Building Strong Semi-Autonomous Systems

Shlomo Zilberstein
School of Computer Science
University of Massachusetts Amherst
shlomo@cs.umass.edu

Abstract

The vision of populating the world with autonomous systems that reduce human labor and improve safety is gradually becoming a reality. Autonomous systems have changed the way space exploration is conducted and are beginning to transform everyday life with a range of household products. In many areas, however, there are considerable barriers to the deployment of fully autonomous systems. We refer to systems that require some degree of human intervention in order to complete a task as *semi-autonomous systems*. We examine the broad rationale for semi-autonomy and define basic properties of such systems. Accounting for the human in the loop presents a considerable challenge for current planning techniques. We examine various design choices in the development of semi-autonomous systems and their implications on planning and execution. Finally, we discuss fruitful research directions for advancing the science of semi-autonomy.

Introduction

A fundamental goal of early AI projects has been to build intelligent systems, such as Shakey the Robot, that would operate in the world *autonomously* and perform some useful tasks (Nilsson 1984). Today, autonomous systems offer transformational impact on society as they help reduce human labor and improve productivity and safety. They have been deployed in a wide range of domains from household products such as the Roomba vacuum cleaners to space exploration vehicles such as Deep Space 1 (Muscettola et al. 1998; Pell et al. 1996). There is no standard definition of *autonomy* in AI, but generally a system is considered *autonomous* if it can construct and execute a plan to achieve its assigned goals, without human intervention, even when it encounters unexpected events (Doyle 2003; Frost 2010).

As the field of AI matured and produced numerous fielded systems and applications, it became apparent that there are still considerable barriers to the deployment of fully autonomous systems. These barriers range from technological and economical constraints to ethical and legal issues (Arkin 2009; Markoff 2012). A good example is *autonomous driving*, which has attracted growing attention since Google revealed its autonomous car. The car can complete an entire trip autonomously, but—at least for now—California's DMV requires it to include a steering wheel and brake pedals, and the presence of a "test driver" who is capable of taking over control at any time. Sergey Brin, the co-founder of Google, predicted in 2012 that self-driving cars will be available for all within five years (Thomson 2012). Nissan expects to market self-driving cars by 2020 (White 2013). But there are significant challenges before such cars can drive *without* human supervision. In his IAAI'12 invited talk, Sebastian Thrun, who led the early stages of the Google project, acknowledged that sensing and real-time inference technologies are not yet reliable enough to handle correctly certain situations, such as a plastic bag blowing in the wind in front of the car (that may be ignored) or a policeman signaling traffic to stop (should not be ignored). To be successful, autonomous cars operating in an environment populated with human drivers will have to be able to read, and eventually respond with, human social cues (Fong, Nourbakhsh, and Dautenhahn 2003).

When full autonomy is not feasible, it is still desirable to implement it even partially. But what does it mean for a system to be partially autonomous? What is the role of the person who controls it? What can the system guarantee without human assistance?

We refer to systems that can operate autonomously *under some conditions*, but cannot always complete an entire task on their own as *semi-autonomous systems* (SAS). There are many reasons for a SAS to require human intervention. For example, a Roomba vacuum cleaning robot could be trapped in a tight corner; or it may need to move to another floor that requires climbing stairs; or it may diagnose a sensor failure that requires a professional technician to repair. In other words, many autonomous systems currently under development are in fact semi-autonomous as they require human intervention under a range of conditions. Examples include semi-autonomous wheelchairs (Marsland, Nehmzow, and Duckett 2001), robotic tractors (Chaffins 2008), and General Motor's EN-V city vehicle (Woodyard 2010).

Semi-autonomous driving, in particular, is progressing rapidly. In some new car models, adaptive forms of cruise control can automatically maintain a desired speed as well as safe distance from the vehicles ahead. Some cars can perform complete maneuvers autonomously, such as parallel parking or lane following (Longjard et al. 2007), allowing drivers to take their hands off the steering wheel. Such maneuvers that can be performed autonomously are expected to grow rapidly in range and complexity, well before fully autonomous driving is realized.

