzione di gesti, il tracciamento di oggetti e l’aptica (cfr. 4.6), Leonardo è stato dotato della facoltà di produrre espressioni facciali, grazie ai 32 gradi di mobilità contenuti nel solo muso. La mimica che è in grado di produrre è estremamente realistica e accattivante, specie se accompagnata dai movimenti del corpo, dimostrando quanto sia importante, dal punto della realizzazione di robot sociali “human-friendly”, la sinergia creativa della scienza e dell’arte.

#### 4.4 Il Tracciamento dello Sguardo

La direzione dello sguardo gioca un ruolo importante nell’interazione sociale umana e, in particolare, nell’identificazione del “focus” di attenzione di una persona. Durante la comunicazione faccia-a-faccia le persone si guardano, tengono d’occhio i movimenti delle labbra altrui, le espressioni facciali e seguono lo sguardo dell’interlocutore. L’informazione derivante dallo sguardo può essere sfruttata come un utile indizio comunicativo all’interno della progettazione di una interfaccia uomo-robot. Un sistema robotico dotato di questa funzionalità, in grado cioè di identificare dove una persona sta guardando e a che cosa sta prestando l’attenzione, sarà capace di capire, ad esempio, se questa persona sta rivolgendosi ad esso o ad un altro essere umano e sarà in grado di interpretare con più facilità a quale oggetto fa riferimento la persona che sta parlando.

Per identificare la direzione dello sguardo di una persona è necessario determinare l’orientazione della testa e quella degli occhi. Mentre l’orientazione della testa determina la direzione approssimativa dello sguardo, mediante l’orientazione degli occhi è possibile stabilire in modo preciso il punto verso il quale la persona sta effettivamente guardando.

Per seguire la direzione dello sguardo è necessario utilizzare dei sistemi di tracciamento dell’occhio (*eye trackers*). Molti dei metodi di tracciamento utilizzati si basano su tecniche intrusive come la misura della riflessione di luce infrarossa puntata verso l’occhio [Haro et al., 2000; Morimoto et al., 2000], la misurazione del potenziale elettrico della pelle intorno all’occhio (elettrooculografia) [Lusted et al., 1996] oppure mediante l’applicazioni di speciali lenti a contatto che facilitano il tracciamento dello sguardo.

Presso gli “Interactive Systems Laboratories” della Carnegie Mellon University, sono stati sviluppati sistemi di tracciamento non-intrusivi in grado di individuare e tracciare la direzione dell’occhio dell’utente in tempo reale non appena la faccia appare nel campo della telecamera, senza bisogno di alcuna illuminazione speciale o particolari segni di riferimento sulla faccia dell’utente [Stiefelhagen, 2001]. Nell’ambito del progetto INTERACT, una unità di



**Figura 19:** Rilevamento di caratteristiche facciali per il tracciamento dello sguardo (CMU).

ricerca specifica si occupa del tracciamento della posa della testa per mezzo del riconoscimento di caratteristiche facciali. La posa della testa è descritta tramite una matrice di rotazione: i nove parametri della matrice possono essere calcolati trovando le corrispondenze tra punti modello della testa ed i punti corrispondenti sull'immagine catturata dalla telecamera. Per stimare la posa della testa, in definitiva, è necessario trovare e seguire sei punti caratteristici sull'immagine, relativi agli occhi, le narici e gli angoli della bocca (Figura 19).

#### 4.5 Segnali Prossemici e Cinesici

Modalità di comunicazione più sofisticate (di solito classificate come “non-verbali”) sono la “prossemica” e la “cinesica”. La prossemica concerne la distanza tra gli interlocutori, la variazione della quale può fornire un utile indizio circa la disponibilità o la reticenza alla conversazione. Esperimenti di prossemica sono stati condotti presso il MIT con il robot Kismet (cfr. 4.3.1) [Breazeal et al., 2000] riportato in Figura 20. La testa robotica reagisce alla distanza del proprio interlocutore: quando questo si avvicina troppo invadendo il suo “spazio personale”, il robot si ritrae per segnalare il proprio disagio.

La kinesica è la modalità che riguarda il movimento e l'assunzione di posizioni: si occupa dei gesti compiuti utilizzando una o più parti del corpo ed in particolare dell'uso delle mani, della mimica facciale e della postura, ovvero della posizione dell'intero corpo e degli atteggiamenti motori.

Questa gestualità (più o meno inconscia) che si produce durante una interazione può costituire, se opportunamente interpretata, una preziosa fonte aggiuntiva di informazione [Ogden et al., 2000]. Tra i movimenti che è possibile individuare nell'ambito di un'interazione riportiamo:



**Figura 20:** Kismet.

- con le mani e le braccia: indicare con l'indice disteso, segnalare un “alt” con la mano, agitare la mano per salutare, sfregarsi le mani, incrociare le braccia, piantare i pugni sui fianchi;
- con la bocca: sorridere, serrare le labbra, digrignare i denti, sollevare un angolo della bocca, umettarsi le labbra;
- con il capo: roteare la testa per dire di “no”, annuire, scuotere la testa per esprimere perplessità, inchinare la testa di lato;
- con il corpo e le gambe: fare un inchino, spostare il peso da una gamba all'altra, passeggiare nervosamente, accavallare le gambe;
- con lo sguardo: strizzare l'occhio, spalancare o sbarrare gli occhi, inarcare un sopracciglio, aggrottare la fronte.

#### 4.6 L'Aptica

Tutto ciò che concerne il senso del tatto può essere classificato sotto la voce “aptica”, che possiamo definire come lo studio dell'acquisizione dell'informazione e della manipolazione attraverso il tatto. Tra le discipline principali che riguardano l'aptica figurano: l'aptica umana, le interfacce aptiche e l'aptica robotica.

