

Founding Agent's 'Autonomy' On Dependence Theory

Cristiano Castelfranchi¹

Abstract.¹ First I shortly recapitulate the claim that autonomy is a relational notion in relation with agenthood and delegation; second, I lead back autonomy to the theory of dependence. Then I deal with the main claim of this paper arguing how *the dimensions of the agent's autonomy derives from its architecture and from the theory of action.*

1. THE PROBLEM OF DEFINING AUTONOMY AND ITS DIMENSIONS

Several people are trying to define several possible 'dimensions' of autonomy [1] [16] [19]. However, these useful attempts proceed in a quite empirical and unprincipled way. In this paper I claim that *the dimensions of the agent's autonomy derives from its architecture and from the theory of action.* More precisely:

each and any component of the architecture or necessary condition for a successful action can define a *dimension/parameter* of autonomy, since it can define an abstract 'resource' or 'power' necessary for the goal achievement, i.e. it can characterise a *specific 'lack of power'* and than a possible dependence and social non-autonomy.

Why should this attempt be useful? Because 'autonomy' is the central issues characterising 'agents' in ALife, in robotics, in DAI, in software-agents [28], but a good unifying conceptual definition is not enough. In fact, people is searching for characterising different *degrees* and *kinds* of autonomy relative to different types of agents, collaborations, and situations. Consider for example the notion of "adjustable autonomy" [11] [12] [16] [19]; or the fact that by definition software agents are *semi-autonomous* since they should act "on behalf of" its user or owner, like "agents" in the economic sense (*agency*). So, degrees and dimensions are important. However, they should not be assessed only empirically and practically. They can be derived from the theory of *agenthood*, from agent' architecture, and from the theory of *agency* (delegation). By doing so, we can also improve our understanding of the true notions of autonomy, of agenthood, and of agency.

2. "AUTONOMY" AS A RELATIONAL NOTION

One is autonomous *as for* a given action or goal (and not for another), and *from* something or somebody [4].

Non-Social Autonomy: Autonomy from the environment

¹ CNR-Institute of Psychology- Divion of AI, Rome, Italy. email: castel@ip.rm.cnr.it. For an extended version of this work: "What Autonomy Eventually Is" -submitted to ATAL'99- see <http://www.ip.rm.cnr.it>

It seems that the relevant notion of autonomy for AI Agents is the "social autonomy": autonomy from other Agents or from the user. But this is not true. The Agents we need are *entities able to act in an environment, on the basis of the perception of this environment* (also during the execution of the action). Being an Agent requires a certain degree of autonomy from these stimuli and from the environment itself [9] [13] [14] [17] [18] [20] [24] [28]. The Agent's behaviour cannot be completely determined and predictable on the basis of the current input, like a billiard ball under mechanical forces. Autonomous means -in general- non hetero-directed; i.e. whose behaviour in not determined and driven from outside. Agents (at this level) are at least Goal-Oriented systems, not simply causal entities [4] [8] [23]. More precisely, their "autonomy from stimuli" is guaranteed

- by the fact that their behaviour is teleonomic: it tends to certain specific results due to internal constraints or representations, produced by design, evolution, or learning, or their previous psychological "history";
- by the fact that they do not simply receive an input but they (actively) "perceive" and interpret their environment and the effects of their actions;
- by the fact that they have "internal states" with their own exogenous and endogenous evolution principles, and their behaviour also depends on such internal states.

The first relevant feature of autonomy in Agents is *in their relation with their environment*. And this is important especially because of their "social autonomy" and of their acting *remotely* i.e. far from and without our monitoring and intervention (control), on the basis of their local and contingent information, reactively and adaptively. They should manage by themselves and in a goal-oriented way (i.e. possibly achieving their task) their relation with the environment (either physical or virtual).

Social Autonomy

Apparently, there are two very different notions of Autonomy at the social level (i.e. relative to other agents):

- one is autonomy as "independence", "self-sufficiency";
- the other is autonomy in delegation (in agency).

