

Absolute Model of Autonomy and Power: Toward Group Effects

Henry Hexmoor

Computer Science & Computer Engineering Department
Engineering Hall, Room 313
Fayetteville, AR 72701
hexmoor@uark.edu

Abstract

We present a model of absolute autonomy and power in agent systems. This absolute sense of autonomy captures the agent's liberty over an agent's preferences. Our model characterizes an affinity between autonomy and power. We argue that agents with similar individual autonomy and power experience an adjusted level of autonomy and power due to being in a group of like agents. We then illustrate our model on the problem of task allocation.

1. Introduction

Interaction among individual agents is an active sub-field of multiagent systems. Social notions such as autonomy, dependence, and power are being modeled between individual agents. Autonomy is a key distinguishing feature of agenthood and is closely related to the concepts of power, control and dependence [Barber, Goel, and Martin 2000; Brainov and Sandholm, 1999; Castelfranchi 2000]. Colloquially, there is a complimentary relationship between autonomy and power. Whereas, power is the experience of social abilities to influence, autonomy is the experience of limits of liberty.

Organizations are also being modeled in various fields [Carley 1999; Schillo, Zinnikus, and Fischer 2001]. The gap between studies of individual agents and issues in organizations is the relationships between agents and groups, and among groups. In this paper we take a preliminary step toward scaling autonomy, power, and dependence to groups of agents.

In this paper, we present a model that approximates absolute autonomy and power in agent systems. This absolute sense of autonomy departs from the relative notion of autonomy and captures the agent's liberty over its own preferences. This model also defines power. We then consider agent groups with shared autonomy and power. This group membership affects agents and alters their individual power and autonomy due to shared attitudes about their liberties. We then illustrate our model through the problem of task allocation and offer concluding remarks.

2. Types of Autonomy

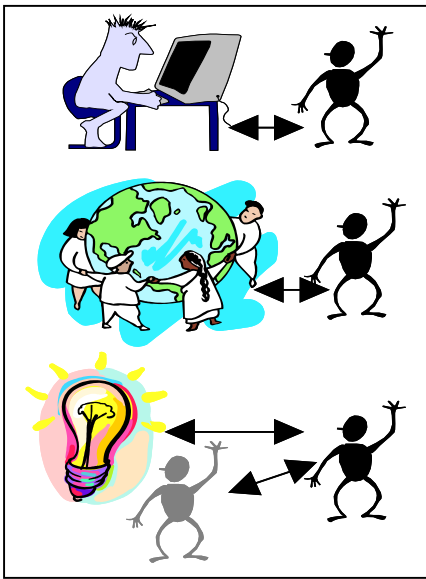
We consider an agent reacting in a rapidly changing environment and thereby consider it *situated*. Previously, we have presented agent autonomy as a relative sense of its individual preference over the intender or desirer of goals over which it has nontrivial abilities [Hexmoor, 2001]. The relative sense was suggested as the primary factor for an agent to choose its manner and level of involvement with other agents about a particular goal. We also presented a quantitative relative measure of autonomy [Brainov and Hexmoor, 2001]. In contrast, here we present an absolute sense and a corresponding measure that accounts for the agent's internal liberties over its preferences. The upshot of this absolute sense of autonomy is that it produces a bias for the agent over certain objects such as decisions, actions, goals, and intentions. This autonomy-induced bias is not the sole reason for the agent's choice but a strong contributing factor to such a reasoning model.

Let's further elaborate the absolute versus relative viewpoints on autonomy. In the relative sense, we are concerned with relative deviations in the agent's attitudes and functioning with respect to other things including other agents. The word "autonomous" connotes a relative sense of autonomy. For instance, consider an agent in service of a human. The agent is said to be fully autonomous when it has access to the complete set of choices and preferences of its user. Here the user is a distinguished entity that might judge or change an agent's autonomy. This is studied under *adjustable autonomy* [Goodrich, et al. 2001]. The agent's absolute autonomy is beyond subservience to the user and is not part of being considered *autonomous*. Considerations of a user are studied under *adjustable autonomy*. For simplicity in this paper, we assume a shared choice set among agents regarding absolute autonomy. Disparity in choice sets can be the basis of a generalized form of adjustable autonomy that examines relative difference in choice sets. This is beyond the scope of this paper and is left for future work.