Per quanto concerne l'Interazione Uomo-Robot, lo studio di interfacce aptiche è cominciato nel 1964, con la realizzazione di sistemi di tele-manipolazione di materiali chimici e nucleari [Goertz, 1964]. La realizzazione di interfacce aptiche più sofisticate, utilizzabili nell'interazione con robot "sociali", è strettamente vincolata alla conoscenza dell'aptica umana: la comprensione delle abilità percettive, motorie e cognitive dell'utente umano sono indispensabili per la realizzazione di un'interfaccia aptica uomo-robot funzionale ed efficiente. Dal momento che il nostro sistema aptico è principalmente composto da sistemi cutanei (relativi alla pelle come sensore di tocco) e cinestetici (relativi alla sensazione di movimento) i dispositivi di interazione aptici possono essere classificati come tattili oppure cinestetici, in relazione all'uso che ne viene fatto.

Presso il MIT, al "Laboratory for Human and Machine Haptics" (informalmente conosciuto come il "Touch Lab"), sono studiati i principi generali utilizzati da esseri umani e robot per esplorare, rappresentare ed interagire con gli oggetti. Gli obiettivi delle ricerche condotte al Touch Lab includono la comprensione dell'aptica umana, lo sviluppo di aptica robotica ed il potenziamento dell'interazione uomo-macchina in sistemi di teleoperazione e realtà virtuale.

Sempre presso il MIT, il robot Leonardo (cfr. 4.3.2) è stato rivestito di una soffice pelle sintetica capace di percepire e localizzare la pressione (Figura 21). La densità dei sensori sparsi sul corpo del robot varia in funzione della frequenza con la quale una certa area entra in contatto con gli oggetti e le persone: ad esempio, maggiore sulle mani e minore sulla schiena. Inoltre, una rete distribuita di piccoli elementi di elaborazione è in preparazione per essere posta sotto lo strato di pelle per acquisire ed elaborare i segnali sensori.

All' "Haptic Exploration Laboratory", della John Hopkins University, si occupano di aptica robotica e interfacce aptiche uomo-macchina. Per quanto riguarda l'aptica robotica, vengono studiati meccanismi che permettano ai robot di esplorare il mondo per mezzo del tatto, utilizzando speciali dita robotiche dotate di sensori [Okamura et al., 2001]. Presso il laboratorio, inoltre, sono utilizzate interfacce aptiche per dotare gli ambienti virtuali e di teleoperazione del senso del tocco.



**Figura 21:** Leonardo e gli esperimenti di aptica.

#### **4.7 Interfacce Multimodali**

Come appare evidente studiando l'Interazione Uomo-Robot, nella maggior parte dei casi l'interfaccia utilizzata per comunicare con un sistema robotico è ottenuta combinando varie modalità di interazione, è quindi un'interfaccia "multi-modale".

Le stesse modalità, inoltre, sfumano spesso l'una nell'altra e richiedono tecniche molto simili, come abbiamo visto a proposito della cinesica, le espressioni facciali, i gesti ed il tracciamento dello sguardo.

Per capire l'importanza della multimodalità è sufficiente analizzare i processi comunicativi umani, e notare come utilizzino, praticamente insieme e spesso contemporaneamente, tutte le modalità di interazione.

Nell'ambito di una qualunque conversazione, ad esempio, sono quasi sempre coinvolti:

- la voce, nelle sue variazioni prosodiche e soprasedimentali (tono, intonazione, cadenza, ecc.);
- lo sguardo, sia nel momento in cui viene rivolto verso l'interlocutore sia quando viene distolto da esso;
- i movimenti delle mani, del corpo, la postura e la distanza dall'interlocutore, ovvero la cinesica e la prossemica;
- le espressioni facciali.

La lista riportata include praticamente tutte le modalità di interazione introdotte in questo capitolo, a riprova di quanto sia indispensabile, al fine di realizzare un robot sociale, progettare un'interfaccia multi-modale capace di incorporare quante più modalità possibili per dotare il robot di un ampio canale di comunicazione.

Ai fini pratici, comunque, è spesso sufficiente realizzare interfacce che combinano i gesti col parlato e, in effetti, la maggior parte delle ricerche svolte nell'ambito della multimodalità si muovono in questa direzione.

Elementi deittici come “questa sedia”, “quel tavolo” oppure “lui” sono incomprensibili senza alcuna informazioni aggiuntiva, in modo analogo lo sono elementi direzionali come “laggiù”, “lì vicino”, ecc. Un comando come “vai là” è ambiguo se non accompagnato da un gesto appropriato che indichi un luogo preciso da raggiungere nell'ambiente operativo. Inoltre, comandi come “gira a sinistra di 20 gradi” possono confondere se accompagnati da gesti inappropriati o contraddittori. Sono perciò necessarie interfacce che gestiscano le ambiguità del linguaggio naturale e gesti appropriati, inappropriati o contraddittori.

Riportiamo di seguito alcuni studi e applicazioni in cui la progettazione di un'interfaccia multi-modale riveste un ruolo primario.

All' “Interaction and Presentation Laboratory” (IPLab) di Stoccolma, è in corso il progetto CERO [Hüttenrauch et al., 2002; Severinson-Eklundh et al., 2003]. Le ricerche sono volte ad analizzare come

le persone possono utilizzare un robot nella loro vita quotidiana e mirano all'indagine degli aspetti sociali e collaborativi che derivano dall'interazione con un robot. Il robot mobile di servizio realizzato nell'ambito del progetto (Figura 22a) è stato principalmente progettato per assistere utenti diversamente abili trasportando piccoli oggetti all'interno di un ufficio. Grande risalto è stato dato allo studio dell'interfaccia multi-modale uomo-robot, realizzata combinando l'uso del parlato con semplici gesti prodotti da un piccolo animatrone posto sulla piattaforma superiore del robot (Figura 22b).



**Figura 22a:** CERO.



**Figura 22b.**

Presso il “Navy Center for Applied Research in Artificial Intelligence” è stata sviluppata un'interfaccia per robot mobili che combina gesti e linguaggio naturale [Perzanowski et al., 1998]. L'interfaccia fa uso di tecniche di comprensione del linguaggio naturale e risolve alcune delle ambiguità del linguaggio sfruttando l'input proveniente dai gesti.

In [Bray et al., 2002] è presentata una tecnica in cui si combina il riconoscimento di gesti con il parlato. Per mezzo di una video camera e un riconoscitore vocale, è stato dimostrato come la

comprensione del sistema incrementi in modo considerevole. Per il riconoscimento dei gesti l'utente deve utilizzare un guanto colorato: il sistema estrae la velocità del centro di gravità della mano e apprende un Modello di Markov Nascosto relativo al determinato gesto. La combinazione di parlato e gesti è ottenuta per mezzo della teoria Bayesiana.

