In: Proceedings of the 3rd European Conference on Cognitive Science, ECCS99, Sienna, Italy, 1999, pages 57-62.

On being responsible: how to be individualistic and smile benevolence to the others

Nuno David
ISCTE/DCTI
Department of Information Sciences and Technologies
Av. das Forças Armadas
1600 Lisbon, Portugal
+351 1 7903911
Nuno.David@iscte.pt
http://www.iscte.pt/~nmcd

Jaime Simão Sichman
Dept. of Computer Engineering
University of São Paulo
Av. Prof. Luciano Gualberto, 158, tv. 3
05508-900 São Paulo, Brazil
+55 11 8185397
jaime@pcs.usp.br
http://www.pcs.usp.br/~jaime

Helder Coelho
Dept. of Computer Science
University of Lisbon
Bloco C5, Piso 1
Campo Grande
1700 Lisbon, Portugal
+351 1 7500122
hcoelho@di.fc.ul.pt
http://www.di.fc.ul.pt/areas/ai

ABSTRACT
In this paper, we discuss the balancing problem relative to the agents' individual and global social achievements in Multi-Agent Systems. Our claim is that any social principle prescribed within a level of abstraction above Newell's Knowledge Level, in order to balance the needs of the individual and the system as a whole, can not adopt an utility-oriented scale of performance but instead must adopt a *substantialist*-oriented (goal-driven) concept of performance. On this perspective, we propose a number of goal-generating principles for social responsible agents, towards an uncompromising individualistic view of autonomous agents that nevertheless attempt to motivate the welfare of collectives.

Keywords
Multi-Agent Systems, benevolence/non-benevolence, individual and social performance, social rationality

BENEVOLENCE DILEMMAS, *DO NOT SMILE*

In *Multi-Agent Systems* (MAS) agents are desirably autonomous, although invariably facing the matter of fact of limited self-sufficiency. Autonomy expresses the agents' ability to act without direct intervention or guidance of others (e.g. humans), primarily according to their adopted goals, relevant competence, and specific problem solving abilities. In addition, while not compromising their autonomy, inasmuch as agents may use each others' capabilities and resources to achieve their ends, agent are also and necessarily interdependent [Castelfranchi, 1990].

Notwithstanding, autonomy and lack of self-sufficiency may even entitle agents, deliberately and selectively, not to interact with others (e.g. co-operate) exclusively or inasmuch as they do not expect to benefit from such pro-activity or external request. In this sense, the social interaction model is a *bottom-up* one: social interactions and organisations are produced as a result of the agents' efforts to achieve their own goals [Conte and Sichman, 1995]. Interactions like co-operation are not assumed *a priori* but the principle of *non-benevolence* is presumed: agents are not prescribed to help each other but decide autonomously whether or not to co-operate with others. As pointed out by Jennings and Campos [1997], in this *contructist* view of multiple agent systems, the measure of *performance* of the system is fundamentally focused on the *individual performance* of the agent, nearly disregarding the achievements of other agents or the overall agency

In effect, if the principle of *non-benevolence* is assumed, we may find two major trends for measuring the agents' *individual performance*.

The first one adopts a utility oriented scale, calculated according to the cost/benefit of the agents' actions against the worth of the corresponding goals, whatever goals these may be. The agents' decision functions are fundamentally concerned with choosing actions that maximise individual utility, according to the classic principle of economic rationality ([e.g. [Zotkin and Rosenchein, 1994; Sandholm and Lesser, 1995]).

On the other hand, an inclined socio-psychological perspective of *performance* will most probably value a *substantialist* [Conte and Pedone, 1998] view of rationality; that is, individual performance measured in terms of the agents' attained goals. Here, similar to Newell's principle of rationality [Newell, 1982], the real *motive* for being rational is focused on the agents' own goals. In this case, the agents' decision functions are essentially concerned with the choice of adequate partners (organisations) and relevant actions in order to better achieve their goals, according to different patterns of dependence and influencing power over the others [Castelfranchi *et al.*, 1992].

On the other extreme of the benevolence vector, the earlier *reductionist* view [Jennings and Campos, 1997]

adopts a *top-down* social interaction model, in which co-operation is taken for granted [Conte and Castelfranchi, 1995]. Here, social interactions are usually constrained by some pre-established organisational structure, where a design team has a particular problem to solve. The problem is then decomposed in a number of sub-problems, which are subsequently assigned to one or more agents (e.g. [Lesser and Corkill, 1983]). Within such an approach, agents are thus *benevolent*: their goals are prescribed and the corresponding actions are performed in order to favour the needs of the overall system rather than favouring the needs of the individual.

