



L'INTERAZIONE UOMO-ROBOT

HUMAN-ROBOT INTERACTION

AMEDEO CAPPELLI
EMILIANO GIOVANNETTI

SOMMARIO/ 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.

Human-Robot Interaction (HRI) is a constantly growing multidisciplinary area rich of cues for advanced researches and technology transfers. It plays a fundamental role in the development of robots that operate in an open environment and cooperate with humans. This task requires the development of techniques that allow inexperienced users to use their robots in an efficient and safe way, using an intuitive and natural interface. In this work, after an introduction to the fundamental issues concerning HRI, we will present the different possible interaction modalities between robot and man followed by a series of advanced interface applications for autonomous mobile robots.

Parole chiave: Robotica, Interazione Uomo-Robot, Interazione Uomo-Macchina.

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. Parallela-

mente 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.

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 in pieno tutti questi requisiti. Uno dei limiti allo sviluppo di questo tipo di robot "sociali" deriva dalla mancanza di interfacce appropriate che permettano un'interazione naturale, intuitiva e versatile, 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, presentando le varie tipologie di robot esistenti, una tassonomia dell'interazione uomo-robot e le varie modalità di interazione possibili nel processo di comunicazione tra robot ed esseri umani*.

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".

Ciò che ha spinto l'uomo a progettare e costruire i robot è la necessità di avere a disposizione dei mezzi che possano sostituirlo, in particolare nelle attività più onerose o rischiose. I primi veri robot sono apparsi negli anni cinquanta ad uso industriale [20]. 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 MIT¹ (Massachusetts Institute of Technology), che ha segnato, per molti, l'inizio dell'era dei robot.

Oggi i robot sono impiegati per diversi scopi: per compiti manifatturieri nell'industria (robot industriali) e per assolvere funzioni nell'ambito dei servizi (robot di servizio) (Figura 1).

2.1 Robot di servizio e robot sociali

I cosiddetti "robot di servizio", o "robot assistenti", sono ancora lungi dal raggiungere il potenziale economico dei propri cugini robotici ad uso industriale. Col termine "robot di servizio" si intende un robot che opera in piena o parziale autonomia per eseguire servizi utili per il bene degli esseri umani e delle cose, con l'esclusione delle operazioni manifatturiere. I robot di servizio comprendono:

- robot destinati ad operare in condizioni pericolose (spegnimento di incendi, disinnescamento di bombe, ispezione di siti pericolosi, manipolazione di materiali radioattivi, ecc.);
- robot per l'esplorazione di ambienti impervi (spazio, profondità oceaniche, ecc.);

* 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).

¹ Ai termini sottolineati corrispondono indirizzi internet riportati in appendice.

- robot per uso medico (operazioni chirurgiche o indagini endoscopiche);
- robot sociali:
 - robot domestici (pulizia della casa, giardinaggio, sicurezza, ecc.);
 - robot di assistenza per disabili e anziani (sedie a rotelle robotiche, robot mobili per l'assistenza alla deambulazione, ecc.);

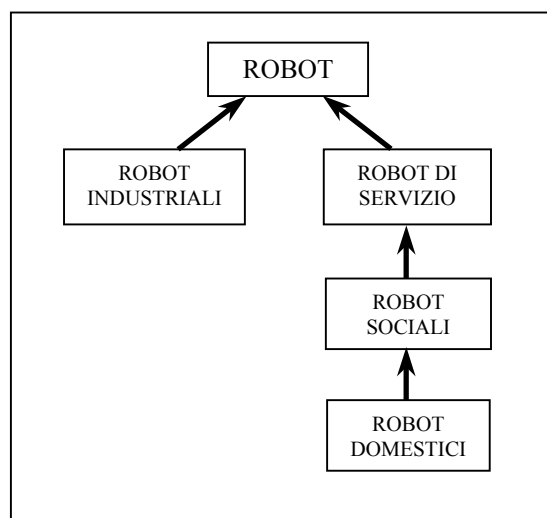


Figura 1: Le principali tipologie di robot.

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, effettuare ragionamenti e di comunicare con altri robot e con gli esseri umani.

Per interagire con un "robot sociale", ovvero un robot di servizio destinato ad operare in un ambiente sociale popolato da 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 sono [39]:

- 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 sosteneva che nel 1995 i robot di tipo diverso da quello industriale avrebbero

bero superato di gran lunga quelli impiegati in ambito manifatturiero [24].

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.

2.1.1 Aspetti sociali e collaborativi

Dal momento che i “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. Una importante distinzione nell’ambito dei robot sociali viene introdotta in [26], tra i cosiddetti robot sociali “collettivi” e i robot sociali “individuali”, intendendo i primi come facenti parte di un gruppo di robot la cui funzione è prevalentemente di squadra, mentre i secondi come agenti individuali più orientati all’interazione con l’uomo.