In **Autonomy as Independence** *an Agent is completely autonomous (relative to a given goal) when does not need the help or the resources of other Agents to achieve its goal.*

This notion of autonomy (it would be possible also to use the term "self-sufficiency" or "independence" as opposed to "dependence") is particularly relevant in M-A systems [3] [25] [26]. But also in interaction between user and Agents, or among independent software Agents it is very relevant. In fact if the Agent is supposed to act on our behalf, if we have to delegate it some task, it must be able to satisfy this task by itself (not necessarily alone, also involving other Agents). It must posses the necessary capability and resources: either practical (to execute the task) or cognitive - to solve the problem -, or social to find out some

cooperation. Thus, relative to the delegated task it should be in some measure non-dependent on the delegating Agent, and able to deal with its dependence relations with others (we will see this in detail).

In other terms we have to *trust* the Agent as for its competence and capability, which is one of the two fundamental part of trust (the other is reliability) [6].

Autonomy in Delegation means how much an Agent is "autonomous" when it is working for another agent: exploited by it, helping it, collaborating with it, obeying to it; i.e. autonomy in "agency" (we will analyse it in section 6).

What I will show here is how these two notions are actually deeply connected one to the other and reducible. The first perspective actually results to be the more general and powerful: *all forms of social autonomy should be defined in terms of different forms of social independence.*

3. COGNITIVE AGENT'S ARCHITECTURE AND POWERS

Let's start our argument by assuming as a good operational reference a BDI architecture - in particular PRS Architecture (Fig. 1)² - which is a minimal, sufficient model of an 'intentional system' or 'cognitive agent', i.e. *an goal-governed system whose actions are internally regulated by explicitly represented goals and whose goals, decisions, and plans are based on beliefs.* Both goals and beliefs are cognitive representations that can be internally generated, manipulated, and subject to inferences and reasoning. Since a cognitive agent may have more than one goal active in the same situation, it must have some form of choice/decision, based on some "reason" i.e. on some belief and evaluation.

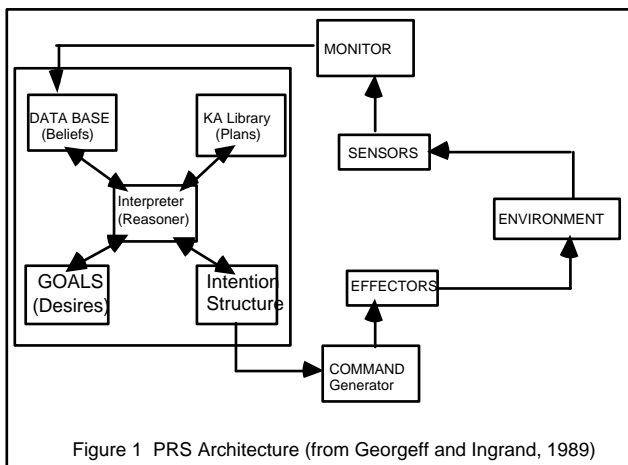


Figure 1 PRS Architecture (from Georgeff and Ingrand, 1989)

This model shows that in order to adaptively 'act' a cognitive agent Y needs:

- information from current environment (sensors), and the ability to interpret/understand this input relative to its knowledge and goals (in this perspective the 'monitor' or the data base modules are a bit reductive);

- a plan library; and the ability to retrieve from it an appropriate plan;
- a knowledge or beliefs base; and the ability to reason on it and to make relevant inferences;
- a set of motives/goals/tasks and the capability to activate them (either internally and pro-actively or by reacting to external stimuli) or to drop them;
- the capability to choose among alternative plans, and among incompatible goals;
- the ability to select an appropriate intention (a preferred, realistic goal endowed with its plan) and to persist in pursuing it (until appropriate);
- the capacity to generate commands to some effectors on the world; and to materially execute this command controlling these effectors;
- the capacity of monitoring the action results during and after its execution and to use this as a feedback on the process, at several levels.

We should add to this components and abilities, the 'external' resources and opportunities that allow or prevent the agent from successfully executing its action. Let distinguish at least between

- material or practical resources and conditions (including agents to be delegated), and
- deontic resources or conditions (permission, authorisation, role, etc.).

In our model this define the set of (internal and external)'powers' [2] that make Y able/in condition to act purposively and autonomously. If Y has these 'powers' it 'can' and 'may'. If some of these internal or external 'powers' or 'resources' is lacking, the agent is not able to (cannot/may not) pursue/achieve its goal. Given this, we have *all the basic dimensions of Y autonomy or dependence* (which is the complement of autonomy [3] [22] [25] [26]).