Unlike the *constructist* approach, the utility-oriented performance of the global system is all that matters, whilst the notion of performance as a *substantialist* measure of the system as a whole can hardly take a place here: the goals of the global system are prescribed in design time.

Unfortunately, these two approaches are not orthogonal. The needs of the individual agent do not always overlap with the needs of the overall system. Presently, the prevailing view of multi-agent modelling is that of accomplishing the agent's own goals or maximising the agent's own gain, without regard for the achievements of the other agents or the overall agency. Frequently, an inflexible non-benevolence view of MAS is adopted at the cost of ignoring the impact of one agent's goals, plans and actions on the goals and plans of other members of the society.

Some authors have readily identified this problem and suggested several different ways of tackling a more responsible view of distributed agents at work, within several different domains (e.g. transportation schedulling [Fischer *et al.*, 1995]). As far as we know, only Jennings and Campos [1997] have tried to approach this problem from an abstract point of view, by using a high-level framework in which the key abstraction mechanism is a new computer level called the Social Level (sitting above Newell's Knowledge Level).

With reference to the mentioned work, in our view there is nevertheless a fundamental incompatibility relative to (i) the level of abstraction above Newell' Knowledge Level in which the authors model the social aspects of multiple agent systems, and (ii) the level of abstraction in which the individual and collective performance of agents is captured. The balance between individual and collective performance of agents is regulated through the so-called *Principle of Social Rationality*, according to the expected equilibrium between individual and social utility of actions. Performance is therefore regulated by the specific processing details of the individual agent, ignoring the motivational or *substantialist* level of abstraction that moves agents into action. This seems to be a contradiction, mainly if such balancing mechanisms are modelled above Newell's Knowledge Level of abstraction.

The purpose of this paper is to analyse some limitations of such an approach and propose some preliminary thoughts and vectors of change, towards an uncompromising individualistic view of autonomous agents which, nevertheless, attempt to motivate the welfare of collectives.

Our claim is that all behavioural laws or social principles stated above Newell's Knowledge Level should not be utility-driven but instead must be necessarily *substantialist* or goal-driven.

In the next section we recall Newell's principle of rationality and discuss some possible shortcomings of Jennings and Campos utility principle of social rationality. Subsequently we propose a set of goal-generating principles that we believe to be able to prescribe in the individual mind a responsible behaviour within multiple agents systems. To be responsible, an agent should abide to the highest possible number of principles to the limits of his available resources. The main advantage of such an approach is that it does not compromise the scale of utility preferences concerning which norms, goals, plans and actions should an agent choose or offer to his partners within an heterogeneous social system.

DISTRIBUTE UTILITY, *ARE THEY SMILING?*

The balancing problem with architectures willing to preserve the individual agents' *autonomy* and therefore benefit from the agents' *interdependence* of skills and resources, facing nevertheless the attempt to provide benefit to the overall system has extensively been identified (e.g. [Gasser, 1990; Jennings and Campos, 1997].

On the basis of Jennings and Campos proposal regarding a new computer level immediately sitting above Newell's *Knowledge Level* (KL) - the so-called *Social Level* (SL) -, further work has suggested the notion of *Socially Rational Agents* [e.g. Hogg and Jennings, 1997]. According to the behavioural laws inherent to the Social Level, socially responsible agents should select an action while striking for a balance between the individual expected utility and the expected utility yielded in the overall agency. The authors present the *Principle of Social Rationality*, meaning that an agent is entitled to perform an action whose joint benefit is greater than its joint loss:

"*If a member of a responsible society can perform an action whose joint benefit is greater than its joint loss, then it may select that action.*".

Joint benefit represents a combined measure incorporating the benefit provided to the individual and the benefit afforded to the overall system. The inherent standard of rationality builds upon the classical economic assumption of rationality - maximising the agent's expected utility *vs*. maximising the society expected utility.

