The Seeing-Eye Robot Grand Challenge: Rethinking Automated Care

Blue Sky Ideas Track

Reuth Mirsky
The University of Texas at Austin
Austin, Texas
reuth@cs.utexas.edu

Peter Stone
The University of Texas at Austin
Sony AI
Austin, Texas
pstone@cs.utexas.edu

ABSTRACT

Automated care systems are becoming more tangible than ever: recent breakthroughs in robotics and machine learning can be used to address the need for automated care created by the increasing aging population. However, such systems require overcoming several technological, ethical, and social challenges. One inspirational manifestation of these challenges can be observed in the training of seeing-eye dogs for visually impaired people. A seeing-eye dog is not just trained to obey its owner, but also to "intelligently disobey": if it is given an unsafe command from its handler, it is taught to disobey it or even insist on a different course of action. This paper proposes the challenge of building a seeing-eye robot, as a thought-provoking use-case that helps identify the challenges to be faced when creating behaviors for robot assistants in general. Through this challenge, this paper delineates the prerequisites that an automated care system will need to have in order to perform intelligent disobedience and to serve as a true agent for its handler.

KEYWORDS

Automated Care; Service Robots; Surrogacy; Grand Challenge

ACM Reference Format:
Reuth Mirsky and Peter Stone. 2021. The Seeing-Eye Robot Grand Challenge: Rethinking Automated Care: Blue Sky Ideas Track. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), Online, May 3–7, 2021*, IFAAMAS, 6 pages.

1 INTRODUCTION

Recent advances in AI and robotics have enabled impressive breakthroughs in automated care design, whether in service [24, 38, 60, 77], rehabilitation [61, 73, 78], socially assistive care [9, 21, 33, 59], or guidance [18, 47, 71]. Each of these areas provides fertile research grounds and can be discussed individually at length. However, in these works, a "good" care system is one that is obedient and works in a predictable manner under the consent of its handler [16, 58]. We propose a different perspective, where a system might choose to intelligently disobey its handler due to a deep understanding of the handler's intentions. With this long-term vision, we propose the seeing-eye robot grand challenge as a guiding use-case that will help imagine how to design an autonomous care system that is able to make decisions as a knowledgeable extension of its handler.

Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), U. Endriss, A. Nowé, F. Dignum, A. Lomuscio (eds.), May 3–7, 2021, Online.
© 2021 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). All rights reserved.

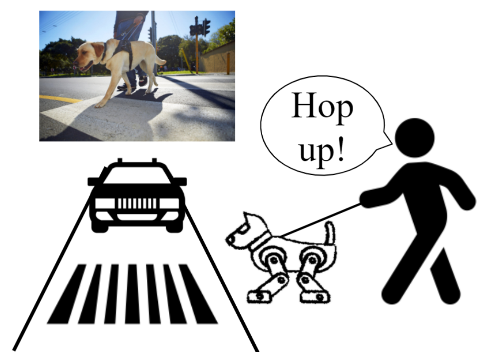

Figure 1: "Hop Up" is the command for a seeing-eye dog to move forward when stalling. These dogs learn to ignore this command if they perceive obeying it to be dangerous.

We emphasize that we focus on the decision making of the robot, not the perception, actuation, or physical design: this challenge deals with the robot's brains, not its eyes, ears, or body. To start discussing these decision processes in existing autonomous systems, we refer to Asimov's laws of robotics [4] as a reasonable opening from which we can then elaborate on necessary research directions:

(1) A robot may not injure a human being or, through inaction, allow a human being to come to harm.
(2) A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
(3) A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

Existing autonomous systems enforce the first law through hard constraints upon the system's abilities. Fully implementing the second law is more intricate, and the main contribution of many modern systems is to provide a better understanding of the instructions given by a human [7, 29], within the limits of the "first law" safety restrictions [1, 35, 40, 55]. Moreover, modern AI systems will need to reason about cases in which the "right" thing to do is the opposite from the instruction given by the handler. Sometimes there is an even more subtle conflict between an instruction given by a human, and the desired outcome due to an imperfect instruction. These conflicts are examples of *the value alignment problem* [22].

Modern autonomous systems do not yet have the required level of context- and state- understanding to be able to reason about the aforementioned conflicts. To address this challenge, many works let a human be the judge in complex decision-making scenarios [31, 69]. Alternatively, Murphy and Woods [49] proposed a modified set for responsible robotics laws. In their proposal, the second law becomes "*A robot must respond to humans as appropriate for their roles*". This modification allows the designers of a robot to bypass the need to understand the meaning and implications of an instruction that conflicts with the first law. While this modification is a pragmatic solution given existing technology, we present a long-term vision where an autonomous system will indeed be able to reason about such cases and perform adequate autonomous decision making.

