Appendix 2

PHILOSOPHY OF TECHNICAL ARTEFACTS

Joint DUT & TU/e Research Programme 2005 - 2010

Programme leaders Peter Kroes and Anthonie Meijers

Introduction

The research programme *Philosophy of Technical Artefacts* focuses on philosophical problems of modern technology and the engineering sciences. It addresses such questions as: How should we conceptualize technical artefacts and their functions? How do norms and values, for example with respect to safety and sustainability, inform design processes? To what extent can agency and responsibility be attributed to artefacts and systems? How to characterize means-ends reasoning as part of technological rationality? What kinds of moral problems are engineers confronted with in their professional practice and how should they deal with these problems? How can the notion of responsibility be upheld given that technology is so overwhelmingly a product of a collective effort?

The programme aims at developing an up to now sparsely inhabited area in the philosophical landscape: an analytically oriented philosophy of technical artefacts. It addresses epistemological, ethical, ontological and conceptual issues related to the design, development, and implementation of technical artefacts. It is not an exaggeration to say that little attention has been paid to these issues in mainstream philosophy. There are a few promising exceptions (e.g. Preston, 1998; Dipert, 1993), but most work on artefacts is motivated by considerations from other disciplines, such as the philosophy of biology or aesthetics. Hardly any attention is given to technical artefacts as a topic for study in its own right. In applied ethics there are many studies on all sorts of technologies, e.g. computer technology, biotechnology, or nuclear power, but the focus has been on the effects of technology and not on the artefacts themselves and their design. Similarly, in the philosophy of technology ethical questions related to technology have been addressed, but the emphasis here has been on the overall impact of technology as such and on issues in the philosophy of culture, inspired by continental philosophers (see Mitcham, 1994). The design and development phase of technology in the engineering sciences has to a large extent been ignored, while the analysis remained on a rather global level.

Remarkably, philosophers of science are also not known for showing an interest in the technological or engineering sciences. Physics, biology and the social sciences have been their primary objects of study. This focus is partly due to the widespread belief that technology is basically applied science and not very interesting from a philosophical point of view. When technology is taken into consideration it is largely because of its role *in* science (instruments, experimentation) and not because it constitutes an interesting subject in its own right.

In this programme we take up fundamental issues in the philosophy of technical artefacts and the engineering sciences. The main focus is on philosophical problems related to the design, development and implementation of technical artefacts. The perspective taken is largely analytic and conceptual, while being empirically informed about the issues at the same time (see Kroes & Meijers, 2000 and forthcoming).

Main themes

The programme starts from key notions in technology and the engineering sciences, such as 'design', 'artefact', 'system', 'value', 'technical function', 'responsibility' and 'means-ends reasoning'. The analysis of these notions involves wider philosophical issues, for example with respect to intentionality, teleology and normativity that are intimately connected with these notions.
The four main themes of the programme are:

1. *Design and moral values*. What types of values are and should be involved in the design of artefacts? How do we cope with conflicting values in a rational way? How can design be made more sensitive to relevant public values? What kinds of moral problems turn up in design practices?
2. *The modeling and design of socio-technical systems.* How should we account for mixed systems of artefacts, human agents and social institutions? To what extent does the notion of design apply to those systems? How to make sense of the notion of responsibility with respect to emergent behaviour of these systems?

3 *Agency and artefacts*. How are the notion of artefact and the notion of technical function related to the notion of human action and the notion of human intentionality?

4 *Technological knowledge and technological rationality.* What types of normativity are involved in the notion of technological knowledge? How should instrumental or means-ends reasoning, as a form of technological rationality, be characterized?

These themes will be discussed in more detail below.

An empirical component

In addition to fundamental philosophical research, the programme aims at active interaction with engineering sciences and practices in well-defined areas and projects. The aim of this is to be inspired by and informed about foundational issues in these disciplines, to base philosophical analysis on empirically accurate descriptions, and to make philosophical insights available for engineering practices where relevant. There are five fields of technology in which researchers are at the moment involved: information and communication technology, biotechnology, biomedical technology, architecture and urban planning, and nanotechnology.

Relevance

Philosophically, the programme is relevant in that the analysis of technical artefacts poses interesting new challenges for existing approaches in various philosophical disciplines; for example, in epistemology (functional knowledge), in philosophy of science (theories of function and explanation), in the theory of action (the agentive function of artefacts), and in ontology (the constitution of artefacts). As regards ethics, it is questionable whether standard ethical theories are able to deal with the kind of moral problems occurring in the design of artefacts and of mixed systems of artefacts and agents (for example, with respect to the attribution of responsibility, or uncertainties and ignorance about first and higher order effects).

