ZÁPADOČESKÁ
UNIVERZITA
V PLZNI

University of West Bohemia in Pilsen
Department of Computer Science and Engineering
Univerzitní 8
306 14 Pilsen
Czech Republic

Modelling and Simulation of Trust Evolution

The State of the Art and Concept of PhD. Thesis

Arnoštka Netrvalová

Technical Report No.DCSE/TR-2006-02
May, 2006

Distribution: public

Technical Report No. DCSE/TR-2006-02
May, 2006

Modelling and Simulation of Trust Evolution

Arnoštka Netrvalová

Abstract

Several studies coming from psychological or social sciences examine the meaning and characteristics of trust. The information technologies significantly influence almost all activities in modern society. This increases the need of studying the concept of trust in the environment of information systems. On the other hand, the methods and tools coming from computing sciences are used for trust modelling and simulation.

This study discusses possible appropriate concepts of modelling and simulation of the trust evolution. The terms trust, and trust representation and visualization are introduced. The approach in human trust modelling is based on the theory of information and social communication knowledge. Some terms of probability theory and information theory are presented. The principle of some phenomena dissemination is demonstrated by epidemic algorithms.

The multi-agent system is the modern technology that is used for modelling of not only the technical systems but also the social ones. Some needed terms from the agent theory are mentioned. The agent approach is chosen for modelling the trust in a community. Fundamentals of trust formation, trust dissemination, and trust evolution are presented by deploying the agent system. Finally, summary of well known tools for modelling and simulation multi-agent systems is presented. Finally, an outlook for the future work is proposed.

Copies of this report are available on
http://www.kiv.zcu.cz/publications/

or by surface mail on request sent to the following address:

University of West Bohemia in Pilsen
Department of Computer Science and Engineering
Univerzitní 8
306 14 Plzeň
Czech Republic

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACL	AGENT COMMUNICATION LANGUAGE
AS	AGENT SYSTEM
BDI	BELIEFS, DESIRES, INTENTIONS
DAI	DISTRIBUTED ARTIFICIAL INTELLIGENCE
FIPA	FOUNDATION FOR INTELLIGENT PHYSICAL AGENT
JADE	JAVA AGENT DEVELOPMENT FRAMEWORK
JAM	JAM INTELLIGENT AGENT ARCHITECTURE
KQML	KNOWLEDGE QUERY MANIPULATION LANGUAGE
MAS	MULTI-AGENT SYSTEM
P2P	PEER-TO-PEER
RETSINA	REUSABLE ENVIRONMENT FOR TASK-STRUCTURED INTELLIGENT NETWORKED AGENTS
TMF	TRUST MANAGEMENT FRAMEWORK
ZEUS	ZEUS AGENT BUILDING TOOL-KIT

Acknowledgements

I would like to thank to my supervisor, professor Jiří Šafařík, for his guidance of the interesting and actual problem of the trust modelling, and for his time, which he has spent with me.

My thanks also belong to doc. František Vávra for his support and help in my particular coordination.

In addition, I would like to thank to my colleagues and friends for their valuable comments, which contributed to the research within this topic.

1 Introduction

There are already several studies coming from psychological or social sciences that are examining the meaning and characteristics of trust. However, just a few works deal with examination of the trust evolution. The method of computational simulation is applied in these works. It will be used also in this study.

1.1 Brief Preview

Trust is a common phenomenon. People (as human entities) and some animals generally help each other hoping that in the future, when they need help, they will be helped back, but this hope is realised with a risk. Moreover, the time matters in the real world and things change over time. This risk of one entity depends on the behaviour of another entity and environment in most of situations.

In the modern world, artificial intelligence societies exist but in a limited fashion. The examples of such an artificial society are Phone network or the Internet, multi-agent systems included. It is needed to understand why some things happen in societies, and the power that group action can give.

The trust plays constantly growing role in constructing, testing, using and securing of those systems at present.

1.2 Aspects of Trust

Trust is a fact of everyday life and plays an explicit role in societies. We all make trust decisions, most of us every day of our lives, and many times per day (Luhmann N., Trust and Power, 1979)[1]. The decision to trust is based on evidence to believe, or be confident in, someone something's good intentions towards us (Yamamoto Y., A Morality Based on Trust, 1990).

How can we see aspects which have effect on these societies? Firstly, there are social, biological and technological aspects.

We attend to social aspects. Firstly, there are groups, their organisation and existence. The group organisation in itself is a complex area and more items have effect on trust. Secondly, the fact of working together with others is how much they need to trust to each other. Working together implies one of the forms - cooperation, coordination, and collaboration and generally, communication. Trust can only concern that which person can rightly demand of another (Hertzberg L., On the Attitude of Trust, 1988).

Further we discuss biological aspects. Is the trust behaviour a uniquely human phenomenon? The answer to that we can found in animal world (Harcourt A. H., Cooperation and Trust in Animals, 1991). Some animals help each other, e.g. chimps help each other in fights. Trust is more likely to be present, since the intellectual capacity of primates is greater than of many other animals. Animals help those who are related to themselves because it perpetuates their genes and chances of survival (Richard Dawkins, The Gene as the Unit of Selection, 1989).

The last one aspect is technological (artificial) aspect. The use of computers enables the growth of artificial societies in the modern world. Telephone network consists of many

[1] Only in this chapter some ideas are taken over from [Marsh] and the original work is cited in the parenthesis.

nodes, each of some intelligence, each deciding which way to route traffic (phone calls, faxes, etc.). These nodes working with each other means that trust is present (MARSH).

Many authors approached trust of many points of view, through many disciplines and academic domains, thus giving birth to the many useful observations and even theories. T^3 Group - Trust: Theory and Technology (T3GRO) tries to organize this huge corpus of contributions and kinds of research. Their table of Trust across the disciplines is in the Appendix. The theory about trust and its related topics can be found on T^3 Group home pages in (T3GRO).

1.3 Work Overview

This study discusses possible concepts appropriate for modelling and simulation of the trust evolution. Firstly, basic terms like trust, trust representation and visualization are discussed. The approach to human trust modelling is based on information theory and social communication knowledge. Some terms of probability theory (e.g. entropy) and information theory are introduced. The epidemic algorithms are presented as an example of phenomena dissemination. Secondly, multi-agent system technology is described. The needed terms of the agent theory are mentioned. The agent approach is chosen for modelling the trust in a community. Fundamentals of trust formation, trust dissemination, and trust evolution are presented by deploying the agent system. Finally, summary of well known tools for modelling and simulation of multi-agent systems is presented.

2 Trust

In this chapter, the term trust, its representation and visualization, trust as a property of relationship, and trust evolution are analyzed. The acceptation of the term trust is wide. The World Book Dictionary [(WBD88)] offers further explanations, as:

1 firm belief in the honesty, truthfulness, justice, or power of person or thing (faith: A child puts trust his parents.) **- 2** a person or thing trusted (God is our trust.) **- 3** confident expectation or hope (Our trust is that she will soon be well.) **- 4** something that is managed for the benefit of another; something committed to one's care (The house is a trust which he holds for his dead brother's children.) **- 5** the obligation or responsibility imposed on one in whom confidence or authority is placed (He will be faithful to his trust.) **- 6** condition of one in whom trust has been placed; being relied on (A guardian is in a position of trust.) **- 7** keeping; care (The farm was left in the caretaker's trust.) **- 8** confidence in the ability or intention of a person to pay at some future time for goods or services; business credit **- 9** law; **- a** confidence reposed in a person by making him nominal owner of property, which he is to hold, use, or dispose of for the benefit for another **- b** an estate or other financial holding committed to a trustee or trustees **- c** the right of a person to enjoy the use or profits of property held in trust for him **- 10** an illegal combination of businessmen or companies having a central committee to control the production and price of some commodity and to eliminate or reduce competition: a steel trust.

Urbánek refers to this term in his work [(URB04)]: "I trust you. It is a phrase. There are people you say that you trust them or do not trust them. However, what does it mean? Merriam-Webster Thesaurus dictionary uses a number of definitions to explain the word "trust". It says that "trust" is: a credit given; especially, delivery of property or merchandise in reliance upon future payment or exchange without immediate receipt of an equivalent. It can be considered trust as behavioural pattern in some system".

2.1 Definition and Properties of Trust

One of the first definition of the trust was formulated by Morton Deutsch in [(DEUT)]. The definition states: "Trusting behaviour occurs when an individual perceives an ambiguous path, the results of which could be good or bad, and the occurrence of the good or bad result is contingent on the actions of another person; finally, the bad result is more harming than the good result is beneficial. If the individual chooses to go down that path, he can be said to have made a trusting choice, if not, he is distrustful".

The similar definition was presented by Golembiewski and McConkie [(GOCO)], "the loss or pain attendant to unfulfilment of the trust is sometimes seen as greater then the reward or pleasure deriving from fulfilled trust. Trust implies some degree of uncertainty as to outcome. Trust implies hopefulness or optimism as to outcome."

As a basic point we review Gambetta's definition of trust[2]:

Trust (or symmetrically, distrust) is a particular level of the subjective probability with which an agent will perform a particular action, both before we can monitor such an action (or independently of our capacity of ever to be able to monitor it) and in a context in which it affects our own action.

[2] Gambetta's definition was derived as a summary of the contributions to the symposium on trust in Cambridge, England, 1988.

We can find other trust description, but we can also say that any universal definition of this term does not exist. The bulk of authors build up their own description or definition of the trust.

Alfarez (ALFAR) in his dissertation summarises the properties of trust as follows:

- Trust is subjective.
- Trust is situation specific.
- Trust is agent specific.
- Trust is not absolute; it exists as levels of trust.
- Trust involves expectations of future outcome.
- Situation of trust can result in positive or negative outcomes, thus involves risk, uncertainty and ignorance.
- Trust gives control to the trustee and an opportunity to betray to truster.
- The inability to verify one's action until after the action has completed requires trust in the trustee prior to the action being taken.
- Trustees are active agents that have the ability to perform with a degree of independence from the truster's control.
- Trust in not a prediction.
- Trust is not transitive.

Briefly we will introduce the contrast between trust and some similar terms. The difference between trust and *confidence* is that trust involves choosing between alternatives while confidence does not. *Reliance* on something or someone is not necessarily an indication of trust. We may rely because we have to or that it is the best to us. Trust also differs from *hope*, because we hope that risk action will result in something satisfactory. The *belief* is acceptance of something as truth. *Trustworthiness* is reputation for being worthy of a certain level of trust in a given situation. To *distrust* is to take an action as if the other one is not trusted with respect context. To distrust is different from having no opinion, it is *ignorance*. *Mistrust* comes in, when a trustee betrays the trust of the truster.

Taking the main social aspects of the definitions above, we can propose our short simple definition of the trust:

> *The trust in an entity is a commitment to an action based on a belief that the future actions of that entity will be make for a good outcome.*

2.2 Representation of Trust

Furthermore we can put a question. Can trust be measured? It is expected that it can. However it has to be done using some simplifications and limiting presumptions. For examining the trust as a behavioural pattern, some ways of representing and visualizing it must be known. It is possible to create some methods that can measure the trust. Some tools that can do the visualization are also constructed.

Trust is a very hazy term. It is not so much inter-subjective as it is widely understood. Its indeterminateness should be taken into account, when we try to represent the trust as a value. Modifying Marsh's way (MARS) of representation, we treat the trust as a value between 0 and 1, where 0 means the complete distrust and 1 means the "blind trust".

The interpretation of the trust value is very model dependent. Generally, the values can be interpreted as in the figure below. A single trust value can be visualized as a point on the line between point 0 and 1.

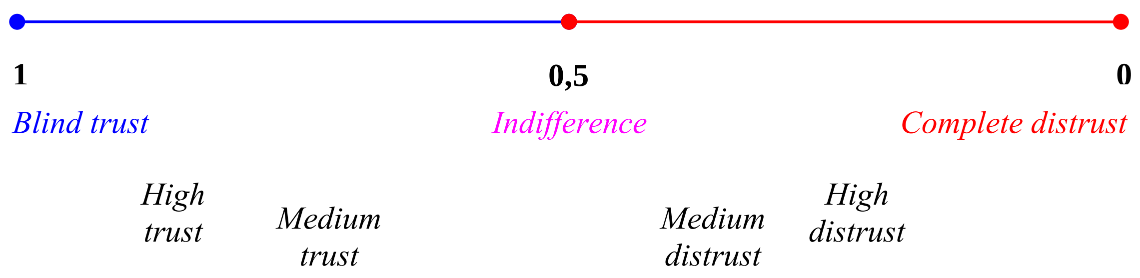

Figure 1 Representation of the trust – the trust value is visualized as a point on the interval <0, 1>

2.3 Trust as a Property of Relationship and its Visualisation

The trust value, when alone, tells us nothing interesting. Predicable like "I trust you" or "he trusts me" suggests us that the trust is usually measured or shared between two or more entities. We can say that it is a property of relationship between entities.

If we look at a community of entities from the point of view of trust then community has to be composed of couples of entities. The trust is usually shared between two or among more entities in some community. Then, we may say that it is a property of the relationship between entities. For the reason of simplicity, let us consider the community of entities to be composed from the couples of single objects. Let us consider a couple with two relationships and one trust value per relationship. In the work [URB03] these trust values are denominated T_L and T_R meaning trust from left to right and from right to left respectively. A square can be drawn in a two dimensional coordinate system (Figure 2). The trust values T_L, T_R are projected onto the two perpendicular sides of the square. Thus, the trust between the objects of the couple may be treated as a two-dimensional vector (T_L, T_R). It is a point in the square, thus visualizing the trust in the couple. It is very simple visualization; therefore it is easily and quickly readable.

If we go further, as Urbánek introduces in his work [URB04], we will see that there can be lots of trust values in a single relationship. For example: "how much do I trust him", "how much I think, he trusts me", "how much I think that how much he thinks that how much I trust him", etc.

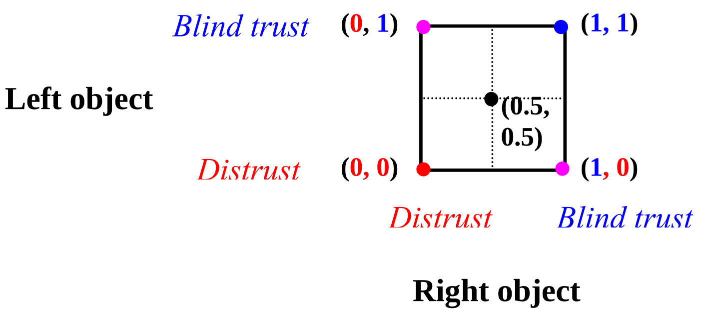

Figure 2 Trust square - very simple trust visualization of couple of entities

There are (URB04) nine basic shapes of trust squares in Figure 3. They visualize either relationships in a single couple or relationships in a population.

Shape 1 denotes a couple with mutual distrust, shape 2 denotes mutual trust. Shapes 3 and 4 represent couples where one entity trusts the other one and the other entity distrusts completely the first one. Shapes 5 and 6 show the situation where one entity trusts and the other one is indifferent. The opposite situation is typified by the shapes 7 and 8 where one entity is indifferent and the other one distrusts the first one. The shape 9 denotes that both entities are indifferent to each other, or that there is no relationship between them.

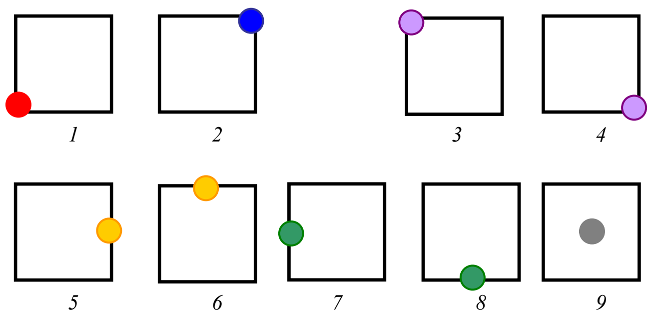

Figure 3 Basic trust square shapes - simple level of trust visualization

When there is a point on the diagonal from the lower left corner to the upper right corner, then we say that the relationships are equal. When the point is close to the diagonal, then we say that the relationships are almost equal. The upper left corner and the lower right corner represent couples with completely unequal, or opposite, trust values.

The square with a single vector gives us some information about the couple. A view of all couples in a population is interesting. We can draw a trust square image by plotting all relationships in a population into the image. We get a square with stippled dots forming brighter and darker regions.

2.4 Reciprocation, Trust and Evolution

What does the term evolution mean? Encyclopaedia Britannica presents the following definition. "Evolution is the theory in biology postulating that the various types of plants, animals, and other living things on the Earth have their origin in other pre-existing types and that the distinguishable differences are due to modifications in successive generations. The theory of evolution is one of the fundamental keystones of modern biological theory".

The process of evolution may be interpreted as a biological term, but also as cultural one or even as a process of learning.

In this context, Marsh introduces following thoughts in his work (MARSH). Interactions between entities proceed in the society quickly. The entities meet more than once, if the society is small, or if they act in the same area. In this case, it is an evolutionary strength to reciprocal cooperation. An individual in the society "will do better" if he supports others, than if he did not. In the society, the reciprocation is a common form of behaviour. Reciprocal altruism is a form of the trust in the world of animals. If trust were not present, reciprocation would not occur, and collective collaboration would be worse off. Thus, there is the meaning that the reciprocation is more likely in trust relationships and reciprocation is also good in evolutionary terms.

2.5 Bounded Rationality

Bounded rationality is important, because entities dealing with one another are not quite rational. Rational behaviour is a property of usually large system. Rationality emerges in an evolutionary process where non rational behaviour leads entity to death and wherefore removes it from the population. If all entities that are not behaving rationally are removed, it seems that system is behaving rationally. But it does not mean that there are no entities behaving differently[UCALCI]. We can say that entities have a bounded rationality.

Simon [SIM79] defines bounded rationality as: "Rationality is bounded when it falls short of omniscience. And the failures of omniscience are largely failures of knowing all the alternatives, uncertainly about relevant exogenous events, and inability to calculate consequences."

We can ask how to model the bounded rationality. We can put a question, what is a model of uncertainty. Can we do the model simply? Entities do not know what can influence their behaviour. It can be a state of their mind, unexpected event in their surroundings, etc. For a distant observer, it can be a random cause for (more or less) unpredicted behaviour. The easiest way to model random causes is to use random number generators.

3 Communication as Information Process

Some necessary terms are described in this chapter. These terms are required for an information description (e.g. entropy) and some further aspects of communication as information process.

Jointly, information can be perceived as a real environment specification, about its state and about running processes. Information decreases, increases, eliminates or disseminates uncertainty of a system (e.g. information receiver). The information quantity is given as a residual between the state of system uncertainty (entropy) before and after receiving the information.

In computer science and cybernetics, we understand the information as a psychological or physiological phenomenon. For instance, Norbert Wiener defines information as an exchange with outward, when we adapt to and our adaptation acts upon it, too.

In Shannon's approach [SHANN], measuring of information is based on the difference between an expectation and a reality. The more "difference" between the expected message (announcement) and the received message the more information we get. It is given by measure of surprise. The more suddenly message is the more information it yields. The less probable announcement was received the greater information was obtained.

3.1 Message

The fundamental problem of communication is that of reproducing at one point either exactly or approximately a received message sent from another point. Frequently, the messages have importance. They refer to or are correlated according to some system with certain physical or conceptual entities. These semantic aspects of communication are sometimes irrelevant to engineering problems. The significant aspect is that the actual message is one selected from a set of possible messages. The system must be designed to operate for each possible selection (no just the one which will actually be chosen since it is unknown at the time of design).

If the set of messages is finite, then their number or any monotonic function of this number can be regarded as a measure of the information. It is produced when one message is chosen from the set. As it was pointed out by Hartley, the most natural choice is the logarithmic function. This definition must be generalized considerably when we consider the influence of the statistics of the messages. When we have a continuous range of messages, we will use a logarithmic measure in all cases.

The logarithmic measure is more convenient for various reasons:

1. It is practically useful. Parameters of engineering importance tend to vary linearly with the logarithm of number of possibilities.
2. It is nearer to our intuitive feeling as to the proper measure. It is related to the previous, since we intuitively measure entities by linear comparison with joint standards.
3. It is mathematically more suitable. Many of the limiting operations are simple in terms of the logarithm (because we are able to work with them) but would require clumsy restatement in terms of number of possibilities.

