InterSymp'2002

Baden Baden, Germany, July 29 – Aug. 3, 2002

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

Tuncer Ören, Professor Emeritus

Ottawa Center of the McLeod Institute of Simulation Sciences of the SCS

University of Ottawa, Ottawa, Ontario, Canada

http://www.site.uottawa.ca/~oren/

Plan:

- **Background**
  - **Ethics/philosophy, ethics/civilization**
  - **Respect for the rights of others**
- **Ethics for individuals**

- **Professional ethics**

- **Ethics for software agents**
  - **Early studies**
  - **Contemporary concerns**

Ethics is one of the fundamental issues of philosophy:

- Existence
- Knowledge
- Values:
  Aesthetics
  Axiology - from Greek axia (value) and logos (study)
  - the study of **ethical values**

Ethics or morality is one of the pillars of civilization:

"*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)

The **respect for the rights of others**:

- is **the essence of right conduct**.

- It is also a fundamental concept in civilized societies, essential for the regulation of the social dynamics in order to sustain the civilization.

Also:

No business, how lucrative it is, can be sustained if one looses the trust of the stakeholders.
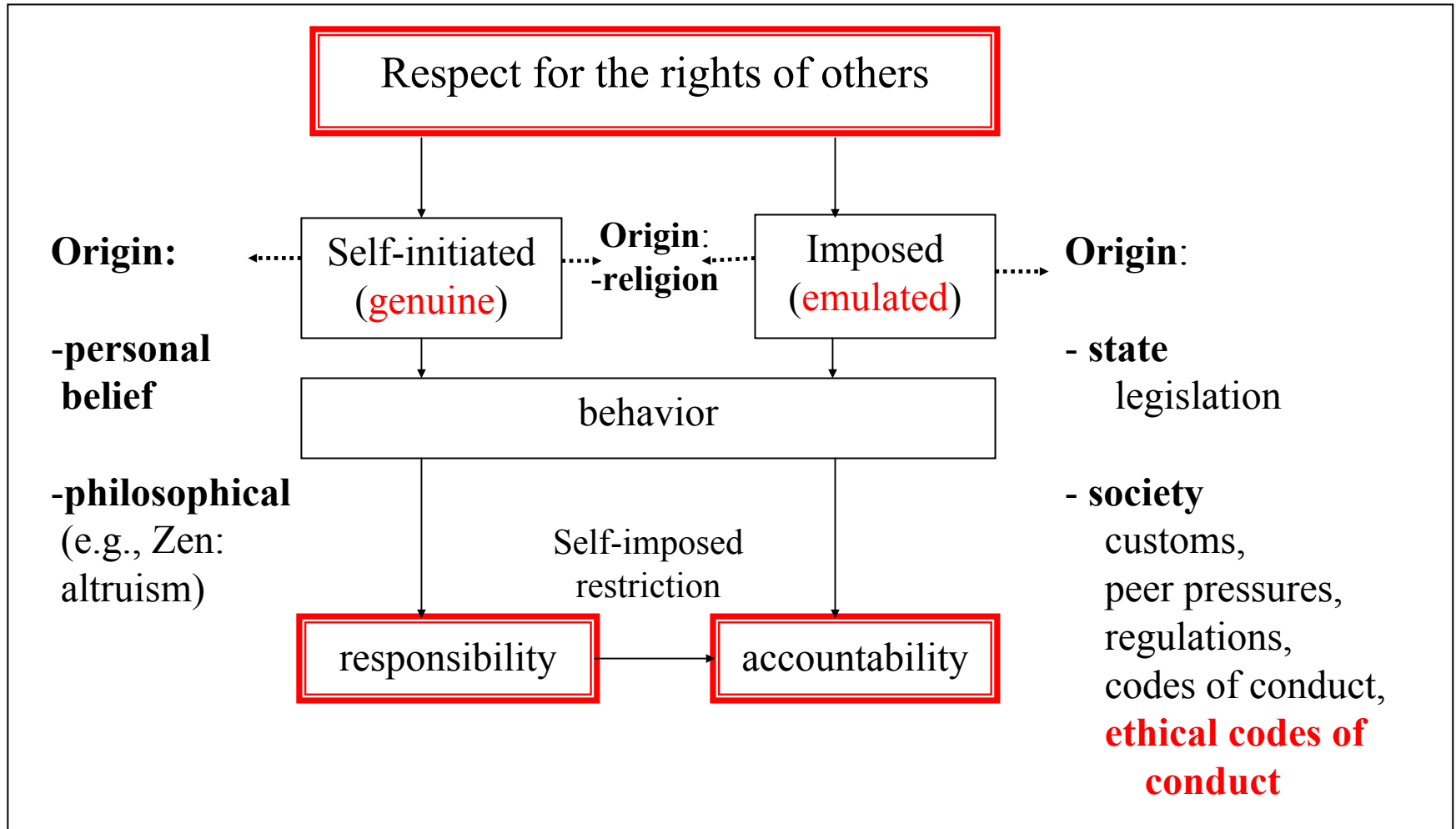
Plan:

- **Background**
  - **Ethics/philosophy, ethics/civilization**
  - **Respect for the rights of others**
- **Ethics for individuals**

- **Professional ethics**

- **Ethics for software agents**
  - **Early studies**
  - **Contemporary concerns**

Golden Rule:

"What you don't want done to yourself, don't do to others."

