AI and Theories of Groups and Organizations: Conceptual and Empirical Research

AAAI'93 Workshop

Commitments:
from individual intentions to groups and organizations

Cristiano Castelfranchi

Project for the Simulation of Social Behaviour
Ist. di Psicologia - CNR
viale Marx 15 - 00137 - Roma - Italy
Email: PSCS @ IRMKANT.bitnet

(Preliminary Report)

Abstract

The aim of these notes is to introduce some notions of Commitment which are crucial for the understanding of groups and organizations' functioning, and of the relations between individual agents and collective activity. I try to identify some of the basic ingredients of such notions, and to make some steps towards their formalization. In particular, I claim that a notion of Commitment is needed as a mediation between the individual and the collective one. Before constructing a notion of "Collective or Group Commitment" a notion of "Social Commitment" is to be defined. "Social commitment" is not an individual Commitment shared by many agents; it is the Commitment of one agent to another agent. I stress the normative contents (entitlements/obligations) of this social relation, and its connections with individual intentions and collective activity. On that basis, a notion of Organizational Commitment is proposed, that could account for the structure of stable Organizations, and the related notions of "role" and of "power of command". This view, based on one side on the intrinsic normative aspects of Social Commitment, and on the other side on the dependence and power relations that found a team, a group, an Organization, is quite in contrast with the "conversational" view of Organization in which agents seem to be completely free to negotiate and establish any sort of Commitment.

1. Introductive remarks

There is an implicit agreement about Organizations in recent computational studies. Either in DAI theories of organization [BON89] [GAS91], or in formal theories of collective activity, team or group work , joint intention, and "social agents" [LEV90] [RAO92], or in CSCW approaches to cooperation [WIN87]. *Organization is accounted for by means of the crucial notion of "commitment".* "Commitment" is seen as the glue of the group, of collective activity: it links the agent with the joint goal and the common solution, it links members' actions with the collective plan, it links the members with each other.

Unfortunately, the current analysis of Organizations in terms of Commitment is quite unsatisfactory, for a number of reasons:
a) the current definitions of Commitment are insufficient to really account for stable group constitution and activity;
b) there is a dangerous confusion between the notion of "social" and the notion of "collective", so there is no theory of "social commitment" as a necessary premise for a theory of collective or group commitment;
c) the relationships among the personal commitment to an action (implied in the notions of "intention" and "intentional act"), the group's commitment to the collective act, and the commitment of a member to its group and the collective activity, are not clearly stated;
d) agents seem to be completely free (also in organizations) to negotiate and establish any sort of Commitment with any partner, without any constraint of

dependence and power relations, of norms and procedures, of pre-established plans and cooperations [CAS92b].

I would like to discuss this problem which is crucial to the relations between "AI and Theories of Group and Organization".

In particular, we would try to analyse a notion of "Social Commitment" as distinct on one side from the notion of "Individual-Internal commitment", and from the notion of "Collective Commitment", on the other side.

This notion of *Social Commitment is a relational one*; it cannot be defined with regard to one single agent (be it an individual or a group); however, this notion is not purely behaviourist, like in [SIN92]: it is not reducible to the individual commitments, but it implies individual Commitment and is analysed in terms of the mental states of the partners.

On such a basis (and on the basis of the notion of GenericCommitment) I would propose:

- a characterisation of the kind of Commitment that supports the structure of an Organization, including:
- the notion of "role";
- the claim that there is *no Organization without Obligations* (norms);
- a criticism of the "conversational view" of organization.

Such an approach is not only necessary for a good definition of our concepts and for developing a formal theory of groups, organizations, collective actions, but also for some consequences in applicative terms. For example, without such a kind of analysis it seems impossible to account for different kinds of organization (e.g. strictly cooperative Vs orchestrated), and consequently for different kinds of commitment (by role or by reciprocation, free, based on interest or benevolence, etc.), and then for different ways to solve the conflicts, both intra- and inter- agents. For ex. how to intervene when one of the agents decides to abandon the group, or the common plan ? We cannot influence appropriately this agent if we don't know exactly its kind of commitment and the related different reasons for defection [CAS93a].