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. Nell’ambito del progetto italiano RoboCare [5] sono stati condotti studi ed effettuati sondaggi d’opinione per valutare le condizioni di accettabilità di robot inseriti nell’ambiente domestico [28]. Particolare riguardo è stato riservato alle persone anziane, il cui numero aumenta costantemente [18], 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.

Presso l’IPLab di Stoccolma 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 [53] 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. Tali studi, fondati sull’assunto che per studiare lo sviluppo cognitivo dei robot sia necessario considerare i “robot nella società”, utilizzano un approccio che considera le interazioni sociali come fondamentali per la progettazione di funzionalità cognitive sempre più ricche [17].

L’HRP-2P è un prototipo di robot umanoide realizzato nell’ambito dello Humanoid Robotics Project giappo-

nese. 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 2), 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 impieghi 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



Figura 2: HELPMATE.

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.



Figura 3: Studi sull’aspetto della testa (Project on People and Robots).

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 3) e sociale, per

andare incontro alle necessità e ai desideri delle persone [19].

L'unità di ricerca del progetto che si occupa di interazione uomo-robot mira alla comprensione 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 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 di come le persone si creino dei modelli mentali 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.

2.1.2 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 multimodali" (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 della bontà dei risultati raggiunti 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 4). 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 loro facce 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.

3. L'interazione uomo-robot

Due enti statunitensi (National Science Foundation e Department of Energy) hanno definito l'interazione uomo-robot di importanza strategica [47] 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 e di entrare in quelle domestiche. A tale scopo, 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." [41].

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



Figura 4: ASIMO.

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 raggiungere questo scopo è attingere alla grande esperienza già maturata nell'ambito della comunità dell'Interazione Uomo-Macchina (HCI, *Human-Computer Interaction*), dove sono state studiate le tendenze dello sviluppo tecnologico e il suo impatto sugli esseri umani.

3.1 Una tassonomia dell'HRI

Per meglio comprendere la natura dell'interazione uomo-robot riportiamo di seguito una particolare tassonomia [63] in cui è possibile inquadrare le possibili situazioni interattive. Altri studi sulle tassonomie hanno interessato l'interazione tra uomo e sistemi intelligenti e l'interazione con robot multipli [1, 22, 6].

La tassonomia presentata verrà costruita sulla base dei seguenti parametri:

- livello di autonomia vs. livello di intervento;
- rapporto tra numero di persone e numero di robot;
- livello di interazione condivisa tra squadre;
- supporto alle decisioni;
- criticità;
- tassonomia spazio-temporale.

Livello di autonomia vs. 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 [53] oppure robot uti-

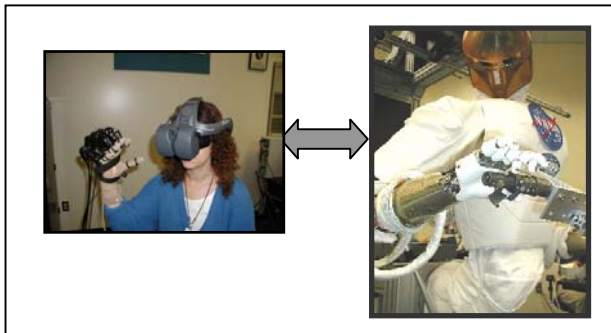


Figura 5: Telepresenza col sistema ROBONAUT.

lizzati per visite guidate come "Polly" che fa visitare il settimo piano dell'Artificial Intelligence Laboratory del MIT [35]. 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 5), della NASA, fornisce un esempio di sistema di "telepresenza" utilizzabile sia da terra che a bordo di una nave spaziale [2].

Rapporto tra numero di persone e 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.

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

Presentando le varie possibilità, cominciamo con l'interazione più semplice, ovvero quella costituita da un operatore umano e un singolo robot (Figura 6a). 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.

Il caso in cui un uomo interagisca con una squadra di robot (un gruppo di robot coordinato) corrisponde alla situazione riportata in Figura 6b: 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 6a: Un robot – un operatore.

Figura 6b: Squadra di robot – un operatore.

Diverso è il caso in cui un singolo operatore umano impartisca compiti differenti a robot differenti (Figura 6c): i singoli robot non sono tenuti necessariamente 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 assegna 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 6d è riportata la situazione interattiva in cui una

squadra 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.

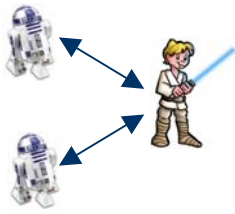


Figura 6c: Robot multipli – un operatore.

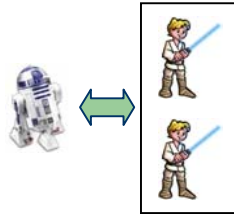


Figura 6d: Un robot – squadra di operatori.

Nella Figura 6e 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.

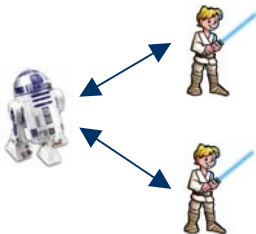


Figura 6e: Un robot – operatori multipli.

