

Ören, T.I. (2002 - Invited Plenary Paper). [Ethics as a Basis for Sustainable Civilized Behavior for Humans and Software Agents](#). *Acta Systemica*, 2:1, 1-5. Also in the Proceedings of the InterSymp 2002 - The 14th International Conference on Systems Research, Informatics and Cybernetics of the IIAS, July 29 - August 3, Baden-Baden, Germany.

## **Ethics as a Basis for Sustainable Civilized Behavior for Humans and Software Agents**

Tuncer Ören  
OC-MISS/SCS

(Ottawa Center of the McLeod Institute of Simulation Sciences  
of the Society for Modeling and Simulation International)  
School of Information Technology and Engineering  
University of Ottawa  
Ottawa, Ontario, Canada

### **Abstract**

Civilization is relatively high level of cultural and technological development. The degree of civility of a society depends on the balance of the rights and privileges of individuals and groups that exist in the society (Durant, 1935). Respect for the rights of others is the foundation of ethics (Ören, 2000). In this article, ethical issues are studied for humans and for software agents. For individuals, sources of the respect for the rights of others are pointed out; they can be self-initiated or externally imposed. Self-initiated respect can be based on personal and philosophical beliefs of civilized individuals. Externally imposed respect for the rights of others can originate from religion, state, or society. In every society, customs, pressures, and regulations influence the behavior of the individuals. Codes of conduct may not be necessarily ethical. In civilized societies, state guarantees the rights of the individuals by jurisdiction. Ethical behavior –for humans as well as for software agents– can be sustained in proper (non corrupt) environments where protective mechanisms would exist. The role of such protective mechanisms in the assurance of the sustainability of civilized behavior is stressed. In an era where software agents are and will be used more and more, mechanisms to monitor their behavior are needed to protect the benevolent agents to sustain the civilization of infohabitants.

### **Keywords:**

Sustainable ethics  
Ethics for infohabitants  
Ethics for software agents  
Ethics for Internet agents

## 1. Ethics as a Basis for Sustainable Civilized Behavior for Humans

Civilization has several aspects and definitions. A classical perception is: “*Civilization is social order promoting cultural creation. Four elements constitute it: economic provision, political organization, moral traditions, and the pursuit of knowledge and the arts.*” (Durant, 1935). According to Durant, the degree of civility of a society depends on the balance of the rights and privileges of individuals and groups that exist in the society. The origin of respect for the rights of others as shown in Figure 1, was summarized by Ören (2000).

