EUROPEAN SCIENCE FOUNDATION
SETTING SCIENCE AGENDAS FOR EUROPE

Exploratory Workshop Scheme

Standing Committees for
- Physical and Engineering Sciences (PESC)
- Social Sciences (SCSS)

ESF Exploratory Workshop on

Science and Technology of Agreement

Barcelona, Spain, 18-21 June 2008

Convened by:

Oscar Vilarroya ①, Ramon Lopez de Mántaras ② and Carles Sierra ②

① "Social Brain Chair".Universitat Autonoma de Barcelona, ES
② Artificial Intelligence Research Institute (IIIA), Spanish Research Council (CSIC), Barcelona, ES

SCIENTIFIC REPORT

EXECUTIVE SUMMARY

The capacity to reach agreements is crucial for all individuals living in society. Indeed, without agreements there is no cooperation and, therefore, societies cannot exist. Until recently, sociologists and philosophers were the only ones in charge of studying the mechanisms envolved in reaching agreements. More recently, other disciplines such as social neuroscience, biological anthropology, and artificial intelligence have also entered into the scene of the study of agreements and are contributing to its understanding. From the artificial intelligence viewpoint even enlarging the concept to model agreements in artificial societies such as multi-nt Systems. The aim of this Workshop has been to bring together researchers from social psychology, biological anthropology, social neuroscience and artificial intelligence to establish a fruitful dialogue among these disciplines. In particular we have discussed which requirements are needed to construct artificial multi-nt Systems based on existing results in the other disciplines as well as the usefulness of artificial multi-nt systems to simulate social systems in general.

The Workshop was structured in two sessions, one for each day. The first session on June 19 was chaired by Prof. Ramon Lopez de Mantaras and was devoted to computacional and logical models of agreement. The first technical presentation of this first session was given by Prof. Carles Sierra and he introduced the concept of "Electronic Institution". Electronic Institutions are computacional environments that organize collective activities by establishing a restrictied environment where all interactions place according to certain conventions. This talk raised questions such as what is and what is not an Electronic instituion? or the difference between dialogical and physical interactions. The second talk was given by Prof. John Debenham and was on the topic "Information-based ncy". He defended the use of information theory as an alternative to game theory. This talk raised questions such as the relation between uncertainty and information. The third talk was on "Conventions and Commitments in nts" and was given by Dr. Pablo Noriega. He put forward the idea that Internet is a turning point in artificial intelligence because it makes possible the concept of "Collective Artificial Intelligence". This talk raised questions such as how to distinguish "convention" from "commitment" and wether or not a norm pressuposes a commitment.

The fourth talk addressed the issue of “Logics, Emotions and Agreements” and was given by Prof. John-Jules Meyer. He described a formal modelization of cognitive nts that extends the traditional one based on beliefs, desires and intentions (BDI). This talk raised questions such as wether the logical specification of emotions captures the definition of those emotions or the role of emotions in negotiation processes and in reaching agreements. The fifth talk was given by Prof. Wiebe van der Hoek and was on multi-nt Systems and negotiation. He actually focused his talk on the issue of Social Laws within the context of a temporal logic called “Alternating Time”. This talk raised questions such as its relation to Deontic Logics, its applicability or wether this approach could model the internal states of nts or just their observed behaviour. The first day ended with a “discusión session” where some of the questions raised during the day were further discussed along with some newly raised questions.

