

SEMINAR: ROBOTERETHIK

WS 2016/17

Dozent: Radin Dardashti	Zeit: 10.00 - 12.00
Email: radin.dardashti@philos.uni-hannover.de	Ort: 1146 - B410

Sprechstunde: Do., 15.00 - 16.00 Uhr oder nach Vereinbarung

Beschreibung

In der modernen Robotik sind Roboter zunehmend autonom handelnde Agenten statt reine Werkzeuge. Als handelnde Agenten stellt sich die Frage, welche Moral geeignet ist das Verhalten von Robotern zu regeln. Die weite Anwendung von Robotern (wie z. B. im Krieg, in der Pflege- oder Sexindustrie) wirft hier sehr unterschiedliche Fragen auf. In diesem Seminar behandeln wir u.a. einfürend die ethischen Fragen, die auftreten in Bezug auf (i) die Anwendung von Robotern in diesen unterschiedlichen Bereichen, (ii) Fragen der Verantwortung, sowie (iii) die Wirkung von Robotern auf die Gesellschaft allgemein.

Literatur:

- Patrick Lin, Keith Abney, and George A. Bekey, eds. (2012): Robot Ethics: The Ethical and Social Implications of Robotics. MIT Press.
- Ronald C. Arkin (2009): Governing Lethal Behavior in Autonomous Robots. CRC Press.
- Weitere Artikel werden zur Verfügung gestellt

Programm

1. Einführung und Vorbesprechung (27.10.2016)

2. Roboter und Ethik (03.11.2016)

Obligatorisch:

- George A. Bekey (2012): "Current Trends in Robotics: Technology and Ethics." in Lin et al. (p.17–34).
- Keith Abney (2012): "Robotics, Ethical Theory, and Metaethics: A Guide for the Perplexed." in Lin et al. (p. 35–47,50–52).

Zusätzliche Literatur:

- Patrick Lin (2012): "Introduction to Robot Ethics" in Lin et al. (p. 3–15)
- Peter M. Asaro (2009): "What Should We Want From a Robot Ethic?" in R. Capurro and M. Nagelborg (eds.): Ethics and Robotics. Heidelberg: AKA Verlag (p. 1–10).

3. Roboter-Design I: Moral programmieren? (10.11.2016)

Obligatorisch:

- Colin Allen & Wendell Wallach (2012): "Moral Machines: Contradiction in Terms or Abdication of Human Responsibility?" in Lin et al. (p. 55–68).
- Selmer Bringsjord & Joshua Taylor (2012): "The Divine-Command Approach to Robot Ethics" in Lin et al. (p. 85–108).

Zusätzliche Literatur:

- Wallach, W., & Allen, C. (2008): *Moral Machines: Teaching Robots Right from Wrong*. Oxford University Press (Ch.6–9, p. 83–138).

4. Roboter-Design II: Das Testen von moralischen Robotern (17.11.2016)

Obligatorisch:

- Allen, C., Varner, G., & Zinser, J. (2000): "Prolegomena to any Future Artificial Moral Agent". *Journal of Experimental & Theoretical Artificial Intelligence*, 12(3), (p. 251–261).
- Gerdes, A., & Øhrstrøm, P. (2013): "Preliminary Reflections on a Moral Turing test", in *Proceedings of ETHICOMP* (p. 167–174).
- Arnold, T., & Scheutz, M. (2016). "Against the Moral Turing Test: Accountable Design and the Moral Reasoning of Autonomous Systems". *Ethics and Information Technology*, 18(2), (p. 103–115).

Zusätzliche Literatur:

- Wallach, W., & Allen, C. (2008): *Moral Machines: Teaching Robots Right from Wrong*. Oxford University Press (Ch.4 "Can (Ro)bots Really be Moral?", p. 55–72).

5. Cyborgs (24.11.2016)

Tag der Philosophischen Fakultät, daher findet das Seminar **um 18 Uhr in Raum B313.**

Obligatorisch:

- Warwick, K. (2003): "Cyborg morals, cyborg values, cyborg ethics". *Ethics and information technology*, 5(3), (p. 131–137).
- Hildebrandt, M., & Anrig, B. (2012). *Ethical implications of ICT implants*. In *Human ICT implants: technical, legal and ethical considerations*, TMC Asser Press, (p. 135–158).

Zusätzliche Literatur:

- Warwick, K. (2015). *The disappearing human-machine divide*. In *Beyond Artificial Intelligence*, Springer, (p. 1–10).

6. Roboter im Krieg I: Krieg und Moral (01.12.2016)

Obligatorisch:

- Lazar, Seth (2016): "War", The Stanford Encyclopedia of Philosophy (Winter 2016 Edition), Edward N. Zalta (ed.). URL = <http://plato.stanford.edu/archives/win2016/entries/war/>.

Zusätzliche Literatur:

- Ronald C. Arkin (2009): "Relevant Laws of War". In *Governing Lethal Behavior in Autonomous Robots*. CRC Press, (p. 225–241).
- Fiala, Andrew (2014): "Pacifism", The Stanford Encyclopedia of Philosophy (Winter 2014 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2014/entries/pacifism/>.