The programme has societal relevance in that it contributes to a better understanding of technology, one of the main driving forces of our modern society, which is continuously reshaping modern individual and social life. We will only understand this force adequately if we also look at the details of ‘technology in the making’. The programme intends to contribute to ongoing discussions about the role of technology in society and about the responsibility of engineers.

Scale of the research programme

The programme combines the research efforts of the philosophy groups of the Universities of Technology at Delft and Eindhoven. In 2005 the equivalent of about 15 full-time researchers (fte) was involved in research in the philosophy of technology at these universities. Both philosophy groups have already done or are doing considerable work in the development of these themes, largely in projects that are financed externally. These include: two NWO programmes ('The dual nature of technical artefacts' and 'Norms in knowledge'), one NWO/VENI project, as well as participations in two NWO/STW programmes and three BSIK programmes. These projects are closely related to the themes described (see below). This effort will be continued and intensified along the lines described in this programme.

Selected references

Baker, L.R. (2004), The ontology of artefacts. *Philosophical Explorations* **7**, 99-111.

Dipert, R.R. (1993), *Artifacts, art works, and agency.* Philadelphia: Temple University Press.

Kroes, P.A., & Meijers, A.W.M., eds. (2000), *The empirical turn in the philosophy of technology.* New York: JAI / Elsevier Science.

Kroes, P.A., & Meijers, A.W.M., eds., (forthcoming), The dual nature of technical artefacts, special issue of *Studies in History and Philosophy of Science*.

Mitcham, C. (1994), *Thinking through technology: the path between engineering and philosophy*. Chicago: University of Chicago Press.

Preston, B. (1998). Why is a wing like a spoon? A pluralist theory of function. *Journal of Philosophy* **95**, 215-54.

1. DESIGN, RISKS AND MORAL VALUES

General characterisation

Engineering design may be conceived as the creation of technical artefacts under the guidance of certain values. A variety of values plays a role here. First of all, one can think of instrumental values like effectiveness and efficiency, which are related to the function of a technical artefact. Secondly, moral values like safety, sustainability, user friendliness, respect for autonomy, privacy and justice often play an important role in design choices and in the formulation of design requirements. Thirdly, aesthetic values usually play their part. This research focuses on the role moral values play and should play in design and how the design process can be made more value-sensitive in this respect.

Values in technology have been an important theme in continental approaches in the philosophy of technology (Heidegger, Ellul, Borgmann, see Mitcham, 1994). Such approaches, however, tend to attribute values to technology as such, while we are interested in how values are embedded in concrete technologies and how these values shape engineering practices. Some work along these lines has been done by Winner (1980) and, from the perspective of Science and Technology Studies, by Latour (1992). There are also investigations on moral issues related to technological risks (e.g. Hansson, 2004; Cranor, 1990; Shrader-Frechette, 1991). In general, however engineering design has received only scant attention in the philosophy of technology; a rare exception is Bucciarelli (1994). In the past few years, the Delft and Eindhoven research groups have made some headway in relevant research themes like the ethical aspects of technological risks, design methodology and moral issues in engineering design (Franssen and Bucciarelli, 2004; Kroes, 2002; van de Poel, 2001; van de Poel et al., forthcoming; Zandvoort, 2000).

This theme of the research programme focuses on moral issues in engineering design. Central questions are: How can public values be accommodated in the design process? What role should public actors play in decision making about technological risks in the design process? Another important issue is how designers could, and should, deal with conflicting values. To that end, investigations into the nature of values are necessary, so that it can be clarified if and where values in the design process are incommensurable. Also the notion of (moral) rationality will be scrutinized in order to suggest rational procedures for making choices in the design process.

The research theme will build on general philosophical insights, e.g. about the nature of values and value incommensurability, (e.g. Chang, 1997), and it aims at applying such insights to the realm of technology. An important notion will be value-sensitive design. Although it has been developed in the past few years for a limited application in information technology (e.g. Friedman, 1996 and http://www.nyu.edu/projects/valuesindesign/index.html), it is a useful concept that sums up what major parts of this research theme are aiming at: suggestions for rational procedures of designing artefacts under the guidance of moral values and of maximal risk-avoidance.