We agree with Jennings and Campos hypothesis relative to the existence of a Social Level (SL)

immediately above the Knowledge Level, being "concerned with the inherently social aspects of multiple agent systems". Following Newell's [1982] intuition, with respect to the advantages given by a level of abstraction that characterise an agent's intelligence and knowledge independently of the specific operational processing and implementation details, the SL may be an adequate candidate to accommodate an abstract characterisation of a MAS behaviour which are social in nature, such as co-operation, competition, etc. It would then be possible to provide various descriptions enabling the overall social system behaviour to be analysed, without "having to delve into the implementation details of the individual agents" [Jennings and Campos, 1997].

With regard to the *Principle of Social Rationality*, we believe nevertheless that what Newell's *Principle of Rationality* confer in simplicity and abstraction to the agent's architecture, the referred social principle looses it in substance and level of detail to the social system. More precisely, and before going into the details of our criticism, let us recall Newell's principle:

"if an agent has knowledge that one of its actions will lead to one of its goals, then the agent will select that action".

One may notice that, unlike Newell's principle, which is inherently *substantialist* ("*if … lead to one of its goals…*"), Jennings and Campos principle is solely and inherently utility oriented (*"if…can perform an action whose joint benefit is greater than its joint loss…"*). It consigns to the agent and his social partners an absent social and motivational stance. In more detail, we may notice a contradiction between:

- the wish to free the social analysis and representation level from the implementation and specific processing details of the individual agent (with the Social Level immediately above the KL), and;
- the utility character of Jennings and Campos' social principle, clearly related to the operational model of the specific social processing (or optimisation) details being employed by the individual agent.

Calling Conte and Pedone [1998] individualistic intuition of rationality: "At most, utility maximisation is a modality of their action, regulating problem solving, planning and decision making, often in interaction with other principles. But action, intention, formation, planning, problem solving, etc. are activated to achieve goals…" - we may recall again Newell's principle: *"if … lead to one of its goals…"*) - "… possibly ***while*** **maximising utility**" (our bold and italics).

Newell's principle is goal governed. It entitles an agent to select an action if there is knowledge that it will *lead to one of its goals*. Its weakness is precisely its strongest value. It does not constraint in any operational matter the agents' processing and decision making. In other words, it is rational but not *statically* normative constraining in terms of the agents' built-in ends.

By analysing Jennings and Campos social principle, we may point out the following shortcomings:

1. **The concept of performance is strictly utility-oriented leaving aside the real motivation of the agent, i.e., its substantive goals**. Performance is not necessarily a utility dimension but also goal fulfilment. Even in the earlier *reductionist* view, where the designer prescribes the goals of the whole system in design time, if the system does not complete its goals then it is fair to say that the system does not perform well.
2. **It is inseparable from possible computational processing power constraints**. Utility maximisation is recognised as computational expensive approach. Newell's abstract principle drives away rationality from the problems of bounded computational power (it may be instantiated with rational agents, as well as agents with bounded rationality [Simon, 1969]). Hence, what is gained in *autonomy* with the motivational perspective, may be lost with Jennings and Campos principle within the *interdependence* matter of fact due to lack of computational power. This is somehow a contradiction. Like many other authors we believe: the fact that agents have interdependent skills does not diminish their autonomy but, on the contrary, the existence of other agents enhances the autonomy and power of individual agents [Sichman, 1995]. One may say limited computational power is inevitable. We agree but let us face it in the right level of abstraction.
3. **It builds upon a static normative action-oriented view of society**. This view may not be so limited if our goal is to build real-world and useful agent based systems (it may be even beneficial). As pointed out in Conte and Castelfranchi [1995] it still has to be tested if norm-governed systems perform better than norm-abiding systems. Nevertheless, to raise the level of abstraction in which the parameters and principles of a responsible agent are statically specified above the KL is only to hide the problem, if a normative motivational view is not accounted nor understood a those higher levels. For instance, according to Jennings and Campos principle, one can not simulate norm-governed and a significant number of norm-abiding systems, since one is not allowed to change an implicit, hardwired, *goal* of maximising utility. Is an agent entitled to break a social principle according to other competing goals if such principle prescribes actions but not goals or beliefs? In our view the answer is no. How could we represent it otherwise above Newell's Knowledge Level? For instance, how could we explain (represent) large economic investments with predicted short- and mid-term utility losses for the sake of other very long-term goals with scattered (possibly negative) economic revenues? Some large

and multi-national companies are willing to spent large quantities of resources in the media (for instance, buying or creating a newspaper company) solely to obtain prestige (as may happen with individuals as well). On the other hand, a motivational principle of social rationality may: (i) treat norms as any other goal; or, (ii) treat norms as prescriptive normative goals[1]. A responsible situated agent shall then abide to the highest number of social normative principles as the limits of his bounded resources allow.