While acknowledging the unique benefits and abilities of modern AI to solve "second law" gaps, the most sophisticated intelligence we know of is still the natural mind, and we take inspiration from it to address the issues discussed above, in the form of the intelligent capabilities of a dog. *Intelligent Disobedience* occurs when a seeing-eye dog, who is trained to help its handler, goes directly against the handler's instructions in an effort to make a better decision [14]. We advocate that a comprehensive treatment of "second law" issues must include a component of intelligent disobedience.

This call is not the first to propose the design of a robot that will replace a service dog: Tachi and Komoriya were the first to construct a robotic guide dog for collision avoidance, and they discussed a specific implementation of intelligent disobedience in the context of obstacle avoidance [67]. Recently, several works proposed various mechanisms that can partially or completely replace the functionality of a white cane by signaling about proximity to obstacles. Sakhardande et al. [57] devised a detachable device that can augment an existing white cane. Additional works proposed enhanced walkers [19, 74, 76]. Other solutions that offer a guiding system for the blind include canes [11, 30], smartphones [2, 37] and wearable augmented reality devices [42]. In a participatory study, Azenkot et al. [5] investigated what people with visual impairments care most about with regards to the behavior of a guide robot in an indoor environment.

However, all of the works mentioned above propose a passive apparatus that can advise using various cues but cannot physically enforce its decisions in dangerous situations [54]. As advocates of embodied intelligent disobedience, we seek a solution in which the robot is an active guide which can have the last word in specific, hazardous scenarios. This requirement might be achieved by extending the application of a personal assistant robot that is already present in the handler's environment [52], but such a solution relies on existing hardware and is likely to be unsatisfactory given the unique requirement of a guide robot: it needs to have excellent maneuverability and robustness, otherwise it will simply replace a visual impairment with a mobility impairment.

2 THE SEEING-EYE ROBOT CHALLENGE

There are more than 1 million people in the US who are blind, millions more with severe visual impairment, and this number is expected to multiply by 5 in the next 30 years. However, only about 2% of blind people work with guide dogs. It costs over $50,000 and up to two years to breed, raise, train, and place one guide dog, and the net time it can serve as a seeing-eye dog is less than ten years before it retires. Moreover, a dog requires constant care from its handler, including daily walks, regardless of weather conditions. Even under these requirements, the use of a service dog can greatly increase a blind person's freedom and ability to integrate in the community. In order to enable more people to enjoy the benefits of a guide dog, even if they cannot afford or take care of one, we propose the following grand challenge:

> Can we design and build a service robot that can replace or surpass the functionalities of a seeing-eye dog?

We define some basic terms to enable a clear discussion about what steps and problems this challenge encompasses: **The Robot** is the combination of the physical embodiment of the seeing-eye robot, along with any cognitive processes it leverages, both locally and remotely. **The Handler** is the person who handles the robot. We assume this person is either blind or has a severe visual impairment. **Passers** are any additional people that might interact with the robot and its handler: other people that accompany the handler, pedestrians, service providers, etc.

Consider the specific scenario depicted in Figure 1. In this case the handler wishes to cross the road and is unaware of the approaching vehicle. Ideally, the robot should refuse to cross the road, even if the handler insists on moving forward, using the command "hop up", which is the common way to coerce a guide-dog to move on. In addition, we lay out the different components of the process of intelligent disobedience, all of which the robot must be capable:

- **Global Objectives** understand a set of standing objectives like "keep the handler safe". This stage requires an understanding of the environment and comprehension of the abilities of the robot under "first law" constraints.
- **Local Objectives** understand what the handler wants to do now. This stage introduces the value alignment problem.
- **Plan Recognition** understand how the handler thinks the local objectives should be accomplished. This step requires yet again an understanding of the environment, its objective limitations, and theory of mind in relation to the handler.
- **Consistency Check** judge whether the given instruction is in conflict with the global objectives.
- **Mediation** if so, make decisions about the given instruction vs. the global objectives. This mediation could range from coming up with a different way to accomplish the local objectives, to ignoring them altogether.

An important *global objective* in our running example from Figure 1 is the safety of the handler. The *local objective* in this scenario is the destination of the handler. Notice that there might be more than one crossings nearby, which means that there might be different plans to achieve this objective. *Plan recognition* is the stage where the robot should disambiguate these different plans and recognize which one the handler is following. The robot will then perform a *consistency check* to find that moving forward, as requested, will conflict with the safety global objective. At this point, the robot should be certain enough in its understanding of the global and local objectives to decide to override the "hop up" command. The robot should then use *mediation* which might just be stopping and waiting for the environment to change, or leveraging alternative plans and guiding the handler towards a different crossing, knowing that there is an alternative path to the handler's goal location.