## 5 L'Interazione Uomo-Robot per la Cura di Anziani e Diversamente Abili.

Tra le numerose tipologie di robot sociali realizzati e in fase di sviluppo, di particolare importanza sono i sistemi robotici destinati al supporto delle persone con problemi motori e di salute, come gli anziani e le persone diversamente abili. Sempre nell'ambito del progetto RoboCare, nel quale si inquadra il presente lavoro, sono stati condotti studi ed effettuati sondaggi d'opinione [Giuliani et al., 2003] per valutare le condizioni di accettabilità di robot inseriti nell'ambiente domestico. Particolare riguardo è stato riservato alle persone anziane, il cui numero aumenta costantemente [Davies, 1999], per le quali l'uso di robot di servizio per l'assistenza sanitaria può costituire un notevole miglioramento della qualità della vita. Come evidenziato in alcuni studi, inoltre, la maggior parte delle persone anziane bisognose di assistenza preferisce restare a casa propria piuttosto che essere ospitata in una casa di cura: la presenza di un robot di servizio per uso domestico, in molti casi, può venire incontro a questa necessità.

Riportiamo di seguito due progetti di robot per l'assistenza a persone anziane e diversamente abili nei quali, proprio per le particolari necessità psico-fisiche dell'operatore umano, l'interfaccia uomo-robot, soprattutto nella sua modalità di interazione vocale, riveste grande importanza.

### 5.1 Progetto "RobChair".

Nell'ambito del progetto RobChair (*Robot and Wheelchair*), dell' "Institute of Systems and Robotics" della Coimbra University, è stata realizzata una sedia a rotelle robotica (Figura 23a), dotata di una interfaccia multimodale, destinata alle persone anziane e con problemi motori gravi (paraplegia, tetraplegia, ecc.) che trovano difficoltà nell'uso di normali sedie a rotelle motorizzate [Pires et al., 2002]. La funzione primaria della sedia robotica è quella di assistere gli operatori in modo da migliorare l'accessibilità della sedia stessa e ridurre le difficoltà nella guida. Per raggiungere questi obiettivi è stato implementato un modulo di navigazione e un'interfaccia uomo-robot basata sulla voce (Figura 23b).

Il parlato è senza dubbio il mezzo di comunicazione più naturale per gli esseri umani ed è considerato ideale per utenti con gravi limitazioni motorie. Nonostante questo, un'interfaccia basata sul parlato, utilizzata in questo caso per far muovere una sedia a rotelle motorizzata, non è sufficiente, da sola, per poter condurre l'operatore all'interno di un ambiente dinamico e irto di ostacoli.

Per mezzo dell'interfaccia vocale, infatti, è possibile impartire a RobChair semplici comandi vocali come "avanti", "indietro", "stop", "sinistra", "destra", "rallenta", ecc. La voce dell'utente è catturata da un microfono "a cuffia" ed elaborata per mezzo di un modulo di riconoscimento vocale. Il modulo di navigazione, a partire dal comando e dalle informazioni sensorie dell'ambiente, calcola infine la migliore traiettoria lungo la quale far muovere la sedia a rotelle.

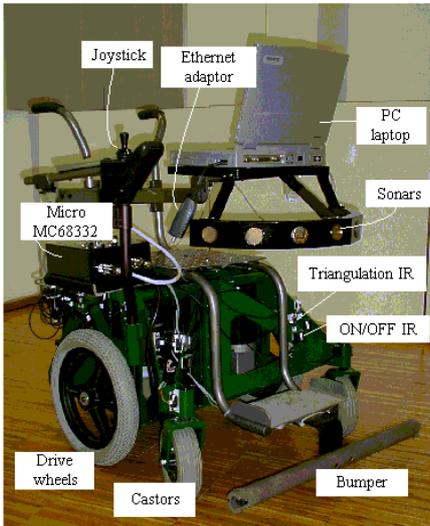


Figura 23a: RobChair.

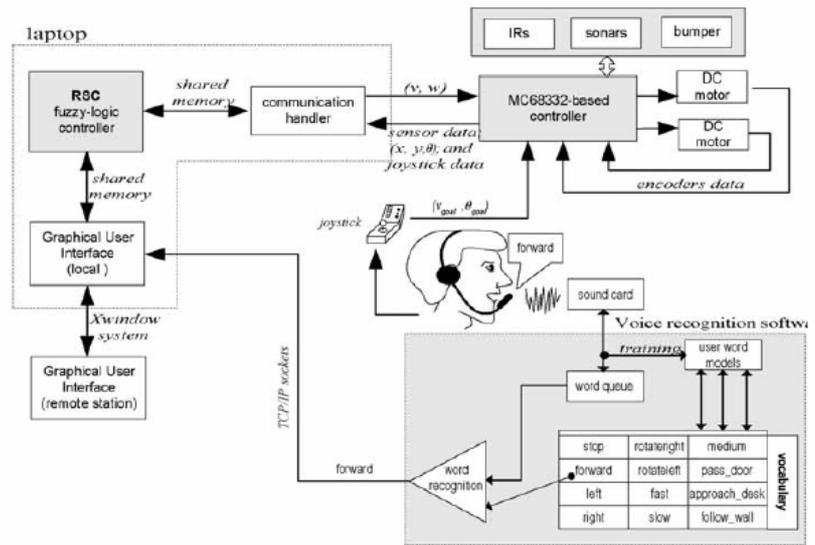


Figura 23b: Architettura del sistema RobChair.

## 5.2 Progetto “NurseBot”

Il progetto “Personal Robotic Assistants for the Elderly” (NurseBot) è nato da un’iniziativa interdisciplinare e multi-universitaria allo scopo di investigare l’applicazione di tecnologie robotiche all’assistenza delle persone anziane [Pollack et al., 2002].

L’obiettivo del progetto è quello di sviluppare un robot mobile di servizio, per uso personale, che assista quotidianamente persone anziane affette da disordini cronici. E’ già in fase di realizzazione un robot autonomo che “vive” nella casa privata di una persona anziana malata cronica.

