

Multimedia Contents



71. Cognitive Human–Robot Interaction

Bilge Mutlu, Nicholas Roy, Selma Šabanović

A key research challenge in robotics is to design robotic systems with the cognitive capabilities necessary to support human–robot interaction. These systems will need to have appropriate representations of the world; the task at hand; the capabilities, expectations, and actions of their human counterparts; and how their own actions might affect the world, their task, and their human partners. *Cognitive human–robot interaction* is a research area that considers human(s), robot(s), and their joint actions as a cognitive system and seeks to create models, algorithms, and design guidelines to enable the design of such systems. Core research activities in this area include the development of representations and actions that allow robots to participate in joint activities with people; a deeper understanding of human expectations and cognitive responses to robot actions; and, models of joint activity for human–robot interaction. This chapter surveys these research activities by drawing on research questions and advances from a wide range of fields including

71.1	Human Models of Interaction	1908
71.1.1	Mental Models of Robots	1908
71.1.2	Social Cognition	1911
71.2	Robot Models of Interaction	1914
71.2.1	Developmental Models	1915
71.2.2	Robot Spatial Cognition	1915
71.2.3	Symbol Grounding	1915
71.3	Models of Human–Robot Interaction	1916
71.3.1	Dialog–Based Models	1916
71.3.2	Simulation–Theoretic Models	1921
71.3.3	Intention– and Activity–Based Models	1922
71.3.4	Models for Action Planning	1923
71.3.5	Cognitive Models of Robot Control	1926
71.4	Conclusion and Further Reading	1927
	Video–References	1927
	References	1928

computer science, cognitive science, linguistics, and robotics.

When people interact with each other, they draw on mental models of themselves, of their interaction partners, of the immediate context of the interaction, and of their broader physical, social, and cultural context. These models help them predict the actions of their interaction partners and make decisions about their own actions. To effectively interact with people, robots need similar models that help them determine their own actions and predict the actions of their users. *Cognitive human–robot interaction* (HRI) is a research area that seeks to improve interactions between robots and their users by developing cognitive models for

robots and understanding human mental models of robots.

A central tenet of cognitive HRI is that humans, robots, and the context of their interaction form a complex cognitive system situated in the real world. A key research activity in the field involves the development of frameworks to represent this system [71.1–3]. This activity is informed primarily by research in cognitive science that develops frameworks to represent human cognitive systems. These frameworks include physical symbol systems [71.4], situated actions [71.5, 6], and those that combine symbolic and situated perspectives,

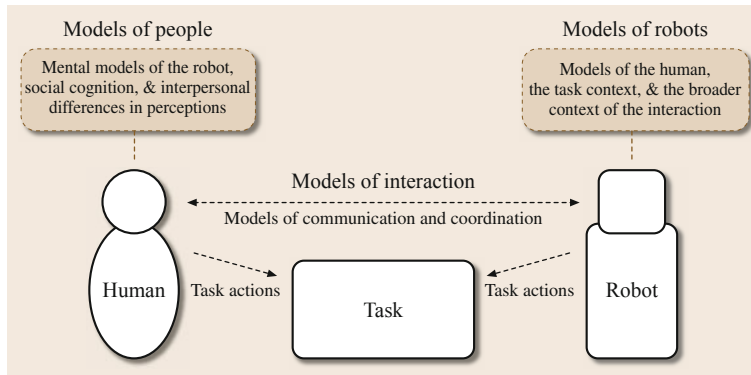


Fig. 71.1 A visual summary of the research activities in cognitive HRI

such as activity theory [71.7] and distributed cognition [71.8]. While the discussion on what framework best represents HRI as a cognitive system is ongoing, research in cognitive HRI involves the development of both symbolic and situated representations. More specifically, research activities in this area include (also illustrated in Fig. 71.1):

1. *Human models of interaction:* Building an understanding of people's mental models of robots, how people perceive robots and interpret their actions and behaviors, and how these perceptions and interpretations change across contexts and user groups.
2. *Robot models of interaction:* The development of models that enable robots to map aspects of the interaction into the physical world and develop cognitive capabilities through interaction with the social and physical environment.
3. *Models of HRI:* Creating models and mechanisms that guide human–robot communication and collaboration, action planning, and model learning.

This chapter surveys existing efforts in these research areas, drawing on research questions and advances from a wide range of fields including robotics, cognitive science, and linguistics.

71.1 Human Models of Interaction

Robots are expected to increasingly enter everyday environments – outside of factories and laboratories – including homes [71.9, 10], offices [71.11], and classrooms [71.12, 13]. In these contexts, robots will need to coexist and collaborate with a wide variety of users, such as children and the elderly, many of whom will not be technically trained. Accordingly, there is a growing research emphasis in cognitive HRI on identifying the mental models people use to make sense of emerging robotic technologies and investigating people's reactions to the appearance and behaviors of robots. This research aims not only to improve the ease of use of robots by designing them to fit human mental models, but also to gain new insights about human cognition and behavior. With the latter goal in mind, researchers also use robots to embody specific theories of human cognition that are then evaluated through HRI studies.

71.1.1 Mental Models of Robots

Research in human–computer interaction has shown that people's attitudes and behaviors toward digital technologies often follow the social rules established in

human–human interaction [71.14]. It is reasonable to expect that people will similarly interpret the interactive behaviors of robots in social ways. Cognitive HRI researchers continue to investigate the extent to and conditions in which this maxim applies to HRI as they use their understanding of human social cognition to develop robotic platforms adapted to users' expectations and behaviors. In the process of evaluating such robotic platforms, researchers explore people's mental models of robots and identify areas in which users' expectations and understandings of robots may not be born out by the robot's appearance or behavior in ways detrimental to the HRI experience. Knowing which mental models people are using to interpret robot behavior not only helps roboticists to understand HRI more deeply, but also helps them in designing appropriate behaviors for the robot.

Models Ascribed to Robots

Extensive research by *Turkle et al.* [71.15, 16] examines how people, including children and older adults, make sense of their novel interactions with social robots such as Kismet, Cog, PARO, Furby, and My Real Baby.

These studies show that people apply a variety of mental models relating to animacy, sociality, affect, and consciousness to explain their experiences and emerging relationships with robots. Some research participants approached robots in a scientific-exploratory mode, interpreting a robot's actions in an emotionally detached and mechanistic manner. Others took a relational-animistic approach, investing in the interactions emotionally and treating robots as if they were living beings, such as babies or pets. The ways in which participants described the robots verbally did not always fit with the way in which they interacted with them – a person who says the robot is *only a mechanical thing* may still act toward the robot in a nurturing manner, such as soothing a crying My Real Baby [71.16, p.118]. This corresponds to previous findings in human–computer interaction (HCI) which suggest that people *mindlessly* apply social characteristics to computers [71.17]. Field studies with the seal-like robot PARO have shown that robots can also act as *evocative objects* that spark reflections on previous relationships and events (e.g., with a grandchild, spouse, or pet), which users then use to make sense of their interactions with the robots [71.18, 19].

In addition to identifying the mental models people use to interpret their experiences with robots, researchers study the effects from deliberately incorporating specific social schemas into robot design. *Anthropomorphism*, or the attribution of human characteristics to nonhuman (e.g., animal or artifact) behavior, is an interpretive schema that has been of particular interest to HRI researchers. Some scholars, such as *Nass* and *Moon* [71.17], critique anthropomorphic explanations as false and misleading. Others, including *Duffy* [71.20] and *Kiesler* et al. [71.21], suggest that the deliberate use of anthropomorphism can benefit social robot design by taking advantage of people's propensity to interpret events and other agents socially to make robot behaviors more understandable to users. This interpretation raises the question of which characteristics of the robot or the interaction are instrumental in inciting people to anthropomorphize robots and has inspired researchers to study a variety of socio-cultural cues, behaviors, and task contexts. *Kiesler* et al. [71.21] showed that people anthropomorphize a physically embodied robot more readily than an on-screen agent, and people behave in a more engaged and socially appropriate manner while interacting with the co-present robot. People also anthropomorphize robots they interact with directly more than they do with robots in general, and with robots that follow social conventions (e.g., polite robots) more than those that do not [71.22]. The personal characteristics, such as personality, of the human interaction partner can also affect their mental models of robots. For example, users with low emotional stability and extraversion

scores were found to prefer mechanical-looking robots to human-like ones [71.23].

As might be expected, a robot's human-like appearance can have a positive effect on people's propensity to anthropomorphize [71.25]. Obversely, too high a level of human-likeness may place the robot in the *uncanny valley* [71.26]. The *uncanny valley* refers to a dip in the hypothetical nonlinear graph describing the relationship between a robot's human-likeness and a human's emotional response to it, suggesting that a robot with a very high degree of human-likeness coupled with some remaining nonhuman qualities will make users uncomfortable. This hypothesized effect essentially describes what happens when a person's mental model of the robot as human is not born out by its interactive capabilities. Various cognitive aspects of this hypothesis have been studied, suggesting that the construct is multidimensional [71.27, 28] rather than two-dimensional (2-D), as depicted by *Mori* [71.26]. Furthermore, research suggests that the mismatch between different dimensions, rather than any quality alone, can cause a dissonance that leads to people's discomfort with robots. *MacDorman* et al. [71.24] show that incongruencies between a robot's appearance and movement can diminish anthropomorphic attributions (Fig. 71.2). A similar result was found by *Saygin* et al. [71.29], who used functional magnetic resonance imaging (fMRI) to show that the human action perception system made distinctive responses to the mismatch between the level of human-likeness of a robot's appearance and motion but not to appearance or motion alone. Mismatches in the human- or robot-like qualities of an on-screen robot's voice and appearance were also shown to heighten people's sense of the character's eeriness – people found

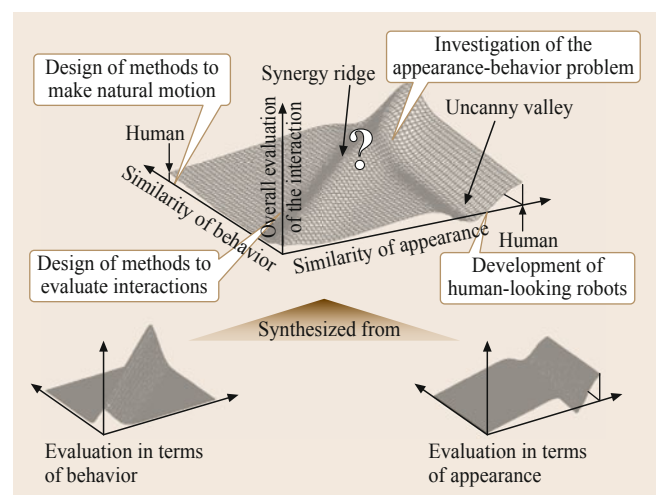


Fig. 71.2 An extended notion of the uncanny valley which includes appearance and behavior as significant variables (after [71.24])

both a robot with a human voice and a human with a robot voice to be creepy [71.30].

Mental Models in Robot Design

Researchers may deliberately include specific anthropomorphic schemas to promote user behaviors that aid robots in performing their tasks. One common example is the use of the *baby schema* – a soft round appearance, large eyes, and proto-verbal utterances – in Kismet [71.31] (Fig. 71.4), Muu [71.32] (Fig. 71.3), and Infanoid [71.33] to encourage people to anthropomorphize robots. This schema is also useful in that it can incite people to behave in a nurturing manner toward robots in the interest of scaffolding the robots' learning in a way similar to infant–parent interactions. A robot's *perceived* gender can also have an effect on people's mental models of the robot's knowledge of certain topics; for example, in one study a female robot was expected to be more knowledgeable about dating than a male robot [71.34]. While certain mental models become operational as soon as a person starts interacting with a robot (e.g., gender, age, human-likeness), people can adapt their mental models of a robot's capabilities when given additional information about the robot's personal characteristics, such as the robot's country of origin or the language it speaks [71.35]. Goetz et al. [71.36] showed that matching a robot's *personality* to the task it is supposed to perform can have a significant effect on its efficacy: people were more responsive to a robot that had a serious, rather than an entertaining, demeanor when its job was to motivate them to exercise. Also focusing on *task models* in HRI, Lee et al. [71.37] showed how people use their existing utilitarian and relational models of service to set expectations for their interactions with a service robot. These models also affected the preferred ways in which the robot should make up for any mistakes it makes in service – people with a utilitarian mental model of service preferred to receive compensation, while those with a relational model responded well to an apology.

As interactive robots are developed and used all over the world, researchers have also started exploring how *cultural models* [71.38] affect people's perceptions of and interactions with robots. Social and behavioral norms are culturally variable, so we can expect users' understanding and adoption of socially interactive robots to differ accordingly. Cross-cultural research in HRI largely supports this expectation. Evers et al. [71.39] showed that users from China and the US respond differently to robots. Further research by Wang et al. [71.40] suggests that specific cultural models regarding communication norms, particularly explicit and implicit modes of communicating information and intent to interaction partners, affect people's

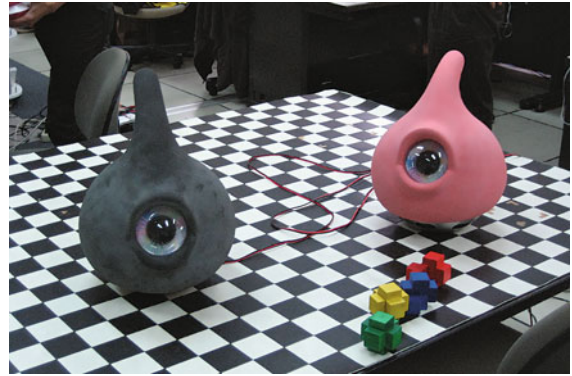


Fig. 71.3 Muu's big eyes and soft round body are designed according to the *baby schema*. Using two robots that can interact with each other instead of one suggests a relational understanding of agency (courtesy of Šabanović)

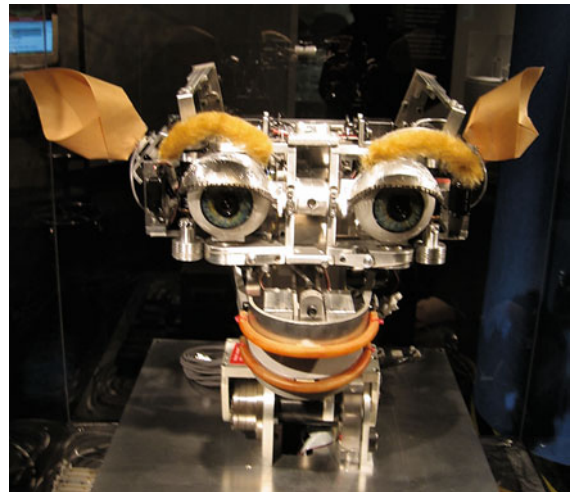


Fig. 71.4 Kismet's big round eyes and infant-like vocalizations are another example of the *baby schema* (courtesy of Šabanović)

perceptions of a robot's trustworthiness and its in-group membership. Researchers have also shown that robot-cists themselves use cultural models unintentionally in their work, including particular models of emotional display [71.41], and cultural models reflecting historical, theological, and popular perceptions of robotic technology [71.42]. Research on cultural models in HRI not only points to the importance of reflexively including such models in robot design, but also allows researchers to do systematic research on culturally situated cognition using robots as stimuli.