2. Kinds of Commitment: Internal, Social, and Collective

We need a notion of Commitment as a mediation between the individual and the collective one. A "social commitment" is not an individual Commitment shared by many agents.

In general, it is absolutely necessary to distinguish between the notion of "social" and the notion of "collective". This is not the case in DAI and MAS: for ex. to designate the notion of a goal or intention shared by many agents and that they can not achieve independent of each other, it is used the notion of "social" intention or goal (ex.[WER88]); to designate an agent formed by many individual agents, i.e. a group or a team, it is used the notion of "social agent" (ex. [RAO92]); to denote the idea of reciprocal commitments within a team of agents it is used the notion of "social Commitment" and "social plan". *"Social" is not a synonym of "collective".*

There is a very important level of "social action", "social agent" and "social mind" where action and mind remain "individual" but they are oriented toward another social entity. At this level, before constructing a notion of "collective or group Commitment" we need a notion of "social Commitment": the Commitment of one agent to an other agent .

So let me step by step to distinguish among: internal, social, collective Commitments.

INTERNAL COMMITMENT (I-Commitment)

Internal Commitment --as [BOU92] calls it-- corresponds to the Commitment defined by Cohen and Levesque (on the basis of Bratman's analysis) [COH90]. It refers to *a relation between an agent and an action.*

The agent has decided to do something, the agent is determined to execute a certain action (at the scheduled time), the goal (intention) that was preferred is a persistent one. The way to capture such a persistence is to establish that the intention will be abandoned only if and when the agent believes that the goal has been reached, or that it is impossible, or that it is no longer motivated.

The term "Internal" is to be preferred to "Individual" because one may attribute I-Commitments also to a group. The term "psychological ", as opposed to "social", is quite misleading because also the Social and the Collective Commitments are relations among minds.

SOCIAL COMMITMENT (S-Commitment)

As said above S-Commitment is a relational concept. *It expresses a relation between at least two agents.* More precisely, S-Commitment is a 4 argument relation:

(S-COMM x y a z)

- who is committed (*x*);
- what action *x* is committed to do (*a*);
- who *x* is committed to (*y*);
- before whom *x* is committed (*z*)

Let us neglect (in this paper) the third agent (*z*), i.e. the witness, who has a very crucial role in normative contexts (norms efficacy) [CON93] and in contractual contexts implying also free riders and cheaters. Let us concentrate on the relation between *x* and *y*.

COLLECTIVE COMMITMENT (C-Commitment)

We should differentiate S-Commitment from the Collective Commitment or Group Commitment. The latter is just *an Internal Commitment of a Collective agent* or Group. In other terms, a set of agents is Internally Committed to a certain intention and (usually) there is mutual knowledge about that.

It remains to be clarified: a) if such conditions are sufficient to account for the commitment to a collective activity; b) which are the relationships between S-Commitment and C-Commitment (later).

3. Toward a Definition of Social Commitment

S-Commitment is much more than an I-Commitment of *x* to *a* known by *y*. Let us suppose that *y* is aware of some intention of *x* to do an action *a*, nevertheless, this does not allow to say that "*x* is committed to *y* to do *a*". Nor is it sufficient to suppose that *x* is aware that *y* knows his intention, etc. (Mutual Knowledge). Let me give an example: I have realised that John intends to play a bad joke on Paul, and John becomes aware of the fact that I understood his intention; nevertheless, John persists in his plan, relying on my silence. I may keep silent and become party to his trick. Even in this case it would be false that "John is committed to me to play a joke to Paul". If John drops out his I-Commitment (that I know), I can be surprised, but I am not entitled to protest .

In this situation two qualifying aspects of the social relation of "being committed to somebody" are lacking:

a) *A social Commitment is a form of "Goal Adoption"*. In other terms: *x* is committed to *y* to do *a*, if *y* is interested in *a*. The result of *a* is one of the goals of *y*; for this reason, *y* has the goal that *x* does *a*.

