Modeling the Impact of Trust and Distrust in Agent Networks

G. Gans M. Jarke[1] S. Kethers G. Lakemeyer

RWTH Aachen, Informatik V

Ahornstr. 55

52056 Aachen, Germany

{gans, jarke, kethers, lakemeyer}@cs.rwth-aachen.de

ABSTRACT

Recently, the role of trust in agent networks has received a lot of attention. In his paper, we argue that an explicit consideration of distrust and its complex interaction with individual trust and confidence in the network as a whole is equally important, especially when designing hybrid networks of human and machine agents. We propose a trust-confidence-distrust (TCD) model of agent network dynamics, and present a multi-perspective methodology according to which such networks can be modeled and simulated. We are currently validating this methodology in the context of designing specialized computer support for networks of organizations, e.g. in the context of high-tech entrepreneurship or continuing education.

1. INTRODUCTION

Recently, there has been a growing interest in modeling trust, mainly driven by the advent of the internet and electronic commerce (see, for example, [Falcone et al. 2000, CACM 2000]). Much of the work is concerned with trust in connection with online interactions, where one of the characteristics is that agents normally do not know each other. In this paper, we focus on formalizing trust in social networks, a recently popular form of "coopetition," that promises to combine the benefits of two traditional coordination mechanisms of modern societies [Powell 1990]: the flexibility and speed of competitive market relationships, and the stability and long duration of cooperative, organizational relationships. More specifically, we are interested in networks created among independent organizations to pursue some shared strategic goals, but always with the risk of falling apart.

The relatively small existing literature in this field typically pursues two avenues: viewing trust as a subjective probability, or modeling it in logic. In the collection of papers in [Gambetta 1990], the prevalent view of trust is that of a subjective probability, which, roughly, amounts to the likelihood (assigned by the trusting agent) that another agent will perform a task or bring about a desired situation on which the trusting agent depends. Other work along this line includes [Coleman 1990], who considers trust as a decision under risk. Trust is given by a trustor if her expectations of gain (G) and the estimated probability of the trustee's trustworthiness (p) are greater than the expectation of loss (L) and the trustee's untrustworthiness (1-p) : $p\ G > (1-p)\ L$. Quantitative measures of trust can also be found in [Marsh 1994, Witkowski et al. 2000], and the game-theoretic approaches to trust. Here trust is analyzed mainly using the iterated Prisoner's Dilemma as a benchmark [Axelrod 1984, Boon and Holmes 1991, Birk 1999]. Rather than condensing trust to a single value, Castelfranchi and Falcone [1999] propose a more fine-grained model. It takes into account the agents' mental attitudes such as the trusting agent's beliefs about the trustee's opportunity, ability, and willingness to perform a desired task.

Trust being a modality, it seems natural to model trust within modal logic. Such approaches include [Demolombe 1998] and [Broersen et al. 2000]. The latter consider the notion of "agent i trusts agent j more after doing A than after doing B," which is formalized within the framework of propositional dynamic and deontic logic. Also, Castelfranchi and Falcone [1999] formalize aspects of the mental state underlying trust using a multi-modal logic [Meyer and van der Hoek 1992, Linder 1996].

---

[1] Also: GMD-FIT, Schloss Birlinghoven, 53754 Sankt Augustin, Germany

Instead of modeling trust as a modality [Yu and Liu 2000] characterize it as a so-called *soft goal* within the i* framework, which will be discussed in more detail in the next section. Among other things, i* allows one to explicitly model goals and tasks of agents and the dependencies between agents as they arise, for example, when a goal of one agent can only be accomplished with the help of another. While i* uses a graphical notation, it has a semantics based on logic [Yu 1995], using the knowledge representation language Telos [Mylopoulos et al. 1990]. It is possible to represent how the fulfillment of trust goals can change indicating an increase or loss of trust. In contrast to most other approaches, Yu and Lin's proposal is purely qualitative and the questions of how trust affects an agent's decisions or how to update trust are left open.

Interestingly, none of the approaches in the literature seem to give distrust a special status. By and large distrust is regarded as just the other side of the coin, that is, there is generally a symmetric scale with complete trust on one end and absolute distrust on the other (see, for example, [Marsh 1994]). As we shall argue below, recent sociological research has shown that the relationship between trust and distrust is much more complicated, and that indeed a reasonable amount of both trust and distrust is necessary to keep a social network successful. We condense this research in a dynamic conceptual model called the TCD model (TCD = Trust, Confidence, Distrust).

The main goal of this paper is to develop a methodology which, based on the TCD model, is suitable for the specification and analysis of social networks. We will argue that an appropriate methodology needs to support multiple perspectives which explicate the interplay between structure, agent planning and action, and communicative actions to manage expectations. Such a methodology is intended, among other things, to support the members of a social network in their own decisions by offering a clear picture of the structure and the dynamics of the network. In this regard an important goal is to provide the means within the formalism to actually simulate different network scenarios. We have chosen to build on previous work on a multi-perspective framework for the modeling and static analysis of cooperation processes [Nissen et al. 1996, Kethers 2000]. Our methodology includes the i* framework, which supports the description of important structural ingredients like dependencies between agents,[2] but also extends it in important ways to take into account the requirements of the TCD model and to integrate it with a planning perspectives [De Giacomo et al. 2000] and a language action perspective [Winograd and Flores 1986].

The rest of the paper is organized as follows. In section 2, we motivate and develop our TCD model and then show, in section 3, how it can be mapped to a practical, multi-perspective modeling environment for the analysis and support of social networks. The concepts are illustrated with an example taken from an ongoing case study in computer support for entrepreneurship networks. Finally, in section 4, we summarize the status of our prototype implementation and outline plans for further research.

2. TRUST, CONFIDENCE AND DISTRUST IN SOCIAL NETWORKS

We follow Weyer's definition of a social network as an autonomous form of coordination of interactions whose essence is the trusting cooperation of autonomous, but interdependent agents who cooperate for a limited time, considering their partners' interests, because they can thus fulfil their individual goals better than through non-coordinated activities [Weyer 2000a, Sydow 1992]. We agree with the recently established network sociology (cf. e.g. [Weyer 2000, Sydow and Windeler 2000]) in that we consider the concept of social networks as a phenomenon in its own right, which has to be dealt with by means of new approaches. The distinguishing factor of social networks is their reliance on the mutual trust of the network partners as the main coordination mechanism. While this idea has been recognized in recent literature, there has been little research on making it fruitful for the design and ongoing support of networked organizations in a similar way that business process modeling and requirements engineering have been attempting this for traditional organizations and human-machine systems. Moreover, the equally important issue of distrust in organizational networks has been largely ignored or at least over-simplified.

A typical definition in the network literature sees trust as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" [Mayer et al. 1995, p. 712]. There is no formal agreement on reciprocity, i.e. the relationship between give and take, investment and return where the partners profit mutually from the other partners' actions [Weyer

[2] An alternative approach to modeling dependencies among agents can be found in [Sichman et al. 1994].

2000a]. Often, the concept of trust is defined in a rather vague and misleadingly standardized way, disregarding the focal point of network research: what is the relationship between trust in a given situation that the trustor exhibits towards concrete persons or organizations, and the confidence in the network as a whole?