Research on mental models applied to interactive robots has not only shown that people use their existing mental models to make sense of these novel artifacts, but also that we may need new *ontological*

categories to accommodate emerging mental models of these entities [71.16, 43]. Kahn's et al. [71.44] studies of children's moral interpretations of interactions with an AIBO robot showed that their mental models of the robots included rationalizations and behaviors related to both inanimate and animate objects. Turkle [71.45] suggests that interactive robots co-opting relational feelings and responses normally reserved for animals and humans call into question the authenticity of relationships. Further, Turkle [71.45] suggests that a more sophisticated new notion of autonomous yet inanimate artifacts has become necessary. Both researchers have suggested that interactive robots might comprise a new ontological category, and that we also need to be conscious of the ways in which interactions with these artifacts affect our mental models of animate beings.

71.1.2 Social Cognition

The development of robots that can interact naturally with humans calls for the detailed study of social activity and the cognitive models that underly such activities. Scassellati [71.46] argues that robots can help us study the limits of human social cognition because they are not alive, yet they can behave in socially appropriate (or inappropriate) and evocative ways. Robots that incorporate social cues such as gaze, proximity, and facial expressions, push our *Darwinian buttons* [71.16, p. 8] and effectively coerce us into interacting with them socially. Studying which cues have these effects is an opportunity to learn more about human social cognition and improve robot design.

Researchers studying the social aspects of cognitive HRI are identifying the minimal cues robots need to evoke social responses from people, including those related to robotic embodiment, gaze, proxemic cues, and interaction rhythms. Current research is also focused on applying and evaluating different models of cognition in the context of HRI. Robots can be unprecedented experimental tools for the study of social cognition. They can be used to provide stimuli in experiments and field studies, since their actions and behaviors can be carefully controlled, finely tuned and varied, and repeated exactly and indefinitely, which is often challenging even for well-trained human confederates [71.47, 48]. Furthermore, robots do not have difficulty acting *unnaturally* (e.g., not reacting to other person's cues) or violating social norms (e.g., being rude) when needed, a source of potential stress in human researchers [71.49].

Minimal and Human-Like Cues in HRI

One approach to studying social cognition has been to try to isolate the *minimal* set of cues that evoke social

responses and perceptions from human interaction partners. The creators of Muu followed a *minimal design strategy* [71.32], using cartoons and children's drawings to develop a robot that can be communicatively engaging to people without relying on overt human-likeness. Kozima et al.'s [71.50] Keepon was designed to include characteristics common to living beings, such as lateral symmetry and two eyes, which are assumed to be important for social interaction (Fig. 71.5). The robot also performs fundamental social behaviors, such as joint attention, eye contact, and emotional expression through bodily posture and movement and using only four degrees of freedom (Fig. 71.6). These minimal cues have been shown to be sufficient for engaging children in short-term interaction in the lab and long-term interaction in more natural environments, such as a classroom [71.50].

Studies with minimalist robots have also underscored the effect of *social context* in people's interpretations of robots. Field studies with Keepon in an elementary school showed that children incorpo-

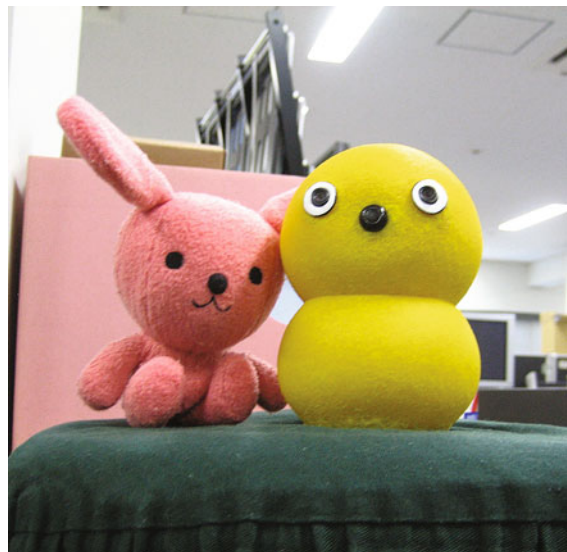


Fig. 71.5 Keepon is a simple robot used to investigate cues such as joint attention, emotive expression, and rhythmicity in HRI (courtesy of Šabanović)

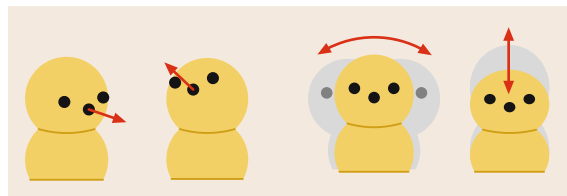


Fig. 71.6 Keepon uses four degrees of freedom to express emotive and attentional cues (after [71.50])



Fig. 71.7 An android robot fabricated by Kokoro Ltd. (courtesy of Šabanović)

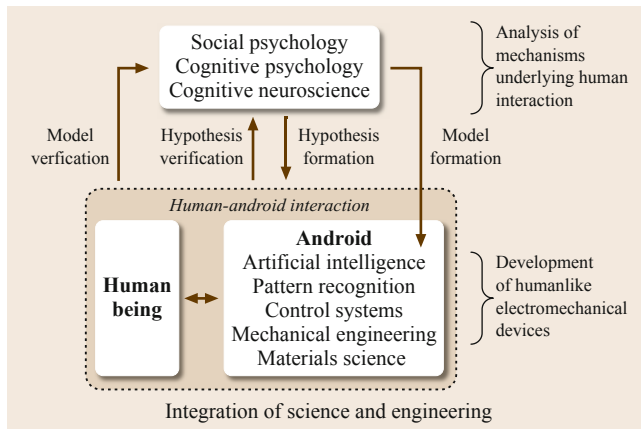


Fig. 71.8 Androids can be used to investigate human cognition analytically as well as synthetically (after [71.51])

a baby or pet, or treating Keepon as another student in the classroom) due not only to its interpretive flexibility, but to the richness of the social environment. This inspired children to engage with the robot over long periods of time, sometimes years, whereas they became bored after 10–15 min when interacting with Keepon in the laboratory. The above mentioned Muu's design was inspired by ecological models of cognition [71.52, 53] suggesting that a robot is inherently incomplete as a communicative device – it needs a human interaction partner to imbue its actions with meaning. Muu therefore relies on the context and the presence of other interactive agents (including people, other Muu, and objects such as blocks displayed in Fig. 71.3 for triadic interaction) to enable people to make sense of its actions and relationally ascribe social agency to the robot. The Social Trash-can project [71.54] similarly explored how minimal social cues, including contingent motion and approaching people, can be used to display the robot's intentions to children and get their assistance in trash collection. Yamaji et al. [71.54] also showed that robots moving together as a group – relationally – were more successful in attracting the children's attention than that moved individually.

An alternative approach to the study of social cognition through HRI focuses on *human-like realism* in appearance and behavior and is proposed by Ishiguro [71.47] and MacDorman and Ishiguro [71.48]. They claim that androids – robots that bear a close and sometimes *uncanny* resemblance to humans (Fig. 71.7, for example) – are unprecedented test beds for the study of social cognition. Used as stand-ins for humans in this *android science*, robots have a twofold function as experimental tools for evaluating hypotheses about human perception, cognition, and interaction, and as a testing ground for various cognitive models (Fig. 71.8). Using an android platform, Ishiguro [71.55] showed the importance of micro movements as a cue that incites people to attribute human-likeness to a robot in short (1–2 s) interactions. Another topic of continuing investigation is the possibility of simulating the personal presence of a remote actor in the local environment using an android platform [71.56, 57]. Shimada et al. [71.58] showed that people evaluated an android as more likable when it mimics them in a way similar to the *chameleon effect* that occurs when two people interact. MacDorman and Ishiguro [71.48] suggested that such androids can be used in research relating to a number of current topics of interest in cognitive science, including the mind-body problem, nature versus nurture, rationality and emotion in human reasoning, and the relationship between social interaction and internal cognitive mechanisms.

rated the simple robot into a wide variety of interaction contexts (e.g., playing house with Keepon as

Embodied Social Cues

Embodiment separates robots from other interactive digital technologies and has been investigated through studies comparing how people interpret and act toward robots that are physically co-present with them and on-screen robots or social agents. *Wainer et al.*'s [71.59] comparison of people's interactions with an embodied robot and a simulated robot, and a co- and tele-present found that people were more engaged with, behaved more appropriately to, and anthropomorphized a co-located robot more than a tele-present robot. People interacting with an embodied robot have also shown tendencies to issue more commands than those interacting with a simulated robot [71.60]. The social effect of embodiment in HRI was further confirmed by *Bainbridge's et al.* [71.61] study showing that people are more likely to comply with requests made by an immediately present robot rather than requests made by a remote robot communicating with them through a television screen. These converging results strongly suggest that a robot's embodied presence has a significant cognitive effect on people's social responses to the robot. The embodied nature of robots also enables the study and use of various other social cues, including proxemic behaviors, gaze, and interaction rhythms, in HRI. A more detailed review of embodied social cues is provided by *Mutlu* [71.62].

Proxemic behaviors [71.63], the study of which is enabled by the embodied nature of robots, not only have a significant effect on people's perceptions of and behaviors toward robots, but have also been used as a measure of people's perceptions of robots as social agents. *Takayama and Pantofaru* [71.64] found that prior experience with pets and robots decreased the distance at which people felt comfortable around robots. Individual traits such as gender and personality also affect people's preferences regarding the distance at which they are comfortable with a robot approaching them [71.64, 65]. Proxemic behavior can be related to other social cues in complex ways. For example, *Mumm's and Mutlu's* [71.66] study showed that people will compensate for the intense gaze in their direction of a robot they do not like by moving away from the robot (Fig. 71.9). While most studies of proxemic systems have been done in the laboratory, recent work is also investigating more natural interactions between humans and robots in open environments [71.67].

Gaze is an important cue in human–human interaction and is also one of the most studied nonverbal social cues in HRI (👁️ VIDEO 128). People use many such seemingly unintentional, unconscious, and automatic nonverbal cues as clues regarding the mental states and intentions of other actors, including robots. Gaze has been shown to be useful for communicat-

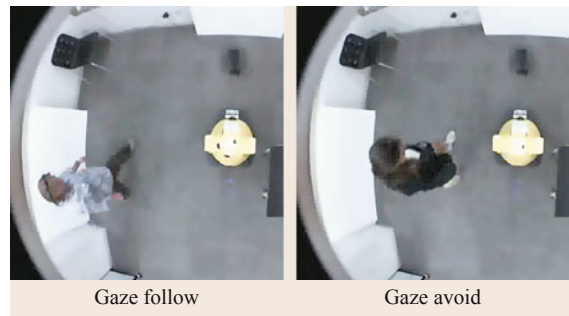


Fig. 71.9 In the study by *Mumm and Mutlu* [71.66], participants maintain a greater distance with the unlikable robot when the robot follows them with its gaze than when its gaze avoids the participant, while their proxemic behavior is not affected by a likable robot's gaze (courtesy of Mutlu)

ing intent, modulating interaction, and even affecting participants' experience and memory of the interaction. Researchers have shown that gaze can be used to engage users [71.69, 70] and to assign them particular roles in and manage the interaction [71.71]. A robot's gaze behavior can affect the human interaction partner's gaze and speech, their comprehension of the robot's speech [71.72], and people's memory of a story narrated by the robot and perceptions of the robotic storyteller [71.73]. Researchers studying the temporal aspects of gaze in HRI found that the timing of gaze behavior provides cues to human intentions while teaching the robot the names of objects, suggesting that properly timed gaze behaviors can have a positive effect on collaborative tasks between a human and a robot [71.68] (Fig. 71.10). *Yu et al.* [71.74] have developed a data-driven approach to analyzing human gaze in the context of HRI, which can be used

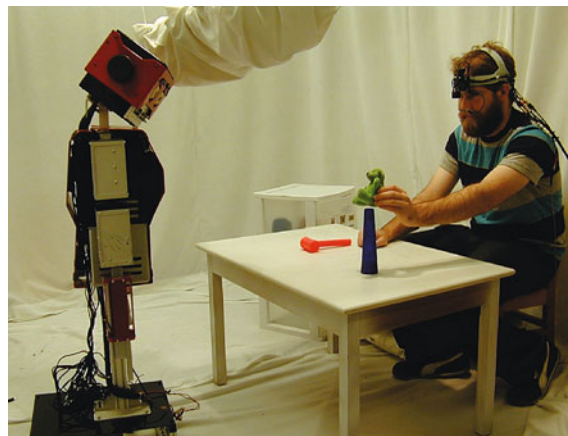



Fig. 71.10 *Yu's et al.* [71.68] HRI studies provide data for developing models of the temporal aspects of interaction (courtesy of Yu)

to develop detailed micro-behavioral gaze models that can guide robot behaviors as well as be used to understand human intentions and behaviors in the course of collaborative activities. A recent study by *Admoni et al.* [71.75], contrary to the assumption that anthropomorphic robots engage us automatically in much the same way we are engaged by people, shows that robot gaze is not necessarily treated by people in the same automatic way that human gaze is treated; we do not, therefore, necessarily perceive robots as social in an automatic and *mindless* way.

Interaction rhythms – nonverbal and largely unconscious temporal coordination between partners in an interaction – enable the exchange of information, anticipation of the interaction partner’s actions, and even positive evaluations of interaction among humans as a fundamental subscript of all human interaction [71.76–78]. The rhythmicity of interaction is therefore also a crucial factor in HRI, both in terms of developing robots that can perceive and respond to people’s rhythmicity, and of understanding how people react to the temporal aspects of robot behaviors. *Michalowski et al.* [71.79] used a dancing robot to explore the rhythmic properties of social interaction and showed that children were more likely to interact with a robot that was synchronized to background music rather than one that was not, and that the children’s own rhythmic behavior was influenced by the robot’s rhythmicity. In further research, *Michalowski et al.* [71.80] suggest that rhythmic interaction can be used as a form of play between children and robots, and that following the robot’s lead in rhythmic entrainment with music causes children to attend more closely to musical rhythm. *Avrunin et al.* [71.81] found that simple changes in a robot’s rhythmic dancing behavior, such as variation of motions, flaws in the robot’s synchrony with music, and coordination of behavior changes with musical dynamics, increased people’s perceptions of the robot’s lifelikeness. *Hoffman and Breazeal* [71.82] used the temporal patterns of interaction – its rhythms – to develop robotic systems that can anticipate a human partner’s actions in collaborative tasks, such as AUR, a robotic desk lamp, and Shimon, a marimba playing robot [71.83] ( VIDEO 236). Along with improving HRI, the use of robots in studying the rhythmic properties of interaction provides a new

tool for cognitive science research on these subtle, fine grained, and unconscious social cues.

Cognitive Development in HRI

A further topic of focus in cognitive HRI has been the study of social and cognitive development through studies of typically developing and autistic children’s interactions with robots. Multiple studies in educational contexts have focused on understanding how children ascribe social agency to robots [71.12, 13]. *Kozima et al.* [71.50] found that children of different ages display varying modes of interaction with the robot, which suggest different levels of comprehension of its ontological status – 0-year-old interacted with *Keepon* as a moving thing, 1–2-year-old interacted with the robot as an autonomous system, and children over 2 years of age treated the robot as a social agent. *Deàk et al.* [71.84] studied the mechanism of joint attention in HRI to explore the importance of contingency and find out which perceptual features infants use to achieve shared attention by modeling these in a robot. Researchers also use robots to study social deficit disorders, particularly autism. Converging results on research using robots to study social deficit disorders show that autistic children respond to robots in a social manner that they do not display with people [71.85–87], inspiring researchers to perform studies with children in the context of HRI. One aim of such research is to try to understand which aspects of a robot’s behavior enable autistic children to participate in social interaction, which may clarify some of the reasons for their difficulties when interacting with humans. HRI researchers have also applied robots to various therapeutic scenarios with autistic children in an effort to provide parents and therapists with a tool to improve communication and understand the children better [71.88, 89]. Studies by *Kozima et al.* [71.90] in which the robot *Keepon* interacts with children with autism suggest that such minimally designed robots can be used to motivate autistic children to share their mental states with others, such as therapists or parents. This work poses a promising possibility for learning more about social deficits and development disorders such as autism, as well as providing tools for diagnosis and therapy using robotic technologies.