(S-COMM x y a) ==> (GOAL y (DONE x a))

So, *x and y share a goal (DONE x a)*. Notice that:
- this goal is an Intention of *x*'s (given that it is his own action);
- *x* and *y* have an identical goal, but *x* adopted this goal from *y* (in our terms [CAS90][CON91], *x* has the supergoal: *(OBTAIN y (DONE x a))*

b) *If x is S-Committed to y, then y can (is entitled to)*:
- control if *x* does what he "promised";
- exact/require that he does it;
- complain/protest with *x* if he doesn't do *a*;
- (in certain cases) make good his losses (pledges, compensations, retaliations, ..)

I will come back later on this necessary and very relevant ingredient of the S-Commitment: the rights of *y* on *x* created by the S-Commitment of *x* to *y*.

Let me introduce, just for the sake of exposition, this quasi-formal and incomplete definition of S-COMM:

$$(MKx\ y\ ((INTEND\ x\ (DONE\ x\ a))\ \&$$
$$(GOAL\ y\ (DONE\ x\ a)))$$
$$\&$$
$$(ENTITLED\ y\ x\ as\ for\ a))$$

where *ENTITLED* means that *y has the rights* of controlling *a*, of exacting a, of protesting (and punishing). *ENTITLED* means also that *x* is S-Commitment to *y* to not oppose to these rights of *y* (in such a way, *x* "acknowledges" these rights of *y*). So, interestingly enough, the definition of S-Commitment seems to require a recursive call of another S-Commitment of *x* itself.

4. Other Social Aspects of Social Commitment

I will consider just three main aspects:

a) *Not all the adoptions of a goal of y by x imply a S-Commitment of x to y*. What else is required ? First, the Mutual Knowledge I already mentioned. Second, *y's agreement* !
In fact, if *x* has just the I-Commitment to favour one of *y*'s goals (to do something so that *(OBTAIN y p)*), this is not sufficient (even if there is common awareness). *Y* should

accept to (is I-Committed to) achieve her goal p by means of x's action. So, the S-Commitment of x implies a S-Commitment of y to x to accept his action (y doesn't refuse, doesn't protest, doesn't say "who toldyou!",..). Without such (often implicit) agreement (which is a reciprocal S-Commitment) no true S-Commitment of x to y has been established.

b) *The very act of committing oneself is a "rights-producing" act*: before the S-Commitment, before the "promise" [SEA69], y has no rights over x, y is not entitled (by x and before z) to exact this action. After the S-Commitment it exists such a new and crucial social relation: y has some rights on x, on the very basis of the entitlement by x himself. The "title" of y is x's act of Commitment. So, the notion of S-Commitment is well defined only if it implies these other relations:

- y is entitled to what I already mentioned;
- x is in debt to y;
- x acknowledges to be in debt to y and y's rights.

In other terms, x cannot protest (or better *he is committed to not protesting*) if y protests (exacts, etc.).

c) What I said just now implies also that *if x is S-Commitment to y, he has a duty, an obligation,* he ought to do what he is Committed to.

$(S\text{-}COMM\ x\ y\ a) ==> (OUGHT\ x\ (DONE\ x\ a))$

So, when x is committed, a is more then an Intention of x, it is a special kind of goal, more cogent [CON93].

5. Relationships between Social and Internal Commitment

S-Commitment it is established through an overt or tacit communicative act. It implies the intention of x that "y believes that x is S-Committed to y to do a" and "that y believes x is I-Committed to do a". This is because in S-Committing himself x want that y intend to accept and to control his behaviour (so that y's knowledge is a necessary condition for x's S-Committing).

This communicative act has very relevant consequences on the relations between S-Commitment and I-Commitment. S-Commitment may be "sincere" or "insincere": x lets y believe that he intends to do what y

needs. If the communicative act and the S-Commitment are "sincere", x's S-Commitment to y to do a also implies an I-Commitment to do a. But, if x is not sincere, if he is deceptive or just a light person, in this case he is not I-Committed to do a.

It is really remarkable that even in that case, x is S-Committed to y to do a, he actually got an Obligation to do a. So, x's I-Commitment on a is neither a necessary nor a sufficient condition for his S-Commitment on a. Just y's belief that x is I-Comm to a, is a necessary condition of x's S-Commitment.