The network as a whole consists of a mesh of dependencies that is not manageable or controllable, nor even completely visible to the trustor, thus requiring *confidence in the system* (”Systemvertrauen” [Luhmann 1988]; cf. also the distinction between personal and institutional trust [Zucker 1986], and between ”facework” and ”faceless commitments” [Giddens 1990]; cf. also [Scheidt 1995, Loose and Sydow 1997]). Thus, participation in a network results in a double vulnerability, on the one hand to identifiable opportunists, on the other to the generally incomprehensible mesh of dependencies of all network partners. This *distinction between trust and confidence* plays an important role for the regulation and control of social networks. Although networks offer the advantages of organizational cooperation without the disadvantages of organizational bureaucracies and hierarchies, networks need to develop binding rules regulating members' behavior. These rules aim at facilitating trust-based interaction, e.g. by ensuring the confidentiality of information exchanged among partners, by supporting network culture (fair play), reputation, regulation of access [Jones 1997, Staber 2000], or by explicitly defining sanctions for breaches of trust [Loose and Sydow 1997, Ortmann and Schnelle 2000].

Finally, although coordination by means of trust and confidence can enable and facilitate cooperation, it has its costs. In networks, trust and confidence need to be *watchful*, i.e. the partners need to be continually aware of their investments and thus the risks that they incur. This watchfulness leads to a continuous (and potentially costly) monitoring of the individual partners' behavior (trust) and the perceived efficiency of the network as a whole (confidence). On the other hand, watchfulness may also be caused by distrust of or against individuals, where distrust is defined as the expectation of opportunistic behavior from partners, thus breaking the reciprocity of trust-based interaction.

Early research on distrust [Luhmann 1989, Gambetta 1988] treated distrust as danger to be avoided (cf. also [Scheidt 1995]). Only recently, distrust is recognized as an opportunity for making network structures less rigid, and thus more suitable for innovations (cf. [Kern 1998]). Recent investigations on conflict and distrust in organizations [Kramer and Tyler 1996, Lewicki et al. 1998] have established the fact that distrust is an irreducible phenomenon that cannot be offset against any other social mechanisms. Distrust is extremely relevant to social networks, as it not only has a negative influence on networks (as described above), but can also influence the network in a positive way.

Summarizing, we need an approach that addresses trust, confidence, and distrust as separate and simultaneous phenomena in a joint framework. We call this framework the Trust-Confidence-Distrust (TCD) model of success or failure of networks. This model is shown in the three “columns” (thick arrows) of Figure 1, each leading up from actions in the network to changes in the structure – with a feedback loop downwards to the actions via rules created by the structure. In the left columns, confidence-based decisions to incur strategic vulnerabilities create mutual dependencies, in the middle trustful decisions for risky and traceable investments increase reputation, goodwill, and moral integrity, whereas the watchful distrust on the right aggregates latent conflicts by collection, storage and (usually negative) interpretation of events. A balanced mix of all three aspects forms the small corridor for success in networks. The upper part of the figures shows three possible ways of failure caused by imbalances. On the upper left, too many dependencies and goodwill without trust may lead to *successful failure*, which refers to scenarios where networks degenerate to family-like or even mafiose relationships. In contrast, on the upper right, over-aggregated distrust may cause a final conflict for the network, which results in members leaving the network or, in the worst case, in the dissolution of the whole network. Finally, a balanced mix cannot be ensured by simply creating a lot of network rules, because this will terminate a network by turning it into an organization-like structure.

Before addressing the question of how to turn the TCD model into a modeling environment for the analysis and support of social networks, let us stress again why we believe that trust and distrust must be considered separately and not just zero-sum as in most previous approaches. For one, the two notions play very different roles in social networks: the level of trust is a key factor when agents decide whether or not to engage in a risky investment; the level of distrust, on the other hand, controls, among other things, the degree to which an agent monitors others, which can lead to a significant overhead. For another, and perhaps more importantly, both trust and distrust co-exist and vary independently of each

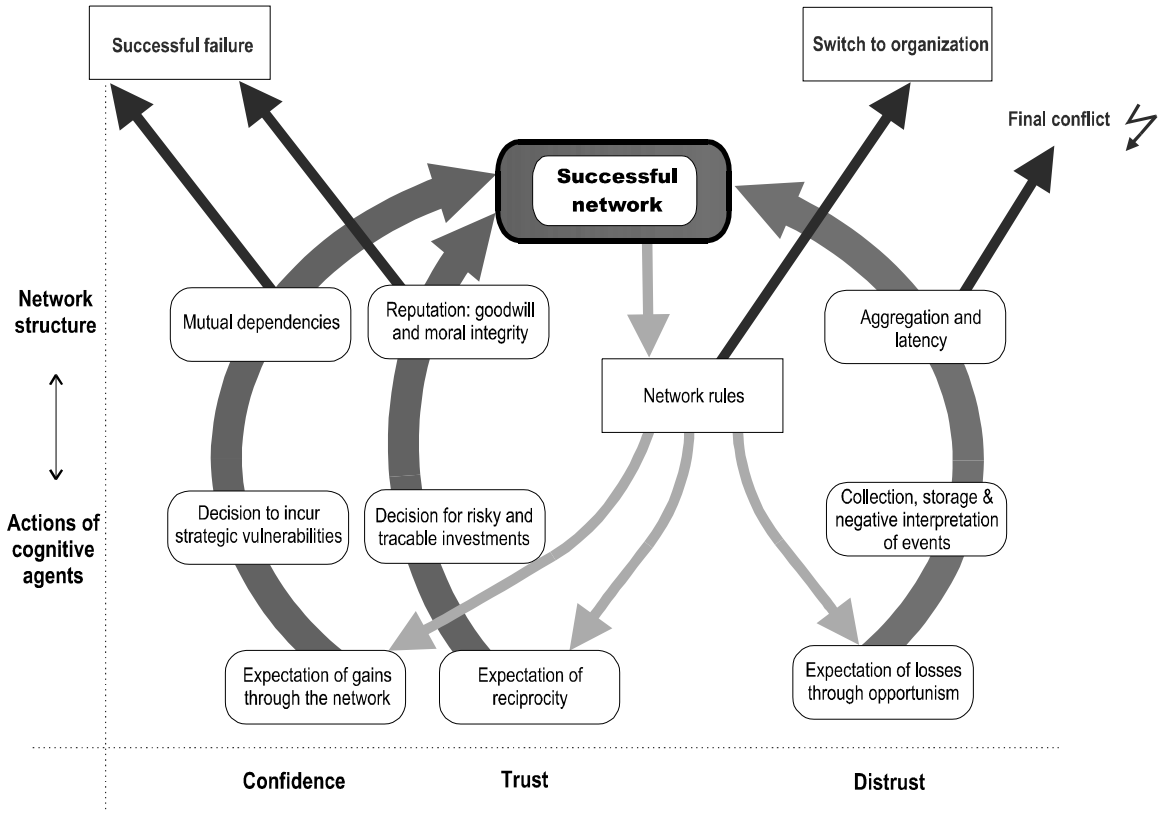


Figure 1: The Trust-Confidence-Distrust (TCD) model of social networks

other, at least to some degree.[3] For example, increasing the level of trust towards another agent and, at the same time, raising the monitoring activity, i.e. distrust, seems plausible when the latter can be viewed as a sign of healthy watchfulness. Also, as explained in [Gans et al. 2001], agents may try to hide rising distrust by not reducing or even increasing their investments, a phenomenon which can be modeled nicely by varying the levels of trust and distrust independently.