The rapid growth of SAS development has far reaching implications on AI, particularly on the field of *automated planning*. Historically, planning has focused on efficient computational paradigms for creating plans that can be executed by autonomous agents, for example, path and motion planning for autonomous robots, plant operation planning in manufacturing, emergency evacuation planning and planning in games (Ghallab, Nau, and Traverso 2004).

Planning for semi-autonomous systems, however, presents significant new challenges and requires new computational and plan execution models, interaction mechanisms and algorithms. The planning process must account for the human in the loop, the *different skills* of the human operator and the semi-autonomous system, and the *decentralized* nature of their operation. The computed *shared plan* must include not only domain-level actions, but also *communication* between the human and the SAS to facilitate smooth *transfer of control*. There could be uncertainty about the underlying state of the SAS as well as the state of human operators and their ability to take over control. A central challenge is to develop mechanisms to allow human operators to *transition smoothly* from being unengaged (during fully autonomous operation) to being engaged (during manual operation) and vice versa. Related questions are how to *measure and monitor human engagement*, how long will the transition take, and how to respond in case it is failing. Finally, human operators may be *prone to error*, unintentionally taking an incorrect action or not taking over control when asked to do so. While some very relevant research has been conducted in AI, robotics, HRI, cognitive science and human factors, robust planning techniques for semi-autonomous systems are sorely lacking.

Background on Semi-Autonomous Systems

Although there is no universally accepted definition of semi-autonomous systems—and no general purpose algorithms exist for planning in this context—substantial relevant research has been conducted in the fields of AI (planning, multiagent systems), robotics (particular HRI), and human factors. This section provides a brief overview of related work.

The multiagent systems community has long been exploring various forms of *adjustable autonomy*, allowing autonomous agents or robots to get help from humans (Bradshaw et al. 2005; Côté, Bouzid, and Mouaddib 2013; Côté et al. 2012; Dorais et al. 1999; Goodrich et al. 2001; Mouaddib et al. 2010). Human intervention could come in different forms such as *teleoperation* (Goldberg et al. 2000) or advice in the form of *goal bias* (Côté et al. 2012). Tools to facilitate human supervision of robots have been developed. Examples include a single human operator supervising a team of robots that can operate with different levels of autonomy (Bechar and Edan 2003), or robots that operate in hazardous environments under human supervision, requiring teleoperation in difficult situations (Ishikawa and Suzuki 1997). There has also been research on mobile robots that can *proactively seek help from people* in their environment to overcome their limitations (Hüttenrauch and Severinson Eklundh 2006; Rosenthal and Veloso 2012; Rosenthal, Veloso, and Dey 2012a; 2012b). In robotics, researchers have started to develop robots that can autonomously identify situations in which a human operator must perform a subtask (Shiomi et al. 2008) and design suitable interaction mechanisms for the collaboration (Yanco, Drury, and Scholtz 2004). However, none of these methods address the full range of challenges discussed above, particularly when there are tight constraints on the timing and conditions for a smooth transfer of control.

There have been previous studies of ways to minimize human effort involved in supervising a semi-autonomous system. One example is the minimization of *neglect time*, which is the period of time during which a request for help can be ignored before performance drops below some threshold (Crandall et al. 2005). Another example is the minimization of *bother cost*, which is the frequency or duration of asking for human help (Cohen et al. 2011). What is still missing is end-to-end task planning techniques that minimize the burden on the human while explicitly factoring the constraints on transfer of control.

In the field of automated planning, there has been interest in *mixed-initiative planning* paradigms that have seen many applications since the 1990's (Burstein and McDermott 1996), such as the MAPGEN system for activity planning for Mars rovers (Bresina et al. 2005). However, the emphasis in this work is on visualization tools that allow people to participate in the planning process itself. That is, the plan itself is built via a collaborative process between a person and an automated system. This is particularly useful when the automated system has a partial model of the environment and human judgment is needed to evaluate candidate plans. While these issues arise in SAS as well, the focus is different in that *plan execution* requires a collaboration between a human and the semi-autonomous system.