4. **It constraints the formation of organisations**. A variable individualistic view of utility may be more or less distributive but, alone, it does not account for the understanding of organisations. It may be a key factor only if maximising joint utility is understood as an *explicit* and competing goal. For instance, if the vector of benevolence calculated on the grounds of joint utility is the single value for the formation of organisations, how is an agent's benevolence "acquired from the outside"? The theory of dependence and social power [Castelfranchi, 1990; Castelfranchi *et al.*, 1992] has extensively identified this problem. One may have notice, that Newell's abstract asocial principle of rationality does not prescribe but is compatible with any lower abstract procedure to maximise internal utility. Similarly, it would be useful to have a social principle that could possibly be instantiated with the *goal* of maximising joint utility. If active, such social goal could call at any moment a lower abstract procedure in order to balance or maximise individual and social utility (whatever decision theory would be used).

If due to an altruistic or instrumentalist motivation a subject agent decides to adopt another agent's goal, can we assume the third-party agent's ends to be maximising utility? Furthermore, can we guarantee that both agents' operational processes for utility maximisation will lead to a same set of actions?

If agents are heterogeneous our answer is no. Our main thesis is the following: a behavioural law prescribed immediately above Newell's Knowledge Level shall not be utility-driven but instead must be necessarily *substantialist* or goal-driven.

DISTRIBUTE GOALS, *CAN THEY SMILE?*

Newell's Theory of Knowledge [Newell, 1982] allows a common abstraction for the agents' knowledge over different physical structures and symbols. Any symbolic intelligent system, ranging from expert systems to pure experimental and scientific research systems, can today be implemented on the grounds of such a theory. The representation schemes are in the Symbolic Level, while the information and processing structures explore a body of knowledge in the Knowledge Level. If shifting to a social apparatus, goals and beliefs in the Knowledge Level should in turn explore social objects and interactions in the Social Level. Conversely, new goals and beliefs may be acquired in the Knowledge Level owing precisely to the Social Level.

On this basis, any principle of social rationality must necessarily capture the motives for being social, in our view, dependence based. They ought to be general enough to build any kind of social system and therefore, like Newell's asocial principle abstract, simple and goal governed.

A *substantialist* vector of approach is intuitive and practical. Imagine an agent immersed on a large society of agents, as we expect the Internet to be in some near future: it is easier to choose an action which we believe will benefit the achievement of a large set of the other agents' goals, against choosing an action which we believe will maximise the other agents' utility. Utility maximisation is recognised as computational expensive approach. Likewise, manipulation of goal adoption and delegation structures are known to be computationally complex as well [d'Inverno, Luck and Wooldridge, 1997], but represent the key aspects of social reasoning [Castelfranchi and Falconi, 1999]. Moreover, goals may be easily stated (e.g collective goals, normative goals), perceived, easily readable, they are essentially desired qualitative states.

Sociality is inherently complex. In our view, not a single but a set of non-mutually inconsistent principles must be established. This set of principles must be sufficiently general in order to capture the various motives for being social, while strong enough so as to produce useful and effective laws for entailing responsiveness. A responsible agent should make all the efforts to act according to the highest possible number of social principles. This may depend on resource bounds.

Our preliminary proposal is focused on the positive assumption that there are no conflicts between agents, that is, if interference between agents exists it will benefit the achievement of their goals. A further set of principles should be added to account for negative social interference between agents. In this paper, three sets of principles are defined, the *individualistic*, the *instrumental social* and *benevolence-end* principles, which progressively increase the deliberative positive impact of one (subject) agent's actions on the other (third-party) agents' goals.

The first principle ***P1*** sets out a weak version of Newell's principle of rationality: a rational agent is allowed to select an action if it has knowledge that it can possibly contribute to one of his goals. It is a mandatory rule since it does not only characterise the primary ingredient for rational behaviour but also since all social external appeals are as effective as they are able to take advantage of this principle. All other existing principles must thus revert to the welfare of society through this first basic principle.

[1] In Conte and Castelfranchi [1995] the reader may find a detailed definition and roles of normative goals.