Specific themes and research questions

Research will focus on three areas:

i) Moral issues in engineering design

One of the key questions here is: In what ways are (moral) values embedded in design and how can design be made more value-sensitive? This raises the further question what methodologies exist or can be developed for value-sensitive design. An important focus will be on technological risks. One of the issues here is how risks are dealt with in engineering design and technological R&D. Relevant questions are: How could technological risks be better addressed from a moral point of view? When are they (morally) acceptable? How could technological risks best be regulated? Another important issue concerns technical codes and standards – like ISO and NEN norms – and the way they can be relevant for the design process. Codes and standards often play a role in safeguarding values like safety and sustainability. They are formulated by the professional group itself and therefore are a form of (moral) self-regulation in engineering.

ii) ***Conflicting values and design***

Usually engineers are confronted with a multiplicity of values, which are in turn translated into design requirements or design criteria. These multiple values often conflict in the sense that different values seem to demand different decisions from the designer. As a result, trade-offs have to be made among the different values. In a first step, existing formal decision models in engineering for multiple criteria decisions will be analysed, in particular with regards to their claim to rationality. Secondly, it will be attempted to improve existing models on the basis of this analysis. The aim is to develop proposals for better procedures of dealing with conflicting values in engineering design.

iii) ***The nature of values***

Meta-ethical research on central notions like value, rationality and moral knowledge will be carried out, as well as foundational research on how to ground values. The aim is to provide tools for the two more practice-oriented research themes described above. This concerns in particular the problem of value incommensurability: are value trade-offs in design problematic in general or does a value hierarchy exist that helps to solve trade-off problems? Furthermore, moral epistemology is highly relevant to ethics and technology. Any rational application of values will have to give some answers as to the epistemology underlying its choice of values. Both are obviously linked; ethicists must have some answer to the question how they arrive at relevant values (Illies, 2003). Empirical research has also shown that the general public relies heavily on emotions in judging risks (e.g. Slovic et al., 2002). Often engineers and policy makers conclude from this that the public is irrational and should be ignored. However, research on the role of emotions in ethics might shed new light on this (cf. Roeser, 2002). A cognitive theory of emotions could allow for the possibility that we need emotions in order to make rational judgements about the acceptability of technological risks. This allows renewed consideration of the role of the public in decision procedures about technological risks (e.g. 'informed consent').

Current or recent projects

- NWO/STW project 'Ethical aspects of risks of the transition from lab-scale model to full-size open plant in bioprocess technology' (Brumsen, Zwart, van Mil, van de Poel).
- NWO/STW project 'Accountability and the use of advanced medical images and the design of hospital picture archive systems' (Lokhorst).
- PhD-project 'Ethical issues in engineering design' (van Gorp).
- PhD-project 'Informed consent in technology development' (Asveld).
- NWO/VENI project 'Emotions and technological risks: emotions as a normative guide in judging the moral acceptability of technological risks' (Roeser).
- project 'Transcendental arguments as a foundation of ethics' (Illies).
- RISKREG project 'Risk regulation and legislation in the EU and USA' (Zandvoort).
- Project 'Accountability for architectures for identity management systems in e-government' (The Dutch Home Office) (van den Hoven).
- IBM/TI/BSIK project 'Alter ego: profiles, privacy and ambient intelligence' (van den Hoven).

- Handbook project *Philosophy of the Technological Sciences* (Meijers et al.), Elsevier Science.
- Online Encyclopaedia *Applied and Professional Ethics* (van den Hoven, Pogge, Miller), Springer.

Selected references

Bucciarelli, L.L. (1994), *Designing engineers*. MIT Press, Cambridge.

Chang, R. (1997), *Incommensurability, incomparability, and practical reasoning*. Harvard University Press, Cambridge.

Cranor, C.F. (1990), Some moral issues in risk assessment. *Ethics* **101**, 123-143.

Franssen, M., & Bucciarelli, L.L. (2004), On rationality in engineering design. *Journal of Mechanical Design* **126**, 945-949.

Franssen, M. (2005), The normativity of artefacts. *Studies in History and Philosophy of Science*, special issue on 'The dual nature of technical artefacts', in print, with a reply by J. Dancy: The thing to use.

Friedman, B. (1996), Value-sensitive design. *Interactions* **3**, 17-23.

Hansson, S.O. (2004), Philosophical perspectives on risk. *Techné: Research in Philosophy and Technology* **8**, 10-35.

Hoven, M.J. van den (1994), Towards ethical principles for designing polito-administrative information systems. *Informatization and the Public Sector* **3**, 353-373.

Hoven, M.J. van den (2005), *Values, design and ICT*. Inaugural address, Delft University of Technology.

Illies, C. (2003), *The grounds of ethical judgement: new transcendental arguments in moral philosophy*. Oxford: Oxford University Press.

Kroes, P.A. (2002), Design methodology and the nature of technical artefacts. *Design Studies* **23**, 287-302.

Latour, B. (1992), Where are the missing masses? The sociology of a few mundane artifacts. In: W.E. Bijker & J. Law, eds., *Shaping technology / building society*, MIT *Press,* Cambridge, Mass., pp. 225-258.

Mitcham, C. (1994), *Thinking through technology: the path between engineering and philosophy*. Chicago: University of Chicago Press.