Formal frameworks for coordination among multiple agents have been studied for many years in economics (e.g., team decision making theories (Marschak 1955)), control theory (e.g., decentralized detection (Tsitsiklis and Athans 1985)), and multiagent systems (e.g., agent coordination (Durfee 1995; Grosz and Kraus 1996)). In recent years, planning under uncertainty for teams of agents has seen much progress thanks to the development of extensions of the *Markov decision process* (MDP), particularly the DEC-POMDP model (Bernstein et al. 2002; Seuken and Zilberstein 2008; Bernstein et al. 2009; Amato, Bernstein, and Zilberstein 2010). Progress in this area could help design coordination techniques for semi-autonomous systems, but there are still substantial computational and modeling challenges that need to be addressed.

Outside of AI, there has been a growing interest over the past decade in developing a special type of semi-autonomous systems design to avoid dangerous human errors by monitoring human operators (Anderson et al. 2009).

Such systems can then either warn the human operator or take over control to prevent an accident. Accident prevention systems have been developed to enhance driving safety, for example by detecting unintended lane departure (Jung and Kelber 2004) or a potential collision at an intersection. In contrast to accident prevention technology that can take over control only momentarily and in exceptionally risky situations, we focus on settings in which the transfer of control between the human and SAS occurs regularly, and is the norm rather than an exception.

AI research of semi-autonomous systems needs to be informed by human factors research in which the impact of state conditions on human judgment has been studied extensively. In the context of driving, for example, it has been shown that unless drivers scan across a curve some 5-8 seconds before approaching it, they will fail to reduce speed sufficiently (Muttart and Fisher 2013). Transferring control to a driver on a road with relatively dense traffic can be dangerous. Similarly, it is clearly undesirable to transfer control to the driver when errors are more likely due to latent hazards such as sudden appearance of pedestrian crossings (Gomez et al. 2011; Pradhan et al. 2005). When the SAS discovers an unexpected object in the roadway, it may require some information from the driver. However, if the available time is too short, the driver is unlikely to become sufficiently aware of the situation to make an informed decision (Borowsky et al. 2013). Thus, any approach to transfer of control must give the driver enough time to scan the environment and gain situational awareness before the driver is required to take action.

Types of Semi-Autonomous Systems

We define a **semi-autonomous system** (SAS) as a system that can operate autonomously under *some* conditions, but may require human intervention in order to achieve its assigned goals. Although we do not target in this paper any particular domain representation, Markov decision process (MDPs) and partially observable MDPs (POMDPs) can be used as representative models to ground the discussion. We consider a variety of limitations on a SAS that may reduce its ability to achieve its assigned goals. In particular, we are interested in domains where (1) a human operator may have superior abilities to observe or infer the current state of the process (e.g. driving a car); (2) a human operator may be able to perform actions that are not available to the SAS (e.g., climb stairs); or (3) a human operator may have a different level of competence in performing certain actions (e.g., removing a stuck light bulb without breaking it).

In general, the scope of automated planning in semi-autonomous systems may include human interventions, or not. A **SAS of type I** (SAS-I) is a semi-autonomous system whose planning process does not factor possible human interventions. In other words, a SAS-I reasons about the world using a *partial model* in which the only actor is the system itself. It may or may not be able to recognize the conditions under which autonomous operation is not feasible. But even if it can recognize such conditions, no knowledge of human interventions is available and therefore there is no way for the system itself to fully analyze goal reachability.

Arguably, many existing "autonomous systems" are in fact semi-autonomous systems of type I. A Roomba vacuum cleaner, for example, may not know that a person could move it to another floor and its plan may not include such actions. The Curiosity Mars rover may stop progressing towards its goal when certain difficulties are encountered (e.g., high resistance to turning wheels), reporting status to earth and waiting for instructions. Curiosity is programmed to enter a "safe mode" whenever it detects certain irregularities and waits for human engineers to assess the situation and construct a new plan. Roomba and Curiosity are therefore examples of a SAS of type I.