3 REQUIREMENTS

As a way to raise points that are relevant to all automated care systems, we now proceed to discuss in depth the scientific requirements of each of these steps, and mention relevant research areas that can be called upon for assistance. In addition, in each component we refer to specific efforts of our research community that can improve the robot's capabilities along the path to intelligent disobedience and better automated care in general. Finally, we briefly discuss how these steps should be assessed and what metrics can be used to create a future standard evaluation.

3.1 Global Objectives

Many global objectives of the robot refer to its safety around its handler and passers [68]. Notice that enforcing these safety constraints might restrict the robot (for example, a smart wheelchair that will keep you from hitting a wall), however these restrictions are not considered intelligent disobedience as they do not require reasoning about the goal of the handler. The robot should also behave in a predictable and explicable way when interacting with new people [13, 16] or following social norms [48]. Lastly, the robot might have additional global interests that go beyond serving the immediate interests of its handler, such as logistical constraints (e.g., battery time) or data collection for self-improvement [15, 64, 65]. Specifying these objectives and balancing between them are closely related to inverse reinforcement learning (IRL) and can benefit from leveraging this approach [27, 51]. Thus, the robot's performance in this step can be evaluated using common metrics used in IRL and other learning approaches: accuracy, precision, and efficiency.

3.2 Local Objectives

This stage in the process of intelligent disobedience requires the robot to take into consideration the current local goals of the handler: most of the time, this goal is to reach a specific location, but it can also encompass opportunistic goals such as letting the handler know that they are passing a new grocery store, that a bus is reaching the station, or that a familiar person is near [12, 34]. Local objectives change, by definition, so the robot will need to progressively assess these objectives during execution [8, 56, 72].

A key step in understanding local objectives is to clearly convey this information to the robot – this communication can be verbal and rely on NLP [43, 70] or by using a controller to portray instructions [20]. Moreover, a potential ability of a robot that can surpass a dog's is that it cannot only understand vocal commands, but also speak or question its handler. This aspect of the problem involves (1) What additional information should be shared with the handler? (2) When will communicating a piece of information be valuable, and when is it interrupting? Lastly, regardless of the model and the modality chosen for conveying local objectives to the robot, the robot will still need to correctly assess these objectives, while avoiding failures due to poor coordination [22, 28]. The evaluation of this step is similar to the first one but will require evaluating the handler's subjective impressions as is often done in HRI research [21, 23]. For example, how does the handler perceive the robot's abilities and how do they convey the goal behind specific actions?

3.3 Plan Recognition

Reasoning about the plans of teammates is one of the biggest areas in which natural intelligence still surpasses AI. The concept of "theory of mind" is often used to model other agents and their goals [6, 36, 41, 62]. This challenge involves not only to understand *what* is the goal of the handler, but also *how* they plan to achieve it. Current state of the art algorithms are both expressive and fast enough to be used in real-world settings to infer human traces in closed environments with predefined settings [44, 53]. In order to be useful "in the wild", the next generation of recognizers will need to overcome the gap between local objective understanding and plan prediction in open-world environments. An additional challenge related to plan recognition that the robot will need to overcome is ambiguity: there can be multiple hypotheses that can explain a sequence of observed actions. An active observer can interact with the actor in order to disambiguate between those hypotheses [45, 46, 63]. Thus, the robot should take into consideration both its ability to correctly recognize plans in new environments and to disambiguate between competing hypotheses.

3.4 Consistency Check

Evidence shows that prelinguistic children will not only recognize the plans of adults they do not know, but will also detect a plan failure and will act to help the adult achieve their plan [39, 75]. The robot will need to evaluate the recognized plan and to see if it fits both the global and the local objectives that were defined.

If the plan does not suit a local objective, there might still be a possibility to resolve the conflict without invoking intelligent disobedience: the robot might communicate with the handler for additional clarifications or explain the inconsistency between the objective and the proposed plan to accomplish it [13, 66]. If this effort fails, or if a plan conflicts with a strict global objective, the proposed plan cannot be executed and a different solution should be found, as we elaborate on next. The evaluation of the robot's capabilities in this step should consider two aspects: capturing inconsistencies efficiently and in a timely manner, and the percentage of cases in which the robot was able to resolve a conflict without reaching intelligent disobedience.