3.2 Measure of Information

We can define a quantity which will measure, i.e. at what rate information is produced. Suppose we have a set of possible events whose probabilities of occurrence are $p_1, p_2, \ldots, p_n$. These probabilities are known, but that is all we know concerning which event will occur. Can we find a measure of how much choice is involved in the selection of the event or of how uncertain we are above the outcome?

If there is such a measure, $H(p_1, p_2, \ldots, p_n)$, it is reasonable to require the following properties:

1. H should be continuous in the p_i.
2. If all p_i are equal, $p_i=1/n$, then H should be a monotonic function increasing of n.
3. If a choice is broken down into two successive choices, the original H should be a weighted sum of the individual values of H.

Then we can define the form:

$$H = -K\sum_{i=1}^{n} p_i \log p_i \tag{3.1}$$

where K is a positive constant.

Quantities of this form play a key role in information theory as measures of information. We shall call H the entropy of the set of probabilities $p_1, p_2, \ldots, p_n$. The details of these problems are described e.g. in the work [POFA04].

3.3 Communication System

In communication conception, a message is transferred between an information source and a receiver subject by a data channel. The form of carrier is a signal. The message can (but need not) produce some information. The gist of information processes consists of transfer and transformation of messages. The messages are transmitted in a coded form. The code represents the transformation rules for obtaining the unique values. The receiver must be able to decode the signal. The receiver must know the message language otherwise the communication is not available.

Generally in engineering, by the communication system we mean the system of the type featured schematically in Fig. 4.

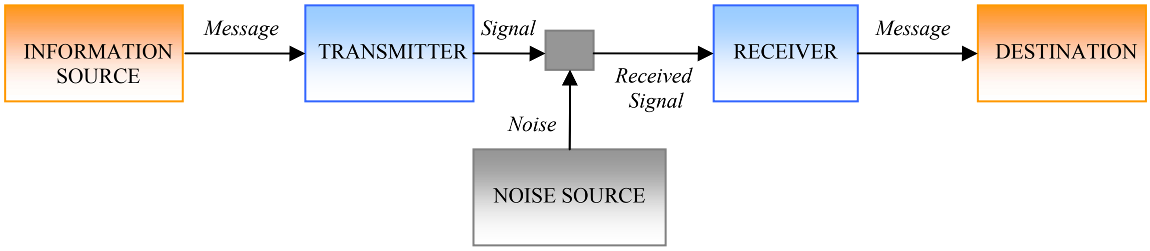

Figure 4 Schematic diagram of general communication system

Essentially, it consists of five parts:

1. An *information source* which produces a message or a sequence of messages.
2. A *transmitter* which operates on the message in some way to produce a signal suitable for transmission over the channel.
3. The *channel* is merely the medium used to transmit the signal from the transmitter to the receiver.
4. The *receiver* ordinarily performs the inverse operation of that done by the transmitter, reconstructing the message from the signal.
5. The *destination* is the person (or thing or subject) for whom the message is intended.

3.4 Principles and Types of Social Communication

Social communication can be defined as the information transfer or the purport change in any social contact. It solves certain technical, semantic and pragmatic problems. For amusement, we can insert the sentence of Harold D. Lasswell (a man of American politics, 1948), that formulated communication process. "Who says what to whom, how and with what effect?"

The elements of social communication system and their equivalents in Shannon-Weaver model are shown in Table 1.

Social information barrier is the set-back in appeasement of information needs. Types of information barrier are time, locality, information competence, real competence and information overloading.

Table 1 Elements of social communication system

Social communication system	***Shannon –Weaver's model***
communicator (source)	coder and transmitter
communicating intent	---
communication	message
sense of announcement for communicator	---
medium	channel
sense of announcement for communicant	---
communicant (recipient)	decoder and receiver
effect of announcement to communicant	---
information barrier	noise
reaction	reaction coupling

We distinguish many various types of social communication.

1. *Direct* (immediate) and *indirect* (intermediate by communication system), where communicant is separated from communicator by space and time (see Fig. 5).
2. *Unilateral* (directed to receiver) and *double sided* (exchange of the role of communicator and communicant)
3. *Verbal* (speech) and *nonverbal* (gesture, mimic)
4. *Formal* (with some rules) and *informal* (without schedules)
5. *Local* (transfer on short distance) and *remote* (telecommunication)
6. *Intrapersonal* (all by himself), *interpersonal* (among entities), *public* (to any group) and *mass* (to mass of entities)
7. *Addressing* (familiar receiver) and *without addressing* (unknown receiver)

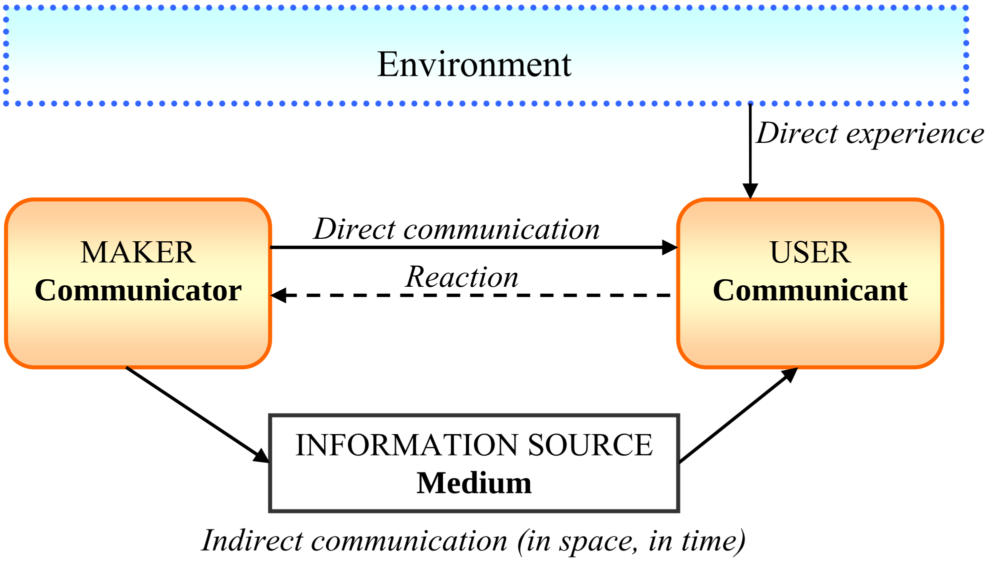

Figure 5 Types of social communication - direct and indirect communication

There are more specific types of social communication as are information propaganda, misinformation, infotainment, edutainment, communication without identifying (e.g. chat), etc. More about social communication is in (DIZ03).

4 Entropy and Information

How to describe the information as a measurable value? We can use the probabilistic and statistic approaches. Information theory applies many methods of the theory of probability.

The concept of information is too broad to be captured completely by a single definition. However, for any probability distribution, we define the quantity called the entropy. It has many properties that agree with the intuitive notion of what a measure of information should be. This notion is extended to define mutual information, which is a measure of the amount of information that one random variable contains about another. Entropy then becomes the self-information of a random variable. Mutual information is a special case of the more general quantity called relative entropy, which is a measure of the distance between two probability distributions. All these quantities are closely related and share a number of properties. We describe some of these properties in this chapter. More of this topic you can found in the introductory works [(COTHO), (RENYI)]. More of disinformation is in the following works [(ROBU04), (VAN04) and (ROBU06)].

4.1 Entropy

As we have mentioned, entropy is a measure of uncertainty of a random variable. Let X be a discrete random variable with the alphabet $\mathcal{X}$ and the probability mass function $p(x)$, $x \in \mathcal{X}$.

Definition: The entropy $H(X)$ of a discrete random variable X is defined by

$$H(X) = -\sum_{x \in X} p(x) \log p(x) \tag{4.1}$$

See (3.1) and more information about this topic can be found in [(SHANN), (POFA04)].

Note that entropy depends on the distribution of X. It does not depend on the actual values of a random variable X, but only on their probabilities.

4.2 Joint Entropy and Conditional Entropy

Now we extend the entropy definition to a pair of random variables.

Definition: The joint entropy $H(X,Y)$ of a pair of discrete random variables (X,Y) with a joint distribution $p(x, y)$ is defined as

$$H(X,Y) = -\sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y), \tag{4.2}$$

which will be also expressed as

$$H(X, Y) = -E \log p(x,y). \tag{4.3}$$

We define the conditional entropy of a random variable as the expected value of the entropies of the conditional distributions, averaged over the conditioning random variable.

Definition: If $p(x, y)$ is the joint distribution of discrete random variables (X, Y), the conditional entropy $H(X|Y)$ is defined as

$$H(Y|X) = \sum_{x \in X} p(x) H(Y|X=x) \tag{4.4}$$

$$= -\sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \log p(y|x) = -\sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(y|x) = -E_{p(x,y)} \log p(y|x) \tag{4.5}$$

4.3 Relative Entropy and Mutual Information

The entropy of a random variable is a measure of the uncertainty of the random variable. It is a measure of the amount of information required on average to describe the random variable. In this section, we introduce two related concepts: relative entropy and mutual information.

The relative entropy is a measure of the distance between two distributions. The relative entropy $D(p||q)$ is a measure of the inefficiency on the assumption that the distribution is q when the true distribution is p.

Definition: The relative entropy (or divergence) between two probability mass functions $p(x)$ and $q(x)$ is defined as

$$D(p||q) = \sum_{x \in X} p(x) \log (p(x)/q(x)) = E_p \log p(x)/q(x) \tag{4.6}$$

Note that $D(p||q) \neq D(q||p)$ in general.

It is often useful to understand the relative entropy as a "pseudo-distance" (do not fulfill triangular inequality) between two distributions.

Now we introduce the mutual information, which is a measure of the amount of information that one random variable contains about another random variable.

Definition: Consider two random variables X and Y with a joint probability mass function $p(x, y)$ and marginal probability mass functions $p(x)$ and $p(y)$. The mutual information $I(X; Y)$ is the relative entropy between the joint distribution and the product of distributions $p(x)p(y)$, i.e.

$$I(X:Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} = D(p(x,y) || p(x)p(y)) = E_{p(x,y)} \log \frac{p(x,y)}{p(x)p(y)} \tag{4.7}$$

4.4 Information and Disinformation

Some unexpected situations with statistical estimations of mutual information are studied e.g. in following works (VAJDA), (COTHO), (CZISZ), (ZHAI). As we introduced above, the mutual information can be described

$$I(X:Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} . \tag{4.8}$$

Because we do not know the $p(x)$, $p(y)$, $p(x,y)$, we must work with their estimations: $e(x)$ is the estimate for $p(x)$, $e(y)$ is the estimate for $p(y)$, $e(x, y)$ is the estimate for the joint distribution $p(x,y)$, or use some of their parametric representation which is also an estimation.

When we respect the reality that we do not have an available correct model (abide by the observed values), we will get for the estimation of $I(X: Y)$

$$\hat{I}_n(X:Y) = \frac{1}{n} \sum_{i=1}^{n} \log \frac{e(x_i, y_i)}{e(x_i)e(y_i)} . \tag{4.9}$$

This estimation can converge to $(n \to \infty)$

$$\hat{I}_n(X:Y) \to \sum_{y \in Y} \sum_{x \in X} p(x,y) \log \frac{e(x,y)}{e(x)e(y)} = I(X:Y;e) , \tag{4.10}$$

where $p(x, y)$ is the real probability and $e(x, y)$ is its model.

Particular description of this problem is in the work (VANO4), where is also introduced the derivation of description of disinformation rate as

$$DI(X:Y;e) = I(X:Y;e) - I(X:Y) \tag{4.11}$$

This reality motivates us to use the entropy and the divergence for measuring of disinformation.

4.5 Duality of Classic Information and Disinformation

Described measures of information have supposed a real distribution of probability. The dual disinformation distribution assumes that heading distribution of probability is not available. In this case, we must work with its model that can be different from real situation. Comparative situation in which the model and reality are different is in Table 2. Heading distribution will be denoted as $p(x)$ (respectively $p(x, y)$), its model (the estimation) as $e(x)$ (respectively $e(x, y)$) and the comparative probability as $q(x)$.

Table 2 Shannon's classical theory in comparison with the concept of disinformation

MEASURE	SHANNON'S, CLASSICAL	CONCEPT OF DISINFORMATION
Entropy	$H(X) = -\sum_x p(x) \log p(x)$	$H(X;\mathrm{e}) = -\sum_x p(x) \log e(x)$
Mutual information	$I(X:Y) = \sum_x \sum_y p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$	$I(X:Y;e) = \sum_x \sum_y p(x,y) \log \frac{e(x,y)}{e(x)e(y)}$
Divergence of probability models	$D(p\Vert q) = \sum_x p(x) \log \frac{p(x)}{q(x)}$	$D(p\Vert q;e) = \sum_x p(x) \log \frac{e(x)}{q(x)}$
Symmetrical divergence of probability models	$J(p\Vert q) = \sum_x (p(x) - q(x)) \log \frac{p(x)}{q(x)}$ $J(p\Vert q) = D(p\Vert q) + D(q\Vert p)$	$J(p\Vert q;e) = \sum_x (p(x) - q(x)) \log \frac{e(x)}{q(x)}$ $J(p\Vert q;e) = D(p\Vert q;e) + D(q\Vert e)$

It is possible to use $p(x)$ as a model of the temporary situation and $e(x)$ as a model of new state after spreading some message, e.g., as a model of trust dissemination.

4.6 Information Control Model

The model of information control (ROBU06) is shown in Figure 6. This model is a transmitting channel that has the same input and the output alphabet. The alphabet X with the probability distribution $p(x)$ is on the input which is joined with the noise with the alphabet X with the probability distribution $r(x)$. The alphabet X with the probability distribution $q(x)$ is on the output. The interference of the input signal (X, p) by the control signal (X, r) to the output signal (X, q) is measured as a symmetrical divergence $J(p||q)$.

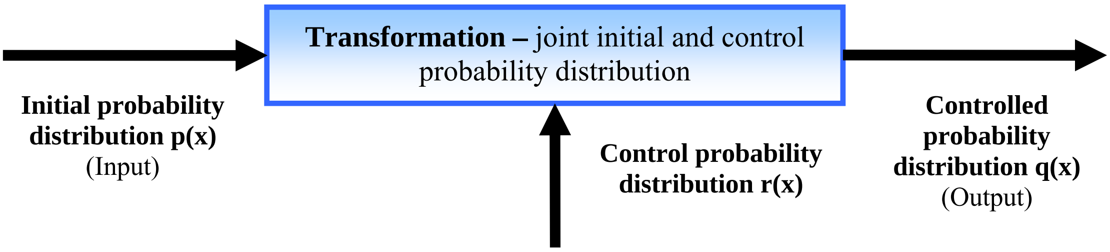

Figure 6 The model of information control

We can adjust to finish this measure by the following:

$$J(p\| q)=\sum_{x\in X}(p(x)-q(x))\log\frac{p(x)}{q(x)}=\sum_{x\in X}(p(x)-r(x)+r(x)-q(x))\log\frac{p(x)}{q(x)}\frac{r(x)}{r(x)}=$$

$$=\sum_{x\in X}((p(x)-r(x))-(q(x)-r(x)))\left(\log\frac{p(x)}{r(x)}-\log\frac{q(x)}{r(x)}\right)=$$

$$=\sum_{x\in X}(p(x)-r(x))\log\frac{p(x)}{r(x)}-\sum_{x\in X}(p(x)-r(x))\log\frac{q(x)}{r(x)}-$$

$$-\sum_{x\in X}(q(x)-r(x))\log\frac{p(x)}{r(x)}-\sum_{x\in X}(q(x)-r(x))\log\frac{q(x)}{r(x)}=$$

$$=J(p\| r)+J(q\| r)-\sum_{x\in X}(p(x)-r(x))\log\frac{q(x)}{r(x)}-\sum_{x\in X}(q(x)-r(x))\log\frac{p(x)}{r(x)}=$$

$$=J(p\| r)+J(q\| r)-J(p\| r;q)-J(q\| r;p) \qquad (4.12)$$

The final form is

$$J(p\,\|\,q)=[J(p\,\|\,r)-J(p\,\|\,r,q)]+[J(q\,\|\,r)-J(q\,\|\,r,p)],\ \text{where} \qquad (4.13)$$

$$J(p\| r;q)=\sum_{x\in X}(p(x)-r(x))\log\frac{q(x)}{r(x)} \qquad J(q\| r;p)=\sum_{x\in X}(q(x)-r(x))\log\frac{p(x)}{r(x)}. \qquad (4.14)$$

The decomposition of "distance = difference" $J(p\| q)$ to both of terms is divided into two parts (the first term $[J(p\| r)-J(p\| r;q)]$ and the second one $[J(q\| r)-J(q\| r;p)]$). The first of both is interference of input reality by the control impulses, i.e. "creation of the information bubble" and in the second one this information bubble is compared with the reality. It can act positively (strengthening) or negatively (correction, i.e. complete or partial reduction). Each of the terms has two components. First one is a symmetrical divergence between input (output) and the active incidence. The second one is a disinformation (information) correction. The former term in both differences is an idealized "distance" between the input (output) and the incidence, the later term is actually a model of the second side reaction, i.e. output on control and input or input on control and output.

We cannot help remarking what the apparatus of information theory (in classical Shannon's version) is able to do. It is convenient for measuring, quantifying and evaluation. Classical theory does not involve the orientation. Mutual information is symmetrical, it does not discern between the input (cause) side and the output (consequence) side. Nevertheless, the classical information theory is able to represent such systems. But the results demonstrate some relationships (binding rate, interconnect), no flux, i.e. running from anywhere to anywhere.

4.7 Demonstration Examples

Two examples present the technique introduced above. They are the examples of recognition the result of fictive aggressive advertisement (puffery). The first one, relatively neutral, where it did not come about the essential interference, i.e. no trust turn, and the second one, successful, where it came about behaviour change, i.e. the trust interference.

The probability distribution $p(x)$ is the model of the market shares before the advertisement; the distribution $r(x)$ is the model of the market shares which is expected by the advertisement. The distribution $q(x)$ is the model of the market shares after advertisement release.

Figure 7 shows the first example, when the neutral advertisement took effect. The probabilities of product *A*, i.e. the probability of purchase ahead of advertisement *p(i)* changed from the value 0,25 to the probability of purchase after advertisement *q (i)* = 0,35 (blue) only. The probabilities of other products stood the same or decreased (especially product *D*). The market did not accept the incidence of advertisement.