71.2 Robot Models of Interaction

Simon [71.91] suggests that the study of human behavior can be approached through synthesis as well as analysis and designs computer simulations as a technique for understanding and predicting the behavior of

natural, social, and cognitive systems. In the spirit of *Simon*’s synthetic approach to the study of human cognition, robotics researchers have been engineering robots as tools for developing and testing a variety of cog-

nitive, behavioral, and developmental models [71.47, 92, 93]. This approach assumes that cognitive models are validated when the implementation of a particular model on a robot produces behavior similar to that produced by humans in the same situation; if this does not occur, it is a sign that there may be something wrong with the model or the way it was implemented in the robot [71.46]. Cognitive HRI research involves the development of robotic platforms based on findings from cognitive science and using such platforms to extend knowledge about human cognitive processes.

71.2.1 Developmental Models

Robots are particularly appropriate for exploring theories of embodied and social cognition, which emphasize the centrality of the agent's interactions with its environment and other agents in that environment to cognitive functioning. In the process of synthesizing a robotic system, the researcher is drawn to focus on the dependency of cognition on noncognitive processes, including the social and physical environment in which cognition takes place. Robots such as Cog and Kismet [71.31] have been used to simulate and validate different theories of cognition, perception, and behavior. Cog was used to implement and test cognitive models relating to reaching behavior, rhythmic motor skills, visual search and attention, and social skill acquisition (e.g., joint attention and theory of mind). In the process they were able to validate, extend, and show the limitations of cognitive, behavioral, and developmental theories. In later projects, researchers have developed models inspired by human cognition and behavior such as social referencing [71.94], perception and action loops [71.95], anticipatory actions in collaborative teamwork [71.96], and others.

Robotics researchers apply the idea that the development of intelligence is embedded in social and cultural environment to the construction of robotic artifacts. For example, *Breazeal* [71.31] applied theories relating to infant social development, psychology, ethology, and evolution to design the robot Kismet, which used infant-like social cues to engage a human participant in interactions that would scaffold the robot's learning, as in the case of infant–parent interactions. Researchers have also developed a variety of robotic systems that exhibit cognitive traits such as imitation [71.97, 98], joint attention [71.99–101], and rhythmic synchrony [71.50, 102]. The Infanoid project [71.33] also used a synthetic approach in which development was understood through studying how the robot learns. Situated and embodied models have been applied to robot learning, particularly through imitation. For example, *Bakker and Kuniyoshi* [71.103] pro-

pose *imitation* as an interaction and learning paradigm in contrast to *robot programming* or *robot learning*. Further, they argue that robot programming is too hard and tedious to specify complex behaviors in sufficient detail and specify how they might be adapted to novel situations.

71.2.2 Robot Spatial Cognition

Systems dedicated to modeling spatial language and interaction, including the theories by *Jackendoff* [71.104], *Landau and Jackendoff* [71.105], and *Talmy* [71.106], have been produced for many years. Several previous works have been computational instantiations of the ideas presented in these theories, in particular the implementation and testing of spatial semantics models. *Regier* [71.107] built a system that assigns labels, such as *through* to a movie showing a figure moving relative to a landmark object. *Kelleher and Costello* [71.108] and *Regier and Carlson* [71.109] built models for the meanings of static spatial prepositions, such as *in front of* and *above*.

Many authors have proposed formalisms for enabling systems to reason about the semantics of natural language use in the context of giving directions. For example, *Bugmann et al.* [71.110] identified a set of 15 primitive procedures associated with clauses in a corpus of spoken natural language directions. *Levit and Roy* [71.111] designed *navigational informational units* that break down instructions into components. *MacMahon et al.* [71.112] represented a clause in a set of directions as a compound action consisting of a simple action (move, turn, verify, and declare-goal), plus a set of pre- and post-conditions. Many of these previous representations are expressive but difficult to automatically extract from text. Some authors avoid this problem by using human annotations [71.111, 112] or by specifying the robot's behavior in a controlled language [71.113]. *Matuszek et al.* [71.114] created a system that follows directions using a machine translation approach. Similarly, *Vogel and Jurafsky* [71.115] used reinforcement learning to automatically learn a model for understanding route instructions.

71.2.3 Symbol Grounding

Mapping language from the human partner to aspects of the external world – locations, objects, or actions the robot should take – described by the language was referred to as an instance of the *symbol grounding* problem [71.116]. There are three different ways people have approached the symbol grounding problem, which is more general than spatial cognition, in robotics. Starting with *Winograd* [71.117], many have created symbol

systems that map between some language and the external world by manually connecting each term onto a pre-specified action space and set of environmental features [71.110, 112, 113, 118–121]. This class of systems takes advantage of the structure of linguistic interaction, but the systems usually do not involve learning, have little perceptual feedback, and have a fixed action space. A second approach involves learning the meaning of words in the sensorimotor space (e.g., joint angles and images) of the robot [71.122–124]. By treating human interaction terms as sensory input, these systems must learn directly from complex features extracted by perceptual systems, resulting in a limited set of commands that can be robustly understood. A third approach is to use learning to convert from an interaction onto aspects of the environment. These approaches may only use linguistic features [71.125, 126], spatial features [71.107] or linguistic, spatial, and semantic features [71.114, 115, 127–129]. These approaches learn the meaning of spatial prepositions (e.g., *above* [71.107]), verbs of manipulation (e.g., *push* and *shove* [71.130]), and verbs of motion (e.g., *follow* and *meet* [71.131]) and landmarks (e.g., *the doors* [71.129]).

Recent progress in probabilistic relational models, such as the generalized grounding graph (G^3), has addressed these issues by exploiting the structure of

spatial discourse, breaking down a natural language command into component clauses and connecting each word to a physical interpretation [71.131, 132]. The grounding graph takes full advantage of the hierarchical and compositional structure of natural language commands and is able to ground landmarks, such as *the computers*, by exploiting object co-occurrence statistics between unknown noun phrases and known perceptual features, spatial relations, such as *past* in the path of an agent relative to an object, and motion verbs, such as *follow*, *meet*, *avoid*, and *go* in the path of a single agent or multiple agents. Once trained, the G3 model can ground spatial discourse in a semantic map of the environment; the map can be given a priori or created on the fly as the robot explores the environment. The G3 model is dynamically instantiated as a hierarchical probabilistic graphical model that connects each element in a natural language command to an object, place, path, or event in the environment. Its structure is created according to the compositional and hierarchical structure of the command, learning the mapping from language onto a continuous robot plan. The G3 model is trained on a corpus of natural language commands paired with groundings, and learns meanings for words and phrases in the corpus, including complex verbs, such as *put* and *take*.

71.3 Models of Human–Robot Interaction

Robotic technologies that interact with people – whether they afford closed-loop teleoperation or collaborate autonomously as peers – need to interpret, make decisions about, and respond to their environment, particularly the physical world, the task that they are expected to support, and the actions, goals, and intentions of the other agents – including people. To achieve these goals, robots need *models* that accurately represent the physical and cognitive characteristics of their environment. These models might outline such characteristics as narrowly as control–action relationships in the context of teleoperation or as comprehensively as human–robot joint activity in the context of peer-to-peer collaboration. Cognitive HRI considers the

robotic system to be a part of a distributed cognitive system and therefore seeks primarily to develop cognitively inspired models [71.2]. These models might draw on knowledge about human cognition to improve the usability of robotic system, mimic human decision making or behavior mechanisms, or represent the complete human–robot cognitive system, offering cognitive representations for different paradigms of HRI (Fig. 71.11).

71.3.1 Dialog-Based Models

Research on human robot interaction across different interaction paradigms from teleoperation [71.133, 134]

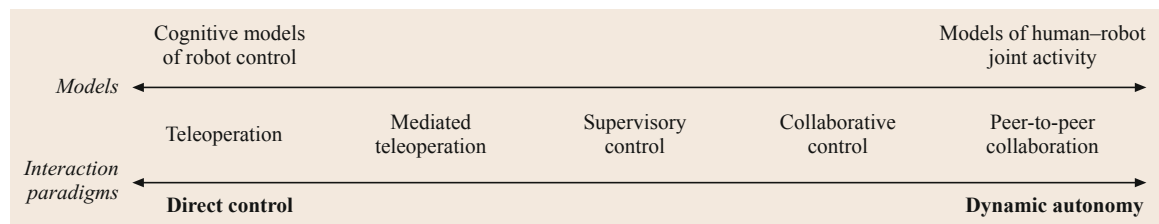


Fig. 71.11 Different paradigms of HRI (after [71.2])

to peer-to-peer interaction [71.135] has highlighted the need for establishing *common ground* [71.136] for effective HRI. In the context of teleoperation, Burke et al. [71.133] found that a lack of appropriate shared representations among human team members and the robot resulted in discrepancies in understanding among team members and breakdowns in perceiving and interpreting data provided by the robot. Stubbs et al. [71.134] observed such lack of common ground between operators and the robot across varying levels of autonomy. In the context of peer-to-peer interaction, Kiesler [71.135] argues that participants in an encounter seek to minimize their collective effort to reach mutual understanding and that the effort needed to establish this understanding between a robot and its users might determine the outcomes and success of HRI. These examples have motivated a large body of research in developing dialog-based models for establishing common ground in human–robot joint activity.

An example of the application of a dialog-based model to a task domain that traditionally involved supervisory control is Fong’s et al. [71.136] *collaborative control* system. In this system, the human and the robot collaborated as partners to perform tasks such as navigation, collaborative exploration, and multirobot teleoperation and achieve shared goals within these tasks. The interaction between the robot and its human counterpart involved engaging in dialog to share information and control at key points in the task. For instance, when the robot encountered an obstacle, it asked the user, *Can I drive through <image>?* along with an image of the obstacle. In asking these questions, the robot drew on specific attributes of the user, such as response accuracy, expertise, availability, efficiency, and preferences to determine whether or not it should direct specific questions to its user.

A number of proposed models and systems take the dialog-based interaction paradigm further to involve the robot and its human counterpart jointly addressing the domain task and dialog itself as *joint action* [71.137, 138]. In this peer-to-peer setup, either party selects goals to address and strategies to be used to address them and either party performs any part of the task. The model proposed by Foster et al. [71.137] includes a semantic interpretation module and a central decision-making module which draw on resources, such as a history of the ongoing discourse between the robot and its user, a world model, a domain planner, and a representation of the plan that is currently being executed, in order to generate action and communication behaviors.

The model proposed by Li et al. [71.138] draws on joint intention theory [71.139], considering the joint activity to involve a common persistent goal of achieving conversational grounding, and explicitly uses elements

of grounding in representing conversational contributions. These contributions involve a *presentation* and an *acceptance* phase. For example, when an agent asks a question and the other agent answers, the question becomes the presentation and the answer becomes the acceptance, forming a *grounded exchange*. The model considers exchanges that involve a presentation without an acceptance to be *ungrounded*. Discourse contributions take place at two layers: *intention* and *conversation*. At the intention layer, the system plans communication intentions based on analyses of previous discourse and the robot’s control system. These intentions can be *self-* or *other-motivated* for each agent. The conversation layer involves the articulation of communication intentions through verbal and nonverbal behaviors. The two layers form an *interaction unit* (IU) in the model. The model determines whether an IU is presentation or acceptance and whether it is grounded or ungrounded by assessing whether it satisfies joint intentions of the agents. Figure 71.12 illustrates how an other-motivated exchange is assessed by the model to determine whether the exchange is a presentation or an acceptance.

Models of Situated Human–Robot Dialog

The models and systems described above consider task-based and communicative exchanges in HRI as a dialog and extend models of spoken dialog to accommodate requirements that are specific to HRI, such as task-management, mixed-initiative dialog management, and physically situated referencing. Research in cognitive HRI has also explored the development of dialog systems that explicitly integrate these mechanisms into dialog modeling and the development of specific models and mechanisms for these requirements.

An example of dialog systems that are specifically developed for situated human–robot dialog is the pattern-based mixed-initiative (PaMini) HRI framework [71.140]. This framework extends spoken dialog systems with two key components: a task-state protocol and interaction patterns. The task-state protocol component explicitly defines *tasks* that either the robot’s perceptual or control subsystems can perform. A task is defined as an execution state and preconditions for execution. The task-state protocol specifies task states and transitions among them to support coordination. The interaction patterns component provides high-level representations of recurring dialog structures such as a *clarification*. A comparison of most commonly used spoken dialog systems and the PaMini framework in the context of a human–robot situated learning scenario is provided by Peltason and Wrede [71.141].

Another example is the Robot Behavior Toolkit developed by Huang and Mutlu [71.3], which supports

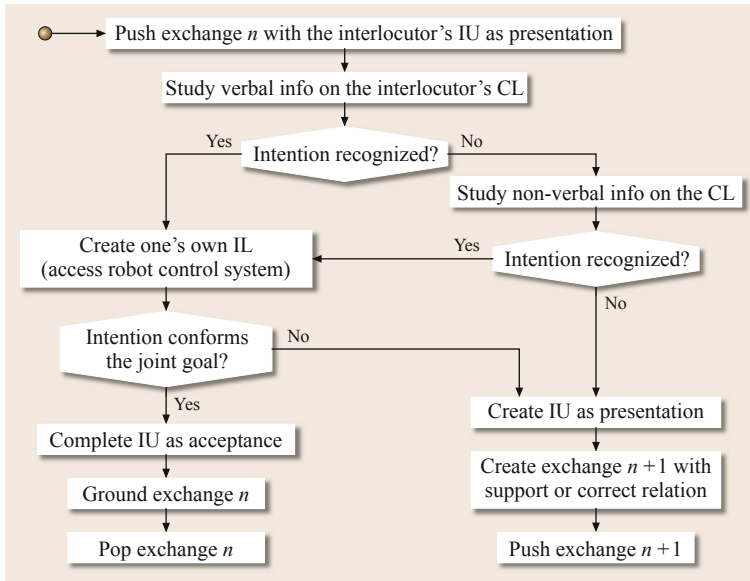


Fig. 71.12 The model evaluates an exchange provided by the interaction partner to determine its *presentation* or *acceptance* status and determine an appropriate action (after [71.138])

situated human–robot dialog by integrating nonverbal cues for task-based referential communication and conversation into the robot’s speech. This system uses a *repository* of specifications of situated communication cues based on models of human interactions and an *activity* model (described in more detail below) that specifies the joint human–robot activity including the agents, task context, shared task goals, and expected

task outcomes to integrate the situated communication cues that are expected to support these outcomes into the robot’s speech. Figure 71.13 displays an example behavior generated by the Toolkit in a collaborative manipulation task. An evaluation of their system showed that interactions in which the robot displayed these situated communication cues as directed by the system more effectively supported desired task outcomes compared with baseline interactions (VIDEO 128).

