INTELLIGENT ETHICS

Penny Duquenoy
School of Computing Science, Middlesex University, London, N14 4YZ.

Abstract;   This paper discusses the impact of envisaged intelligent applications on the lives of the individuals who may be using them, and investigates the ethical implications of autonomous decision-making that is beyond the control of the user. In an increasingly networked world we look beyond the individual to a social picture of distributed multi-agent interaction, and in particular at the concepts of rules and negotiation between these virtual social agents. We suggest that the use of such agents in a wider social context requires an element of ethical thinking to take place at the grass roots level – that is, with the designers and developers of such systems.

Keywords;   Ethics, Ambient Intelligence, intelligent agents, multi-agent systems, negotiation.

1. INTRODUCTION

Developments in technology have seen an increasing trend towards "smart" and "intelligent" applications. This trend is a natural progression from the research and work in Artificial Intelligence and the availability of enhanced computing infrastructures, combined with economic need and the marketing drive of the computer industry. The benefits of such applications are easy to see – fast and autonomous action in complex situations beyond the abilities of the human mind. As the use of technology generally has spread beyond the work environment into the lives of ordinary people the benefits are held to be increased user-friendliness - thus addressing the difficulties of the novice user - and reducing the user's cognitive load.

Duquenoy, Penny, (2004) "Intelligent Ethics". In *Building the Information Society*, René Jacquart (Ed.). Kluwer Academic Publishers. Pp. 597-602. IFIP World Computer Congress 2004, Toulouse, France.

This trend for smart appliances has gained momentum with recent concepts in ubiquitous and pervasive computing. These terms describe the move towards computer technology operating "in the background" in an invisible and non-intrusive way[1]. The combination of smart, or "intelligent", technology with ubiquitous computing generates the notion of "Ambient intelligence" – that is, intelligent systems that operate in our surrounding environment.

Whilst fully integrated and compatible intelligent systems are still a long way off, there are some ethical concerns regarding their implementation. We can predict (based on current technologies) that issues of privacy and security will continue to prevail (cf. Petrio, 2003), but there are also deeper issues at stake. In a situation where a number of technologies are competing for bandwidth, airwaves, processing power, storage and memory, choices will need to be made. Priorities will have to be allocated, and trade-offs accepted. Who decides the rules, and makes the choices?

This short paper discusses the impact of envisaged intelligent applications on the lives of the individuals who may be using them, and investigates the ethical implications of autonomous decision-making that is beyond the control of the user. In an increasingly networked world we look beyond the individual to a social picture of distributed multi-agent interaction, and in particular at the concepts of rules and negotiation between these virtual social agents. We suggest that the use of such agents in a wider social context requires an element of ethical thinking to take place at the grass roots level – that is, with the designers and developers of such systems.

2. BACKGROUND

The background to this discussion is founded upon earlier reflections on ambient intelligence inspired by the work carried out by the European Commission Information Society Technology Advisory Group (ISTAG) in 2000/2001 (Ducatel et al. 2001). In that document four futuristic scenarios were put forward as a device for considering the future direction of research and development of intelligent systems. It is clear from these scenarios that

---

[1] Pervasive computing is described as providing a paradigm "for all the time, everywhere services" (Second IEEE International Conference on Pervasive Computing and Communications. PerCom2004. The call for papers includes topics such as "Intelligent environments, wearable computers, smart devices and smart spaces" amongst others. Note also 4th International Workshop on Smart Appliances and Wearable Computing (www.unl.im.dendai.ac.jp/IWSAWC/).

Duquenoy, Penny, (2004) "Intelligent Ethics". In *Building the Information Society*, René Jacquart (Ed.). Kluwer Academic Publishers. Pp. 597-602. IFIP World Computer Congress 2004, Toulouse, France.

intelligent devices will be expected to interact with others in different contexts. The devices are described as communicating systems exchanging information with each other on behalf of human individuals. Whilst the use of intelligent devices for the individual user are projected as beneficial – that is, as facilitating the management of various lifestyles (such as business, personal, and information management) – the implications of agents acting in a broader social context are not apparent.

Our previous work examined the ethical implications in the application of these devices (Duquenoy and Masurkar, 2004). This current work looks at a more fundamental issue, that is, the infrastructure of multi-agent networking. Networked agents must necessarily include decision-making procedures that deal with negotiation and priorities of operation, and in situations where decisions and judgements are made by third parties (be they computer-mediated or human) there are also inherent ethical issues.

3. ETHICS, JUDGEMENT AND CHOICE

Ethics is founded upon the principle of intelligent rational agency, and is essentially about judgements and choices that lead to actions. Living a "good" life is to live a life based on good choices and actions (whatever we may determine "good" to be). Taking two of the most used ethical theories we can, in simple terms, say that actions can be good or bad in themselves, or good or bad in their consequences. That is not to say that all actions have ethical implications – but any that cause harm to others, are unjust, or take advantage of another to their detriment (exploit others), are usually considered unethical.