Poel, I.R. van de (2001), Investigating ethical issues in engineering design. *Science and Engineering Ethics* **7**, 429-446.

Poel, I.R, van de, Zwart, S.D., Brumsen, M., & Mil, H.G.J. van (2005), Risks of aerobic granular sludge technology: ethical and methodological aspects. In: S. Bathe et al, eds., *Aerobic granular sludge*, London: IWA Publishing, in print.

Poel, I.R. van de, & Verbeek, P.P., eds., (2006) Ethics and engineering design, special issue of *Science, Technology and Human Values*, forthcoming.

Roeser, S. (2002), *Ethical intuitions and emotions: a philosophical study*. PhD-thesis, Free University Amsterdam

Shrader-Frechette, K.S. (1991), *Risk and rationality*. Berkeley: University of California Press.

Slovic, P., Finuncane, M., Peters, E., & MacGregor, D.G., (2002), The affect heuristic. In: T. Gilovich, D. Griffin & D. Kahnemann, eds., *Intuitive judgment: heuristics and biases*, Cambridge: Cambridge University Press, pp. 397-420.

Winner, L. (1980), Do artefacts have politics? *Daedalus* **109**, 121-136.

Zandvoort, H. (2000), Codes of conduct, the law, and technological design and development. In: P.A. Kroes & A.W.M. Meijers, eds., *The empirical turn in the philosophy of technology*, JAI / Elsevier, Amsterdam, pp. 193-205.

2. THE MODELING AND DESIGN OF SOCIO-TECHNICAL SYSTEMS

General characterization

This theme moves from the artefact level to the system level in which artefacts are embedded. Questions addressed at the artefact level in the other research themes are here addressed at the system level. The systemic character of modern technology is not confined to complex aggregates of technical artefacts, but stretches out into the social domain. The functioning of many modern, large-scale technological systems depends as much on all kinds of social institutions – that is, social infrastructure – and on the behaviour of agents, as it depends on the technical hardware. To stress the entanglement of technical and social elements, such complex technological systems are often conceived as socio-technical systems. Both technological elements and non-technical (intentional, social) elements are considered to be constitutive parts of socio-technical systems and to be open to (re)design. A better insight into the nature and dynamics of socio-technical systems may contribute to a clarification of longstanding problems in the philosophy of technology concerning the ideas of technological determinism and of the social construction of technology.

In the 1950s and 1960s the discipline of systems theory rapidly developed in response to an increasing recognition that technology operates and develops in a context where almost everything is connected to almost anything and where the consequences of our actions are difficult to foresee. A systematic scientific approach to technological design was therefore necessary, incorporating human action and decision making (Miser & Quade, 1985; Wilson, 1990; Jackson, 1991). Systems theory never succeeded in making good on its initial promises. Theorizing has remained on a too general level and has in particular not succeeded in developing a clear account of the various ways in which the human agents and social institutions are part of technological systems (Kroes et al., forthcoming, Ottens et al., 2005 and forthcoming). The lack of conceptual clarity is, however, felt as an urgent problem by those who are occupied with the design of large-scale, complex systems (cf. Moses, 2004). There is a challenge to investigate the precise nature of the relations between the technical, individual, and social aspects of these systems, and also what the control or directed change of such complex systems would involve.

Research questions

Three lines of research will be pursued with regard to socio-technical systems:

i) Conceptual clarification of socio-technical systems

A conceptual clarification of the notion of socio-technical system involves a clarification of the nature of its constitutive elements, as well as of the relations between these elements. In particular the relations between heterogeneous elements pose conceptual problems. Questions here include: How should the technical and social elements and their relations be modelled from a formal, systems-theoretic point of view? What kinds of models are used in the engineering and social sciences to represent technical and social systems, respectively, and what are the similarities and differences, strengths and weaknesses of these models? What kinds of elements and what kinds of relations between those elements – physical, functional, intentional, normative – are to be considered as constitutive for socio-technical systems, and on what grounds? How are different ways to conceptualize socio-technical systems related to the way the design of such systems is structured?

ii) Moral issues concerning socio-technical systems

The introduction of elements from the social world as an integral part of complex technological systems implies the introduction of moral/public values as integral elements of these systems. Infrastructures, for instance, raise all kinds of questions about public values, such as autonomy and privacy, which may touch upon the technical subsystems involved. Questions addressed include: In what sense can it be said that (safeguards for) values are embedded in or designed into socio-technical systems? How to conceive of moral issues in the context of networks of actors? To what extent does the notion of (collective) responsibility makes sense at the level of socio-technical systems, given the alleged emergent character of their behaviour?

iii) Emergence and control

Due to the relative freedom of action of individual actors within the system, the precise behaviour of a socio-technical system is difficult to predict or to control. In such systems, one is likely to be confronted with what seem to be emergent phenomena. This emergence may extend to aspects such as reliability or safety. This gives rise to the following questions: to what extent can socio-technical systems and their behaviour be said to be designed, given that the actors within the systems are only to a limited extent under the control of the system's designers and operators? What does this mean for the traditional engineering approach of total design/operation control? To what extent can these systems be controlled?