Research in cognitive HRI has also explored the development of models for specific communication and coordination mechanisms in situated interaction, such as perspective-taking, spatial referencing, reference resolution, and joint attention (VIDEO 129).

Perspective-Taking. A core process in situated interaction toward establishing common ground is *perspective-taking* [71.142]. Research in social cognition has shown that the ability to take another’s perspective and share common ground significantly improves collaborative performance in human teams [71.143]. Research in HRI has also explored how robots might employ this core mechanism to establish common ground with their users in situated interactions and has proposed several models that supported perspective taking.

Trafton et al. [71.144] studied interactions among astronauts in a naturalistic collaborative assembly task and found that a quarter of the utterances in the data involved taking the perspective of another and that participants frequently switched among egocentric, exocentric, addressee-centered, and object-centered perspectives. Based on their results, they developed a cog-

```

<behaviors>
<channel type='gaze'>
<action endTime='214.5' startTime='0' target='unspecified' />
<action endTime='1160' startTime='214.5' target='the green
object with one peg' />
<action endTime='2735.4' startTime='1160' target='unspecified' />
<action endTime='3597' startTime='2735.4' target='the red box' />
<action endTime='4308' startTime='3597' target='unspecified' />
<action endTime='4963' startTime='4308' target='listener' />
</channel>
<channel type='speech'>
    Could you help me put the green object with one peg into the red
    box, please?
</channel>
</behaviors>
    
```

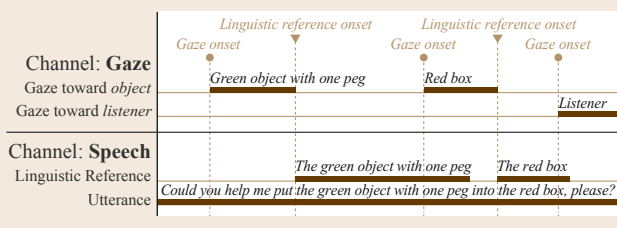


Fig. 71.13 The robot behavior toolkit uses specifications from a repository and a model of the joint activity to integrate effective multimodal task-related dialog behaviors (after [71.3])

nitive model of perspective taking that allowed the robot to maintain multiple perspectives – or *alternative worlds* – at once and explore propositions about these worlds, such as the perspective of an interaction partner. This exploration allowed the robot to make inferences about the perspective of its partner by simulating this alternative world and act on the world from this perspective. The following sequences of actions illustrate the simulations that the robot might carry out based on the command *go to the cone* (adapted from Trafton et al. [71.2]). Underlined text describes components of the system implementation:

Simulate current real world (i. e., perceive it)
Perception specialist notices the existence and location of *person*, *cone₁*, *cone₂*, and *obstacle*
Language specialist hears *Coyote*, *go to the cone* and infers that there is an object, *C*, that is a cone and that the *person* wants it to go to
Identity hypothesis specialist infers that *C* can be identical to *cone₁* or *cone₂*
C = *cone₁*, *C* = *cone₂*
Identity constraint specialist notices a contradiction
 This contradiction triggers the counterfactual simulation strategy

Simulate the world where *C* = *cone₁*
 Because in this world *person* has referred to *cone₁*, the perspective-simulation strategy is triggered

Simulate the world where *C* = *cone₁* and *robot* = *person*
 The spatial reasoning perspective indicates that *cone₁* does not exist in this world because *person* cannot see it
 Thus, *C* ≠ *cone₁*

Simulate the world where *C* = *cone₂*
 Because in this world *person* has referred to *cone₂*, the perspective-simulation strategy is triggered

Simulate the world where *C* = *cone₂* and *Robot* = *Person*
 Because *cone₂* is visible in this world, there is no contradiction in this world
 Infer that *C* = *cone₂* (i. e., *the cone* refers to *cone₂*)

Following a *counterfactual simulation strategy* provides the robot with the ability to make inferences about situated actions across alternative scenarios with alternative physical (e.g., whether or not an object is present) and cognitive (e.g., whether or not the object is visible to the human counterpart) characteristics and determine appropriate next actions, such as carrying out a request or seeking clarification from its human counterpart.

Figure 71.14 illustrates four alternative scenarios with different physical and cognitive properties explored by Trafton et al. [71.144]. In each scenario, the robot assesses these properties to determine its next actions, as illustrated below.

Algorithm 71.1

```

function: Scenario(nCones = 1 ⇒ conea)
if conea = visiblerobot ∧ conea = visiblehuman then
    Go to conea
end if
function: Scenario(nCones = 2 ⇒ conea, coneb)
if conea, coneb = visiblerobot ∧ conea = visiblehuman then
    Go to conea
end if
function: (Scenario(nCones = 1 ⇒ conea)
if conea, coneb ≠ visiblerobot ∧ conea = visiblehuman then
    Check hidden location
end if
function: Scenario(nCones = 2 ⇒ conea, coneb)
if conea, coneb = visiblerobot ∧ conea,
    coneb = visiblehuman then
    Request clarification
end if
    
```

Berlin et al. [71.145] developed a similar model that enabled the robot to understand its environment from the perspective of an interaction partner by maintaining

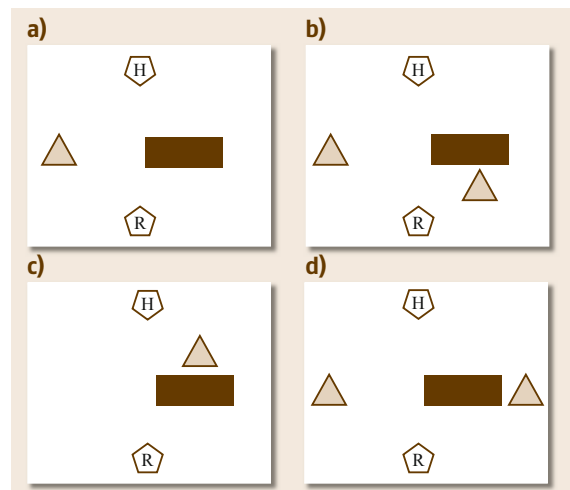


Fig.71.14a–d The alternative scenarios considered by the system in which the robot and its human counterpart are in a room with several objects and possible occlusions from the perspectives of the robot or the human (after [71.144])

separate and potentially different sets of beliefs in its belief system for itself and for its interaction partner. To construct a model of the beliefs of its interaction partner, the robot employed the same mechanisms it used to model its own beliefs but transformed the data it perceived from the world to match the reference frame of its interaction partner. These two sets of beliefs were maintained separately so that the robot can compare differences between its beliefs and its interaction partner's beliefs and plan actions in order to establish common ground or identify discrepancies in its learning in the context of task learning. Figure 71.15 illustrates parallel beliefs maintained by the robot in a button-pressing task.

Spatial Referencing. Moratz et al. [71.146] proposed a cognitive model of spatial reference that represented different kinds of spatial reference systems and allowed the robot to interpret instructions from an interaction partner. This model mapped the locations of all objects as projections on a plan view, considering the robot's point of view as *origin* and the location of the object that will be used as *relatum* to determine the reference axis. This axis enabled the robot to interpret directions such as *left of*, *right of*, *in front of*, and *to the back* in relation to the relatum, providing the robot the ability to interpret natural language references to objects in the environment.

Reference Resolution. Ros et al. [71.147] extended these approaches to develop a model that enabled the robot to clarify references made by its interaction partner. This model employed several mechanisms including visual perspective taking, spatial perspective taking, symbolic location descriptors, and feature descriptors to determine whether it needed any clarification on its interaction partner's references. The visual perspective taking mechanism allowed the robot to determine whether or not objects in the environment were in its interaction partner's focus of attention (FOA), in its part-

ner's field of view (FOV), or out of its partner's field of view (OOF). The spatial perspective taking mechanism maintained egocentric and addressee-centered perspectives to determine ambiguities in object references. The system also included symbolic location descriptions such as *is in*, *is on*, and *is next to* to determine spatial relationships between objects and the environment. Finally, the robot used feature descriptors such as color and shape to identify ambiguities in the references of its interaction partner. Once the robot determined the need clarification in its partner's references, it used an ontology-based clarification algorithm to ask questions to its partner about the object of reference.

Joint Attention. Another key mechanism in situated interaction is *joint attention* – the ability to use non-verbal cues, such as gaze and pointing, to establish common ground on what referents in the environment are under consideration in the dialogue [71.149]. Scassellati [71.99] proposed a task-based decomposition of joint attention skills, including mutual gaze, gaze following, imperative pointing, and declarative pointing, and implemented these skills in a robot as stages for establishing joint attention with a human counterpart. The mutual gaze skill provided the robot with the ability to recognize and maintain eye contact with its interaction partner. At the gaze following stage, the robot followed the eyes of its partner to direct its attention to the object of its partner's attention. Imperative pointing involved pointing at an object that is out of reach in order to request the object. Finally, the declarative pointing stage involved extending an arm and index finger to draw attention to an object that is out of reach without necessarily requesting the object.

Connection Events. Rich et al. [71.150] argued that mechanisms such as joint attention serve as *connection events* in situated dialog and establish and maintain engagement among interaction partners. From data on human interactions, they identified a set of key connec-

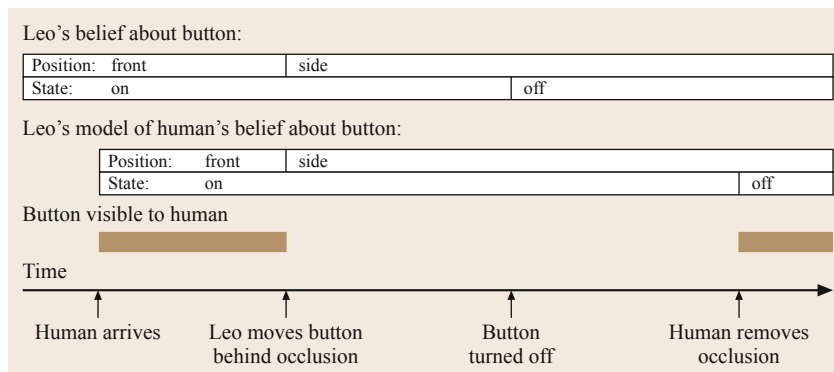


Fig. 71.15 In the system proposed by Berlin et al. [71.145], the robot maintains parallel beliefs for itself and for its human counterpart for the task and updates its beliefs based on sensory input and those of the user based on the user's awareness (after [71.145])

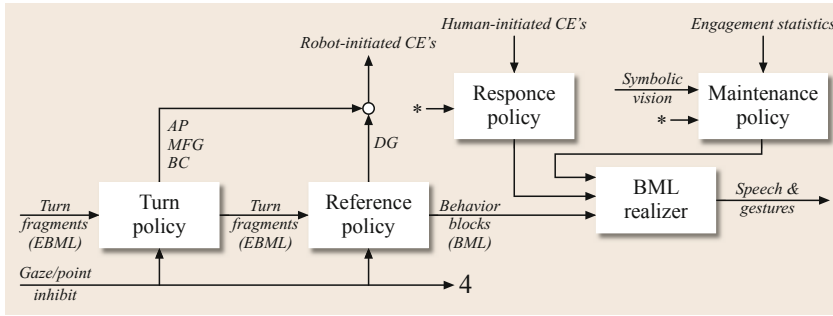


Fig. 71.16 The policy components, the BML realizer, and the generated connection events for the robot in the engagement generator module (after [71.148])

tion events, including mutual gaze, directed gaze, adjacency pairs, and backchannels, and developed a system that recognized these events in human counterparts and generated them for a robot (see Rich et al. [71.150] for details on the recognizer and Holroyd et al. [71.148] for details on generation). The recognizer module included dedicated recognizers for each type of connection event and an estimator for engagement levels for the robot’s human counterpart, while the generation module included four policy components and a behavior mark-up language (BML) realizer for generating robot behaviors toward establishing and maintaining engagement. The components of this engagement generator are illustrated in Fig. 71.16.

71.3.2 Simulation-Theoretic Models

Research in cognitive HRI has also been inspired by neurocognitive mechanisms in developing models of human–robot joint activity, building particularly on *simulation theory*, which suggests that people (and primates) represent other people’s mental states by adopting their perspective, specifically by tracking or matching their states with resonant states of their own [71.151]. This simulation-theoretic approach led to several models of robot behavior and human–robot joint action that involve the robot imitating or simulating the behaviors of its interaction partner in order to learn from or make inferences about its partner’s goals.

As an example of this approach, Bicho et al. [71.152] proposed a model for action preparation and decision-making in cooperative human–robot tasks that is inspired by the finding that action observation elicits an automatic activation of motor representations associated with the execution of the observed action. This *motor-resonance mechanism* allows people to internally simulate action consequences using their own motor repertoire and predict the consequences of action of others. In the proposed model, a perception–action linkage enables efficient coordination of actions and decisions between the agents in a human–robot joint action task.

The model integrates a mapping between observed actions and complementary actions in memory, while taking into account the inferred goals of the actions of the interaction partner, contextual cues, and shared task knowledge.

Building on simulation theory, Gray et al. [71.153] proposed a similar system in which the robot parses user actions and matches the user’s movements to movements in its own repertoire toward making inferences about the user’s goals and perform a task-level simulation (Fig. 71.17). This simulation allows the robot to determine the preconditions of the schemas that represent the task and track its human partner’s progress over the course of the task in order to anticipate its partner’s needs and offer relevant help accordingly. The simulation also provided the robot with the ability to make inferences on the beliefs of its partner and simulate its partner’s perspective in a fashion similar to the perspective-taking mechanisms proposed by Trafton et al. [71.144] and Berlin et al. [71.145] (VIDEO 130).

Aspects of the simulation-theoretic approach explicitly taken in these examples can also be seen in other control architectures developed for HRI. Nicolescu and Mataric [71.154] proposed a control architecture that unifies perception and action to achieve *action-based*

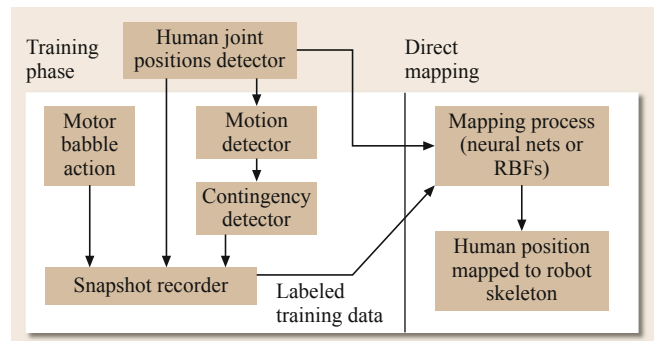


Fig. 71.17 The mapping of perceived human actions onto the robot’s body in order to make comparisons and task-level inferences (after [71.153])

interaction. In this architecture, behaviors are built from *perceptual* and *active* components. Perceptual components allow the robot to link its observations and actions and thus to learn to perform a task from the experiences it gains from its interactions with people. Active components enable task-based behaviors that also serve as implicit communication rather than explicit behaviors such as speech and gestures. Behavior representation in the architecture captures two types of behaviors: *abstract* and *primitive*. Abstract behaviors are explicit specifications of the behaviors' activation conditions (preconditions), goals in the form of abstracted environmental states, and effects (postconditions), while primitive behaviors are those that the robot performs to achieve these effects. By linking perceptions and actions, the robot learns what actions of its own might achieve the same observed effects.