Making ethical decisions can be difficult - there are often competing ethical choices, and often a lack of knowledge about circumstances that may affect a judgement.

Even from this extremely simplistic picture of ethics, we can see that intelligent devices acting on our behalf will be making judgements and choices to inform actions, and that some of these actions are likely to have ethical implications. If we are to achieve the potential of ambient intelligence, and have intelligent devices operating in an "invisible and non-intrusive way" then it seems apparent that some thought should be given to the instructions we give them.

Duquenoy, Penny, (2004) "Intelligent Ethics". In *Building the Information Society*, René Jacquart (Ed.). Kluwer Academic Publishers. Pp. 597-602. IFIP World Computer Congress 2004, Toulouse, France.

4. AGENTS, NEGOTIATIONS AND DECISIONS

It is clear that for the successful co-ordination of networked agents a co-operative system must be in place. A key element of co-operation is negotiation between participants - which necessarily implies choice. Offers are made, assessed, accepted or rejected. As a result of the negotiation a decision is reached, and appropriate action taken.

Negotiating programmes are not a new idea, and have been developed to promote ethical "good" – the World Wide Web consortium's Platform for Privacy Preferences Project (P3P) is a good example of a simple interaction[2].

A more complex "social" example of negotiation between agents is a simulation developed by Clarke and McCabe (2003). The authors demonstrate the application of the programme by setting it in the context of a "ballroom scenario" thereby allowing elements of negotiation, choice and action. Their agents are able to distinguish between male and female partners, negotiate with prospective partners for dance engagements, and make choices between alternative dances and activities. By attributing attribute the concepts of belief, desire and intention to the agents, the authors can encompass "key behavioural features of agents: autonomy, adaptability and responsibility." This scenario describes a familiar context and incorporates social norms, such as male and female dance partners, the male approaching the female to request a dance, and with the alternative to dancing given as going to the bar! So, with a given set of social norms in a familiar, and fairly restricted context, negotiations can be pursued. Ethical considerations in this case could concern the application of social conventions – what happens for user's who do not conform?

5. SCARCE RESOURCES

If it is true that agents will be sharing and/or competing for finite resources such as operational space (bandwidth, signal processing, processing power, memory etc.) then some rules of precedence must be followed – some functions and operations must take priority over others. Fixed rules under these circumstances are not appropriate, as the available resources are likely to be in a continual state of flux. Negotiations in such a

---

[2] The concept behind this is to aid users with privacy requirements. P3P will negotiate on behalf of a user with a web site, checking the web site's privacy policy against the user's requirements, and notify of any discrepancies. A warning is given however, that this does not imply the web site will adhere to their privacy policy.

Duquenoy, Penny, (2004) "Intelligent Ethics". In *Building the Information Society*, René Jacquart (Ed.). Kluwer Academic Publishers. Pp. 597-602. IFIP World Computer Congress 2004, Toulouse, France.

situation need to be on-going and adaptive. Some preliminary work in this area has been carried out by Rzevski et al. (2003) and is intended to address the changing needs of supply and demand in virtual organisations. Their guiding principle is based on "incessant negotiation between (1) internal VE resources and (2) external demands for these resources".

The above model, set as it is in a context of virtual enterprises within a global marketplace, involves an element of competition. Put simply, whilst agents within an organisation may be working as a team for the benefit of the organisation, external operators may be ignored or rejected as they become less useful. Applying this model to an individual user operating in a wider social network could be beneficial, in that the agents working for the user will always act in the user's best interests, discarding non-useful connections. We should remember though that commercial models are not always appropriate in a social context – if those who are not useful, or unsuccessful get excluded. This is not to say however, that the principles of "usefulness" cannot be changed to incorporate different values.

6. CODE AS REGULATION

Lawrence Lessig (1999) shows how programming code can, and does, act as a regulatory force. In basic terms programmes allow users to perform certain actions, and prohibit them from pursuing others. Programmes can be used to uphold laws and values – as we have seen with the W3C P3P project above. We also know that other operational programmes have preferences set – for example, browsers can be set to accept or deny cookies, and there are other security settings and levels that are optional for the user. The virtual enterprise model described above is designed to allow users to set and change priorities.

But, how many ordinary users are aware of these options? How many are aware of the rationale behind the options? Some users may say that they prefer to be left out of the loop, and not to have to make decisions. They may prefer to leave these decisions in the hands of agents.