3.5 Mediation

Once the robot decides that intelligent disobedience is needed, it will require a reasonable amount of force to stop the handler from executing the plan, while still keeping in mind the handler's safety. Impedance control, an approach to dynamically control the forces and position of the robot, will be needed to be personalized and adaptable so that the robot's actions remain safe and efficient [1]. With respect to personalization, the *context* in which the robot is acting can influence the type and the intensity of the disobedience act. For example, ignoring a drift to the left that might cause the handler to get off the sidewalk will require a different mechanism than an emergency brake to avoid a passing car [17, 25].

No matter what action the robot chooses to take, it will need to predict or estimate how its actions will affect the handler, in order to avoid any further conflicts with the global and local objectives [50]. In addition, as the final step in the process of intelligent disobedience, evaluation of this step can also encompass the general ability of the robot in this task, which means measuring the robot's success at achieving safe, efficient, and explainable disobedience.

4 ADDITIONAL MILESTONES

So far, we discussed the main components that the seeing-eye robot will need to have in order to be a useful proxy for its handler while employing intelligent disobedience. All of the identified landmarks and the accomplishments mentioned above are just the beginning – each solution that was presented opens up many new questions, each of which supports new fascinating research directions.

There is also an abundance of challenges related to the deployment of a seeing-eye robot, that are not of immediate use in the intelligent disobedience process but will need to be considered: the robot form factors, perception, battery life, navigation, and more. These issues can stimulate and motivate interdisciplinary efforts towards the core decision-making challenge. Moreover, since they can be decoupled from the decision making challenge, different solutions can be developed for different hardware and for varying environments. We provide a more detailed discussion of specific challenges that are independent from the intelligent disobedience discussion but are necessary for the realization of fully functional seeing-eye robots.

Rehabilitation and Teaching. Diabetes is the leading cause of blindness in American adults, and diabetic people with blindness might have other health issues that can also affect their mobility. In other cases, people's visual impairment can interfere with their balance, but if assisted by a guide dog, they can rest their leg against the dog as a balance cue. These types of behaviors are dynamic and fast, which can make it hard for the robot to respond properly. A challenging task can be to model this problem as a constraint satisfaction problem and use existing techniques to solve it. In addition, the robot can be used for long-term rehabilitation where it changes the level of support it provides the person. For this aspect, curriculum learning can be used to model the learning progress and to design the robot's behavior [10]. This challenge will require interdisciplinary collaborations with rehabilitation therapists and mechanical engineers that will aim to enhance the robot with balancing and rehabilitation abilities.

Social Companionship. It is important to note that being a social companion is not the main goal of the seeing-eye robot. However, it is also not the main goal of seeing-eye dogs and yet they are perceived as companions. An artificial guide robot that is explicitly meant to replace a seeing-eye dog is thus also an opportunity to investigate its performance as a social companion. A unique challenge in this context will be to investigate the dynamics between the robot's functional performance and its social behavior and perceived social abilities. This type of project will require assistance from HRI researchers and psychologists.

Ethics. The seeing-eye robot can be fertile ground for discussing the role of ethics in AI. Any care system should respect and support autonomy, which might conflict with the need to perform intelligent disobedience [32]. Once the robot will be given the ability to override the decisions of its handler, its designers become accountable for the consequences of these decisions, and hence there is a need to define its morals in a more crisp sense than Asimov's rules [3]. Moreover, in the design of any particular solution, the engineering process will require incorporating a diversity of perspectives: the cost of a functioning robot that is capable of the proposed feats is not negligible, and will not be accessible to all. The community should aim to produce a robot that costs significantly less than the cost of raising and training a seeing-eye dog.

Human and Environment Engineering. In addition to enhancements and improvements that will be applied to the robot, this challenge can include comprehensive solutions that change the environment in such a way that will take the burden off the robot. For example, Bluetooth beacons can assist the robot with identifying its current location without the need to use expensive sensors or GPS signals [26]. In a similar fashion, the human mind and body can adapt greatly to changes, and some of the challenges discussed in this paper will be most efficiently solved by human training, such that the handler or passers will adapt to the robot's abilities instead of the other way around. For comparison, the training process of a handler with a new seeing-eye dog takes about two weeks and might require occasional guidance after that.

5 DISCUSSION

We have introduced a long-term challenge to design a robot that can be at least as intelligent and capable as a seeing-eye dog. We discussed the different steps this robot will need to go through in order to perform intelligent disobedience and proposed initial ways to approach these steps and how to assess the robot's performance. This challenge crosses disciplines and motivates new ideas and directions for the Autonomous Agents and Multi-Agent Systems community to explore. The ultimate goal of this challenge will be to create robots that can make better decisions than seeing-eye dogs. However, even partial solutions may have immediate utility, especially for people who don't have access to seeing-eye dogs or people who use non-robotic guidance systems.

