Slightly Subversive Methods for Promoting Use of Autonomy in Robots

Aaron Steinfeld
Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, steinfeld@cmu.edu

Abstract— **Roboticists have frequently found that users are notoriously unwilling to pass control to autonomy, even when appropriate. It is important to find ways to promote the use of autonomy when switching out of manual mode is needed. This becomes important when autonomy is inherently better equipped to support the task at hand, especially when autonomy is safer than manual control. This paper proposes a selection of slightly subversive interaction design approaches that leverage knowledge about human behavior and may lead to greater utilization of autonomy by robot users.**

Keywords-component; human-robot interaction, autonomy, interaction design

I. Introduction

There are many situations where users should activate autonomy sparingly due to the well-established loss of situation awareness and problems with vigilance during autonomous modes. However, there are certainly cases where the value of using robot autonomy outweighs these concerns, especially when autonomy is inherently safer than manual control.

For example, there are known problems with manual teleoperation. McGovern [1, 2] studied rate-controlled ground vehicles (e.g., dune buggies with video cameras) and reported operator problems including slow driving, imprecise control, loss of situational awareness, poor attitude and depth judgment, and failure to detect obstacles. He also concluded that many vehicle failures (collision, roll-over, etc.) were traceable to these operator problems. Enthusiasts have integrated smartphones with an actuated car and demonstrated intuitive teleoperation but their released videos clearly show the interaction requires significant attention by the operator [3]. Thus, basic teleoperation of vehicles may be greatly aided by systems that reduce the need for continuous manual control. These same methods may also reduce errors in real-time manual teleoperation [1, 2, 4] and reduce weaknesses under communication constraints [5].

These findings also suggest that, in some applications, manual teleoperation may not be adequately robust or safe for the general population. This concern is reinforced by the lack of human safety monitors typically found during many robotics experiments. Capable autonomy can address these weakness and lead to safer robot handling.

Unfortunately, roboticists have found that users are notoriously unwilling to pass control to autonomy, even when appropriate. It is important to find ways to promote the use of autonomy when we want users to switch out of manual mode. This paper proposes a handful of interaction design approaches that leverage knowledge about human behavior and should lead to greater use of autonomy by robot users. In most cases, deception is part of the proposed approach.

II. Methods for Promoting Autonomy Use

A. Familiarity

The easiest way to encourage autonomy utilization over manual control is to increase familiarity. Past use of automated systems has been shown to increase likelihood of continued use, even when operators think they might be able to perform the task better than an automated component [6]. Research on other intelligent systems has produced comparable findings where operators with more exposure time reported higher favorable opinions [7]. Lee and Moray [6] referred to this tendency as "inertia" and recommended designing systems so that reallocation of control is simple. They also suggested that new operators should be encouraged to use automation in order to mitigate their inherent bias towards manual control.

The automotive industry is, intentionally or not, pursuing a strategy that promotes familiarity. Due to concerns over the ability to leap directly to fully autonomous cars after successful demonstrations by the National Automated Highway System Consortium (NAHSC) in the late 1990's, the U.S. Department of Transportation moved to an incremental model. Research and development funds were directed at precursor technologies with the expectation that incremental advances would eventually lead to autonomous vehicles. Autonomy is desired by safety and congestion stakeholders since it removes the large human element from crash risk and congestion.

The first incremental advance seen widely in the market is adaptive cruise control – a variation of traditional cruise control where the car slows down to match forward vehicles rather than require driver intervention. This feature was initially sold as a "convenience feature" rather than as safety or autonomy to preemptively skirt liability. While first developed for highway speeds, most users rapidly express interest in having the capability during slow-speed congestion. The obvious reason for this request is to free up attention for secondary tasks, thus demonstrating a desire to cede more control to the autonomy. These baby steps are priming drivers to actively request full autonomy and increase market demand.

Therefore, there is real potential to promote autonomy use in HRI by paying attention to how users are introduced to autonomy in advance of system deployment. It is possible that users can be guided into utilizing more autonomy through precursor systems or simulated experiences that increase overall familiarity with autonomy. Interestingly, Hollywood's optimistic depiction of robot capability may lead to more autonomy use as compared to other autonomous systems.

B. *Perception of Having Control*

A longstanding principle within the field of human-computer interaction is the importance of user control [8]. Whether they admit it or not, most people want the ability to shape outcomes and will gravitate towards interaction designs which provide user control. The core psychology behind this is likely related to the previously mentioned bias towards manual control.

Similar evidence has been seen in HRI when comparing interaction methods for self-parking cars. A survey of regular drivers comparing manual parking, parking by taxi drivers, and autonomous parking revealed that apprehension to autonomous parking pivoted on the ability of the driver to override vehicle behavior [9].