Product	Probability of purchase ahead of advertisement	Probability of purchase pretended by advertisement	Probability of purchase after advertisement	Terms J(p‖r) and J(p‖r)	Terms J(p‖r;q) and J(p‖r;q)	Difference J(p‖r)-J(p‖r;q)	Terms J(q‖r) and J(q‖r)	Terms J(q‖r;p) and J(q‖r;p)	Difference J(q‖r)-J(q‖r;p)	Terms J(p‖q) and J(p‖q)
	p(i)	r(i)	q(i)							
A	0,2500	0,9000	**0,3500**	1,2012	0,8857		0,7494	1,0164		0,0485
B	0,1667	0,0250	0,1500	0,3877	0,3662		0,3231	0,3421		0,0025
C	0,1250	0,0250	0,1200	0,2322	0,2263		0,2150	0,2206		0,0003
D	0,4000	0,0250	0,3000	1,5000	1,3444		0,9859	1,1000		0,0415
E	0,0583	0,0250	0,0800	0,0407	0,0559		0,0923	0,0672		0,0099
				3,3619	2,8785	**0,4834**	2,3657	2,7463	**-0,3807**	**0,1027**

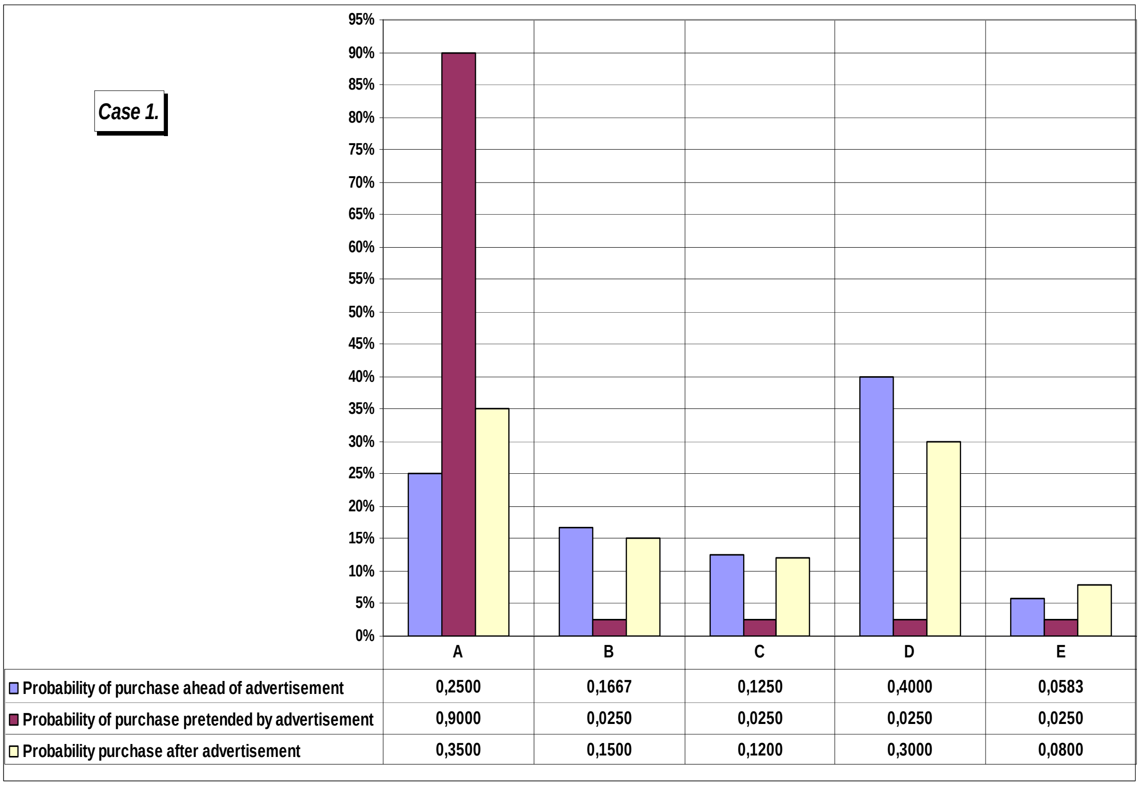

Figure 7 The case of neutral advertisement

Figure 8 illustrates the second example, when the aggressive advertisement took effect. The probability of purchase ahead of advertisement *p(i)* of the product *A* increased from the value 0,25 to the probability of purchase after advertisement *q (i)* = 0,55 (blue). The probabilities of other products decreased (especially product *D* over again). The market accepted the incidence of advertisement.

Product	Probability of purchase ahead of advertisement	Probability of purchase pretended by advertisement	Probability of purchase after advertisement	Terms J(p‖r) and J(p‖r)	Terms J(p‖r;q) and J(p‖r;q)	Difference J(p‖r)-J(p‖r;q)	Terms J(q‖r) and J(q‖r)	Terms J(q‖r;p) and J(q‖r;p)	Difference J(q‖r)-J(q‖r;p)	Terms J(p‖q) and J(p‖q)
	p(i)	*r(i)*	*q(i)*							
A	0,2500	0,9000	**0,5500**	1,2012	0,4618		0,2487	0,6468		0,3413
B	0,1667	0,0250	*0,1000*	0,3877	0,2833		0,1500	0,2053		0,0491
C	0,1250	0,0250	*0,1200*	0,2322	0,2263		0,2150	0,2206		0,0003
D	0,4000	0,0250	*0,1800*	1,5000	1,0680		0,4414	0,6200		0,2534
E	0,0583	0,0250	*0,0500*	0,0407	0,0333		0,0250	0,0306		0,0019
				3,3619	2,0728	**1,2891**	1,0801	1,7232	**-0,6431**	**0,6460**

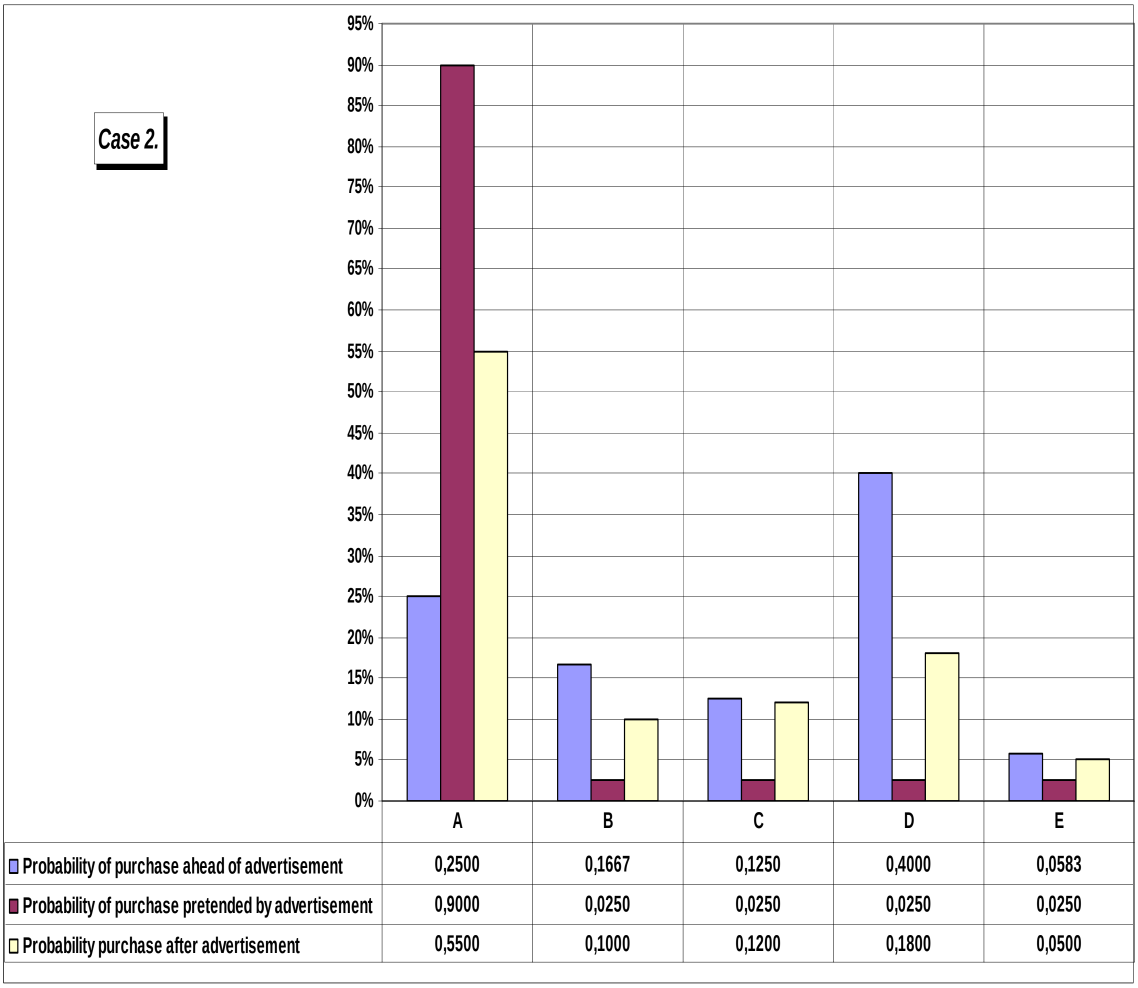

Figure 8 The case of aggressive advertisement

Both examples represent the following situation. The bubble is injected by a control action. It is consecutively corrected by interaction with the environment, in which it took effect. Analogous mechanisms perceptibly operate also on the stock market.

5 Dissemination and Epidemic Algorithms

One way, how to describe the information dissemination, is an application of the processes seen in the nature. For example, it can be the dissemination of epidemic diseases for which many epidemic algorithms have been constructed, e.g. [EGKM]. They have recently gained popularity as a potentially effective solution for disseminating of information in large-scale systems. In addition to their inherent scalability, they are easy to deploy, robust, and resilient to failure.

Epidemic algorithms mimic the spread of a contagious disease - infected individuals pass on a virus (or a microbe) to those with whom they come into contact. Similarly, each process in a distributed receives new information from randomly chosen peers. In turn, each of these processes forwards the information to other randomly selected processes, and so on.

Once an epidemic has started, it is hard to eradicate. It only takes a few people to spread a disease, directly or indirectly, to the community at large. An epidemic is also highly resilient – even if many infected people die before they transmit the contagion or are immunized, the epidemic will reliably propagate throughout the population.

5.1 Dissemination in Distributed System

In an epidemic algorithm, all system processes are potentially involved in the information dissemination. Basically, every process buffers every message it receives up to a certain buffer capacity b and forwards that message a limited number of times t. The process forwards the message each time to a randomly selected set of processes of limited size f.

Many variants of epidemic algorithms exist and are typically distinguished by the values of b, t, and f. These parameters may be fixed independently of the number n of processes in the system, in which case the load imposed on every process remains bounded. The reliability of information delivery will then depend both on these values, as well as on the system size. Alternatively, the dissemination parameters can evolve with n. In this case, a reasonable load could be maintained, if the parameters increase slowly with n – e.g., logarithmically.

Every process that receives a message to be disseminated forwards it by default to a randomly chosen subset f of other processes. Each of these infected processes in turn forwards the information to another random subset. Thus, unlike reactive algorithms, in which processes react to failures by retransmitting missing information, epidemic algorithms do not require a mechanism to detect and reconfigure from failures.

As Figure 9 shows, a multicast source, represented by the black circle, sends a message to be disseminated in a system of size n. Each infected process (e.g. each process that receives the message) forwards it to a random subset of size $O(\log n)$. Eventually, the message will reach all members of the system with a high probability after $O(\log n)$ rounds. The failure of one of several communication links or processes does not significantly affect propagation of the message to live processes.

In addition, epidemic algorithms exhibit bimodal behaviour. They either achieve successful delivery to almost all processes or only reach a negligible portion of the processes.

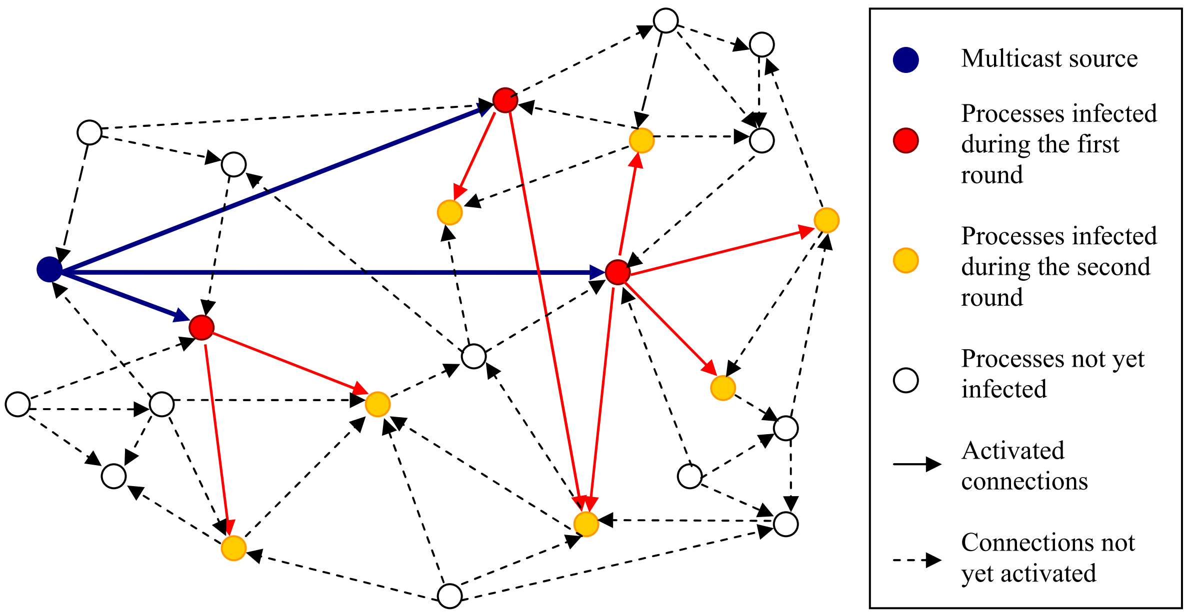

Figure 9 Schema of the mechanism of epidemic infection dissemination

Implementing an epidemic algorithm requires addressing specific design constraints with respect to:

- *membership* - how processes get to know each other, and how many they need to know
- *connecting awareness* – how to make connections among processes to ensure acceptable performance
- *buffer management* – which information to drop at a process when its storage is full
- *message filtering* – how to take into account the actual interest of processes and decrease the probability that they receive and store information of no interest

Although studies of natural epidemics can provide useful insights into these issues, innovative solutions are required, because such studies have primarily focused on quenching epidemics rather than facilitating their spread, which is the goal of epidemic algorithms.

5.2 Simple Epidemic Models

Several mathematical models have been proposed as the following simple epidemic models will show. They contain many limiting assumptions and are posted here as an overview of the epidemic problems. The subsequent description was inspired by the paper (EGKM) and by the works (ATNE77), (BAIL75), (PIT87), (WS98), (KKD01) and also by the realization of the project (HOFHA), which I managed, and the work (CIPRA).

5.2.1 Branching Processes

Let us assume that X is a random variable interpreted as the number of infectious descendants of the individuals in a given population in some generation. Each individual in each generation gives the birth with some probability p_k to k descendants ($k \geq 0$).

$$P\{X = k\} = p_k \tag{5.1}$$

Then, $P\{Z_n = k\}$ is the probability of the occurrence of k infectious individuals in the n-th generation. Usually, $P\{Z_0 = 1\} = 1$ is considered.

Let

$$m = E\{X\} = \sum_{i=0}^{+\infty} ip_i \tag{5.2}$$

is an expected value of a number of infected descendants of some individual in some generation.

The variance of the number of infected descendants in some generation is

$$\sigma^2 = E\{(X - E\{X\})^2\} = \sum_{i=0}^{+\infty} (i - E\{X\})^2 p_i \tag{5.3}$$

Thus, we can write by using (CIPRA)

$$E\{Z_n\} = m^n \tag{5.4}$$

that is the expected value of the number of infected individuals in the n-th generation, and

$$\begin{aligned} \sigma^2\{Z_n\} &= \frac{m^{n-1}(m^n - 1)}{m-1}\sigma^2 \Leftrightarrow m \neq 1 \\ \sigma^2\{Z_n\} &= n\sigma^2 \Leftrightarrow m = 1 \end{aligned} \tag{5.5}$$

is the variance of the number of infected individuals in the n-th generation.

Now, we can describe the probability generation functions

$$g(z) = \sum_{i=0}^{+\infty} p_i z^i \,,\; g_n(z) = \sum_{i=0}^{+\infty} P\{Z_n = i\} z^i \tag{5.6}$$

and then

$$g_{n+1}(z) = g_n(g(z)) = g(g_n(z)),\; g_0(z) = g(z). \tag{5.7}$$

As the formula of the probability $P\{Z_n = k\}$ is very complicated, the recursive evaluation of the probability generation function is relatively simple and explicit.

The probability of the extinction in the *n*-th generation

$$P\{Z_n = 0\} = q_n \tag{5.8}$$

is the important term in the analyze of the branching processes.

Neglecting the cases $p_0 = 1$, $p_0 = 0$, the sequence q_n is increasing with the limit value equal to the smallest positive root of the equation $z = g(z)$, especially, for $m \le 1 \Rightarrow \lim_{n\to\infty} q_n = 1$ and for $m > 1 \Rightarrow \lim_{n\to\infty} q_n = \varsigma \quad 0 < \varsigma < 1$.

Some examples of the branching processes for the various parameters changing are presented on three graphs in the Fig. 10. The models based on the basic type of the branching process describe the behaviour of the infected population only. They do not respect the population which is not infected. The simple probable and statistic apparatus is the main advantage of these models. These models are known as "infect and die". The infected individual does not penetrate in the next traced generation.

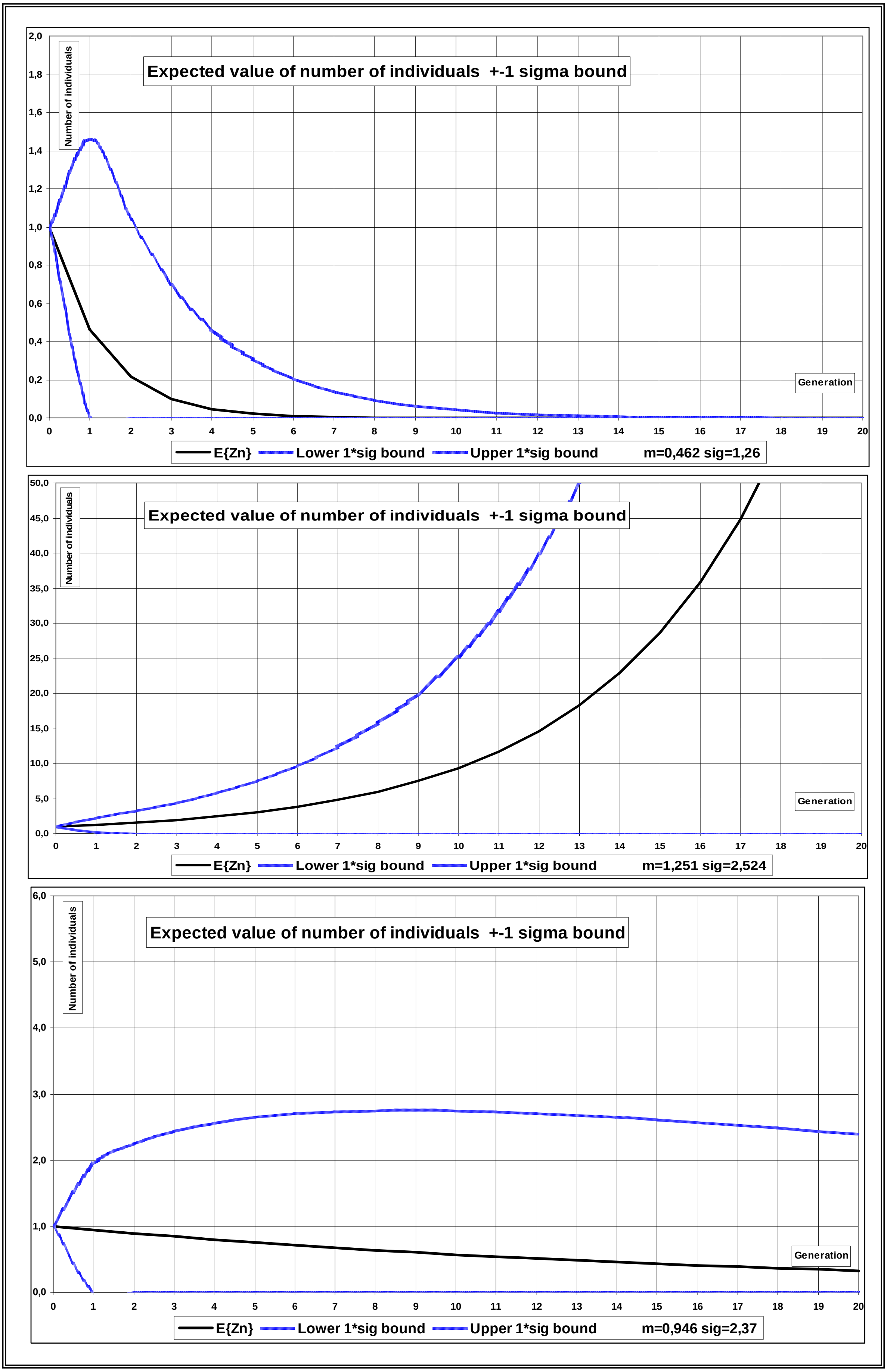

Figure 10 Three examples of the branching processes

5.2.2 Expected Value Models

Let us denote $Y_n = E\{Z_n\}$ the expected value of the number of the individuals in the n-th generation, some models of infection spread can be described. We can distinguish four special cases of the models – the geometrical spread, the finite population geometrical spread with constant intensity, the geometrical spread with variable intensity and the logistic spread.