Even when a semi-autonomous system does not possess operational knowledge of possible human interventions, it's useful for it to maintain a safe state until such interventions may take place. We say that a SAS reaches a *dead end* when it can no longer reach its assigned goal, even with human intervention. Otherwise, we say that the SAS is in a *live state*. For example, a Mars rover in a safe mode is designed to simply maintain a live state by keeping its battery charged, communicating with the control center, and waiting for further instructions. If a Mars rover depletes its battery and can no longer charge it, that would be considered a dead end.

Because a SAS-I has no ability to reason about human interventions, inherently, it cannot guarantee maintaining a live state using its own planning and reasoning capabilities. In some cases, the designer of the system may be able to show that it is going to maintain a live state using knowledge that is outside the scope of what the system can reason about.

We define a **SAS of type II** (SAS-II) as a semi-autonomous system whose planning process includes knowledge about possible human interventions and how they can be used to complete the assigned task. In other words, the scope of planning in a SAS-II includes human actions and the uncertainty associated with them. Modeling human actions can be simple or complex depending on the context. For a Roomba vacuum cleaner, it may be sufficient to text its owner and ask "please move me to floor 2" and wait for that action to be performed. However, effective general frameworks for representing human actions and the associated uncertainty are yet to be developed.

There are many challenges associated with modeling human actions. To start, there are no complete operational models of human cognition, so modeling must be based largely on experience. Human competence is not stationary because people often learn over time and can sometimes learn a lot from limited experience, just by thinking about the task. Humans need time for context switching, and the amount of time depends on the activity they perform while neglecting the SAS and how tired they are. Even people with high competence at a certain task are prone to error. And their response could be delayed due to other cognitive demands. All these aspects make planning with a human in the loop much more challenging than planning in fully autonomous systems.

While the assumptions about human responsiveness may be different in different contexts, the uncertainty it introduces raises a fundamental question: can a SAS survive human negligence to respond in a timely manner? Suppose that we assume that human actions are performed *eventually* and that a SAS cannot be neglected by the human operator forever. Given a particular human action model that satisfies this assumption, we define a **strong SAS** as a semi-autonomous system of type II that always maintains a safe state. Otherwise, we refer to the systems as a **weak SAS**.

The ability to maintain a safe state is therefore relative to a *given* model of human actions and a *given* range of objectives. A realistic model of human action is likely to include the possibility of delay in action as well as human errors. A strong SAS must therefore be robust to these characteristics of human behavior, for example, by overriding actions that lead to dead ends and allowing idle actions that prolong the time available for human intervention. Various approaches to *fault-tolerant planning* (Jensen, Veloso, and Bryant 2004; Pineda et al. 2013) offer some useful mechanisms for designing strong SAS.

Building Strong Semi-Autonomous Systems

Building strong semi-autonomous systems requires progress on several fronts. First, it is necessary to develop representations for modeling human actions and the associated uncertainty. Such models have been studied for many years in the area of human factors, but not as much within the context of automated planning. Second, realtime activity and intent recognition techniques are needed to monitor the human's state (Freedman, Jung, and Zilberstein 2014). Third, the ability for people to cooperate seamlessly with a semi-autonomous system depends largely on the availability of suitable interfaces that facilitate communication between the human and machine and fast transfer of control. Additionally, the underlying execution architecture should support semi-autonomy via increased robustness to various faults (Fernández et al. 2001).

Crucially, it is also necessary to develop new general-purpose planning and execution algorithms that can maintain *live state*. When there is enough time to construct a complete plan that covers every reachable state, detection and avoidance of dead ends is relatively straightforward. But in many real-world applications, creating a complete plan before taking the first action is not feasible due to the large size of the state space. Hence, planning and execution must be interleaved and actions must be taken based on an incomplete and often approximate plan. Maintaining live state under these conditions is particularly challenging. Our proposed approach to this problem is based on a *multi-criteria optimization* method where the planning process optimizes first the maintenance of live state and as a secondary goal minimizes the cost of reaching the goal (Wray, Zilberstein, and Mouaddib 2015). Our hypothesis is that often a reasonably small partial plan can be constructed that is complete with respect to safety (live state), but possibly incomplete with respect to goal reachability.

