

Collaboration, Dialogue, and Human-Robot Interaction

Terrence Fong^{1,2}, Charles Thorpe¹, and Charles Baur²

¹ The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA

² Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

Abstract. Teleoperation can be improved if humans and robots work as partners, exchanging information and assisting one another to achieve common goals. In this paper, we discuss the importance of collaboration and dialogue in human-robot systems. We then present collaborative control, a system model in which human and robot collaborate, and describe its use in vehicle teleoperation.

1 Introduction

1.1 Robot as Partner

A robot is commonly viewed as a *tool*: a device that performs tasks on command. As such, a robot has limited freedom and will perform poorly whenever it is ill-suited for the task at hand. Moreover, if a robot has a problem, it has no way to ask for assistance. Yet, frequently, the only thing a robot needs to work better is some advice (even a small amount) from a human.

Consider the situation in which a mobile robot is driving outdoors when it encounters tall grass in its path. Depending on its sensors, the robot's perception system may have difficulty deciding if the grass poses a danger. Thus, the robot may be unable to proceed or may take a long, resource consuming detour. If, however, the robot is able to discuss the situation with a human, a better solution can be found. For example, if the robot asks "Is there an obstacle ahead?" and shows a camera image, the human can help decide that it is safe to drive forward.

Generally speaking, robots are more adept at making some decisions by themselves than others. For example, structured planning (for which algorithms or well-defined solutions exist) has proven to be quite amenable to automation. Unstructured decision making, however, remains the domain of humans, especially whenever common sense is required[6]. In particular, robots continue to perform poorly at high-level perceptual functions, including object recognition and situation assessment[14].

In order for robots to perform better, therefore, they need to be able to take advantage of human skills (perception, cognition, etc.) and to benefit from human advice and expertise. To do this, robots need to function not as passive tools, but rather as active partners. They need to have more freedom of action, to be able to drive the interaction with humans, instead of merely waiting for (or blindly executing) human commands.

1.2 Collaborative Control

To address this need, we have developed a new system model for teleoperation called collaborative control[7]. In this model, a human and a robot work as partners (if not peers), *collaborating* to perform tasks and to achieve common goals. Instead of a supervisor dictating to a subordinate, the human and the robot engage in *dialogue* to exchange ideas, to ask questions, and to resolve differences.

With collaborative control, the human functions as a resource for the robot, providing information and processing just like other system modules. In particular, the robot can ask questions to the human as it works, to obtain assistance during task execution. With this approach, the robot has more freedom in execution and is more likely to find good solutions when it has problems.

Collaborative control is a radical departure from traditional teleoperation, from the conventional "robot as tool" system model. Collaborative control encourages human-robot interaction to be more natural, more balanced, and more direct. Collaborative control also allows robots to benefit from human assistance during perception and cognition, and not just planning and command generation.

1.3 Key Issues

To build a collaborative control system, we have found that there are four key issues that must be addressed. First, the robot must have *self-awareness*. This does not imply that the robot needs to be fully sentient, merely that it be capable of detecting limitations (in what it can do and what the human can do), determining if it should ask for help, and recognizing when it has to solve problems on its own.

Second, the robot must be *self-reliant*. Since the robot cannot rely on the human to always be available or to provide accurate information, it must be able to maintain its own safety. Specifically, the robot should be capable of avoiding hazards, monitoring its health, and taking action to "safe" itself when necessary.

Third, the system must support *dialogue*. That is, the robot and the human need to be able to communicate effectively with each other. Each participant must be able to convey information, to ask questions and to judge the quality of responses received. To an extent, traditional teleoperation has dialogue (i.e., the feedback loop), but the conversation is limited. With collaborative control, dialogue is two-way and requires a richer vocabulary.

Finally, the system must be *adaptive*. By design, collaborative control provides a framework for integrating users with varied skills, knowledge, and experience. As a consequence, however, the robot has to be able to adapt to different operators and to adjust its behavior as needed, e.g., asking questions based on the operator's capacity to answer.

2 Dialogue

2.1 Communication and Conversation

Dialogue is the process of communication between two or more parties. Dialogue is a joint process: it requires sharing of information (data, symbols, context) and of control. Depending on the situation (task, environment, etc.), the form or style of dialogue will vary. However, studies of human conversation have revealed that many properties of dialogue, such as initiative taking and error recovery, are always present[13].