We also acknowledge that many of the approaches proposed throughout the paper may require expensive, durable, and robust hardware as well as massive computation power, which could cause useful robots to be drained of battery power quickly. However, given the increasing speed in which the robotics field is advancing, we remain optimistic and hope that this challenge will inspire the community towards groundbreaking advances. We therefore put forth this seeing-eye robot as a new grand challenge for the community to address, that will ultimately improve handlers' autonomy, while leading to new, broadly applicable agent technologies.

ACKNOWLEDGMENTS

The authors thank Scott Niekum and Aaron Steinfeld for helpful discussions on the ideas presented in this paper. This work has taken place in the Learning Agents Research Group (LARG) at UT Austin. LARG research is supported in part by NSF (CPS-1739964, IIS-1724157, NRI-1925082), ONR (N00014-18-2243), FLI (RFP2-000), ARO (W911NF-19-2-0333), DARPA, Lockheed Martin, GM, and Bosch. Peter Stone serves as the Executive Director of Sony AI America and receives financial compensation for this work. The terms of this arrangement have been reviewed and approved by the University of Texas at Austin in accordance with its policy on objectivity in research.

REFERENCES

[1] Priyanshu Agarwal and Ashish D Deshpande. 2015. Impedance and force-field control of the index finger module of a hand exoskeleton for rehabilitation. In *2015 IEEE International Conference on Rehabilitation Robotics (ICORR)*. IEEE, 85–90.

[2] Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris Kitani, Hironobu Takagi, and Chieko Asakawa. 2016. NavCog: a navigational cognitive assistant for the blind. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services*. 90–99.

[3] Thomas Arnold, Daniel Kasenberg, and Matthias Scheutz. 2017. Value alignment or misalignment-what will keep systems accountable?. In *AAAI Workshops*.

[4] Isaac Asimov. 2004. *I, robot*. Spectra.

[5] Shiri Azenkot, Catherine Feng, and Maya Cakmak. 2016. Enabling building service robots to guide blind people a participatory design approach. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 3–10.

[6] Chris Baker, Rebecca Saxe, and Joshua Tenenbaum. 2011. Bayesian theory of mind: Modeling joint belief-desire attribution. In *Proceedings of the annual meeting of the cognitive science society*, Vol. 33.

[7] Kim Baraka, Francisco S Melo, Marta Couto, and Manuela Veloso. 2020. Optimal action sequence generation for assistive agents in fixed horizon tasks. *Auton. Agents Multi Agent Syst.* 34, 2 (2020), 33.

[8] Andrea Bauer, Dirk Wollherr, and Martin Buss. 2008. Human–robot collaboration: a survey. *International Journal of Humanoid Robotics* 5, 01 (2008), 47–66.

[9] Roger Bemelmans, Gert Jan Gelderblom, Pieter Jonker, and Luc De Witte. 2012. Socially assistive robots in elderly care: A systematic review into effects and effectiveness. *Journal of the American Medical Directors Association* 13, 2 (2012), 114–120.

[10] Yoshua Bengio, Jérôme Louradour, Ronan Collobert, and Jason Weston. 2009. Curriculum Learning. In *Proceedings of the 26th Annual International Conference on Machine Learning*. ACM, 41–48.

[11] Johann Borenstein and Iwan Ulrich. 1997. The guidecane-a computerized travel aid for the active guidance of blind pedestrians. In *Proceedings of International Conference on Robotics and Automation*, Vol. 2. IEEE, 1283–1288.

[12] John F Canny and Ming C Lin. 1993. An opportunistic global path planner. *Algorithmica* 10, 2-4 (1993), 102–120.

[13] Tathagata Chakraborti, Anagha Kulkarni, Sarath Sreedharan, David E Smith, and Subbarao Kambhampati. 2019. Explicability? legibility? predictability? transparency? privacy? security? the emerging landscape of interpretable agent behavior. In *Proceedings of the international conference on automated planning and scheduling*, Vol. 29. 86–96.

[14] Ira Chaleff. 2015. *Intelligent disobedience: Doing right when what you're told to do is wrong*. Berrett-Koehler Publishers.

[15] Zhiyuan Chen and Bing Liu. 2018. Lifelong machine learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning* 12, 3 (2018), 1–207.

[16] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. 2013. Legibility and predictability of robot motion. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 301–308.