3. A MULTI-PERSPECTIVE MODELING METHODOLOGY

The previous discussion has shown that trust, confidence, and distrust in social networks are complex phenomena which are not easily captured by simplistic, single-faceted models. Previous work in requirements engineering has attempted to address such complex multi-viewpoint situations by explicitly modeling multiple, possibly conflicting perspectives or viewpoints [Nuseibeh et al. 1996, Nissen et al. 1996], and by managing their static and dynamic inter-relationships through reasoning and/or simulation mechanisms. In this section, we describe such a methodology for the TCD approach. Excessive learning efforts by requirements engineers would be sure to prevent adoption of such a methodology. We have therefore taken care to support our methodology by extended versions of well-known modeling notations, rather than inventing completely new ones.

3.1. Overview of the Methodology

The proposal developed below builds on experiences with a multi-perspective framework for the modeling and (static) analysis of cooperation processes developed in [Nissen et al. 1996, Kethers 2000] where perspectives are integrated under a meta modeling and perspective transformation and consistency checking mechanism offered by the metadata repository ConceptBase [Jarke et al. 1995], using the Telos formalism [Mylopoulos et al. 1990].

The problem at hand strongly generalizes this setting. Our goal is to formulate a methodological

[3] We do not deny that the two notions are correlated. It is just that the correlation is not as strong as often assumed.

and technically supported multi-perspective framework which includes the aspects of core/individual trust, confidence/system trust, and distrust.

Firstly, the discussion of the previous two sections showed that trust, confidence, and distrust manifest themselves in specific behavior patterns, which need to be adequately represented when analyzing such systems. In particular, we feel that an appropriate methodology must enable a *dynamic, simulation-oriented analysis* of social networks in addition to a static one. For that purpose, we integrate the logic-based high-level planning mechanism ConGolog [de Giacomo et al. 2000] in our methodology to make the related modeling and simulation capabilities available to our framework.

Secondly, the discussion above shows that the dynamics of trust, confidence, and distrust are heavily influenced by the *perceived relationships between communication acts of the agents and real action* done with respect to these communication acts. From this observation, we conclude the need to include an explicit speech-act perspective in our framework. This speech-act perspective interacts with the planning perspective provided by ConGolog.

Finally, we agree with Yu, Coleman, and many others that *explicit modeling of goals and dependencies* is crucial to understand networks in general and, in particular, the role of trust, confidence, and distrust within these networks. We therefore include Yu's strategic rationale model as well as his strategic dependency model as perspectives in our approach. However, our view here is again more dynamic than in previous work which leads to a much closer integration with the other two perspectives than investigated in previous research: Strategic dependencies are treated as the reasons for speech-act based delegations, and the latter are evaluated partially with respect to the former. Conversely, planning from strategic goals (captured in the strategic rationale submodel) may generate strategic dependencies to other actors if certain subgoals or tasks turn out to be inefficient to handle for the planning agent itself.

Thus, we have a dynamic mutual influence among the perspectives. Moreover, this mutual influence is mediated by trust, confidence, and distrust. For example, our methodology supports patterns such as the following :

- Existing core trust towards specific network agents will enhance the possibility for network action rather than individual action, and thus increase the capabilities of the network (modeled by creating more strategic dependencies and more speech-act commitments)
- Existing network trust (confidence) will enable agents to commit more rapidly to actions requested by customers, without prior communicative acts with possible subcontractors/collaborators. This significantly increases the responsiveness of the network as a whole. In contrast, lack of trust will lengthen the offer phase within a speech act, and make the network slow and bureaucratic.
- Both of the above will have an impact on the complexity, reliability and speed of collaborative action plans generated.
- Performance monitoring and thus the evolution of trust, distrust, and confidence will be based on relationships between goals, expectations (defined by communicative situations in speech-acts), plans and actual processes. A certain degree of institutionalized network distrust will be offered by the monitoring rules.

Individual distrust is not symmetric to lack of trust but will instead again change plans by adding monitoring actions to it, thus creating overhead and reducing network effectiveness in the long run. From a bird's eye view, our methodology based on the TCD model can be summarized as follows (cf. Figure 2):

1) Goal hierarchies for each agent following Yu's Strategic Rationale Approach are not just created and maintained dynamically but also mapped to operational (base) plans using the ConGolog formalism which composes a plan from declarative building blocks with pre- and post-conditions.
2) Strategic dependencies, following Yu's Strategic Dependency modeling formalism, are in part pre-existing from a-priori goal and capability analysis, in part created dynamically from the recognition that certain parts of a plan are better delegated to others.
3) Plans and dependencies, initially often based on required agent role types rather than concrete network partners, are mapped to specific communicative actions (speech-acts) in order to ensure a strategic dependency with a specific contract, which can be an informal agreement.
4) The modalities of trust, confidence, and distrust will shape the way how this will be done. They then indirectly also shape how these contract patterns refine the base plans gradually into cooperation plans in the network.

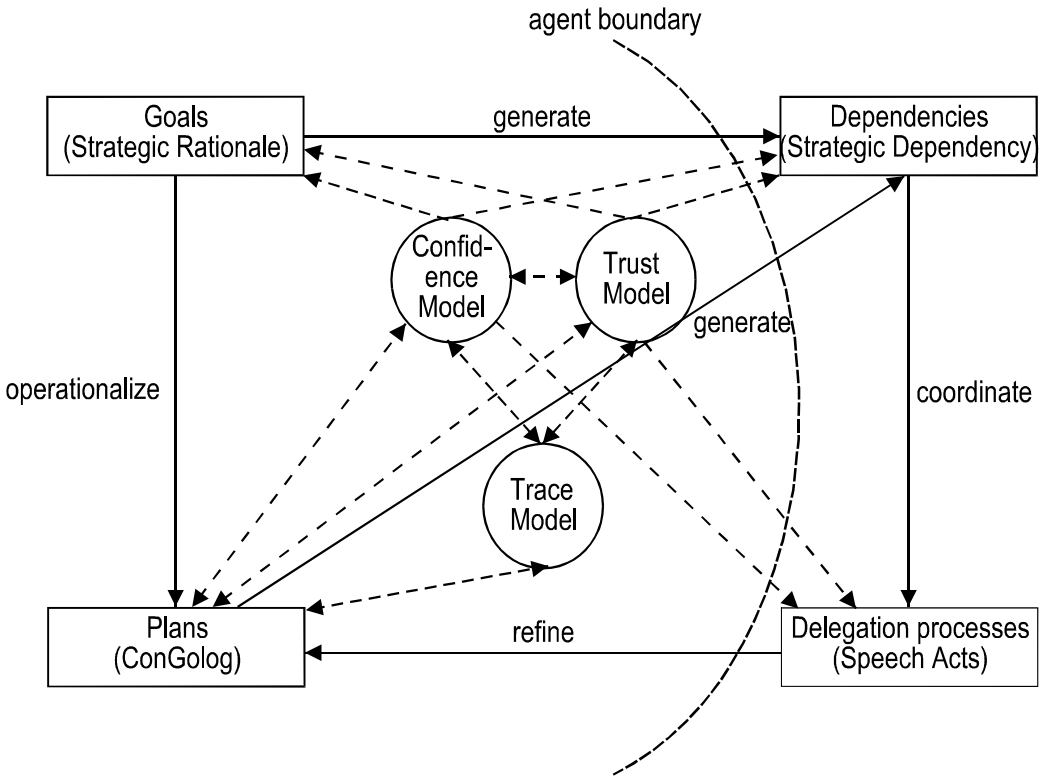


Figure 2: Overview of the Perspectives of our Framework and Types of Interrelationships