When humans and machines (computers, robots, etc.) communicate, dialogue is usually mediated by an interface. Some interfaces (e.g., computer command languages) offer great power and flexibility, though at an associated high learning cost. Other interfaces, such as menus, are easier for novices because they make few assumptions about what the user knows. Regardless the form, however, a good interface provides structure that facilitates human-machine dialogue and information exchange.

2.2 Dialogue Management

Unless the interaction is simple (e.g., fixed grammar), human-computer systems require dialogue management. The basic function of dialogue management is to translate user requests into a language the computer understands and the system's output into a language that the user understands[10]. In addition, dialogue management must be capable of performing a variety of tasks including disambiguation, error handling, and role switching[1].

Role switching occurs because at any stage in a dialogue, one participant has the initiative (control) of the conversation. In a sense, initiative is a function of the roles of the participants. Dialogue systems may allow the user or the computer to take the initiative, or may allow both to switch roles as required. By far, the hardest dialogues to model are those in which the initiative can be taken at any point in the dialogue[5].

2.3 User Model

Dialogue cannot make sense unless the user and the system have a reasonable understanding of each other. Given a user model, dialogue adaptation can be performed by referring to a user's expertise, knowledge, and preferences [10]. For example, the way in which information is collected (filtering, classification, etc.) and presented (text, graphics, speech) can be adapted to the user.

The stereotype approach is the most popular user modeling method. With the stereotype approach, a designer defines appropriate subgroups of the user population (the stereotypes), identifies user behaviors that enable a system to categorize users into a subgroup, and represents the set of features (a "user profile") that characterizes each stereotype[19].

3 System Design

3.1 Architecture

We have implemented collaborative control as a distributed set of modules, connected by a message-based architecture (Fig. 1). The architecture includes a safeguarded teleoperation controller that supports varying degrees of cooperation between the operator and robot [9]. Our primary user interface is the PdaDriver, which runs on WindowsCE-based PocketPC's[8]. PdaDriver provides a variety of command modes including rate, position and image/map waypoint control. We are currently using collaborative control to operate Pioneer mobile robots, which are equipped with a variety of sensors including ultrasonic sonar, color CCD camera, and differential GPS[8].

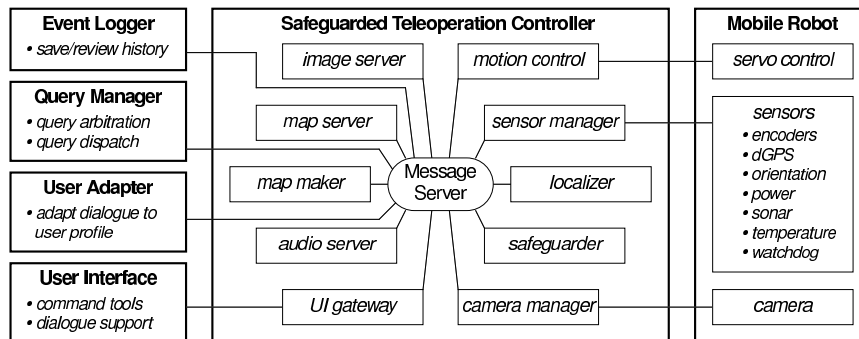


Fig. 1. Collaborative control architecture

3.2 User Model

Our collaborative control system manages dialogue with attribute-based, stereotype user profiles. We currently use three *user attributes* (accuracy, expertise, and query interval) to describe users, chosen because they are well-suited for vehicle teleoperation in unknown environments and because they provide a sufficiently rich basis for experimentation.

Accuracy estimates how accurately a user answers questions. If a user is highly accurate, then the robot can place greater confidence in responses to safety-critical questions. *Expertise* estimates the task skill (or domain knowledge) that the user possesses. This is valuable for adapting dialogue and autonomy. *Query interval* indicates how often the user can answer questions. Query interval may indicate: availability (amount of time a user can dedicate to questions), efficiency (how quickly he responds), and preference (how often he prefers to be interrupted by questions).