Absolute autonomy considers the agent's internal manipulation of its own capabilities, its own liberties and what it allows itself to experience about the outside world

as a whole. Therefore, the agent formulates a liberty over how it functions in the world. It is not nuanced based on one thing or another but perhaps it is affected by its entire perception of its world. Absolute autonomy is different than *operational autonomy*, where the concern is the agent's ability to generate a reply plus capacity to be independent [Ziemke, 1998]. Absolute autonomy is also different than *behavioral autonomy*, where the concern is the agent's ability to originate behavior. With behavioral autonomy we are concerned with the agent's capacity to be original and not guided by outside sources [Boden 1996].

When an agent is viewed as a member of a large community such that the agent views the community as a single entity, absolute autonomy gains meaning. In contrast, in relative autonomy, an agent considers itself with respect to a specific agent(s) or object(s). Therefore, contexts set forth the meaning of specific type of autonomy. The following Figure summarizes these perspectives. The top drawing shows a user-human perceptible where adjustable autonomy makes sense. The middle drawing depicts user and the world perspective where absolute autonomy makes sense. We are not capturing self-sufficiency of the agent from its world but its freedom to choose [Mele, 2001]. The bottom drawing shows the agent with respect to another agent and a light bulb where relative autonomy makes sense.



2. A Model of Absolute Power and Autonomy

Let's consider an agent's choice set C at a particular moment. Objects in a choice set are of the same type and this might be a set of actions, a set of goals, a set of tasks, or sets of other mental notions. Apart from our autonomy considerations, the agent has non-deterministic choice

over the objects of its choice set. The agent might consider a context for its choice set. When this context is another agent or an object, then autonomy is a relative notion. With respect to the elements in the context, we can measure the agent's independence. This is exactly what it means to be *autonomous*. However, if this context is limited to agent internal notions, then autonomy is absolute. With respect to agent internal contexts, measures of autonomy are derived from agent internal notions. For example, an action a (e.g., taking the elevator), in the context of a corresponding goal g (e.g., going up 2 flights of stairs) and a particular situation (e.g., a typical hotel and typical circumstances of going up to the room alone) is treated as an internal context. If this context set is empty, then the agent's measure of autonomy is purely internal and measures the agent's own capacity for autonomy.

Let's consider an agent's preferences as a function denoted by P . Instead of preferences establishing order among an agent's choices, let's imagine the preference function to assign each choice in C a level of appropriateness purely from a means-end stand-point. This can be a number in the range -1.0 to 1.0 denoted by a function $P(c, Co)$. This preference might consider some choices highly appropriate and the agent would wish to promote it, while other choices might be considered highly undesirable and the agent would wish to suppress it. An agent's preferences may change from time to time based on her analysis of the situation. In the foregoing, we consider the agent to have nontrivial ability and readiness for its choices. Variations in an ability and readiness may contribute to changing situations that affect its P .

In addition to preferences, we consider liberties an agent experiences for adoption and suppression of preference for a choice. These liberties reflect contextual forces beyond means-end analysis such as agents' principles and conventions (i.e. values and norms), and emotions. An agent's liberties might change from time to time independent of its preferences.

We divide liberties into endogenous and exogenous types. Endogenous liberty is the forces derived from an agent's individualistic sources. We model this with a range 0.0 to 1.0 . An agent's endogenous liberty is 0.0 when it inhibits or prohibits the preference over choice, whereas liberty is 1.0 when it feels no inhibition but instead feels free toward the choice. Let function $Endo(c, P, Co)$ produce such a value for choice c in the context Co .

Exogenous liberty is the force derived from an agent's outside sources, that is, influences as the agent regards the world at large. We model this with range 0.0 to 1.0 with similar meaning as before, except this is about the agent's

social attitude. Let function $\text{Exo}(c, P, Co)$ produce such a value.

Agents differ in their application of endogenous versus exogenous liberties, and we model this by weights. $C1$ and $C2$ are the agent's independent relative weights to Individual Rationality and Social Rationality. Individual rationality is when an agent considers its own welfare to the exclusion of other agents. Social rationality is when an agent pays attention to welfare of other agents. Social rationality is used to develop a theory of joint responsibility [Jennings and Campos, 1997]. To summarize, our model of autonomy contains the following 7-tuple model: $\langle C, Co, P, \text{Endo}, \text{Exo}, C1, C2 \rangle$.