Confucius (551-479) BC)

The origins for the respect for the rights of others:

Respect for the rights of others

Origin: ← Self-initiated (genuine) → **Origin:** **-religion** ← Imposed (emulated) → **Origin:**

-**personal belief**

behavior

-state legislation

-**philosophical** (e.g., Zen: altruism)

Self-imposed restriction

- **society** customs, peer pressures, regulations, codes of conduct, **ethical codes of conduct**

responsibility → accountability

Why?

"We all think electricity is entirely governed by natural laws, and yet we think it is rational to put up lightning conductors. Well, I should say that an **ethic is**, as it were, **a lightning conductor for human passions**, to enable them within a deterministic world to work in a way that produces a minimum of disaster."

Bertrand Russell, 1942 (Spinoza'a Ethics)

Ethics and Courage:

"To the extend to which a man has freedom, he needs a personal morality to guide his conduct. There are some who would say that a man need only obey the accepted moral code of his community. … . Such practices as cannibalism, human sacrifice, and head hunting have died out as a result of moral protests against conventional moral opinion. If a man seriously desires to live the best life that opens to him, he must learn to be critical of the tribal customs and tribal beliefs that are generally accepted among his neighbours" (Russell).

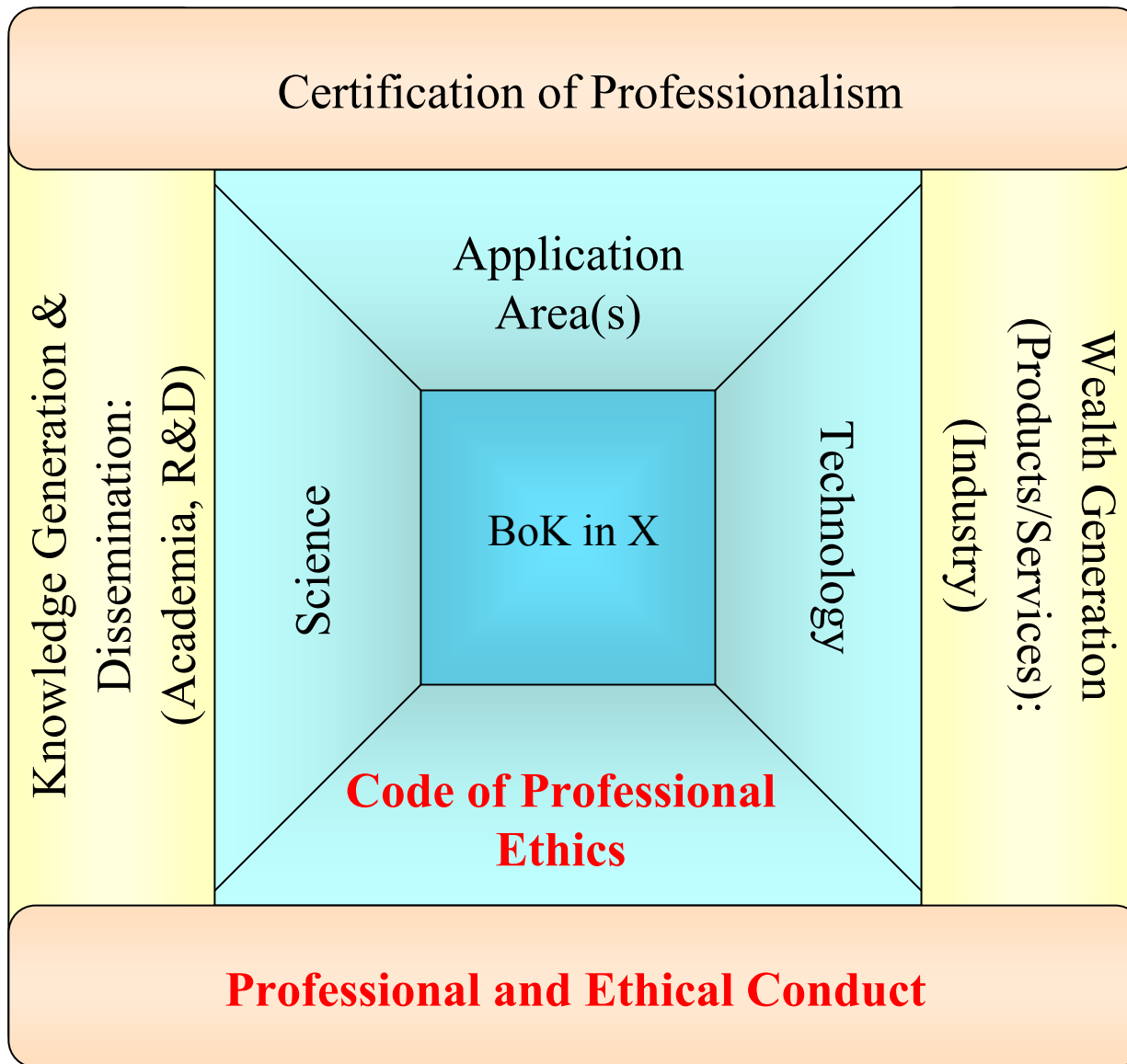> *"The unexamined life is not worth living."*
>
> Socrates (469-399 BC)

And "life" may well cover professional societies and professions as well.