To support our initial experiments and evaluation of collaborative control, we defined three user stereotypes. A *novice* has no training or teleoperation experience and is not expected to answer questions well. A *scientist* is also untrained and inexperienced, but is able to answer domain specific questions. An *expert* knows everything: he is experienced, understands how the system is designed, and can answer questions quickly and accurately.

For each user stereotype, we assigned the attribute values shown in Table 1. In our current system, all user attributes are constant (i.e., no adaptation occurs during use) and reflect *a priori* assumptions about each user stereotype (e.g., novices take longer to respond than experts).

Table 1. Stereotype attributes

Stereotype	Accuracy	Expertise	Query interval
Novice	30	30	60
Scientist	50	100	30
Expert	100	100	0

We use each user profile in three ways. First, we modify interaction by configuring the user interface for each type of user (e.g., selecting which control modes and displays are available). Second, we adapt dialogue by filtering messages: only messages that are appropriate for the user are selected. Finally, we modify how the robot acts by dynamically varying its autonomy.

3.3 Query Manager

Under collaborative control, multiple robot modules may ask questions at the same time. Thus, a collaborative control system needs query arbitration: a mechanism for choosing which questions to ask based on immediate (local) needs and overall (global) strategy. In our system, the QueryManager performs this task with an attribute-based filtering scheme [7].

Whenever a robot has a question to ask the human, it sends a message to the QueryManager. A message is defined by user attributes, query attributes (type, priority level, expiration time), and question-specific data (image, text, etc.) Our collaborative control system currently supports two query types: *y/n* (user must answer *y* or *n*) and *value* (user must provide a numeric value).

The QueryManager stores incoming messages into a cache. When the human indicates that he is available to answer a question, the QueryManager selects a message by filtering the cache. Because the cache is priority-sorted, urgent questions have preference. Expired questions are discarded undelivered (i.e., the user is never asked a question which is no longer valid).

3.4 Dialogue

In our system, dialogue arises from an exchange of messages between human and robot. Effective dialogue does not require a full language, merely one that is pertinent to the task at hand and that efficiently conveys information. Thus, we do not use natural language and we limit the vocabulary and grammar to vehicle mobility issues (navigation, obstacle avoidance, etc).

Table 2. Robot-to-user queries

Query	Type	Accuracy	Expertise
Can I drive through (<i>image</i>)?	y/n	50	50
Is this a rock (<i>image</i>)? If you answer <i>y</i> , I will stay here.	y/n	0	50
The environment is very cluttered (<i>map</i>). What is the fastest I should translate?	value	0-100	50
My motors are stalled. Can you come over and help?	y/n	0	0
Motion control is currently turned off. Shall I enable it?	y/n	50	0
Safeguards are currently turned off. Shall I enable it?	y/n	50	0
Stopped due to collision danger. Disable safeguards?	y/n	100	0
Stopped due to high temperature. What should the safety level be?	value	0-100	0
Stopped due to low power. What should the safety level be?	value	0-100	0
Stopped due to rollover danger. Can you come over and help?	y/n	0	0

Table 2 lists the queries that a robot can ask. Two queries have variable accuracy levels because the importance of these questions can change with time or situation. Low accuracy values means that the robot is willing to accept any response. High accuracy values, however, indicate that the setting is critical to the robot’s continued health.

Three of the queries in Table 2 have non-zero expertise values. To answer these queries, the human must have a certain level of expertise. For our experiments, we did not distinguish between different types of experts (e.g., skilled pilot vs. geologist). In practice, however, we would use additional attributes to target queries to specific task or domain expertise.

4 Results

To gain insight into collaborative control, we examined three vehicle teleoperation scenarios. In each case, we found that collaborative control enabled the robot to perform better. In particular, we observed that when the robot is operating poorly, or when it does not know what to do, a simple human answer can lead to a significant improvement.

4.1 “A to B”

Perhaps the most basic task in vehicle teleoperation is “A to B”: controlling the robot so that it moves from point A to B. As simple as this may seem, successful execution is critical to many applications. In reconnaissance, for example, performance is often governed by how well the robot moves from point to point. Thus, we need to make “A to B” as successful as possible[9].

The most effective command mechanism currently used in vehicle teleoperation is waypoint driving: the operator specifies a series of points that must be passed *en route* to a target position. Waypoint driving has many advantages over direct (manual) control, e.g., it can tolerate significant delay. However, waypoint driving is not without problems. For example, if the robot has trouble deciding if an obstacle is in the way, it may operate slowly or be forced to stop. Collaborative control remedies this by allowing the robot to confer with the human before proceeding.