Next, we define absolute autonomy and power.

Definition 1: The amount of Absolute Autonomy of an agent's choice c , which belongs to choice set C in the context of Co and preference P , is the sum of the endogenous and exogenous freedoms over c weighted by the agent's independent relative weights to individual and social rationalities. I.e., $[(1+C1-C2) * \text{Endo} + (1+C2-C1) * \text{Exo}] / 2.0$.

The values of absolute autonomy range between 0.0-1.0. At 0.0, the agent has no autonomy, whereas at 1.0, the agent's autonomy is greatest. For example, if c is an action and Co is a set of goals, P is how appropriate the action is in the context of the goals.

When there is a sharp discrepancy between Endo and Exo values, there is power between agents. This is not individual capability (i.e., personal power) but the amount of influence the agent perceives. This is consistent with viewing power as a property of a social relation rather than a characteristic of an individual [Emerson, 1962]. Here we are not considering effects of organizational structure and the agent's position in the structure, which can further affect the agent's power. The following power is the power an agent experiences and not the power it exerts.

Definition 2: The amount of Absolute Power of an agent's choice c , which belongs to choice set C in the context of Co and preference P , is the difference between endogenous and exogenous freedoms over c weighted by the agent's independent relative weights to individual and social rationalities. I.e., $|(1+C1-C2) * \text{Endo} - (1+C2-C1) * \text{Exo}| / 2.0$.

The values of potential amount of power range between 0.0-1.0. At 0.0, the agent has no power difference with other agents. At 1.0, this power difference is greatest.

By focusing on relations between individuals, Emerson suggests a direct relationship to dependence between those individuals [Emerson, 1962]. However, since our definition does not pick specific individuals but treats others as a unit, we conjecture that that dependence is a form of the agent's generalized dependence on the larger community.

Definition 3: The Potential of amount of Absolute Dependence about choice c in the context of Co and preference P is the amount of absolute power the agent has conceded.

An agent that goes along with the group despite the discrepancy between Exo and Endo values is conceding to the absolute power. Such an agent keeps its difference in the "closet" (i.e., the power difference does not cause the agent to dissent). In contrast to a dissenting agent, an agent with a power difference, which chooses to go against the group is defending its decision by the amount of absolute power and is a "rebel". Let's focus on multiple agents and build a notion of a power group.

3. Group Effects

Let's consider a number of agents that perceive social pressures, where the balance of their Exo and Endo values are similar. We assume the group is entirely within a larger group, which we will call a society. An example is a nudist colony. Members of this group all allow themselves nudity and feel similar social pressures from outside their group. This group is defending against powers of clothing conventions in the larger society and has come together. For another example, consider a robot Mars explorer encountering an interesting feature in a somewhat steep slope. Whereas all robots are instructed to avoid steep slopes of the kind encountered, the robot feels some power to abandon exploration of the slope. Imagine a second robot encountering the same situation and experiencing the same power. Together they are in the same power group with respect to the larger society of their space mission. However, since they are in a group, it is conceivable that they individually feel that power is reduced (see [Abrams and Hogg] for more on groups effects).

Definition 4: A *single-issue power group* is a number of agents with similar Endo values about a decision c , which either unanimously concede or defend absolute power about c , given that Exo values are similar for the group members and Exo is derived from outside the group.

¹ We use Endo and Exo for values returned to corresponding functions.

A special case of a single-issue power group is one composed of non-members of the group that experiences pressures about the decision c . This group has power level of 0.0. Agents might share Endo and Exo levels for more than one decision and this will be an n -issue power group. Agents who share power differentials with the general population of agents also have the same levels of dependence with the population. A group effect from sharing dependence is co-dependence, which amplifies members' dependence more than their individual levels of power. We will not formalize this notion but mention it passing.

Let's consider an agent's entire choice set, where an agent has no power difference with a few choices, concedes with a few, yet rebels against other choices. Agents might also form groups that may not have similar power differences on specific choices, but might be similar due to the balance of their power ratios. For example, a number of "rebels", "closets", or good citizens might form groups. This type of group formation might also provide its members with power amelioration bonuses. This is a preliminary observation and needs to be further explored.

