Cognitive network dynamics in chatlines

Available online at www.sciencedirect.com
ScienceDirect
Procedia Computer Science 1 (2010) 2355–2362
Procedia Computer Science
www.elsevier.com/locate/procedia

International Conference on Computational Science, ICCS 2010

Cognitive network dynamics in chatlines

Andrea Guazzini[a,*,1], Pietro Lió[b], Franco Bagnoli[c], Andrea Passarella[a], Marco Conti[a]

[a]*Institute of Informatics and Telematics - IIT - CNR, via G. Moruzzi,1 - 56124 Pisa, Italy*
[b]*Computer Laboratory, University of Cambridge, 15, J.J. Thompson Avenue, Cambridge, CB30FD, UK*
[c]*Dept. Energy and Centre for the Study of Complex Dynamics - CSDC, University of Florence, Via S. Marta, 3 - 50100 Florence, Italy; also INFN, sez. Firenze.*

Abstract

Beyond their common use for interpersonal communication, chatlines (also chatrooms) can be formalised as dynamic systems with heuristics. We have studied chatlines in the framework of social networks. The design and data analysis of chatlines opens a new interesting research direction in social network studies. It provides the opportunity of studying the dynamics of human social behaviour in experimental 'controlled' (or nearly controlled) conditions. Our study aims to point out both the analogy with physical systems of interacting objects and the social network emerging properties linked to the existence of different communication patterns and usage of different heuristics in the participants. We describe guidelines for effective implementation of a chatline in controlled experimental conditions. We identified several parameters which represent meaningfull statistical estimators of the activity of the network and we computed the correlation of these parameters and measures of network statistics.

Key words: Heuristics, Social dynamics, Complex network analisys, Cognitive Psychology, Chatlines

1. Background

Social relationships are commonplace in human and animal societies, with dynamic characteristics, massive scalability, inclusivity, self organisation and remarkable topological and structural properties. While communication and knowledge exchange are fundamental and natural tenets of animal behaviour, as technology progresses, we are seeing increasing numbers of communication devices that have the capability to store, process and forward information in human everyday physical environments. Increasingly, this is opening up the opportunity to generate, collect and communicate new knowledge and information in diverse aspects of human life. Multiple human relationships imply explicit and implicit social structures that have further desirable properties, such as highly connected structures with distinct topological traits such as short paths. Conversely several current researches focus on investigating and exploiting the structures defined by human social relationships to design better adaptive networking systems and integrated communication devices. The tools of complex network analysis are today fundamental to validate the social networking paradigm from the technological standpoint (feasibility and effectiveness), and to identify the technical potential for integrating communication, data and knowledge management, and security techniques within social

*Andrea Guazzini
Email addresses: andrea.guazzini@iit.cnr.it (Andrea Guazzini)
[1]Corresponding author