Geometrical Spread

This model is one of the simplest models (Malthusian model) and it is known as the "infected forever" model. It can be used as the base for the construction of the spread models, which describe the reality more truly.

Let r is the relative increase (intensity) of the population in some generation. Thus,

$$\frac{Y_{n+1}}{Y_n} = 1 + r \quad \Leftrightarrow \quad \frac{Y_{n+1} - Y_n}{Y_n} = r \tag{5.9}$$

The explicit form of expected value of the number of individuals in the n-th generation is

$$Y_n = Y_0 (1 + r)^n \tag{5.10}$$

We can consider the spread intensity

$$r \approx \frac{Y_{n+1} - Y_n}{Y_n} = \rho \tag{5.11}$$

as the random variable with some type of the distribution and the expected value r. The random process is

$$\frac{Z_{n+1} - Z_n}{Z_n} = \rho_n \ , \tag{5.12}$$

where ρ_n is the realization of the spread intensity upon the transition from the n-th generation to the (n+1)-th generation. This augmentation of the spread model assumes the independence of the random variables Z_n and ρ_n , i.e. the spread intensity and the size of infected population are independent. The simulation of the geometrical spread model is shown in Fig. 11.

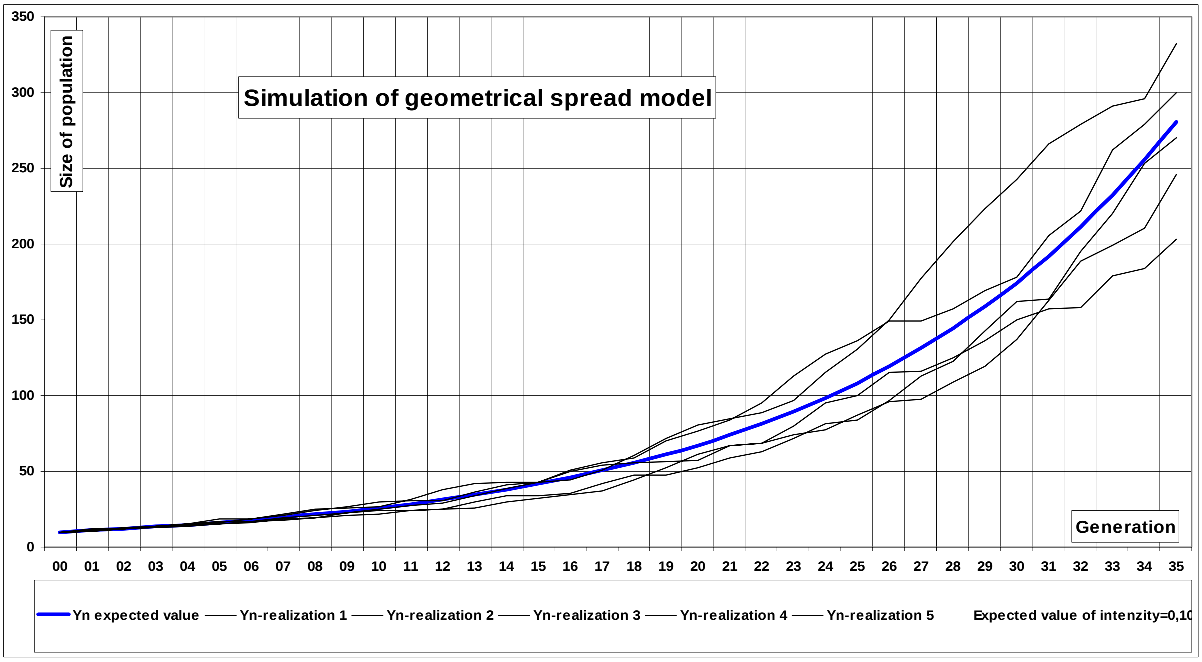

Figure 11 Simulation of the geometrical spread model

Geometrical Proportional Spread Model

We consider the finite population and the constant spread mean intensity in the remaining healthy population.

Again, $Y_n = E\{Z_n\}$ and N is the whole number of individuals in the considered population. The expected value of the number of infected individuals is given by the formula

$$Y_{n+1} = Y_n + r(N - Y_n);\ \ Y_0 < N;\ \ 0 \le r \le 1 \tag{5.13}$$

The following formula describes the explicit solution of this spread

$$Y_n = N - (1-r)^n (N - Y_0). \tag{5.14}$$

On contrary to the previous model, this model does not describe non-restricted spread. It cannot model the situation, when the epidemic has extinct before infection of all individuals.

In parallel to the previous model, the spread intensity

$$r \approx \frac{Y_{n+1} - Y_n}{N - Y_n} = \rho \tag{5.15}$$

can be perceived as the random variable with some type of the distribution and the expected value *r*. The random process is

$$\frac{Z_{n+1} - Z_n}{N - Z_n} = \rho_n, \tag{5.16}$$

where ρ_n is the realization of the spread intensity upon the transition from the *n*-th generation to the (n+1)-th generation. This augmentation of the spread model assumes the independence of the random variables Z_n and ρ_n, i.e. the spread intensity and the size of infected population are independent. The simulation of the proportional spread model shows Fig. 12.

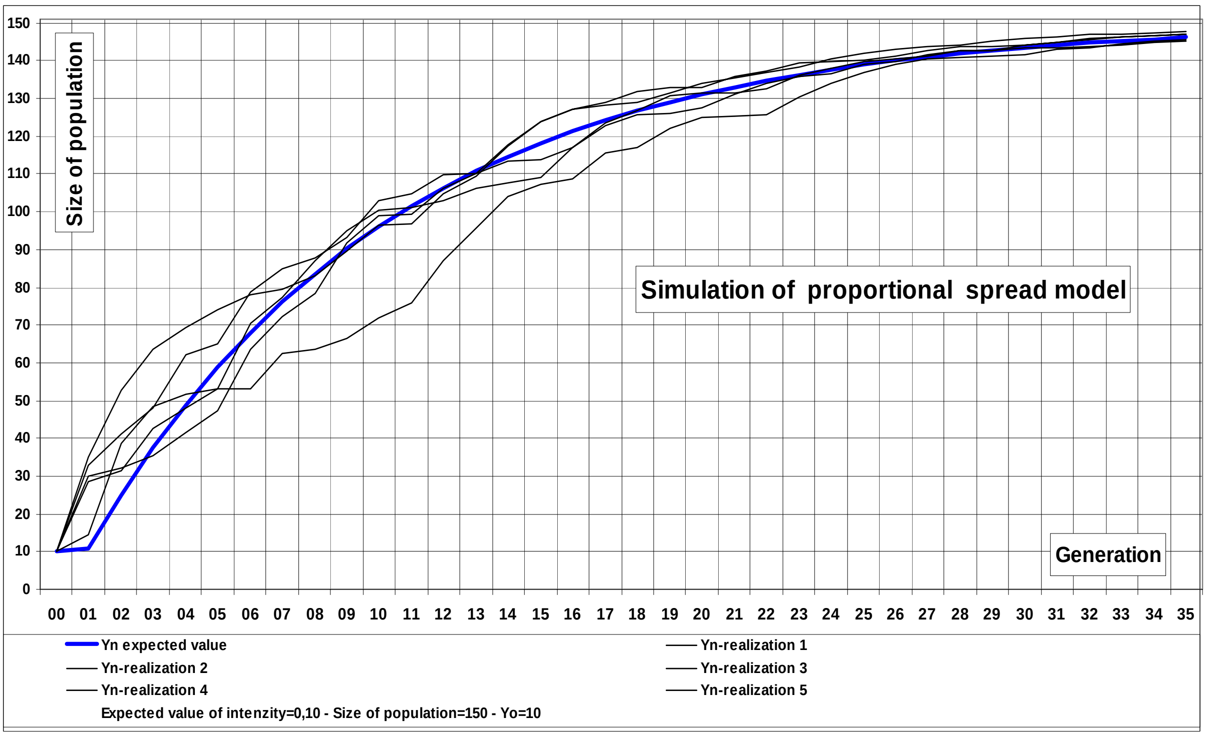

Figure 12 Simulation of the proportional spread model

Geometrical Spread Variable Intensity

Relative increase (spread intensity) r_n of the population from the n-th to the (n+1)-th generation is

$$\frac{Y_{n+1}}{Y_n} = 1 + r_n \quad \Leftrightarrow \quad \frac{Y_{n+1} - Y_n}{Y_n} = r_n \tag{5.17}$$

The explicit solution has the form

$$Y_n = Y_0 \prod_{i=0}^{n-1} \left(1 + r_i\right) \tag{5.18}$$

or the logarithmic one

$$log\, Y_n = log\, Y_0 + \sum_{i=0}^{n-1} log(1 + r_i\,)\,. \tag{5.19}$$

The following convention

$$log(1 + r_i\,) = q^i \Rightarrow r_i = e^{q^i} - 1 \tag{5.20}$$

may be used for some kind of the spread. Thus,

$$Y_n = Y_0 e^{\frac{1-q^n}{1-q}} \quad \text{by} \quad log_e\,. \tag{5.21}$$

In some tasks, the natural requirement is $0 \le r_i \le 1; \forall i \ge 0 \Leftrightarrow 0 \le e^{q^i} - 1 \le 1$, we get $0 \le q^i \le log_e 2 \Rightarrow 0 \le q \le 1$. While we omit the limit states, we get $0 < q < 1$ and $\lim_{n \to +\infty} Y_n = Y_0 e^{\frac{1}{1-q}}$. In this case, we have the model with the saturation of spread, which can be lower than the whole size of population N.

It is up to $$Y_0 e^{\frac{1}{1-q}} < N \Leftrightarrow e^{\frac{1}{1-q}} < \frac{N}{Y_0} \Leftrightarrow \frac{1}{1-q} < log_e \frac{N}{Y_0}\,. \tag{5.22}$$

Solving (5.22) we have $$q < \frac{log_e \frac{N}{Y_0} - 1}{log_e \frac{N}{Y_0}} \tag{5.23}$$

and $$0 < q < \frac{log_e \frac{N}{Y_0} - 1}{log_e \frac{N}{Y_0}}\,. \tag{5.24}$$

Special type of this model is the geometrical spread variable intensity model described above.

Logistic Spread Model

Again, $Y_n = E\{Z_n\}$ and N is the whole number of individuals in the considered population and $Y_{\max}$ is the maximum number of non immune individuals from the population to the given infection. The expected value of the number of infected individuals is given by the formula

$$\frac{Y_{n+1}-Y_n}{Y_n}=r\left(1-\frac{Y_n}{Y_{\max}}\right);\quad Y_0, Y_{\max}<N;\ 0\le r\le 1. \tag{5.25}$$

The exact form of the spread is

$$Y_{n+1}=Y_n\left[1+r\left(1-\frac{E\{Z_n^2\}}{Y_nY_{\max}}\right)\right]. \tag{5.26}$$

The spread described by this model is limited by the value $Y_{\max}$ and by the spread intensity on the other hand. Then, the spread approximately measures the relative increase, as can be seen from the definition equation

$$\frac{Y_{n+1}-Y_n}{Y_n}=r\frac{Y_{\max}-Y_n}{Y_{\max}}. \tag{5.27}$$

The spread intensity

$$r\approx\frac{\dfrac{Y_{n+1}-Y_n}{Y_n}}{\dfrac{Y_{\max}-Y_n}{Y_{\max}}}=\rho \tag{5.28}$$

can be considered as the random variable with some type of the distribution and the expected value r. The random process is

$$Z_{n+1}=Z_n\left[1+\rho_n\left(1-\frac{Z_n}{Y_{\max}}\right)\right] \tag{5.29}$$

where ρ_n is the realization of the spread intensity upon the transition from the n-th generation to the (n+1)-th generation. This augmentation of the spread model assumes the independence of the random variables Z_n and ρ_n, i.e. the spread intensity and the size of infected population are independent. See the simulation in Fig. 13.

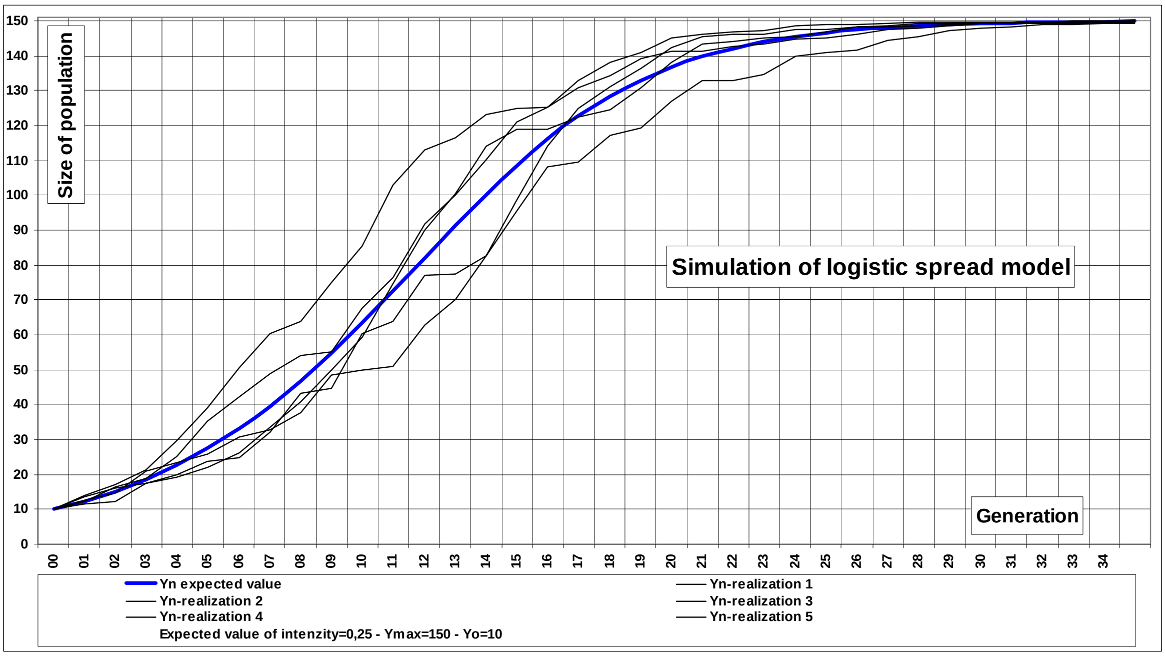

Figure 13 Simulation of the logistic spread model

5.2.3 Latent Infection Models

There is an implicit assumption in the models described above, namely the spread begins immediately. In some situations, the propagation delay of L generations exists between the contact and exhibition of the infection. This situation is described in the proportional model by

$$Y_{n+L} = Y_n + r(N - Y_n); \quad Y_0 < N; \; 0 \le r \le 1; \; L \ge 1 \tag{5.30}$$

Then the explicit formula is more complicated.

The spread intensity

$$r \approx \frac{Y_{n+L} - Y_n}{N - Y_n} = \rho \tag{5.31}$$

is the random variable with some type of the distribution and the expected value *r*. The random process is

$$\frac{Z_{n+L} - Z_n}{N - Z_n} = \rho_n \tag{5.32}$$

where ρ_n is the realization of the spread intensity upon the transition from the *n*-th generation to the (*n*+1)-th generation. This augmentation of the spread model assumes the independence of the random variables Z_n and ρ_n, i.e. the spread intensity and the size of infected population are independent. The simulation of the proportional model with the latency is shown in Fig. 14

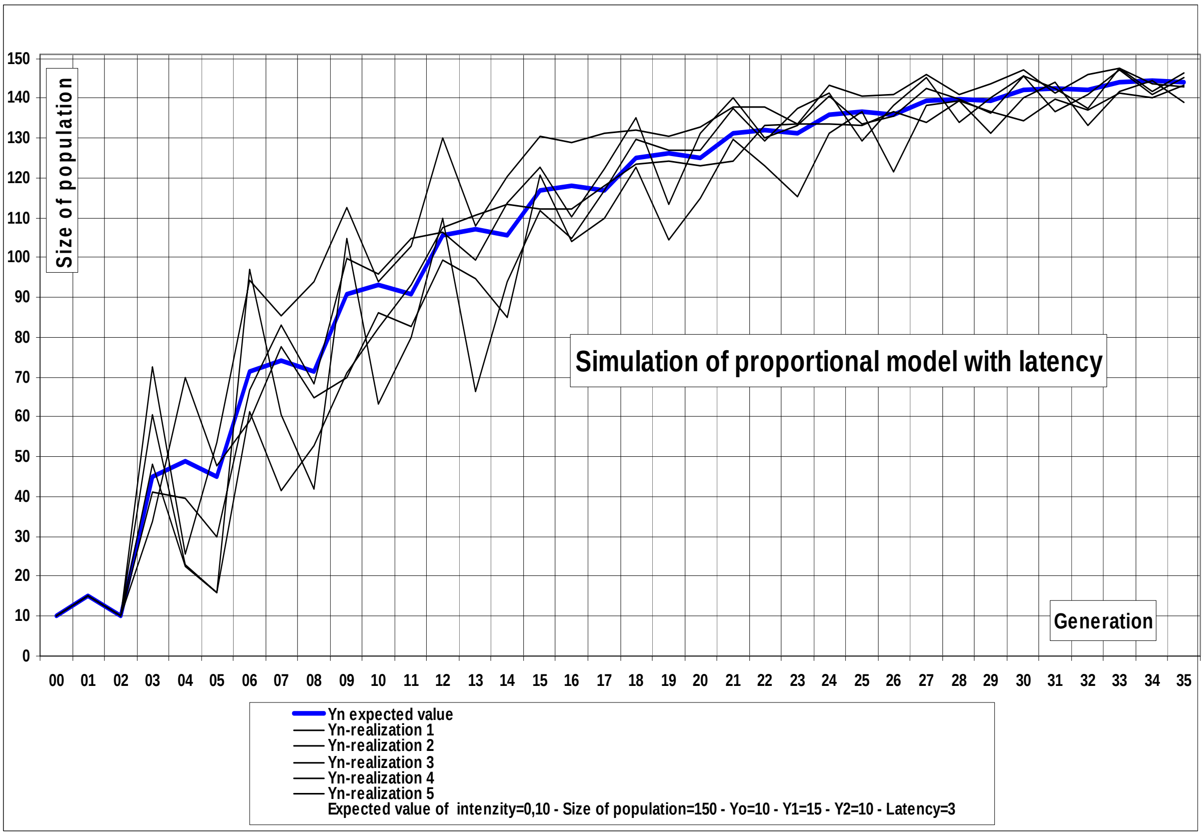

Figure 14 Simulation of the proportional model with the latent infection

6 Agent and Multi-Agent Systems Technology

Agent and Multi-Agent Systems (AS and MAS), as decentralized distributed systems, are very suitable for creating models in diverse scientific areas. Adapting some agent definitions [(WOOL), (SWARM), (SIGI), (FERB99)], agents are sophisticated computer programs that act autonomously on behalf of their users, across open and distributed environments, to solve a growing number of complex problems. But most of applications require multiple agents working together. This way a multi-agent system is a loosely coupled network of software agents [(WOJE95)]. These agents interact to solve problems that are beyond the individual capacities or knowledge of each problem solver. Agent based architectures are becoming much more simple to construct due to development of object oriented programming languages. Consequently, multi-agent, or Distributed Artificial Intelligence (DAI), models have become increasingly popular. Multi-agent approaches can be an effective alternative to other modeling approaches. On the other hand, if a mathematical model using differential equations is established, then it can be embedded into a broader multi-agent approach.

The advantages of the multi-agent approach over a single agent or centralized approach are as follows. A multi-agent system distributes computational resources and capabilities across a network of interconnected agents. It models problems in terms of autonomous interacting component–agents. A multi-agent system efficiently retrieves, filters, and globally coordinates information from distributed sources. It provides solutions in situations of distributed expertise and enhances overall system performance, along with computational efficiency, robustness, flexibility etc.

In addition, multi-agent simulations can be used to model continuous or discrete state variables, lending themselves equally to linear or non-linear modeling tasks. They can be readily assembled at single or multiple levels. They may be designed to provide simple report including visual representation of whatever state conditions are of interest to the researcher [(FERB99)]. There is an eventuality how to design the trust evolution model by applying the modern agent technology.