Anyway, let me simplify. Let me consider that our agents are always sincere (like in other models of Commitment). That is why I put Mutual Knowledge in my temptative definition of S-Commitment. Given this postulate, we may remark that *the S- Commitment of x to y to a, implies an I-Commitment of x to a.*

In our approach, it is possible to state such a precise relation between the two notions, because also the S-Commitment is analysed in terms of the mental states of the partners. It is not a primitive notion (like in [SIN92]), but is not "reducible" to the I-Commitment, because it is an intrinsically relational/social notion (among agents), and it contains much more than the I-Commitment of the involved agents.

I should also mention the **Self-Commitment**, which is the S-Commitment of an agent to itself ($x=y$), with the same deontic consequences of the S-Commitment to another agent. This kind of Commitment is specially interesting in human psychology, I mentioned it to avoid confusion with the simpler I-Commitment.

6. Relationships among Internal, Social, and Collective Commitment

Is a true CollectiveCommitment of a group of agents just the sum of the I-Commitments of the members, or does it require S-Commitments among those members ? I think that there is no univocal answer. It depends on the kind and nature of the group (there are many kinds of groups or collective agents [CON91] [CAS93a]).

Commitments in strictly "cooperative" groups

Let us consider first a true "cooperative" group which in our sense is based on a Common Goal and Mutual Dependence.

More precisely, *true cooperation is defined in terms of Mutual Knowledge of Mutual Dependence* relative to an Identical goal of x and y.

Mutual Dependence is defined as follows: x Depends on y for doing the action $a1$ relative to his goal p, and y Depends on x for doing the action $a2$ relative to her goal p.

An agent x *Depends on an agent y for doing an action a* relative to a state p [CAS92a] when:

- p is a goal of x's
- x is unable to do a (and to realise p)
- y is able to do so
- action a is in a plan useful to achieve p

If everybody knows to be dependent on the actions of other agents, everybody wants that these others do their share and wants to do its own, for the common goal. Then, in a fully cooperative group, a S-Commitment of everybody to everybody arises: each one has to do its own job. Given that the members form the group, we may say that *each member is S-Committed to the group* to do his share [SIN92].

So, the Collective-Commitment (defined as the I-Commitment of a collective agent) will imply (at least in the case of a fully cooperative group):

a) the S-Commitment of each member to the group: x is S-Committed not simply to another member y, but to the all set/group X he belongs to;

b) the S-Commitment of each member to each member (included itself); then also many Reciprocal Commitments and the SelfCommitments of each member;

c) the I-Commitment of each member to do his action.

Who is entitled to protest, given that S-Commitment is characterized by entitlements ? The group: each member or the authority in charge (if such authority exists).

Commitment and the notion of Collective Agent

To define a notion of Collective Agent just on the basis of the I-Commitment (or of the individual intention), of its sharing among the members, and on the members' mutual knowledge, is, in my view, a fallacious attempt.

Members of a "social agent" -- in [RAO92]'s terms -- or of a group or "team" -- in [LEV90]'s -- have a JointPersistentGoal, that is, a realizable achievement goal associated with a mutual belief that other members of the team have an equal goal and belief. Now, it is possible to show that this notion is not sufficient to account for a teamwork [CAS93b].

Consider this real life example: the case of a couple of people stranger to each other standing at the same bus stop. It is an optional stop for the driver (i.e. he will stop only if people make some signal to stop). They know that both are waiting for the same bus: so, they have an identical intention ("to take the bus n^ 3") and they mutually know that. Is this a sufficient basis for a collective activity ? Do they form a collective agent, a team ? Unfortunately, each of them (because each knows the goal of each other and that it is identical with its own goal) has got the expectation that the other does the necessary actions; each of them relies upon each other for achieving the shared goal. The result is that the bus arrives, nobody does the signal, the bus doesn't stop: both of them lose the opportunity to reach the shared goal. What is needed ? It is needed an (implicit or explicit) agreement about a common activity, based on the awareness of the reciprocal dependence relations.