71.3.3 Intention- and Activity-Based Models

The models and systems described above are concerned primarily with establishing and maintaining common ground and coordinating actions in task-based interactions using dialog- and simulation-theoretic approaches with limited consideration of the broader context of these interactions as complex activities involving multiple agents with common goals and commitments to these goals. A number of models and systems sought to address this limitation, building on models and theories of human joint activity such as joint intention theory [71.139] and activity theory [71.7].

Building on joint intention theory, *Breazeal et al.* [71.155] proposed a model of human-robot collaboration that involved dynamically meshing subplans into joint activity toward achieving common goals of the human-robot team. In this model, task and goal representations have a goal-centric view, employing an action-tuple data structure that captures preconditions, executables, until-conditions, and goals. Tasks are represented in a hierarchical structure of actions and recursively defined subtasks. Goals are also represented hierarchically as overall intent rather than a chain of low-level goals. The implemented joint intention model dynamically assigns tasks to members of the human-robot team. These intentions are derived based on the robot's actions and abilities, the actions of the human partner, the robot's understanding of the common goal of the team, and its assessment of the current task state. At every stage of the interaction, the robot negotiates who should complete the task. Action at these points might look like turn-taking or simultaneous action (the robot and the human working on different parts of the task).

Alami et al. [71.156, 157] similarly built on joint intention theory to propose a human-robot decision framework in which team members are committed to a *joint persistent goal* and follow cooperation schemes to contribute toward achieving this goal. The framework involves a goal planner called *the agenda* for the robot and human collaborators to pursue, a proxy representation of the human in the robot called *Interaction Agents (IAA)*, *task delegates* that monitor and control the task commitment of the human or the robot for each active, inactive, or suspended goal, and a *robot supervision kernel* that monitors and controls robot activities. For each new active goal, the Robot Supervision Kernel creates a Task Delegate, selects or elaborates a plan, and allocates the roles of each team member.

Fong et al. [71.1] proposed a similar system called the HRI operating system (HRI/OS) to support human-robot teamwork. The system involves a task manager, resource manager, interaction manager, spatial reasoning agent, context manager, human and robot agents, and an open agent architecture (OAA) facilitator. The task manager decomposes the overall goal of the system into high-level tasks and assigns to humans or robots for execution. The manager relies on the agents to complete the low-level steps of the tasks. It communicates with the Resource Manager to find an agent capable of performing the work. Resource manager processes all agent requests, prioritizing the list of agents to be consulted when a task needs to be performed. Interaction manager coordinates dialog-based communication between agents. Context manager keeps track of everything that occurs while the system is running including task status and execution, agent activities, agent dialogue, etc. Spatial reasoning agent (SRA) is used to resolve spatial ambiguities in human-robot dialog through mechanisms such as perspective taking and frames of reference, resolving ambiguities among as ego-, addressee-, object-, and exo-centric references. To do this, SRA transforms the spatial dialog into a geometric reference and perform a *mental simulation* of the interaction to explore how ambiguities might be resolved through multiple references. Finally, the OS includes a software representation of the human – a human proxy agent that represents user capabilities and accepts task assignments in the way that robot agents do. These proxies represent task capabilities, including domains of expertise, and provide health monitoring feedback.

Huang and Mutlu [71.3] built on an alternative model of human activity – activity theory [71.7] – to develop a model of human-robot joint activity. Their model builds on five key constructs from activity theory including consciousness, object-orientedness, hierarchical structure, internationalization and external-

ization, and mediation. The *consciousness* construct pertains to attention, intention, memory, reasoning, and speech and includes specific representations for attention and intention. The *object-orientedness* construct describes material artifacts, plans of action, or common ideas to be shared by the members of the joint activity. Following the *hierarchical structure* construct, the model organizes joint activity into three layers: activity, action, and operation. An activity consists of a series of actions that share the same goal, and each action has a defined goal and a chain of operations that are regular routines performed under a set of conditions. *Internalization and externalization* describes cognitive processes; internalization involves transforming external actions or perceptions into mental processes, while externalization is the process of manifesting mental processes in external actions. Finally, the *mediation* construct defines several external and internal tools, such as physical artifacts that might be used in an activity and cultural knowledge or social experience that an individual might have acquired, as mediators of human–robot joint activity. These constructs and their corresponding system elements allow the construction of and planning for joint human–robot activities. For each activity, a *motive* governs actions. Each action, by achieving its corresponding goal, helps to fulfill the motive of the activity. Each action may have several *operations* that are constrained by a set of conditions and that can be executed only when all the conditions are met. Actions have predefined *outcomes*, which specify the orientation of an action. Figure 71.18 shows the XML (extensible markup language) representation of a model of a collaborative manipulation task.

71.3.4 Models for Action Planning

The models described above primarily enable communication and coordination between humans and robots toward planning and carrying out joint tasks. In order to successfully contribute to these tasks, robots also need models for planning their actions in a dynamic physical and cognitive environment. Research in cognitive HRI seeks to develop models for *action planning* that help robots estimate the actions that they have to take in order to achieve task goals and learn the parameters of the tasks space. The paragraphs below review research in two common approaches to building such models: *decision-theoretic models* and *model learning*.

Decision-Theoretic Models

One of the simplest approaches to control and decision-making in HRI is to define the interaction as a decision-theoretic planning problem, such as a Markov decision process (MDP). Formally, an MDP consists of the n -

tuple $\{S, A, T, R, \gamma\}$. The set S is a set of states, which in the HRI setting typically correspond to the combination of state variables, such as the robot state and the desired outcome of the interaction. For example, if the interaction model allows a human partner to instruct the robot to move to different locations in the environment, one state variable may correspond to different current locations of the robot and another state variable may correspond to the goal states intended by the human partner. The full state space S is given by the combination of possible values for the different state variables.

The action set A represents actions that the robot may take. The actions may include asking a question, performing some physical movement, or even doing nothing. Each action has a cost R depending on the current state, which rewards the robot for performing useful actions, and penalizes the robot for taking actions that either make no immediate progress toward the specified goal (typically a small penalty) or completely unhelpful (a large penalty).

Lastly, the transition function T provides a notion of the dynamics of the environment in terms of how the state changes as robot takes actions, and especially how a human partner's state variables may change as the robot takes actions. The transition function $T(s'|s, a)$ places a probability distribution over the states to which the user in state s may transit if the robot takes action a . The MDP formulation is very appealing, because there exist efficient techniques for solving for interaction policies. Once the policy is computed, the interaction can be managed simply by querying the policy for the

```
<Activity id='1'>
  <Motive>clear(table)</Motive>
  <Description>Clear objects on table</Description>
  <Participants>Self, User1</Participants>
  <Action id='1'>
    <Outcome>Task</Outcome>
    <Goal>disappear(object)</Goal>
    <Description>
      Instruct User1 to categorize object
    </Description>
    <Operation type='utterance'>
      <Condition>present(User1)</Condition>
      <Condition>
        known(the blue object with two pegs)
      </Condition>
      <Condition> known(the blue box)</Condition>
      <Info turn='end'>
        Could you help me put the blue object with
        two pegs into the blue box, please?
      </Info>
    </Operation>
    ...
  </Action>
</Activity>
```

Fig. 71.18 The XML representation of the activity-theory-based model for a collaborative manipulation task (after [71.3])

appropriate action in response to the current state of the robot and the human partner.

A limitation of the MDP approach is that some of the state variables may not be directly observable, in particular the state variables corresponding to human intentional states, such as intended goal locations of the robot. The values of the state variables must be inferred from observations, such as speech acts performed by the human partner, which are inherently noisy. For example, the system may hear the words *coffee machine* when the user asks the robot to go to *copy machine*. While speech recognition errors may be mitigated to some extent by asking the user to use only acoustically distinct keywords when speaking to the system, a system that does not model the likelihood of recognition errors and act accordingly will be brittle; a robust system must be able to infer user intent under uncertainty.

The observations are rarely sufficient to uniquely determine the current state, but more commonly are used to compute a *belief*, or probability distribution over dialog states. If the agent takes some action a and hears observation o from an initial belief b , it can easily update its belief using Bayes rule

$$b^{a,o}(s) = \frac{\Omega(o|s', a) \sum_{s \in S} T(s'|s, a)b(s)}{\sum_{\sigma \in S} \Omega(o|\sigma, a) \sum_{s \in S} T(\sigma|s, a)b(s)} \tag{71.1}$$

This probability distribution will evolve as the dialog manager asks clarification questions and receives responses. In Fig. 71.19, we show a cartoon of a simple dialog model. Initially, we model the user as being in a *start* state. Then, at some point in time, the user speaks to the robot to indicate that he or she wants it to perform a task. We denote this step by the set of vertical stack of nodes in the center of the model. Each node represents a different task. The dialog manager must now interact with the user to determine what is wanted. Once the task is successfully completed, the user transitions to the right-most *end* node, in which he or she again does not desire anything from the robot. We note that it can be easily augmented to handle more complex scenarios. For example, by including the time of day as part of the state, we can model the fact that the user may usually wish to go to certain locations in the morning and other locations in the afternoon.

Intuitively, we can see how the belief can be used to select an appropriate action. For example, if the dialog manager believes that the user may wish to go to either the coffee machine or the copy machine (but not the printer), then it may ask the user for clarification before commanding the wheelchair to one of the locations. More formally, we call the mapping from beliefs

to actions a policy. We represent this mapping using the concept of a value function $V(b)$. The value of a belief is defined to be the expected long-term reward the dialog manager will receive if it starts a user interaction in belief b . The optimal value function is piecewise-linear and convex, so we represent V with the vectors V_i ; $V(b) = \max_i V_i \cdot b$. The optimal value function satisfies the Bellman equation [71.158]

$$V(b) = \max_{a \in A} Q(b, a),$$

$$Q(b, a) = R(b, a) + \gamma \sum_{o \in O} \Omega(o|b, a)V(b_a^o), \tag{71.2}$$

where $Q(b, a)$ represents the expected reward for starting in belief b , performing action a , and then acting optimally. The belief b_a^o is b after a Bayesian update of b using (71.1), and $\Omega(o|b, a)$, the probability of seeing o after performing a in belief b ($\sum_{s \in S} \Omega(o|s, a)b(s)$).

There are also non-Bayesian approaches for acting in uncertain environments. Many interaction systems provide the dialog manager with a set of rules to follow given particular outputs from a speech recognition system. The drawback to rule-based systems is that they often have difficulty managing the many uncertainties that stem from noisy speech recognition or linguistic ambiguities. The ability to manage the trade-off between gathering additional information and servicing a user's request have made partially observable Markov decision process (POMDP) planners particularly useful in dialog management; applications include a Nursebot robot, designed to interact with the elderly in nursing homes [71.159], a vision-based system that aids Alzheimer's patients with basic tasks such as hand-washing [71.160], an automated tele-

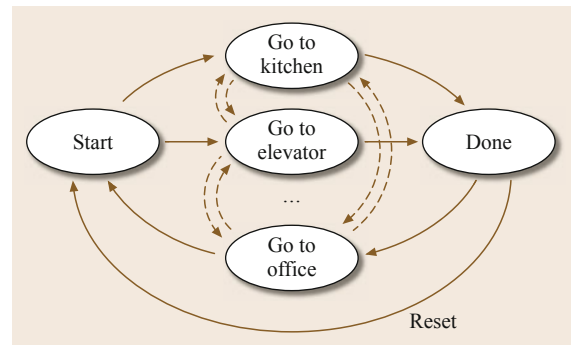


Fig. 71.19 A toy example of a dialog POMDP. The *nodes* in the graph are different states of the dialog (i. e., user requests). *Solid lines* indicate likely transitions; we assume that the user is unlikely to change their request before their original request is fulfilled. The system automatically resets once we reach the end state

phone operator [71.161], and a tourist information kiosk [71.162].

Beyond the initial formulations of cognitive HRI as a decision-theoretic problem, there have been a number of algorithmic improvements that increase the domains of applicability of this approach. For example, the conventional MDP and POMDP algorithms have typically assumed that each observation and action takes approximately the same amount of time, which can lead to an implicit bias toward longer actions. Representing time explicitly leads to computational intractability, but Broz et al. [71.163] demonstrated that the similar states that vary only by the time-index can be aggregated, leading to reduced-order models that can be solved very efficiently. Similarly, Doshi and Roy [71.164] showed that symmetries in human intentional states could be exploited to dramatically reduce the size of the planning problem, also leading to very efficient solutions. Most recently, again in the non-Bayesian line, Wilcox et al. have shown that the temporal dynamics of task-based HRI can be formulated as a scheduling problem [71.165].

Model Learning

The behavior of the dialog manager derived from solving (71.2) depends critically on accurate choices of the transition probabilities, observation probabilities, and the reward. For example, the observation parameters affect how the system associates particular keywords with particular requests. Similarly, the reward function affects how aggressive the dialog manager will be in assuming that it understands a user's request, given limited and noisy information. An incorrect specification of the dialog model may lead to behavior that is either overly optimistic or conservative, depending on how accurately the model captures the user's expectations on the interaction.

A common approach in other domains is to collect data using a fixed policy, typically referred to as system identification. In HRI, this is easiest to perform using so-called *Wizard of Oz* studies where a human experimenter executes the policy unseen to generate data or evaluate a policy. Prommer et al. [71.166] showed that Wizard-of-Oz studies could be used effectively not only to learn model parameters for an MDP dialog model, but also to learn an effective policy.

At the same time, learning all the parameters required to specify a rich dialog model can require a prohibitively large amount of data. While the model parameters may be difficult to specify exactly, either by hand or from data, we can often provide the dialog manager with an initial estimate of the model parameters that will generate a reasonable policy that can be executed while the model is improved. For example, even

though we may not be able to attach an exact numerical value to driving a wheelchair user to the wrong location, we can at least specify that this behavior is undesirable. Similarly, we can specify that the exact numerical value is initially uncertain. As data about model parameters accumulate, the parameter estimates should converge to the correct underlying model with a corresponding reduction in uncertainty.

Figure 71.20a depicts the conventional model, where the arrows in the graph show which parts of the model affect each other from time t to $t + 1$. Although the variables below the *hidden* line in Fig. 71.20a are not directly observed by the dialog manager, the parameters defining the model (i.e., the parameters in the function giving the next state) are fixed and known

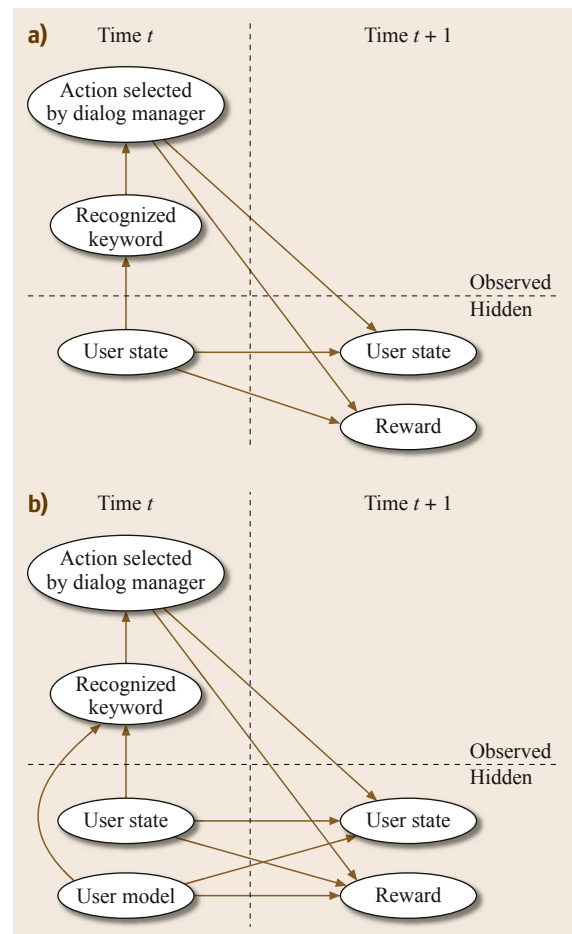


Fig. 71.20 (a) The standard POMDP model. (b) The extended POMDP model. In both cases, the arrows show which parts of the model are affected by each other from time t to $t + 1$. Not drawn are the dependencies from time $t + 1$ onward, such as the user state and user model's effect on the recognized keyword at time $t + 1$

a priori. For instance, the reward at time t is a function of the state at the previous time and the action chosen by the dialog manager.