Il robot costituisce una piattaforma di ricerca per diverse sperimentazioni:

- “Reminding” intelligente. Molti pazienti anziani sono costretti a rinunciare ad una vita indipendente perché dimenticano di prendere medicine, andare al bagno, bere oppure andare dal medico. Il progetto esplora la possibilità di fornire al paziente un robot che lo segua e gli ricordi impegni e scadenze.
- Telepresenza. Gli operatori sanitari possono utilizzare il robot per stabilire una “tele-presenza” ed interagire direttamente con i pazienti a distanza: in questo modo molte visite mediche diventano superflue.
- Raccolta dati e sorveglianza. Una vasta gamma di situazioni di emergenza (come certi problemi cardiaci) possono essere evitate con una sistematica raccolta di dati.
- Manipolazione mobile. Tra le ragioni che costringono le persone anziane a rinunciare alla propria indipendenza spiccano i problemi alle articolazioni delle mani, come l’artrite. Per ovviare al problema di manipolare oggetti (frigoriferi, lavatrici, forni, ecc.) è possibile utilizzare un manipolatore mobile semi-intelligente capace di integrare la propria forza robotica con i sensi e l’intelletto di una persona.

- Interazione sociale. Un gran numero di persone anziane sono costrette a vivere da sole e sono quindi private di contatti sociali. Il progetto cerca di esplorare in che modo i robot possono svolgere determinate funzioni sociali.

Un prototipo di questo robot, battezzato Pearl (Figura 24a), è stato realizzato e messo alla prova presso il Longwood Retirement Community a Oakmont, Pennsylvania (Figura 24b). Il robot è capace, al momento, di eseguire due funzioni principali:

- ricordare alle persone di prendere medicine, andare in bagno, ecc.;
- guidarle nel loro ambiente.

Pearl è equipaggiato con un sistema di guida differenziale, due PC Pentium “on-board”, rete wireless ethernet, un *laser range finder* SICK, sensori sonar, microfoni per il riconoscimento vocale, altoparlanti per la sintesi vocale, display grafici sensibili al tocco, una testa robotica ed un sistema di



**Figura 24a:** Pearl.



**Figura 24b:** Esperimenti con Pearl.

telecamere stereo. Un’attenzione particolare è stata messa nel creare l’aspetto estetico di Pearl, in special modo nel design della testa.

Dal punto di vista del software, il robot è dotato di un sistema di navigazione per robot mobili autonomi *off-the-shelf*, software di riconoscimento e sintesi vocale, software per la compressione e la cattura di immagini per *online video streaming*.

Il modulo software centrale di Pearl è costituito da un algoritmo probabilistico per il controllo ad alto livello e la gestione del dialogo. Questo modulo utilizza le informazioni provenienti dai moduli di più basso livello per selezionare le risposte e i comportamenti appropriati. L’architettura ad alto livello di Pearl è modellata come un POMDP (*Partially Observable Markov Decision Process*), per calcolare le azioni di controllo ottimali in condizioni di incertezza: la

decisione di controllo è basata su una “credenza” probabilistica su un certo numero di stati possibili. Questa distribuzione è definita in base ad una collezione di variabili di stato multi-valore:

- locazione del robot (approssimazione discreta);
- locazione della persona (approssimazione discreta);
- stato della persona (inferito dal riconoscitore vocale);
- Obiettivo di movimento (dove andare);
- Obiettivo di *reminder* (di che cosa informare l’utente);
- Obiettivo iniziato dall’utente (ad es., una richiesta di informazione).

In totale ci sono 516 stati possibili: l'incertezza nello stato corrente deriva prevalentemente dai moduli di localizzazione e dal sistema di riconoscimento del parlato. E' stata realizzata una versione gerarchica del POMDP che scompone il problema decisionale in un insieme di problemi più piccoli che possono essere risolti più efficientemente: lo spazio delle azioni risulta così partizionato e strutturato gerarchicamente (Figura 25a) e comprende "azioni astratte" (mostrate nei cerchi) introdotte per sussumere sottogruppi logici di azioni di più basso livello. In Figura 25b è riportato un esempio di dialogo in cui si mostra l'interazione tra il robot ed uno sperimentatore. In base al sistema di gestione dell'incertezza dei POMDP il robot interroga l'operatore per dei chiarimenti in tre occasioni diverse. Il numero di tali interrogazioni dipende da quanto è comprensibile e chiaro il parlato dell'operatore umano, elaborato dal sistema di riconoscimento della Sphinx. Nell'esempio, la colonna *Observation* riporta gli output del riconoscitore vocale, *True State* indica lo stato attuale del robot, *Action* riporta l'azione scelta dalla politica POMDP: in grassetto sono evidenziate le azioni di chiarificazione scelte a fronte di un segnale vocale ad alta incertezza. *Reward*, infine, è un indicatore che descrive l'appropriatezza dell'azione per il dato stato.

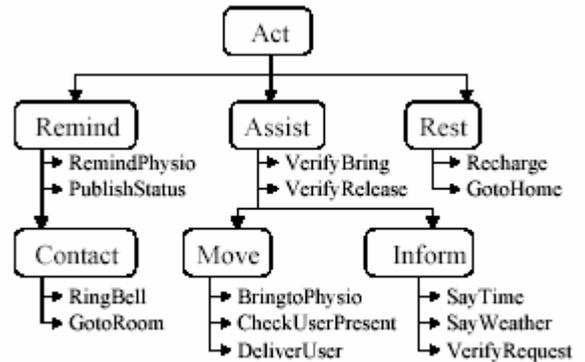


Figura 25a: Gerarchia di azioni (*Dialog Problem Action Hierarchy*).