Current or recent projects

- PhD project 'The (re)design of socio-technical systems' (Ottens)
- NGI/BSIK) postdoc project 'The design of hybrid (social/technical) systems' (Pedersen, Franssen).
- NGI/BSIK postdoc project 'Modeling infrastructures as socio-technical systems' (Kroes, Pedersen, van de Poel)
- NGI/BSIK postdoc project 'Apportioning responsibility in operation and management of infrastructures' (Zandvoort, postdoc).
- PhD-project 'Informed consent in technology development' (Asveld).
- RISKREG project 'Risk regulation and legislation in the EU and USA' (Zandvoort).
- Handbook project *Philosophy of the Technological Sciences* (Meijers et al.), Elsevier Science.

Selected references

Jackson, M.C. (1991), *Systems methodology for the management sciences*. New York: Plenum Press.

Kroes, P.A., Franssen, M., Poel,I.R. van de, & Ottens, M.M. (forthcoming), Treating socio-technical systems as engineering systems: some conceptual problems. *Systems Research and Behavioral Science*.

Meijers, A.W.M. (2003), Can collective intentionality be individualized? *American Journal of Economics and Sociology* **62**, 167-183.

Miser, H.J., & Quade, E.S., eds. (1985-1995), *Handbook of systems analysis*. 3 vol., Chichester: John Wiley.

Moses, J., Foundational issues in engineering systems: a framing paper. In: *Engineering systems monograph*, web document at MIT at http://esd.mit.edu/symposium/pdfs/monograph/.

Ottens, M.M., Franssen, M., Kroes, P.A., & Poel, I.R. van de, (forthcoming), Modeling infrastructures as socio-technical systems. *International Journal of Critical Infrastructures*.

Ottens, M.M., Franssen, M., Kroes, P.A., & Poel, I.R. van de, (2005), Systems engineering of socio-technical systems. Paper for the 2005 INCOSE International Symposium, Rochester, N.Y., 10-15 July 2005.

Royakkers, L.M.M., & Buskens, V. (2002), Collective commitment: a theoretical understanding of human cooperation. *ProtoSociology* **16**, 215-240.

Searle, J.R. (1995), *The construction of social reality*. New York: Free Press.

Wilson, B. (1990), *Systems: concepts, methodologies and applications*. 2nd edition. Chichester, etc.: John Wiley.

3. AGENCY AND ARTEFACTS

General characterisation

This part of the research programme focuses on three related topics. The first one is the nature of technical artefacts. Contrary to non-animated natural objects, artefacts can be said to have a dual nature: a physical nature on the one hand, and a functional nature on the other. The functional nature specifies what the artefact is meant for. This *for-ness* gives artefacts a teleological aspect. The functional nature of artefacts also involves a form of normativity in that function attributions specify what the artefact is *supposed* to do (and it subsequently malfunctions if it does not). Theories of artefacts aim at the integration of the physical and functional aspects that are conceptually very different.

The second topic explores how the nature of artefacts is related to human action and human agency. For an agent an artefact is a standing possibility to *do* something in order to realize certain practical aims. These aims and actions are often entirely new in the sense that they cannot be realized without artefacts. Artefacts thus make actions possible, while they can also prevent agents from doing certain things. This is sometimes conceptualized in terms of a script that is part of the artefact (Akrich, 1992; Latour, 1992).

The functional/teleological nature of artefacts and their intimate connection to human action give artefacts an intentional aspect. In case of complex artefacts (expert systems or control systems, for example) one might also argue that artefacts not only have an intentional aspect, but seem to instantiate a certain form of agency themselves. This raises fundamental questions about the grounds for attributing actions and agency to objects in the world. This is the third topic that will be explored.

With respect to the notion of function, the emphasis in the philosophy of science in recent decades has clearly been on biological functions (Millikan, 1984; Neander, 1991; Ariew et al., 2002). An exception is Preston (1998), who developed a pluralist theory of function that includes artefacts, using Cummins' (1975) notion of system function. Theories that aim at a systematic account of technical functions have been virtually nonexistent until the recent development of such a theory in the Delft NWO research programme 'The dual nature of technical artefacts' (Vermaas & Houkes, 2003 and 2005; Houkes & Vermaas, 2004). There is an important analogy between the mind-body problem and the function-structure problem in the case of artefacts. Much work has been done in the philosophy of mind, though little work has been done to connect this to the theory of artefacts.