Individualistic principle

P1 *if the subject agent has knowledge that one of its actions will lead to one of its goals, then he may select that action.*

Instrumental social principles

P2 *if the subject agent has knowledge that one of its goals makes him dependent on a third-party agent, then he may adopt a third-party agent's goal.*

P3 *if the subject agent has knowledge that, according to a third-party agent's knowledge, one of its own actions will lead to one of its own goals, then he may select that action.*

Benevolence end principles

P4 *if the subject agent has knowledge that one of its actions will lead to a known third-party agent's goal, then he may adopt that goal.*

P5 *it the subject agent has knowledge that a third-party agent wants him to select one of its actions, then he may select that action.*

The second set of principles assumes that an agent can not act exclusively according to his own needs. He can not achieve his goals without providing some benefit to the others. These principles apply for instance if two agents are bilaterally dependent for one same or two different goals, which can then lead to interactions such as co-operation or social exchange [Castelfranchi *et al.* 1992].

Principle ***P2*** allows an agent to adopt instrumentally another agent's goal to achieve a goal of his own. Consequently, the adoption of a third-party agent's goal through principle ***P2*** may in turn revert to the third-party according to principle ***P1***, since the subject agent is now entitled to select an action which he knows that will lead to his new adopted goal, i.e., the third-party agent's goal. However, it may be the case that the subject agent knows that, according and exclusively to the third-party agent's knowledge, one of its own actions will lead to the new adopted goal. This situation may happen for example if the subject agent ignores the outcome of the action or disagrees with the third-party agent's expected outcome. Principle ***P2*** can thus also revert to the third-party due to principle ***P3***, which explains interactions such as co-cooperation or social exchange with subjective *locally* believed unilateral dependencies [David *et al.*, 1999].

The last set of principles encompasses rules adopting benevolence as a "hardwired" end. Principle ***P4*** entitles the subject agent to perform one of its actions simply because he knows that will benefit a third-party agent's goal. Here, goal adoption is not yet a "hardwired" end but a limited dynamic mechanism. It may be used within a Contract Net Protocol framework [Smith, 1980]: the subject agent adopts a third-party agent's goal according to his own *local* believed competence. Conversely, the last principle ***P5*** may be relevant within a typical master-slave situation: the subject agent tries to execute a requested action without questioning its use or ends. Unlike the previous principle, goal adoption is completely "hardwired".

Now, how should a responsible agent proceed? By adopting to all of his resources all the given social principles. For example, an action that is commonly accepted by principles ***P1***, ***P3***, ***P4*** and ***P5*** is a social responsible action, with both a positive outcome for the individual and the society. An agent willing to adopt principles ***P1*** to ***P3*** is a rational agent yet a non-benevolent one. This same agent will be guaranteed to be sincere to a same third-party agent if he does not break principle ***P4***. A benevolent agent within a Contract Net Protocol would adopt principles ***P1*** and ***P4***. A pure slave would stick to principle ***P5***. Another responsible agent may prefer to abide to a different number of principles according to his available resources at different points in time: if the agent has available resources beyond the ones needed to achieve his own goals, then the agent may use instrumental and benevolent-end principles. Otherwise, he should stick exclusively to individualistic and instrumental social principles.

Note that this normative framework does not prescribe any operational level of choice or multi-partner negotiation. While we believe that it may prescribe an individualistic rational or responsible behaviour through "hardwired" goal-generating laws, it does not compromise the scale of utility preferences concerning which norms, goals, plans and actions should an agent choose or offer to his partners. These may be concerned with subjective choices, either on a lower social or individual processing level, which **eventually may** be given the end to maximise utility.

By using a set of goal-generating social principles, the MAS designer is able to tune in design time or dynamically to which principles an agent shall adhere. However, the agent's motives and preferences are still dynamic, susceptible to change on-line according to relevant different contexts.

DISCUSSION

We have discussed the shortcomings of social utility-oriented principles, which are prescribed above Newell's Knowledge Level.

Our arguments suggest that utility-oriented principles must be prescribed within lower levels of abstraction. Conversely, an eventual social principle prescribed within a level of abstraction above Newell's Knowledge Level, in order to balance the needs of the individual and the system as a whole, must adopt a *substantialist*-oriented (goal-driven) concept of individual performance, in accordance with Newell's principle of rationality.

On this perspective, these principles should be defined as *goal-generating* rules, from which we have identified three distinct categories: individualistic, instrumental social and benevolence-end principles.

