

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/367068966>

Social Robots in Constructive Conflicts

Chapter · January 2023

DOI: 10.3233/FAIA220640

CITATIONS

0

READS

40

4 authors, including:



Dave Miller

New Zealand Quality Research and Innovation (NZQIR)

165 PUBLICATIONS 2,441 CITATIONS

SEE PROFILE



Ben D Sawyer

University of Central Florida

78 PUBLICATIONS 1,161 CITATIONS

SEE PROFILE



James Intriligator

Tufts University

90 PUBLICATIONS 4,052 CITATIONS

SEE PROFILE

Social Robots in Constructive Conflicts

Dave B MILLER ^{a, 1}, Ben D SAWYER ^b, James INTRILIGATOR ^a, Byron BLAND ^c

^a*Tufts University*

^b*University of Central Florida*

^c*Stanford University*

Abstract. As social robots increase in capabilities and become ubiquitous parts of the environment, there will be more conflicts between humans and technological agents. Conflict is not necessarily bad: it can provide opportunities for sharing information, calibrating trust, and establishing common situation awareness—provided the conflict plays out in a reasonable manner. Social robots should be designed to act in conflict with humans gracefully and artfully, in order to use conflict as a mode of communication, limiting the adverse effects of confusion, frustration, and deadlock.

Keywords. Human-robot conflict, human-technology conflict, human-computer conflict, HRI, HCI

1. Introduction

Conflict isn't necessarily a bad thing—it can be an avenue for learning about others, calibrating trust [1–3], and developing an ongoing relationship. Social robots are nonconsciously treated as humanlike entities [4–6], and with technological agents becoming ubiquitous parts of the everyday environment as robots, autonomous, automated, and artificially intelligent (A³) systems [7] proliferate, opportunities for conflict will only increase. Some of these conflicts will result from the “growing pains” as humans (and robots) become accustomed to a landscape that includes artificial agents. Yet other conflicts will be both ongoing and inevitable: anytime you have two things (be they people, machines, systems, etc.) in an interaction, conflicts will arise, but conflicts are not by definition adverse events.

The Merriam-Webster dictionary defines conflict as “struggle resulting from incompatible or opposing needs, drives, wishes, or external or internal demands.” Other works on conflict and conflict resolution [8,9] add additional antecedents to this list, such as: disagreement over facts, goals, means employed, and aspects of the relationship between the interactants. Furthermore, conflict is often part of a dynamic relationship [10] that requires continuous construction, maintenance, and repair.

Any relationship that persists over time inevitably involves conflict, and this can be a good thing [11]. Conflict can help a relationship evolve: to stretch, grow, and deepen. In addition, conflict impacts those involved: the interactants may learn about how the other actor reacts, thinks about, and wrestles with the conflict. Both the relationship-impact and the

¹ Corresponding author, Dave B Miller, Tufts University; E-mail: dave.miller@tufts.edu.

individual-impact can be motivators towards subsequent interactions, for example, “Why do you want to take the scenic route?” Or “Why must we take the efficient route?” [12].

Human-robot conflict, just like human-human conflict, is an inevitable and essential part of a relationship between a person and a machine, and therefore should be treated as something to be designed for. By reducing the adverse consequences that can result from conflicts, these situations can be used as teaching opportunities where humans can learn the limits of the robot’s capabilities, and potentially adapt to the robot’s behavior—or a human interactant can take actions to adapt the robot to the needs of the situation. If all agents or parties in the conflict adapt, as often occurs in human-human conflicts, the conflict can be managed, transformed, or resolved [13].

The challenge is to design robots capable of navigating conflicts such that the conflicts help humans to calibrate their trust and reliance [1] and build accurate mental models [14,15] of its capabilities and limits. This can be achieved through design by engineering the system to limit the negative consequences of conflicts, and through the more complex challenge of building robots that can engage in dialogue or adaptation. These systems should not operate in such fixed ways that would make for an intractable conflict where a human must take or leave what the robot offers.

1.1. Conflict Management, Conflict Resolution, and Conflict Transformation

Conflict management aims to prevent adverse outcomes, and this should probably be the goal under time pressure, in situations with high uncertainty, or where the multiple agents’ aims are unable to be reconciled. Many human-robot conflicts are likely to fall into this regime, due to the inability of robots (at this time) to engage in dialogue. Even robots designed for social interaction or home use are still limited in their negotiating capacities [16,17], understanding natural language, and processing emotional cues.