Philosophical analyses of actions such as making, using or designing are also scarce (as opposed, for example, to planning, deciding, or raising one's arm). In the context of artificial intelligence much work has been done on artificial agency and even some work on artificial moral agency (see for example Allen et al., 2000). The attribution of agency to artefacts is controversial, though. There are general accounts of agency that include artefacts, such as Dennett's (1989) theory of intentional systems, or the actor-network theory (Law and Hassard, 1999). Finally, work has been done on the analysis of social artefacts (for example, money) in which action and functions also play important roles (Searle, 1995).

Specific themes and research questions

Corresponding to the three topics mentioned, there are three clusters of problems and research questions that are being addressed in this part of the research programme:

i) The nature of artefacts

This cluster addresses mainly epistemological and ontological questions concerning artefacts. They include: How to account for the physical and functional aspects of artefacts? To what extent does the social context of designers and users codetermine the functional properties of artefacts? Is the notion of constitution suitable to capture the specific ontological characteristics of artefacts? What are conditions for function attribution? How do humans recognize artefacts and of what kind is the knowledge of them? Are current theories of function generic or do they apply only to specific kinds or artefacts (do they apply for example, to materials and basic components)? In what sense is a functional decomposition of artefacts different from one in terms of physical properties? If function theories based on selection history are hard to apply to artefacts, as Houkes and Vermaas (2003) have argued, what, if any, is the relevance of evolutionary theory for understanding the function of technical artefacts?

ii) Agents, actions and artefacts

The wide-spread idea that artefacts are mere instruments ignores the fact that they often impose constraints on human actions, or the reverse that they make new types of action possible; that they influence the user's beliefs, desires and intentions; that they support particular values; that they form part of all kinds of social and institutional arrangements; and that they interact in unforeseen ways with human beings. The challenge is to develop a more substantial notion of the agentive function of technical artefacts, while retaining the possibility of making morally relevant distinctions between (i) human agency, (ii) technology assisted human agency, (iii) causal efficacy of artefacts on human agents and (iv) artificial agency. In addition, this cluster includes research questions such as: How does the use of artefacts affect our notion of (moral) action itself? How should responsibility be conceived in actions that include artefacts? To what extent can artefacts be said to embody a script for action and thus to have normative/moral properties?

iii) Artefacts as agents

Complex artefacts challenge our attributions of actions and agency. What are the grounds for eventually attributing agency to these artefacts? Would such an attribution imply that we also have to attribute intentionality to more simple artefacts (such as thermostats) as Dennett claims? Adherents of the actor-network theory even go a step further. They accept the 'principle of generalized symmetry', according to which what is human and non-human should be integrated into the same conceptual framework. They are both agents, or what is called 'actants'. This idea has raised concerns about human agency and identity, and about volunteerism and determinism. The challenge here is to develop a notion of artificial agency without entirely blurring the (morally relevant) distinction between humans and artefacts. Another set of questions concerns the role of artefacts as *epistemic* agents. Given that artefacts are increasingly important in the generation of knowledge, in what sense do artefacts co-determine and bias our knowledge of the world, for example in monitoring and control activities, or the acquisition of data?

Current or recent projects:

- NWO PhD project 'The proper use of artefacts' (*Dual Nature* programme, Scheele)
- PhD project 'The slippery slope of intentionality' (van Amerongen)
- NWO postdoc project 'Intentionality and technical functions' (*Dual Nature* programme, Houkes)
- PhD project 'Thoughtful things' (Romano)
- Handbook project *Philosophy of the Technological Sciences* (Meijers et al.), Elsevier Science.

Selected references

Akrich, M. (1992), The de-scription of technical objects. In: Bijker and Law, eds., pp. 205-224.

Allen, C., Varner, G., & Zinser, J. (2000), Prolegomena to any future artificial moral agent. *Journal of Experimental and Theoretical Artificial Intelligence* **12**, 251-261.

Ariew, A., Cummins, R., & Perlman, M., eds. (2002), *Functions: new essays in the philosophy of psychology and biology*. New York: Oxford University Press.

Baker, L.R. (2004), The ontology of artefacts. *Philosophical Explorations* **7**, 99-111.

Bijker, W.E., & Law, J., eds. (1992), *Shaping technology / building society: studies in sociotechnical change*. Cambridge, Mass.: MIT Press.

Cummins, R. (1975), Functional analysis. *Journal of Philosophy* **72**, 741-765.

Dennett, D.C. (1987), *The intentional stance*. Cambridge, Mass.: MIT Press.