[17] Anca D Dragan and Siddhartha S Srinivasa. 2013. A policy-blending formalism for shared control. *The International Journal of Robotics Research* 32, 7 (2013), 790–805.

[18] Bosede I Edwards and Adrian D Cheok. 2018. Why not robot teachers: artificial intelligence for addressing teacher shortage. *Applied Artificial Intelligence* 32, 4 (2018), 345–360.

[19] Vitor Faria, Jorge Silva, Maria Martins, and Cristina Santos. 2014. Dynamical system approach for obstacle avoidance in a Smart Walker device. In *2014 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*. IEEE, 261–266.

[20] Ildar Farkhatdinov, Jee-Hwan Ryu, and Jury Poduraev. 2009. A user study of command strategies for mobile robot teleoperation. *Intelligent Service Robotics* 2, 2 (2009), 95–104.

[21] David Feil-Seifer and Maja J Mataric. 2005. Defining socially assistive robotics. In *9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005*. IEEE, 465–468.

[22] Jaime F Fisac, Monica A Gates, Jessica B Hamrick, Chang Liu, Dylan Hadfield-Menell, Malayandi Palaniappan, Dhruv Malik, S Shankar Sastry, Thomas L Griffiths, and Anca D Dragan. 2020. Pragmatic-pedagogic value alignment. In *Robotics Research*. Springer, 49–57.

[23] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. A survey of socially interactive robots. *Robotics and autonomous systems* 42, 3-4 (2003), 143–166.

[24] Jodi Forlizzi and Carl DiSalvo. 2006. Service robots in the domestic environment: a study of the roomba vacuum in the home. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 258–265.

[25] Joseph Funke, Matthew Brown, Stephen M Erlien, and J Christian Gerdes. 2016. Collision avoidance and stabilization for autonomous vehicles in emergency scenarios. *IEEE Transactions on Control Systems Technology* 25, 4 (2016), 1204–1216.

[26] Aura Ganz, James M Schafer, Yang Tao, Larry Haile, Charlene Sanderson, Carole Wilson, and Meg Robertson. 2015. PERCEPT based interactive wayfinding for visually impaired users in subways. (2015).

[27] Yang Gao, Jan Peters, Antonios Tsourdos, Shao Zhifei, and Er Meng Joo. 2012. A survey of inverse reinforcement learning techniques. *International Journal of Intelligent Computing and Cybernetics* (2012).

[28] Dylan Hadfield-Menell, Smitha Milli, Pieter Abbeel, Stuart J Russell, and Anca Dragan. 2017. Inverse reward design. *Advances in neural information processing systems* 30 (2017), 6765–6774.

[29] Dylan Hadfield-Menell, Stuart J Russell, Pieter Abbeel, and Anca Dragan. 2016. Cooperative inverse reinforcement learning. In *Advances in neural information processing systems*. 3909–3917.

[30] Marion Hersh and Michael A Johnson. 2010. *Assistive technology for visually impaired and blind people*. Springer Science & Business Media.

[31] Andreas Holzinger. 2016. Interactive machine learning for health informatics: when do we need the human-in-the-loop? *Brain Informatics* 3, 2 (2016), 119–131.

[32] Michael N Huhns and Munindar P Singh. 2005. Service-oriented computing: Key concepts and principles. *IEEE Internet computing* 9, 1 (2005), 75–81.

[33] Reza Kachouie, Sima Sedighadeli, Rajiv Khosla, and Mei-Tai Chu. 2014. Socially assistive robots in elderly care: a mixed-method systematic literature review. *International Journal of Human-Computer Interaction* 30, 5 (2014), 369–393.

[34] Ece Kamar, Ya'akov Gal, and Barbara J Grosz. 2009. Incorporating helpful behavior into collaborative planning. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. Springer Verlag.

[35] Ankur Kapoor, Ming Li, and Russell H Taylor. 2006. Constrained Control for Surgical Assistant Robots.. In *ICRA*. 231–236.

[36] Henry A Kautz and James F Allen. 1986. Generalized Plan Recognition.. In *AAAI*, Vol. 86. 5.

[37] Seita Kayukawa, Hironobu Takagi, João Guerreiro, Shigeo Morishima, and Chieko Asakawa. 2020. Smartphone-Based Assistance for Blind People to Stand in Lines. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.

[38] Thomas Keller, Patrick Eyerich, and Bernhard Nebel. 2012. Task planning for an autonomous service robot. In *Towards Service Robots for Everyday Environments*. Springer, 117–124.