In conflict resolution, the goal is for the parties to achieve constructive solution which places them in a better position than they could realize alone. Negotiation needs to lead to a solution at least acceptable to all parties, and which is preferable to continued conflict. How is an algorithm to determine what is an acceptable outcome, for itself (or the interests it represents) or for the humans on the other side? This may be an engineering problem of building mechanisms for engaging in different calculations than one-sided optimization, modeling human desires, creating a dialogue or negotiation engine [16], or perhaps ‘solving’ for the needs of multiple agents, but this is far from a solved engineering problem.

Conflicts can also be transformed: rather than leading to an impasse or a zero-sum game where one agent wins at the expense of the other, there is often the possibility of reframing the situation so to realize a gain for all parties, or at least a minimization of the losses that may be incurred. Again, this can be determined through discussion in human-human or intergroup conflict, but in the case of human-robot conflict, the technologies needed are still immature.

Without technological developments enhancing the capabilities of social robots, there are still some avenues open. One option is for humans to learn to deal with the limits of the robotic system, but this offers little but cold comfort if the conflict is more than just an

annoyance. The other option is for the robot to adapt its behavior—this of course requires an adaptive [18,19] or adaptable system capable of self-reconfiguration.

1.2. Conflict With vs. Conflict Through

With a true agent, you can have a direct conflict with the other interaction partner: I can criticize the barista for how they made my coffee, and they can respond. The barista can also mediate my conflict with others—they stand between me and the grower, the roaster, the engineers of the espresso machine, the maker of the grinder ... A conflict with a robot does not completely differ with this type of human-human conflict. If the robot has significant agency, one could have a conflict directly with it. Certainly, the fictional Commander Data from *Star Trek the Next Generation* is fallible in his decisions, is fully responsible for them, and can argue with other members of the crew—and they can argue with him. A humble toaster has much more limited agency, and therefore a conflict between a hungry and impatient person and the rather simple robot that toasts the bread can be better characterized as a conflict *mediated* by the robot. The conflict in this case is partly with oneself, who set the darkness; the designer of the controls; the engineers who specified the heating element and controller; and perhaps even the baker of the bread. If the toaster could talk back, it might say “hey, you set the knob too dark, I just heated up and stayed on for as long as specified!”

This paradigm of technology as conflict mediator is the dominant mode at this time, and of course needs to be considered in design. Many of the conflicts mediated by a robot result from poor usability or poor design, and thus resolution, transformation, or management can be achieved through improving the design and engineering of the robotic agent. Nonetheless, we curse the toaster, exemplifying the social perception that the conflict is with the robot itself, and demonstrating that the needed solution is in part one of social design.

2. Humans and Robots in Conflict Situations

Human-human conflict often results from multiple agents pursuing different goals or employing different means. In the film *2001: A Space Odyssey*, HAL 9000 states “I have the greatest enthusiasm for the mission”—which the crew likely shares ... But in the pursuit of mission success and maintaining congruence with inbuilt directives, HAL determines the best course of action is to kill the crew. That is certainly not what anyone wants to have happen, even if it does resolve the conflict in terms of eliminating it. This situation involves both a conflict over ends and means, and in this case, HAL takes actions that while purely logical, are inhuman and immoral. HAL’s intended aim of eliminating the conflict rather than transforming it resolving it in some other way leads to the means violation of killing the crew, rather than an alternative such as negotiating with those who created the conflict through incompatible directives.

While robots are swiftly moving into roles beyond the “dirty, dangerous, and dull” described by Takayama et al. [20], they are still needed in these roles, as they are both expendable compared to humans (at this time, as they are not yet true moral patients), and can have superhuman abilities in terms of specialization. Bomb disposal robots, which are

far from anthropomorphic and are intended to do the dangerous work of defusing or destroying explosive devices, potentially sacrificing themselves to protect human operators, have been granted medals of honor for their “sacrifice” and have been described by their human comrades as having developed personalities [21, see also 22]. Even though humans may nonconsciously treat robots as social agents with perceived capacities beyond their actual capacities [4–6], technical advances will make it possible for robots to achieve a greater degree of agency [23,24]. Even if the strength of this effect has perhaps declined [6] due to increased contact with robotic and agentic technologies, or if norms in human-human behavior are less strongly followed in human-robot interaction [25,26], we still should consider the similarities between human-human and human-robot interaction, and design accordingly. Encouraging people to espouse different norms when interacting with robots, especially as robots increase in patience and agency [27], could be disastrous for interactional success, and could train humans to act badly in both human-human and human-robot contexts.