To conclude, semi-autonomy presents new challenges for artificial intelligence. In this paper, we focus particularly on challenges in automated planning. With some careful separation of objectives, planning algorithms could address one of the key problems of maintaining live state, even when goal reachability relies on timely human interventions.

Acknowledgment

Thanks to Avinoam Borowsky, Claudia Goldman, Donald Fisher, Abdel-Illah Mouaddib, Luis Pineda, and Kyle Wray, for fruitful discussions of semi-autonomy. This research was supported in part by NSF grant number IIS-1405550.

References

Amato, C.; Bernstein, D. S.; and Zilberstein, S. 2010. Optimizing fixed-size stochastic controllers for POMDPs and decentralized POMDPs. *Autonomous Agents and Multi-Agent Systems* 21(3):293–320.

Anderson, S. J.; Peters, S. C.; Iagnemma, K. D.; and Pilutti, T. E. 2009. A unified approach to semi-autonomous control of passenger vehicles in hazard avoidance scenarios. In *Proc. of the IEEE Int'l Conference on Systems, Man and Cybernetics*, 2032–2037.

Arkin, R. C. 2009. Ethical robots in warfare. *IEEE Technology and Society Magazine* 28(1):30–33.

Bechar, A., and Edan, Y. 2003. Human-robot collaboration for improved target recognition of agricultural robots. *Industrial Robot: An International Journal* 30(5):432–436.

Bernstein, D. S.; Givan, R.; Immerman, N.; and Zilberstein, S. 2002. The complexity of decentralized control of Markov decision processes. *Mathematics of Operations Research* 27:819–840.

Bernstein, D. S.; Amato, C.; Hansen, E. A.; and Zilberstein, S. 2009. Policy iteration for decentralized control of Markov decision processes. *Journal of Artificial Intelligence Research* 34:89–132.

Borowsky, A.; Horrey, W. J.; Liang, Y.; Garabet, A.; Simmons, L.; and Fisher, D. L. 2013. The effects of momentary visual disruption on hazard anticipation. In *Proc. of Seventh Int'l Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*.

Bradshaw, J. M.; Jung, H.; Kulkarni, S.; Johnson, M.; Feltovich, P.; Allen, J.; Bunch, L.; Chambers, N.; Galescu, L.; Jeffers, R.; Suri, N.; Taysom, W.; and Uszok, A. 2005. Kaa: policy-based explorations of a richer model for adjustable autonomy. In *Proc. of the Fourth Int'l Joint Conference on Autonomous Agents and Multiagent Systems*, 214–221.

Bresina, J.; Jónsson, A.; Morris, P.; and Rajan, K. 2005. Mixed-initiative activity planning for Mars rovers. In *Proc. of the Nineteenth Int'l Joint Conference on Artificial Intelligence*, 1709–1710.

Burstein, M., and McDermott, D. 1996. Issues in the development of human-computer mixed-initiative planning. In Gorayska, B., and Mey, J. L., eds., *Cognitive Technology*. Elsevier. 285–303.

Chaffins, G. 2008. Robotrac is at home down on the farm. NerdBeach.

Cohen, R.; Jung, H.; Fleming, M. W.; and Cheng, M. Y. 2011. A user modeling approach for reasoning about interaction sensitive to bother and its application to hospital decision scenarios. *User Modeling and User-Adapted Interaction* 21(4-5):441–484.

Côté, N.; Canu, A.; Bouzid, M.; and Mouaddib, A.-I. 2012. Humans-robots sliding collaboration control in complex environments with adjustable autonomy. In *Proc. of Intelligent Agent Technology*, 146–153.

Côté, N.; Bouzid, M.; and Mouaddib, A.-I. 2013. Integrating human recommendations in the decision process of autonomous agents. In *Proc. of Intelligent Agent Technology*, 195–200.