6.1 Agent Systems

We can think of agents as of living entities. M. E. Bratman in his work Intention, Plans, and Practical Reasoning in 1987 described behaviour of an agent using the following three parts - the beliefs (this part contents elements which the agent treats as true), the desires (this part contents elements which the agent wants to perform), and the intentions (this part contents elements which are agent's goals). Then the formal description of the agent can be expressed by BDI logic, the example is shown in Figure 15.

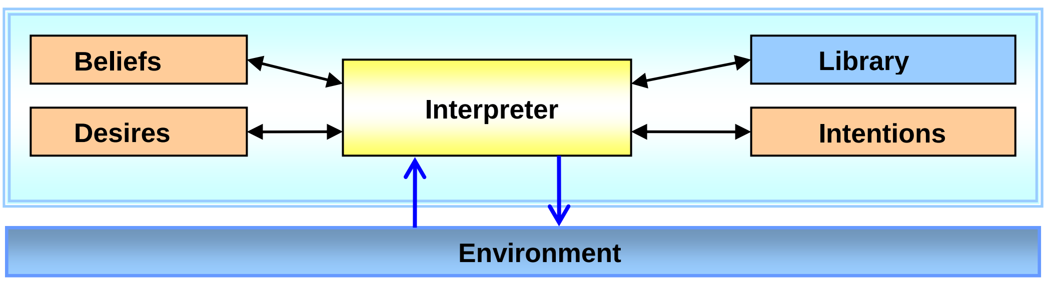

Figure 15 BDI description of the agent

Our agents are traditional agents with memory, energy supply, receptors and effectors. They have ability to observe, act, remember, reproduce and die. Agent's energy supply is a simplified concept of the life energy. Basically, the energy is used for performing agent actions. By running out of the energy, the agent dies. Memory is the agent's organ that has an ability to collect, store and forget observed information.

The agent system is characterized by the environment, where the agent operates. This environment may be accessible or inaccessible, dynamic or static, non-deterministic or deterministic and discrete or continuous. The agent is an autonomous unit that is furnished with a quantum of intellect and is able to solve some specific problems. The result of the agent action is the transition from initial to the required state. By Wooldridge [(WOOL)]: "The agent is an encapsulated computational system that is situated in some environment, and that is capable of flexible, autonomous behaviour in order to meet its design objective".

6.1.1 Characteristics

We can introduce basic characteristics of agents [(PECHO)]. They are:

- *autonomous* - agents are proactive, goal directed and acting on their own, performing tasks without necessarily requiring user initiation, confirmation and notification, do not have to be benevolent, have free will, can cheat, can leave or join the community
- *reactive* - agents are triggered by events and sensitive to real-time domain considerations able to sense and act
- *intentional* - agents have the ability to maintain long term intentions, organize their behaviour in order to meet targeted goals, use speech-act-based communication, formulate plans in pursuit of their own agenda, and use reflective reasoning
- *social* - agents collaborate together in communities to achieve a shared goals, they are aware one of the other, they perform reasoning about each other, can group into coalitions, teams and they can benefit from this

6.1.2 Behaviour and Knowledge

Each agent enters into a partnership with a group and respects the specified rules, shares resources, offers services, accepts commitments and coordinates its activity in concordance with a global group intention. This intention we denominate as agent's cooperation. The coordination specifies generally agent's behaviour in order to avoid conflicts and to implement coherent interaction. Cooperation is task oriented, directed towards a specified goal of a team or coalition.

Agent has usually limited knowledge about the state of environment. However, this knowledge widely conduces to its rational behaviour. Consequently, the exchange of knowledge between agents can be a benefit to the whole group work.

We can divide agent's knowledge into:

- *problem oriented knowledge* – guides agent's autonomous local decision making processes
- *self knowledge* – knowledge of agent's behaviour, internal status and commitments
- *social knowledge* – knowledge of other agents, their behavioural patterns, their capabilities, loading, experiences, commitments and belief

Social knowledge enables responsibility delegation, simple tasks decomposition, contracting of optimal working agents, grouping teams and coalitions and searching missing information.

6.1.3 Knowledge Management and Models

Agent acquaintance model, shown in Figure 16, is a computational model of agent's mutual awareness. It stores and maintains social knowledge, based on the 3bA model. It composes Cooperator Base that contains permanent knowledge, State Base, whose one part is a permanent problem section and the second part is maintained by the plan, and Task Base, which is maintained by periodical revisions, subscription based maintenance, blackboard based maintenance and non-cooperative knowledge maintenance.

Knowledge improvement is one of the forms of "social thinking". It can be on the level of single agents – agent learns itself, optimizes and reorganizes its activity and modifies permanent knowledge. It can improve on meta-level. Meta-agent is an independent agent that follows the objectives of the whole community or only of its part and generalizes collected data and gained knowledge.

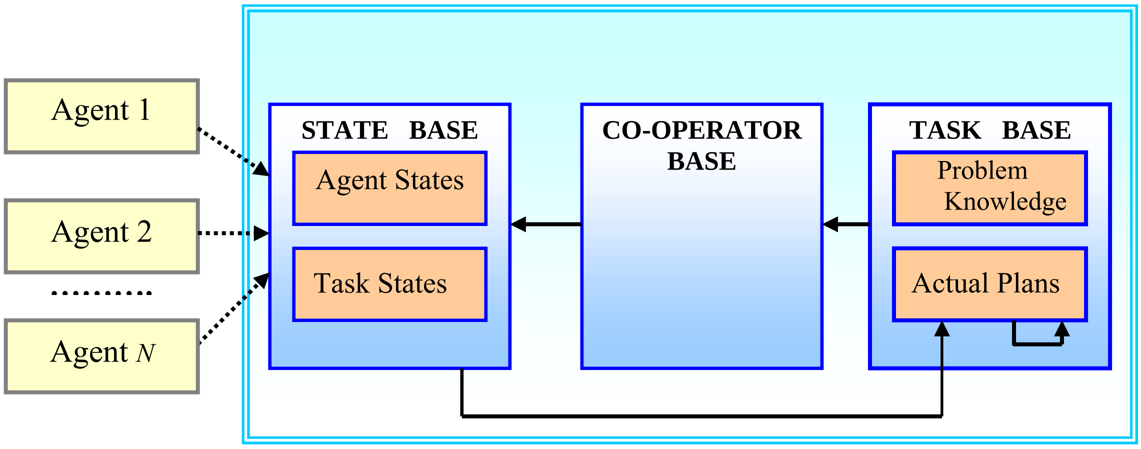

Figure 16 Agent acquaintance model composed of state base, co-operator base and task base

6.1.4 Communication and FIPA

There are several possible ways how agents may communicate. It can be through a shared memory – blackboard, or by means of a communication facilitator (see Fig. 17) (RETSI). It is a component that organizes communication on a platform of broadcasting or peer-to-peer communication. The broadcasting sends every massage to the entire community. The peer-to-peer is direct, non transparent communication between two agents.

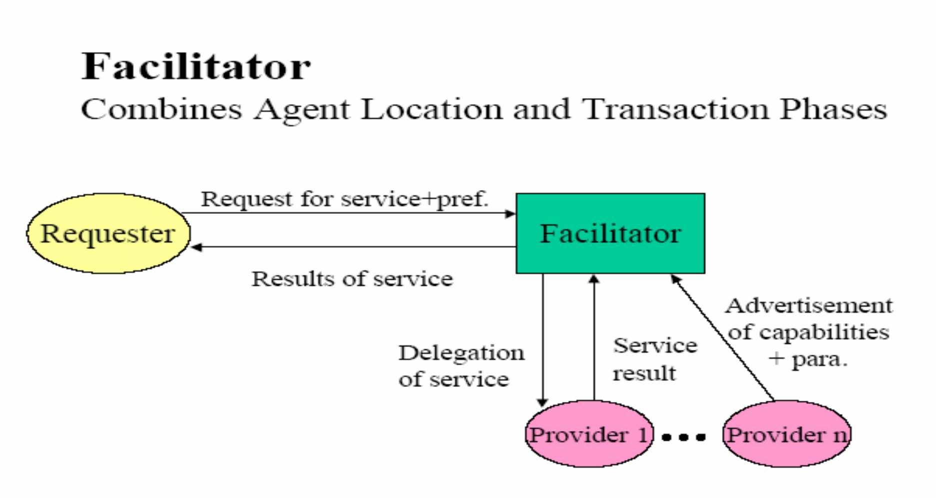

Figure 17 Communication facilitator function scheme (RETSI)

The languages, e.g. KQML (Knowledge Query Manipulation Language) or ACL (Agent Communication Language specified by FIPA (Foundation for Intelligent Physical Agent), are used by communicating agents.

FIPA is a non-profit association registered in Geneva, founded in 1995. The main goal is to maximize interoperability across agent based applications, services and equipment. It is done through FIPA specifications. FIPA specifies the set of interfaces which the agent uses for interaction with various components in environment. It focuses on specifying external communication among agents rather than the internal processing of communication at the receiver.

Clearly, a specification of how the agent will treat communication must exist. This is the interaction protocol, which is a specification of only possible and required response to a message. It specifies both, the very simple protocols such as requesting an action, querying the information, or informing others as much as complex interaction patterns such as contract-net-protocol, auctioning or voting. In Figure 18 (RETSI), one of often used treat mechanisms – the matchmaker is shown.

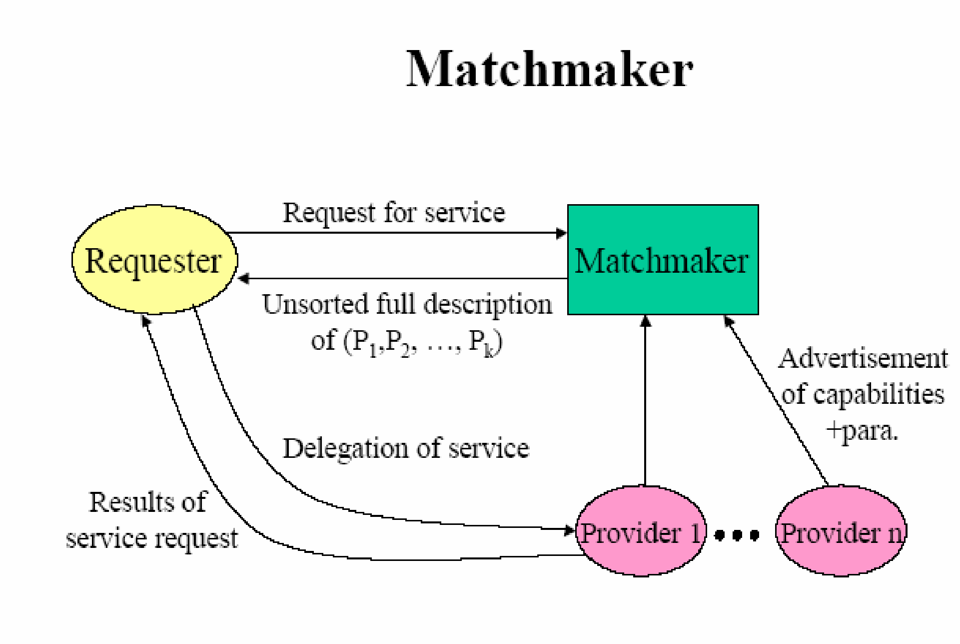

Figure 18 Matchmaker function scheme (RETSI)

6.1.5 Planning

One of agent's instances of intelligent behaviour is a rational interaction among the agent, the environment and the task, as shown in Fig. 19.

Figure 19 Interrelationship between agent, environment and task

Planning (ZBOR) is nothing but reasoning about a hypothetical interaction among the agent and the environment with respect to a given task. Motivation for the planning process is the reason of the possible action that will change the environment in order to reach the goal.

How do we draw the distinction between planning and scheduling? While the scheduling assigns in time resources to separate processes, planning considers possible interaction among components of the plan.

The plan is correct if it provides a solution of the problem. The correct plan is complete and consistent. Planning has a hierarchy. Thus, the planning problem is decomposed to the sub-problems. The agent has not always full and complete information about the evolving environment. Then we try to use a conditional planning, which requires the actions execution monitoring. When the preconditions do not hold, the plan must be repaired. We realize re-planning. In multi-agent systems, the team action plan is the result of inter-agent negotiation and mutual agreement upon joint commitments.

6.2 Social Aspects in Agent Systems

6.2.1 Strategy, Agreements, Negotiation, Coalitions and Commitments

Agent's strategy describes that an action will be done as an actual reaction status of the environment. The dominant strategy is the best individual strategy without seeing the strategy of others. The rational agent chooses always the dominant strategy. The strategy of a group is Nash equilibrium, which describes that each of strategies is the best individual strategy of the competent agent due to selected strategies of the other agents.

Generally, the strategy choice leading to the optimal benefit of the whole group requires coordination of negotiation among all agents. They must communicate with each other and need the will to benefit to the whole group. Agent's group can have joint mental poses defined by formulas and all of agents must know them.

The common mental poses are the background for making agreements and coalitions. The agents that create a group accept the commitments and general rules, and they abide by the norms. The agents that collaborate must have the capability of communication with each other. It enables coordination of their actions and searching for the joint strategies to acquire their joint interests. The negotiation is a technique for reaching an agreement on a matter of mutual interest. The negotiation may be

- *one-to-one* - symmetric preferences scenarios
- *many-to-one* - auctions (e.g. English, Dutch, Vickery); contract-net-protocol
- *many-to-many* - a special social choice action; P2P possible computational threads

The basic rules of collaboration are saved in their knowledge bases. Furthermore, the agents are able to plan their activity.

The rational agent enters the collaboration with the other agents with their commitments only then if it may look to any profit. The agents reach by agreement better environment status than they would have reached by an autonomous non-coordinated action, or they reach a compromise in the course of a conflict of their interests. The collaboration and the creation of compromises bear on the shared goals, resources and the conflicts of interest. It calls into existence of cooperation agreements and conflict agreements.

The agents opine their will to collaborate by the way of the commitments. The commitment is the maintenance of the mental pose. In more occasions, the agents concert the conditional agreements, but the norms. The agents must share the norms for all the time of their existence.

We categorize the agent groups based on collaboration interests to the other groups they share goals or resources, or furnish information with.

6.2.2 Introduction to Game Theory

The agent always chooses the dominant strategy, if this strategy exists. However, the agent must know the strategies of others. Appling the game theory may help us to find the dominant strategy.

Osborne (OSBO97) writes: "Game theory is a set of analytical tools designed to improve our understanding of situations in which decision-makers interact. Two assumptions underlie the theory: each decision-maker pursues a well-defined exogenous objective, and takes into account his knowledge or expectations of the other decision-makers' behavior. The second assumption leads us to refer decision-makers as players. The models of game theory are abstract representations of situations in which decision-makers interact. The theory aims to help us see how the outcome of an interaction depends on its structure."

For example, Pěchouček (PECHO) introduces the well-known prisoner's dilemma game. We insert this example for illustration how to find the dominant strategy. What does the example concern? Two men are collectively charged on a crime and held in separate cells with no way of communicating. They are told that if one confesses and the other does not, the confessor will be freed, and the other will be jailed for ten years. Both prisoners know that if neither confesses then they will each be jailed for one year. If both prisoners confess then they will each be jailed for five years. The following table determines the possibilities for both men.

Table 3 The example of Prisoner's dilemma with possibilities of two men

		Man *Y*	
		does not confess to	**confesses to**
Man *X*	**does not confess to**	$X = 1;\ Y = 1$	$X = 10;\ Y = 0$
	confesses to	$X = 0;\ Y = 10$	$X = 5;\ Y = 5$

Thus, the game pay off matrix is

$$X = \begin{Bmatrix} 1 & 10 \\ 0 & 5 \end{Bmatrix} \qquad Y = \begin{Bmatrix} 1 & 0 \\ 10 & 5 \end{Bmatrix} \tag{6.1}$$

and

$$\begin{aligned} &X\!: \{C, \neg C\} \succ \{\neg C, \neg C\} \succ \{C, C\} \succ \{\neg C, C\} \\ &Y\!: \{\neg C, C\} \succ \{C, C\} \succ \{\neg C, \neg C\} \succ \{C, \neg C\}, \end{aligned} \tag{6.2}$$

where the symbol $\succ$ represents the fact that the strategy on the left is better than on the right.

The individual rational action is confession. This guarantees no worse than 5 years, whereas not confessing guarantees at most 10 years. Such, confession is the best response to the all of possible strategies. Both men confess and get 5 years. This strategy is dominant strategy and also it is Nash equilibrium (OSBO97). But the intuition says that it is not the best outcome. Surely they should both cooperate and neither confessed each get 1 year. It should be the optimal strategy for both. This apparent paradox is the fundamental problem of multi-agent interactions.

The possibility of applying the game theory is wide. We can use it, where the choice is not necessarily the results of conscious decisions. Using of the scientific language has advantage in investigation of hypotheses.

6.2.3 Behaviour as Trust Game

We know that the agents interact with the other agents. The agents play the cooperation game that we have described above. The actions of the agents are cooperation or defection. The cooperative action is an expression of trust to others and defective action is an expression of distrust to the other agents. The rules of our game are following. We assume that the couple exists in the population of the agents, which play this game. Each agent from the couple invests energy to the interaction with the companion agent. Each of the agents performs an action, which could be the cooperation or the defection. The agents play the game while they have some energy supply, then they die.

Agent's behaviour is simple. It is based on its previous experience. If agent *X* trusts to agent *Y*, and agent *X* thinks that agent *Y* trusts to agent *X*, then agent *X* cooperates, otherwise agent *X* will defect. The predication that agent *X* trusts to agent *Y* means that the trust value of agent *X* is greater than 0.5, in case that the trust is the value between 0 and 1.

Thus, we can compute (URB04) two trust values for a relationship. These trust values are the trust to the other one and the trust from the other one. We can give a memory *M* of trusted or distrusted actions *A* to the agents. The formula, how the trust is from agent *X* to agent *Y*, is

$$T_X(y) = \frac{\sum_{i=0}^{c_{X(y)}-1} m_{X(y)}(i)}{c_{X(y)}} , \tag{6.3}$$

where $m_X(t) = A_X(t\text{-}1)$. The *c* variable is the number of the actions remembered in the memory M_{XY}.

To model behaviour with the bounded rationality, we must add a random number generator with the equal distribution *G*. It generates value 1 for the trust and value 0 for the distrust. Now, we may formulate agent's decision function

$$a(t) = \frac{T_X(t-1) + T_Y(t-1) + G(t)}{3} \tag{6.4}$$

It calculates the average of the sum of the trust and the value of the random number generator. Then the choice of the action depends on the result of the decision function. If it is greater than 0.5 the agent will cooperate, otherwise the agent will defect.

7 Engineering Human Trust in Multi-Agent System

In this chapter, we continue in trust explanation. Despite extensive studies from sociology and other disciplines a unique, precise and universal definition of the trust is missing yet. One of accepted and the most cited definitions is from sociologist Diego Gambetta that we cited in Chapter 2.

7.1 Principles of Trust

We can write that the trust is

- *subjective* - it is the degree of belief about the behaviour of other entities upon which we depend (in the definition mentioned above the entity is called agent)
- *asymmetric* - the agents need not have similar trust in each others
- *context-dependent* - trust in a specific environment does not necessarily transfer to another agent
- *dynamic* – it tends to be reduced or increased on the dependence of agent behaviour

Now, we want to use modern technologies to promote non-trivial interactions among the agents and reduce the risk transactions as much as possible. Thus, the trust-based collaboration development is necessary. This requires some trust management framework that enables to form, maintain and evolve trust opinions.

An overview of the Trust Management Model that was developed by Capra (Cap01) is shown in Figure 20. Three components form the model. They are the trust formation, the trust dissemination and the trust evolution. If the agent *a* which is called the "truster"[3] is decided for another agent *b* which is called "trustee" trust information about agent *b* has to be collected.