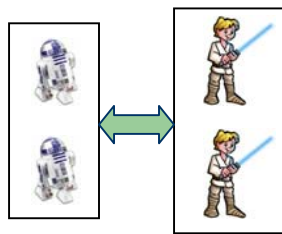


Figura 6f: Squadra di robot – squadra di operatori.

Gli ultimi tre casi di interazione da analizzare coinvolgono operatori umani multipli che interagiscono con robot multipli. In Figura 6f 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.

Il caso in cui una squadra di operatori impartisca ordini diversi a robot diversi corrisponde alla situazione riportata in Figura 6g. 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 6h). 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.

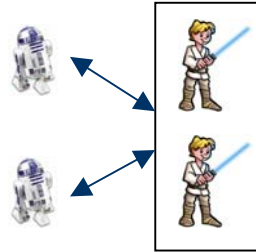


Figura 6g: Robot multipli – squadra di operatori.

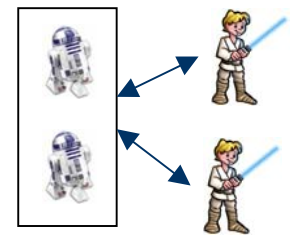


Figura 6h: Squadra di robot – operatori multipli.

Supporto alle decisioni. 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 disponibili sul robot:* indica l'ubicazione e la tipologia dei sensori disponibili sulla piattaforma robotica;
- *informazione dei sensori fornita agli operatori:* 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 di 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

causandogli gravi lesioni. All'estremo opposto, un cagnolino robotico giocattolo che non funziona bene non provoca alcun danno.

Tassonomia spazio-temporale. Questa tassonomia 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, come schematizzato con alcune esemplificazioni in Figura 7 [23].

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

Figura 7: Categoria tassonomica spazio-temporale con esempi.

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-macchina (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 [39] la maggior parte delle persone, interrogate a proposito di quali tipi di interazione uomo-robot prediligere, propende nettamente per 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 [21, 44]. 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 [31].

Per quanto riguarda le Interfacce in Linguaggio Naturale sono già stati effettuati diversi studi [48] e sviluppati diversi sistemi di dialogo telefonici [7] 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. Questo problema può essere risolto utilizzando più canali comunicativi, ad esempio 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 grazie anche ai 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 che l'utente e il robot possono condividere [10]. 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 multimodale.

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 [4].

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. Il caso in cui il robot sia utilizzato come strumento di supporto rientra fra le situazioni del primo tipo.

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 sia 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 8) [40]. 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.

L'architettura del sistema KANTRA prevede 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.

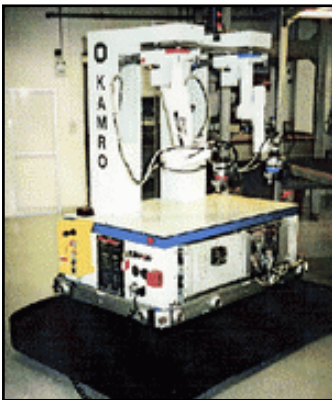


Figura 8: KAMRO.

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 [34]. 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.

All'Università di Edimburgo, il robot mobile Godot [60] 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 in miniatura [14]. 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 [59]. Nello spazio virtuale possono trovarsi diversi "softbot" (software robots) contemporaneamente: si consideri, ad esempio, la situazione rappresentata in Figura 9 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,



Figura 9: Il sistema Kairai in azione.

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.

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 [56]. 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 10: AIBO.

Tra i robot disponibili sul mercato spicca AIBO (Figura 10), 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 comporta-

mento 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 11) 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 [8]. HERMES (*Humanoid Experimental Robot for Mobile Manipulation and Exploration Services*) è in grado di 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. Ecco di seguito riportato un esempio di dialogo che Hermes è in grado di intrattenere con un operatore umano:

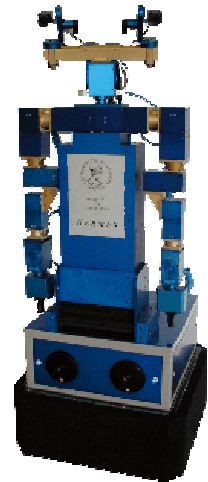


Figura 11: HERMES.

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?"

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 di questo tipo. 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 [55] 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 gesti

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

Diversi studi si sono interessati al ruolo dei gesti nell'interazione uomo-robot [13, 38]. Molti ricercatori si sono interessati dell'aspetto collaborativo e del dialogo tra uomo e robot [25] e sono stati effettuati diversi studi con agenti bidimensionali in grado di riprodurre gesti durante una conversazione [15, 37] pur non incorporando nel sistema anche una fase di riconoscimento.

4.2.1 Il riconoscimento di gesti

Per quanto riguarda le tecniche utilizzate per il riconoscimento dei gesti sono stati adottati diversi approcci: per mezzo di analisi di sequenze video [16], mediante tecniche di riconoscimento in tempo reale [9] 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 [57] 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 [52] 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.