1877-0509 © 2010 Published by Elsevier Ltd.
doi:10.1016/j.procs.2010.04.265

communication networks [4, 5, 8, 3]. Complex network theory has also proved important for its application to ICT [11, 7]. Within the framework of human-centred networking, i.e., communication networks made up from people with their hand-held devices, the exploitation of concepts such as social communities and social hubs has been the key strategy for designing routing and content dissemination algorithm. This is the case of the ContentPlace dissemination protocol [6], whose main idea is to use nodes that are central from the social network standpoint in order to carry and distribute content items across the network. Specification of the common solutions (i.e., prototype future applications of the social networking paradigm), will involve specification of typical real-world environments and scenarios where exploitation of social networking under human mobility can be performed [1]. To achieve these targets we need to define conditions of reproducibility and select a set of meaningful human communication markers or parameters to characterize issues ranging from trust, communication and content provision. We identified in chatlines a tool providing an optimal environment where to analise entirely decentralised formation and maintenance of social networks, with heterogeneity of participants; massively scalability and inclusivity with embedded hierarchies and communication protocols. Chatline could be considered in the frame of dynamical networks which constitute a very wide class of complex and adaptive systems. A good understanding of network theory is therefore of basic importance for complex systems such as a chatlines. We may push the tool further by trying to use dynamical parameters and measures of centrality, such as the betweeness, to characterise the chatline. This would provide an estimate of the reproducibility of experiments while we perturb initial conditions, change partecipants, etc. What we have found is that this approach is not sufficient to understand the phenomenology observed and needs to be linked to cognitive psychological aspects such as heuristics. Cognitive psychology is the discipline that more than others has developed scientifical paradigms to investigate how human beings face with *social problem solving*, avoiding to start from the neural structures, but rather dealing with the empirical relationship among ”atomic” identifiable processes into the human behavior [13]. Within this framework, the basic blocks are represented by mental schemes and cognitive heuristics [16, 17, 18, 19, 20, 15]. A mental scheme is a cognitive procedure activated by the context or by interactions. Schemes are innate (reflexes actions) or learned (training) and in some context more than one scheme may be activated. When no schemes are sufficiently activated, and maybe time pressure requires that an action has to be initiated, the choice of the scheme is decided by heuristics. The activation of a scheme or of a heuristic can be detected by reaction times. The key concept characterizing the cognitive approach is the cognitive heuristic. We can define the cognitive heuristics as those procedures adopted by an individual to force the activation of one and only one scheme, to learn/build new schemes and to manage existing schemes [10]. From this point of view heuristics are a sort of meta-schemes that alter the (perceived) context. For instance, the elements causing ambiguity in the context are removed (ignored). Cognitive sciences have shown how the evolution of the above mentioned cognitive structures (e.g. heuristics and mental schemes) through diversification and selection processes [14], have addressed people’s lack of hability to make completely rational judgements/choices, due to insufficient time and information, with sets of systematic biases, rules of thumb, or heuristics. The study of such biases has allowed the identification and description of the principal cognitive heuristics. Among others some heuristics have received in recent years significanta ttention in the research community, and sometimes a computational definition has been proposed. *Anchoring* is a kind of heuristics that simply states ”take just one aspect of the context and decide upon it”, generally the as pect with high emotional content. The reaction may be adjusted considering other aspects of the context. *Availability* is related to examples that can be retrieved quickly from memory, sometimes overestimating the probability of the occurrence of a situation. *Representativeness* is related to similarities between hypothesis and memorized dependencies. Finally chat line could represent the ideal environment to study the cognitive strategies for the *social problem solving*. Since we know the human tendency to structure very fast a mental representation of the environment in order to predict it, is it possible to search for regularities among inviduals behaviours and their mental representation of the community. Moreover a chat session among strangers users represents the ideal condition to study both the individual strategies of network survey and the process of formation of mental schemes for the interactors.

2. A cognitive inspired framework to study human heuristics by web-based communication

In order to analyze and interpret experimental data, it is necessary to have in mind a theoretical framework, i.e., a model. The ingredients of the model to be examined are: the individuals, the messages among individuals and the network structure of the resulting interactions.

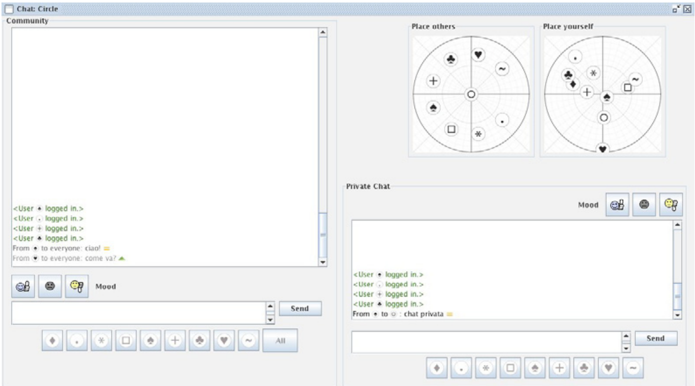

Figure 1: Chat interface

Real humans are indeed complex, so that one has to deal with rather simplifications. The crudest one is represented by a linear automaton without memory. In this approximation, a message $\mathbf{y}$ that an individual receives is decomposed into a certain number M of independent factors j, so $\mathbf{y} = (y_1, y_2, \dots, y_M)$. This decomposition can be different from individual to individual (misunderstanding), but for the moment we just suppose it to be unique. The processing of a message is given by a linear superposition of weigths w_j, that represent the "tastes" of an individual, i.e., his psychological profile. Therefore, the response or opinion x of an individual i to a message $\mathbf{y}$ (that can also be a question in a questionnaire or a good in a supermarket shelf) is $x_i = \sum_j w_{ij} y_j$. As usual, one includes all aspects not captured by this model into a "noise" term.

Such a model is at the basis of factorial analysis. The linear dependence of outputs from inputs allow the use of principal component analysis, and in general variance analysis. Indeed, it can be shown that a simple linear model can capture also a moderate amount of nonlinearities [12]. Such a model allows the anticipation of individuals preferences based on correlations, without the need of knowing the number of "hidden" dimensions M or the individual's tastes [2]. Essentially, an individual is represented as a perceptron; and factors might be identified with functional areas in the brain.