All instrumental social and benevolence-end principles can and must revert to the welfare of society through the first principle of individual rationality. For instance, a responsible agent concerned with both his individual needs and the society as a whole should abide to the highest number of *goal-generating* principles to the limits of his available resources, possibly in different points in time.

Within this context, the MAS designer will be able to specify statically or dynamically to which social principles an agent shall obey, without compromising the scale of utility preferences, concerning which norms, goals, plans and actions should an agent internally choose or offer dynamically to his partners.

To be easily scaleable is one of the advantages of this kind of prescriptive architecture. The MAS designer may have the possibility to add other principles, without compromising the lower symbolic and knowledge level basic representations and processing details.

ACKNOWLEDGMENTS

Thanks to Mónica and João Pedro for their valuable contribution towards a good and ironic title.

REFERENCES

Castelfranchi C. (1990) ***Social power: a point missed in multi-agent, DAI and HCI***, Proceedings of the First European Workshop on Modelling Autonomous Agents in a Multi-Agent World (MAMAAW'90), pages 49-62.

Castelfranchi C., Miceli M. and Cesta A (1992) ***Dependence relations among autonomous agents***, Proceedings of the Third European Workshop on Modeling Autonomous Agents in a Multi-Agent World (Amsterdam) (MAMAAW'92), Elsevier Science Publishers B. V., pages 215-227.

Castelfranchi, C. and Falcone, R. (1999) ***Basic mental attitudes of a collaborating agent: Cognitive primitives for MAS***. Proceedings of the Ninth European Workshop on Modeling Autonomous Agents in a Multi-Agent World (Valencia) (MAMAAW'99), Springer Verlag, pages 188-209.

Conte R. and Pedone R. (1998) ***Finding the best partner: The PART-NET system.*** First International Workshop on Agent-Based Simulation (MABS98, Paris), Sichman J., Conte R. and Gilbert N. editors, pages 156-168.

Conte R. and Castelfranchi C. (1995) ***Cognitive and Social Action.*** UCL Press.

Conte R. and Sichman J. (1995) ***DEPNET: How to benefit from social dependence.*** Journal of Mathematical Sociology (20), pages 161-177.

David, N. Sichman, J. and Coelho, H. (1999) ***Extending social reasoning to cope with multiple partner coalitions***. Proceedings of the Ninth European Workshop on Modeling Autonomous Agents in a Multi-Agent World (Valencia) (MAMAAW'99), Springer Verlag, pages 175-187.

d'Inverno M., Luck, M. and Wooldridge, M. (1997) ***Cooperation structures***. Proceedings of IJCAI'97, pages 600-605.

Fischer K., Müller J., Pischel M. and Schier D. (1995) ***A model for cooperative transportation schedulling.*** Proceedings of the First International Conference on Multi-Agent Systems. AAAI Press, pages 109-136.

Gasser L. (1991) ***Social conceptions of knowledge and action: DAI foundations and open systems semantics.*** Artificial Intelligence (47), pages 107-138.

Hogg L. and Jennings N. (1997) ***Socially rational agents.*** Proceedings of the AAAI fall symposium on Socially Intelligent Agents, pages 61-63.

Jennings N. and Campos J. (1997) ***Towards a social level characterisation of socially responsible agents.*** IEE Proceedings of Software Engineering, 144 (1), pages 11-25.

Lesser, V. and Corkill D. (1983) ***The distributed vehicle monitoring testbed: a tool for investigating distributed problem solving networks.*** AI magazine, Fall, pages 15-33.

Newell A. (1982) ***The Knowledge Level.*** Artificial Intelligence, (18), pages 87-127.

Sandholm T. and Lesser V. (1995) ***Coalition formation among bounded rational agents.*** Proceedings of IJCAI, pages 662-669.

Sichman J. (1995) ***Du raisonnement social chez les agents: Une approche fondée sur la Théorie de la Dependence.*** Thèse de Doctorat, Institut National Polytechnique de Grenoble. France.

Simon, H. (1957). ***A behavioural model of rational choice***. Quartery journal of economics, (69), pages 99-118.

Smith, R. (1980) ***The contract net protocol: High-level communications and control is a distributed problem solver.*** IEEE Transactions on computers, vol.29 (12), pages 1104-1113.

Zlotkin G. and Rosenschein J. (1994) ***Coalition, cryptograpphy and stability: Mechanisms for coalition formation in task oriented domains.*** Proceedings of AAAI (Seattle, WA), pages 432-437.