3. Design Guidelines for Conflict Resolution, Transformation, and Management

Education and training can be valuable mechanisms to help humans to understand what robots will do in conflict situations [28], but that will not be sufficient: robots will need to be designed with conflict resolution in mind. Some of the approaches useful for human-human conflict resolution are useful for resolving human-robot conflicts, some must be adapted to the differences between humans and robots, and some of the approaches may not be at all appropriate. Considering both the features of the agents and the particulars of the conflict are necessary to cope with these factors, not that that will be easy.

From an applied perspective, how should we conceptualize and incorporate conflict into design processes? The initial response may be along the lines of *we should design systems to avoid conflicts*. We contend that the answer is definitely *not* to avoid or minimize conflicts. Instead, conflict should be seen as a frequent (and often even desirable) byproduct of good design that aims to optimize overall system performance while leveraging the unique capabilities of the parties involved [29]. Of course, conflict that leads to a roadblock, standoff, inordinate delay, or sub-optimal performance is a bad thing. The key is to develop conflict resolution behaviors that can move towards overall beneficial actions, behaviors, and decisions. Robots have been employed in mitigating and coping with human-human conflict [30,31], and work has been done studying human-robot teaming [32–34], but further work needs to be done exploring how robots can deal with direct human-robot conflicts, including human failures and robot malfunctions [35], extending current work [36–40].

When a conflict does occur, sometimes the resolution might be “the thing was right” and sometimes the optimal resolution is to let the human “win”. In other situations, a compromise might be discovered. The point being that it is not necessarily about resolving the conflict, instead it is about advancing towards desirable or at least acceptable final outcomes. The social role of a robot should not be to be a “silicon sycophant” [32]: in some circumstances, a robot must advise and inform a human [41,42], convince them [43,44], negotiate with them [24], or even act against human desires [36,40,45]. For these to be feasible actions, robots need to be designed with such capacities, and they require the communication abilities with

which to properly interact with and understand humans. To quote Nyman [46], “we need to teach AI manners,” and AI needs to teach us some as well [47].

Conflicts need not be purely adversarial. The question of how designing robots to engage in constructive conflict behavior will be one of the next great challenges for social robotics, and this will consume the energies of practitioners in many fields. Valuing inputs from the social science and humanities, as well as from engineering and mathematics will be critical for the success of robots and agents in social environments, as they will co-construct the reality of the human future. To this end, our key recommendation is design robots with conflict in mind: conflict is inevitable, if the robots are designed appropriately, communication channels are available, and humans are educated to expect robots to exert independent agency and thus interact appropriately, we can harness the combined abilities of humans and robots. Drawing from traditions such as value sensitive design [48,49] and robot ethics [50,51] will aid in creating robotic systems that help us pursue positive goals. While of course values shift and advance over time, keeping ethical principles at the top of mind will help in ensuring that the future where robots are even more deeply enmeshed in human society and our environment are doing so in ways that are socially beneficial, even when they are in conflict with us.

References

1. Lee JD, See KA. Trust in Automation: Designing for Appropriate Reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 2004 Mar 1;46(1):50–80.
2. Hoff KA, Bashir M. Trust in Automation Integrating Empirical Evidence on Factors That Influence Trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 2015 May 1;57(3):407–34.
3. Merritt SM, Ilgen DR. Not All Trust Is Created Equal: Dispositional and History-Based Trust in Human-Automation Interactions. *Hum Factors*. 2008 Apr 1;50(2):194–210.
4. Reeves B, Nass C. *The Media Equation: How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge university press; 1996.
5. Nass C, Steuer J, Tauber ER. Computers Are Social Actors. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems [Internet]*. New York, NY, USA: ACM; 1994 [cited 2015 Sep 6]. p. 72–8. (CHI '94). Available from: <http://doi.acm.org/10.1145/191666.191703>