Let me give another crucial example. Prof. Montaigner, of the Institute Pasteur in France, and Prof. Gallo in the US both have the final goal p "vaccine anti-AIDS be found out" relative to the belief q that "if vaccine is discovered, AIDS is wiped out". They share all three mental attitudes described [LEV90] as necessary and sufficient conditions for a Joint Persistent Goal and then for a team:

1) they mutually believe that p is currently false;

2) they mutually know they all want p to eventually be true;

3) it is true (and mutually knowledge) that until they come to believe either that p is true, that p will never be true, or that q is false, they will continue to mutually believe that they each have p as a "weak achievement goal" relative to q and with respect to the team.

Where a "weak achievement goal" with respect to a team has been defined as "a goal that the status of p be mutually believed by all the team members".

But no-one would stretch oneself up to saying that Prof. Gallo and Prof.Montaigner form a team. Indeed, given their "parallel goals" ("I discover the vaccine"), they might come to strongly compete with each other.

What else is needed for them to form a teamwork?

Without the belief about the mutual dependence also the Commitment to participate, to do one's own share is unmotivated, it is irrational. The belief in an dependence relation or in a "necessity to collaborate" is a fundational condition for a really cooperative work and team [JEN92] [CON91].

Now, my point is the following: if the members acknowledge their mutual dependence and the need to collaborate, if consequently they (implicitly or explicitly) agree about a (already specified or to be specified) common activity, have they got only an I-Commitment or better have they got S-Commitments to the others ? *Is the S-Commitment a necessary condition for the constitution of a Collective Agent ?*

Commitments in other kinds of collective activity and group structure

In my view, the current characterization of the group activity and collective agents is neither sufficient (as we saw) nor necessary. In many kinds of natural group, team, organizations people participate in a collective activity without sharing the same goals or the final end of the group or organization. Nor are they even expected to have such joint mental states. Let me mention just an other "cooperation" model I call "orchestrated cooperation" [CAS92c]. Suppose there are three agents : a boss A, and two executors B and C. Suppose that only A knows the final goal of the coordinated activity he requires from B and C, only A knows the complete plan and the reciprocal dependence relations between B and C. It is even possible that B and C ignore each other. The cooperative plan is in one and only one mind: the boss'. Between respectively A and B, and A and C, there are S-Commitments based on social exchange relations (not on true cooperation): they are not interested in the result of the plan but just in their personal benefits (rewards). There is no S-Commitment between B and C, who in fact collaborate in a coordinate way, and are members of a team.

Even in these cases we may say that the group is I-Commmmitted to do *a* (or to achieve a certain goal), because this is the explicit plan and intention that organizes the action of the participants and determine their S-Commitments and their I-Commitments. But, in these case, the I-Commitment of the group (C-Commitment) doesn't correspond to identical I-Commitments of the members.

However, there is an inportant constraint about the relation between the members' Commitments and the C-Commitment: the former should be "instrumental" to the latter. In other terms, the S-Commitments of the members (and their consequent *I*-Commitments) should be Commitments to an action wich is part of the plan (complex action) the group is C-Committed to achieve.

7. Generic Commitment, and Structural or Organizational Commitment

A *GenericCommitment is a Commitment to a class of actions*: x is Committed to do any instance of such a class A, any action of that kind.

(GenericCOMM x A) = for all *a* (where *a* is an instance of A) *(COMM x a)*

True Organizations are not extemporary, built up at the moment. They are not made of Commitments to do a specific action at a specific moment: so they use GenericCommitments. More precisely, they are made of Commitments to Commit oneself to do the right thing at the right moment.

These Commitments to Commit oneself determine the "structure" of the Organization. They are different from the running Commitments involved in the structuring of the collective activities of the Organization. So, *the structure of the Organization is different from the structures of its activities.* The former partially determines the latter.

These **Organizational Commitments** are also Generic S-Commitments:

(OrgCOMM xX A) =
for all *a* (((REQUEST X x (DONE x a))* ==>
==> *(S-COMM x X (DONE x a))*

where *a* is an instance of *A*, and *x* is a member of *X* (set, collective agent).