The second day started with a welcome by the chairman of the second session, Dr. Oscar Vilarroya. He also reminded the scope of the day which focused more on approaches from social psychology, biological anthropology and social neurosciences along with a talk on computacional models of social interaction and another talk on computacional models of argumentation and negotiation.
The first talk was on “Trust and Misunderstanding in Social Interaction” and was given by Dr. Jan-Willem van Prooijen. He actually focused his talk on the issue of procedural justice and trust in social interaction. This talk raised again the issue of the importante of norms because norms are closely related to the concept of justice. The second talk was on “Social Cognition in Primates” and was presented by Prof. Josep Call. He focused on the issue of agreement and cooperation in animals. This talk raised a very interesting discussion about the importante of having motivations (or goals) to eng in joint activities. This is a crucial lacking notion in artificial nts. The third talk was on “Computational Models of Social Interaction” and given by Prof. Cristiano Castelfranchi. He focused on the cognitive foundation of the notions of agreement and trust. This talk raised the question wether “Contracts” and “Norms” replace Trust.

The fourth talk was given by Dr. Oscar Vilarroya.The fifth, and last talk, was on “Argumentation and Negotiation” presented by Prof. Katia Sycara. She focused on computacional models of argumentation and negotiation. This talk raised important questions regarding how far the computacional models of cooperation and negotiation are from the human counterparts.

The subsequent discussions, at the end of the day, clearly showed that although the computacional models started inspired by human cooperation and negotiation, we are still very far from the human (or even other animals) capabilities but time has come to pay more attention to human negotiation and cooperation. From these two days of fruitful discussions, we ended up with many questions and a lack of answers but we think that this is what we really were aiming at. As a matter of fact, in science it is crucial to ask ourselves pertinente questions. For those of us involved on building artificial intelligence multi-nt systems it became very clear that time has come to pay a lot of attention not only to human but also to other animal societies if we really want to come up with believable multi-nt systems capable of solving problems cooperatively. The participants from other disciplines have seen that approaches based on logic and artificial intelligence might be a good alternative/complement to game-theoretic approaches to model and simulate societies. Finally, given the amount of discusión raised by concrete issues such as “norms” or “trust”, we think that a future workshops should focus on specific aspects of the science and technology of agreement instead of trying to grasp the whole issue of agreement. This is perhaps one of the main conclusions of this first workshop.

SCIENTIFIC CONTENT OF THE EVENT

The capacity to reach agreements is crucial for all individuals living in society. Indeed, without agreements there is no cooperation and, therefore, societies cannot exist. Until recently, sociologists and philosophers were the only ones in charge of studying the mechanisms envolved in reaching agreements. More recently, other disciplines such as social neuroscience, biological anthropology, and artificial intelligence have also entered into the scene of the study of agreements and are contributing to its understanding. From the artificial intelligence viewpoint even enlarging the concept to model agreements in artificial societies such as multi-nt Systems.

The aim of this Workshop has been to bring together researchers from social psychology, biological anthropology, social neuroscience and artificial intelligence to establish a fruitful dialogue among these disciplines. In particular we have

discussed which requirements are needed to construct artificial multi-nt Systems based on existing results in the other disciplines as well as the usefulness of artificial multi-nt systems to simulate social systems in general.

FIRST DAY

The Workshop was structured in two sessions, one for each day. The first session on June 19 was chaired by Prof. Ramon Lopez de Mantaras and was devoted to computacional and logical models of agreement. The session started by welcoming the participants and reminding the aim of the Workshop. After that, the programme had forseen the presentations from two representatives of the PESC and SCSS comittees of the ESF but, unfortunately, they could not come.

The first technical presentation of this first session was given by Prof. Carles Sierra and he introduced the concept of "Electronic Institution". Electronic Institutions are computacional environments that organize collective activities by establishing a restrictied environment where all interactions place according to certain conventions. They act as an interface between decision-making mechanisms of individuals and social tasks and mimic tradicional institutions. This talk raised questions such as what is and what is not an Electronic instituion? or the difference between dialogical and physical interactions

The second talk was given by Prof. John Debenham and was on the topic "Information-based ncy". He defended the use of information theory as an alternative to game theory to model negotiation and agreements. This talk raised questions such as the relation between uncertainty and information (too much information may increase uncertainty) or where do probabilities come from?