6. Gambino A, Fox J, Ratan R. Building a Stronger CASA: Extending the Computers Are Social Actors Paradigm. *Human-Machine Communication* [Internet]. 2020 Feb 3;1(1). Available from: <https://stars.library.ucf.edu/hmc/vol1/iss1/5>
7. Sawyer BD, Miller DB, Canham M, Karwowski W. Human Factors and Ergonomics in Design of A3: Automation, Autonomy, and Artificial Intelligence. In: *Handbook of Human Factors and Ergonomics* [Internet]. 5th ed. John Wiley & Sons, Ltd; 2021 [cited 2021 Dec 20]. p. 1385–416. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119636113.ch52>
8. Mostert E. A Framework for Conflict Resolution. *Water International*. 1998 Dec 1;23(4):206–15.
9. Flemisch F, Pacaux-Lemoine MP, Frédéric V, Itoh M, Saito Y, Herzberger N, Wasser J, Grislin-Le Strugeon E, Baltzer M. Conflicts in Human-Machine Systems as an Intersection of Bio- and Technosphere: Cooperation and Interaction Patterns for Human and Machine Interference and Conflict Resolution. In 2020.
10. Pondy LR. Organizational Conflict: Concepts and Models. *Administrative Science Quarterly*. 1967;12(2):296–320.
11. Gallo A. Why We Should Be Disagreeing More at Work. *Harvard Business Review* [Internet]. 2018 Jan 3 [cited 2022 Mar 7]; Available from: <https://hbr.org/2018/01/why-we-should-be-disagreeing-more-at-work>
12. Martelaro N, Nneji VC, Ju W, Hinds P. Tell Me More: Designing HRI to Encourage More Trust, Disclosure, and Companionship. In: *The Eleventh ACM/IEEE International Conference on Human Robot Interaction* [Internet]. Piscataway, NJ, USA: IEEE Press; 2016 [cited 2016 Apr 12]. p. 181–8. (HRI '16). Available from: <http://dl.acm.org/citation.cfm?id=2906831.2906863>
13. Bland B, Powell B, Ross L. Barriers to Dispute Resolution. In: Goodman R, Jinks D, Woods AK, editors. *Understanding Social Action, Promoting Human Rights* [Internet]. Oxford University Press; 2012 [cited 2022 May 31]. p. 265–90. Available from: <https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780195371895.001.0001/acprof-9780195371895-chapter-10>
14. Johnson-Laird PN. Mental Models in Cognitive Science. *Cognitive Science*. 1980;4(1):71–115.
15. Jones NA, Ross H, Lynam T, Perez P, Leitch A. Mental Models: An Interdisciplinary Synthesis of Theory and Methods. *Ecology and Society* [Internet]. 2011 [cited 2022 Jan 30];16(1). Available from: <https://www.jstor.org/stable/26268859>

16. Krishna A. AI Learns the Art of Debate with IBM Project Debater [Internet]. IBM Blog Research. 2018 [cited 2018 Oct 22]. Available from: <http://www.ibm.comhttps://www.ibm.com/blogs/research/2018/06/ai-debate/>
17. Lin R, Kraus S. Can automated agents proficiently negotiate with humans? *Communications of the ACM*. 2010 Jan 1;53(1):78.
18. Kaber DB, Riley JM, Tan KW, Endsley MR. On the Design of Adaptive Automation for Complex Systems. *International Journal of Cognitive Ergonomics*. 2001 Mar;5(1):37–57.
19. Fereidunian A, Lehtonen M, Lesani H, Lucas C, Nordman M. Adaptive autonomy: Smart cooperative cybernetic systems for more humane automation solutions. In: 2007 IEEE International Conference on Systems, Man and Cybernetics [Internet]. Montreal, QC, Canada: IEEE; 2007 [cited 2020 Aug 30]. p. 202–7. Available from: <http://ieeexplore.ieee.org/document/4413874/>
20. Takayama L, Ju W, Nass C. Beyond dirty, dangerous and dull: What everyday people think robots should do. In: 2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI). 2008. p. 25–32.
21. Garber M. Funerals for Fallen Robots [Internet]. *The Atlantic*. 2013 [cited 2022 Mar 7]. Available from: <https://www.theatlantic.com/technology/archive/2013/09/funerals-for-fallen-robots/279861/>
22. Barley SR. The Social Construction of a Machine: Ritual, Superstition, Magical Thinking and other Pragmatic Responses to Running a CT Scanner. In: Lock M, Gordon D, editors. *Biomedicine Examined* [Internet]. Springer Netherlands; 1988 [cited 2016 Aug 19]. p. 497–539. (Culture, Illness and Healing). Available from: http://link.springer.com/chapter/10.1007/978-94-009-2725-4_19
23. Miller D. AgentSmith: Exploring Agentic Systems. In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* [Internet]. ACM Press; 2016 [cited 2018 Jul 23]. p. 234–8. Available from: <http://dl.acm.org/citation.cfm?doid=2851581.2859025>
24. Lee M, Lucas G, Gratch J. Comparing mind perception in strategic exchanges: human-agent negotiation, dictator and ultimatum games. *J Multimodal User Interfaces*. 2021 Jun 1;15(2):201–14.
25. Bartneck C, C R, R M, I D. Robot Abuse - A Limitation of the Media Equation. 2005 [cited 2022 May 31]; Available from: <https://ir.canterbury.ac.nz/handle/10092/16925>