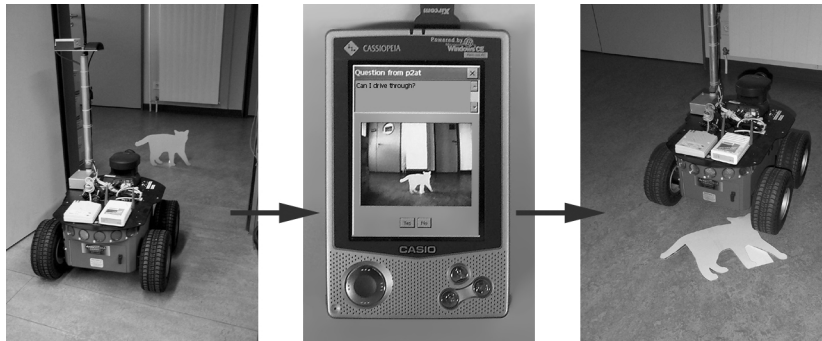


Fig. 2. Query to the human: “Can I drive through?”

Figure 2 shows an example of this interaction occurring in an experiment we performed in an office environment. During this test, we placed a cardboard cat in the path of the robot. The cardboard cat was detected as an obstacle by the robot’s sonar, thus forcing it to stop. At this point, the robot sent a camera image to the human and asked if was safe to drive forward. When the human answered “yes” (based on his interpretation of the image), the robot was able to proceed without delay.

4.2 Collaborative Exploration

Although considerable research has focused on using human and robotic systems for planetary exploration, scant attention has been given to developing joint human-robotic systems. Yet, such systems offer significant potential to improve planetary missions by increasing productivity while reducing cost, particularly for surface operations such as material transport, survey and sampling[8].

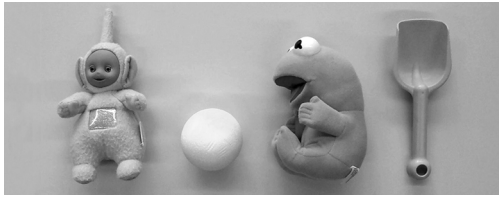


Fig. 3. Some “rocks”

To examine collaborative human-robot exploration, we developed a perception module, *RockFinder*, that autonomously locates “rocks”. Our intent was to study the assistance an exploration rover might provide. Thus, rather than examine morphology, *RockFinder* simply searches for objects (Fig. 3) that match a color signature. Whenever *RockFinder* detects a potential “rock”, it converses with the human to decide what to do.



Fig. 4. Collaborative exploration

Figure 4 shows an example of this interaction. With human and robot collaborating, exploration becomes very efficient: the human can say, “Tell me when you find something interesting” and the robot can do the search even if its autonomy (e.g., *RockFinder*) is limited.

4.3 Multi-Robot Teleoperation

The American military is currently developing mobile robots to support future combat systems. These robots will be used to perform a variety of reconnaissance, surveillance and target acquisition (RSTA) tasks. Because these tasks have traditionally required significant human resources (manpower, time, etc.), one of the primary areas of interest is determining how a small number of operators can control a larger number of robots. We believe that collaborative control provides an effective solution for this problem.

For example, consider the situation in which a single operator needs to control multiple robots, each of which is capable of limited autonomous RSTA functions (e.g., “move to point Tango and collect imagery”). As they traverse unknown, unexplored or changing terrain each robot will likely have questions such as: “Is it safe to continue driving at this power level?”, “Is this obstacle dangerous?”, and “Is it safe to drive over this terrain?”.

Since the human can only focus his attention on one robot at a time, we can use dialogue to unify and coordinate the multiple requests. Specifically, we arbitrate among the questions so that the human is always presented with the one that is most urgent (in terms of safety, timeliness, etc). This helps to reduce the level of attention and control the operator must dedicate to each robot. Consequently, the human can more effectively perform simultaneous, parallel control. In addition, because each robot is aware that the human may not be able to respond, it can still try to resolve problem on its own.