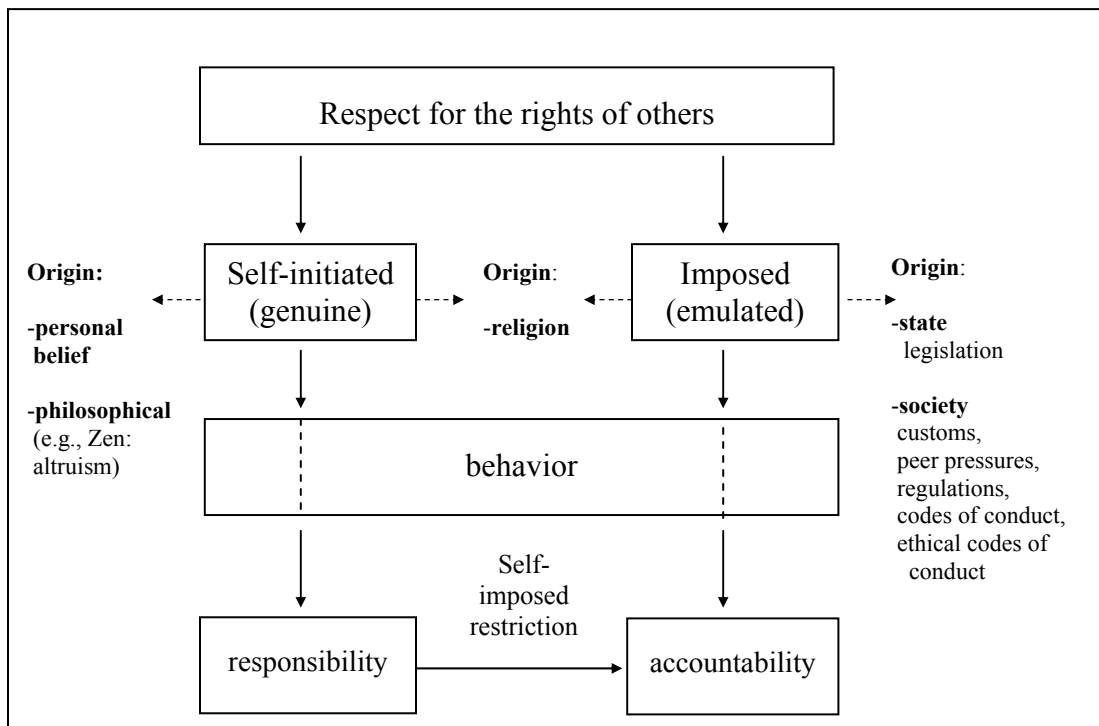


Figure 1. Origins of the respect for the rights of others

Respect for the rights of others can be self-initiated or externally imposed. Self-initiated respect for the rights of others can be based on personal and philosophical beliefs of civilized individuals. Self-imposed restriction on the responsible behavior implies accountability. Externally imposed respect for the rights of others can originate from religion, state, or society. As far as religious origin of respect for the rights of others is concerned, unfortunately, some people may think that the respect is due to only members of the same religion and/or sect. Codes of conduct may not be necessarily ethical. Accountability is a required characteristic of the behavior when it is imposed. Ethical conduct is essential in sustaining a civilization. In practically every society, customs, pressures, and regulations influence the behavior of the individuals. Whenever, respect for the rights of others does not stem from individuals, it has to be enforced by external sources. Every source of ethical behavior ought to be encouraged.

These include promoting ethics as the rising value in the society, educating (as opposed to mere training) the members of the society, enforcing legislation consistently, promoting professional ethical codes of conduct.

## **2. Ethics as a Basis for Sustainable Civilized Behavior for Software Agents**

*Software agents* (IAP) are software modules that can work on behalf of humans, organizations, smart systems, and other agents. They have cognitive abilities such as autonomy (usually limited for the purpose of not creating havoc in software environments), perception, goal processing and goal-directed knowledge processing; and they can affect their knowledge environment –directly if their environment is purely software, or indirectly through actuators. Software agents are already very important and promising for several application areas.

### **2.1 Early studies**

Problems associated with software agents have been elaborated on since early 1990s. For example Croft (1997) starts his article with “agent abuses” and states that: “The evolution of Internet Agents, bred by commercial interests to produce the quick and the thorough, must be tempered by the necessity to respect the boundaries of the environment.” Etiquette is the conduct or procedure required by good breeding or prescribed by authority to be observed in social or official life (Merriam-Webster). “Internet etiquette, which is often called netiquette, is similar to and often overlaps with ethical issues; however, netiquette is less a strict code of conduct and more an understanding of what constitutes good manners and electronic community expectations.” (Hinchliffe, 1997). Netiquette applies both to humans as well as to software agents. Several studies elaborate on Internet agent etiquette. Early netiquette studies for software agents concentrated on issues such as safety, tidiness, moderation, and vigilance. (Etzioni and Weld, 1994). *Safety*: The agent should not alter its environment in a destructive way. *Tidiness*: The agent should leave its environment as it found it. *Moderation*: The agent should limit its use of scarce resources. *Vigilance*: The agent should not perform user requests to generate detrimental consequences. Eichmann’s netiquette rules include identification, moderation, self monitoring, and information sharing (Eichmann, 1994). *Identification*: Agents must identify themselves. *Moderation*: Agents must not overload servers from which they request information. *Information sharing*: Agents should share their information with others.

Nwana (1996) categorized some issues that the society has to grapple with through various legislations. She categorized the issues as privacy, responsibility, and ethical issues. *Privacy*: Ensuring the anonymity of the user while an agent is working. However, this is in conflict with identification rule. *Responsibility*: The relinquished authority of the user to the software agent redefines user’s responsibility in several ways: responsibility to others (Is it the agent or its owner who will be responsible? What would it mean agent’s responsibility?) Responsibility of the agent to its owner: What happens when an agent makes an agreement –unacceptable to its user- with another agent? *Legal issues*: Most often, agents will be COTS (commercial off-the-shelf) products with possible customization. If an agent does not disclose some vital information for a thorough assessment of a situation, in other words, if an agent misguides another agent, who

will be responsible and/or accountable: the company which sells the agent and/or whoever does the customization of the agent.

## 2.2 Contemporary concerns

The players in the distributed and connected information systems are not limited to individuals and software agents, but include infohabitants in general. *Infohabitants* of the connected information systems include individuals, organisations, smart appliances, smart buildings, and other smart systems, as well as virtual entities acting on their behalf. Hence their behavior are important for the sustainability of the overall system. The virtual entities acting on behalf of individuals and organisations are (or can be) implemented as software agents (and avatars).