It is also worth noting that in the field of Human Computer Interaction (HCI) the user is central to the design, that is, "users" are given a valued status. This might not necessarily be the case in some cultures and states. In answer to the question "who is the user" we might reply: governments, law enforcement agencies, commercial enterprises, illegal enterprises, etc. In other words it is equally possible to devise agents that operate out of sight of

Duquenoy, Penny, (2004) "Intelligent Ethics". In *Building the Information Society*, René Jacquart (Ed.). Kluwer Academic Publishers. Pp. 597-602. IFIP World Computer Congress 2004, Toulouse, France.

the user, that exploit individuals, and cause harm. As well as "good" ethical agents we could have unethical or "rogue" agents.

7. CHALLENGES

The challenge for technologists will be to develop compatible, safe and robust multi-agent systems. These systems will be responding to given rules; negotiating; arranging tasks according to perceived worth and timeliness; matching data and profiles; and determining when tasks are complete. All of these activities clearly have an impact on the user – but consideration should be given to the extent of that impact on personal life-styles. The anticipation is that ambient intelligence will "make life better" – with some advance thought this could be the case.

The examples above illustrate the immediate impact of negotiation strategies on outcomes. In an agent-mediated environment the consequences of the negotiations are ethically relevant to the user. For example: In competing for real-time information delivery – who wins? Who (designer) or what (intelligent agent) will prioritise – will my grocery need take precedence over my medical appointment? Amongst the many likely stakeholders (individuals, commercial enterprises, social services, government) who takes priority? Will commercial interests impinge on personal management of information (as for example the current intrusions based on cookies, or spam). Will government needs override personal needs? In some cases user preferences can be set – would there be any circumstances where those preferences could, or should, be over-ridden?

The developers of such models and programmes are clearly in a position to regulate not only the operation of agents, but by default the way individuals may operate in their lives. The task is how best to serve the developers and designers, so that the considerations indicated above can be taken into account proactively in the development process – rather than trying to reactively put matters right after the event.

8. CONCLUSION

In the ambient intelligent environment we have, potentially, a number of interactions and negotiations taking place. The results of these interactions are informed by values (beliefs), motivated by need (desires), and directed by goals (intentions) - characteristics which are likely to be embedded in the

Duquenoy, Penny, (2004) "Intelligent Ethics". In *Building the Information Society*, René Jacquart (Ed.). Kluwer Academic Publishers. Pp. 597-602. IFIP World Computer Congress 2004, Toulouse, France.

early stages of development, and may not necessarily be to the benefit of users.

The examples used give some indication of the different ethical implications of the models, and the regulatory aspects of coding. Negotiations coupled with scarce resources raise a number of questions with regard to priorities, and beneficiaries. Designers and developers will be influential in determining the benefits of ambient intelligence in the future, and attending to the ethical aspects of these technologies is a vital ingredient to their success.

Intelligent devices acting as agents and pursuing negotiations with other agents and making decisions are, by definition, acting on our behalf as third parties. The decisions and judgements they make carry ethical implications. Human agents, such as lawyers, financial agents, estate agents, etc. have codes of conduct and practice to inform their practice – ethics is as important in intelligent agents, and even more important in an environment of ambient intelligence as envisaged by the EU, simply because of their implicit "invisibility".

REFERENCES

Clark, K.L. and F.G. McCabe (2003) "Go! For multi-threaded deliberative agents". Available from http://www.doc.ic.ac.uk/~klc/ Accessed: 02/03/04

Ducatel, K., M. Bogdanowicz, F.Scapolo, J. Leijten and J-C Burgelman. "Scenarios for Ambient Intelligence in 2010", *ISTAG (Information Society Technologies Advisory Group) Final Report, February 2001*, IPTS Seville. http://www.cordis.lu/ist/istag.htm

Duquenoy, Penny and Vijay Masurkar (2004) "Surrounded by Intelligence" in *Risks and Challenges of the Networked Society.* Proceedings of 2nd IFIP Summer School. Forthcoming.

Lessig, Lawrence (1999) *Code and other laws of cyberspace*. Basic Books.

Petrio, Roberto Di and Luigi V. Mancini. (2003) "Security and Privacy Issues of Handheld and Wearable Wireless Devices". *Communications of the ACM*, Vol.46 No. 9. Pp75-79.

Rzevski, G, Skobelev, P, Batishchev, S, Orlov, A. (2003) "A Framework for Multi-Agent Modelling of Virtual Organisations". In Camarinha-Matos, L M and Afsarmanesh, H (eds), *Processes and foundations for Virtual Organisations*, Kluwer Academic Publishers. Pp. 253-260.

Duquenoy, Penny, (2004) "Intelligent Ethics". In *Building the Information Society*, René Jacquart (Ed.). Kluwer Academic Publishers. Pp. 597-602. IFIP World Computer Congress 2004, Toulouse, France.

World Wide Web Consortium. Platform for Privacy Preferences (P3P), http://www.w3.org/TR/P3P/ accessed 21/04/04.