Even assuming the simple linear superposition model, one has to design or discover the internal dynamics of weights w_{ij}. These weights have a slow dynamics (learning) and a fast one (boring). Indeed, and psychology experimentalists know that even the order of questions in a questionnaire is important, and can affect the results.

We can assume that interpersonal communication follow similar rules. The advantage in analyzing interpersonal communication is that messages are accessible to experimentalists, especially if they are in a digital form like in a computer chat. We are therefore lead to analyzing the dynamics of small groups.

A group is formed by three or more persons in interaction. It is assumed that up to roughly 10 persons it is possible for an individual in the groups to have a mental representation for each others. One can call this a small group.

On the other hand, measurements on agendas and cellular phones [9] have shown that the usual number of persons that are personally known by an individual is about 150. This could be defined as a medium group. In such a group (say, a classroom), one is unable to know what all other people are doing, but he can maintain an personal affinity with any other individual.

Larger groups are crowds. Crowds have are felt as "distinctive" objects, with a given personality, that may result from an average evaluation of individual behavior. Clearly, even in a crowd one can be surrounded by individuals belonging to his small or medium group.

3. Chatline as a laboratory to study small group communication dynamics

We have developed a chatline environment that allows the recording of the interactions among a group of ten participants. We have tried to arrange an environment that mimics that of an unstructured group of stranger participants. Each participant has at his disposal a console with two textual windows, one for communicating with the rest

Figure 2: Chat panel for message typing. Above the message textbox are allocated the mood buttoms, while below the textbox are present the bottom to select the message addressees.

of participants in a public way, and one to communicate to a selected subset in a private way. This should model loud conversations and whispering. Since people do not see each other, we have included two "radars" in the interface, in which the symbols representing other participants may be placed (1). One radar is public, seen by others, and user can only move his own symbol. The other is private, and one can move all others symbols while his own is always at the center. In this way we are trying to offer an equivalent to external non-verbal communications (the public radars) similar to changing place in order to be closer to a given person, and a mnemonic aid (the private radar) for the representation of others' identities and their perceived social proximity, as seen by each user. In order to corroborate this interpretation, messages coming from a given individuals are darker is the individual is close in the public radar, and vanishingly clear if she is far. Finally, each message can have a "mood", represented by a a small icon with thumb up, down or neutral (2). This should condense the non-verbal content of a message (as usual in textual chats). Since the semantic content of a message depends crucially on the cultural context, we are trying to extract informations only from the timings of messages and the network of connections [12]. For each of the communication channels (private and public textual chats, public radar displacements, moods) we have extracted the relevant events from the log file, and computed the activity indicators defined below. The activity indicator is a matrix W^t whose element W^t_{ij} gives the number of messages sent by individual i to individual j at each time step.

$$W^t_{ij} = \sum_t W^t_{ij}. \tag{1}$$

The matrix W is not symmetric. In order to taken into account the psychological qualities of the interactions, different matrices W_{ij} are defined by mean considering the mood and the *space* (e.g. public or private) where the interaction take place. Subsequently the users activity produces nine different spaces. The global messages space W^{GM}_{ij} is obtained by considering all the interactions in every spaces and with every mood, while the spaces identified by the messages exchanged into the public, labeled as community, and the private environment are respectively indicate as W^{CM}_{ij} and W^{PM}_{ij}. Finally considering the different possible moods of a message as positive (W^{CposM}_{ij}), neutral (W^{CneuM}_{ij}) and negative (W^{CnegM}_{ij}) we have defined three different subspaces for the CM and the PM respectively. From the matrix W one can construct a stochastic matrix a normalized on the columns, so that a^t_{ij} is the probability of sending a message from j to i in the unit of time at time t.

$$a^t_{ij} = \frac{W^t_{ij}}{\sum_j Wi.^t}. \tag{2}$$

The first quantity to be considered is the in-degree k_i of a given individual i, defined as $k_i = \sum_j W_{ij}$. A similar quantity of the out-degree: $k_j = \sum_i W_{ij}$. The quantity $(a^2)_{ij}$ is the probability that a message from j reaches i in two steps, summed over all the possible intermediate. The diagonal of a^2 gives the probability that a message travels back to i in two time steps (a two-loop). This measure represents a good equivalent of the connectivity degree for the directed weighted network.