<i>Observation</i>	<i>True State</i>	<i>Action</i>	<i>Reward</i>
pearl hello	request_begun	say_hello	100
pearl what is like	start_meds	<b>ask_repeat</b>	-100
pearl what time is it for will the	want_time	say_time	100
pearl was on abc	want_tv	ask_which_station	-1
pearl was on abc	want_abc	say_abc	100
pearl what is on nbc	want_nbc	<b>confirm_channel_nbc</b>	-1
pearl yes	want_nbc	say_nbc	100
pearl go to that pretty good what	send_robot	ask_robot_where	-1
pearl that that hello be	send_robot_bedroom	<b>confirm_robot_place</b>	-1
pearl the bedroom any i	send_robot_bedroom	go_to_bedroom	100
pearl go it eight a hello	send_robot	ask_robot_where	-1
pearl the kitchen hello	send_robot_kitchen	go_to_kitchen	100

Figura 25b: un esempio di dialogo tra Pearl ed un operatore.

## 6 Conclusioni.

Potremmo chiederci, a questo punto, se esista un' interfaccia uomo-robot "ottimale" e quali caratteristiche debba avere, ma riteniamo, da parte nostra, che ogni situazione richieda caratteristiche specifiche e differenti strategie di interazione.

La scelta dell'interfaccia in un sistema robotico, per sua natura molto complesso, deve essere il frutto di un'attenta valutazione di diversi fattori..

In primo luogo, la tassonomia descritta, soprattutto per quanto riguarda gli attori, il luogo e il tempo della comunicazione, impone delle restrizioni sull'interazione che sono molto stringenti, sia per motivi di funzionalità degli apparati meccanici che per la natura del processo comunicativo.

Il livello di interazione condivisa costituisce uno dei primi elementi da tener presente: mentre tra l'uomo ed i robot, interagenti sia a livello individuale che di team, l'interazione naturale è preferibile, tra i robot stessi, sempre operanti nello stesso ambiente, non ci sembra ragionevole pensare a modalità comunicative di tipo umano, pena la perdita di rapidità ed efficacia che vengono invece garantite utilizzando modalità e strumenti artificiali.

La copresenza dell'uomo e del robot nello stesso spazio di interazione – quindi *in praesentia* – rende possibili molti tipi di interazione perché in tali condizioni tutti i canali comunicativi possono essere sfruttati. L'interazione *in absentia* impone, invece, la rinuncia a molti di questi canali, in particolare i non verbali come, espressioni facciali, gestualità, tracciamento dello sguardo, ecc.

La non sincronicità dell'interazione fa perdere la consapevolezza in tempo reale delle reazioni dell'interlocutore che normalmente sono prodotte attraverso un feedback interazionale che è alla base della dinamicità del processo comunicativo stesso. In questo modo si perde la possibilità di recepire e quindi di produrre molti segnali comunicativi, reazioni mimiche e cinesiche, ma anche verbali, con i quali è possibile determinare l'andamento e, quindi, il risultato della comunicazione.

Un'altra classe di fattori da tener presente è legata alla natura ed agli scopi dei robot stessi. In genere, un robot è un artefatto costruito appositamente per risolvere qualche problema per e insieme all'uomo: il suo utilizzo in ambiente reale e per scopi specifici è quindi una caratteristica imprescindibile. La sua funzionalità deve essere quindi garantita e con successo indipendentemente dalle diverse situazioni in cui si trovi ad operare. L'interfaccia deve quindi essere parte della sua robustezza e fornire tutte le garanzie possibili per quanto riguarda la sua efficacia comunicativa che significa la copertura adeguata dei segni, della loro combinazione e del loro uso dinamico e variabile in una pluralità di ambienti.

Inoltre, per poter soddisfare molte esigenze di carattere ambientale ed ergonomico, i robot, soprattutto quelli sociali, hanno dimensioni ridotte nelle quali si cerca di integrare il maggior numero di abilità, quindi di strumenti e processi che implementano quelle abilità, ivi comprese quelle interazionali. E' necessario, quindi, che tutti i processi possano essere agevolmente integrati, per cui se ne richiede, in genere, l'ottimizzazione e la miniaturizzazione. Per quanto riguarda la scelta dell'interfaccia, allora, una caratteristica fondamentale che viene richiesta è costituita dal suo grado di affidabilità che dipende in larga misura anche dallo stato di maturità delle tecnologie con le quali è possibile implementarla.

Tra le diverse modalità di interazione che abbiamo introdotto, la prossemica, la cinesica e l'aptica sono ancora ad uno stadio iniziale e richiedono ulteriori studi. In ogni caso, il parlato sembra essere la modalità di interazione più adeguata nella maggior parte delle situazioni: inoltre, lo sviluppo di interfacce basate sul linguaggio naturale può trarre beneficio da una lunga tradizione di ricerca nel contesto dell'Interazione Uomo-Computer, Elaborazione del Linguaggio Naturale e Tecnologie del Parlato, per di più, alcuni prodotti commerciali sono già disponibili sul mercato e permettono di realizzare sistemi di dialogo dotati di un alto grado di robustezza e di affidabilità. La combinazione del parlato con altre modalità, come i gesti ed il tracciamento dello sguardo, inoltre, sembra cruciale per lo

sviluppo di robot sociali capaci di instaurare un dialogo e, per questo motivo, le interfacce multi-modali meritano maggiore interesse.

### **Bibliografia.**

[Agah, 2001] A. Agah. Human interactions with intelligent systems: research taxonomy. *Computers and Electrical Engineering*, vol. 27, pp. 71 – 107, 2001

[Ambrose et al., 2000] R. O. Ambrose, H. Aldridge, R. S. Askew, R. Burrige, W. Bluethman, M. A. Diftler, C. Lovchik, D. Magruder e F. Rehnmark. ROBONAUT: NASA's Space Humanoid, in *IEEE Intelligent Systems Journal*, Agosto 2000.

[Amigoni et al., 2002] F. Amigoni, V. Schiaffonati e M. Somalvico. A Theoretical Approach to Human-Robot Interaction Based on the Bipolar Man Framework. Proceedings of the “Eleventh IEEE International Workshop on Robot and Human Interactive Communication 2002 (ROMAN2002), IEEE Press, Berlino, Germania, 25-27 settembre 2002, pagg. 183-188.

[Aryananda, 2001] L. Aryananda. Online and Unsupervised Face Recognition for Humanoid Robot: Toward Relationship with People, in *Proceedings of the 2001 IEEE – RAS International Conference on Humanoid Robots*, 2001.

