

Eine bearbeitete Version davon erschien in *der AISB Quarterly* Issue 133 (Frühjahr 2012), Seite 14-15.

Die Entstehung der EPSRC-Prinzipien der Robotik

Joanna J. Bryson

Artificial Models of Natural Intelligence (AmonI) Group
Department of Computer Science, University of Bath
<http://www.cs.bath.ac.uk/~jjb>

5. Dezember 2011

Ende 2010 lud mich das EPSRC unerwartet zu einem Treffen im New Forest zum Thema Ethik und Roboter ein. Ich schreibe seit 1996 gelegentlich Artikel zu diesem Thema, als Reaktion auf meine Erfahrung bei Cog (einem humanoiden Roboterprojekt) und der Erkenntnis, wie bereitwillig und sogar eindringlich Menschen moralische Verpflichtungen einem völlig nicht funktionsfähigen (im Jahr 1993) aber vage zugeschrieben humanoider Roboter. Seit 1998 betreibe ich auch eine Webseite zu diesem Thema. Trotzdem war es das erste Mal, dass eine Regierungsbehörde auf mich zukam, und natürlich sagte ich ja.

Bei dem Treffen handelte es sich um ein dreitägiges Offsite-Treffen unter der (fachmännischen) Leitung der Journalistin Vivienne Parry. Zu den weiteren Teilnehmern gehörten neben mir auch Margaret Boden, Darwin Caldwell, Kerstin Dautenhahn, Paula Duxbury, Lilian Edwards, Ann Grand, Hazel Grian, Sarah Kember, Stephen Kemp, Paul Newman, Geoff Pegman, Andrew Rose, Tom Rodden, Tom Sorell und Mick Wallis, Shearer West, Blay Whitby und Alan Winfield sowie kompetente Unterstützung von Ian Baldwin, Denise Dabbs und Paul O'Dowd. Die Teilnehmer kamen überwiegend aus der Robotik, aber auch aus den Geistes-, Rechts- und Sozialwissenschaften. Wir waren hauptsächlich im akademischen Bereich beschäftigt, aber auch in der Industrie und bei Forschungsräten (einschließlich des damaligen Leiters des AHRC, Shearer West).

Ich war überrascht, dass sich die EPSRC so lange für so viele Menschen in einem so schönen Hotel zu diesem Thema eingesetzt hat, aber sie haben den Gegenstand ihrer Besorgnis sehr früh deutlich gemacht. Das EPSRC betrachtet die Robotik als eine entscheidende Technologie für das Vereinigte Königreich und möchte nicht, dass ihr das gleiche Schicksal widerfährt wie anderen „futuristischen“ Technologien, nämlich einer öffentlichen Abneigung, die an Hysterie grenzt und nicht mehr durch Maßnahmen bewältigt werden kann wissenschaftliche Beurteilung. Das EPSRC möchte die Roboterethik von Anfang an berücksichtigen, um sowohl die Sicherheit als auch die Akzeptanz von Robotertechnologien zu gewährleisten.

Es zeigte sich fast sofort, dass es ihnen gelungen war, eine sehr pragmatische und sozial engagierte Expertengruppe auszuwählen. Die Gruppe vertrat eine sehr klare Linie hinsichtlich der moralischen und ethischen Rolle der Robotik, und ich würde nicht sagen, dass sie bei typischen AISB-Treffen die dominierende Rolle spielt.

Am letzten ganzen Tag waren die Internet-Rechtsprofessorin Lilian Edwards und ich in einem

Kleines Breakout-Treffen in einer gemeinsamen Sitzung mit dem Ziel, Ergebnisse als Ergebnisse für das Meeting zu entwerfen. Wir haben beschlossen, ein „echtes“ Gesetzeswerk für Roboter zu erlassen. Lillian wollte, dass sie die Gesetze von Asimov klar befolgen, aber korrigieren, während ich daran interessiert war, einige einzubeziehen, die ich bereits beim Schreiben von A Proposal for the Humanoid Agent-builders League (HAL) (Bryson, 2000) entwickelt hatte. Am Ende haben wir uns für fünf entschieden, von denen die ersten drei Asimov widerspiegeln und auf die Anliegen der Gruppe eingehen. Die gesamte Gruppe verfeinerte dann nicht nur unsere „Gesetze“, sondern auch deren gewöhnliche Sprachversionen und entwickelte eine weitere Liste von Konzepten, die wir Ihnen, unseren Kollegen, mitteilen möchten.