4. FROM POWERS TO DEPENDENCE

We already defined the relation of Dependence [2] [3] [8] [25]. This was mainly denoting practical or behavioural or external dependence for an action or a material resource of X: Y is internally able to want, decide, plan, intend, but it lacks some skills or some resource for behaving. Now, we have to abstract and generalise this notion to the 'cognitive' dependence: the lack of cognitive skills or resources (or permissions). By abstracting the notion of resource and of power (like in the original theory [21]) we can generalise the Dependence notion to cover the broad spectrum of dependence relations. Since we already defined Autonomy as the complement of Dependence, we will generalise at the same time the notion of Autonomy.

Let's call *p* any needed action (mental or practical) or resource (mental or practical) or condition (mental or practical):

If an agent Y depends on an agent X for its internal or external power/resource p relative to its goal G

(Dep Y X p G)

then Y is not autonomous from X

relative to its goal G and resource p

(Autonomous Y X p G) ==> (Not (Dep Y X p G))³

² We might use any other BDI model [15] or any other agent architecture - for example Chaib-draa's more "psychological" model [7]- provided that it is minimally complex and articulated.

³ The reverse is not true. In fact, Social Dependence is a sub-set of Non-Autonomy. Y can be non-autonomous (not able, needing something) without being socially dependent, since social dependence presupposes that

If, in order to successfully act, I have to receive either information, or a plan, or a goal, or a decision, or a resource, or some control, or an action I'm not able to perform, etc., I'm not completely autonomous. As many necessary powers as many potential limits to autonomy. As many *lacks of power*, as many potential dependence relations.

4.1 Generalising previous definitions

If we use Sichman's formulation of the Dependency theory [25] [26] (where Dep is relativised to available plans), we find the following notions of autonomy and the complementary notions of dependence.

An agent ag_i will be **a-autonomous (action autonomous)** for a given goal g_k , according to a set of plans P_{qj} if there is a plan that achieves this goal in this set and every action appearing in this plan belongs to ag_i 's action repertoire $A(ag_i)$:

$$a_{aut}(ag, g_k, P_{qk}) \stackrel{def}{=} \exists g_k \in G(ag) \exists p_k \in P_{qk} \forall a_{m, p_k}^{inst} \in A_{p_k}^{inst} a_m \in A(ag) \quad (1)$$

Analogously an agent ag_i will be **r-autonomous (resource autonomous)** if there is a plan that achieve this goal in this set and every resource appearing in this plan belongs to ag_i 's resources $R(ag_i)$.

Finally, an agent ag_i will be **s-autonomous (social autonomous)** if he is both a-autonomous and r-autonomous.

$$s_{aut}(ag, g_k, P_{qk}) \stackrel{def}{=} a_{aut}(ag, g_k, P_{qk}) \wedge r_{aut}(ag, g_k, P_{qk}) \quad (2)$$

Correspondingly, an agent ag_i **action-depends (a-dep)** on another agent ag_j for a given goal g_k according to a set of plans P , if g_k is in his set of goals; he is not a-autonomous for g_k ; there is a plan in P that achieves g_k and at least one action used in this plan is in ag_j 's set of actions. Analogously ag_i **resource-depends (r-dep)** if he is not resource-autonomous for g_k and there is a plan in P that achieves g_k and at least one resource used in this plan is in ag_j 's set of resources. Finally, an agent ag_i **social-depends (s-dep)** on another agent ag_j if it is either action-or resource-dependent on ag_j :

$$s_{dep}(ag, ag_j, g_k, P_{qk}) \stackrel{def}{=} a_{dep}(ag, ag_j, g_k, P_{qk}) \vee r_{dep}(ag, ag_j, g_k, P_{qk}) \quad (3)$$

We have simply to generalise the notion of '*what is needed*' (either resource or action, or whatever). As we saw, an agent can be dependent, non-autonomous from another agent, not only for a practical action he is not able to perform or for a practical resource, but for several material, mental, or deontic conditions. Let's introduce in our ontology the set or class of abstract resources or powers P needed for a given plan; in this way **a_{aut}** and **r_{aut}**, **a_{dep}** and **r_{dep}** are simply sub-cases respectively of **P_{aut}** and **P_{dep}**. An agent can be social autonomous or social dependent relative to whatever p he needs and the other can/has.

there is at least another agent X in the same word able to do the needed action [3].