[Asimov, 1995] I. Asimov. The Complete robot - The Definitive Collection of Robot Stories. London: Harper Collins, 1995.

[Balch, 2002] T. Balch. Taxonomies of multirobot task and reward, in *Robot Teams*, Balch, T. e Parker, L.E., eds., Natick, MA: A K Peters, 2002, pp. 3 - 22.

[Bernsen et al, 1997] N. O. Bernsen, H. Dybkjær e L. Dybkjær. What Should Your Speech System Say. *IEEE Computer*, 30(12), dicembre, 1997.

[Bischoff et al., 2003] R. Bischoff e V. Graefe. HERMES – an Intelligent Humanoid Robot, Designed and Tested for Dependability. In: B. Siciliano and P. Dario (eds) *Experimental Robotics VIII, Proceedings of the 8th International Symposium ISER02*, volume 5 of Springer Tracts in Advanced Robotics (STAR), Springer, Heidelberg, 2003.

[Bobick et al, 1996] A. Bobick e J. Davis. An Appearance-Based Representation of Action. In Proc. IEEE Int. Conf. in Pattern Recognition (ICPR-96), pagg. 307-312, Vienna, Austria, Agosto 1996.

[Bolt, 1980] R. A. Bolt. „Put-That-There”: Voice and Gesture at the Graphics Interface. *Computer Graphics*, 14(3), 262 – 70, 1980.

[Bray et al., 2002] M. Bray, H. Sidenbladh e J. O. Eklundh. Recognition of gestures in the context of speech, International Conference on Pattern Recognition (ICPR02), Quebec City, Agosto 2002.

[Breazeal et al., 2000] C. Breazeal e B. Scassellati. Infant-like social interactions between a robot and a human caretaker, *Adaptive Behavior*, 8(1), (2000), pp. 49–74.

- [Breazeal, 2001] C. Breazeal. Affective interaction between humans and robots, Proceedings of the 2001 European Conference on Artificial Life (ECAL2001), Praga, Repubblica Ceca, 2001.
- [Cassell et al., 2000] J. Cassell, J. Sullivan, S. Prevost e E. Churchill, Embodied Conversational Agents. MIT Press, Cambridge, MA, 2000.
- [Cawsey et al, 1995] A.Cawsey, K.Binsted & R.Jones, “Personalised explanations for patient education”, in Proceedings of the 5th European Workshop on Natural Language Generation, 1995.
- [Cedras et al, 1995] C. Cedras e M. Shah. Motion-Based Recognition: A Survey. Image and Vision Computing, 13(2):129-155, 1995.
- [Coch, 1996] J. Coch, “Evaluating and comparing three text production techniques”, in Proceedings of COLING, 1996.
- [Dautenhahn et al., 1999] K. Dautenhahn e A. Billard. Bringing up robots or – The psychology of socially intelligent robots: From theory to implementation, in: *Proceedings of the Autonomous Agents*, 1999.
- [Davies, 1999] A. M. Davies. Aging and Health in the 21<sup>st</sup> century, in *Aging and Health: A Global Challenge for the Twenty-First Century*, Ginevra, Svizzera: World Health Organization.
- [de Rosis et al, 1996] F. de Rosis et al. “Generating Recipient-Centred Explanations About Drug Prescription”, *Artificial Intelligence in Medicine* 8, 1996.
- [DiSalvo et al., 2002] C. DiSalvo, F. Gemperle, J. Forlizzi, e S. Kiesler. All Robots Are Not Created Equal: Design and the Perception of Humanness in Robot Heads, *DIS2002 Conference Proceedings*, pp. 321-326.
- [Donnellan, 1966] K. Donnellan. Reference and Definite Descriptions. *Philosophical Review*, LXXV, 281-304, 1966.
- [Dudek et al., 2002] G. Dudek, M. Jenkin e E. Milios. A Taxonomy of Multirobot Systems, in *Robot Teams*, Balch, T. e Parker L.E., eds., Natick, MA: A K Peters, 2002, pp. 3 - 22.
- [Ellis et al., 1991] C. A. Ellis, S. J. Gibbs e G. L. Rein. Groupware: some issues and experiences. *Communications of the ACM*, 34(1): 39 – 58.
- [Fong et al., 2001] T. Fong, C. Thorpe e C. Baur. Collaboration, Dialogue and Human-Robot Interaction, 10th International Symposium of Robotics Research, Lorne, Victoria, Australia, novembre, 2001.
- [Fong et al., 2003] T. Fong, I. Nourbakhsh e K. Dautenhahn. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42:143--166, 2003.

- [Giuliani et al., 2003] M. V. Giuliani, M. Scopelliti, F. Fornara, E. Muffolini e A. Saggese. Human-Robot Interaction: How People View Domestic Robots, in: *Proceedings of the First RoboCare Workshop*, Roma, Ottobre 2003.
- [Goertz, 1964] R. C. Goertz, R.C. Manipulator systems development at ANL, *Proceedings of the 12th Conference on Remote Systems Technology*, ANS. Teleoperation Systems (1964). pp. 117-136.
- [Goldberg et al, 1994] E. Goldberg, N.Driedgar, R.Kittredge, “Using natural language processing to produce weather forecasts”, *IEEE Expert* 9, pp.45-53, 1994.
- [Grant, 1998] D. Grant. An Integrated Human Computer Interface using Eye Gaze Tracking and Facial Feature Recognition, project proposal, MSc Computational Intelligence, University of Plymouth, 1998.
- [Grice, 1975] H. P. Grice. Logic and Conversation. In: P. Cole & J. L. Moorgan (red.), *Syntax and Semantics — III: Speech Acts*, New York: Seminar Press, 1975.
- [Gutta et al, 1996] S. Gutta, J. Huang, I. Imam e H. Wechsler. Face and Han Gesture Recognition Using Hybrid Classifiers, Department of Computer Science, George Mason University, Fairfax, 1996.
- [Haro et al., 2000] A. Haro, M. Flickner e I. Essa. Detecting and Tracking Eyes By Using Their Physiological Properties, Dynamics, and Appearance, *IEEE CVPR 2000*, pagg. 163-168, 2000.
- [Herskovits, 1986] A: Herskovits. *Language and Spatial Cognition. An Interdisciplinary Study of the Propositions in English*, Cambridge, Londra: Cambridge University Press, 1986.
- [Horswill, 1995] I. Horswill. Analysis of adaptation and environment. *Artificial Intelligence*, 73(1-2), pp. 1-30.
- [Hüttenrauch et al., 2002] H. Hüttenrauch and K. Severinson-Eklundh. Fetch-and-carry with CERO: Observations from a long-term user study with a service robot, 2002.
- [Ishii et al., 1997] H. Ishii e B. Ullmer. Tangible Bits: Towards Seamless Interfaces Between People, Bits and Atoms. *Proceedings of the ACM Conference on Human Factors in Computing Systems*, 1997, pp. 234-241.
- [Johnson et al., 2000] W.L. Johnson, J. W. Rickel, J. W. e J.C. Lester. Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments, *International Journal of Artificial Intelligence in Education*, 11: 47-78, 2000.
- [Kanda et al., 2002] T. Kanda, H. Ishiguro, M. Imai. T. Ono e K. Mase. A constructive approach for developing interactive humanoid robots. *Proceedings of IROS 2002*, IEEE Press, NY, 2002.
- [Khan, 1998] Z. Khan. Attitudes towards Intelligent Service Robots, IpLab, Nada, Royal Institute of Technology, 1998.
- [Lorek, 2001] L. Lorek. "March of the A.I. Robots", *Interactive Week*, 30 Aprile, 2001.