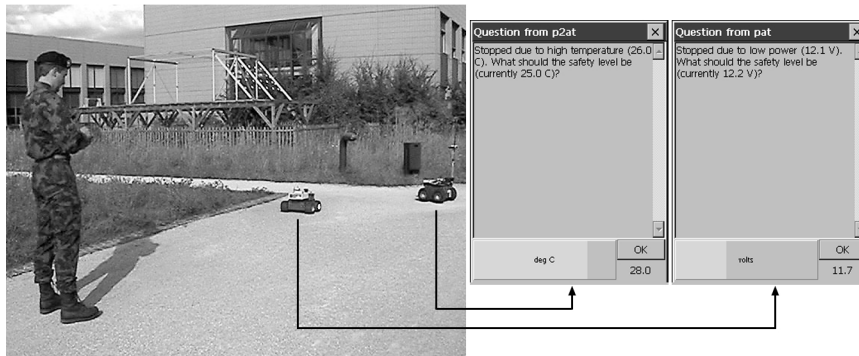


Fig. 5. Multi-robot teleoperation

Figure 5 shows an example of this behavior occurring in multi-robot teleoperation. In this experiment, an operator is using two robots for reconnaissance. Collaborative control allows the human to quickly switch his attention between the two, directing and answering questions from each as needed. In our testing, we found that this to be an effective way to interact with independently operating robots. In particular, coordination arises naturally from query arbitration.

5 Discussion

5.1 Benefits

By enabling humans and robots to work as partners, we have found that teleoperation is easier to use and more productive. Collaboration enables the human and robot to complement each other, as well as allowing the robot to proceed when the human is unavailable. Dialogue lets us build systems that are user adaptive and that encourage teamwork.

We have observed that dialogue makes human-robot interaction adaptable. Since the robot is “aware” of to whom is speaking, it can dynamically decide whether or not asking a question to the human is worthwhile. We have also found that there are situations for which dialogue enables the robot to perform significantly better than without human intervention. Moreover, this is true regardless of whether the human is a novice or an expert. In other words, even a novice can help compensate for inadequate sensing and autonomy.

Lastly, it seems evident (though we have not yet confirmed this) that specific combinations of collaboration and dialogue are appropriate for multiple situations. In other words, it may be possible that the interaction used for a specific user could be appropriate for all users when the system is constrained by factors such as bandwidth and delay.

5.2 Limitations

Although collaboration and dialogue provide significant benefits, we recognize that there are limitations. First, identifying which parameters are well-suited to a given task and assigning appropriate values for each query is difficult. If there are many tasks to perform or if task execution creates many questions, then dialogue may add considerable complexity to system design.

Second, if human-robot interaction is adaptive, then the flow of control and information through the system will vary with time and situation. This may make debugging, validation, and verification harder because it becomes more difficult (though not impossible) to precisely identify an error condition or to duplicate a failure situation.

Finally, working in collaboration requires that each partner trust and understand the other. To do this, each collaborator needs to have an accurate model of what the other is capable of and of how he will carry out a given assignment. If the model is inaccurate, or if the partner cannot be expected to perform correctly (e.g., a novice answering a safety critical question), then care must be taken. For the robot, this means that it may need to weigh human responses instead of accepting them at face value. For the human, this means the robot may not always behave as expected.

6 Related Work

6.1 Human-Robot Interaction

Human-robot interaction (HRI) can be defined as *the study of humans, robots, and the ways they influence each other*. Sheridan notes that one of the challenges for HRI is to provide humans and robots with models of each other[18]. In particular, he claims that the ideal would be analogous to two people who know each other well and who can pick up subtle cues from one another (e.g., musicians playing a duet).

In recent years, much effort has focused on developing robots that work directly with humans, as assistants or teammates [3,16,12]. The problem with most of these human-robot systems is that they are limited by their user-centric designs. Since effective collaboration requires an exchange and sharing of information between parties, collaborative control considers both user and robot needs.

6.2 Human-Robot Control Architectures

Numerous robot control architectures have addressed the problem of mixing humans with robots. The two most common approaches are prioritized control[2] and command fusion (arbitration)[17]. In such systems, the human is limited to providing command input to the system. Collaborative control, however, relaxes this restriction and allows the human to also contribute high-level planning or perception input to robot modules.