If the model parameters are not known a priori because the model is uncertain – for example, how much reward is received by the agent given the previous state and the action selected – then the concept of the belief can be extended to also include the agent’s uncertainty over possible models. In this new representation, which we call the *model-uncertainty POMDP*, both the user’s request and the model parameters are hidden. Figure 71.20b shows this extended model, in which the reward at time t is still a function of the state at the previous time and the action chosen by the dialog manager, but the parameters are not known a priori and are therefore hidden model variables that must be estimated along with the user state. The system designer can encode their knowledge of the system in the dialog manager’s initial belief over what dialog models it believes are likely – a Bayesian prior over models – and let the agent improve upon this belief with experience.

Poupart et al. treated the unknown MDP parameters as hidden state in a larger POMDP and derived an analytic solution (based on [71.167]) for a policy that will trade optimally between learning the MDP and maximizing reward. Unfortunately, these techniques did not extend tractably to the model-uncertainty POMDP, which is continuous in both the POMDP parameters (like the MDP) and the belief state (unlike the MDP). Doshi and Roy [71.168, 169] provided an approximate, Bayes risk action selection criterion that allows the dialog manager to function in this complex space of dialog models. This approach was applied to the intelligent wheelchair assistant shown in Fig. 71.21. Their goal was to design an adaptable HRI system, or dialog manager, that allows both the user of the wheelchair and a caregiver to give natural instructions to the wheelchair, as well as ask the wheelchair computer for general information that may be relevant to the user’s daily life.

In contrast to the Bayesian approach, Cakmak and Thomaz [71.170] pursued an active learning approach and identified three types of queries that a robot could generate while learning a new task (VIDEO 237). While this result does not provide a comparison to an approach embedded in an ongoing dialogue, their results do provide guidelines for model designers.

71.3.5 Cognitive Models of Robot Control

A final line of research in cognitive HRI seeks to achieve greater task efficiency in human–robot teams, thus addressing common problems between opera-

tors and robots such as those identified by Burke et al. [71.133] and Stubbs et al. [71.134], by developing models and control interfaces that exploit mechanisms of human cognition such as working memory and mental models [71.171, 172]. This research includes formalisms such as *neglect time*, the amount of time that an operator can neglect a robot before the robot’s performance drops below a certain threshold [71.171], and *fan out*, a measure of how many robots an operator can effectively manage in a human–robot team [71.172]. Such formalisms inform the development of guidelines for designing effective control mechanisms such as the following principles proposed by Goodrich and Olsen [71.171]:

1. *Implicitly switch interfaces and autonomy modes.* Context determines the mode of use. For instance, the user starts using a joystick and the interaction modality automatically switches, rather than the user explicitly selecting a modality.
2. *Let the robot use natural human cues.* The robot uses the cues to provide feedback and present information that the human uses to provide the commands or present information to the robot.
3. *Manipulate the world instead of the robot.* Control interfaces integrate knowledge about the task and the world to minimize low-level control of the robot and maintaining of a mental model of the robot’s functioning.
4. *Manipulate the relationship between the robot and world.* Control interfaces provide real-world

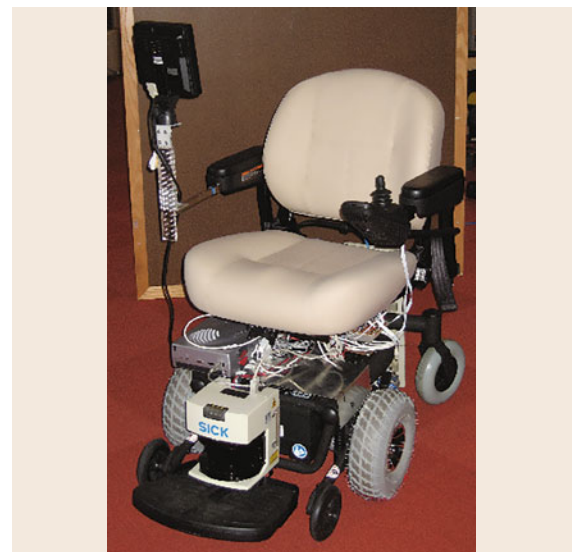


Fig. 71.21 Our dialog manager allows for more natural human communication with a robotic wheelchair (after [71.168, 169])

- representations for control to minimize low-level control.
5. *Let people manipulate presented information.* Interfaces present information in a way that represents the real world and allows users to provide input directly into the representation rather than translating information readings to a different modality or representation.
 6. *Externalize memory.* Different types of information are integrated into a single representation to reduce the working memory load for the user.
 7. *Help people manage attention.* The robot provides appropriate indicators to capture the attention of the operator.
 8. *Learn.* Control mechanisms adapt system activity to the user’s mental models.

71.4 Conclusion and Further Reading

This chapter presented an overview of research in cognitive human–robot interaction, the area of research concerned with modeling human, robot, or joint human–robot cognitive processes in the context of HRI. This research seeks to gain a better understanding of people’s interaction with robots and build robotic systems with the necessary cognitive mechanisms to communicate and collaborate with their human counterparts. Three key themes fall within this research area. The first theme seeks to build a better understanding of human cognition in HRI; specifically, people’s mental models of robots as ontological entities, social cognition of robot behaviors, and the use of robots as experimental platforms to study cognitive development in humans. The second theme includes research that seeks to build models for simulating human cognition in robots, gaining cognitive capabilities through imitation and interaction with the physical environment, and mapping aspects of interaction, such as commands from or references by human counterparts to objects in the environment. The final theme seeks to build models that support human–robot joint activity, including dialog-, simulation-theoretic-, joint-intention-, activity- and action-planning-based models that enable robots to reason about the physical and cognitive properties of the environment and the actions of their human counterparts and to plan actions toward achieving communicative or collaborative goals. The common thread among these three themes of research is the consideration of humans and robots as part of a cognitive



system in which cognitive processes – natural or designed – shape how humans and robots communicate and collaborate.

As an interdisciplinary area of research, cognitive human-robot interaction receives contributions from a diverse set of research fields including robotics, cognitive science, social psychology, communication studies, and science and technology studies. Further reading on the topic is also available in a diverse set of venues such as:

- *The Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction (HRI)*
- *The Proceedings of the Annual Meeting of the Cognitive Science Society (CogSci)*
- *International Conference on Epigenetic Robotics (EpiRob)*
- *The Proceedings of the AAAI Conference on Artificial Intelligence*
- *The Proceedings of the IEEE International Symposium on Robots and Human Interactive Communication (RO-MAN)*
- *The Proceedings of the Robotics: Science and Systems (RSS) Conference*
- Sun, [71.173]
- *Journal of Human–Robot Interaction*
- *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems.* John Benjamins
- *International Journal of Social Robotics.* Sage.

Video-References

- 📺 VIDEO 128 Gaze and gesture cues for robots available from <http://handbookofrobotics.org/view-chapter/71/videodetails/128>
- 📺 VIDEO 129 Robotic secrets revealed, Episode 1 available from <http://handbookofrobotics.org/view-chapter/71/videodetails/129>
- 📺 VIDEO 130 Robotic secrets revealed, Episode 2: The trouble begins available from <http://handbookofrobotics.org/view-chapter/71/videodetails/130>
- 📺 VIDEO 236 Human-robot jazz improvisation available from <http://handbookofrobotics.org/view-chapter/71/videodetails/236>

-  VIDEO 237 Designing robot learners that ask good questions available from <http://handbookofrobotics.org/view-chapter/71/videodetails/237>
-  VIDEO 238 Active keyframe-based learning from demonstration available from <http://handbookofrobotics.org/view-chapter/71/videodetails/238>

References

- 71.1 T. Fong, C. Kunz, L.M. Hiatt, M. Bugajska: The human-robot interaction operating system, Proc. 1st ACM SIGCHI/SIGART Conf. HRI, Salt Lake City (2006) pp. 41–48
- 71.2 J.G. Trafton, A.C. Schultz, N.L. Cassimatis, L.M. Hiatt, D. Perzanowski, D.P. Brock, M.D. Bugajska, W. Adams: Communicating and collaborating with robotic agents. In: *Cognition and Multi-Agent Interaction*, ed. by R. Sun (Cambridge Univ. Press, New York 2006) pp. 252–278
- 71.3 C.-M. Huang, B. Mutlu: Robot behavior toolkit: Generating effective social behaviors for robots, Proc. 7th ACM/IEEE Intl. Conf. HRI, Boston (2012) pp. 25–32
- 71.4 A.H. Vera, H.A. Simon: Situated action: A symbolic interpretation, *Cogn. Sci.* **17**(1), 7–48 (1993)
- 71.5 T. Winograd, F. Flores: *Understanding Computers and Cognition: A New Foundation for Design* (Ablex Publ., New York 1986)
- 71.6 L.A. Suchman: *Plans and Situated Actions: The Problem of Human-Machine Communication* (Cambridge Univ. Press, Cambridge 1987)
- 71.7 A.N. Leont'ev: The problem of activity in psychology, *J. Russ. East Eur. Psychol.* **13**(2), 4–33 (1974)
- 71.8 E. Hutchins: The social organization of distributed cognition. In: *Perspectives on Socially Shared Cognition*, ed. by L.B. Resnick, J.M. Levine, S.D. Teasley (American Psychological Association, Washington, DC 1991)
- 71.9 B. Gates: A robot in every home, *Sci. Am.* **296**, 58–65 (2007)
- 71.10 C. Pantofaru, L. Takayama, T. Foote, B. Soto: Exploring the role of robots in home organization, Proc. 7th Annu. ACM/IEEE Intl. Conf. HRI, Boston (2012) pp. 327–334
- 71.11 K.M. Tsui, M. Desai, H.A. Yanco, C. Uhlik: Exploring use cases for telepresence robots, ACM/IEEE 6th Intl. Conf. HRI, Lausanne (2011) pp. 11–18
- 71.12 F. Tanaka, A. Cicourel, J.R. Movellan: Socialization between toddlers and robots at an early childhood education center, Proc. Natl. Acad. Sci. USA **104**(46), 17954–17958 (2007)
- 71.13 T. Kanda, R. Sato, N. Saiwaki, H. Ishiguro: A two-month field trial in an elementary school for long-term human-robot interaction, *IEEE Trans. Robotics* **23**(5), 962–971 (2007)
- 71.14 B. Reeves, C. Nass: *The Media Equation: How People Treat Computers, Television and New Media Like Real People and Places* (Cambridge University Press, Cambridge 1996)
- 71.15 S. Turkle, O. Daste, C. Breazeal, B. Scassellati: Encounters with Kismet and Cog: Children respond to relational artifacts, Proc. IEEE-RAS/RJS Int. Conf. Humanoid Robots, Los Angeles (2004) pp. 1–20
- 71.16 S. Turkle: *Alone Together: Why We Expect More from Technology and Less from Each Other* (Basic Books, New York, 2011)
- 71.17 C. Nass, Y. Moon: Machines and mindlessness: Social responses to computers, *J. Soc. Issues* **56**(1), 81–103 (2000)
- 71.18 S. Turkle: *Evocative Objects: Things We Think With* (MIT Press, Cambridge 2011)
- 71.19 K. Wada, T. Shibata, Y. Kawaguchi: Long-term robot therapy in a health service facility for the aged – A case study for 5 years, Proc. 11th IEEE Int. Conf. Rehabil. Robotics, Kyoto (2009) pp. 930–933
- 71.20 B.R. Duffy: Anthropomorphism and the social robot, *Robotics Auton. Syst.* **42**(3–4), 177–190 (2003)
- 71.21 S. Kiesler, A. Powers, S.R. Fussell, C. Torrey: Anthropomorphic interactions with a software agent and a robot, *Soc. Cogn.* **26**(2), 168–180 (2008)
- 71.22 S.R. Fussell, S. Kiesler, L.D. Setlock, V. Yew: How people anthropomorphize robots, Proc. 3rd ACM/IEEE Int. Conf. HRI, Amsterdam (2008) pp. 145–152
- 71.23 D.S. Syrdal, K. Dautenhahn, S.N. Woods, M.L. Walters, K.L. Koay: Looking good? Appearance preferences and robot personality inferences at zero acquaintance, AAAI Spring Symp.: Multidiscip. Collab. Socially Assist. Robotics, Stanford (2007) pp. 86–92
- 71.24 K.F. MacDorman, T. Minato, M. Shimada, S. Itakura, S. Cowley, H. Ishiguro: Assessing human likeness by eye contact in an android testbed, Proc. XXVII Annu. Meet. Conf. Cogn. Sci. Soc., Stresa (2005) pp. 1373–1378
- 71.25 F. Hegel, S. Krach, T. Kircher, B. Wrede, G. Sagerer: Understanding social robots: A user study on anthropomorphism, Proc. 17th IEEE Int. Symp. Robot Hum. Interact. Commun., Munich (2008) pp. 574–579
- 71.26 M. Mori: The uncanny valley, *Energy* **7**(4), 33–35 (1970)
- 71.27 C. Bartneck, T. Kanda, H. Ishiguro, N. Hagita: My robotic Doppelgänger – A critical look at the Uncanny Valley Theory, IEEE 18th Intl. Symp. Robot Hum. Interact. Commun., Toyama (2009) pp. 269–276
- 71.28 M.L. Walters, D.S. Syrdal, K. Dautenhahn, R. te Boekhorst, K.L. Koay: Avoiding the uncanny valley: Robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion, *Auton. Robots* **24**(2), 159–178 (2008)