Presso il Robotics Institute della Carnegie Mellon University è stato proposto un nuovo paradigma per il rilevamento e il tracciamento dei gesti e, in generale, di tutti i movimenti della parte alta del corpo umano [27]. Il problema è stato modellato come un problema di inferenza Bayesiano in una rete di Markov e sono stati proposti metodi per la cattura di vincoli di configurazioni umane da differenti angoli di visuale e per la combinazione di vincoli temporali (dinamici) e strutturali nella rete di Markov. Per inferire il movimento del corpo nella rete di Markov è utilizzata la *belief propagation* in grado di eseguire simultaneamente inferenze spaziali e temporali.

Sempre presso la Carnegie Mellon University è stata realizzata un'interfaccia basata sui gesti per il controllo di un robot mobile equipaggiato con un manipolatore [61]. L'interfaccia è basata sulla visione e permette di istruire il robot sia mediante i gesti che la postura del corpo. I gesti sono riconosciuti in due fasi: nella prima le immagini catturate dalle telecamere individuali sono mappate in un vettore che specifica le probabilità per le posture individuali. Nella seconda fase è utilizzato

l'algoritmo di Viterbi per far corrispondere dinamicamente il flusso di immagini con alcuni *template* gestuali temporali.

La capacità di riconoscere gesti da parte di un robot permette anche l'osservazione e la replicazione di attività manipolative nell'ambito dell'apprendimento automatico e della pianificazione dei compiti. Presso la Bielefeld University sono state svolte ricerche sull'interazione uomo-robot mediante l'addestramento del robot GRAVIS per la manipolazione di oggetti [43]. Il sistema combina un'interfaccia visuale per il riconoscimento dei movimenti della mano (Figura 12) con interfacce intelligenti per il riconoscimento del parlato mediante le quali è possibile istruire il robot a compiere complesse operazioni manipolative.

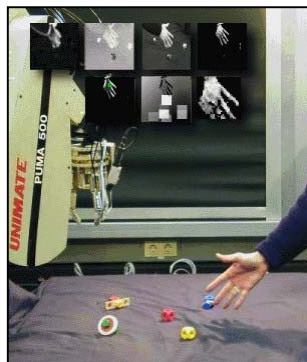


Figura 12: Apprendimento di attività manipolative.

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.

Tra i numerosi studi sull'interpretazione dei gesti da parte di macchine si vedano, in particolare [62, 64].

4.2.2 La produzione di gesti

Sempre presso il MERL sono studiati principi di interazione uomo-robot (Figura 13) in cui è possibile integrare riconoscimento e produzione di gesti in fase di conversazione e collaborazione [54]. 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 13: 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 gover-

nano le emozioni e la vita sociale [32]. 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 modellando 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 (*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-macchina [30]. 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.

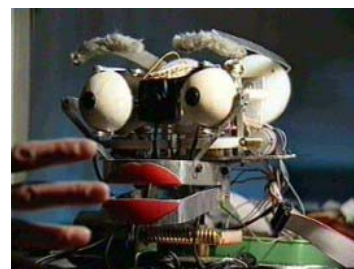


Figura 14: Aryan.

Hossein Mobahi, dell'Università di Tehran, ha progettato

e realizzato Aryan (Figura 14), una faccia robotica in grado di riconoscere volti, interpretare le espressioni facciali e di produrne per esprimere emozioni [45]. 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.

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 [3].

4.3.2 La produzione di espressioni facciali

K-Bot (Figura 15) 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 15: K-Bot.

Leonardo 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 quella 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 [33, 46], la misurazione del potenziale elettrico della pelle intorno all'occhio (elettrooculografia) [42] 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 [58]. Nell'ambito del progetto INTERACT, un'unità di ricerca specifica si occupa del tracciamento della direzione della testa per mezzo del riconoscimento di caratteristiche facciali. La direzione 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 cor-



Figura 16: Rilevamento di caratteristiche facciali per il tracciamento dello sguardo.

rispondenti sull'immagine catturata dalla telecamera. Per stimolare la direzione 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 16).

4.5 Segnali prossemici e cinesici

Modalità di comunicazione più sofisticate 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) [12] riportato in Figura 17. La testa robotica reagisce alla distanza del

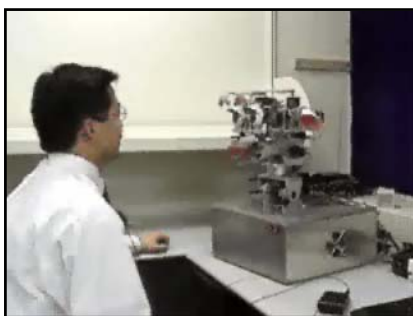


Figura 17: Kismet.

proprio interlocutore: quando questo si avvicina troppo invadendo il suo "spazio personale", il robot si ritrae per segnalare il proprio disagio.

La cinesica 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 [49]. Tra i movimenti che è possibile individuare nell'ambito di un'interazione riportiamo:

- 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 [29]. 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 18). 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 sensoriali.