[39] Birgit Knudsen and Ulf Liszkowski. 2012. Eighteen- and 24-month-old infants correct others in anticipation of action mistakes. *Developmental science* 15 (01 2012), 113–22. https://doi.org/10.1111/j.1467-7687.2011.01098.x

[40] Axel Lankenau, Oliver Meyer, and Bernd Krieg-Bruckner. 1998. Safety in robotics: The bremen autonomous wheelchair. In *AMC'98-Coimbra. 1998 5th International Workshop on Advanced Motion Control. Proceedings (Cat. No. 98TH8354)*. IEEE, 524–529.

[41] Clayton Lewis and Donald A Norman. 1995. Designing for error. In *Readings in Human–Computer Interaction*. Elsevier, 686–697.

[42] Yang Liu, Noelle RB Stiles, and Markus Meister. 2018. Augmented reality powers a cognitive assistant for the blind. *ELife* 7 (2018), e37841.

[43] Cynthia Matuszek. 2018. Grounded Language Learning: Where Robotics and NLP Meet.. In *IJCAI*. 5687–5691.

[44] Reuth Mirsky, Ya'akov Gal, and Stuart M Shieber. 2017. CRADLE: an online plan recognition algorithm for exploratory domains. *ACM Transactions on Intelligent Systems and Technology (TIST)* 8, 3 (2017), 1–22.

[45] Reuth Mirsky, William Macke, Andy Wang, Harel Yedidsion, and Peter Stone. 2020. A Penny for Your Thoughts: The Value of Communication in Ad Hoc Teamwork. *International Joint Conference on Artificial Intelligence (IJCAI)* (2020).

[46] Reuth Mirsky, Roni Stern, Kobi Gal, and Meir Kalech. 2018. Sequential plan recognition: An iterative approach to disambiguating between hypotheses. *Artificial Intelligence* 260 (2018), 51–73.

[47] Aaron Morris, Raghavendra Donamukkala, Anuj Kapuria, Aaron Steinfeld, Judith T Matthews, Jacqueline Dunbar-Jacob, and Sebastian Thrun. 2003. A robotic walker that provides guidance. In *2003 IEEE International Conference on Robotics and Automation (Cat. No. 03CH37422)*, Vol. 1. IEEE, 25–30.

[48] Yoram Moses and Moshe Tennenholtz. 1995. Artificial social systems. *Computers and Artificial Intelligence* 14 (1995), 533–562.

[49] Robin Murphy and David D Woods. 2009. Beyond Asimov: the three laws of responsible robotics. *IEEE intelligent systems* 24, 4 (2009), 14–20.

[50] Amal Nanavati, Xiang Zhi Tan, Joe Connolly, and Aaron Steinfeld. 2019. Follow The Robot: Modeling Coupled Human-Robot Dyads During Navigation. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*.

[51] Andrew Y Ng, Stuart J Russell, et al. 2000. Algorithms for inverse reinforcement learning.. In *Icml*, Vol. 1. 2.

[52] Jordi Palacín, Eduard Clotet, Dani Martínez, David Martínez, and Javier Moreno. 2019. Extending the application of an assistant personal robot as a walk-helper tool. *Robotics* 8, 2 (2019), 27.

[53] Ramon Fraga Pereira, Nir Oren, and Felipe Meneguzzi. 2017. Landmark-based heuristics for goal recognition. In *Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*. AAAI Press.

[54] Santiago Real and Alvaro Araujo. 2019. Navigation systems for the blind and visually impaired: Past work, challenges, and open problems. *Sensors* 19, 15

(2019), 3404.

[55] Leonel Rozo Castañeda, Sylvain Calinon, Darwin Caldwell, Pablo Jimenez Schlegl, and Carme Torras. 2013. Learning collaborative impedance-based robot behaviors. In *Proceedings of the twenty-seventh AAAI conference on artificial intelligence*. 1422–1428.

[56] Sandhya Saisubramanian, Ece Kamar, and Shlomo Zilberstein. 2020. Mitigating the negative side effects of reasoning with imperfect models: A multi-objective approach. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*. 1984–1986.

[57] Jayant Sakhardande, Pratik Pattanayak, and Mita Bhowmick. 2012. Smart cane assisted mobility for the visually impaired. *International Journal of Electrical and Computer Engineering* 6, 10 (2012), 1262–1265.

[58] Vasanth Sarathy, Thomas Arnold, and Matthias Scheutz. 2019. When Exceptions Are the Norm: Exploring the Role of Consent in HRI. *ACM Transactions on Human-Robot Interaction (THRI)* 8, 3 (2019), 1–21.