$$k^t_i = (a^2)^t_{ii}. \tag{3}$$

Similarly, one can define the probability of larger loops or paths, possibly avoiding already visited sites (by putting at all steps $(a^n)_{ii} = 0$. It is also easy to extract the most probable path that connects two given sites, and from that one can define the diameter of the graphs (the largest number of steps of the most probable paths among all pairs of sites), the most central node (that belonging to the largest number of most probable paths), the betweennes (the average number of nodes in the most probable paths). The same measures of connectivity and centrality have been computed

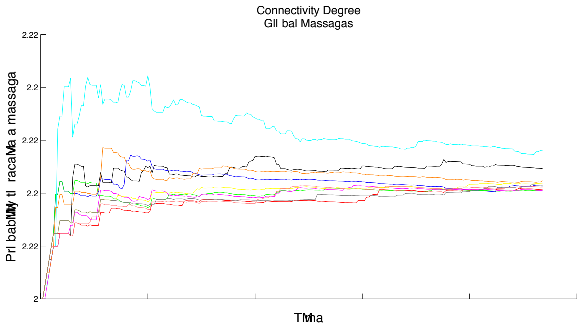

Figure 3: Centrality degree of individuals for Global Messages Space. The centrality degree is here intended as the probability to receive a message from another one into the network in the unit of time. Lines identify different individuals.

for each subject also considering their private radar distances, obtaining the matrix d_{ij}. Finally a standard linear correlation is used to investigate the relation between the interaction pattern among individuals and their psychological representation of the virtual community represented by the final state of their private radar (4). In order to correct the potential bias affecting this kind of experiment, for each subject we have computed the correlation between the radar distances and either with the absolute probability of interaction (a_{ij}^{GM}) and with the probability of a positive interaction (a_{ij}^{CposM}) (e.g. interactions characterized by a positive mood).

$$r_i = \frac{\sum_j^n (a_{ij} - \overline{a_{i.}})(d_{ij} - \overline{d_{i.}})}{(n-1) s_{a_{i.}} s_{d_{i.}}}. \tag{4}$$

where $\overline{a_{i.}}$ and $\overline{d_{i.}}$ are respectively the mean of the vector of the probability to have at least one communication and the mean of the vector of the radar distances for the subject i. This quantity is the covariance of $a_{i.}$ and $d_{i.}$ for the subject i; $s_{a_{i.}}$ and $s_{d_{i.}}$ are the standard deviation of $a_{i.}$ and $d_{i.}$ respectively.

4. Results and Discussion

We have conducted a series of experiments considering 5 communities of 10 different subjects in interaction for 45 minutes. The subjects did not know each other and a randomly assigned nickname were associated by the system to each of them.

Before the experiments all subjects were instructed in the same way and underwent the same training on the system. The instructions was symply to freely interact with others and structuring within the end of the experiment a representation of the *affinity* with others using the private radar tool.

Figure 3 describes the centrality degree, where the *y axis* represents the probability to exchange a message with another participant at each time step, and the *x axis* represents time. Figure 4 describes the centrality degree with respect to the private messages space, again the *y axis* represents the probability to exchange a private message with another participant at each time step.

Comparison figures 3 and 4 provide a measure based on centrality degree for public and private messages spaces respectively, which follow qualitatively different dynamics. While the global space is characterized by a quasi stationary and stable trend, the dyamics of private messages shows a more rich dynamics where individual move towards an increasing differentiation. From a psychological point of view is remarkable that for all the five experiments the trend of the centrality degree for all the subjects were stabilized at least before the third of the session. Moreover for all the experiments the network of total communications at the end of the session were full connected with the weights ranging between 0.09 and 0.16.

Figure 5 shows the dynamics of the mean distances between each individual and all the others in their private radars. Distance among the individual i and individual j, sayng d_{ij} is defined as the geometrical distance of their icons into the private radar of i. Also in this case a process of differentiation can be observed and it is possible

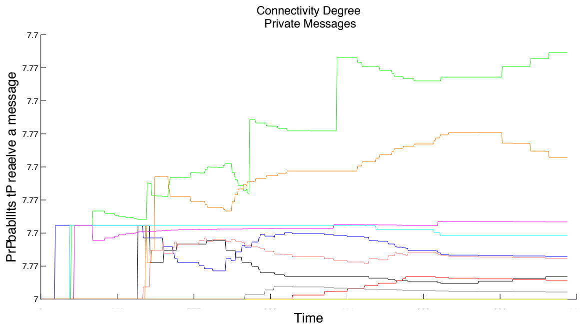

Figure 4: Centrality degree of individuals for Private Messages Space. Probability to receive a private message from another one into the network in a unit of time. Lines identify different individuals.