5) Monitoring in the network (Trace Model) can either be done systematically, based on network rules, again brought as modalities into preconditions of actions or speech-act patterns. Or monitoring can be done by individual agents outside the agreed procedures due to distrust. In both case, monitoring results help to re-compute, among others, the values for trust, distrust, and confidence – the influence function being a variable of the methodological framework which requires further research.

In the following, we discuss in detail two components of the methodology. The first concerns the delegation processes, the other the network structure. We also briefly touch on the issue of how to obtain ConGolog plans suitable for the simulation of network processes.

3.2. The Action Workflow Approach

Delegation processes and monitoring activities are reflected in the speech acts used in the network processes. For modeling speech acts, we use the notation of [Schäl 1996], which is based on the Action Workflow approach [Medina-Mora et al. 1992] which in turn is based on the Language-Action Perspective [Winograd and Flores 1986].

An action workflow consists of four phases. An initial *customer request*, whereby the customer states his or her condition of satisfaction, is followed by a *commitment phase*, at the end of which the supplier has agreed to accept the – possibly changed - customer request under certain conditions. During the subsequent *performance phase*, the supplier performs whatever actions are necessary to fulfill the customer request. The phase ends when the supplier reports that the commitment has been fulfilled. Finally, in the *evaluation p*hase, the customer has to evaluate whether the supplier really fulfilled the commitment.

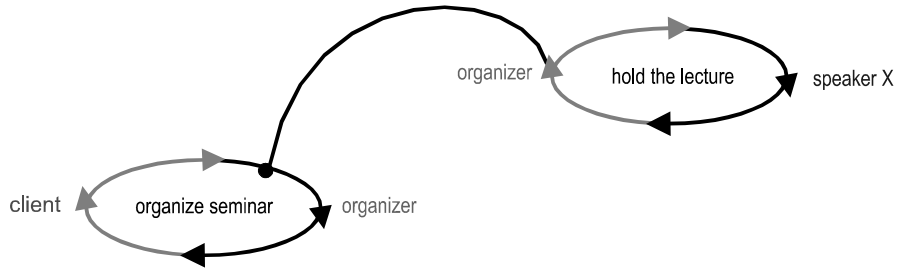
A key property of the this approach is that it works recursively, i.e. in each phase sub-work-flows of the same pattern can be spun-off with sub-contractors. If we consider the network situation, delegation can be looked at in two ways, especially in the early phases of a workflow. On the one hand, we can look at delegation or expected delegation to individual subcontractors – this is the traditional case studied in speech act theory; on the other, at expected availability and reliability of subcontractors within the network itself.

Pursuing this idea, we obtain a mapping of trust, confidence, and distrust to typical delegation patterns in the workflow model. These are shown in

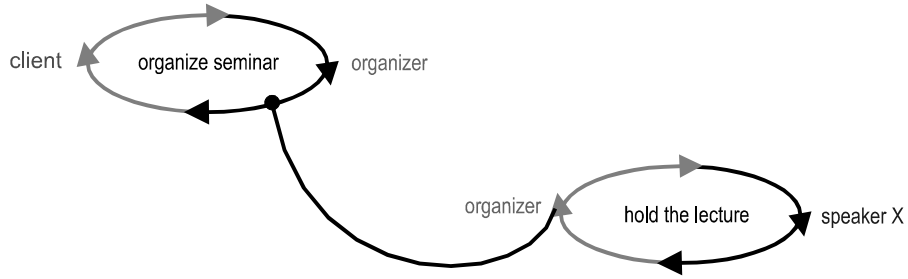
the four cases of Figure 3 below, using as an example a network of specialists who intend to cooperate in performing seminars for outside customers. The specialists considered here are a seminar organizer and a number of speakers available to hold seminars in their respective areas of expertise.

- If there is little or no trust, the seminar organizer will either do everything himself or have an elaborate delegation process in the offer phase, to ensure commitment of subcontractors (speakers) prior to committing to the customer (cases (a) and (b)).
- With growing core trust in specific individuals, the delegation activity starts to shift from the offer phase to the performance phase, but always with the same group of sub-contractors – the seminar organizer trusts in the willingness and ability of speakers to fulfill their role and can thus commit faster (case (b), case (c) with specific subcontractors only)
- Network trust/confidence is characterized by concentrating delegation on the performance phase only, with a broad range of subcontracting agents and significant sub-subcontracting which is not monitored in great detail by the original contractor (case (c)).
- Distrust within or beyond network delegation is characterized by a significant amount of activity in the evaluation phase, in particular in the delegation phase of specific subcontracts (against whose performers the distrust exists). This again demonstrates that distrust is different from the negation of trust (case (d)).

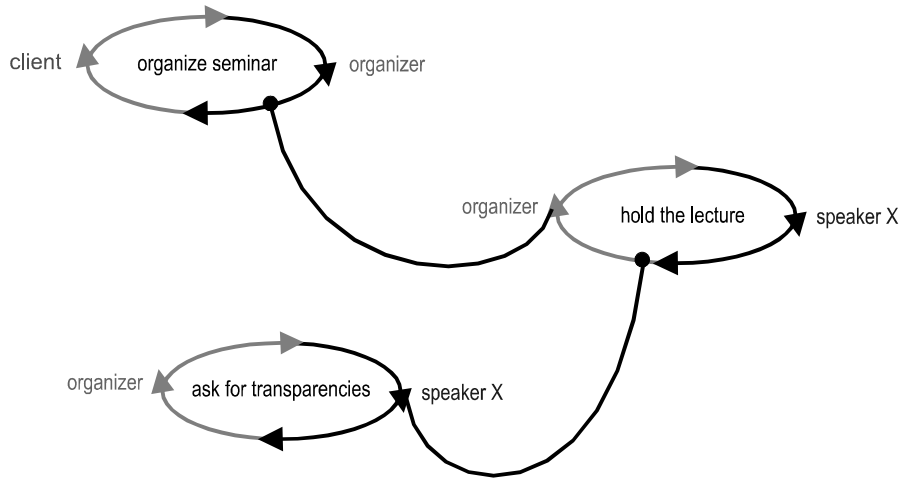
(a) no delegation



(b) limited trust requires pre-commitment of sub-contractors



(c) high confidence allows commitment prior to subcontracting



(d) distrust implies intensive evaluation

Figure 3: Relationships between trust, confidence, and distrust with expected speech-act delegation patterns

3.3. Extending the i* Framework for a Trust-Based Approach

The i* framework proposed in [Yu 1995] offers a conceptual framework for modeling social settings, based on the notions of actor and goal. It assumes that social settings involve social actors who depend on each other for goals to be achieved, tasks to be performed, and resources to be furnished. The i* framework includes the *strategic dependency (SD) model* for describing the network of relationships among actors, as well as the *strategic rationale (SR) model* for describing and supporting the reasoning that each actor performs concerning his relationships with other actors. We will not go over the details of the SD and SR models here, but instead illustrate some of their key features by way of our example.