If your activities have no (serious) implications to others,

it does not matter whether or not you feel responsible towards them or you should be held accountable.

Plan:

- **Background**
  - **Ethics/philosophy, ethics/civilization**
  - **Respect for the rights of others**
- **Ethics for individuals**

- **Professional ethics**

- **Ethics for software agents**
  - **Early studies**
  - **Contemporary concerns**

3 Aspects of Professionalism in X: Knowledge, Activities, Behavior

Certification of Professionalism

Knowledge Generation & Dissemination: (Academia, R&D)

Science

Application Area(s)

BoK in X

Technology

Wealth Generation (Products/Services): (Industry)

Code of Professional Ethics

Professional and Ethical Conduct

Knowledge:
To solve problems:
- BoK of X
- Science
- Technology
- Application Area(s)
How to solve them (Attitude):
- **Code of Professional Ethics**

Activities:
- Knowledge Generation and Dissemination: (Academia, R&D)
- Wealth Generation (Products/Services): (Industry)

Behavior:
- Professional and Ethical Conduct
- Certification of Professionalism

Professional societies

- **Need** and often have regulations and codes to guide and regulate the conduct of their members.

- Some professional codes refer *only to ethical duties*.

- Some other codes cover *both ethical and purely professional issues* with or without referring to "ethics" in their title.

- The term "*professional ethics*" implies ethics and professional conduct.

- Hence, the term "***code of professional ethics***" is a conveniently short term to denote "*code of ethics and professional conduct*."

Some References for Codes of ethics:

- Center for the Study of Ethics in the Professions Illinois Institute of Technology (in 24 categories):

  http://www.iit.edu/departments/csep/PublicWWW/codes/codes.html

Plan:

- **Background**
  - **Ethics/philosophy, ethics/civilization**
  - **Respect for the rights of others**
- **Ethics for individuals**

- **Professional ethics**

- **Ethics for software agents**
  - **Early studies**
  - **Contemporary concerns**

Software agents:

are software modules that can work on behalf of humans, organizations, smart systems, and other agents.

are already very important and promising for several application areas.

Software agents 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.

Ethics as a Basis for Sustainable Civilized Behavior for Software Agents: **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

➢ Netiquette (Internet etiquette)

- ▪ Not a code of conduct; an understanding of what constitutes good manners and electronic community expectations.

- ▪ Applies both to humans as well as to software agents.

- ▪ Several studies elaborate on Internet agent etiquette.

Early netiquette studies for software agents concerns:

- *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.

Responsibility: The relinquished authority of the user to the software agent *redefines responsibility*:

- *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?

Contemporary concerns:

Infohabitants of the connected information systems:

- individuals,

- organisations,

- smart appliances,

- smart buildings, and

- other smart systems, as well as

- virtual entities (agents) acting on their behalf.

Hence their behavior are important for the sustainability of the overall system.

Relationship of infohabitants:

Goals of infohabitants			Type of agent / infohabitant:	Sustainability depends on:
	unrelated		agents	prevention of / protection from **accidents**
	related	**harmony / cooperation** - non-zero-sum game - optimization of collective (global) utility	cooperative agents benevolent agents holonic agents	- **trust** is essential & should not be breached - re-establishment of trust
		non cooperation / conflict - zero-sum game - optimization of individual (local) utility	self-interested agents (catallactic information systems)	**conflict resolution**
		Hostility / caution - for (possible) enemy - for even unknowns	antagonistic agents (infowar agents (time /logical bombs); viruses, worms, Trojan horses)	**survivability** - prevention / protection **from assaults**

		Coordination	Type of conflict resolution
Information system	closed	centralized	arbitration
		decentralized	auctioning
	open	decentralized	self-regulation self-organization free-market economy (catallaxy, Hayek, 1945)

We have seen:

- **Background**
  - **Ethics/philosophy, ethics/civilization**
  - **Respect for the rights of others**
- **Ethics for individuals**

- **Professional ethics**

- **Ethics for software agents**
  - **Early studies**
  - **Contemporary concerns: infohabitants**

A Quotation as a Conclusion:

`Would you tell me, please, which way I ought to go from here?'
`That depends a good deal on where you want to get to,' said the Cat.

`I don't much care where--' said Alice.
`Then it doesn't matter which way you go,' said the Cat.

`--so long as I get SOMEWHERE,' Alice added as an explanation.
`Oh, you're sure to do that,' said the Cat, `if you only walk long
   enough.'

(Lewis Carroll, Alice in Wonderland, 1872).