- 71.29 A.P. Saygin, T. Chaminade, H. Ishiguro, J. Driver, C. Frith: The thing that should not be: Predictive coding and the uncanny valley in perceiving human and humanoid robot actions, *Soc. Cogn. Affect. Neurosci.* **7**(4), 413–422 (2012)
- 71.30 W. Mitchell, K.A. Szerszen Sr., A.S. Lu, P.W. Schermerhorn, M. Scheutz, K.F. MacDorman: A mismatch in the human realism of face and voice produces an uncanny valley, *i-Perception* **2**, 10–12 (2011)
- 71.31 C. Breazeal: *Designing Sociable Robots* (MIT Press, Cambridge 2002)
- 71.32 N. Matsumoto, H. Fujii, M. Okada: Minimal design for human-agent communication, *Artif. Life Robotics* **10**(1), 49–54 (2006)
- 71.33 H. Kozima, H. Yano: A Robot that Learns to Communicate with Human Caregivers, *Proc. 1st Int. Workshop Epigenetic Robotics*, Lund (2001) pp. 47–52
- 71.34 A. Powers, A.D.I. Kramer, S. Lim, J. Kuo, S.-I. Lee, S. Kiesler: Eliciting information from people with a gendered humanoid robot, *IEEE 14th Int. Workshop Robot Hum. Interact. Commun.*, Nashville (2005) pp. 158–163
- 71.35 K.M. Lee, N. Park, H. Song: Can a robot be perceived as a developing creature?: Effects of a robot's long-term cognitive developments on its social presence and people's social responses toward it, *Human Commun. Res.* **31**(4), 538–563 (2005)
- 71.36 J. Goetz, S. Kiesler, A. Powers: Matching robot appearance and behavior to tasks to improve human-robot cooperation, *Proc. 12th IEEE Int. Workshop Robot Hum. Interact. Commun.*, Silicon Valley (2003) pp. 55–60
- 71.37 M.K. Lee, S. Kiesler, J. Forlizzi, S. Srinivasa, P. Rybski: Gracefully mitigating breakdowns in robotic services, *Proc. 6th ACM/IEEE Int. Conf. HRI*, Lausanne (2010) pp. 203–210
- 71.38 B. Shore: *Culture in Mind: Cognition, Culture, and the Problem of Meaning* (Oxford Univ. Press, Oxford 1996)
- 71.39 V. Evers, H. Maldonado, T. Brodecki, P. Hinds: Relational vs. group self-construal: Untangling the role of national culture in HRI, *Proc. 3rd ACM/IEEE Int. Conf. HRI*, Amsterdam (2008)
- 71.40 L. Wang, P.-L.P. Rau, V. Evers, B.K. Robinson, P. Hinds: When in Rome: The role of culture and context in adherence to robot recommendations, *Proc. 5th ACM/IEEE Int. Conf. HRI*, Osaka (2010) pp. 359–366
- 71.41 S. Sabanovic: Robots in society, society in robots – Mutual shaping of society and technology as a framework for social robot design, *Int. J. Soc. Robotics* **2**(4), 439–450 (2010)
- 71.42 G. Shaw-Garlock: Looking forward to sociable robots, *Int. J. Soc. Robotics* **1**(3), 249–260 (2009)
- 71.43 P.H. Kahn, A.L. Reichert, H.E. Gary, T. Kanda, H. Ishiguro, S. Shen, J.H. Ruckert, B. Gill: The new ontological category hypothesis in human-robot interaction, *Proc. 6th ACM/IEEE Int. Conf. HRI*, Lausanne (2011) pp. 159–160
- 71.44 P.H. Kahn, N.G. Freier, B. Friedman, R.L. Severson, E.N. Feldman: Social and moral relationships with robotic others?, *IEEE 13th Int. Workshop Robot Hum. Interact. Commun.*, Kurashiki (2004) pp. 545–550
- 71.45 S. Turkle: *A Nascent Robotics Culture: New Complicities for Companionship* (AAAI, Boston 2006)
- 71.46 B. Scassellati: How developmental psychology and robotics complement each other, *NSF/DARPA Workshop Dev. Learn.* (MIT Press, CSAIL, Cambridge 2006)
- 71.47 H. Ishiguro: Android science – toward a new cross-interdisciplinary framework, *ICCS/CogSci Workshop Toward Soc. Mech. Android Sci.*, Stresa (2005) pp. 1–6
- 71.48 K.F. MacDorman, H. Ishiguro: The uncanny advantage of using androids in cognitive and social science research, *Interact. Stud.* **7**(3), 297–337 (2006)
- 71.49 M. Stanley, J. Sabini: On maintaining social norms: A field experiment in the subway. In: *Advances in Environmental Psychology: The Urban Environment*, ed. by A. Baum, J.E. Singer, S. Valins (Erlbaum Associates, Hillsdale 1978) pp. 31–40
- 71.50 H. Kozima, M.P. Michalowski, C. Nakagawa: Keepon: A playful robot for research, therapy, and entertainment, *Int. J. Soc. Robotics* **1**(1), 3–18 (2009)
- 71.51 K.F. MacDorman: Introduction to the special issue on android science, *Connect. Sci.* **18**(4), 313–317 (2006)
- 71.52 J.J. Gibson: *The Ecological Approach to Visual Perception* (Houghton Mifflin, Boston 1979)
- 71.53 E.S. Reed: *Encountering the World: Toward an Ecological Psychology* (Oxford Univ. Press, Oxford 1996)
- 71.54 Y. Yamaji, T. Miyake, Y. Yoshiike, P.R.S. De Silva, M. Okada: STB: Human-dependent sociable trash box, *Proc. 5th ACM/IEEE Int. Conf. HRI*, Osaka (2010) pp. 197–198
- 71.55 H. Ishiguro: Android science: Conscious and sub-conscious recognition, *Connect. Sci.* **18**(4), 319–332 (2006)
- 71.56 S. Nishio, H. Ishiguro, N. Hagita: Geminoid: Teleoperated android of an existing person. In: *Humanoid Robots, New Developments*, ed. by A.C. De Pina Filho (InTech, Vienna 2007) pp. 343–352
- 71.57 S. Nishio, H. Ishiguro, N. Hagita: Can a teleoperated robot represent personal presence? – A case study with children, *Psychologia* **50**(4), 330–342 (2007)
- 71.58 M. Shimada, K. Yamauchi, T. Minato, H. Ishiguro, S. Itakura: Studying the influence of the chameleon effect on humans using an android, *IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nice (2008)
- 71.59 J. Wainer, D.J. Feil-Seifer, D.A. Shell, M.J. Mataric: Embodiment and human-robot interaction: A taskbased perspective, *Proc. 2nd ACM/IEEE Int. Conf. HRI*, Washington (2007) pp. 872–877
- 71.60 P. Schermerhorn, M. Scheutz: Disentangling the effects of robot affect, embodiment, and au-

- tonomy on human team members in a mixed-initiative task, Proc. 4th Int. Conf. Adv. Comput.–Hum. Interact., Gosier (2011) pp. 235–241
- 71.61 W.A. Bainbridge, J.W. Hart, E.S. Kim, B. Scasselati: The benefits of interactions with physically present robots over video–displayed agents, Int. J. Soc. Robotics 1(2), 41–52 (2010)
- 71.62 B. Mutlu: Designing embodied cues for dialog with robots, AI Magazine 32(4), 17–30 (2011)
- 71.63 E.T. Hall: *The Hidden Dimension* (Anchor Books, New York 1966)
- 71.64 L. Takayama, C. Pantofaru: Influences on proxemic behaviors in human–robot interaction, IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS), St. Louis (2009) pp. 5495–5502
- 71.65 D.S. Syrdal, K.L. Koay, M.L. Walters, K. Dautenhahn: A personalized robot companion? The role of individual differences on spatial preferences in HRI scenarios, Proc. 16th IEEE Int. Symp. Robot Hum. Interact. Commun., Jeju Island (2007) pp. 1143–1148
- 71.66 J. Mumm, B. Mutlu: Human–robot proxemics: Physical and psychological distancing in human–robot interaction, Proc. 6th ACM/IEEE Int. Conf. HRI, Lausanne (2011) pp. 331–338
- 71.67 M.L. Walters, M.A. Oskoei, D.S. Syrdal, K. Dautenhahn: A long–term human–robot proxemic study, Proc. 20th IEEE Int. Symp. Robot Hum. Interact. Commun., Atlanta (2011) pp. 137–142
- 71.68 C. Yu, M. Scheutz, P. Schermerhorn: Investigating multimodal real–time patterns of joint attention in an HRI word learning task, Proc. 5th ACM/IEEE Int. Conf. HRI, Osaka (2010) pp. 309–316
- 71.69 T. Yonezawa, H. Yamazoe, A. Utsumi, S. Abe: Gaze–communicative behavior of stuffed–toy robot with joint attention and eye contact based on ambient gaze–tracking, Proc. 9th Int. Conf. Multimodal Interfaces, Nagoya (2007) pp. 140–145
- 71.70 Y. Yoshikawa, K. Shinozawa, H. Ishiguro, N. Hagita, T. Miyamoto: Responsive robot gaze to interaction partner, Robotics Sci. Syst., Philadelphia (2006)
- 71.71 B. Mutlu, T. Shiwa, T. Kanda, H. Ishiguro, N. Hagita: Footing in human–robot conversations: How robots might shape participant roles using gaze cues, Proc. 4th ACM/IEEE Int. Conf. HRI, San Diego, California (2009) pp. 61–68
- 71.72 M. Staudte, M.W. Crocker: Visual attention in spoken human–robot interaction, Proc. 4th ACM/IEEE Int. Conf. HRI, San Diego (2009) pp. 77–84
- 71.73 B. Mutlu, J. Forlizzi, J.K. Hodgins: A storytelling robot: Modeling and evaluation of human–like gaze behavior, IEEE–RAS Conf. Humanoid Robots, Genoa (2006) pp. 518–523
- 71.74 C. Yu, P. Schermerhorn, M. Scheutz: Adaptive eye gaze patterns in interactions with human and artificial agents, ACM Trans. Interact. Intell. Syst. 1(2), 1–25 (2012)
- 71.75 H. Admoni, C. Bank, J. Tan, M. Toneva, B. Scasselati: Robot gaze does not reflexively cue human attention, Proc. 33rd Annu. Conf. Cogn. Sci. Soc., Boston (2011) pp. 1983–1988
- 71.76 W.S. Condon: Cultural microrhythms. In: *Interaction Rhythms: Periodicity in Communicative Behavior*, ed. by M. Davis (Human Sciences Press, New York 1982) pp. 53–76
- 71.77 E. Goffman: Some context for content analysis: A view of the origins of structural studies of face–to–face interaction. In: *Conducting Interaction: Patterns of Behavior in Focused Encounters*, ed. by A. Kendon (Cambridge Univ. Press, Cambridge 1990) pp. 15–49
- 71.78 C. Trevarthen: Can a robot hear music? Can a robot dance? Can a robot tell what it knows or intends to do? Can it feel pride or shame in company? – Questions of the nature of human vitality, Proc. 2nd Int. Workshop Epigenet. Robotics, Edinburgh (2002)
- 71.79 M. Michalowski, S. Sabanovic, H. Kozima: A dancing robot for rhythmic social interaction, Proc. 2nd ACM/IEEE Int. Conf. HRI, Washington DC (2007) pp. 89–96
- 71.80 M.P. Michalowski, R. Simmons, H. Kozima: Rhythmic attention in child–robot dance play, Proc. 18th IEEE Int. Symp. Robot Hum. Interact. Commun., Toyama (2009) pp. 816–821
- 71.81 E. Avrunin, J. Hart, A. Douglas, B. Scasselati: Effects related to synchrony and repertoire in perceptions of robot dance, Proc. 6th ACM/IEEE Int. Conf. HRI, Lausanne (2011) pp. 93–100
- 71.82 G. Hoffman, C. Breazeal: Anticipatory perceptual simulation for human–robot joint practice: Theory and application study, Proc. 23rd AAAI Conf. Artif. Intell., Chicago (2008) pp. 1357–1362
- 71.83 G. Hoffman, G. Weinberg: Interactive improvisation with a robotic marimba player, Auton. Robots 31(2–3), 133–153 (2011)
- 71.84 G. Deàk, I. Fasel, J. Movellan: The emergence of shared attention: Using robots to test developmental theories, Proc. 1st Int. Workshop Epigenet. Robotics, Lund (2001) pp. 95–104
- 71.85 K. Dautenhahn: Roles and functions of robots in human society: Implications from research in autism therapy, *Robotica* 21(4), 443–452 (2003)
- 71.86 H. Kozima, C. Nakagawa, Y. Yasuda: Wowing together: What facilitates social interactions in children with autistic spectrum disorders, Proc. 6th Int. Workshop Epigenet. Robotics Model. Cogn. Dev. Robotics Syst., Paris (2006) p. 177
- 71.87 B. Scasselati: How social robots will help us to diagnose, treat, and understand autism, Proc. 12th Int. Symp. Robotics Res., San Francisco, ed. by S. Thrun, R.A. Brooks, H. Durrant–Whyte (Springer, Berlin, Heidelberg 2005) pp. 552–563
- 71.88 D.J. Feil–Seifer, M.J. Mataric: B3IA: An architecture for autonomous robot–assisted behavior intervention for children with autism spectrum disorders, Proc. 17th IEEE Int. Workshop Robot Hum. Interact. Commun., Munich (2008) pp. 328–333
- 71.89 H. Kozima, C. Nakagawa, Y. Yasuda: Interactive robots for communication–care: A case–study in autism therapy, Proc. 14th IEEE Int. Workshop Robot Hum. Interact. Commun., Nashville (2005) pp. 341–346