Next we will extend the single-issue power to groups with multiple issues.

Definition 5: An n -issue power group is a number of agents with similar Endo values about n decisions, which either unanimously concede or defend absolute power about c given that Exo value is similar for the group members and Exo is derived from outside the group.

Agents can be compared on differences in their power over choices. Agents in n -power group have n decisions in common.

Definition 6: The power distance between two agents A_1 and A_2 who form an n -power group, from a total of m choices available, is $(m-n)/m$.

Agents with no choices in common have a distance of 1.0. Agents who have every decision in common have a distance of 0.0. The group, which has 0 absolute power level is considered as a *base population*.

Definition 7: A base population is a group of agents whose power distance is 0 and all individually have 0 absolute power level.

This group is homogenous and distance among agents is 0. In addition to the base population, there might be other populations where distance among agents in the group is zero. All agents in a 0-distance group will have the same

distance from agents in the base population. In the following theorem, we make the observation that there is transitivity in distances in two exceptional cases where distances are 0 or 1.

Theorem 1: Consider two groups A and B with 0-distance in each group and a distance of 0 (or 1) between the base population and group A (or B) and between group A and group B . It follows transitively that the distance between the base population and B (or A) is 0 (or 1).

Proof: In the 0-distance case, transitivity is obvious. In the 1-distance, if either group A (or B) differs from the base group, members of the group all have non-zero powers. If the groups A and B have distance 1, it means that the difference between powers levels are maximal. When one decision is conceded it is defended by the other and vice versa. Therefore, they are similarly different with the base population.

In the following theorems, we make two other observations about distances among 0-distance groups who have certain distances with the base population.

Theorem 2: Consider two 0-distance groups: A where a member has a distance of x (less than $\frac{1}{2}$) from any member of the base population, and B where a member has a distance of y (less than $\frac{1}{2}$) from any member of the base population. The distances x and y can be the same. Then the distance between a member of group A is at most $x+y$ from any member of group B and vice versa between B and A .

Proof: There is overlap in choice sets of A and B when the distance with each group and the base population is greater than $\frac{1}{2}$.

Theorem 3: Consider two 0-distance groups: A where a member has a distance of x (greater than $\frac{1}{2}$) from any member of the base population, and B where a member has a distance of y (greater than $\frac{1}{2}$) from any member of the base population. The distances x and y can be the same. Then the distance between a member of group A is possible to be the maximum of x and y from any member of group B and vice versa between B and A .

Proof: There might be complete overlap between A and B in their similarity of choices with the base population.

Naturally, members of a power group will feel higher autonomies in their own group than with respect to outsiders. But there is more to the story. The group affords their members a softening effect in the power and autonomy. Apart from increased autonomy and lowered

power an agent may experience within a group, a group member will feel a group effect that ameliorates the power and lowered autonomy it feels toward outsiders. In general, this is a difficult effect to model. But a crude measure is ratio of number of people in the community to the group in question. The next definition captures this notion. To lesser degree, the group effect also affects an agent outside the power group. General public, not members of a power group will have somewhat diminished autonomy and increased sense of power differential due to large groups.

Definition 8: A group of n agents with similar Endo value about a decision c and absolute power level P , in a community of m agents (not including the group members) with similar Endo values about c will experience a normalized $(n/m) P$.

The group effect that is captured in the above definition as a multiplication factor of n/m can be adjusted to control group behavior. It appears desirable to enable a human operator in charge of a group of agents the ability to adjust group effect. We have only pointed the way to the notion of a sense of power not possessed by an individual but a power that belongs to a group. We will leave this for future work.

In designing and deploying a system of agents, we must envision effects of groups of agents on their individual autonomy and power.