- [Lusted et al., 1996] H. S. Lusted, R. B. Knapp. Controlling Computers with Neural Signals, *Scientific American*, ottobre 1996.
- [Milde et al., 1997] J. T. Milde, K. Peters e S. Strippgen. Situated communication with Robots: First International Workshop on Human-Computer Conversation. Bellagio, Italy, 1997.
- [Mobahi, 2003] H. Mobahi. Building an Interactive Robot Face from Scratch, Bachelor of Engineering Final Project Report, Azad University, Tehran-South Campus, Tehran, Iran, Maggio 2003.
- [Moore et al, 1995] J. Moore et al “An intelligent interactive system for delivering individualized information to patients” *Artificial Intelligence in Medicine* 7, 1995.
- [Morimoto et al., 2000] C. Morimoto, D. Koons, A. Amir e M. Flickner. Pupil Detection and Tracking Using Multiple Light Sources, *Image and Vision Computing*, Special issue on Advances in Facial Image Analysis and Recognition Technology, Vol.18, No.4, pp.331-335, 2000.
- [Murphy et al, 2001] R. R. Murphy e E. Rogers. Final Report for DARPA/NSF Study on Human-Robot Interaction.
- [Nunes et al., 2000] U. Nunes, R. Cortesao, J.L. Cruz e P. Coelho, Shared-Control Architecture: concepts and experiments. “Service Robotics – Applications and Safety Issues in an Emerging Market”, 14th European Conference on Artificial Intelligence, Berlino, 22 Agosto 2000.
- [Ogden et al, 1997] W. C. Ogden e P. Bernick. Using Natural Language Interfaces. In: M. Helander, T. K. Landauer e P. Prabhu (red.), *Handbook of Human-computer Interaction*, Amsterdam, Elsevier Science Publishers B.V., 1997.
- [Ogden et al., 2000] B. Ogden and K. Dautenhahn. Robotic Etiquette: Structured Interaction in Humans and Robots, in *Proceedings of SIRS2000, Symposium on Intelligent Robotic Systems*, Reading, UK, 2000, pp. 353-361.
- [Okamura et al., 2001] A. M. Okamura e M. R. Cutkosky. Feature Detection for Haptic Exploration with Robotic Fingers, *International Journal of Robotics Research*, Vol. 20, No. 12, pp. 925-938, 2001.
- [Paris et al, 1996] C. Paris e K. V. Linden. DRAFTER: An Interactive Support Tool for Writing Multilingual Instructions. *IEEE Computer*, Special Issue on Interactive NLP, July 1996.
- [Perzanowski et al., 1998] D. Perzanowski, W. Adamse e A. Schultz, (23-25 Marzo 1998). Communicating with a Semi-Autonomous Robot. In *Workshop on Integrating Robotics Research: Taking the Next Leap*. AAAI98 Spring Symposium Series. Menlo Park, CA: AAAI Press, pp. 55-59.
- [Pires et al., 2002] G. Pires e U. Nunes. A Wheelchair Steered through Voice Commands and Assisted by a Fuzzy-Logic Controller, in *Int. Journal of Intelligent and Robotic Systems*, 34(3), 301-314, Luglio 2002.
- [Pollack et al., 2002] M. Pollack et al. Pearl : Mobile Robotic Assistant for the Elderly, AAAI Workshop on Automation as Eldercare, agosto 2002.