⁴ Where a^{inst} is each instantiated action and A^{inst} is a (possibly empty) sequence of instantiated actions used in a given plan.

$$s_{aut}(ag_j, g_k, P_{qk}) \stackrel{def}{=} \exists \pi \in \Pi \pi_{aut}(ag_j, g_k, P_{qk}) \quad (4)$$

$$s_{dep}(ag, ag_j, g_k, P_{qk}) \stackrel{def}{=} \exists \pi \in \Pi \pi_{dep}(ag, ag_j, g_k, P_{qk}) \quad (5)$$

4.2 Resource or Action Dependence?

To be true, whatever I lack or need (an abstract 'resource' or 'power') when I'm *dependent* on another agent (i.e. somebody else can produce/provide what I need) *necessarily I depend on some action of this agent*. In fact, to *provide* (me) something requires some action. Any social dependence (dependence on another agent) is reducible to a dependence for an action of it [3]. I need (also) an action of it. This is also one criticism of d'Inverno [11] to Sichmans' formulation of Dep theory, and it is basically correct. However, this does not make redundant and superfluous to specify the kind of p, the specific lack of power that makes necessary X's help, i.e. the source of the Dep. One thing is the fact that for my own reasons I have to do some practical action I'm not skilled for; a different thing is that I need training or permission or passing a resource by X. I always need an action of X, but the action I need depends on my specific lacking power.

Moreover, we have to understand what level of abstraction we need about the notion of 'action' when claiming that any Dep is an action-dep. There is not only the practical case of a material resource that must be dislocated from X to me ("to give" "to pass"). The general notion is that of either

- to make it accessible to me; or
- to let /permit me to access it.

It depends on the situation of interference.

- a) I cannot practically access the resource, an action of X is needed to change the world in order I can access the resource; I need X's *active help* (he has to move away, to let r free, or to pass r to me, etc.)
- b) I can practically access the resource but X can prevent me from accessing it. What I need is a *passive help* (or goal-adoption), a *passive-action* by X (doing nothing), his omitting to do, his decision of not preventing me, of letting me to do [5]. I do not consider here the deontic (authorisation, prohibition) aspects (see later).

In conclusion, when we will say that agent Y depends on agent X as for information, or control, or recipe, or decision we mean that Y depends for an *action* of x: providing him that information (or let him access it), providing the recipe, or the control, the decision (either informing him about the decision or materially take the decision within his "cognitive" machinery, see section 8). In other words, X must in any case "empower" Y, either in strict deontic sense, by authorisation and responsibility-power delegation, or by training and instruction (giving abilities and knowledge), or in a more broad sense providing Y with the necessary material resources and conditions [5].

5. DIMENSIONS OF AUTONOMY

Let eventually examine the different forms and dimensions of autonomy/dependence as predicted by this approach (compare it with section 3):

- **Information autonomy/dependence:** if in order to act the agent needs perceptual information from its environment, it does not

access this information directly (independently) while another agent X has this information and provides it to Y, then Y is not 'information autonomous' [22] and is dependent on X as for information. It is the case of a captain depending on a look-out man in order to know where the enemy is.