Crandall, J. W.; Goodrich, M. A.; Olsen, Jr., D. R.; and Nielsen, C. W. 2005. Validating human-robot interaction schemes in multitasking environments. *IEEE Transactions on Systems, Man and Cybernetics, Part A* 35(4):438–449.

Dorais, G. A.; Bonasso, R. P.; Kortenkamp, D.; Pell, B.; and Schreckenghost, D. 1999. Adjustable autonomy for human-centered autonomous systems. In *IJCAI Workshop on Adjustable Autonomy Systems*, 16–35.

Doyle, R. J. 2003. Autonomy needs and trends in deep space exploration. *Intelligent Systems for Aeronautics* 4(1):4–12.

Durfee, E. H. 1995. Blissful ignorance: Knowing just enough to coordinate well. In *Proc. of the First Int'l Joint Conference on Autonomous Agents and Multiagent Systems*, 406–413.

Fernández, J. L.; Simmons, R.; Sanches, R.; and Diéguez, A. R. 2001. A robust stochastic supervision architecture for an indoor mobile robot. In *Proc. of the Int'l Conference on Field and Service Robotics*, 269–274.

Fong, T.; Nourbakhsh, I.; and Dautenhahn, K. 2003. A survey of socially interactive robots. *Robotics and Autonomous Systems* 143–166.

Freedman, R. G.; Jung, H.-T.; and Zilberstein, S. 2014. Plan and activity recognition from a topic modeling perspective. In *Proc. of the Int'l Conference on Automated Planning and Scheduling*, 360–364.

Frost, C. R. 2010. Challenges and opportunities for autonomous systems in space. In *Proc. of the National Academy of Engineerings U.S. Frontiers of Engineering Symposium*.

Ghallab, M.; Nau, D.; and Traverso, P. 2004. *Automated Planning: Theory and Practice*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Goldberg, K.; Chen, B.; Solomon, R.; Bui, S.; Farzin, B.; Heitler, J.; Poon, D.; and Smith, G. 2000. Collaborative teleoperation via the Internet. In *IEEE Int'l Conference on Robotics and Automation*, 2019–2024.

Gomez, R.; Samuel, S.; Roman Gerardino, L.; Romoser, M. R. E.; and Fisher, D. L. 2011. Do advance yield markings increase safe driver behaviors at unsignalized, marked midblock crosswalks: A simulator study. *Transportation Research Record* 2264:27–33.

Goodrich, M. A.; Olsen, Jr., D. R.; Crandall, J. W.; and Palmer, T. J. 2001. Experiments in adjustable autonomy. In *IEEE Int'l Conference on Systems, Man and Cybernetics*, 1624–1629.

Grosz, B. J., and Kraus, S. 1996. Collaborative plans for complex group action. *Artificial Intelligence* 86:269–357.

Hüttenrauch, H., and Severinson Eklundh, K. 2006. Beyond usability evaluation: Analysis of human-robot interaction at a major robotics competition. *Interaction Studies* 7(3):455–477.

Ishikawa, N., and Suzuki, K. 1997. Development of a human and robot collaborative system for inspecting patrol of nuclear power plants. In *Robot and Human Communication*, 118–123.

Jensen, R. M.; Veloso, M. M.; and Bryant, R. E. 2004. Fault tolerant planning: Toward probabilistic uncertainty models in symbolic non-deterministic planning. In *Proc. of the Fourteenth Int'l Conference on Automated Planning and Scheduling*, 335–344.

Jung, C. R., and Kelber, C. R. 2004. A lane departure warning system based on a linear-parabolic lane model. In *Proc. of IEEE Intelligent Vehicles Symposium*, 891–895.

Longjard, C.; Kumsawat, P.; Attakitmongkol, K.; and Srikaew, A. 2007. Automatic lane detection and navigation using pattern matching mode. In *Proc. of the Seventh WSEAS Int'l Conference on Signal, Speech and Image Processing*, 44–49.

Markoff, J. 2012. Collision in the making between self-driving cars and how the world works. New York Times.

Marschak, J. 1955. Elements for a theory of teams. *Management Science* 1:127–137.