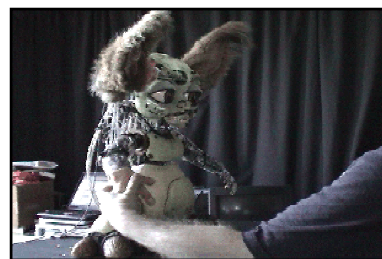


Figura 18: Leonardo e gli esperimenti di aptica.

All' "Haptic Exploration Laboratory", della John Hopkins University, si occupano di aptica robotica e in-

terfacce 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 [50]. Presso il laboratorio, inoltre, sono utilizzate interfacce aptiche per dotare gli ambienti virtuali e di teleoperazione del senso del tatto.

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 multimodale.

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 soprasegmentali (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 multimodale capace di incorporare quante più modalità possibili per dotare il robot di un ampio canale di comunicazione e per permettere di interpretare correttamente alcuni comandi che gli sono stati rivolti.

Elementi deittici come “questa sedia”, “quel tavolo” oppure “lui” sono incomprensibili senza alcuna informazione 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 dei gesti appropriati, inappropriati o contraddittori e che tengano conto

dell'efficacia della combinazione delle varie modalità comunicative che abbiamo visto.

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.

Un gran numero di studi teorici sono volti all'indagine della multimodalità, non solo nell'ambito dell'interazione uomo-robot ma, più in generale, nel campo dell'interazione uomo-macchina.

Presso la Tilburg University il programma di ricerca “Communication and Cognition” si occupa dell'analisi, della valutazione e della progettazione di informazione e comunicazione multimodali in contesti di diverso tipo. In particolare, il programma mira alla creazione di conoscenza e competenza necessarie per la gestione degli strumenti tecnologici di nuova generazione in grado di fornire interazioni di tipo multimodale tra l'uomo e la macchina. L'uso effettivo di questi strumenti richiede risposte precise a domande di questo tipo: in che modo le persone utilizzano ed elaborano informazione presentata mediante modalità differenti; come è possibile determinare la combinazione ottimale di modalità in relazione a particolari condizioni comunicative, compiti e utenti; come esplicitare teoricamente e rilevare sperimentalmente i pro e i contro di ciascuno dei sistemi modali coinvolti, l'interazione tra di essi e la suddivisione dei loro compiti nella



Figura 19a: CERO.

trasmissione dei messaggi. Domande di questo tipo oltre a costituire il tema centrale del progetto della Tilburg University sono fondamentali in generale nel trattamento della multimodalità all'interno dell'interazione uomo-robot.

Riportiamo di seguito alcuni studi e applicazioni in cui la progettazione di un'interfaccia multimodale riveste un ruolo primario nella comunicazione tra uomo e robot.



Figura 19b: animatrone di CERO.

All' “Interaction and Presentation Laboratory” (IPLab) di Stoccolma, è in corso il progetto CERO [36, 53]. 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 19a) 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 multimodale uomo-robot, realizzata combinando l'uso del parlato con semplici gesti prodotti da un piccolo animatrone posto sulla piattaforma superiore del robot (Figura 19b).

Presso il "Navy Center for Applied Research in Artificial Intelligence" è stata sviluppata un'interfaccia per robot mobili che combina gesti e linguaggio naturale [51]. 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 [11] 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 ad un determinato gesto. La combinazione di parlato e gesti è ottenuta per mezzo della teoria Bayesiana.

5. 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-macchina, 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 multimodali meritano maggiore interesse.

BIBLIOGRAFIA

- [1] A. Agah. Human Interactions with Intelligent Systems: Research Taxonomy. *Computers and Electrical Engineering*, 27: 71-107, 2001.