The third talk was on "Conventions and Commitments in nts" and was given by Dr. Pablo Noriega. He put forward the idea that Internet is a turning point in artificial intelligence because it makes possible the concept of "Collective Artificial Intelligence" and collective AI needs conventions and commitments from the component nts. This talk raised questions such as how to distinguish "convention" from "commitment" and wether or not a norm pressuposes a commitment (the concept of "Norm" played a very central role in the discussion session at the end of this first session).

The fourth talk addressed the issue of “Logics, Emotions and Agreements” and was given by Prof. John-Jules Meyer. He described a formal modelization of cognitive nts that extends the traditional one based on beliefs, desires and intentions (BDI). It extends BDI along two main directions: Adding “emotions” (that have an influence on *deliberation*), and programming normative Systems (that include regimentation and enforcement of norms as well as a sanctioning mechanism). This talk raised questions such as wether the logical specification of emotions captures the definition of those emotions or the role of emotions in negotiation processes and in reaching agreements.

The fifth talk was given by Prof. Wiebe van der Hoek and was on multi-nt Systems and negotiation. He actually focused his talk on the issue of Social Laws within the context of a temporal logic called “Alternating Time”. He argued that it is a crucial aspecto for achieving coordination in multi-nt systems. It is woth noticing that social laws are actually norms that constrain the behaviour of the nts by forbidding performing actions in certain cases. He defended that Alternating Time Temporal Logic captures social laws in which legality of actions can be expressed. This talk raised questions such as its relation to Deontic Logics, its applicability or wether this approach could model the internal states of nts or just their observed behaviour.

The last talk of the first day was on “Logic of Information in Distributed Environments” and was presented by Dr. Marco Schorlemmer. He argued that the current approaches to perform semantic alignment (or semantic “agreement”), which assume that the semantic agreement is prior to the interaction, has important limitations because often semantics is interaction-specific (context specific) and, therefore, interaction should be prior to semantic agreement. This talk raised the question of how to have a successful interaction without any prior common understanding. Indeed, some shared commodity is required but this does not need to be a shared semantics in the sense of existing approaches to semantic alignment.

The first day ended with a “discusión session” where some of the questions raised during the day were further discussed along with some newly raised questions. As we mentioned above, possibly the most debated question was that of the concept of “Norm” and in particular how do norms raise in societies? As we will see later, this concept was again extensively discussed during the session of the second day.

SECOND DAY

The second day started with a welcome by the chairman of the second session, Dr. Oscar Vilarroya. He also reminded the scope of the day which focused more on approaches from social psychology, biological anthropology and social neurosciences along with a talk on computacional models of social interaction and another talk on computacional models of argumentation and negotiation.

The first talk was on “Trust and Misunderstanding in Social Interaction” and was given by Dr. Jan-Willem van Prooijen. He actually focused his talk on the issue of procedural justice and trust in social interaction and he addressed the fundamental question: “Can we trust others not to take advant of us?” He enphasized that this is a asymmetrical relation because there are “authorities” and “subordinates” and there is a need for mutual trust. To establish mutual trust we need procedural justice: The extent to which subordinates believe that authorities take decisions by means of fair procederes. He then focused his talk on the influence of procedural justice on the extent to which subordinates trust authorities and on the extent to which subordinates behave in a trustworthy manner. For both cases, he presented experimental results showing that procedural justice shapes trust particularly among people who are dispositionally distrustful of others (proselfs) and that procedural justice also has substancial consequences for cooperation in a group setting where trust is required. This talk raised again the issue of the importante of norms because norms are closely related to the concept of justice.