- 71.90 H. Kozima, Y. Yasuda, C. Nakagawa: Social interaction facilitated by a minimally-designed robot: Findings from longitudinal therapeutic practices for autistic children, Proc. 16th IEEE Int. Symp. Robot Hum. Interact. Commun., Jeju Island (2007) pp. 599–604
- 71.91 H.A. Simon: *The Sciences of the Artificial* (MIT Press, Cambridge 1969)
- 71.92 B. Adams, C.L. Breazeal, R.A. Brooks, B. Scassellati: Humanoid robots: A new kind of tool, IEEE Intell. Syst. Appl. **15**(4), 25–31 (2000)
- 71.93 L.W. Barsalou, C. Breazeal, L.B. Smith: Cognition as coordinated non-cognition, Cogn. Process. **8**(2), 79–91 (2007)
- 71.94 A.L. Thomaz, M. Berlin, C. Breazeal: An embodied computational model of social referencing, Proc. 14th IEEE Int. Workshop Robot Hum. Interact. Commun., Nashville (2005) pp. 591–598
- 71.95 G. Hoffman, C. Breazeal: Robotic partners? Bodies and minds: An embodied approach to fluid human-robot collaboration, Proc. 5th Int. Workshop Cogn. Robotics, Boston (2006) pp. 95–102
- 71.96 G. Hoffman: Effects of anticipatory action on human-robot teamwork efficiency, fluency, and perception of team, Proc. 2nd ACM/IEEE Int. Conf. HRI, Washington D.C. (2007) pp. 1–8
- 71.97 Y. Demiris, A. Meltzoff: The robot in the crib: A developmental analysis of imitation skills in infants and robots, Infant Child Dev. **17**(1), 43–53 (2008)
- 71.98 C. Nehaniv, K. Dautenhahn (Eds.): *Imitation and Social Learning in Robots, Humans and Animals: Behavioural, Social and Communicative Dimensions* (Cambridge Univ. Press, Cambridge 2009)
- 71.99 B. Scassellati: Imitation and mechanisms of joint attention: A developmental structure for building social skills on a humanoid robot, Lect. Notes Comput. Sci. **1562**, 176–195 (1999)
- 71.100 Y. Nagai, K. Hosoda, A. Morita, M. Asada: A constructive model for the development of joint attention, Connect. Sci. **15**(4), 211–229 (2003)
- 71.101 F. Kaplan, V. Hafner: The challenges of joint attention, Proc. 4th Int. Workshop Epigenet. Robotics, Lund (2004) pp. 67–74
- 71.102 C. Crick, M. Munz, B. Scassellati: Synchronization in social tasks: Robotic drumming, Proc. 15th IEEE Int. Workshop Robot Hum. Interact. Commun., Hatfield (2006) pp. 97–102
- 71.103 P. Bakker, Y. Kuniyoshi: Robot see, robot do: An overview of robot imitation, AISB-96 Workshop Learn. Robots Animals, Brighton (1996) pp. 3–11
- 71.104 R.S. Jackendoff: On beyond zebra: The relation of linguistic and visual information, Cognition **26**, 89–114 (1987)
- 71.105 B. Landau, R.S. Jackendoff: *What and where* in spatial language and spatial cognition, Behav. Brain Sci. **16**, 217–265 (1993)
- 71.106 L. Talmy: The fundamental system of spatial schemas in language. In: *From Perception to Meaning: Image Schemas in Cognitive Linguistics*, ed. by B. Hamp (Mouton de Gruyter, Berlin 2005)
- 71.107 T.P. Regier: The Acquisition of Lexical Semantics for Spatial Terms: A Connectionist Model of Perceptual Categorization, Ph.D. Thesis (University of California at Berkeley, Berkeley 1992)
- 71.108 J.D. Kelleher, F.J. Costello: Applying computational models of spatial prepositions to visually situated dialog, Comput. Linguist. **35**(2), 271–306 (2008)
- 71.109 T.P. Regier, L.A. Carlson: Grounding spatial language in perception: An empirical and computational investigation, J. Exp. Psychol. **130**(2), 273–298 (2001)
- 71.110 G. Bugmann, E. Klein, S. Lauria, T. Kyriacou: Corpus-based robotics: A route instruction example, Proc. 8th Conf. Intell. Auton. Syst. (IAS-8), Amsterdam (2004) pp. 96–103
- 71.111 M. Levit, D. Roy: Interpretation of spatial language in a map navigation task, IEEE Trans. Syst. Man Cybern. B **37**(3), 667–679 (2007)
- 71.112 M. MacMahon, B. Stankiewicz, B. Kuipers: Walk the talk: Connecting language, knowledge, and action in route instructions, Proc. Natl. Conf. Artif. Intell., Boston (2006) pp. 1475–1482
- 71.113 H. Kress-Gazit, G.E. Fainekos: Translating structured English to robot controllers, Adv. Robotics **22**, 1343–1359 (2008)
- 71.114 C. Matuszek, D. Fox, K. Koscher: Following directions using statistical machine translation, Proc. 5th ACM/IEEE Int. Conf. HRI, Nara (2010) pp. 251–258
- 71.115 A. Vogel, D. Jurafsky: Learning to follow navigational directions, Proc. 48th Annu. Meet. Assoc. Comput. Linguist., Uppsala (2010) pp. 806–814
- 71.116 S. Harnad: The symbol grounding problem, Physica D **43**, 335–346 (1990)
- 71.117 T. Winograd: Procedures as a Representation for Data in a Computer Program for Understanding Natural Language, MIT Tech. Rep. TMAC-TR-84 (MIT, Cambridge 1971)
- 71.118 K.Y. Hsiao, N. Mavridis, D. Roy: Coupling perception and simulation: Steps towards conversational robotics, Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS), Las Vegas (2003) pp. 928–933
- 71.119 D. Roy, K.Y. Hsiao, N. Mavridis: Conversational Robots: Building blocks for grounding word meanings, Proc. HLT-NAACL 2003 Workshop Learn. Word Mean. Non-Linguist. Data, Stroudsburg (2003) pp. 70–77
- 71.120 D. Roy: Semiotic schemas: A framework for grounding language in action and perception, Artif. Intell. **167**(1–2), 170–205 (2005)
- 71.121 J. Dzifcak, M. Scheutz, C. Baral, P. Schermerhorn: What to do and how to do it: Translating natural language directives into temporal and dynamic logic representation for goal management and action execution, IEEE Int. Conf. Robotics Autom. (ICRA), Kobe (2009) pp. 4163–4168
- 71.122 Y. Sugita, J. Tani: Learning semantic combinatoriality from the interaction between linguistic and behavioral processes, Adapt. Behav. – Animals Animat. Softw. Agents Robots Adapt. Syst. **13**(1), 33–52 (2005)
- 71.123 J. Modayil, B. Kuipers: Autonomous development of a grounded object ontology by a learning

- robot, Proc. 22nd AAAI Conf. Artif. Intell., Vancouver (2007) pp. 1095–1101
- 71.124 D. Marocco, A. Cangelosi, K. Fischer, T. Belpaeme: Grounding action words in the sensorimotor interaction with the world: Experiments with a simulated iCub humanoid robot, *Front. Neurobotics* **4**, 1–15 (2010)
- 71.125 R. Ge, R.J. Mooney: A statistical semantic parser that integrates syntax and semantics, Proc. 9th Conf. Comput. Nat. Lang. Learn., Ann Arbor (2005) pp. 9–16
- 71.126 N. Shimizu, A. Haas: Learning to follow navigational route instructions, Proc. 21st Int. Jt. Conf. Artif. Intell., Pasadena (2009) pp. 1488–1493
- 71.127 S.R.K. Branavan, H. Chen, L.S. Zettlemoyer, R. Barzilay: Reinforcement learning for mapping instructions to actions, Proc. 47th Jt. Conf. Annu. Meet. Assoc. Comput. Linguist. 4th Int. Jt. Conf. Nat. Lang. Process. (AFNLP), Singapore (2009) pp. 82–90
- 71.128 S.R.K. Branavan, D. Silver, R. Barzilay: Learning to win by reading manuals in a Monte-Carlo framework, Proc. 49th Annu. Meet. Assoc. Comput. Linguist. Hum. Lang. Technol., Portland (2011)
- 71.129 T. Kollar, S. Tellex, D. Roy, N. Roy: Toward understanding natural language directions, Proc. 5th ACM/IEEE Int. Conf. HRI, Osaka (2010) pp. 259–266
- 71.130 D. Bailey: When Push Comes to Shove: A Computational Model of the Role of Motor Control in the Acquisition of Action Verbs, Ph.D. Thesis (Univ. of California, Berkeley 1997)
- 71.131 T. Kollar, S. Tellex, D. Roy, N. Roy: Grounding verbs of motion in natural language commands to robots, Proc. Int. Symp. Exp. Robotics, New Delhi (2010) pp. 31–47
- 71.132 S. Tellex, T. Kollar, S. Dickerson, M.R. Walter, A.G. Banerjee, S. Teller, N. Roy: Understanding natural language commands for robotic navigation and mobile manipulation, Proc. Natl. Conf. Artif. Intell., San Francisco (2011)
- 71.133 J.L. Burke, R.R. Murphy, M.D. Coovert, D.L. Riddle: Moonlight in Miami: Field study of human-robot interaction in the context of an urban search and rescue disaster response training exercise, *Hum.-Comput. Interact.* **19**(1/2), 85–116 (2004)
- 71.134 K. Stubbs, P.J. Hinds, D. Wettergreen: Autonomy and common ground in human-robot interaction: A field study, *IEEE Intell. Syst.* **22**(2), 42–50 (2007)
- 71.135 S. Kiesler: Fostering common ground in human-robot interaction, Proc. 14th IEEE Int. Workshop Robot Hum. Interact. Commun., Nashville (2005) pp. 729–734
- 71.136 T. Fong, C. Thorpe, C. Baur: Collaboration, dialogue, human-robot interaction, *Robotics Res.* **6**, 255–266 (2003)
- 71.137 M.E. Foster, T. By, M. Rickert, A. Knoll: Human-robot dialogue for joint construction tasks, Proc. 8th Int. Conf. Multimodal Interfaces, Banff (2006) pp. 68–71
- 71.138 S. Li, B. Wrede, G. Sagerer: A computational model of multi-modal grounding for human robot interaction, Proc. 7th SIGdial Workshop Discourse Dialogue, Sydney (2009) pp. 153–160
- 71.139 P.R. Cohen, H.J. Levesque: Teamwork, *Nous* **25**(4), 487–512 (1991)
- 71.140 J. Peltason, B. Wrede: Pamini: A framework for assembling mixed-initiative human-robot interaction from generic interaction patterns, Proc. 11th SIGdial Annu. Meet. Special Interest Group Discourse Dialogue, Tokyo (2010) pp. 229–232
- 71.141 J. Peltason, B. Wrede: The curious robot as a case-study for comparing dialog systems, *AI Magazine* **32**(4), 85–99 (2011)
- 71.142 M.F. Schober: Spatial perspective-taking in conversation, *Cognition* **47**(1), 1–24 (1993)
- 71.143 J.E. Hanna, M.K. Tanenhaus, J.C. Trueswell: The effects of common ground and perspective on domains of referential interpretation, *J. Mem. Lang.* **49**(1), 43–61 (2003)
- 71.144 J.G. Trafton, N.L. Cassimatis, M.D. Bugajska, D.P. Brock, F.E. Mintz, A.C. Schultz: Enabling effective human-robot interaction using perspective-taking in robots, *IEEE Trans. Syst. Man Cybern. A* **35**(4), 460–470 (2005)
- 71.145 M. Berlin, J. Gray, A.L. Thomaz, C. Breazeal: Perspective taking: An organizing principle for learning in human-robot interaction, Proc. 21st Natl. Conf. Artif. Intell., Boston (2006) p. 1444
- 71.146 R. Moratz, K. Fischer, T. Tenbrink: Cognitive modeling of spatial reference for human-robot interaction, *Int. J. Artif. Intell. Tools* **10**(04), 589–611 (2001)
- 71.147 R. Ros, S. Lemaignan, E.A. Sisbot, R. Alami, J. Steinwender, K. Hamann, F. Warneken: Which one? Grounding the referent based on efficient human-robot interaction, Proc. 19th IEEE Int. Symp. Robot Hum. Interact. Commun., Viareggio (2010) pp. 570–575
- 71.148 A. Holroyd, C. Rich, C.L. Sidner, B. Ponsler: Generating connection events for human-robot collaboration, Proc. 20th IEEE Int. Symp. Robot Hum. Interact. Commun., Atlanta (2011) pp. 241–246
- 71.149 G. Butterworth, L. Grover: Joint visual attention, manual pointing, and preverbal communication in human infancy. In: *Attention and Performance, Vol. 13: Motor Representation and Control*, ed. by M. Jeannerod (Lawrence Erlbaum Assoc., Mahwah 1990) pp. 605–624
- 71.150 C. Rich, P. Ponsler, A. Holroyd, C.L. Sidner: Recognizing engagement in human-robot interaction, Proc. 5th ACM/IEEE Int. Conf. HRI, Osaka (2010) pp. 375–382
- 71.151 V. Gallese, A. Goldman: Mirror neurons and the simulation theory of mind-reading, *Trends Cogn. Sci.* **2**(12), 493–501 (1998)
- 71.152 E. Bicho, W. Erlhagen, L. Louro, E. Costa e Silva: Neuro-cognitive mechanisms of decision making in joint action: A human-robot interaction study, *Hum. Mov. Sci.* **30**(5), 846–868 (2011)
- 71.153 J. Gray, C. Breazeal, M. Berlin, A. Brooks, J. Lieberman: Action parsing and goal inference using self as simulator, Proc. 14th IEEE Int. Workshop

- Robot Hum. Interact. Commun., Nashville (2005) pp. 202–209
- 71.154 M.N. Nicolescu, M.J. Mataric: Linking perception and action in a control architecture for human–robot domains, Proc. 36th Annu. Hawaii Int. Conf. Syst. Sci., Big Island (2003) pp. 10–20
- 71.155 C. Breazeal, G. Hoffman, A. Lockerd: Teaching and working with robots as a collaboration, Proc. 3rd Int. Jt. Conf. Auton. Agents Multiagent Syst., New York, Vol. 3 (2004) pp. 1030–1037
- 71.156 R. Alami, A. Clodic, V. Montreuil, E.A. Sisbot, R. Chatila: Task planning for human–robot interaction, Proc. 2005 Jt. Conf. Smart Obj. Ambient Intell. Innov. Context-Aware Serv. Usages Technol., Grenoble (2005) pp. 81–85
- 71.157 R. Alami, A. Clodic, V. Montreuil, E.A. Sisbot, R. Chatila: Toward human–aware robot task planning, AAAI Spring Symp.: To Boldly Go where No Human–Robot Team Has Gone Before, Palo Alto (2006) pp. 39–46
- 71.158 R. Bellman: *Dynamic Programming* (Princeton Univ. Press, Princeton 1957)
- 71.159 N. Roy, J. Pineau, S. Thrun: Spoken dialog management for robots, Proc. Assoc. Comput. Linguist., Hong Kong (2000) pp. 93–100
- 71.160 J. Hoey, P. Poupart, C. Boutilier, A. Mihailidis: POMDP models for assistive technology, Proc. AAAI Fall Symp. Caring Mach., AI in Eldercare (2005)
- 71.161 J. Williams, S. Young: Scaling up POMDPs for dialogue management: The *summary POMDP* method, Proc. IEEE Autom. Speech Recognit. Underst. Workshop, Cancun (2005)
- 71.162 D. Litman, S. Singh, M. Kearns, M. Walker: NJFun: A reinforcement learning spoken dialogue system, Proc. ANLP/NAACL 2000 Workshop Conversat. Syst., Seattle (2000) pp. 17–20
- 71.163 F. Broz, I. Nourbakhsh, R. Simmons: Planning for human–robot interaction using time–state aggregated POMDPs, Proc. 23rd Conf. Artif. Intell., Chicago (2008) pp. 1339–1344
- 71.164 F. Doshi, N. Roy: The permutable POMDP: Fast solutions to POMDPs for preference elicitation, Proc. 7th Int. Conf. Auton. Agents Multiagent Syst., Estoril (2008) pp. 493–500
- 71.165 R. Wilcox, S. Nikolaidis, J. Shah: Optimization of temporal dynamics for adaptive human–robot interaction in assembly manufacturing, Proc. Robotics Sci. Syst., Sydney (2012) p. 441
- 71.166 T. Prommer, H. Holzapfel, A. Waibel: Rapid simulation–driven reinforcement learning of multimodal dialog strategies in human–robot interaction, 9th Int. Conf. Spoken Lang. Process., Pittsburgh (2006)
- 71.167 J.M. Porta, N. Vlassis, M. Spaan, P. Poupart: Point-based value iteration for continuous POMDP, J. Mach. Learn. Res. **7**, 2329–2367 (2006)
- 71.168 F. Doshi, N. Roy: Efficient model learning for dialog management, Proc. 2nd ACM/IEEE Int. Conf. HRI, Arlington (2007) pp. 65–72
- 71.169 F. Doshi, N. Roy: Spoken language interaction with model uncertainty: An adaptive human–robot interaction system, Connect. Sci. **20**(4), 299–319 (2008)
- 71.170 M. Cakmak, A.L. Thomaz: Designing robot learners that ask good questions, Proc. 7th Annu. ACM/IEEE Int. Conf. HRI, Boston (2012) pp. 17–24
- 71.171 M.A. Goodrich, D.R. Olsen: Seven principles of efficient human robot interaction, IEEE Int. Conf. Syst. Man Cybern., Washington D.C. (2003) pp. 3942–3948
- 71.172 J.W. Crandall, M.A. Goodrich, D.R. Olsen, C.W. Nielsen: Validating human–robot interaction schemes in multitasking environments, IEEE Trans. Syst. Man Cybern. A **35**(4), 438–449 (2005)
- 71.173 R. Sun (Ed.): *Cognition and Multiagent Interaction: From Cognitive Modeling to Social Simulation* (Cambridge Univ. Press, Cambridge 2005)