- **Interpretation autonomy/dependence:** Y is able (and permitted, see later) and relies on itself as for interpreting the data and the sensory information. For ex. nurses or patients are not able (or permitted) to interpret by themselves the clinical data.
- **Know-how autonomy/dependence:** Y knows recipes and plans for that goal. For example when delegated, Y does not need to ask X "how" to do *a*, what to do for G.
- **Plan-discretion autonomy/dependence:** Y is able (and entitled) to choose among different possible plans for G.
- **Planning autonomy/dependence:** Y is able (and entitled) to develop a plan and apply it to a given circumstance.⁵
- **Motivational autonomy/dependence:** Y has its own goals, it does not depends on X in order to have goals, to know what to pursue.
- **Goal-dynamics autonomy/dependence:** Y is able (and entitled) to suspend a given plan or drop a given intention, to change its preferences; to react to novelties.
- **Goal-discretion autonomy/dependence:** Y is able (and entitled) to choose among different active goals (of its own and/or assigned tasks); its can/may have preferences.
- **Reasoning autonomy/dependence:** Y is able (and entitled) to make its own inferences and reasoning and to rely on its own conclusions.
- **Monitoring autonomy/dependence:** Y does not depend on X for monitoring its action and checking if it is regular (as expected) and successful. The system Y is a 'self-governing' systems.⁶
- **Skill autonomy/dependence:** Y do not depends on X as for some action/skill necessary in the plan for G; its action repertoire is sufficient. (notice that skills include 'mobility')
- **Resource autonomy/dependence:** Y do not depends on X as for some material resource necessary in the plan for G; its own resources⁷ are sufficient (notice that resources include space and time to act)
- **Enable or Condition autonomy/dependence:** Y do not depends on X as for some necessary condition for performing its actions; X has not to enable, allow or permit (in practical sense - [5]) Y's action.

This list is not exhaustive. In fact, clearly enough, *the dimensions of dependence/autonomy just depend on the granularity of our model of mind and action.*

The greater the number of distinctions relative the mental processes and products, and to action conditions, the larger the number of possible *lacks of powers* (inabilities) and of possible "interventions" of X. In fact, notice that one can delegate or non-delegate very specific and subtle part of the cognitive-practical task. One can for example delegate to find some plan without deciding among or without creating a new plan; or one can delegate to evaluate between two alternative what is better, but

⁵ In (4) I called 'plan autonomy' the conjunction of the last three autonomies.

⁶ In Castelfr Falcon 'Open Delegation' combines several of these autonomies (know-how, discretion, reasoning, etc.). Also the Webster's Dictionary (II -1988) definition of Autonomy as "the quality or condition of being *self-governing*" covers and mixes up several of these features.

⁷ Of course, the notion of "its own resources" is everything but trivial!

without finding the alternatives and without choosing (just a simple advise) (see later). So, in this case we might for example split our Plan-discretion and Goal-discretion into an Evaluation-autonomy and a true Discretion or Choice autonomy. And so on.

5.1 Autonomy as Self-Motivation

The role of autonomous goal in autonomy deserves special attention because of the special role of goal in the definition of an interesting notion of 'agent' (sect. 2). As I said a *goal-autonomous Agent* is an Agent endowed with its own goals. I claimed [4] that an Agent is fully socially autonomous if:

- It has its own Goals: endogenous, not derived from other Agents' will.
- It adopts goals from outside, from other Agents; it is liable to influencing. To be motivation-autonomous does not means to be autarchic or a-social; the agent can accept goals from others.
- *It adopts other Agents Goals only if it sees the adoption as a way of enabling itself to achieve some of its own goals* (i.e. the Autonomous Agent is a Self-Interested Agent).⁸

In sum, an autonomous agent should not be *provided* by other agents of its goals [17]. "Its own goals" mean non adopted from others, endogenous motives and preferences⁹.

However, this is not "autonomy" but is "goal-autonomy" or -better- motivation-autonomy (self-interesting). In fact -as we are arguing- there is no a unique notion and a unique dimension of autonomy. Also a motivation-autonomous agent can be non autonomous under other respects, for example it can be dependent as for material resources, or abilities, or information and so on. Of course, it is true that autonomy in motivation is autonomy *par excellence*, is a fundamental and high level dimension of autonomy¹⁰.

6. PRACTICAL VS DEONTIC AUTONOMY

We might also put another kind of autonomy in the list after the Enable autonomy:

- **Entitlement autonomy/dependence:** Y do not depends on X as for some necessary deontic condition for performing its actions (permission, authorisation, role, entitlement, etc.); X has not to formally enable, allow or permit (in the institutional/normative sense - [5]) Y's action.

However, this would be in part misleading. In fact the dimension of entitlement -as we remarked step by step- is orthogonal to all the previous dimensions. I mean that for each of

⁸ For a more complete list of conditions see [Cas95].