Dennett, D.C. (1990), The interpretation of texts, people and other artifacts. *Philosophy and Phenomenological Research* **50**, Supplement, 177-194.

Dipert, R.R. (1993), *Artifacts, art works, and agency*. Philadelphia: Temple University Press.

Houkes, W.N. & Vermaas, P.E. (2004), Actions versus functions: a plea for an alternative metaphysics of artifacts. *The Monist* **87**, 52-71.

Houkes, W.N., & Meijers, A.W.M. (2005), The ontology of artefacts: the hard problem', *Studies in History and Philosophy of Science*, special issue on 'The dual nature of technical artefacts', in print, with a reply by L.R. Baker: On the twofold nature of artefacts.

Kroes, P.A. (2005), Coherence of structural and functional descriptions of technical artefacts. *Studies in History and Philosophy of Science*, special issue on 'The dual nature of technical artefacts', in print, with a reply by R.R. Dipert: Coherence and engineering design.

Latour, B. (1992), Where are the missing masses? The sociology of a few mundane artifacts. In: Bijker & Law, eds., pp. 225-258.

Law, J. & Hassard, J., eds. (1999), *Actor network theory and after*. Oxford: Blackwell.

Millikan, R.G. (1984), *Language, thought, and other biological categories: new foundations for realism*. Cambridge, Mass.: MIT Press.

Neander, K. (1991), Functions as selected effects: the conceptual analyst's defense, *Philosophy of Science* **58**, 168-184.

Preston, B. (1998), Why is a wing like a spoon? A pluralist theory of function. *Journal of Philosophy* **95**, 215-254.

Scheele, M. (2005), Function and use of technical artefacts: the social conditions of function ascription. *Studies in History and Philosophy of Science*, special issue on 'The dual nature of technical artefacts', in print, with a reply by B. Preston: Social context and artefact function.

Searle, J.R. (1995), *The construction of social reality*. New York: Free Press.

Vermaas, P.E, & Houkes, W.N. (2003), Ascribing functions to technical artefacts: a challenge to etiological accounts of functions. *British Journal for the Philosophy of Science* **54**, 261-289.

Vermaas, P.E. (2005), The physical connection: engineering function ascriptions to technical artefacts and their components. *Studies in History and Philosophy of Science*, special issue on 'The dual nature of technical artefacts', in print, with a reply by S. Mumford: Function, structure, capacity.

Vermaas, P.E., & Houkes, W.N. (2005), Technical functions: a drawbridge between the intentional and structural natures of technical artefacts. *Studies in History and Philosophy of Science*, special issue on 'The dual nature of technical artefacts', in print, with a reply by S.O. Hansson: Defining technical functions.

4. TECHNOLOGICAL KNOWLEDGE AND TECHNOLOGICAL RATIONALITY

General characterisation

What is special about technological knowledge as compared to knowledge in the natural sciences is that it not just describes the world as it is, but that it is specifically directed at the design, manufacture, use and maintenance of artefacts and systems. Technological knowledge, therefore, is action-oriented. It concerns knowledge of technical functions, of operational principles, of how to design, make or use artefacts. What is especially interesting is that this type of knowledge involves normativity in various ways. Functional knowledge, for example, *prima facie* specifies, among other things, what an artefact *should* do. Codes, standards or manuals contain normative requirements with respect to either the artefacts themselves, or to specific actions with artefacts or systems.

The design, manufacture and use of artefacts is not only guided by theoretical rationality, but also by practical rationality. The distinction between these two forms of rationality is common within philosophy. An analysis of the nature of technological rationality requires an understanding of how technological knowledge and technological action build upon both forms of rationality, an understanding that philosophy is not yet able to give. An important aspect is instrumental or means-ends reasoning, where one reasons from ends and plans to decisions about actions and the design and use of artefacts. Current formalizations of practical rationality, such as the theory of rational choice, cannot be seen as an adequate explication of means-ends reasoning. The notions of means and ends are difficult to retrieve from these formalizations, and the idea of the appropriateness of the means and the task of getting to know the means to accomplish an end have no place in these approaches. A key problem here is how to explicate means-ends reasoning in technology as a way of rationally employing knowledge for the sake of action.

Limited work has been done on the analysis of technological knowledge. Bunge (1985), Vincenti (1990) and Mitcham (1994) are among the few examples. There are some interesting analyses of normative aspects that could be part of analyses of technological knowledge by Broome (2000) and by Pollock (2001).The analysis of means-ends reasoning is also underdeveloped, despite the attention given in the philosophical literature to practical rationality. It has enjoyed renewed interest in recent years, beginning with Von Wright's (1969) seminal work and continuing with applications in artificial intelligence, including Pollock (2002) and Castilho (1999). A general account of engineering rationality has been developed by Walton (1990).