26. Rosalia C, Menges R, Deckers I, Bartneck C. Cruelty towards robots. 2005 [cited 2022 May 31]; Available from: <https://ir.canterbury.ac.nz/handle/10092/16847>
27. Lee M. Interactional Morality: Technology as Our Moral Mirror [Phd Thesis 1 (Research TU/e / Graduation TU/e)]. [Eindhoven]: Technische Universiteit Eindhoven; 2021.
28. Miller DB. Simulated Extreme Experiential Training for Engaging with Automation. In: Proceedings of AutomationXP22: Engaging with Automation [Internet]. IOS Press; 2022 [cited 2022 Jun 1]. p. 6. Available from: https://matthiasbaldauf.com/automationxp22/papers/AutomationXP22_paper7.pdf
29. Sheridan TB, Verplank WL. Human and Computer Control of Undersea Teleoperators. 1978 Jul.
30. Jung MF, Martelaro N, Hinds PJ. Using Robots to Moderate Team Conflict: The Case of Repairing Violations. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction [Internet]. ACM; 2015 [cited 2015 Sep 13]. p. 229–36. Available from: http://www.researchgate.net/profile/Malte_Jung/publication/275522525_Using_Robots_to_Moderate_Team_Conflict_The_Case_of_Repairing_Violations/links/553e9a620cf20184050f89be.pdf
31. Druckman D, Adrian L, Damholdt MF, Filzmoser M, Koszegi ST, Seibt J, Vestergaard C. Who is Best at Mediating a Social Conflict? Comparing Robots, Screens and Humans. *Group Decis Negot*. 2021 Apr;30(2):395–426.
32. Nass C, Fogg BJ, Moon Y. Can computers be teammates? *International Journal of Human-Computer Studies*. 1996 Dec;45(6):669–78.
33. Sangseok Y, Lionel R. Teaming up with Robots: An IMOI (Inputs-Mediators-Outputs-Inputs) Framework of Human-Robot Teamwork. *Int J Robot Eng* [Internet]. 2017 Dec 31 [cited 2022 Jun 1];2(1). Available from: <https://www.vibgyorpublishers.org/content/ijre/fulltext.php?aid=ijre-2-003>
34. Cuevas HM, Fiore SM, Caldwell BS, Strater L. Augmenting Team Cognition in Human-Automation Teams Performing in Complex Operational Environments. *Aviation, Space, and Environmental Medicine*. 2007 May 1;78(5):B63–70.
35. Turkle S. *Evocative Objects: Things We Think With*. MIT Press; 2011. 397 p.
36. Takayama L, Groom V, Nass C. I'm Sorry, Dave: I'm Afraid I Won't Do That: Social Aspects of Human-Agent Conflict. In: Proceedings of the SIGCHI Conference on

Human Factors in Computing Systems [Internet]. ACM; 2009 [cited 2015 Sep 16]. p. 2099–108. Available from: <http://dl.acm.org/citation.cfm?id=1519021>