7. Roboter im Krieg II: Autonome Systeme in der Militärtechnik (08.12.2016)

Obligatorisch:

- Arkin, R. C. (2010): "The Case for Ethical Autonomy in Unmanned Systems". *Journal of Military Ethics*, 9(4), (p. 332–341).
- Robert Sparrow (2016): "Robots and Respect: Assessing the Case Against Autonomous Weapon Systems". *Ethics & International Affairs*, 30, no. 1, (p.93–116)

Zusätzliche Literatur:

- Ronald C. Arkin (2009): *Governing Lethal Behavior in Autonomous Robots*. CRC Press, (Ch.3 p. 29–36; Ch.4 p. 37–48).
- Human Rights Watch (2012): "Losing Humanity: The Case against Killer Robots". URL = <https://www.hrw.org/report/2012/11/19/losing-humanity/case-against-killer-robots>.

8. Roboter in der Pflege (15.12.2016)

Obligatorisch:

- Jason Borenstein and Yvette Pearson (2012): "Robot Caregivers: Ethical Issues across the Human Lifespan" in Lin et al. (p. 251–266).
- Noel Sharkey and Amanda Sharkey (2012): "The Rights and Wrongs of Robot Care", in Lin et al. (p. 267–282).

Zusätzliche Literatur:

- Steve Petersen (2012): "Designing People to Serve", in Lin et al. (p. 283–298)

9. Die KI-Singularität (22.12.2016)

Obligatorisch:

- Muehlhauser, L., & Helm, L. (2012). “The singularity and machine ethics”, in Singularity Hypotheses, Springer Berlin Heidelberg (p. 101-126).

Zusätzliche Literatur:

- Allen, P., & Greaves, M. (2011). “The singularity isnt near.” MIT Technology Review.
- Kurzweil, R. (2011). “Kurzweil Responds: Dont Underestimate the Singularity.” MIT Technology Review.
- Walsh, Toby (2016). “The Singularity May Never Be Near”. arXiv preprint:1602.06462.

10. Roboter in der Sexindustrie(12.01.2017)

Obligatorisch:

- Mathias Scheutz (2012): “The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots”, in Lin et al. (p 205–222)
- David Levy (2012): “The Ethics of Robot Prostitutes’, in Lin et al. (p 223–232)
- Kathleen Richardson (2016). “Sex Robot Matters: Slavery, the Prostituted, and the Rights of Machines”. IEEE Technology and Society Magazine, 35(2), (p. 46–53).

Zusätzliche Literatur:

- Blay Whitby (2012): “Do You Want a Robot Lover? The Ethics of Caring Technologies”, in Lin et al. (p. 233–248)

11. Roboter im Verkehr (19.01.2017)

Obligatorisch:

- Patrick Lin (2016): “Why Ethics Matters for Autonomous Cars”. In C. Gerdes et al. (eds.). Autonomous Driving, Springer, (pp. 69-85).
- Gerdes, J. C., & Thornton, S. M. (2016): “Implementable Ethics for Autonomous Vehicles”. In C. Gerdes et al. (eds.). Autonomous Driving, Springer, (p. 87-102).

Zusätzliche Literatur:

- Hevelke, A., & Nida-Rümelin, J. (2015): “Responsibility for crashes of autonomous vehicles: an ethical analysis”. Science and Engineering Ethics, 21(3), (p. 619-630).
- Patrick Lin (2013): “The Ethics of Saving Lives With Autonomous Cars Are Far Murkier Than You Think”, Wired. URL = <https://www.wired.com/2013/07/the-surprising-ethics-of-robot-cars/>

12. Roboter und Ihre Rechte (26.01.2017)

Obligatorisch:

- Robert Sparrow (2012): “Can Machines Be People? Reflections on the Turing Triage Test”, in Lin et al. (p. 301–316).
- Singer P, Sagan A (2009): Do humanoid robots deserve to have rights? The Japan Times, 17 Dec 2009.
- Darling Kate (2012): “Extending Legal Rights to Social Robots”.
- Bryson J. J. (2008): “Robots Should be Slaves”. In: Wilks Y. (Ed), Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issue. John Benjamins, Amsterdam.

Zusätzliche Literatur:

- Anthony F. Beavers (2012): “Moral Machines and the Threat of Ethical Nihilism”, in Lin et al. (p. 333–344)
- Kevin Warwick (2012): “Robots with Biological Brains”, in Lin et al. (p. 317–332).

13. Roboterethik in Japan (02.02.2017)

Obligatorisch:

- Kitano, N. (2006). “ ‘Rinri’: An Incitement towards the Existence of Robots in Japanese Society”. Ethics in Robotics, 6, (p. 78–83).
- Makoto Nakada (2010): ”Different Discussions on Roboethics and Information Ethics based on Different Cultural Contexts (Ba)”. In F. Sudweeks, H. Hrachovec and C. Ess (eds.). Proceedings Cultural Attitudes Towards Communication and Technology, Murdoch University, Australia, (p. 300-314).

Zusätzliche Literatur:

- Robert E. Carter (2011): ”Japanese Ethics”. In W. Edelglass and J. L. Garfield, The Oxford Handbook of World Philosophy, (p. 303–316).
- Kitano N (2007): “Animism, Rinri, modernization: the base of japanese robotics”, Workshop on Roboethics: ICRA07.