The second talk was on “Social Cognition in Primates” and was presented by Prof. Josep Call. He focused on the issue of agreement and cooperation in animals. He defended the hypothesis that chimpanzees and humans share an appreciation of others’ psychological states of attention and intention and he described several exprimentally supported results indicating that chimps indeed have a theory of mind. However he also argued that chimps and humans differ in the sharing of those psychological states and the motivation to eng in a variety of joint activities. In summary: agreement and cooperation is widespread in the animal kingdom; the mechanisms supporting agreement and cooperation differ between species; humans and chimpanzees share an appreciation of others’ psychological states but they differ in sharing them and the

motivation to eng in joint activities; and such ultra-sociality is key to understand the emergence of human culture. This talk raised a very interesting discussion about the importante of having motivations (or goals) to eng in joint activities. This is a crucial lacking notion in artificial nts. We will see below that this issue was also raised by the next talk.

The third talk was on “Computational Models of Social Interaction” and given by Prof. Cristiano Castelfranchi. He focused on the cognitive foundation of the notions of agreement and trust. He argued that the notion of “agreement” has different meanings and levels and it is intrinsically based on the mental states of the nts. Within these different levels and meanings one can distinguish between sharing beliefs (epistemic agreement) and sharing goals (motivational agreement). However, “full” agreement requires modeling goal-adoption and this adopted goal is a common goal shared by the nts and therefore some form of communications (interaction) is needed. Regarding “trust” he argued that is a fundamental base for arriving to an agreement and, viceversa, the existence of agreements is a very substantial ground for Trust. He also raised the question wether “Contracts” and “Norms” replace Trust. That is, people put contracts in place precisely because they do no trust the others.

The fifth, and last talk, was on “Argumentation and Negotiation” presented by Prof. Katia Sycara. She focused on computacional models of argumentation and negotiation. She argued that in order to cooperate and negotiate, nts need to be able to reason about their own beliefs and about the beliefs of others. In her own work, the computacional models are embedded into nts that interact in a simulated environment in various scenarios of interest. The nts can cooperate and negotiate with one another and have the capability of performing “what-if” analysis. Human users can interact with the nts that embody such computacional models. This talk raised important questions regarding how far the computacional models of cooperation and negotiation are from the human counterparts. The subsequent discussions, at the end of the day, clearly showed that although the computacional models started inspired by human cooperation and negotiation, we are still very far from the human (or even other animals) capabilities but time has come to pay more attention to human negotiation and cooperation.

ASSESSMENT OF THE RESULTS, FUTURE DIRECTIONS

From these two days of fruitful discussions, we ended up with many questions (see above) and a lack of answers but we think that this is what we really were aiming at. As a matter of fact, in science it is crucial to ask ourselves pertinente questions. For those of us involved on building artificial intelligence multi-nt systems it became very clear that time has come to pay a lot of attention not only to human but also to other animal societies if we really want to come up with believable multi-nt systems capable of solving problems cooperatively. The participants from other disciplines have seen that approaches based on logic and artificial intelligence might be a good alternative/complement to game-theoretic approaches to model and simulate societies. Finally, given the amount of discusión raised by concrete issues such as “norms” or “trust”, we think that a future workshops should focus on specific aspects of the science and technology of agreement instead of trying to grasp the whole issue of agreement. This is perhaps one of the main conclusions of this first workshop.

FINAL WORKSHOP PROGRAMME:

DAY 1:

Agreement Technology. Agreement models			
Timetable	**Speaker**	**Discipline**	**Topic**
9:00 – 9:15	Ramon L. de Mántaras	IA	Welcome and objectives of AT
9:15 – 10:45	Carles Sierra	IA	Electronic institutions
10:45 – 11:30	John Debenham	Engineering	Information-based ncy
11:30 – 11:45	Coffee Break		
11:45 – 12:30	Pablo Noriega	IA	Conventions and Commitments in nts
12:30 – 13:15	John Jules Meyer	Computer Science	Logics, emotions and agreements
15:00 – 15:45	Wiebe van der Hoek	Computer Science	Multi-nts Systems and Negotiation
15:45 – 16:30	Marco Schorlemmer	IA	Logic of Information in Distributed environments
16:30 – 16:45	Coffee Break		
16:45 – 18:45	Discussion		