Graphically, an SD model features actors (drawn as circles) which are connected according to the dependencies they engage in. There are various types of dependencies. In Figure 4, we only show so-called goal dependencies among actors. For example, a speaker depends on the seminar organizer for payment and for the opportunity to hold seminars in the future ("further jobs").

.

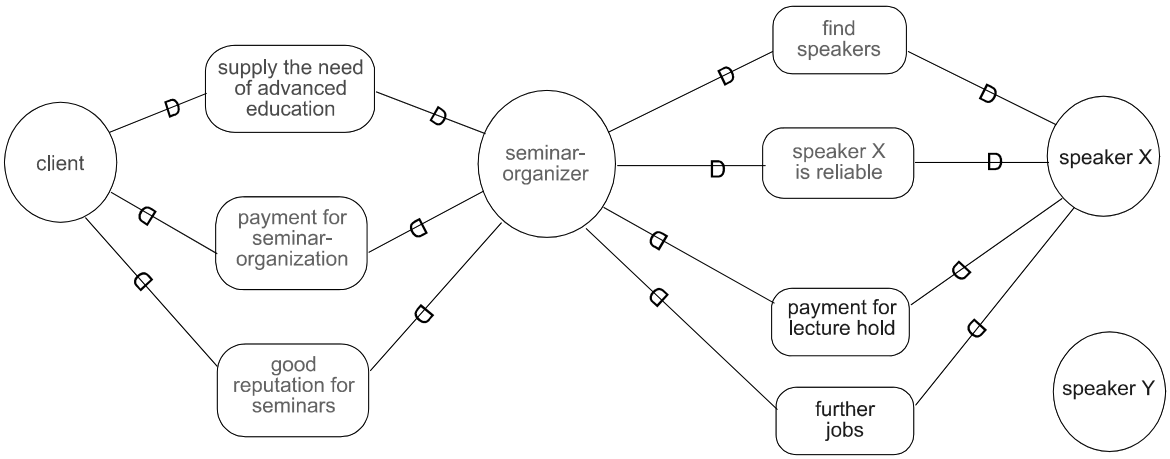


Figure 4: Strategic dependency model for seminar organization example

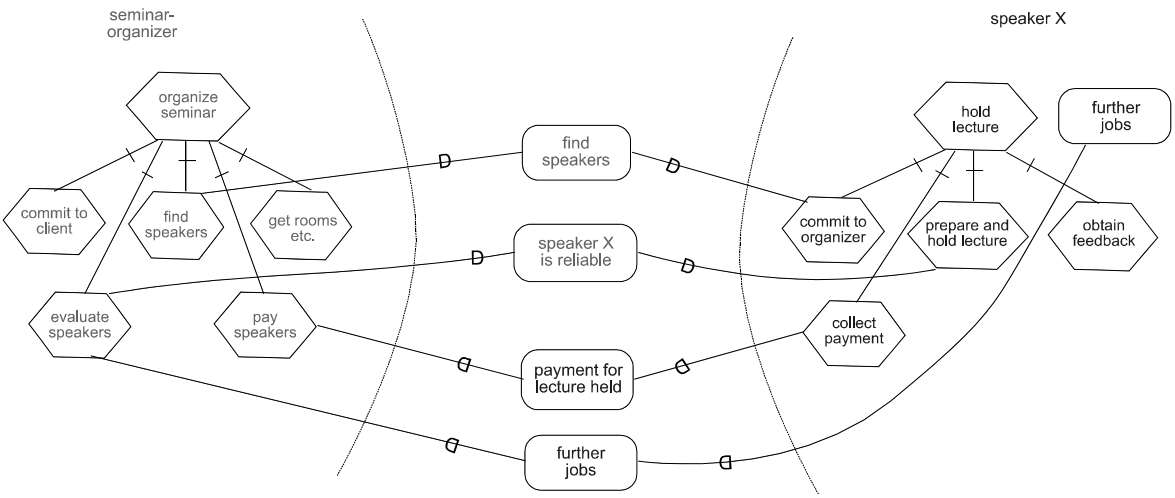


Figure 5: Strategic Rationale model of the "seminar" example

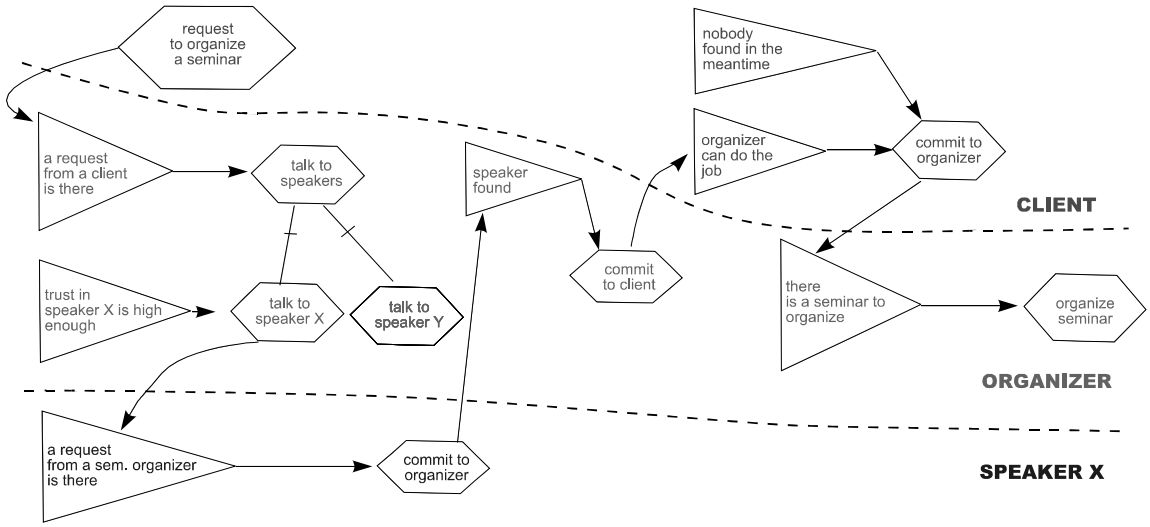


Figure 6: Organizer's plan with trust in individual speakers

The strategic rationale (SR) model describes the intentional relationships that are internal to actors so that they can be reasoned about. Process alternatives can be generated and evaluated. Elements of the model are nodes representing goals, tasks, resources, and softgoals, respectively, and links, representing either means-ends links, or task decomposition links. In Figure 5, we essentially confine ourselves to task decomposition links. The dependencies among agents are included as well. Since agents now have structure, they can be specified at a more fine-grained level of detail.