However, humans are not always able to perceive differences between their expected level of control and reality. This is especially true for remote robots where situation awareness is compromised. A clever interaction designer can leverage this gap during initial exposure in order to generate autonomy inertia. It is probably a good idea to relax this distortion once the user is familiar with the autonomy.

C. *Perception of Capability*

There is a famous scene in *Spinal Tap* where Nigel insists his guitar amplifier is louder than others since the volume knob labels "go to eleven" [10]. This humorous example starkly illustrates how subtle interface features can lead to improper perception of capability. Similar findings within the HRI community have demonstrated that users can be easily deceived by incorrect labels of robot competence and preexisting expectations of robot fairness [11, 12].

Interaction design can also actively sway perception of autonomy capability. Parasuraman and Miller [13] showed how automation etiquette has an impact on user trust and overall human-automation performance. Their study also showed how "good automation etiquette can compensate for low automation reliability" (p. 55).

These findings imply that end user perception and preexisting expectations can be co-opted to promote higher autonomy utilization.

D. *Interaction Models*

Most users of robot systems will not be developers or receive extensive training. Therefore, we can expect significant gaps between user mental models and implemented models. Inaccurate metal models are a significant concern when dealing with semi and fully autonomous robots [14]. This suggests that complex interaction models should be abstracted into simpler forms for most end users.

This simplification process has potential for promoting autonomy use. Generic "safety" modes have promise in this respect. For example, the aforementioned self-parking car might switch into pause mode for both nearby pedestrians and transient problems in the perception sub-system. Reporting a transient sub-system failure would likely lead to an unnecessary request for manual control. Thus, concealing system faults of marginal consequence under a less alarming state abstraction may be appropriate in some scenarios.

III. Discussion

It is important to emphasize that the methods described in this paper should be applied carefully and sparingly. Research on humans and autonomy has shown that it is not advisable to define the operator's role as a byproduct of how automation is implemented [15]. Users should not be steered into autonomy just for the sake of autonomy.

Acknowledgements

Support for this work was provided by the National Science Foundation under Grant No. IIS-0905148 and IIS-1012733.

References

[1] D. McGovern, "Human Interfaces in Remote Driving," Technical Report SAND88-0562, Sandia National Laboratory, Albuquerque, NM, 1988.

[2] D. McGovern, "Experiences and Results in Teleoperation of Land Vehicles," Technical Report SAND 90-0299, Sandia National Laboratories, Albuquerque, NM, 1990.

[3] Waterloo Labs, http://www.waterloolabs.com

[4] J. Casper and R. Murphy, "Workflow Study on Human-Robot Interaction in USAR," IEEE Conference on Robotics and Automation (ICRA), 2002.

[5] J. van Erp, "Trade-offs between spatial and temporal resolution in driving unmanned ground vehicles," Human Factors and Ergonomics Society Annual Meeting, pp. 1550-1554, 1998.

[6] J. D. Lee and N. Moray, "Trust, self-confidence, and operators' adaptation to automation," International Journal of Human-Computer Studies, vol. 40, pp. 153-184, 1994.

[7] A. Steinfeld, and H.-S. Tan, Preliminary Findings for a Lane- Keeping and Collision-Warning Driver Interface for Snowplow Operations (PATH Working Paper PWP-99-6), Richmond, CA: University of California, Partners for Advanced Transit and Highways, 1999.

[8] P. W. Jordan, An Introduction to Usability. Bristol, PA: Taylor & Francis, 1998.

[9] M. Desai, K. Stubbs, A. Steinfeld, and H. Yanco, 2009. "Creating trustworthy robots: Lessons and inspirations from automated systems," Proceedings of the AISB Convention: New Frontiers in Human-Robot Interaction, 2009.

[10] Wikipedia, Up to eleven, http://en.wikipedia.org/wiki/Up_to_eleven

[11] C. Bartneck, M. Verbunt, O. Mubin, and A. Al Mahmud, "To kill a mockingbird robot," ACM/IEEE International Conference on Human-Robot Interaction (HRI). pp. 81-87, 2007.

[12] M.Vázquez, A. May, A. Steinfeld, and W.-H. Chen, "A Deceptive Robot Referee in a Multiplayer Gaming Environment," International Workshop on Collaborative Robots and Human-Robot Interaction. 2011.

[13] R. Parasuraman and C. A. Miller, "Trust and etiquette in high-criticality automated systems," Communications of the ACM vol. 47, no. 4, pp. 51-55, 2004.

[14] A. Steinfeld, "Interface lessons for fully and semi-autonomous mobile robots," IEEE International Conference on Robotics and Automation (ICRA), pp. 2752-2757, 2004.

[15] R. Parasuraman, "Human use and abuse of automation," in M. Mouloua and J. Koonce, Eds. Human-Automation Interaction, Mahwah, NJ: Lawrence Erlbaum Associates, 1997, pp. 42-47.