One of the projects of the Information Society Technologies Programme (IST) of the European Commission is Universal Information Ecosystems (UIE). ALFEBIITE is a European multinational consortium focused on the development of “A logical framework for ethical behavior between infohabitants in the information trading economy of the universal information ecosystem.” To assure reliability of communications, mechanistic procedures such as encryption, passwords, and firewalls are necessary but insufficient. The ALFEBIITE consortium aims to supply enhanced reliability with formally specified, socially motivated, anthropomorphic relations, such as trust, authority, reputation, etc. Eymann et al. (2000) also elaborates on the need of trustworthy agents by elaborating on what they call “catallactic information systems.” Infohabitants are autonomous (or quasi-autonomous since each has to act within given constraints). Goals of the infohabitants can be unrelated or related. Related goals can be harmonious, conflicting, or hostile. *Harmonious relations* are represented by holonic agents (or holons (Schillo et al), or cooperative agents). The aim for each holon is to reduce its autonomy, when need arise, to collectively achieve a goal. *Conflicting relations* are represented by self-interested agents. The aim is to optimize an individual (local) utility. Problem arises when there is a conflict. In closed information systems, two possibilities exist: centralized and decentralized authorities can resolve conflicts by arbitration or by auctioning, respectively. However, Universal Information Ecosystems are by definition very large heterogeneous dynamic information systems. Hence, decentralized conflict resolution mechanisms should exist to avoid collapse of the systems as well as attack of the self-interested agents from possible hostile behavior of other agents, including from other self-interested agents. In harmonious and conflicting relations assurance of the trustworthiness of agents is of primordial importance. *Hostile relations* are represented by antagonistic agents. The aim is to harm: cripple or eliminate; the counter aim is to minimize the threat to a nuisance. Some examples are viruses, worms, and trojan horses. Infowar agents, including logical and time bombs, are for the electronic wars and may threaten civilizations by attacking civilian as well as military targets.

Since agent technology is an important technology, a critical question is: “does blind trust ensures reliability of an agent system?” Accordingly, there has to be ways to monitor and/or limit the behavior of agents to assure that they don’t breach the rights of other agents. A recent study focuses on artificial moral agents (Allen et al, 2000). Fritz’ (2002) interest is to create a scientific ethics for artificial intelligent systems consistent with ethics applicable to humans. While

Internet is transforming into Universal Information Ecosystems, appropriate research and development of corresponding technology is paramount.

## References

- (ALFEBIITE); <http://alfebiite.ee.ic.ac.uk/Templates/overview.htm>
- Allen, C., Varner, G. and J. Zinser (2000). Prolegomena to Any Future Artificial Moral Agent; J. Expt. Theor. Artif. Intell., Vol. 12 (pp. 251-261)
- Croft, D.W. (1997); Internet Agent Etiquette;  
<http://www.alumni.caltech.edu/~croft/research/agent/etiquette/>
- Durant, W. (1935); The Story of Civilization, Vol. 1 (pp. 36-71), Simon and Schuster.
- Eichmann, D.; Ethical Web Agents; <http://mingo.info-science.uiowa.edu/eichmann/www-f94/ethics/ethics.html#FN1> In Electronic Proc. of the 2<sup>nd</sup> World Wide Web Conference'94: Mosaic and the Web. Elsevier.
- Etzioni, O. and D. Weld (1994); A Softbot-Based Interface to the Internet", *Communications of the ACM* Vol. 37, No. 7 (pp. 72-76)
- Eymann, T. et al. (2000); The Catallaxy as a New Paradigm for the Design of Information Systems; <http://www.iig.uni-freiburg.de/telematik/forschung/publikationen/pubfiles/EyPaSc2000.pdf> In: Proc. of the 16<sup>th</sup> IFIP World Computer Congress, Conf. on Intelligent Information Processing, Beijing, PR China.
- Fritz, W. (2002); Ethics as a Science; <http://www.anice.net.ar/intsys/ethics.htm>
- Hinchliffe, L.J. (1997); Electronic Culture, Ethics, and Etiquette.  
<http://alexia.lis.uiuc.edu/~janicke/Culture.html>
- (IAP); Intelligent Agent Repository <http://www-cia.mty.itesm.mx/~lgarrido/Repositories/IA/agents.html>
- (IST); Information Society Technologies Programme of the European Commission  
<http://www.cordis.lu/ist/home.html>
- Merriam-Webster; <http://www.m-w.com/cgi-bin/dictionary?book=Dictionary&va=etiquette>
- Nwana, H.S. (1996); Software Agents: An Overview, Knowledge Engineering Review, Vol. 11, No 3 (pp.1-40); <http://agents.umbc.edu/introduction/ao/>
- Ören, T.I. (2000); Responsibility, Ethics, and Simulation; Transactions of the Society for Modeling and Simulation International, Vol. 17, No. 4 (pp. 165-170)
- Schillo, M. et al. (2001); Towards a Theory of Flexible Holons: Modelling Institutions for Making Multi-Agent Systems Robust.  
<http://ccs.mit.edu/dell/aa2001/aa2001papers%5Cpaper8.pdf>
- (UIE); Universal Information Ecosystems <http://www.cordis.lu/ist/fetuie.htm>