The Delft and Eindhoven groups have started to develop these themes further, primarily in the recent NWO research programme 'Norms in knowledge'.

Specific themes and research questions

Research focuses on two related areas:

i) The analysis of technological knowledge

Several types of knowledge in the engineering sciences do not have a clear equivalent in the natural sciences. How to characterize, for example, functional knowledge of technical artefacts? Furthermore, manuals, instructions, codes and standards suggest that there is a type of technological knowledge that contains *ought to do* and *ought to be* statements. How to analyse such prescriptive knowledge? Again, another type of technological knowledge concerns the operational principles of artefacts. Is it different from our knowledge of the mechanisms in nature? What role do models play in engineering, for example scale models or computer models, and do these models differ from those in the natural sciences? The same question applies to explanations in the technological sciences. Are they different? And what is the (applied) ontology that corresponds to knowledge statements in the engineering sciences, that is often operationalized in information systems or expert systems?

ii) Technological knowledge and means-ends reasoning.

This area of research investigates the nature of means-ends reasoning and additionally the extent to which means-ends reasoning can serve as a unifying framework for justifying the various forms of technological knowledge. This leads first of all to the research question what a formal semantics for means-end relations must look like. Particular interest has been paid in the literature to arguments involving means-end relations, since such arguments are needed to design intelligent agents, but this concentrated effort leaves aside many interesting related questions. How should a means-ends semantics account for a notion of efficacy (propensity to bring about an end) and how to distinguish this from related notions of efficiency (ability to avoid burdensome costs and side effects)? Second, functional knowledge seems to have an obvious relation to means-ends reasoning: that an artefact is for doing X implies that the artefact can be used as a means for realizing X. It is not obvious, however, that this can exhaust the notion of an artefact's proper function, since there are strong indications that proper functions also depend on social factors (Scheele, 2005). So how exactly is means-ends reasoning related to functional knowledge? A third type of question concerns the role of means-ends reasoning in the design of technical artefacts: To what extent can design processes be reconstructed as rational processes and can their outcome be justified rationally? What are the success criteria for a design?

Current or recent projects:

- NWO PhD project ‘Designing technical artefacts’ (*Dual Nature* programme, de Ridder)
- NWO PhD project ‘Normativity and prescriptive knowledge’ (*Norms in Knowledge* programme, Vaesen)
- NWO Postdoc project ‘Functional knowledge and normativity’ (*Norms in Knowledge* programme, Hughes)
- NWO/STW project ‘Ethical aspects of risks of the transition from lab-scale model to full-size open plant in bioprocess technology’ (Brumsen, Zwart, van Mil, van de Poel)
- Handbook project *Philosophy of the Technological Sciences* (Meijers et al.), Elsevier Science.

Selected references

Broome, J. (2000), Normative requirements. In: J. Dancy, ed., *Normativity*, Oxford: Blackwell, pp. 78–99.

Bunge, M. (1985), Technology: from engineering to decision theory. *Treatise on basic philosophy* vol. 7, Epistemology and methodology, part 3, Philosophy of science and technology. Dordrecht etc.: Reidel.

Castilho M.A., Gasquet, O., & Herzig, A. (1999), Formalizing action and change in modal logic I: the frame problem. *Journal of Logic and Computation* **9**, 701-735.

Franssen, M., & Bucciarelli L.L. (2004), On rationality in engineering design. *Journal of Mechanical Design* **126**, 945-949.

Houkes, W.N. (2005), Knowledge of artefact functions. *Studies in History and Philosophy of Science*, special issue on ‘The dual nature of technical artefacts’, in print, with a reply by A. Morton: Finding the corkscrew.

Pollock, J.L (2001), Evaluative cognition. *Nous* **35**, 325-364.

Pollock, J.L (2002), The logical foundations of means-end reasoning. In R. Elio, ed., *Common sense, reasoning, and rationality*, New York: Oxford University Press.

Raz, J. (1975/1990), *Practical reason and norms*. Oxford: Oxford University Press.

Scheele, M. (2005), *The proper use of artefacts: a philosophical theory of the social constitution of artefact functions*. PhD thesis, Delft University of Technology.

Vincenti, W.G. (1990), *What engineers know and how they know it: analytical studies from aeronautical history*. Baltimore: Johns Hopkins University Press.

Vries, M.J. de, Cross, N., & Grant, D.P. (1993), *Design methodology and relationships with science*. Dordrecht: Kluwer Academic Publishers.

Walton, D.N. (1990), *Practical reasoning: goal-driven, knowledge-based, action-guiding argumentation*. Savage, Md.: Rowman and Littlefield.

Wright, G.H. von (1963), Practical inference. *Philosophical Review* **72**,159-179.