Die vollständige Version dieser Dokumente kann jetzt durch Googeln der EPSRC-Grundsätze für Robotik gefunden werden. Seit April 2011 sind sie zur EPSRC-Richtlinie geworden. Die Grundprinzipien sind in der Box wiedergegeben. Nachfolgend finden Sie sieben hochrangige Ideen, die die Gruppe Ihnen, unseren Kollegen, mitteilen möchte. Ausführliche Erläuterungen finden Sie auf der Website, die Highlights habe ich hier aufgeführt.

1. Wir glauben, dass Roboter das Potenzial haben, einen enormen positiven Einfluss auf die Gesellschaft zu haben. Wir wollen eine verantwortungsvolle Roboterforschung fördern. **Wir sind kein Haufen Idioten, die das wahre Potenzial der KI „nicht verstehen“. Wir sind besorgte Fachleute, die wirklich wollen, dass KI funktioniert und Roboter Wirklichkeit werden.**
2. Schlechte Praktiken schaden uns allen. **Wir können die Situation nicht ignorieren, wenn einige unserer Kollegen tun Dinge, die uns alle schlecht aussehen lassen.**
3. Die Auseinandersetzung mit offensichtlichen Bedenken der Öffentlichkeit wird uns allen dabei helfen, Fortschritte zu machen.
4. Es ist wichtig zu zeigen, dass wir uns als Robotiker dazu verpflichten die bestmöglichen Praxisstandards.
5. Um den Kontext und die Konsequenzen unserer Forschung zu verstehen, sollten wir mit Experten aus anderen Disziplinen zusammenarbeiten, darunter Sozialwissenschaften, Recht, Philosophie und Kunst. **Wir waren alle beeindruckt, wie viel wir von diesem multidisziplinären Arbeitsteam gelernt haben.**
6. Wir sollten über die Ethik der Transparenz nachdenken: Gibt es Grenzen dessen, was offen zugänglich sein sollte? **Alle Teilnehmer des Treffens setzten sich für Open-Source-Software-Lösungen und -Ansätze ein, aber wir kamen zu der Erkenntnis, dass wir bei Robotern und KI im Allgemeinen die Pflicht haben, dafür zu sorgen, dass sich nicht jeder „Skript-Kiddy“ in eine Software hacken kann System, das Informationen oder Erinnerungen über das Privatleben von Menschen enthält.**
7. Wenn wir fehlerhafte Berichte in der Presse sehen, verpflichten wir uns, uns die Zeit zu nehmen, die berichtenden Journalisten zu kontaktieren. **Die meisten Wissenschaftsreporter möchten wirklich nicht albern aussehen, indem sie über einen „Experten“ berichten, der sich als Selbstdarsteller oder Sensationsjournalist herausstellt. Ein leises Wort oder eine E-Mail kann oft die Hysterie dämpfen, die durch unverantwortliche Aussagen entsteht.**

Ich möchte dem EPSRC und auch unseren Kollegen danken, die sich für dieses Treffen eingesetzt haben. Zwei dieser Letzteren waren Alan Winfield und Tom Rodden. Persönlich bin ich für meinen Beruf und meine Nation äußerst stolz und glücklich, dass das Vereinigte Königreich nun über eine offizielle Reihe von Robotikprinzipien verfügt, die sich mit so wichtigen Themen befassen. Aber wir sind nur ein Land, und es gibt noch viel zu tun und zu befürworten, um sicherzustellen, dass intelligente Robotik in unserer Gesellschaft angemessen eingesetzt wird.

Kasten!