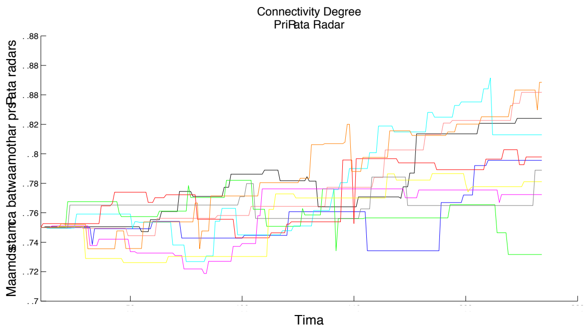

Figure 5: Mean distance between an individual and others in others' private radars. The range of the distances into radars is (0, 1). Lines identify different individuals

to hypothesize that it represents the management of the newborn "mental schemes of others", intended as a stable structure of preattentive analysis activated selectively for a specific interactor. This space allows to observe from another perspective the process of leader formation. The agent with the greater connectivity degree is the leader, and the dynamics of his election by the group can be monitorized and estimated through this space.

Figure 6 shows the correlation between the final representation of the community on the private radar and the probability to have a communication. Time proximity of changes on private radar and communication messages rate suggest that all subjects actively structure their cognitive representation of environment (mental schemes) according with the exchanged messages. Moreover the largest part of the sample shows a significant correlation between the distance of an interactor on the own radar and the probability to send him a message. The accord between the two correlations represented in 4 is here used to evaluate both the coherence of subjects and/or their comprehension of experimental instructions. A subject can be considered as coherent if the sign of correlation between the radar distance and the probability to have a positive interaction is less than zero. Obviously since the subjects were asked to communicate with others freely, it is surprising to discover that only 6 subjects over 50 have adopted the eclectic behaviour of placing those individuals with whom they interact more in distant positions in their private radars.

It is necessary to underline that four of the incoherent subjects had managed very little their own radar, suggesting a probable misunderstanding of the instructions.

Almost all subjects show a characteristic pattern of interaction with others during the early stage (i.e. the first 45 minutes of interaction among strangers) of the communication dynamics. This aspect suggest that chatline can be an effective tool to investigate the dynamics of an unstructured small group of stranger participants and study yje emergence of heuristics.

The design and data analysis of chatline opens a new interesting research direction in social network studies. It

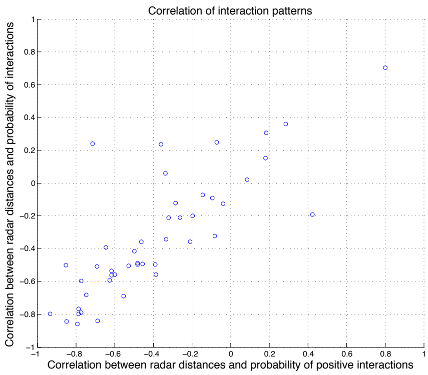

Figure 6: Correlation between private radar distances and probability of interaction. The X axis represents the correlation between the radar distances and the probability of a positive interaction for each subject. The Y axis represents the correlation between the probability of a generical communication with the radar distances.

provides the opportunity of studying dynamics of human social behaviour with an high degree of reproducibility and nearly controlled environmental conditions.

Our study would like to highlight both the analogy with physical systems of interacting particles and the social network emerging properties linked to the existence of different ”*thresholds*” and usage of different heuristics in the participants.

Althought the absence of a specific task, the exploration of the chatlines' dynamics for a small region of space parameters has produced intresting evidencies. The achieved results have underlined that the individual communication strategies seem to evolve in a very similar way.

Two different aspects of results suggest that subjects adapt ”common” heuristics in order to interact effectively with the artificial environment: the time needed from the subject to reach a stable state of the communication dynamics and the presence of a little number of outlier using radars in an opposite way compared to the others.

We believe that our efforts have two principal feed-backs: firstly they can be considered as suggestions or guidelines for building effective chatlines to the investigation of the human social behaviours, then, given the absence of a specific task (blank modality), the results can be used as a baseline for future and more complex experiments.

Finally we are now improving the interface in order to promote the use of nonverbal add-in, private radars and targeted messages, and we are extending this instrument by embedding task oriented tool (e.g. games, problem solving, etc) in order to link more closely psychological and network measurement with heuristics.