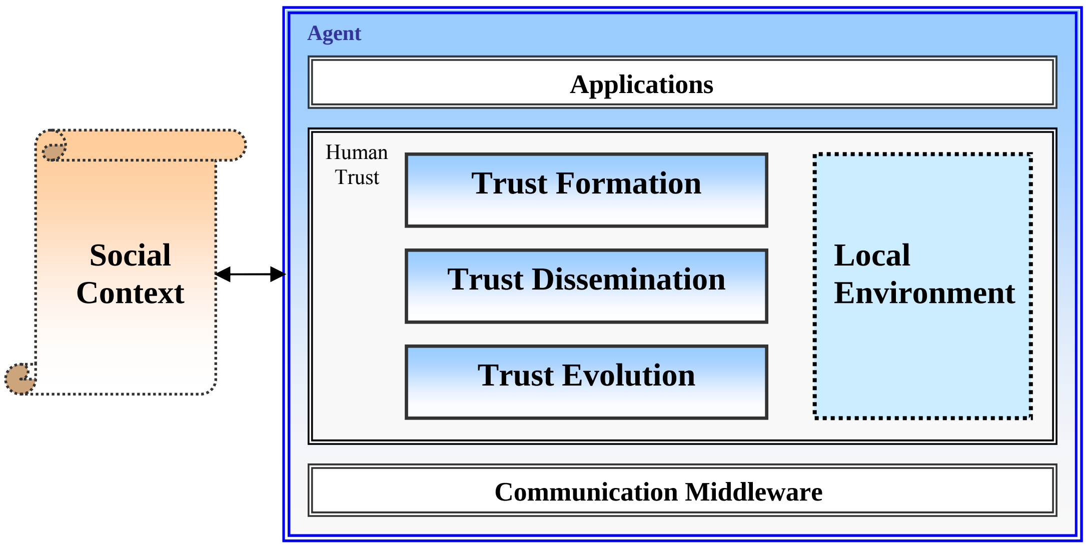

Figure 20 Overview of the Trust Management Model

[3] Here we used the received terminology. The truster is the same term that was introduced in Chapter 2. The author, Licia Capra, used the term trustor in her original terminology.

The experience represents the history of the agent which is saved in the local environment. The recommendations from the other agents are propagated by the means of the trust dissemination component. The trust information is processed by the trust formation component. These facts help to predict the trustee's trustworthiness.

In the next chapters we will introduce description of three core components of the Trust Management Model that was developed by Capra [Cap01].

7.2 Trust Formation

Trust formation is called the process that enables a truster agent to predict a trustee's trustworthiness before the interaction takes place. There is a need to have information that is used to predict trustworthiness and the trust formation function that is used to compute a prediction.

7.2.1 Trust Data Model

A truster forms a trust opinion about a trustee by aggregated trust information. Aggregated trust information is the information locally kept by Trust Management Framework (TMF) and mainly based on past direct experience with other agent (from their transactional context) and recommendations that sent to the agent by others in the social context (only those that in the past interacted with him).

Aggregated trust information is created in the local environment of truster as the set

$$[a, b, l, s, c, k, t]\ . \tag{7.1}$$

The meaning of characters is that the agent a trusts the agent b at the level l to carry on the service s in the context c with degree of knowledge k. But then aforesaid representation of value of trust as a point on the interval <0, 1> there is trust level in the interval <-1, 1> in this model, where the value -1 means total distrust. Because the agents can have only a partial knowledge of their surroundings, the degree of knowledge k is expressed in the trust opinion. The degree of knowledge varies in the interval <0, 1>, where the value 0 means unknown and value 1 represents perfect knowledge. The third important parameter (except l and k) is time. The truster's knowledge decays also with time t that indicates at which time the knowledge k refers to.

The recommendation is signed with the private key. The recommendation sent by the agent x about the agent b can be described as

$$[x, b, l, \mathrm{s}, \mathrm{c}, k, t]_{SKx}\ . \tag{7.2}$$

This has to be interpreted as: the agent x trusts the b at the level l to carry on the service s in the context c at time t and x is confident in the trust opinion given at degree k and x signs recommendation with the private key SK_x.

Recommendations are used to form trust opinion to predict the trustee's behaviour and for trust delegation. In human interactions, there is a tendency to weighting recommendations; some of them (from people with divergence of opinion) are discarded. Information about trustworthiness of an agent as a recommender is a set

$$[a, x, l, s, c, k, t]\ . \tag{7.3}$$

The meaning is that a truster a trusts the agent x at the level l to provide recommendations (service s) in a certain context c and the truster has knowledge k at time t about this information. Thus, we can describe recommender's trust as

$r = [a, x, l, k, t] \in \mathcal{R}$, where $\mathcal{R}$ is the set of all trust opinions about recommenders.

7.2.2 Trust Formation Function Υ

A truster willing to predict trustworthiness of trustee uses trust formation function, whose formal definition is described in Fig. 21 by Capra (Cap01).

$$\begin{aligned}
\Upsilon_{op} &: \mathcal{O} \rightarrow \mathcal{E} \rightarrow [-1,1] \times [-1,1] \\
\Upsilon_{op}[\![[a,b,l,k,t]]\!]_e &= [\max(-1, l - |l-f|), \min(l + |l-f|, 1)] \ , \quad f = l * k * \max\left(0, \frac{T - (t_{now} - t)}{T}\right) \\
\Upsilon_{rec} &: \wp(\mathcal{O}) \rightarrow \mathcal{E} \rightarrow [-1,1] \times [-1,1] \\
\Upsilon_{rec}[\![\{o_i | i \in [1,m]\}]\!]_e &= [l_{low}, l_{high}] \ , \\
& \quad l_{low} = \frac{\sum_i \{\pi_1(\Upsilon_{op}[\![o_i]\!]_e) * q_i | q_i > \eta\}}{\sum_i (q_i | q_i > \eta)}, \quad l_{high} = \frac{\sum_i \{\pi_2(\Upsilon_{op}[\![o_i]\!]_e) * q_i | q_i > \eta\}}{\sum_i (q_i | q_i > \eta)} \\
& \quad q_i = \max(\eta, l'_i * k'_i * \max\left(0, \frac{T - (t_{now} - t'_i)}{T}\right)) \ , \quad r'_i = lookup(i, e) = [a, i, l'_i, k'_i, t'_i] \\
\Upsilon &: \mathcal{O} \times \wp(\mathcal{O}) \rightarrow \mathcal{E} \rightarrow [-1,1] \times [-1,1] \\
\Upsilon[\![(o, O)]\!]_e &= h_1(\Upsilon_{op}[\![o]\!]_e, \Upsilon_{rec}[\![O]\!]_e)
\end{aligned}$$

Figure 21 Trust formation function Υ (Cap01).

Given the aggregated trust opinion taken from agent's local environment, and the set of recommendations coming from social context, the trust formation function returns a range of predicted trust values.

The trust formation function is used to derive a predicted range of trust values. The customising function h_1 synthesises a trust range that is given by two different trust ranges. It may be chosen to consider the local aggregated trust information only (i.e. trust reflexivity). The recommendations alone can be used (trust transitivity) when the truster has no previous knowledge of trustee.

7.3 Trust Dissemination

The trust formation function, as it was described above, uses recommendations to predict trustworthiness of a trustee. These recommendations are important when the trustee is unknown to the truster. Then, some protocol for dissemination of recommendation is indispensable. This recommendation exchange protocol guarantees truster minimum set information to create the prediction.

How does the exchange of recommendation run? Each trustee carries a portfolio of credentials. The portfolio is a set of letters of presentation that represent the history of the agent itself. Each letter comprises information as was described in (7.2). The letter is authentic, i.e. it is signed with his private key. How the exchange protocol runs is extended by the four following steps.

1. $a \rightarrow b$: *reg_for_creditals(m)* (7.4)

   This command means that *a* sends *b* a request to see its portfolio of credentials and the parameter *m* is the maximum number of letters received from *b*.

2. $b \rightarrow a$: $(o_i)_{Ski}$, where $i \in$ <1, *m*> (7.5)

   The trustee b replies with a set of letters of presentation.

3. TMF decrypts received letters by public-key infrastructure from agent signs of letters.

   The local trust formation function forms a trust opinion about *b* from decoded letters. If this information has not been allowed enough then the TMF queries the social context to obtain further recommendations about *b* and the information function one last time to predict trust interval.

4. The interaction between agent *a* and the agent *b* does not only depend on result of transformation function but also on the risks from the transactions. In this case the protocol demands that *a* and *b* exchange a letter of a presentation

$$a \rightarrow b : [a, b, l', k', t]_{SKa.} \quad \text{and} \quad b \rightarrow a : [b, a, l'', k'', t]_{SKb.} \tag{7.6}$$

If *b* has given negative feedback to *a* then *a* can decide to discard this letter without including in its portfolio.

Thus, in the exchange protocol the trust information function is computed in three different events: prior to its execution, after to obtain portfolio of credentials and once again if further recommendations are received from the social context.

There is necessity to notify that when entering a social context for the first time, an agent has no history and thus also no portfolio. After the start-up of an agent *x* without past, the agent *a* that is a member of the social context may send out, an introductory message $[a, x, l, s, c, 0,,t]_{SKa}$ to the community. The knowledge parameter $k = 0$ warns the community that the trust opinion is not based on direct experience. Thus, the acceptation of newcomer depends on the trustworthiness of agent *a*. If no introductory message has come, the newcomer may offer incentives to solicit trust (that is service-specific).

7.4 Trust Evolution

As it was discussed above, the trustworthiness of trustee is based on past experiences as perceived by truster. A fundamental component of TMF is trust evolution in this case. The evolution is the continuous self-adaptation of trust information that is kept in the local environment of the agent. There is need to introduce two further functions. These functions are an aggregation function Φ that is used to maintain information about the trustworthiness of an agent as a service provider and a tacit information extraction function Ψ that is used to maintain information about agent's trustworthiness as a recommender.

7.4.1 Aggregation Function *Φ*

The aggregation function is used to update the perceived trustworthiness of trustee when a new direct experience between two agents occurs. Only if there is no interaction then the trustworthiness of trustee may be updated. This is based on the recommendations received about trustee from trusted recommenders. Thus, the trust information for each agent is minimal. The following figure shows a formal definition of the aggregation function.

$$\begin{aligned}
\Phi &: [-1,1] \times \wp(\mathcal{O}) \rightarrow \mathcal{E} \rightarrow \mathcal{E} \\
\Phi[\![(\bar{l}, O)]\!]_e &= e \setminus \{[a,b,l,k,t]\} \cup \{[a,b,l',k',t']\} \mid o = lookup(b,e) = [a,b,l,k,t] \ \wedge \\
&\quad l' = h_4(\bar{l}, l, h_2(\Upsilon_{rec}[\![O]\!]_e)) \ \wedge \ k' = \min(k + k_{min}, 1) \ \wedge \ t' = t_{now} \\
\Phi[\![(\varepsilon, O)]\!]_e &= e \setminus \{[a,b,l,k,t]\} \cup \{[a,b,l',k',t']\} \mid o = lookup(b,e) = [a,b,l,k,t] \ \wedge \\
&\quad l' = h_4(\varepsilon, l, h_2(\Upsilon_{rec}[\![O]\!]_e)) \ \wedge \ k' = k \ \wedge \ t' = \max(t, \max_i(\{\pi_{time}(o_i), o_i \in O\}))
\end{aligned}$$

Figure 22 Aggregation FunctionΦ (Cap01).

The first equation describes the case that aggregated trust information kept in truster local environment is updated as a result of interaction occurred between truster and trustee. The old trust opinion is replaced by trustee trustworthiness as perceived by truster in just completed information.

The second equation describes the case that trust information about trustee is updated only based on lately received recommendations without an influence of interaction. This is useful for collecting information about trustworthiness about other agents with no previous interaction. Thus, a new trust opinion is computed based only on the old one and the predicted trustworthiness are computed by lately received recommendations. But in this case knowledge of truster has no increase.

The aggregated information is signed by a private key of truster. Then it is used to provide the trustee a letter of presentation at the end of the exchange protocol. Likewise it is used to answer request for recommendations that come from other agents in social context.

7.4.2 Tacit Information Extraction Function Ψ

When a truster has to make a trust decision about trustee (without previous direct experiences) there are only recommendations to rely on. Because the trust is subjective these recommendations can be conflicting with each other. In this case weighting function is used. The recommendation, coming from the agent with no share opinion, has lower weight or it is even discarded. Thus, TMF maintains a set that assesses the trustworthiness of agents as recommenders. This set serves as tacit information. When the interaction has occurred, the content of this set is updated. The formal definition of a tacit information extraction function is shown in Fig. 23.

$$
\begin{aligned}
\Psi \quad &: \quad [-1,1] \times \wp(\mathcal{O}) \rightarrow \mathcal{E} \rightarrow \mathcal{E} \\
\Psi[\![(l', O)]\!]_e \quad &= \quad e \setminus \{r_i = lookup(i,e) = [a, i, l_i, k_i, t_i], \forall o_i \in O\} \cup \{r'_i = [a, i, l'_i, k'_i, t_i], \forall o_i \in O\} \mid \\
& \quad k'_i = \min(k_i + k_{min}, 1) \ \wedge l'_i = \begin{cases} \max(-1, h_5(l_i, \delta l_i)) \text{ if } \delta l_i > \delta_{max} \\ \min(h_5(l_i, \delta l_i), 1) \text{ if } \delta l_i \leq \delta_{max} \end{cases}, \\
& \quad \delta l_i = |l' - h_2(l_i - |\pi_l(o_i) - \pi_l(o_i) * \frac{T - (t_{now} - t_i)}{T}|, l_i + |\pi_l(o_i) - \pi_l(o_i) * \frac{T - (t_{now} - t_i)}{T}|)|
\end{aligned}
$$

Figure 23 Tacit Information Extraction Function Ψ (Cap01).

The tacit information contains information on the trustworthiness of the agent as the recommender. This information is updated based on the perceived trustworthiness of trustee with whom truster has just interacted and the recommendation about trustee. A new trust value level is computed based on its past trustworthiness and the discrepancy δl_i by customising function h_5.

Both functions (Φ, Ψ) adjust the value of an agent's trustworthiness based on behaviour of the agents. The trust is changed dynamically – it will increase when behaviour is good and it will decrease when it is misbehaving. We may say that the more accurate the agent's knowledge of the surroundings becomes, the more frequently the agent has interacted, and conversely.

7.4.3 Malicious Agents Detection

In the social context, malicious behaviour refers to the spreading of fake bad recommendations and fake good recommendations. The bad ones that the agents start spreading bad recommendations to damage some others and the good ones that the agents aggregate and support each other to create a false good reputation. The feeling of these behaviours is difficult, because there is no definite way to distinguish between opinion's difference and real treat. The punishment is difficult to use in respect of agent's anonymity.

In this case Capra cogitates using an anarchic model, where each agent is responsible for its own fate. In this model the feeling of malicious agents and the punishment are concerned.

TMF supports the agent by providing a conflict detection mechanism. The punishment in this model is the loss of trust which results in isolation from future interactions. Then an agent may create a new identity. But it has no history and the others are afraid to trust to new agent. At first they must make some new interactions with this agent.

For using of this model it is assumed that the number of honest agents must be higher than the number of malicious ones. This information would be in the social context. Thus, the exchange protocol can end with lately created letters sent to social context. Detection and isolation of malicious agents would be thus at the boundary of the social and the transactional context.

As a summary, the following table shows the information that forms local environment of agents. The data is updated by TMF that used the aggregation function to maintain aggregated trust information, the tacit information extraction function to maintain tacit information, and during interactions of agents to maintain a portfolio of credentials.

Table 4 Truster Local Environment (Cap01)

Data	
Aggregate Trust Information	$[a, x, l, k, t]$
Tacit Information	$[a, x, l, k, t]$
Portfolio of Credentials	$[x, a, l, k, t]_{SK_x}$
Parameters	
Time Interval of Relevant Observations	T
Maximum Tolerate Discrepancy of Opinions	δ_{max}
Single Increment of Knowledge	k_{min}
Minimum Trust Level	η
Customising Functions	
Given two trust ranges, compute a trust range (used by Υ)	h_1
Given a trust range, compute a trust opinion (used by Υ)	h_2
Given a trust range, decide whether the prediction is precise enough (used by the recommendations exchange protocol)	h_3
Given three trust opinions, compute a new one (used by Φ)	h_4
Given a trust opinion and a discrepancy, compute a new trust opinion (used by Ψ)	h_5

8 Tools for Modelling and Simulation MAS

For the trust model implementation, we will need a simulation tool in which we can implement the model and observe the results. We can excerpt from various multi-agent modelling tools and development environments for agent based modelling or we can apply some software package for simulation of a complex system, e.g. Swarm or RETSINA.

8.1 Tools Based on Agent Principle

In this chapter we refer about some tools for implementation of the model based on the agent principle. These tools implement a base of knowledge, intentions, desire and planning. The following first three tools are created on the Java language base. It enables to take advantage of object approach and exploitation of network communication facility.

8.1.1 JAM

JAM language does not come under the well known agent implementation tools. This language consists of three components which are knowledge, intentions and plan. Knowledge is written like facts (likewise in PROLOG language) and intention is the status which agent seeks to get. Pending this plan realization sub-plan can be run. Procedures definitions contain duly intention, name, plan body and like facultative running condition, context, plan execution effect, failure effect and priority. Plan body is constructed by JAM actions i.e. insert, delete, fact changes, conditions of the environment status test, iteration, branching, etc. The primitive actions can be defined in this language as the methods and in this manner they utilize the comfort of Java language. Agent programming in JAM language diverge thereby that the environment will be transformed if the plan aborts. Thus the new intention is chosen and the plan for fulfilment is created owing to this new environment status.

8.1.2 JADE

The next tool for creating rational agents is JADE. It is not a real language, but it is the library of Java's classes. It contains software and packages of classes for creating agent platforms, for communication in ACL language, for agent life cycle definition, for folder of services, etc. JADE is a suitable tool for the implementation software mobile agents and a very popular realization resource.

8.1.3 ZEUS

ZEUS is a tool used for creating multi-agent systems. It includes the possibility of representation of knowledge, planning, communication and social commitment. Project visualization is one of ZEUS's advantages. A multi-agent system is hierarchically constructed in three layers which are definition of the social and the organisation layers. Agent's definition and its knowledge, intentions, resources and abilities are maintained in the definition layer. Configuration of agents that contains knowledge also of other agents in this group and information of its availability are defined in the social layer. The organisation layer defines the way of communication and negotiation, distribution task strategy, etc.

Furthermore we can touch on some ulterior tools.

Stella is a program for system dynamics that allows creation of models with static structure. This is suitable for modelling relationships between agents, but it is not suitable for creating models with dynamic relationships.

Squeak and other Smalltalk dialects are good for quick model prototyping and fast problem implementation.

At last, there is **Agent Farms** which is a suite of applications and frameworks for multi-agent based systems for object network oriented problems. This tool is not mature, but it provides mechanisms to explore structure of complex system by means of agents and relationships between them [URB03, URB04].

8.2 Software Packages for Simulation of Complex Systems

The software packages for multi-agent simulation of the complex systems could be divided in two approach branches, RETSINA and Swarm Intelligence.

8.2.1 RETSINA

The Intelligent Software Agents Lab at Carnegie Mellon University's Robotic Institute has developed the RETSINA multi-agent system infrastructure [SYCAR]. That infrastructure was applied also in this Lab. RETSINA is abbreviation for the Reusable Environment for Task-Structured Intelligent Networked Agents. It is an open multi-agent system that supports communities of heterogeneous agents. Intelligent software programs are known as software agents. The soft-agents undertake many of the operations as well as a multitude of other tasks. These operations and tasks are performed by human users of the World Wide Web.

An agent [RETSI] is defined as an autonomous, (preferably) intelligent, collaborative, adaptive computational entity. Here, intelligence is the ability to infer and execute needed actions, and seek and incorporate relevant information and it is given by the explicit goals.

The RETSINA system has been implemented on the next premise. The agents in a system should form a community of peers that engage in peer to peer interactions. Any coordination structure in the community of agents emerges from the relations between agents. Thus RETSINA does not employ centralized control within the MAS.

The research focuses on the problem of how to facilitate communications among agents of different types. The middle or the matchmaker agents are proposed that they serve the liaisons between requesting agents and provide agents services. It is developed to increase inter-agent communication by an agent capability description language (ACL). This language allows to communicate otherwise incompatible agents. The RETSINA Individual Agent Architecture [SYGI] is shown in Figure 24.