When the member *x* of the group *X* is organizationally Committed to his group, he is Committed to accept the requests of the group (or of a specific member, e.g. the boss) within a certain class of actions (his *office*). Then, *x*'s OrgCommitment to *X* implies that if there is a request of *X* to *x* about an action of the class *A*, *x* is automatically S-Committed to *X* to this action, he automatically gets an obligation to do *a*.

The nature of the GenericCommitment may give rise to controversies. In fact, there may be different points of view between *x* and *X* (or *y*) as to whether a specific requested action *a* is or isn't a member of the class *A*.

The "Role" of an agent is relative to the group or the organization. It is *the set of the structural Commitments of the member to his group*. So, the organizational role is neither a functional behavior, nor a task or a set of tasks. It is a "normative" notion: it is a set of behavioral obligations based on the structural agent's Commitments to the group relative to certain classes of possible actions to be requested or expected (tasks)

A GenericS-Commitment of x to y to A (like the OrgCommitment), gives to y a special *"power of influencing"* x [CAS90], a "Command Power" over x. In fact, x is committed to comply with any request of y as for A: x is "benevolent" for special reasons towards y as for A. This prevents y from the need for negotiating each time x's compliance with y's requests. This kind of "benevolence" based on such a Commitment-obligation is what we call "obedience".

This "Command Power" gives y (or X) a very special faculty: that of taking decisions for x, of deciding that x does something. This possibility of deciding about other agents' actions is quite strange : y has an "intention" about x's action.

More than this, y can take a S-Commitment for x, she can Commit x to w to do something. To be more precise: y S-Commits herself to w about an action of x, or better to require and to influence x to do this action; y is "responsible" for x's action. This S-Commitment of y to w implies an I-Commitment of y to require/influence x to do a; then, y will require x to do a. Given that x is OrgCommitted to y, from y's request a S-Commitment derives of x to y to do a. So, in fact, y has indirectly Committed x to do a.

This kind of Indirect Commitments, as the Command Power and the faculty to decide the actions of other people, are obviously a crucial and very well known feature of Organizations. This feature depends on the OrgCommitments of the members, and on their being bound by obligations

Let me, at this point, just mention a criticism about the "conversational view" of Organization. In this view, agents seem to be completely free to negotiate and establish any sort of Commitment with any partner, without any constraint of dependence and power relations, of norms and procedures, of pre-established plans and cooperations.

Not only behavior in Organizations is bound by "external" procedures, norms and rules, but even in a purely contractual perspective --based on agents' Commitments and their mental attitudes-- we saw that the agent is not free to establish his specific and extemporary Commitments. He is bound by his role (previous OrgCommitments which are Generic) and by the consequent expectations and duties. In this sense, there is no Organization without Obligations. For this reason too, the Commitments an agent establishes in his cooperative work in an Organization, are not all equivalent, nor can be handled in the same way. For example, it should be distinguished whether a Commitment is merely "personal" (both by friendship or by social exchange) or it is "by role", or it is "free" or it is "due" (by role or just by reciprocation).

Conclusions

AI current interest in group work and Organization is basically motivated by the fact that the introduction of computers and their role in supporting and mediating cooperation, absolutely requires a theoretical understanding of human cooperative activity and of Organizations. I think that AI is right in identifying the core of this problem in the notion of Commitment and in the mental representations of the agents. I tried to show that current notions are insufficient for accounting for the link between individual mind and collectivity. In particular, I introduced the intrinsically relational notion of Social Commitment (which is neither mutual nor collective) as an intermediary between the Internal Commitment and the Collective one. I analysed many relevant aspect of this social relation, with particular attention to its normatives ingredient (obligations, expectations). I think that a formalization of these notions, as well as of the notions of Organizational Commitment and of Role, could help both in the theory of Organizations and Groups, and in the computational supporting of cooperative work. My attempt has been just to clarify some useful concepts. Current views of Organization risk to be too "subjective" and too based on communication. They risk to neglect, on one side, the objective basis of social interaction (dependence and power relations), on the other side, its normative components. The notions I proposed try to relate also to these levels of analysis.