SR models contain strategically relevant elements only, hence are not suitable for operational use. Perhaps the main deficiency of SR models is that they do not provide the means for specifying an ordering of tasks. Yet in scenarios like social networks where one actor depends on another to achieve certain subgoals or tasks, the order in which things happen is essential. Other aspects not currently covered in SR models include an explicit model of time or the conditional execution of a task. The extensions we propose to SR models draw their inspiration directly from the plan language ConGolog, which has at least two advantages. For one, the concepts ConGolog provides are well-

understood and they come equipped with a formal semantics. For another, as we would like to map SR models into ConGolog plans, which are executable and thus usable for simulations, enriching SR model with some of ConGolog's features narrows the gap between the two formalisms and, therefore, eases the task of mapping one into the other.

It turns out that only very few extensions to the original SR formalism are needed. The most important new feature is the ability to specify tasks which have *preconditions* as well as postconditions.[4] This allows us to capture the ordering of events in a natural way: A task which has preconditions attached to it can only be performed if all its preconditions are satisfied, that is, multiple preconditions of a task are understood conjunctively. Only tasks can satisfy preconditions of other tasks. In our formalism, preconditions are viewed as being both necessary and sufficient for the task to be executed. In other words, as soon as all the prerequisites of a task are satisfied, the task will be performed. Graphically, (see figures below), task preconditions are denoted as triangles labeled with assertions. While the assertions are written here in natural language for readability, they should be thought of as sentences in some formal declarative language like first-order logic. Unlabeled directed edges from a task (hexagon) to a precondition (triangle) are implicitly meant to satisfy or achieve the precondition.

As examples, we consider extended SR models corresponding to the speech-act scenarios of Figure 3b and d, where the delegation of subtasks plays a key role and which illustrates how trust, confidence, and distrust crucially affect agent plans.

In Figure 6, the seminar organizer delegates holding the seminar to one or more outside speakers, provided he has enough core trust in them. An advantage of delegation in this case is that the competence is enlarged due to the accumulated competencies of the different potential trustworthy speakers. The downside is that having to find speakers before committing to the client may be time consuming. When the organizer finally is able to commit to the client, the client may already have found another organizer. In the corresponding action workflow (Figure 3, case b) the "overhead" of finding speakers first is exhibited perspicuously by the additional speech-act cycle emanating from the negotiation edge of the main loop.

Note that, in contrast to [Yu and Liu 2000], trust is not represented structurally as a soft goal but appears only as part of preconditions. (When mapping SR models into ConGolog for simulations, trust will be denoted by real-valued terms in logic corresponding, roughly, to subjective probabilities which will be updated during simulations.) Note also that, in contrast to Yu's formulation of SR models of different actors, the dependencies between the actors need not be stipulated, but are now derivable from the SR model. For example, the case where the seminar organizer is the depender, the speaker is the dependee and finding a speaker is the dependum, is now reflected in the fact that the precondition "speaker found," which belongs to the seminar organizer, can only be satisfied by the commitment of the respective speaker. We believe that the initial SD model is usually the result of a preliminary analysis of the dependencies between actors. Once we take a closer look at the agents themselves by designing their corresponding SR models, the dependencies will follow from them. We expect that, in many cases, new dependencies will be discovered in this process which were not considered in the initial SD model. The type of a new dependency and whether to classify it as strategic in the first place is so far left to the designer.

In Figure 7, we illustrate the situation where watchful confidence exists. As described earlier, confidence is different from interpersonal trust in that it expresses trust in the network as a whole being beneficial. The two problems of both narrowed competence and the delayed commitment are now eliminated. The organizer and a pool of speakers are actors in a network. Since the organizer has trust in this network, he can commit to the client immediately. On the other hand, it may also be the case that the seminar organizer distrusts some speakers to a certain extent. As explained earlier, there is no contradiction in trustful actions and hidden distrust at the same time. In our view, distrust is reflected by a certain amount of overhead due to the monitoring of other actors. In Figure 7, a monitor is invoked by the seminar organizer once a speaker he distrusts commits to holding the seminar. The monitor will watch for critical deadlines like the time when the seminar organizer expects the speaker to deliver transparencies to be included in the seminar folder. Should a deadline be missed (or even before that, depending on the level of distrust), the monitor would alert the seminar organizer or send a reminder to the speaker. Since the monitor keeps an eye on the speaker, this also seems to be the natural place where eventually an update of the trust/distrust values occurs, after the overall task of organizing and holding the seminar has ended. Besides its function as a "watchdog," the monitor is

[4] In Yu's original formalism, postconditions are already captured to some extent in that tasks can achieve or break goals.

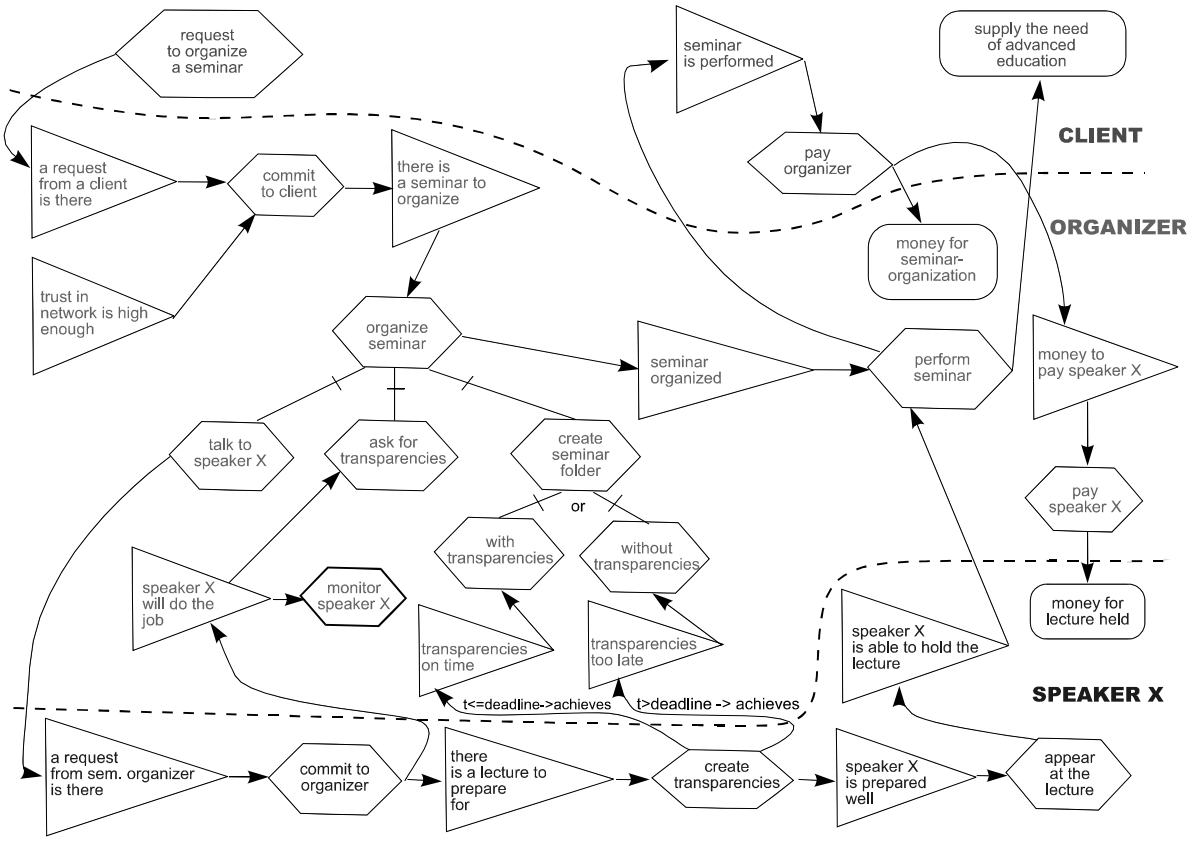


Figure 7: The organizer trusts the network in a watchful way

also in charge of bookkeeping, that is, it keeps a log of the interaction with the speaker and the eventual outcome of the overall task. Having a log of past interactions with another actor seems to play an important role when assessing the trustworthiness of the actor or, when aggregated over many actors, confidence in the whole network. In our diagram, the details of the monitor are left out for readability.