- [2] R. O. Ambrose, H. Aldridge, R. S. Askew, R. Burrige, W. Bluethman, M. A. Diftler, C. Lovchik, D. Magruder and F. Rehnmark. ROBONAUT: NASA's Space Humanoid. *IEEE Intelligent Systems Special Issue on Humanoids*, 2000.
- [3] L. Aryananda. Online and Unsupervised Face Recognition for Humanoid Robot: Toward Relationship with People. *Proceedings of the 2001 IEEE – RAS International Conference on Humanoid Robots*, 2001.
- [4] I. Asimov. *The Complete robot - The Definitive Collection of Robot Stories*, Harper Collins, Londra, 1995.
- [5] S. Bahadori, A. Cesta, G. Grisetti, L. Iocchi, R. Leone, D. Nardi, A. Oddi, F. Pecora and R. Rasconi. RoboCare: an Integrated Robotic System for the Domestic Care of the Elderly. In P. Remagnino and G. L. Foresti (Eds.), *Proceedings of Workshop on Ambient Intelligence AI*IA-03*, pp. 81 – 96, Pisa, Italia, 2003.
- [6] T. Balch. Taxonomies of Multirobot Task and Reward. In T. Balch and L. E. Parker (Eds.), *Robot Teams*, pp. 3- 22, Natick, MA, 2002.
- [7] N. O. Bernsen, H. Dybkjær and L. Dybkjær. What Should Your Speech System Say, *IEEE Computer*, 30(12), 1997.
- [8] R. Bischoff and 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 (Springer Tracts in Advanced Robotics, 5)*, Heidelberg, Germania, 2003.
- [9] A. Bobick and J. Davis. An Appearance-Based Representation of Action. *Proceedings of the IEEE International Conference in Pattern Recognition (ICPR-96)*, pp. 307- 312, Vienna, Austria, Agosto 1996.
- [10] R. A. Bolt. "Put-That-There": Voice and Gesture at the Graphics Interface. *Computer Graphics*, 14(3):262-270, 1980.
- [11] M. Bray, H. Sidenbladh and J. O. Eklundh. Recognition of Gestures in the Context of Speech. *Proceedings of the International Conference on Pattern Recognition (ICPR02)*, Quebec City, Agosto 2002.
- [12] C. Breazeal and B. Scassellati. Infant-like Social Interactions Between a Robot and a Human Caretaker. *Adaptive Behavior*, 8(1):49-74, 2000.
- [13] C. Breazeal. Affective Interaction Between Humans and Robots. *Proceedings of the 2001 European Conference on Artificial Life (ECAL2001)*, Praga, 2001.
- [14] G. Bugmann. Instructing Robots. *AISB Quarterly*, vol. 113, pp. 6-7, 2003.
- [15] J. Cassell, J. Sullivan, S. Prevost and E. Churchill. *Embodied Conversational Agents*, MIT Press, Cambridge, MA, 2000.
- [16] C. Cedras and M. Shah. Motion-Based Recognition: A Survey. *Image and Vision Computing*, 13(2):129-155, 2000.
- [17] K. Dautenhahn and A. Billard. Bringing up Robots or – The Psychology of Socially Intelligent Robots: from Theory to Implementation. *Proceedings of the Autonomous Agents (Agents '99)*, pp. 366-367, Seattle, Washington, USA, 1999.
- [18] A. M. Davies. Aging and Health in the 21st Century. *Aging and Health: A Global Challenge for the Twenty-First Century*, Ginevra, Svizzera, 1999.
- [19] C. DiSalvo, F. Gemperle, J. Forlizzi, and S. Kiesler. All Robots Are Not Created Equal: the Design and the Perception of Humanness in Robot Heads, *DIS2002 Conference Proceedings*, pp. 321- 326, 2002.
- [20] L. Donati. *La storia dei robot*, Franco Cosimo Panini Editore, Modena, 1995.
- [21] K. Donellan. Reference and Definite Descriptions, *Philosophical Review*, LXXV, pp. 281-304, 1966.
- [22] G. Dudek, M. Jenkin and E. Milios. A Taxonomy of Multirobot Systems. In T. Balch and L. E. Parker (eds.), *Robot Teams*, pp. 3-22, Natick, MA, 2002.
- [23] C. A. Ellis, S. J. Gibbs and G. L. Rein. Groupware: some Issues and Experiences. *Communications of the ACM*, 34(1):39-58, 1991.
- [24] J. F. Engelberger. *Robotics in Service*, MIT Press, Cambridge, MA, 1989.
- [25] T. Fong, C. Thorpe and C. Baur. Collaboration, Dialogue and Human-Robot Interaction. *Proceedings of the 10th International Symposium of Robotics Research*, Lorne, Victoria, Australia, 2001.
- [26] T. Fong, I. Nourbakhsh and K. Dautenhahn. A Survey of Socially Interactive Robots. *Robotics and Autonomous Systems*, 42:143-166, 2003.
- [27] J. Gao and J. Shi. Inferring Human Upper Body Motion Using Belief Propagation. Technical report CMU-RI-TR-03-06, Robotics Institute, Carnegie Mellon University, Giugno 2003.
- [28] M. V. Giuliani, M. Scopelliti, F. Fornara, E. Muffolini and A. Saggese. Human-Robot Interaction: How People View Domestic Robots. *Proceedings of the First RoboCare Workshop*, Roma, Ottobre 2003.
- [29] R. C. Goertz. Manipulator Systems Development at ANL. *Proceedings of the 12th Conference on Remote Systems Technology*, pp. 117-136, American Nuclear Society, 1964.
- [30] D. Grant. An Integrated Human Computer Interface using Eye Gaze Tracking and Facial Feature Recognition. Project proposal, MSc Computational Intelligence, University of Plymouth, 1998.
- [31] H. P. Grice. Logic and Conversation. In P. Cole and J. L. Moorgan (Eds.), *Syntax and Semantics 3: Speech Acts*, pp. 41-58, Academic Press, New York, 1975.
- [32] S. Gutta, J. Huang, I. Imam and H. Wechsler. Face and Hand Gesture Recognition Using Hybrid Classifiers. *Proceedings of the International Conference on Automatic Face and Gesture Recognition (ICAFGR 96)*, pp. 164-169, Killington, 1996.
- [33] A. Haro, M. Flickner and I. Essa. Detecting and Tracking Eyes By Using Their Physiological Properties, Dynamics, and Appearance. *IEEE CVPR 2000*, pp. 163 – 168, 2000.
- [34] A. Herskovits. Language and Spatial Cognition. An Interdisciplinary Study of the Prepositions in English, Cambridge University Press, Cambridge, UK, 1986.
- [35] I. Horswill. Analysis of Adaptation and Environment, *Artificial Intelligence*, 73(1-2):1-30, 1995.