37. Sasabuchi K, Ikeuchi K, Inaba M. Agreeing to Interact: Understanding Interaction as Human-Robot Goal Conflicts. In: Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction [Internet]. New York, NY, USA: Association for Computing Machinery; 2018 [cited 2022 Jun 1]. p. 21–8. (HRI '18). Available from: <https://doi.org/10.1145/3173386.3173390>
38. Dehais F, Causse M, Tremblay S. Mitigation of Conflicts with Automation: Use of Cognitive Countermeasures. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. 2011 Oct;53(5):448–60.
39. Dehais F, Causse M, Vachon F, Tremblay S. Cognitive conflict in human–automation interactions: A psychophysiological study. *Applied Ergonomics*. 2012 May 1;43(3):588–95.
40. Miller DB. Human-computer conflicts in partially-automated driving [Internet] [Dissertation]. [Palo Alto, CA, USA]: Stanford University; 2019 [cited 2020 Jul 27]. Available from: <https://searchworks.stanford.edu/view/13426012>
41. St. Amant R. Resolving conflicts in collaborative human-computer interaction [Internet]. AAAI; 1999 [cited 2017 Nov 29] p. 58–9. Report No.: WS-99-08. Available from: <https://www.aaai.org/Papers/Workshops/1999/WS-99-08/WS99-08-009.pdf>
42. Koo J, Kwac J, Ju W, Steinert M, Leifer L, Nass C. Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *Int J Interact Des Manuf*. 2014 Apr 30;9(4):269–75.
43. Kaptein M, Eckles D. Selecting Effective Means to Any End: Futures and Ethics of Persuasion Profiling. In: Ploug T, Hasle P, Oinas-Kukkonen H, editors. *Persuasive Technology* [Internet]. Springer Berlin Heidelberg; 2010 [cited 2016 Jan 2]. p. 82–93. (Lecture Notes in Computer Science). Available from: http://link.springer.com/chapter/10.1007/978-3-642-13226-1_10
44. Kaptein M, Markopoulos P, Ruyter B de, Aarts E. Can You Be Persuaded? Individual Differences in Susceptibility to Persuasion. In: Gross T, Gulliksen J, Kotzé P, Oestreicher L, Palanque P, Prates RO, Winckler M, editors. *Human-Computer Interaction – INTERACT 2009* [Internet]. Springer Berlin Heidelberg; 2009 [cited 2016 Jan 2]. p. 115–8. (Lecture Notes in Computer Science). Available from: http://link.springer.com/chapter/10.1007/978-3-642-03655-2_13
45. Johns M, Mok B, Sirkin DM, Gowda NM, Smith CA, Talamonti Jr WJ, Ju W. Exploring Shared Control in Automated Driving. In: *The Eleventh ACM/IEEE*

International Conference on Human Robot Interaction [Internet]. Piscataway, NJ, USA: IEEE Press; 2016 [cited 2018 Aug 15]. p. 91–8. (HRI '16). Available from: <http://dl.acm.org/citation.cfm?id=2906831.2906847>

46. Nyman G. Teaching manners to AI: Internet of good Behavior [Internet]. Gote Nyman's (gotepoem) Blog. 2018 [cited 2022 Jun 1]. Available from: <https://gotepoem.wordpress.com/2018/02/05/teaching-manners-to-ai-internet-of-good-behavior/>
47. Nyman G. Internet of Behaviors (IOB) With a Human Touch. Gote Nyman; 2022. 204 p.
48. Friedman B. Value-sensitive Design. *interactions*. 1996 Dec;3(6):16–23.
49. Friedman B, Kahn P, Borning A. Value sensitive design: Theory and methods. University of Washington technical report. 2002;02–12.
50. Deng B. Machine ethics: The robot's dilemma. *Nature*. 2015 Jul 2;523(7558):24–6.
51. Lin P, Abney K, Bekey GA. Robot Ethics: The Ethical and Social Implications of Robotics [Internet]. The MIT Press; 2012 [cited 2015 Oct 9]. Available from: http://books.google.com/books?hl=en&lr=&id=oBb-lt3l4oYC&oi=fnd&pg=PR5&dq=info:ADj64LLD-hcJ:scholar.google.com&ots=yueOx-zX1l&sig=DXthq_nCwx11k0zF70YLJlCuQs