Notice that this postulate does not necessarily coincide with a "selfish" view of the Agent. To be "self-interested" or "self-motivated" is not the same of being "selfish". The Agent's "own" goals, for the purpose of which he decides to adopt certain aims of someone else, may include "benevolence" (liking, friendship, affection, love, compassion, etc.) or impulsive (reactive) behaviours/goals of the altruistic type.

⁹ I do not agree with Luck and d'Inverno' definition of motivation or desires as something "qualitatively different from goals"; to me they are just kinds of goals. Goals can be either very specific or quite vague and generic, chosen and pursued or put aside, related to action or just waiting for satisfaction, etc. We are just speaking of top-level or terminal goals (Conte and Castelfranchi, 1995) Vs instrumental goals.

¹⁰ Very important is also 'norm-autonomy' (Castelfr, Conte, Dignum, 1998) HARKO

the previous *lacks of power* we can wonder if it is due to material, practical impossibility or to deontic impossibility (lack of permission/entitlement; prohibition, non delegation, etc.). Y can practically lack or not be able to access a given resource or it can be not permitted/ prohibited to use it. It can be unable to build a plan or to decide among several alternative plans, or it can be able but not delegated and entitled to do so. And so on. Thus, let us say that there are two kinds of lacks of powers (and then of dependence and of autonomy): one based on practical conditions the other based on deontic conditions.

Deontic Autonomy: Y is permitted to do/ to decide/ to interpret/ to infer/ etc. Not only it is practically able and in condition to, but it can do this *without violating* a social or legal norm, or the user/designer prescriptions.

As there are two kinds of autonomy there are two kinds of "empowerment" (giving autonomy): deontic empowerment Vs practical, material empowerment¹¹.

7. A SYNTHESIS: AGENTS IN THE LOOP OF OTHER AGENTS

Let me recapitulate all we have said in other terms, leading it back to the initial definition of agenthood.

Given the abstract control loop of any control system which purposely [23] relates information, knowledge, goals, situation, and action, or given the general sensing-acting-environment loop of any interacting agent [13] (including a simple reactive agent),

if another agent interferes in, is inserted in this control loop, if the flow of the information/control or the flow of causal effects on the world is interrupted and pass through another agent and needs it,

the system is no more completely independent or autonomous: it depends on X's intervention.

This in fact corresponds to the well known view of Autonomy as acting "without the *intervention* of human beings or other agents" (Wooldridge) [29], where 'intervention' is simply a different label/view about that 'interference' or 'insertion'. Depending on which part of the goal-governed or goal-oriented (even a simple reactive) action loop X is inserted in, Y's dependence/autonomy is behavioural or cognitive (or more specifically resource based, skill-based, decision-based, etc.). Y depends on and is not autonomous from X as for *p*, i.e. for that part of the flow.

In conclusion -as I argued- the dimensions of the agent's autonomy derives from its architecture and from the theory of action; *any needed resource and power within the action-perception loop of an agent define a possible dimension of dependence or of autonomy.*