- [36] H. Hüttenrauch and K. Severinson-Eklundh. Fetch-and-carry with CERO: Observations from a Long-term User Study with a Service Robot. *Proceedings of the 11th IEEE International Workshop on Robot and Human Interactive Communication (ROMAN 2002)*, pp. 158-163, 2002.
- [37] W.L. Johnson, J. W. Rickel, J. W. and 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.
- [38] T. Kanda, H. Ishiguro, M. Imai, T. Ono and K. Mase. A constructive approach for developing interactive humanoid robots. *Proceedings of IROS 2002*, IEEE Press, New York, 2002.
- [39] Z. Khan. Attitudes towards Intelligent Service Robots, IpLab, Nada, Royal Institute of Technology, 1998.
- [40] T. Leangle, T. C. Lueth, E. Stopp, G. Herzog and G. Kamstrup. KANTRA - A Natural Language Interface for Intelligent Robots. *Intelligent Autonomous Systems (IAS 4)*, pp. 357- 364, 1995.
- [41] L. Lorek. March of the A.I. Robots. *Interactive Week*, 8(17):46, 2001.
- [42] H. S. Lusted and R. B. Knapp. Controlling Computers with Neural Signals. *Scientific American*, Ottobre 1996.
- [43] P. McGuire, J. Fritsch, J. J. Steil, F. Röthling, G. A. Fink, S. Wachsmuth, G. Sagerer and H. Ritter. Multi-modal human-machine communication for instructing robot grasping tasks. *Proceedings of IROS 2002*, pp. 1082- 1089, IEEE, 2002.
- [44] J. T. Milde, K. Peters and S. Strippgen. Situated Communication with Robots. *Proceedings of the First International Workshop on Human-Computer Conversation*, Bellagio, Italia, 1997.
- [45] H. Mobahi. Building an Interactive Robot Face from Scratch. Bachelor of Engineering Final Project Report, Azad University, Tehran-South Campus, Tehran, Iran, Maggio 2003.
- [46] C. Morimoto, D. Koons, A. Amir and 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*, 18(4):331-335, 2000.
- [47] R. R. Murphy and E. Rogers. Final Report for DARPA/NSF Study on Human-Robot Interaction, 2001.
- [48] W. C. Ogden and P. Bernick. Using Natural Language Interfaces. In M. Helander, T. K. Landauer and P. Prabhu (Eds.), *Handbook of Human-computer Interaction*, Elsevier Science Publishers B.V., Amsterdam, 1997.
- [49] B. Ogden and K. Dautenhahn. Robotic Etiquette: Structured Interaction in Humans and Robots. *Proceedings of SIRS2000, Symposium on Intelligent Robotic Systems*, pp. 353-361, Reading, UK, 2000.
- [50] A. M. Okamura and M. R. Cutkosky. Feature Detection for Haptic Exploration with Robotic Fingers. *International Journal of Robotics Research*, 20(12):925-938, 2001.
- [51] D. Perzanowski, W. Adamse and A. Schultz. Communicating with a Semi-Autonomous Robot Combining Natural Language and Gestures. *Workshop on Integrating Robotics Research: Taking the Next Leap*, pp. 55-59, AAAI98 Spring Symposium Series, AAAI Press, Menlo Park, CA, 1998.
- [52] G. Rigoll, A. Kosmala and S. Eickeler. High Performance Real-Time Gesture Recognition Using Hidden Markov Models. *Proceedings of the Bielefeld Gesture Workshop*, pp. 69-70, Bielefeld, Germania, Settembre 1997.
- [53] K. Severinson-Eklundh, A. Green and H. Hüttenrauch. Social and Collaborative Aspects of Interaction with a Service Robot. *Robotics and Autonomous Systems, Special Issue on Socially Interactive Robots*, 42(3-4), 2003.
- [54] C. L. Sidner and M. Dzikovska. Engagement between Humans and Robots for Hosting Activities. *IEEE International Conference on Multimodal Interfaces*, pp. 123-128, Ottobre 2002.
- [55] C. L. Sidner, C. Lee and N. Lesh. The Role of Dialogue in Human Robot Interaction. Mitsubishi Electric Research Laboratories, Giugno 2003.
- [56] D. Spiliotopoulos, I. Androutopoulos and C. D. Spyropoulos. Human-Robot Interaction Based on Spoken Natural Language Dialogue. *Proceedings of the European Workshop on Service and Humanoid Robots*, 2001.
- [57] T. Starner and A. Pentland. Visual Recognition of American Sign Language Using Hidden Markov Models. *International Workshop on Automatic Face and Gesture Recognition*, Zurigo, Svizzera, 1995.
- [58] R. Stiefelhagen, J. Yang and A. Waibel. Tracking Focus of Attention for Human-Robot Communication. *Proceedings of the International Conference on Humanoid Robots*, 2001.
- [59] H. Tanaka, T. Tokunaga and Y. Shinyama. Animated Agents that Understand Natural Language and Perform Actions. *Proceedings of Lifelike Animated Agents (LAA)*, Tokyo, Giappone, 19 Agosto 2002.
- [60] C. Theobalt, J. Bos, T. Chapman, A. Espinosa-Romero, M. Fraser, G. Hayes, E. Klein, T. Oka and R. Reeve. Talking to Godot: Dialogue with a Mobile Robot. *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2002)*, pp. 1338-1343, 2002.
- [61] S. Waldherr, R. Romero and S. Thrun. A Gesture Based Interface for Human-Robot Interaction. *Autonomous Robots*, 9(2):151-173, 2000.
- [62] Y. Wu and T. Huang. Vision-Based Gesture Recognition: A Review. *Gesture-Based Communications in HCI*. Lecture Notes in Computer Science, 1739, Springer, Berlino, 1999.
- [63] H. A. Yanco and J. L. Drury. A Taxonomy for Human-Robot Interaction. *Proceedings of the AAAI Fall Symposium on Human-Robot Interaction*, pp. 111-119, AAAI Technical Report FS-02-03, Falmouth, Massachusetts, Novembre 2002.
- [64] J. Zelek, D. Bullock, S. Bromley and H. Wu. What the Robot Sees & Understands Facilitates Dialogue. *Human-Robot Interaction, 2002 AAAI Fall Symposium*, North Falmouth, Massachusetts, USA, 15 – 17 Novembre 2002.