4. A task delegation algorithm

Let's imagine a shared choice set among a group of agents is the steps of a plan and agents are trying to delegate the steps among themselves with the least amount of potential autonomy/power conflict. Here the plan steps are generic and not the actual actions an agent may take. For example, this can be plans to repair a jammed printer with steps such as uncover paper loading compartment, examine paper path, etc. We could imagine three agents attempting to correct this situation and each will be assigned to a single plan step. We can construct a matrix with each row corresponding to each agent's autonomy power with respect to a plan step and columns are the plan steps. A "0" would indicate the agent has no power issue about that plan step and has the most autonomy. A "+x" indicates that the agent has a tendency to concede with the power it experiences from other agents. A "-x" indicates that the agent has a tendency to go against the power it experiences from other agents. The following algorithm assigns the task in the least conflicting manner. We assume the base population is substantially larger than other groups.

While there are unassigned plan steps,

1. Pick the column(s) with the most 0's first, most amount of +x's next and then most amount of -x's.
2. Among the selected column(s), select the row with the least 0's first, least amount of +x's next, and least amount of -x's and assign it to the corresponding agent.
3. If more than agent, consider group effects, otherwise pick randomly.

This algorithm assigns plan steps to agents with least power bias and then proceeds assigning steps with more bias. A side effect is that agents in a power group will be assigned similarly biased plan steps.

Agents may belong to multiple groups. With a decision set C of cardinality n , the power set will be 2^n decision group an agent may belong.

5. Conclusion

We have developed a model of absolute autonomy and power. This enabled us to consider groups with shared power differentials. Common sensically, group memberships alters absolute power and autonomy levels. We remarked on this effect and hope to investigate how it can benefit agents systems. An algorithm for task allocation was discussed that showed a use of grouping agents into power groups. We believe reasoning about absolute autonomies and power is useful for teams of agents.

Acknowledgement

This work is supported by AFOSR grant F49620-00-1-0302.

References

- D. Abrams and M. A Hogg (Editors), *Journal of Group Processes & Intergroup Relations*, Sage publications.
- K. S., Barber, A. Goel, and C.E. Martin, 2000. *Dynamic Adaptive Autonomy in Multi-Agent Systems*, In *Journal of Experimental and Theoretical Artificial Intelligence*, 12(2): 129-147.
- M.A. Boden, 1996. Autonomy and artificiality. In Boden, *The Philosophy of Artificial Life*, Alife.
- S. Brainov, H. Hexmoor, 2001. *Quantifying Relative Autonomy in Multiagent Interaction*, In *IJCAI-01 Workshop, Autonomy, Delegation, and Control*.
- S. Brainov, T. Sandholm, 1999. *Power, Dependence and Stability in Multiagent Plans*. *AAAI/IAAI 1999*: 11-16.
- K. M Carley, 1999. On the Evolution of Social and Organizational Networks, In *Research in the Sociology of Organizations* 16, 3-30.
- C. Castelfranchi, 2000. *Founding Agent's Autonomy on Dependence Theory*, In *proceedings of ECAI'01*, pp. 353-357, Berlin.
- R. Emerson, 1962. Power-Dependence Relations, *American Sociological Review*. 1962; 27: 31-41.
- H. Hexmoor, 2001. Stages of Autonomy Determination, *IEEE Transactions on Man, Machine, and cybernetics-Part C (SMC-C)*, Vol. 31, No. 4, Pages 509-517, November 2001.
- M. A. Goodrich, D. R. Olsen, J. W. Crandall, and T.J. Palmer, Experiments in Adjustable Autonomy, In *proceedings of the IJCAI01 workshop on Autonomy, Delegation, and Control: Interacting with Autonomous Agents*, Seattle, WA.
- N. R. Jennings and J. R. Campos, 1997. Towards a Social Level Characterisation of Socially Responsible Agents, In *IEEE Proceedings on Software Engineering*, 144 (1), 1997, 11-25.
- A. Mele, 2001. *Autonomous Agents: From Self-Control to Autonomy*, Oxford University Press.
- M. Schillo, I. Zinnikus, and K. Fischer: 2001. Towards a Theory of Flexible Holons: Modelling Institutions for Making Multi-Agent Systems Robust. 2nd Workshop on Norms and Institutions in MAS.
- E.R. Smith, and D.M. Mackie, 2000. *Social psychology* (second edition). Philadelphia: Psychology Press.
- T. Ziemke, 1998. Adaptive Behavior in Autonomous Agents. *PRESENCE*, 7(6), special issue on *Autonomous Agents, Adaptive Behaviors, and Distributed Simulations*, 564-587.