REFERENCES

- [1] AAAI Spring Symposium on Agents With Adjustable Autonomy, March 22-24, 1999, Stanford University.
- [2] Castelfranchi, C., Social power: a point missed in Multi-Agent, DAI, and HCI. In Y. Demazeau and J. P. Muller (eds.), *Decentralized A.I.*, North-Holland, Amsterdam, 1990
- [3] 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)
- [4] Castelfranchi, C., Guaranties for Autonomy in Cognitive Agent Architecture. In M.J. Woolridge and N. R. Jennings (eds.) *Intelligent Agents I*, Berlin, Springer, 1995
- [5] Castelfranchi, C., Practical "Permission": Dependence, Power, and Social Commitment. In J. Bell, Z. Huang, S. Parsons (eds.) *Proceedings of "Practical Reasoning and Rationality" PRR-2*, Manchester, April 1997, 171-82
- [6] Castelfranchi C. & Falcone, R., Principles of Trust for MAS: Cognitive Anatomy, Social Importance, and Quantification. ICMAS'98, Paris 2-8 July 98, AAAI-MIT Press.
- [7] Chaib-draa, B. Coordination between agents in routine, familiar, and unfamiliar situations. *Int. Journ. of Intelligent & Cooperative Information Systems*, 1995.
- [8] R. Conte and C. Castelfranchi. *Cognitive and Social Action*, UCL Press, London, 1995.
- [9] [Cra96] Crabtree B., Wiegand M., Davies J., Building Practical Agent-based Systems, PAAM Tutorial, London, 1996.
- [10] d'Inverno M., PhD. Thesis, Westminster, Univ. UK, 1998
- [11] Falcone, R., Castelfranchi, C., (2000), Levels of Delegation and Levels of Adoption as the basis for Adjustable Autonomy, *Lecture Notes in Artificial Intelligence* n°1792, pp.285-296.
- [12] Falcone, R., Castelfranchi, C., (2000), Grounding Autonomy Adjustment on Delegation and Trust Theory, *Journal of Experimental and Theoretical Artificial Intelligence*, (accepted).
- [13] S. Franklin and A. Graesser, Is it an Agent, or just a Program: A Taxonomy for Autonomous Agents. In J.P. Muller, M.J. Wooldridge, N.R. Jennings (eds.) *Intelligent Agents III*, Berlin, Springer, LNAI 1193, 1997.
- [14] Goodwin, R., Formalizing Properties of Agents. Technical report, CMU-CS-93-159, 1993.
- [15] Haddadi, A. and Sundermeyer, K. Belief-Desire-Intention Agent Architectures. In G.M. O'Hare and N.R. Jennings (eds.) *Foundations of Distributed Artificial Intelligence*, Wiley & Sons, London, 1996.
- [16] H. Hexmoor and D. Kortenkamp, (eds.), *Journal of Experimental and Theoretical Artificial Intelligence*, Special Issue on "Adjustable Autonomy" (in press)
- [17] Luck, M. and d'Inverno M. (1995) A Formal Framework for Agency and Autonomy. ICMAS-95, pp. 254-60
- [18] P. Maes, Situated agents can have goals. In P. Maes, editor, *Designing Autonomous Agents*, pp. 49-70. The MIT Press, 1990.
- [19] Martin C. E., Barber K. S., (1996), Multiple, Simultaneous Autonomy Levels for Agent-based Systems, in Proc. Fourth International Conference on Control, Automation, Robotics, and Vision, Westing Stamford, Singapore, pp.1318-1322.
- [20] Nwana, H.S. 1996 Software Agents: An Overview. *Knowledge Engineering Review*, 11 (3), 205-44.
- [21] Parisi, D. and Castelfranchi, C., Potere. *Rassegna Italiana di Sociologia*, 19, 1978, 555-592.
- [22] Plu, M., agents-list: <michel.plu@cnet.francetelecom.fr>To: <agents@cs.umbc.edu> Subject: autonomy Date: Mon, 28 Sep 1998 10:46:48 +0200
- [23] A. Rosenblueth and N. Wiener. Purposeful and Non-Purposeful Behavior. In *Modern systems research for the behavioral scientist*, Buckley, W. (ed.). Chicago: Aldine, 1968.
- [24] Shoham, Y., Agent-oriented programming. In *Artificial Intelligence*, 60, 1993.
- [25] Sichman, J.S., Du Raisonnement Social Chez les Agents. PhD. Thesis, Polytechnique de Grenoble, 1995.
- [26] Sichman, J.S., Conte, R., Castelfranchi, C., Demazeau, Y. A Social Reasoning Mechanism Based On Dependence Networks. In

¹¹ Deontic Dependence is a case of Jennings' "weak-dependence" [29]. In fact Y is practically able to do *a* but it *cannot do a without violating*, and since/if it is a normative agent it has the goal of not violating; if it does not want to violate, then it cannot do *a*. This is a very special and important form of lack of power, of being unable, of 'cannot'; what Parisi and Castelfranchi [21] called "lack of power by conflict or decision balance".

Proceedings of the European Conference on Artificial Intelligence - ECAI'94, Amsterdam August 8-12, 1994. (pp. 188-192.)

- [27] Verhagen, H. *Normative Autonomous Agents*. PhD. Thesis, University of Stockholm, May, 2000
- [28] Wooldridge M., and Jennings N., *Intelligent Agents: Theory and Practice*, The knowledge Engineering Review, Vol. 10, N.2, pp. 115-152, 1995.