Specifically, the work presented in this paper starts exploring how humans build social structures from scratch in ICT systems. From a user behavioural standpoint, this certainly involves cognitive processes responsible for building ”self-awareness”, i.e., by way of which the users understand the (social) environment around them, based on the information they can get. As subject of future work, we are investigating how to model these processes through heuristics, as mentioned in Section 1. This looks a very promising direction. Heuristics are actually quick and efficient ways to process partial information and provide approximate answers. Therefore, heuristics seem particularly suitable to model the cognitive processes providing self-awareness, which, from a purely functional standpoint, provide ”quick and approximate” answers starting from partial information about the surrounding environment.

Acknowledgments.
This research is supported by the European Commission under the Socialnets (217141) FET Project.

References

[1] S. M., Allen, M., Conti, J., Crowcroft, R., Dunbar, P., Lio, J. F., Mendes., R., Molva, A., Passarella, I., Stavrakakis, R. M., Whitaker, *Social Networking for Pervasive Adaptation*. Perada Workshop on Pervasive adaptation, (2008).

[2] F. Bagnoli, A. Berrones and F. Franci, *De Gustibus Disputandum - (forecasting opinions by knowledge networks)*. Physica A **332**, 509 (2004).

[3] R. Albert, A., Barabasi, *Statistical mechanics of complex networks*. Rev. Mod. Phys. **74**, 47 (2002).

[4] A., Barrat, M., Barthelemy, R., Pastor-Satorras, A., Vespignani, *The architecture of complex weighted networks*. PNAS **101**, 3747 (2004).

[5] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, *Complex Networks: Structure and Dynamics*. Physics Reports **424**, 175 (2006).

[6] C., Boldrini, M., Conti, A., Passarella, *Design and performance evaluation of Content Place, a social-aware data dissemination system for opportunistic networks*. Elsevier Computer Networks **54(4)**, 589 (2010).

[7] M., Cha, A., Mislove, B., Adams and K.P., Gummadi, *Characterizing social cascades in Flickr*. Proc. ACM workshop of Online Social Networks, (2008).

[8] S. N. Dorogovtsev, J. F. F. Mendes and A. N. Samukhin, *Structure of Growing Networks with Preferential Linking*. Phys. Rev. Lett. **85**, 4633 (2000).

[9] R.I.M Dunbar, *Neocortex size as a constraint on group size in primates*. Journal of Human Evolution **20**, 469 (1992).

[10] G. Gigerenzer, P.M. Todd and the ABC Research Group, *Simple Heuristics That Make Us Smart*, (Oxford University Press, Oxford UK 1999).

[11] Newman ME, Girvan M. *Finding and evaluating community structure in networks*. Phys Rev E Stat Nonlin Soft Matter Phys. **69**, 26113 (2004).

[12] V Nguyen, Z Koukolikova, F Bagnoli, P Lio, Noise and nonlinearities in high-throughput data, J. Statistical Mechanics (2009) P01014

[13] M.R. Rosenzweig, S.M. Breedlove, N.V. Watson, *Biological Psychology: An Introduction to Behavioral and Cognitive Neuroscience*, (Sinauer Associates, Sunderland Mass. 2004).

[14] P. Slovic, B. Fischoff and S. Lichtenstein, *Facts versus Fears: Understanding Perceived Risk, in Judgment Under Uncertainty: Heuristics and Biases* in emphJudgment Under Uncertainty: Heuristics and Biases, D. Kahneman, P. Slovic, and A. Tversky (Eds.), (Cambridge University Press, New York 1982).

[15] E. Shelley Taylor, *The Availability Bias in Social Perception and Interaction*, in *Judgment Under Uncertainty: Heuristics and Biases*, D. Kahneman, P. Slovic and A. Tversky (Eds.) (Cambridge University Press, New York 1982)

[16] A. Tversky and D. Kahneman, *Availability: A heuristic for judging frequency and probability*. Cognitive Psychology **5**, 207 (1973).

[17] A. Tversky and D. Kahneman, *Judgment under uncertainty: Heuristics and biases*, Science **185**, 1124 (1974).

[18] D. Kahneman and A. Tversky, *Prospect theory: An analysis of decision under risk* Econometrica **47**, 263 (1979).

[19] A. Tversky and D. Kahneman, *The framing of decisions and the psychology of choice*, Science **211**, 453 (1981).

[20] D. Kahneman, P. Slovic and A. Tversky (Eds), *Judgment under uncertainty: heuristics and biases* (Cambridge University Press, New York 1982).