

Machine Ethics

Edited by

Michael Anderson

University of Hartford

Susan Leigh Anderson

University of Connecticut



CAMBRIDGE
UNIVERSITY PRESS

Contents

<i>General Introduction</i>	page 1
PART I THE NATURE OF MACHINE ETHICS	
<i>Introduction</i>	7
1. The Nature, Importance, and Difficulty of Machine Ethics James H. Moor	13
2. Machine Metaethics Susan Leigh Anderson	21
3. Ethics for Machines J. Storrs Hall	28
PART II THE IMPORTANCE OF MACHINE ETHICS	
<i>Introduction</i>	47
4. Why Machine Ethics? Colin Allen, Wendell Wallach, and Iva Smit	51
5. Authenticity in the Age of Digital Companions Sherry Turkle	62
PART III ISSUES CONCERNING MACHINE ETHICS	
<i>Introduction</i>	79
6. What Matters to a Machine? Drew McDermott	88
7. Machine Ethics and the Idea of a More-Than-Human Moral World Steve Torrance	115

8.	On Computable Morality: An Examination of Machines as Moral Advisors Blay Whitby	138
9.	When Is a Robot a Moral Agent? John P. Sullins	151
10.	Philosophical Concerns with Machine Ethics Susan Leigh Anderson	162
11.	Computer Systems: Moral Entities but Not Moral Agents Deborah G. Johnson	168
12.	On the Morality of Artificial Agents Luciano Floridi	184
13.	Legal Rights for Machines: Some Fundamental Concepts David J. Calverley	213
PART IV APPROACHES TO MACHINE ETHICS		
	<i>Introduction</i>	231
	<i>a. Overview</i>	
14.	Towards the Ethical Robot James Gips	244
	<i>b. Asimov's Laws</i>	
15.	Asimov's Laws of Robotics: Implications for Information Technology Roger Clarke	254
16.	The Unacceptability of Asimov's Three Laws of Robotics as a Basis for Machine Ethics Susan Leigh Anderson	285
	<i>c. Artificial Intelligence Approaches</i>	
17.	Computational Models of Ethical Reasoning: Challenges, Initial Steps, and Future Directions Bruce M. McLaren	297
18.	Computational Neural Modeling and the Philosophy of Ethics: Reflections on the Particularism-Generalism Debate Marcello Guarini	316

19. Architectures and Ethics for Robots: Constraint Satisfaction as a Unitary Design Framework Alan K. Mackworth	335
20. Piagetian Roboethics via Category Theory: Moving beyond Mere Formal Operations to Engineer Robots Whose Decisions Are Guaranteed to be Ethically Correct Selmer Bringsjord, Joshua Taylor, Bram van Heuveln, Konstantine Arkoudas, Micah Clark and Ralph Wojtowicz	361
21. Ethical Protocols Design Matteo Turilli	375
22. Modeling Morality with Prospective Logic Luís Moniz Pereira and Ari Saptawijaya	398
<i>d. Psychological/Sociological Approaches</i>	
23. An Integrated Reasoning Approach to Moral Decision Making Morteza Dehghani, Ken Forbus, Emmett Tomai and Matthew Klenk	422
24. Prototyping N-Reasons: A Computer Mediated Ethics Machine Peter Danielson	442
<i>e. Philosophical Approaches</i>	
25. There Is No "I" in "Robot": Robots and Utilitarianism Christopher Grau	451
26. Prospects for a Kantian Machine Thomas M. Powers	464
27. A Prima Facie Duty Approach to Machine Ethics: Machine Learning of Features of Ethical Dilemmas, Prima Facie Duties, and Decision Principles through a Dialogue with Ethicists Susan Leigh Anderson and Michael Anderson	476
PART V VISIONS FOR MACHINE ETHICS	
<i>Introduction</i>	495
28. What Can AI Do for Ethics? Helen Seville and Debora G. Field	499
29. Ethics for Self-Improving Machines J. Storrs Hall	512

30. How Machines Might Help Us Achieve Breakthroughs
in Ethical Theory and Inspire Us to Behave Better 524
Susan Leigh Anderson
31. *Homo Sapiens 2.0*: Building the Better Robots of Our Nature 531
Eric Dietrich