Adjustable autonomy and mixed initiative systems have recently received considerable research attention[4,15]. Although both approaches share some aspects of collaborative control, neither completely addresses the idea of peer interaction between humans and robots. The distinctive feature of collaborative control is that it uses human-robot dialogue as a mechanism for adaptation and a framework for coordination.

6.3 Human-Computer Collaboration

There are two major approaches to human-computer collaboration[19]. Human Emulation (HE) assumes that the way to get computers to collaborate with humans is to endow them with human-like abilities, to enable them to act like humans. Human Complementary (HC) assumes that computers and humans have fundamentally asymmetric abilities. Thus, the focus of HC is to develop techniques that make the computer a more intelligent partner.

Collaborative control is a hybrid approach that exhibits characteristics of both HE and HC. Collaborative control is HE-like because it emphasizes human-robot dialogue and because it includes techniques to model and adapt behavior to different users. Collaborative control is HC-like because it recognizes that humans and robots have different capabilities, skills, and needs. As a consequence, it provides different types of support to each.

Acknowledgements

We would like to thank the EPFL Institut de production et robotique for providing research facilities and infrastructure. This work was partially supported by a grant from SAIC and the DARPA ITO Mobile Autonomous Robot Software (MARS) program.

References

1. Abella, A., Gorin, A. (1999) Construct Algebra: analytical dialog management. In: Proc Annual Meeting of the ACL. Morgan-Kaufmann, San Francisco
2. Albus, J., et al. (1987). NASA/NBS Standard reference model for telerobot control system architecture (NASREM). Technical Note 1235, NIST.
3. Baltus, G., et al. (2000) Towards personal service robots for the elderly. In: Proc WIRED. Carnegie Mellon University, Pittsburgh
4. Bonasso, P. (1999) Issues in providing adjustable autonomy in the 3T architecture. In: Proc Spring Symposium on Agents with Adjustable Autonomy. AIAA, Menlo Park
5. Churcher, G., et al. (1997) Dialogue management systems: a survey and overview. Report 97.6. School of Computer Studies, University of Leeds
6. Clarke, R. (1994) Asimov's Laws of Robotics: implications for information technology. *IEEE Computer* **26(12)** and **27(1)**
7. Fong T. (2001) Collaborative control: a robot-centric model for vehicle teleoperation. CMU-RI-TR-01-34. Ph.D. thesis, Carnegie Mellon University, Pittsburgh
8. Fong, T., et al. (2001) A personal user interface for collaborative human-robot exploration. In: Proc iSAIRAS. Canadian Space Agency, Ottawa
9. Fong, T., et al. (2001) A safeguarded teleoperation controller. In: Proc Intl Conf on Advanced Robotics. IEEE, Piscataway
10. Goren-Bar, D. (2001) Designing model-based intelligent dialogue systems. In: Information Modeling in the New Millennium. Idea Group, London
11. Kiriki, T., et al. (1999) A 4-legged mobile robot control to observe a human behavior. In: Proc RO-MAN. IEEE, Piscataway
12. Laengle, T., et al. (1997) Cooperation in human-robot-teams. In: Proc Intl Conf on Informatics and Control. Russian Academy of Science, Moscow
13. Lansdale, M., Ormerod, T. (1994) Understanding interfaces: a handbook of human-computer dialogue. Academic Press, London
14. Milgram, P., et al. (1993) Applications of augmented reality for human-robot communication. In: Proc Intl Conf on Intelligent Robots and Systems. IEEE, Piscataway
15. Myers, K., Morley, D. (2001) Human directability of agents. In: Proc. Intl Conf on Knowledge Capture. ACM, New York
16. Nourbakhsh, I., et al. (1999) An affective mobile robot educator with a full-time job. *Artificial Intelligence* **114(1-2)**
17. Rosenblatt, J. (1997) DAMN: A Distributed Architecture for Mobile Navigation. CMU-RI-TR-97-01. Ph. D. thesis, Carnegie Mellon University, Pittsburgh.
18. Sheridan, T. (1997) Eight ultimate challenges of human-robot communication. In: Proc RO-MAN. IEEE, Piscataway
19. Terveen, L. (1994) An overview of human-computer collaboration. *Knowledge-Based Systems* **8(2-3)**