1. Roboter sind vielseitig einsetzbare Werkzeuge. Roboter sollten nicht ausschließlich oder in erster Linie dazu konzipiert sein, Menschen zu töten oder ihnen Schaden zuzufügen, außer im Interesse der nationalen Sicherheit. **Obwohl wir anerkennen, dass selbst tote Fische von kreativen Menschen als Waffen verwendet werden können, war es uns ein Anliegen, die Entwicklung und den Einsatz autonomer Roboter als Waffen zu verbieten. Auch wenn wir pragmatisch anerkannt haben, dass dies im militärischen Kontext bereits geschieht, möchten wir nicht, dass diese in anderen Kontexten genutzt werden.**
2. Menschen, nicht Roboter, sind verantwortliche Akteure. Roboter sollten so weit wie möglich so konzipiert und betrieben werden, dass sie den bestehenden Gesetzen und Grundrechten und -freiheiten, einschließlich der Privatsphäre, entsprechen. **Wir waren sehr besorgt, dass jede Diskussion über „Roboterethik“ dazu führen könnte, dass Einzelpersonen, Unternehmen oder Regierungen ihre eigene Verantwortung als Hersteller, Käufer und Betreiber von Robotern aufgeben. Wir hatten das Gefühl, dass die Konsequenzen dieser Sorge den „Vorteil“ an der Freude, etwas zu erschaffen, das die Gesellschaft als empfindungsfähig und verantwortungsbewusst ansieht, bei weitem überwiegen. Wir wussten, dass dies das Gesetz war, das einige unserer Kollegen in der AISB am meisten abwehren würde – folglich (mit Dav** Ich veranstalte auf der AISB 2012 ein Symposium, um zu prüfen, ob dies eine vernünftige Regel ist. Das Symposium trägt den Titel „Die Maschinenfrage: KI, Ethik und moralische Verantwortung“.
3. Roboter sind Produkte. Sie sollten mithilfe von Prozessen entworfen werden, die ihre Sicherheit gewährleisten. **Dieses Prinzip erinnert uns erneut daran, dass es bei uns als Roboterentwicklern und nicht bei den Robotern selbst liegt, dafür zu sorgen, dass Roboter keinen Schaden anrichten.**
4. Roboter sind hergestellte Artefakte. Sie sollten nicht so konzipiert sein, dass sie gefährdete Benutzer ausnutzen; Stattdessen sollte ihre Maschinennatur transparent sein. **Es war am schwierigsten, sich auf die Formulierung dieser Regel zu einigen. Die Idee ist, dass jeder, der einen Roboter besitzt, wissen sollte, dass er nicht „lebt“ oder „leidet“, doch die Täuschung des Lebens und emotionale Bindung ist genau das Ziel vieler Therapie- oder Spielzeugroboter. Wir kamen zu dem Schluss, dass, solange die verantwortliche Person, die einen Roboter kauft, auch nur indirekt (z. B. Internetdokumentation) Zugang zu Informationen über die Funktionsweise seines „Geistes“ hat, dies für eine ausreichend informierte Bevölkerung sorgen würde, um zu verhindern, dass Menschen ausgebeutet werden.**

5. Die Person, die die rechtliche Verantwortung für einen Roboter trägt, sollte zugewiesen werden.
Es sollte immer möglich sein, herauszufinden, wem ein Roboter gehört, so wie es immer möglich ist, herauszufinden, wem ein Auto gehört. Dies erinnert uns erneut daran, dass, was auch immer ein Roboter tut, ein Mensch oder eine menschliche Institution (z. B. ein Unternehmen) für seine Handlungen haftbar ist.

Siehe auch (Hinweis an den Herausgeber, **bitte kombinieren Sie die Webseiten und die Referenzen in einer Liste für insgesamt fünf Elemente**):

- Die EPSRC-Website „Grundsätze der Robotik“.
- Die Maschinenfrage: KI, Ethik und moralische Verantwortung.
- Ethik: KI, Roboter und Gesellschaft.
- (Wilks, 2010)

Verweise

Bryson, JJ (2000). Ein Vorschlag für die Humanoid Agent-Builders League (HAL). In Barnden, J., Herausgeber, AISB'00 Symposium on Artificial Intelligence, Ethics and (Quasi-)Human Rights, Seiten 1–6.

Wilks, Y., Herausgeber (2010). Enge Auseinandersetzung mit künstlichen Begleitern: Wichtige soziale, psychologische, ethische und gestalterische Fragen. John Benjamins, Amsterdam.