This architecture implements hierarchical task network planning, scheduling and execution monitoring, each of them in parallel thread. The fourth thread that is the Communicator ensures the aids for communication with network. The Communicator provides a level of abstraction. The components are point off issues of communication language, communication session management, the location of agent services, the logging and visualization of agent messages and the information of the state. The Planner thread receives plan objectives from the Communicator. It extracts the information and instructions and attempts to apply the extracted data to all the plans in the library of plans. After data applying plan actions are partially enabled. Once all actions of a plan are enabled completely and scheduled by the Scheduler, Scheduler puts the enabled actions in a priority queue. After it, the Execution Monitor actually executes the enabled actions, monitors the execution, and handles failures. The coordination among three modules is done by high-priority actions with interrupt.

The RETSINA MAS is a collection of heterogeneous software entities. These entities collaborate with each other and provide a result or a service to other users. Individual agents take up the roles which represent their commitment to participate on acquiring of a team goal. These roles may be described by considering the functional contributions that an agent makes.

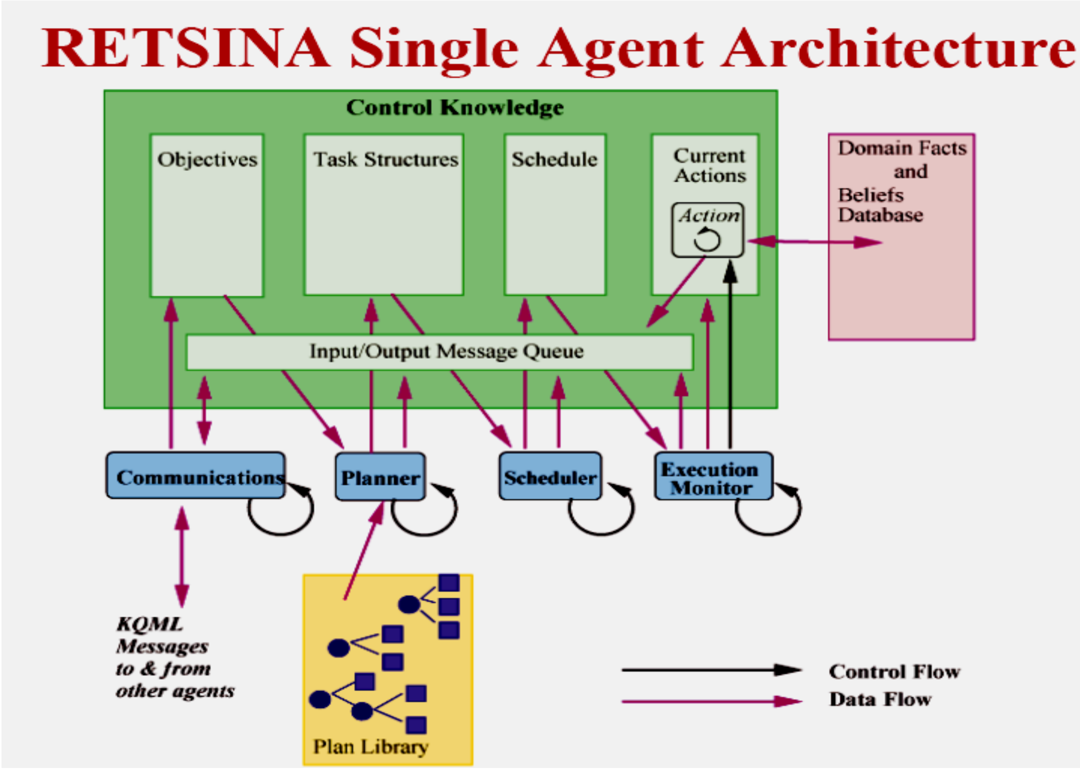

Figure 24 Schematic diagram of the RETSINA Agent Architecture (SYGI)

Graphical representation of the RETSINA MAS functional architecture is illustrated in Figure 25.

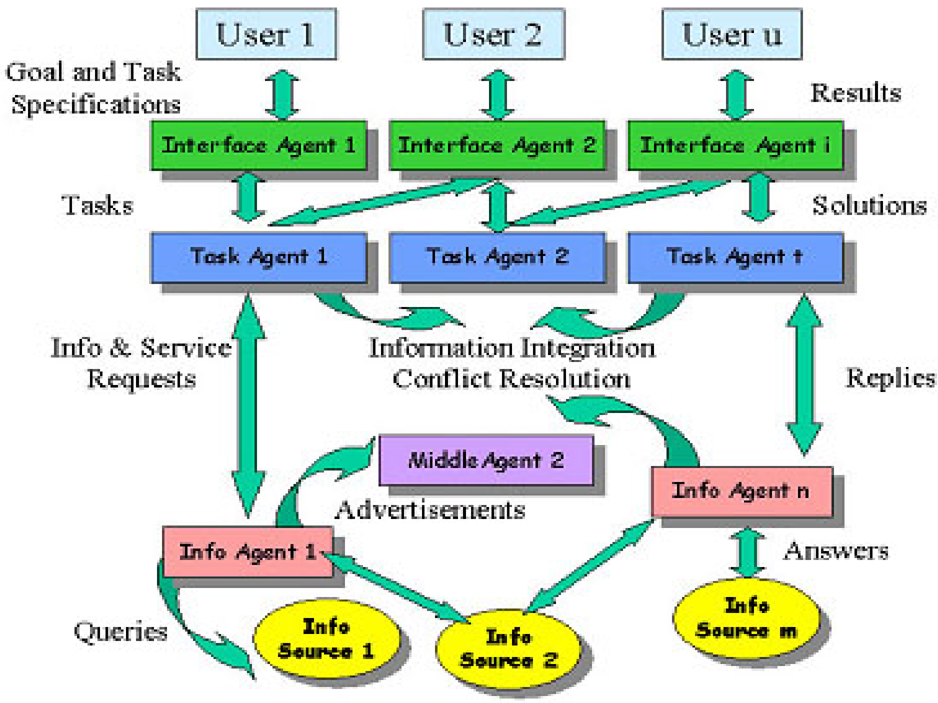

Figure 25 The RETSINA Functional Architecture (SYGI)

Four basic agent types compose the RETSINA functional architecture. These are:

- *Interface agents* that interact with users, receive user's input and display purchase results.
- *Task agents* that help users to perform tasks, and formulate problem solving plans. Thus they carry these plans by coordinating and exchanging information with other agents.
- *Middle agents* that help to make a match between agents that request and provide services.
- *Information agents* that provide intelligent access to a heterogeneous information sources collection.

RETSINA agents typically use coordination technique to ask each other. One agent dynamically discovers and interacts with the others by their needs. RETSINA agents support also other forms of coordination techniques such as team-oriented, auction-based, contract net protocol, etc.

The agents need an infrastructure of services. It permits them to find each other, to change the environment, to communicate, to warrant for satisfying proper security constrains, etc. Agents need also conventions such as ACL, conversational policies and ontology, sharing knowledge of infrastructure use and protocols. The description of organizational architecture of RETSINA with comparison between MAS and individual infrastructure is shown in Table 5. More on this problem is in (SYGI).

Table 5 RETSINA MAS Infrastructure and Individual Agent Infrastructure (SYGI)

RETSINA MAS INFRASTRUCTURE	INDIVIDUAL AGENT INFRASTRUCTURE IN RETSINA
MAS Interoperation	
RETSINA-OAA Interoperator	
Capability to Agent Mapping	**Capability to Agent Mapping**
Matchmaker	Matchmaker Module
Name to Location Mapping	**Name to Location Mapping**
ANS	ANS Module
Security	**SECURITY**
Certificate of Authority/Cryptography Services	Security Module/ private/public keys
Performances Services	**Performances Services**
Failure Monitoring	Self-Monitory/Cloning
MAS Management Services	**Management Services**
Logger/Activity Visualizer/Launcher	Logger Module
ACL Infrastructure	**ACL Infrastructure**
Public Ontology/Protocols Servers	ACL Parser/Private Ontology/Protocol Engine
Communication Infrastructure	**Communication Modules**
Discovery/Message Transfer	Discovery Module/RETSINA Communicator
Operating Environment	
Machines, OS, Network Multicast Transport Layer: TCP/IP, Wireless, Infrared, SSL	

The applications of MAS cover a variety of domains, e.g. aircraft maintenance, military electronic book buying coalitions, wireless collaboration and communications, military logistic planning, supply-chain management, joint mission planning, financial portfolio management, etc. RETSINA is currently interested in inter-agent communication and coordination, building reusable multi-agent applications that facilitate interaction among different kinds of agent systems. Some of the concrete examples (RETSI) of applications are the following:

- *Agent Storm* is an agent scenario where agents autonomously coordinate their team-oriented roles and actions while executing a mission in the simulation environment.
- *Calendar Agent* provides interoperability between web-calendar and personal information manager.
- *The Aircraft Maintenance System* is the software used by agents to assist in the documenting and making repairs to aircraft.
- *The Coala book buying coalitions* is an initial application of virtual coalitions in E-commerce.
- *The RETSINA Demining System* is a robotic de-mining system developed for assisting human commanders.
- *Joccasta* is a multi-disciplinary research project for increasing the effectiveness of a team decision making in joint planning tasks.
- *MINTEC* is a multi-agent based dynamic supply chain management system.
- *MOCHA* is a multi-agent system for "any-ware" communication and for displaying a mobile communications network.
- *MokSAF* is a software environment for route planning and team coordination.
- *The MORSE Simulation Environment* is a distributed agent-based system simulating a team-oriented task of range operations during space launch that must be completed by a team of human subjects.
- *NEO* is a multi-agent system demonstration of agent technology in non-combatant evacuation operation.
- *Text Miner* is a text classification agent application for intelligent portfolio management.
- *Urban Search and Rescue* is a hybrid system for addressing the challenges of urban search and rescue.
- *Visitor-Hoster* is designed to help a human secretary to organize a visit in an academic environment.
- *Warren* is a portfolio management application using the distributed agent architecture to access information resources already available over the Internet.
- *WebMate* is a personal agent for World-Wide Web browsing that enhances searches and learns users' interests.

8.2.2 Swarm Intelligence

The Santa FE Institute is devoted for creating a new kind of scientific research community. That is multidisciplinary collaboration in pursuit of understanding the common themes. These themes arise in natural, artificial and social systems. Swarm intelligence (SI) is an artificial intelligence technique which is based on the study of collective behaviour in decentralised and self-organised systems. The expression "swarm intelligence" was introduced by Beni & Wang in 1989 in the context of cellular robotic systems. SI systems typically contain simple agent population that interacts locally with one another in their environment. Normally there is no centralised control structure, which prescribes individual agent behaviour. Local interactions between such agents can lead to the emergence of global behaviour. We can found these systems in the first place in nature, e.g. ant colonies, bacteria moulding, animal herding, etc.

Swarm is a software package for multi-agent simulation of complex systems. It is a useful tool for researches that construct and study agent based models. Swarm software consists of a set of code libraries. These libraries enable to write simulations of agent based models in object-oriented languages like Java. The advantage of this way of solution is that it works on a very wide range of computer platforms. The Swarm basic architecture is a simulation of an agent community in which agents concurrently interact. Thus, a large variety of tasks based on agent principles can be implemented.

The software is available to the public under GNU licensing terms. It is necessary to keep in view that it is the experimental software. It is enough to be useful, but yet under development.

The agent definition by Swarm is the following. "An agent is an autonomous entity with an ontological commitment and agenda of its own." Each agent possesses the ability to act autonomously, but in business an in law an agent acts often in the best interest. An agent interacts or negotiates with other agents and also with the environment. The agent can make decisions such as whether to trust and to cooperate with others or whether not to do this.

There are two strategies useful to researchers. The first one is an empirical evaluation of dynamics. The combination of autonomous entities in a shared environment and in a recursive process can be used. The second idea is the synthesis. Some knowledge can be extrapolated in order to suggest new experiments. Swarm enables this extrapolation via computer simulation.

Engineers are increasingly interested in swarm behaviour since the results of this research can be applied in optimization, robotics, traffic patterns (in transportation systems), and military applications. Learning and evolution are the basic features of living creatures. Various genetic and evolution algorithms have been proposed in the field of artificial life. Swarm robotics is currently one of important application areas. Swarm robots can be applied to many fields, e.g. flexible manufacturing systems, spacecraft, inspection or maintenance, construction, agriculture, medicine, etc. The research of decentralized autonomous robotic systems can help in several areas of agent based modelling like agent planning or group behaviour establishing and evolution of group behaviour. There are several optimization methods proposed for the group behaviour. Thus, we will introduce a distributed genetic algorithm, a concept of the self-recognition for learning and adaptation strategy.

The following example [SWARM] is one of sample applications of Swarm software. This example is an old, basic school demonstration of nuclear fission. Let a bunch of mousetraps lay on the floor in a regular grid. Each mousetrap is loaded with two ping-pong balls. We drop one ping-pong ball in the middle. Then we watch the reaction.

This example is implemented as a simple discrete event simulation. Time stepping each mousetrap would be ineffective, because in the start of reaction most of mousetraps are doing nothing for relatively long time. It is better to simulate each ping-pong ball as an event on the schedule. This way of simulation is more effective. The computation runs only when something has to be done. One of Swarm advantages is support of discrete events modelling and time stepped models. In Figure 26 there are snapshots of the reaction all over time. The mousetrap changes colour from white to black after fire. When the mousetrap is triggered, two trigger events will schedule for two random nearby mousetraps sometime in the near future. We assume that the balls fly only a limited distance. This one induces a spreading aspect to the reaction.

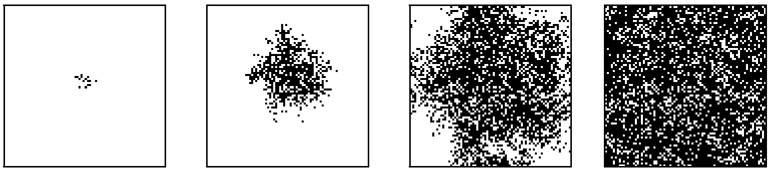

Figure 26 Time evolution of the mousetrap triggering (SWARM)

Two graphs in Figure 27 show detail of the reaction rate. The top graph shows the classic result. The reaction spreads rapidly at the beginning. But it slows down and stops when the number of loaded mousetraps derogates. The bottom graph shows the number of events pending on the schedule. This situation corresponds to the number of ping-pong balls currently in the air.

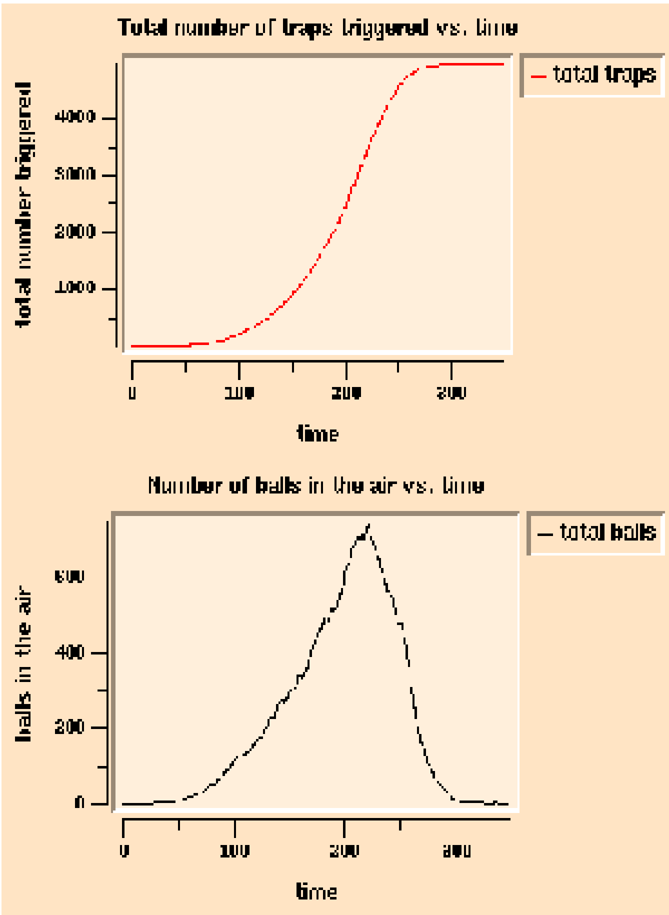

Figure 27 Model data (SWARM)

8.3 Comparison of Tools

We present the comparison of described tools according their various evaluations in (ZBOR), (SWARM), (RETSI).

The tools based on the agent principle mentioned above are rather small software instruments that enable object modelling and use modern Java language for the implementation. JAM is a language based on components with the methods of primitive actions. It is not well known implement tool. JADE is the library of Java's classes with packages for creating agent platforms using ACL language. It is a very popular tool and it is suitable for using in mobile agents implementation. ZEUS is useful for multi-agent modelling. Its advantage is visualization project.

Both introduced software packages RETSINA and Swarm have various area using and they are well known tools. Their development is running permanent. RETSINA is wide tool for MAS modelling and focuses to communication among different types of agents with using ACL language. Its applications cover a variety of domains. All about RETSINA is published on web pages, where one can find also software downloads.

Swarm is software package for multi-agent simulation of complex system. It is a useful tool for researches that construct study agent based models. The software is available to public under GNU licensing terms. Swarm software is available in three modifications: in Objective C (and other dynamically-typed languages), Java (and statically typed languages), and Functional programming languages (like ML, Haskell, and Mercury).

Both tools mentioned above are suitable for modelling and simulation of trust evolution. We will choose one of them after their testing on our examples.

9 Conclusions and Future Work

In this work, we have discussed one particular phenomenon which is called trust. The trust is emerging as an important facet of the relationships of the entities. It is plentifully used in many areas of both engineering systems and non technical systems.

The trust and its representation and visualization have been introduced. We have described some approaches to the measurement of the trust and outlined their deployment for the modelling of the trust in the community. We have made the acquaintance of some knowledge from the social communication.

We were inspired with the description of the information dissemination of the processes which are known in nature. These processes model epidemic algorithms, where the theory of probability is used. Some results of the simulation of simple epidemic algorithms were achieved.

The approach in human trust modelling can be based as well on the theory of information. Here we have presented the concept of disinformation and the comparison of the Shannon's classical theory with the concept of disinformation. We have proposed the model of information control and some results of the demonstration examples were shown.

Some terms from the agent theory have been mentioned. Fundamentals of the trust formation, trust dissemination, and trust evolution were presented deploying the agent system. The existing model of the trust management using agent technology is considered as possible environment for exploring the different measures of the trust.

We have discussed of some tools for modelling and simulation for multi-agent systems in the last chapter. Likewise, we compared some of the well-known simulation tools.

Based on rationale above, the aims of the doctoral thesis are following:

1. Development of the novel agent based model of trust evolution deploying the concept of disinformation.
2. Implementation of the developed model using selected MAS tool.
3. Evaluation of the experiments with the model and its possible application.

10 References

[ALFAR] Alfarez A.: *A Framework for Decentralised Trust Reasoning*, Dissertation, Department of Computer Science, University College London, 2005.

[ATNE72] Athreya K., Ney P.: *Branching Processes*, Springer – Verlag, 1972.

[BAIL75] Bailey N.: *The Mathematical Theory of Infectious Diseases and Its Applications,* 2nd ed., Hafner Press, 1975.

[CAP01] Capra L.: *Engineering Human Trust in Mobile System Collaborations*, Dept. of Computer Science, University College London, UK, 2001.

[CIPRA] Cipra T.: *Matematické metody demografie a pojištění*, SNTL, Praha, 1990.

[COTHO] Cover T., Thomas J.: *Elements of Information Theory*, J. Wiley & Sons, Inc., United States of America, 1991, ISBN 0-471-06259-6.

[CZISZ] Cziszár I., Körner J.: *Information Theory, Coding Theorems for Discrete Memory-less Systems.* KIADÓ, Budapest, 1986.

[DEUT] Deutsch M.: *Cooperation and Trust: Some Theoretical Notes*, Nebraska Symposium on Motivation, Nebraska University Press, 1962.

[DIZ03] Kučerová H.: *Data – informace – znalosti*, Úvod do informačních systémů, VOŠ informačních služeb, Praha, 2003.