APPENDICE

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

Artificial Intelligence Laboratory: www.ai.mit.edu
 Aryan: www.digibrain.org
 ASIMO: world.honda.com/ASIMO
 CAS: www.cas.edu.au
 Center for Work Technology and Organization:
www.stanford.edu/group/WTO
 Communication and Cognition:
<http://www.tilburguniversity.nl/faculties/flw/cc/research/goals.html>
 Godot: www.ltg.ed.ac.uk/godot
 Haptic Exploration Laboratory: www.haptics.me.jhu.edu
 HELPMATE:
www.pyxis.com/products/newhelpmate.asp
 HERMES: www.unibw-muenchen.de/hermes
 Humanoid Robotics Project: www.mstc.or.jp/hrp
 Institute for Interactive Arts and Engineering:
iiae.utdallas.edu/projects
 Instruction-Based Learning for Mobile Robots:
www.tech.plym.ac.uk/soc/staff/guidbugm/ibl
 Interactive Systems Laboratories:
www.is.cs.cmu.edu/mie
 IPLab (The Interaction and Presentation Laboratory):
www.nada.kth.se/iplab
 JSK Laboratory: www.jsk.t.u-tokyo.ac.jp

KANTRA: www.dfki.uni-sb.de/vitra/papers/ro-man94
 Leonardo:
robotic.media.mit.edu/projects/Leonardo/Leo-intro.html
 MERL: www.merl.com
 MIT (Massachusetts Institute of Technology):
web.mit.edu
 MIT Media Laboratory: www.media.mit.edu
 National Centre for Scientific Research "Demokritos":
www.demokritos.gr
 Navy Center for Applied Research in Artificial
 Intelligence: www.aic.nrl.navy.mil
 Project on People and Robots:
www.peopleandrobots.org
 ROBONAUT: robonaut.jsc.nasa.gov
 Robotic Intelligence Laboratory:
www.tech.plym.ac.uk/iq4bots
 Robotics Institute: <http://www.ri.cmu.edu/home.html>
 SpeechLinks (CMU):
www.speech.cs.cmu.edu/comp.speech/Section6/speechlinks.html
 Stan Winston Studio: www.stanwinstonstudio.com
 Tokyo Institute of Technology: www.titech.ac.jp
 Touch Lab: touchlab.mit.edu

CONTATTI



Amedeo Cappelli, Emiliano Giovannetti

KDD Laboratory

Istituto di Scienza e Tecnologie dell'Informazione "A. Faedo", CNR
 via G. Moruzzi 1

56124 San Cataldo (PI)

Email: {*Amedeo.Cappelli* | *Emiliano.Giovannetti*}@isti.cnr.it

	<p>Amedeo Cappelli is a senior research scientist at the Institute of Information Science and Technology of the Italian Research Council. He coordinates the Group on Natural Language Processing of the Knowledge Discovery and Delivery Laboratory. He is one of the founding members of the AI*IA.</p>
	<p>Emiliano Giovannetti received a Laurea degree in computer science from the University of Pisa in 2003. His current research interest lies in the field of natural language processing and knowledge representation and reasoning.</p>