Figure 7 contains two more features not present in Yu's original proposal: For one, we explicitly refer to a simple (linear) model of time in the *conditional achieve* links. The other new feature concerns the *conditional execution* of tasks, exhibited in the example of creating a seminar folder. The idea is that if the speaker manages to deliver his transparencies on time, then these are included in the folder, otherwise they are not. Rather than introducing conditional tasks as a primitive, we have opted to define them instead in terms of a (non-deterministic) choice between tasks (denoted by "or" in the diagram). Together with the mutually exclusive preconditions attached to those subtasks, we obtain the desired effect.

3.4. Mapping SR models into ConGolog

ConGolog is a logic-based plan language suitable to model and simulate dynamic domains [Lesperance et al. 1999]. In contrast to most plan languages, ConGolog features concurrent actions and comes equipped with control structures from imperative programming languages such as conditionals, loops, and recursive procedures. Facilitated by the fact that our extensions of SR models are directly inspired by ConGolog, it turns out to be fairly straightforward to map SR models into ConGolog plans. For example, tasks are mapped into ConGolog procedure definitions. Task preconditions correspond to conditionals or interrupts. The latter is useful when a task is to be performed repeatedly (like organizing a seminar), triggered by its preconditions being satisfied (like a new request from a client). At the moment, the designer needs to decide when to use a precondition or an interrupt. In the future, we may make this distinction explicit already in the SR models. Time can be added very easily by introducing a discrete clock which is treated as a separate process run concurrently and with lowest priority with respect to all other processes. Monitors

such as the one mentioned in Figure 7, also correspond to ConGolog procedures. They run concurrently to the other agent processes, waiting for certain conditions like deadlines to arrive, at which time appropriate actions are initiated. Going into more detail of the mapping from SR models into ConGolog is outside the scope of this paper and the reader is referred to [Gans et al. 2001] instead.

4. SUMMARY AND CONCLUSION

In this paper, we have extended previous work on viewpoint resolution in requirements engineering for distributed organizations by novel dynamic mechanisms for organizational network settings for which the previously proposed methods turn out to be too static and situation-insensitive. The key sociological idea underlying our solution is a three-column success model for social networks, resting on the columns individual core Trust, network Confidence, and Distrust – hence the name TCD model.

The TCD model requires an extended modeling formalism which does not restrict itself to structural aspects of trust as Yu's i* formalism does, but explicitly models the interplay between structure, agent planning and action, and communicative actions to manage expectations. Our approach extends i* mainly by pre-/postcondition and time mechanisms that enable a componentized planning mechanism (supported by ConGolog) interacting with a Speech Act formalism whose variants express trust, confidence, and distrust in a manner compatible with the TCD predictions. We remark that, although we have chosen ConGolog for simulation and analysis, other methods may be employed as well. In fact, [Mylopoulos et al. 2000] combine i* with a temporal logic and propose to use a model-checking mechanism for analysis. We plan to compare the two methodologies in the future.

We are presently completing a first implementation of this integrated formalism, based on an embedding of ConGolog with the ConceptBase metadata manager. This implementation will then be used with examples from our ongoing case study in cross-Atlantic entrepreneurship networks to validate the TCD model itself and predictions resulting from it. Specific TCD-based lifecycle hypotheses as discussed in [Gans et al. 2001] can, after this initial validation and calibration, be used for problem analysis in ongoing requirements management efforts in networks, e.g. concerning the appropriateness of proposed network rules, the situation-dependent optimal mix of trust, confidence, and distrust, evaluation of specific strategic actions, and the like, as suggested by the nodes and links in the TCD model of Figure 1. In the longer run, we plan to integrate these mechanisms in the design an infrastructure for computer-supported cooperative work for distributed internet-based communities [Appelt et al. 2001] which is tuned to the specific needs of organization networks, as seen from the viewpoint of the TCD model.

Acknowledgments. This work is supported in part by the Deutsche Forschungsgemeinschaft in its Focused Research Program on Socionics. We would like to thank the anonymous referees, whose comments have helped us improve the presentation. Feedback by our project partners as well as numerous discussions in the emerging MIT-NRW entrepreneurship network are gratefully appreciated.

5. REFERENCES

Appelt, W., Jarke, M. and Prinz, W. (2001): *Emergence of a global cooperation network infrastructure*, submitted to a special issue on globalization in CSCW, Comm. ACM.

Axelrod, R (1984): *The Evolution of Cooperation*, New York: Basic Books.

Birk, A. (1999): *Learning to Trust*, in [Castelfranchi et al. 1999]

Boon, S.D. and Holmes, J.G. (1991): *The dynamics of interpersonal trust: resolving uncertainty in the face of risk*, Hinde, R.A. and Groebel, J. (eds.): *Cooperation and Prosocial Behaviour*, Cambridge University Press, 190-211.

Broersen, J., Dastani, M. and van der Torre, L. (2000): *Leveled Commitment and Trust Negotiation*, in [Falcone et al. 2000].

CACM (2000): *Special Issue on "Trusting Technologies"*, Comm. ACM 43, 12, Dec. 2000.

Castelfranchi, C. and Falcone, R. (1999): *Social Trust: A Cognitive Approach*, invited talk at *International Workshop on Agent-Oriented Information Systems*, CAiSE 1999, Heidelberg. http://www.aois.org/99/castelfranchi-Social-Trust-paper.doc

Castelfranchi, C., Tan, Y.H., Falcone, R. and Firozabadi, B.S. (eds.) (1999): *Proceedings of the Autonomous Agents Workshop on Deception, Fraud and Trust in Agent Societies,* Seattle, Wa, 1999.

Coleman, J.S. (1990): *Foundations of Social Theory*, Harvard University Press, Cambridge/Mass.

De Giacomo, G., Lespérance, Y. and Levesque, H. (2000): *ConGolog, a concurrent programming language based on the situation calculus*. Artificial Intelligence 121, 1-2, 109-169.

Demolombe, R. (1998): *To trust information sources: a proposal for a modal logical framework* in Proceedings of the Autonomous Agents Workshop on Deception, Fraud and Trust in Agent

Societies, Roma, Italy, 1998, 9-19.

Falcone, R., Singh, M. and Tan, Y.H. (eds.) (2000): *Proceedings of the Autonomous Agents Workshop on Deception, Fraud and Trust in Agent Societies*, Barcelona, Spain, June, 2000.