[EGKM] Eugster P., Guerraoui R., Kermarrec A., Massoulié L.: *Epidemic Information Dissemination in Distributed Systems*, Computer, Vol. 37, May 2004, pp. 60-67.

[FERB99] Ferber J.: *Multi-agent Systems: An Introduction to Distributed Artificial Intelligence*, Addison – Wesley Pub Co, 1999.

[GOCO] Golembiewski R. T., McConkie M.: *The Centrality of Interpersonal Trust in Group Processes*, Theories of Group Processes, Wiley, 1975.

[HOFHA] Hofhans T.: *Matematické modely pro epidemické algoritmy*, Projekt 3, KIV, FAV ZČU, Plzeň, 2005.

[KKD01] Kempe D., Kleinberg J. and Demers A.: *Spatial Gossip and Resource Location Algorithms*, Proc. 33rd Ann. ACM Symposium Theory of Computing, ACM Press, 2001, pp. 163-171.

[MAR02] Mařík V.: *MAS-řešení složitých a rozsáhlých úloh*, přednášky Umělá inteligence, FEL ČVUT, Praha, 2002.

[MARSH] Marsh S.: *Formalising Trust as a Computational Concept*, Ph.D. Thesis, Department of Mathematics and Computer Science, University of Stirling, 1994.

[NESA] Netrvalová A., Šafařík J.: *Human Trust Modelling.* Aplimat 2006, 5th International Conference, Department of Mathematics, Faculty of Mechanical Engineering, Bratislava, Slovenská republika, 2006, ISBN 80-967305-7-6.

[OSBO97] Osborne M.: Draft chapter from: *An Introduction to Game Theory*, URL: http://www.socsci.mcmaster.ca/~econ/faculty/osborne .

[PECHO] Pěchouček M.: *Introduction to Multi-Agent Systems*, Agent Technology Group, Gerstner Laboratory, Department of Cybernetics, Czech Technical University URL: http://{agents|gerstner|cyber}.felk.cvut.cz .

[PIT87] Pittel B.: *On Spreading a Rumor,* SIAM J, Applied Mathematics, vol.47, no.1, 1987, pp. 213 – 223.

[POFA04] Po-Ning Ch., Fady A.: *Lecture Notes in Information Theory*, Vol. I, department of Communication Engineering, National Chiao Tung University, Taiwan, Republic of China; Department of Mathematics & Statistics, Queen's University, Kingston, Canada, 2004.

[RENYI] Rényi A.: *Teorie pravděpodobnosti*, Academia, Praha 1972.

[RETSI] RETSINA - The Robotics Institute Carnegie-Mellon University: URL: http://www.cs.cmu.edu/~softagents .

[ROBU04] Kotlíková M., Mašková H., Netrvalová A., Nový P., Vávra F., Zmrhal D.: *Informace a dezinformace – statistický pohled*, Konference ROBUST, Třešť, 2004, ISBN 80-7015-972-3.

[ROBU06] Vávra F., Nový P., Neumanová M., Netrvalová A.: *Model informačního vlivu a* dezinformace. Conference Robust 2006, Lhota nad Rohanovem, 2006, ISBN 80-967305-7-6.

[SHANN] Shannon C.: *A Mathematical Theory of Communication*, The Bell System Technical Journal, Vol. 27, pp. 379-423, 623-656, July, October, 1948.

[SIM79] Simon Herbert A.: *Rational Decision Making in Business Organisations*, The American Economic Review, Volume 69, Issue 4, 1979, 493-513.

[SWARM] Swarm Development Group - URL: http://www.swarm.org/, and Swarm Intelligence - URL: http://www.santafe.edu/swarm.

[SYCAR] Sycara K.: *Multi-Agent Infrastructure for Agent Interoperatibility in Open Computational Environments*, School of Computer Science, Carnegie Melon University, Pittsburgh, 2004.

[SYGI] Sycara K., Giampapa J., Langley B., Paolucci M.: *The RETSINA MAS, a Case Study*, The Robotic Institute, Carnegie Mellon University, Pittsburgh, 2005.

[UCALCI] Urbánek Š.: *Learning of a System with Hidden Economy by Using Simulation with Minimal Assumptions*, CALCI 2003, ISBN 80-890666-64-X.

[URB03] Urbánek Š.: *Virtual Laboratory for Evolution of Trust Modelling in Complex Systems*, Master Thesis - Project, STU - Faculty of Informatics and Information Technologies (FIIT), Bratislava, 2003.

[URB04] Urbánek Š.: *Modelling and Simulation of Trust Evolution in Complex Systems*, Master Thesis, STU – Faculty of Informatics and Information Technologies, Bratislava, 2004.

[VAJDA] Vajda I.: *Teória informácie a štatistického rozhodovania*, Alfa, Bratislava, 1982.

[VANO4] Vávra F., Nový P.: *Informace a dezinformace*, přednáška na Semináři z aplikované matematiky, Katedra aplikované matematiky, Přírodovědecká fakulta Masarykovy univerzity, Brno, 2004.

[WBD88] The World Book Dictionary, World Book, Inc. a Scott Fetzer company, The World Book Encyclopaedia, Chicago 1988.

[WOJE95] Wooldridge M., Jennings N.: *Intelligent Agents: Theory and Practice*, Knowledge Engineering review, 1995.

[WOOL] Wooldridge M.: *An Introduction to Multi Agent Systems*, John Wiley & Sons (Chichester, England), ISBN 0 47149691X.

[WS98] Watts D., Stroganz S.: *Collective Dynamics of "Small World" Networks*, Nature, vol. 363, no. 6684, 1998, pp. 440-442.

[ZBOR] Zbořil F.: *Plánování a komunikace v multiagentních systémech*, Disertační práce, Fakulta informačních technologií, VUT Brno, 2004.

[ZHAI] Zhai C. X.: *Essential Probability and Statistics*, Department of Computer Science, University of Illinois, Urbana Champaign, Lecture for CS397-CXZ Algorithms in Bioinformatics, USA, 2004.

11 Publications

[APL03] Vávra F., Nový P., Kotlíková M., Mašková H., Netrvalová A., Zmrhal D.: *Application of Insurance Methods in Power Engineering*, Aplimat 2003, 2nd International Conference, Department of Mathematics, Faculty of Mechanical Engineering, STU, Bratislava, Slovakia, February 2003, ISBN 80-227-1813-0.

[APL04] Vávra F., Nový P., Mašková H., Kotlíková M., Netrvalová A.: *Morphological filtration for time series*. Aplimat 2004, 3rd International Conference, Department of Mathematics, Faculty of Mechanical Engineering, STU, Bratislava, Slovakia, February 2004, ISBN 80-227-1995-1.

[APL05] Vávra F., Nový P., Mašková H., Netrvalová A.: *Distortion of probability models*. Aplimat 2005, 4th International Conference, Department of Mathematics, Faculty of Mechanical Engineering, STU, Bratislava, Slovakia, February 2005, ISBN 80-969264-3-8.

[APL06] Vávra F., Nový P., Netrvalová A., Neumanová M., Vokáčová K.: *Transformation and probability models.* Aplimat 2006, 5th International Conference, Department of Mathematics, Faculty of Mechanical Engineering, STU, Bratislava, Slovakia, February 2006, ISBN 80-967305-7-6.

[NEMA] Netrvalová A., Mašková H.: *Přírůstkové statistické modely časových řad*. Managing and Modelling of Financial Risks 2004, 2nd International Conference, Department of Finance, Faculty of Economics, VŠB – TU Ostrava, September 2004, ISBN 80-248-0618-5.

[NESA] Netrvalová A., Šafařík J.: *Human Trust Modelling*. Aplimat 2006, 5thInternational Conference, Department of Mathematics, Faculty of Mechanical Engineering, STU, Bratislava, Slovakia, February 2006, ISBN 80-967305-7-6.

[ROBU04] Vávra F., Nový P., Kotlíková M., Mašková H., Netrvalová A., Spíralová D., Zmrhal D.: *Informace a dezinformace – statistický pohled*. Conference Robust 2004, Třešť, June 2004, ISBN 80-7015-972-3.

[ROBU06] Vávra F., Nový P., Neumanová M., Netrvalová A.: *Model informačního vlivu a dezinformace*. Conference Robust 2006, Lhota nad Rohanovem, January 2006.

12 APPENDIX

Table 6 Trust across disciplines (T3GRO)

TRUST ACROSS DISCIPLINES	ECONOMICS / ORGANIZATIONS	SOCIOLOGY	PSYCHOLOGY	COMPUTER SCIENCE	SOCIO-COGNITIVE APPROACH
REPUTATION IN TRUST	KRAMER: Trust is a function of past interactions between subjects. Wide perspective (organizations).	GOOD: Trust is the result of several interactions between individuals (personal perspective) . Cognitive inertia tends to confirm trust or distrust behaviours. DASGUPTA: There is trust only if we know the partner. Reputation is a capital asset.		RESNICK, ZECKHOUSER, FRIEDMAN, KUWABARA: Reputation systems can solve the problem of dealing with strangers in online environments. SEN: Computational systems for reputation management: an agent gives ratings to other agents' behaviour. BARBER/KIM: Distributed solutions for reputation management. SABATER/SIERRA: When the interactions with another agent are scarce it is not possible to assign a reputation based just on direct experiences. It is in these situations when the social dimension of an agent may help by using information coming from	Reputation is one of the sources for the beliefs that contribute to create the feeling of trust within an agent's mind. Besides direct experiences, categorization and reasoning, reputation can produce the necessary beliefs for trust. (more)

				other agents. YOLUM/SINGH: Peer-to-peer service networks consist of autonomous agents who seek and provide information services among their neighbours to fulfil their needs. Information services can be cached by agents, so that the providers need not generate them again. PAVLOV/ ROSENSHEIN/ TOPOL: When feedback providers' identities are publicly known, reputation ratings can be provided in a strategic manner for reasons of reciprocation and retaliation, not properly reflecting the trustworthiness of the rated parties. SHROBE: Active Trust Management and Survivable systems: reputation as computers' strength against	

				viruses and hackers attacks.	
SOURCES ROLE IN TRUST			BJ FOGG: While surfing the Web, people assess credibility of information considering elements that are prominent.	YU/SINGH: Referral systems for trust and reputation management. They are multi-agent system whose member agents are capable of giving and following referrals. Agent s cooperate by giving and taking referrals so each can better help its user locate relevant information. BARBER/KIM: Belief revision process supported by algorithms; an agent is capable of evaluating incoming information and generating a consistent knowledge base to reason on, and avoiding fraudulent information from unreliable, incompetent, or deceptive agents. FULLAM/BARBER: Policies for belief revision and for assessing source	Trustworthiness is a property of a source while credibility should be considered a property of a piece of information. These two concepts are linked. The credibility of a piece of knowledge is a function of its sources. (more)

				reliability.	
RISK IN TRUST	WILLIAMSON: Trust is a benefits / risks balance. The goal is to maximize profits and to minimize costs. SLOVIC: Trust is a critical factor in risk assessment and management. Social relationships of all kinds are strongly influenced by trust.	LUHMANN: Trust is based on a circular relation between risk and action, both being complementary requirements. COLEMAN: The decision of an actor to trust or not is a function of the expected gain and loss involved and it is like to place a bet. SZTOMPKA: Trust is a bet on others' future behaviour; it is composed of 7 factors: regularity, efficiency, reliability, representative ness, fairness, accountability and benevolence LAGERSPETZ: Trust is a tacit mental state; if we consider trust, it is a sign of distrust.	Risk inclination is one of the characteristics of every individual. DEUTSCH: Trust is a risky choice: the trusting agent knows that he can be damaged by his trust decision.		Trust cannot be reduced to a simple and opaque index of probability because agents' decisions and behaviours depend on the specific, qualitative evaluations and mental components. (more)

CONTROL AND CONTRACTS IN TRUST			Control can reduce anxiety because it brings some guarantees and makes individual feel much secure. CIALDINI: control can diminish trust within organizations: de-motivation, resentment, etc.	Computer scientists developed and implemented systems that are able to monitor and control users' behaviour. BONS et alii: Trust can be created by forced exchanges of documents between partners that do not trust each other.	The relation between trust and control is quite complex and contracts can produce not only positive effects like stability in interactions but also counter-productive results due to the reactions of the agents against the rules. (more)
SECURITY AND PRIVACY IN TRUST	CAMP: Security and privacy are essential features for trust.			IBM LABS: Technology can easily provide security and access control: cryptography, protocols, digital signatures, etc. SHROBE: Active Trust Management and Survivable systems. PAVLOV/ ROSENSHEIN/ TOPOL: it is possible to consider the feedback provided by each witness to be his private information, or secret. This would provide	Security is one of the trust sources but they are distinct entities. (more)

				benefits in their authenticity. SEIGNEUR /JENSEN: there is a trade-off between trust and privacy: we should allow the benefit of adjunct trust when entities interact without too much privacy loss. FRIEDMAN/ RESNICK: in the Internet negative reputations do not stick; if a player used a once-in-a-lifetime identifier would effectively commit to having his reputation spread through the arena he is acting in.	
LAWS AND AUTHORITIES IN TRUST	Economists focused on a macro level, which regards state laws, government authorities and interrelations among them. Relevant problems deal with the super-national nature of the Internet, which eludes common corpora of laws	SARTOR: Laws may provide the basis for trust to exist and support those trust relationships which autonomously emerge. COLEMAN : Trust in both public and private institutions has been		Protocols and standards which are able to control online activities and data exchanges, such as TCP/IP, digital signature, and so on. Trusted Third Parties approach. MASS/SHEHORY: a new mechanism that allows agents in	Shift of the trust problem. The presence of a third party doesn't eliminate the need of trust but it merely shifts the attention of the agents towards different kinds of

	and local forms of government, and the growing menace to people's freedom and privacy.	declining for several decades. NORRIS: citizens have become critical of the representative government. BOUCKAERT /VAN DE WALLE: the nature of trust in government is changing because points of reference have changed. HARDIN: trust is largely irrelevant for the loyalties of most citizens. WARREN: Trusting an institution depends on shared knowledge of legal rules that establish institutions.		an open system to establish trust among themselves and to dynamically update this trust. Although we rely on certificates for our solution, we do not require (in contrast to previous solutions) any centralized certificate authority system.	trust, even more complex and specific. (more)
MENTAL ATTITUDES IN TRUST		LUHMANN: Distinction between trust, familiarity, confidence, faith, hope. HARDIN: Individuals	DEUTSCH: Several aspects of trust: as despair, faith, social conformity, impulsiveness,	DEMOLOMBE: Trust is an attitude of an agent who strongly believes that another agent has a given	We tried to explain the peculiarity of the trust feeling with a complex mental architecture,

		can trust someone if they have adequate grounds for believing it will be in that person's interest to be trustworthy in the relevant way at the relevant time. Trust as encapsulated interest. MARGALIT: Analysis of trust, distrust and what is in between. TUOMELA: trustworthiness, predictive trust and thick (rational social normative) trust must be distinguished.	confidence, etc. Suspicion is one of the central cognitive components of distrust.	property. Trust is given by the sincerity, credibility, cooperation, vigilance, validity and completeness of trustor's beliefs. Modal logic.	where some specific beliefs are linked together in order to describe what an agent thinks about himself, the partner and the environment during a relationship. (more)
EXPECTATIONS IN TRUST		GAMBETTA: Trust as subjective probability with which an agent assesses that another agent or group of agents will perform a particular			An expectation is not just an evaluation about the future but a combination of a belief with a goal: it is only with respect

		action. BARBER: Trust is a set of social expectations about partners, about organizations and institutions in which we live, and about natural and moral social orders. BACHARACH: The expectations about an agent are based on the questions the trustor has to answer by reading that agent's "manifest".			to a goal that an agent can have a belief about the future which is relevant for trust decisions. (more)
ROLE AND CATEGORY IN TRUST	The propensity for trust is given on the basis of membership, such as social roles, social categories, jobs etc. KRAMER: Category-based trust refers to trust predicated on information regarding a trustee's membership in a social or organizational category. Role-based trust is	BARBER: Individuals can adopt a sort of presumptive trust based upon knowledge of role relations, even in the absence of personalized knowledge or history of prior interaction. BREWER: Membership in a salient			We consider the role and the category of an agent as sources of trust: role and category can be vessels for information about the competence of an agent (mechanic), about his willingness (priest) or about both of them

	predicated on knowledge that a person occupies a particular role in the organization rather than specific knowledge about the person's capabilities, dispositions, motives, and intentions.	category can provide a basis for presumptive trust. ORBELL: Gender is a social category which discriminates between trust propensities. BACHARACH: To trust is therefore a problem of signs interpretation: considering what we know and what we see about a person, we can decide if we should trust him or not.			(doctor). This is a very easy falsification source of trust. (more)
DYNAMICS AND LEARNING IN TRUST	KRAMER: Analysis about how trust grows or dies within organizations. MILLER: Organizational rules can foster trust through their effects on individuals' self-perceptions and their shaping of expectations about other organizational members. SLOVIC: Trust	GAMBETTA: Trust is a peculiar belief predicated not on evidence but on the lack of contrary evidence makes it vulnerable to deliberate destruction. In contrast, deep distrust is very difficult to invalidate through			Trust can spread among agents: trust elicits trust, reciprocity and transitivity effects, trust atmosphere, etc. (more)

	is easier to destroy than create: negative events are more visible and influential than positive events.	experience, for either it prevents people from engaging in the appropriate kind of social experiment or, worse, it leads to behaviour which bolsters the validity of distrust itself.			
TRUST & ENVIRONMENT	FUKUYAMA: The presence of high trust in a society is followed by a greater economic wealth. A country's economic prosperity is correlated with the amount of social capital existing within that country.	YAMAGISHI: Cross-cultural study about trust in USA and Japan. BUCHAN/CROSON: Cross-cultural study about trust in US and China.			We distinguish between trustor's beliefs concerning the partner (trustee) and beliefs related to the environment where the partner acts and to the temporal dimension in which the trustee plan will develop. (more)
.TRUST & (IR)RATIONALITY	HOLLIS: It is not easy to explain why we trust: social or economical theories developed since the Enlightenment haven't	GAMBETTA: Even in the absence of 'thick' trust it may be rational to trust and distrust. If we do not, we shall			We should distinguish between rationality as intended by economics (trust can be an irrational behaviour

	succeeded in resolving the contrast between trust and rationality.	never find out.			then) and complete irrationality (trust is not irrational). (more)
TRUST & GAMES THEORY	AXELROD: The PD illustrates the paradox that if everybody were to behave in such a way, they would risk actually ending up with less than they might have obtained had they chosen to cooperate with each other.	GOOD: Sociological analysis of numerous games developed by economists and psychologists.	DEUTSCH: The Trucking Game allows to the study the role of communication in two individual dilemmas.	SEN: Classical dilemmas applied to MAS. Simulations regarding negotiation, learning and cooperation.	
TRUST & SOCIAL CAPITAL	KRAMER: Trust is one feature of the social capital of any organisation. It allows three benefits: 1) transaction costs reduction (WILLIAMSON); 2) spontaneous sociability among members increasing (FUKUYAMA); 3) authority credibility increasing.	COLEMAN: Social capital is a set of social resources that facilitates resolution of collective action problems. PUTNAM: Social capital is features of social organization, such as trust, norms, and networks, which can improve the efficiency of society by facilitating coordinated action.			

TRUST & TECHNOLOGY (HCI, CSCW)		NISSENBAUM: Technology and security are not a panacea for trust problems.		TAN: Technology has a great influence on trust therefore to understand it could help in building trust towards technical products.	Technology can be a base for trust but it cannot resolve the problem. Moreover, much technical solution can have counterproductive effects. (more)
TRUST & MORALITY		BARBER: trust as a fiduciary obligation transcends technically competent performance and implies a moral dimension in social interaction. USLANER: Trust in other people, is based upon a fundamental ethical assumption: that other people share your fundamental values. O'NEILL: trust is needed precisely because there are no guarantees.			The morality of a partner Y is to be considered. The motivation belief is one of the beliefs involved in the trust framework: X believes that Y has some motives to help him (i.e. to adopt his goal), and that these motives will probably prevail - in case of conflict - on other motives, negative for him. (more)