Marsland, S.; Nehmzow, U.; and Duckett, T. 2001. Learning to select distinctive landmarks for mobile robot navigation. *Robotics and Autonomous Systems* 37:241–260.

Mouaddib, A.-I.; Zilberstein, S.; Beynier, A.; and Jeanpierre, L. 2010. A decision-theoretic approach to cooperative control and adjustable autonomy. In *Proc. of the Nineteenth European Conference on Artificial Intelligence*, 971–972.

Muscettola, N.; Nayak, P. P.; Pell, B.; and Williams, B. C. 1998. Remote agent: to boldly go where no AI system has gone before. *Artificial Intelligence* 103(1-2):5–47.

Muttart, J. W., and Fisher, D. L. 2013. Comparing response behaviors of newly licensed drivers to exemplary drivers at high risk curves and turning locations. In *Proc. of Seventh Int'l Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design*.

Nilsson, N. J. 1984. Shakey the robot. Technical Note 323, AI Center, SRI International.

Pell, B.; Bernard, D. E.; Chien, S.; Gat, E.; Muscettola, N.; Nayak, P. P.; Wagner, M. D.; and Williams, B. C. 1996. Remote agent prototype for spacecraft autonomy. In *Proc. of the SPIE Conference on Space Sciencecraft Control and Tracking in the New Millennium*, 74–90.

Pineda, L.; Lu, Y.; Zilberstein, S.; and Goldman, C. V. 2013. Fault-tolerant planning under uncertainty. In *Proc. of the Twenty-Third Int'l Joint Conference on Artificial Intelligence*, 2350–2356.

Pradhan, A. K.; Hammel, K. R.; DeRamus, R.; Pollatsek, A.; Noyce, D. A.; and Fisher, D. L. 2005. The use of eye movements to evaluate the effects of driver age on risk perception in an advanced driving simulator. *Human Factors* 47:840–852.

Rosenthal, S., and Veloso, M. M. 2012. Mobile robot planning to seek help with spatially-situated tasks. In *Proc. of the Twenty-Sixth AAAI Conference on Artificial Intelligence*.

Rosenthal, S.; Veloso, M. M.; and Dey, A. K. 2012a. Acquiring accurate human responses to robots' questions. *Int'l Journal of Social Robotics* 4(2):117–129.

Rosenthal, S.; Veloso, M. M.; and Dey, A. K. 2012b. Is someone in this office available to help me? – proactively seeking help from spatially-situated humans. *Journal of Intelligent and Robotic Systems* 66(1-2):205–221.

Seuken, S., and Zilberstein, S. 2008. Formal models and algorithms for decentralized decision making under uncertainty. *Autonomous Agents and Multi-Agent Systems* 17(2):190–250.

Shiomi, M.; Sakamoto, D.; Kanda, T.; Ishi, C. T.; Ishiguro, H.; and Hagita, N. 2008. A semi-autonomous communication robot: a field trial at a train station. In *Proc. of the 3rd ACM/IEEE Int'l Conference on Human Robot Interaction*, 303–310. New York, NY, USA: ACM.

Thomson, I. 2012. Google promises autonomous cars for all within five years. *The Register*, `http://www.theregister.co.uk/2012/09/25/google_automatic_cars_legal/`.

Tsitsiklis, J. N., and Athans, M. 1985. On the complexity of decentralized decision making and detection problems. *IEEE Transactions on Automatic Control* 30(2):440–446.

White, J. B. 2013. Nissan expects to market self-driving cars by 2020. The Wall Street Journal.

Woodyard, C. 2010. GM takes its EN-V two-wheeler out for a spin. USA Today.

Wray, K. H.; Zilberstein, S.; and Mouaddib, A.-I. 2015. Multi-objective MDPs with conditional lexicographic reward preferences. In *Proc. of the 29th Conference on Artificial Intelligence*.

Yanco, H. A.; Drury, J. L.; and Scholtz, J. 2004. Beyond usability evaluation: Analysis of human-robot interaction at a major robotics competition. *Human–Computer Interaction* 19(1-2):117–149.