DAY 2:

Agreement Technology. Computational models of trust and reputation			
Timetable	**Speaker**	**Discipline**	**Topic**
9:00 – 9:15	Oscar Vilarroya	Cognitive Science	Welcome and objectives of Day 2
9:15 – 10:45	Jan-Willem van Prooijen	Cognitive Science	Trust and misunderstanding in social interaction
10:45 – 11:30	Josep Call	Cognitive Science	Social cognition in primates
11:30 – 11:45	Coffee Break		
11:45 – 12:30	Christiano Castelfranchi	Cognitive Science	Computational Models for Social Interaction
12:30 – 13:15	Oscar Vilarroya	Cognitive Science	Neurobiology and Decision Making
15:00 – 15:45	Katia Sycara	Robotics	Argumentation and negotiation
15:45 – 16:15	Coffee Break		
16:15 – 18:15	Discussion		

STATISTICAL INFORMATION OF PARTICIPANTS.

Sascha Ossowski (age: 42) Universidad Rey Juan Carlos -Madrid-

José Santiago Pérez (age: 30) Universidad Rey Juan Carlos -Madrid-

Moser Fagundes (age: 28) Universidad Rey Juan Carlos -Madrid-

Viviane Torres (age: 30) Universidad Rey Juan Carlos -Madrid-

Mª Carmen Romero (age: 28) Escuela Superior de Ingenieros -Sevilla-

Mari Carmen Delgado (age: 32) Escuela Superior de Ingenieros –Sevilla-

Adolf Tobeña (age: 58) Autonomus University of Barcelona

Carlos Carrascosa (age: 40) University of Valencia

Stella Heras (age: 28) University of Valencia

José Such (age: 26) University of Valencia

Juan M. Alberola (age: 30) University of Valencia

Natalia Criado (age: 29) University of Valencia

Àngels Fàbregues (age: 25) Institute of the Spanish Research Council

Enric Plaza (age: 51) Institute of the Spanish Research Council

Jordi Sabater (age: 33) Institute of the Spanish Research Council

Teresa Garzón (age: 33) The Social Brain Chair

Josep M. Rabanal (age: 50) The Social Brain Chair

Susanna Carmona (age: 30) The Social Brain Chair

Joan Duran (age: 32) The Social Brain Chair

FINAL LIST OF PARTICIPANTS:

-Ramón López de Mántaras.	**Professor of the Artificial Intelligence Research, Institute of the Spanish Research Council, Barcelona-**
-Josep Call Balaguer	**Wolfgang Kölher Primate Research Center, Germany-**
-Christiano Castelfranchi	**Director of the Instituto di Scienze e Tecnollogia della Cognizione, Roma, Italy-**
-John Debenham	**Engineering University of Technology, Sydney-**
-Òscar Vilarroya Oliver	**Director of the Chair "The Social Brain" Autonomus University of Barcelona**
-John Jules Meyer	**Professor of Computer Science Utrech University, Netherlands-**
-Pablo Noriega	**Researcher of the Artificial Intelligence Research Institute of the Spanish Research Council, Barcelona-**
-Carles Sierra	**Professor of the Artificial Intelligence Research Institute of Spanich Research Council, Barcelona-**
-Katia Sycara	**Director of the Intelligent Software nts Lab. In Carnegy Mellon University, U.S.A.**
-Marco Schorlemmer	**Researcher of the Artificail Intelligence Research Institute of the Spanish Research Council, Barcelona-**
-Wiebe Van der Hoek	**Head of the nt "ART" University of Liverpool, United Kingdom.**
-Jan Willem van Prooijen	**Department of Social Psycology of Free University of Amsterdam, Netherland**

Total participants: 12

Male 11; Female 1

Countries:

Netherlands: 2	**Spain: 5**	**Germany: 1**	**Italy: 1**
United Kingdom: 1	**USA: 1**	**Australia: 1**	