[59] Brian Scassellati, Henny Admoni, and Maja Matarić. 2012. Robots for use in autism research. *Annual review of biomedical engineering* 14 (2012).

[60] Kerstin Severinson-Eklundh, Anders Green, and Helge Hüttenrauch. 2003. Social and collaborative aspects of interaction with a service robot. *Robotics and Autonomous systems* 42, 3-4 (2003), 223–234.

[61] Takanori Shibata and Joseph F Coughlin. 2014. Trends of robot therapy with neurological therapeutic seal robot, PARO. *Journal of Robotics and Mechatronics* 26, 4 (2014), 418–425.

[62] Maayan Shvo, Toryn Q Klassen, Shirin Sohrabi, and Sheila A McIlraith. 2020. Epistemic Plan Recognition. In *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*. 1251–1259.

[63] Maayan Shvo and Sheila A. McIlraith. 2020. Active Goal Recognition.. In *The AAAI conference on artificial intelligence*. 9957–9966.

[64] Daniel L Silver, Qiang Yang, and Lianghao Li. 2013. Lifelong machine learning systems: Beyond learning algorithms. In *2013 AAAI spring symposium series*.

[65] Grant P Strimel and Manuela M Veloso. 2014. Coverage planning with finite resources. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2950–2956.

[66] Katia P Sycara. 1989. Argumentation: Planning Other Agents' Plans.. In *IJCAI*, Vol. 89. 20–25.

[67] Susumu Tachi and Kiyoshi Komoriya. 1984. Guide dog robot. *Autonomous mobile robots: Control, planning, and architecture* (1984), 360–367.

[68] Tadele Shiferaw Tadele, Theo de Vries, and Stefano Stramigioli. 2014. The safety of domestic robotics: A survey of various safety-related publications. *IEEE robotics & automation magazine* 21, 3 (2014), 134–142.

[69] Kazuaki Takeuchi, Yoichi Yamazaki, and Kentaro Yoshifuji. 2020. Avatar Work: Telework for Disabled People Unable to Go Outside by Using Avatar Robots. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 53–60.

[70] Jesse Thomason, Aishwarya Padmakumar, Jivko Sinapov, Nick Walker, Yuqian Jiang, Harel Yedidsion, Justin Hart, Peter Stone, and Raymond Mooney. 2020. Jointly improving parsing and perception for natural language commands through human-robot dialog. *Journal of Artificial Intelligence Research* 67 (2020), 327–374.

[71] Sebastian Thrun, Maren Bennewitz, Wolfram Burgard, Armin B Cremers, Frank Dellaert, Dieter Fox, Dirk Hahnel, Charles Rosenberg, Nicholas Roy, Jamieson Schulte, et al. 1999. MINERVA: A second-generation museum tour-guide robot. In *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*, Vol. 3. IEEE.

[72] Viet Trinh and Roberto Manduchi. 2020. Feeling your way around: Assessing the perceived utility of multi-scale indoor tactile maps. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–8.

[73] BT Volpe, HI Krebs, N Hogan, L Edelstein, C Diels, and M Aisen. 2000. A novel approach to stroke rehabilitation: robot-aided sensorimotor stimulation. *Neurology* 54, 10 (2000), 1938–1944.

[74] Andreas Wachaja, Pratik Agarwal, Mathias Zink, Miguel Reyes Adame, Knut Möller, and Wolfram Burgard. 2017. Navigating blind people with walking impairments using a smart walker. *Autonomous Robots* 41, 3 (2017), 555–573.

[75] Felix Warneken and Michael Tomasello. 2007. Helping and Cooperation at 14 Months of Age. *Infancy, v.11, 271-294 (2007)* 11 (05 2007). https://doi.org/10.1080/15250000701310389

[76] Christian Werner, George P Moustris, Costas S Tzafestas, and Klaus Hauer. 2018. User-oriented evaluation of a robotic rollator that provides navigation assistance in frail older adults with and without cognitive impairment. *Gerontology* 64, 3 (2018), 278–290.

[77] D Mitchell Wilkes, A Alford, Robert T Pack, T Rogers, RA Peters, and Kazuhiko Kawamura. 1998. Toward socially intelligent service robots. *Applied Artificial Intelligence* 12, 7-8 (1998), 729–766.

[78] Jungwon Yoon, Bondhan Novandy, Chul-Ho Yoon, and Ki-Jong Park. 2010. A 6-DOF gait rehabilitation robot with upper and lower limb connections that allows walking velocity updates on various terrains. *IEEE/ASME Transactions on Mechatronics* 15, 2 (2010), 201–215.