References

[BON89]   Bond, A.H., *Commitments, Some DAI insigths from Symbolic Interactionist Sociology*, 9^ International AAAI Workshop on Distributed Artificial Intelligence, 1989, pp.239-261

[BOU92]   Bouron, Th., *Structures de communication et d'organisation pour la cooperation dans un univers multi-agents.* , These de Doctorat, Universite' Paris 6, nov. 1992

[CAS90]   Castelfranchi, C., *Social Power. A point missed in Multi-Agent, DAI and HCI*. In Y.Demazeau & J.P.Muller (Eds.), Decentralized A.I, Amsterdam, Elsevier, 1990 (pp. 49-62)

[CAS92a] Castelfranchi,C., Miceli,M., Cesta,A., *Dependence Relations among Autonomous Agents.* In Y.Demazeau & J.P.Muller (Eds.), Decentralized A.I.3, Amsterdam, Elsevier, 1992 (pp. 49-62)

[CAS92b] Castelfranchi,C., Conte,R., Cesta,A., *The Organization as a Structure of Negotiated and Non Negotiated Binds*. In G.C.van der Veer, M.J.Tauber, S.Bagnara, M.Antalovits (Eds.), Human Computer Interaction: Tasks and Organization. Proceedings of the Sixth European Conference on Cognitive Ergonomics (ECCE6), Balatonfured, Hungary, September 6-11, 1992

[CAS92c] Castelfranchi,C., Conte,R., *Emergent Functionality Among Intelligent Systems: Cooperation Within and Without Minds*, AI & Society, 6, 1992, pp.78-93

[CAS93a] Castelfranchi, C., Cesta, A., *On Rational Reasons for Dropping Commitments in Joint Activities*, IJCAI 1993. Workshop on "Conflicts management in cooperative problem solving" (submitted)

[CAS93b] Castelfranchi, C., Cesta, A., Conte,R., Miceli, M., *Foundations for Interaction: the Dependence Theory*, AI*IA 3^ Conference. Torino, 1993 (submitted)

[COH90]   Cohen, P.R., Levesque, H.J., *Intention is Choice with Commitmentitment*, Artificial Intelligence, 42, 1990, pp.213-261

[CON91]   Conte, R., Miceli,M., Castelfranchi,C., *Limits and Levels of Cooperation: Disentangling various types of prosocial interaction*. In Y.Demazeau & J.P.Muller (Eds.), Decentralized A.I.2, Amsterdam, Elsevier, 1991 (pp. 147-157)

[CON93]   Conte, R., Castelfranchi, C., *Norms as mental objects. From normative beliefs to normative goals. "Reasoning about Mental States: Formal Theories & Applications"*, AAAI Spring Symposium Series, March 23-25 1993, Stanford, CA

[GAS91]   Gasser, L., *Social conceptions of knowledge and action: DAI foundations and open systems semantics*. Artificial Intelligence, 47, 1991, pp. 107-138

[JEN92]   Jennings, N.R., Mamdani, E.H., Using Joint Responsibility to Coordinate Collaborative Problem Solving in Dynamic Environments, *Proceedings of AAAI-92.*

[LEV90]   Levesque,J.H., Cohen,P.R., Nunes,J.H., *On Acting Together*, Proceedings of AAAI-90, 1990

[RAO92]   Rao,A.S., Georgeff, M.P., Sonenberg, E.A., *Social Plans: A preliminary Report*. In E. Werner & Y.Demazeau (Eds.), Decentralized A.I.3, Amsterdam, Elsevier, 1992

[SEA69]   Searle, J.R., *Speech Acts*, Cambridge University Press, Cambridge, 1969.

[SIN92]   Singh, M.P., *Social and Psychological Commitmentitments in Multiagent Systems*. In "Knowledge and Action at Social & Organizational Levels", AAAI Fall Symposium Series, 1991, pp.104-106

[WER88]   Werner, E., *Social Intentions*, ECAI-88, Munich, WG, pp. 719-723, 1988.

[WIN87]   Winograd,T.A.,*Language/Action perspective on the Design of Cooperative Work*, Human-Computer Interaction, vol.3, 1, pp.3-30, 1987