Gambetta, D. (1988): *Mafia: the price of distrust*, in Ders (ed.): *Trust. Making and Breaking Cooperative Relations*, New York, 158-175.

Gambetta, D. (1990): *Trust. Making and Breaking Cooperate Relations*, Oxford: Basil Blackwell.

Gans, G., Jarke, M., Kethers, S., Lakemeyer, G., Ellrich, L., Funken, Ch., and Meister, M. (2001): *Towards (Dis)Trust-Based Simulations of Agent Networks*, to appear in: Proceedings of the 4th International Workshop on Deception, Fraud,and Trust in Agent Societies, Montreal.

Giddens, A. (1990): *The Consequences of Modernity*, Oxford.

Jarke, M., Gallersdörfer, R., Jeusfeld, M.A., Staudt, M. and Eherer, S. (1995*): ConceptBase: A deductive object base for metadata management*, International Journal of Intelligent Information Systems 4, 2, 67-95.

Kern, H. (1998*): Lack of Trust, Surfeit of Trust: Some Causes of the Innovation Crises in German Industry*, in Lane, C. and Bachmann, R. (eds.): *Trust Within and Between Organizations*, Oxford, 203-213.

Kethers, S. (2000): *Multi-Perspective Modeling of Cooperation Processes*. Doctoral Thesis, RWTH Aachen, Germany.

Kramer, R.M. and Tyler, T.R. (eds.) (1996): *Trust in Organizations: Frontiers of Theory and Research*, Thousand Oaks, Sage

Lespérance, Y., Kelley T.G., Mylopoulos, J. and Yu, E. (1999): *Modeling dynamic domains with ConGolog*, in Jarke, M. and Oberweis, A. (eds.): *Advances in Information Systems Engineering.* Proceedings CAiSE 1999. LNCS 1626, Springer Verlag, 365-380.

Lewicki, R.J., McAllister, D.J. and Bies, R.J. (1998): *Trust and Distrust: New Relationsships and Realities*. Academy of Management Review 23, 3: 438-458.

Loose, A. and Sydow, J. (1997): *Vertrauen und Ökonomie in Netzwerkbezeihungen – Strukturationstheoretische Betrachtungen*, in: [Sydow et al. 2000], 161-192.

Luhmann, N. (1988): *Familiarity, confidence, trust: problems and alternatives*, in: Gambetta, D. (ed.): *Trust. Making and Breaking Cooperative Relations*, Oxford, 94-107.

Malone, T.W., Yates, J. and Benjamin, R.B. (1987): *Electronic Markets and Electronic Hierarchies*. Communications of the ACM 30 (6), 484-497.

Marsh, S. (1994): *Trust in Distributed Artificial Intelligence*, in: Castelfranchi and Werner (eds.): *Artificial Social Systems*, Berlin/New York, 94-111.

Mayer, R.C., Davis, J.H. and Schoorman, F,.D. (1995): *An integrative model of organizational trust*, Academy of Management Review 20, 3, 709-734.

Medina-Mora, R., Winograd, T., Flores, R. and Flores, C.F. (1992): *The Action Workflow Approach to Workflow Management Technology*, Proceedings 4th ACM Conf. *Computer-Supported Cooperative Work*, Toronto, 281-288.

Meyer, J.J. Ch. and van der Hoek, W. (1992): *A modal logic for nonmonotonic reasoning*, in van der Hoek, W., Meyer, J.J. Ch., Tan, Y. H. and Witteveen, C., (eds.): *Non-Monotonic Reasoning and Partial Semantics*, 37-77. Ellis Horwood, Chichester.

Mylopoulos, J., Borgida, M., Jarke, M. and Koubarakis, M. (1990*): Telos: a language for managing knowledge about information systems*, ACM Trans*. Information Systems 8*, 4, 327-362.

Mylopoulos, J. , Fuxman, A. and Paolo G. (2000): *From Entities and Relationships to Social Actors and Dependencies*, in Proceedings of the *19th International Conference on Conceptual Modeling*, ER 2000, Salt Lake City, USA.

Nissen, H.W., Jeusfeld, M.A., Jarke, M. and Huber, H. (1996*): Managing multiple requirements perspectives with meta models*, IEEE Software, March 1995.

Nuseibeh, B., Finkelstein, A. and Kramer, J. (1996): *Method engineering for multi-perspective software development*, Information and Software Technology 38, 4, 267-274.

Powell, W. (1990): *Neither market nor hierarchy: Network forms of organization*, in Staw, B. and Cummings, L.L. (eds.): *Research in Organizational Behavior*, Greenwich, 295-336.

Schäl, T. (1996): *Workflow Management Systems for Process Organizations*. LNCS 1096, Berlin: Springer.

Scheidt, B. (1995): *Die Einbindung junger Technologieunternehmen in Unternehmens- und Politiknetzwerke. Eine theoretische, empirische und strukturpolitische Analyse*, Berlin.

Sichman, J., Demazeau, Y., Conte, R. and Castelfranchi, C. (1994): *A Social Reasoning Mechanism Based On Dependence Networks*,

Proceedings 11th European Conference on Artificial Intelligence, 188-192.

Staber, U. (2000): *Steuerung von Unternehmensnetzwerken: Organisationstheoretische Perspektiven und soziale Mechanismen*, in Sydow, J., Windeler, A. (eds.): *Steuerung von Netzwerken.* Opladen: Westdeutscher Verlag, 58-87.

Sydow, J. (1992): *Strategische Netzwerke. Evolution und Organisation*, Wiesbaden: Gabler.

Sydow, J., Windeler, A., (eds.) (2000): *Steuerung von Netzwerken*, Opladen/Wiesbaden: Westdeutscher Verlag.

Tan, Y.H. and Thoen W.A.W. (2000): *Formal aspects of a generic model of trust for electronic commerce*, in [Falcone et al. 2000].

van Linder, B. (1996): *Modal Logics for Rational Agents*, Ph.D. Thesis, Department of Computing Science, University of Utrecht.

Weyer, J., (ed.) (2000): *Soziale Netzwerke. Konzepte und Methoden der sozialwissen-schaftlichen Netzwerkforschung*, München.

Weyer, J. (2000a): *Zum Stand der Netzwerkforschung in den Sozialwissenschaften*, in: [Weyer 2000], 1-34.

Williamson, O.E. (1975): *Markets and Hierarchies: Analysis and Antitrust Implications*, New York 1975.

Winograd, T. and Flores, F. (1986): *Understanding Computers and Cognition: A New Foundation for Design*, Norwood, NJ: Ablex.

Witkowski, M., Artikis A. and Pitt J. (2000): *Trust and Cooperation in a Trading Society of Objective-Trust Based Agents*, in [Falcone et al. 2000].

Yu, E., Liu, L. (2000): *Modelling Trust in the i* Strategic Actors Framework,* in [Falcone et al. 2000].

Yu, E. (1995): *Modeling Strategic Relationships for Process Reengineering*. Ph.D. Thesis, University of Toronto, Canada. Also published as Technical Report No. DKBS-TR-94-6.

Zucker, L. (1986): *Production of Trust: Institutional Sources of Economic Structure*, Research in *Organizational Behaviour 8*, 53-111.