- [Reiter et al, 1995] E. Reiter, C. Mellish & J. Levine, “Automatic generation of technical documentation”, *Applied Artificial Intelligence* 9, 1995.
- [Rigoll et al, 1998] G. Rigoll, A. Kosmala e S. Eickeler. High Performance Real-Time Gesture Recognition Using Hidden Markov Models. Department of Computer Science, Faculty of Electrical Engineering, Duisburg, Germania, 1998.
- [Severinson-Eklundh et al., 2003] K. Severinson-Eklundh, A. Green e H. Hüttenrauch. Social and collaborative aspects of interaction with a service robot, *Robotics and Autonomous Systems*, Special Issue on Socially Interactive Robots, vol. 42, no. 3-4, 2003.
- [Sidner et al., 2002] C. L. Sidner e M. Dzikovska. Engagement between Humans and Robots for Hosting Activities, *International Conference on Multimodal Interfaces*, October 2002
- [Sidner et al., 2003] C. L. Sidner, C. Lee e N. Lesh. The Role of Dialogue in Human Robot Interaction, Mitsubishi Electric Research Laboratories, giugno 2003.
- [Spiliotopoulos et al., 2001] D. Spiliotopoulos, I. Androutsopoulos e C. D. Spyropoulos. Human- Robot Interaction Based on Spoken Natural Language Dialogue, in *Proceedings of the European Workshop on Service and Humanoid Robots*, 2001.
- [Springer et al, 1991] S. Springer et al. “Automatic letter composition for customer service”, *Proceedings of the Conference on Innovative Application of Artificial Intelligence*, 1991.
- [Starner et al, 1995] T. Starner e A. Pentland. Visual Recognition of American Sign Language Using Hidden Markov Models. In *International Workshop on Automatic Face and Gesture Recognition*, Zurigo, Svizzera, 1995.
- [Stiefelhagen, 2001] R. Stiefelhagen, J. Yang e A. Waibel, Tracking focus of attention for human–robot communication, in *Proceedings of the International Conference on Humanoid Robots*, 2001.
- [Tanaka et al., 2002] H. Tanaka, T. Tokunaga e Y. Shinyama. Animated Agents that Understand Natural Language and Perform Actions, in *Proceedings of Lifelike Animated Agents (LAA)*, Tokyo, 2002.
- [Terdada et al., 2002] K. Terdada e T. Nishida. Active Artifacts: for New Embodiment Relation between Human and Artifacts, in M. Gini et al. (editors), *Intelligent Autonomous Systems 7*, pagg. 333-340, IOS Press, 2002.
- [Tews et al., 2003] A. D. Tews, M. J. Matari e G. S. Sukhatme. A Scalable Approach to Human-Robot Interaction, to appear in *Proceedings of the International Conference on Robotics and Automation (ICRA 2003)*, 2003.
- [Ullmer et al., 1997] B. Ullmer e H. Ishii. The metaDESK: Models and Prototypes for Tangible User Interfaces. *Proceedings of the ACM Conference on User Interface and Software Technology*, 1997, pp. 223-232.

[Yanco et al., 2002] H. A. Yanco e J. L. Drury. A Taxonomy for Human-Robot Interaction, in: *Proceedings of the AAAI Fall Symposium on Human-Robot Interaction*, AAAI Technical Report FS-02-03, Falmouth, Massachusetts, Novembre 2002, pp. 111-119.

[Zepek et al., 2002] J. Zepek, D. Bullock, S. Bromley e H. Wu, What the Robot Sees & Understands Facilitates Dialogue, *Human-Robot Interaction, 2002 AAAI Fall Symposium*, November 15-17, North Falmouth, Massachusetts, USA, 2002.

## Appendice

### Indirizzi Internet.

AIBO: <http://www.sony.net/Products/aibo>

Artificial Intelligence Laboratory: [www.ai.mit.edu](http://www.ai.mit.edu)

Aryan: [www.digibrain.org](http://www.digibrain.org)

ASIMO: [world.honda.com/ASIMO](http://world.honda.com/ASIMO)

CAS: [www.cas.edu.au](http://www.cas.edu.au)

Center for Work Technology and Organization: [www.stanford.edu/group/WTO](http://www.stanford.edu/group/WTO)

Godot: [www.ltg.ed.ac.uk/godot](http://www.ltg.ed.ac.uk/godot)

Haptic Exploration Laboratory: [www.haptics.me.jhu.edu](http://www.haptics.me.jhu.edu)

HELPMATE: [www.pyxis.com/products/newhelpmate.asp](http://www.pyxis.com/products/newhelpmate.asp)

HERMES: [www.unibw-muenchen.de/hermes](http://www.unibw-muenchen.de/hermes)

Humanoid Robotics Project: [www.mstc.or.jp/hrp](http://www.mstc.or.jp/hrp)

Institute for Interactive Arts and Engineering: [iiae.utdallas.edu/projects](http://iiae.utdallas.edu/projects)

Institute of Systems and Robotics: [www.isr.uc.pt](http://www.isr.uc.pt)

Instruction-Based Learning for Mobile Robots: [www.tech.plym.ac.uk/soc/staff/guidbugm/ibl](http://www.tech.plym.ac.uk/soc/staff/guidbugm/ibl)

Interactive Systems Laboratories: [www.is.cs.cmu.edu/mie](http://www.is.cs.cmu.edu/mie)

IPLab (The Interaction and Presentation Laboratory): [www.nada.kth.se/iplab](http://www.nada.kth.se/iplab)

JSK Laboratory: [www.jsk.t.u-tokyo.ac.jp](http://www.jsk.t.u-tokyo.ac.jp)

Leonardo: [robotic.media.mit.edu/projects/Leonardo/Leo-intro.html](http://robotic.media.mit.edu/projects/Leonardo/Leo-intro.html)

Project on People and Robots: [www.peopleandrobots.org](http://www.peopleandrobots.org)

MERL: [www.merl.com](http://www.merl.com)

MIT (Massachusetts Institute of Technology): [web.mit.edu](http://web.mit.edu)

MIT Media Laboratory: [www.media.mit.edu](http://www.media.mit.edu)

National Centre for Scientific Research “Demokritos”: [www.demokritos.gr](http://www.demokritos.gr)

Navy Center for Applied Research in Artificial Intelligence: [www.aic.nrl.navy.mil](http://www.aic.nrl.navy.mil)

NurseBot: [www-2.cs.cmu.edu/~nursebot/](http://www-2.cs.cmu.edu/~nursebot/)

RobChair: [www.isr.uc.pt/~urbano/robchair/robchair.html](http://www.isr.uc.pt/~urbano/robchair/robchair.html)

ROBONAUT: [robonaut.jsc.nasa.gov](http://robonaut.jsc.nasa.gov)

Robotic Intelligence Laboratory: [www.tech.plym.ac.uk/iq4bots](http://www.tech.plym.ac.uk/iq4bots)

SpeechLinks (CMU): [www.speech.cs.cmu.edu/comp.speech/Section6/speechlinks.html](http://www.speech.cs.cmu.edu/comp.speech/Section6/speechlinks.html)

Stan Winston Studio: [www.stanwinstonstudio.com](http://www.stanwinstonstudio.com)

Things that Think: [ttt.media.mit.edu](http://ttt.media.mit.edu)

Tokyo Institute of Technology: [www.titech.ac.jp](http://www.titech.ac.jp)

Touch Lab: [touchlab.mit.edu](http://touchlab.mit.edu)