Conceptualizing Social Power for Agents

Gonçalo Pereira, Rui Prada, and Pedro A. Santos

INESC-ID and Instituto Superior Técnico and Technical University of Lisbon
Avenida Professor Cavaco Silva, Porto Salvo, Portugal
goncalo.pereira@gaips.inesc-id.pt, rui.prada@ist.utl.pt,
pasantos@math.ist.utl.pt
http://gaips.inesc-id.pt/

Abstract. One of the most pervasive concepts in human interactions is social power since many social situations entail disputes of social power. These disputes are power games and range from simple personal reasoning to the exercise of specific power strategies, which enhance or assert one's power. Recognizing the importance of such interactions and how they can enhance autonomous agents' socially intelligent behaviors, we present a formalization of the fundamental bases of power and conceptualize the diverse forces that should underlie an agent's deliberative decision process. Different bases of power stem from diverse factors and have particular dynamics and effects. The objective of this work is to establish a theoretical basis for social intelligent agents capable of both being aware of and manipulating social power.

Keywords: social power, autonomous agents, social intelligence.

1 Introduction

The motivation for studying social power, relates to its ability to act as a *social heuristic* [10] in many social situations such as friends' interactions [10], organizations [15] or even laboratory experiments [14] among others. Consequently, the behavior effects of power are broad, extending to social processes such as coordination, delegation, cooperation, hierarchy formation, alliance formation, resources allocation, conflict resolution and negotiation [5,10]. Given the widespread impact of power in people's attitude and behavior, it is fundamental to understand and emulate such power-based social dynamics in multi-agent systems in order to build socially intelligent agents.

Even though power has been previously explored as a social heuristic for agent behavior [1,3,19,18], these approaches have shortcomings when considered in the context of social intelligence for agent believability, namely inter-agent [13] and agent-human interactions [16]. Current approaches do not take into account the different bases of power as established by French and Raven [7] the particular dynamics and the contrasting effects associated with each one. Modeling the different bases of social power is crucial to simulate the complex processes behind social power and its ubiquitous influence in social interactions. For instance, even in a relatively simple situation where there are just two agents,

playing the role of father and son, and a single decision, the son arriving home at the ordered time, many different types of power are at play. If one wants to take into consideration the possibility of power related to reward, punishment, legitimacy, rebel tendencies and love, the current models are too simplistic.

Our objective is to define a model of social power that can cover all these possibilities. To address this we propose a conceptual framework to support agent perception, reasoning and intelligent use of social power aimed at multi-agent and agent-human interactions. The framework was developed from fundamental concepts grounded in the seminal work from social psychology of French and Raven [7] addressing the cognitive ingredients of social power. By using those concepts, we argue that a broader range of social settings can be more accurately modeled in agent simulations through power-based reasoning and interaction.

This document is structured as follows. In Related Work we review previous contributions regarding power in multiagent systems. Next, we present our framework for social power aware agents, followed by an example scenario of its application. Finally we offer some conclusions and future directions.

2   Related Work

The subject of social power, namely the representation of power and the formalization of the associated dynamics have been previously researched from several approaches. The fundamental difference between them is the origin of the represented power, i.e. the main focus of the approach being described.

One approach to power is based on autonomy and is rooted in an agent's capability of pursuing its goals without the intervention of other agents. This approach's central work by Hexmoor [9] addresses "absolute autonomy" which is described as the measure of an agent's liberties (internal and external) over preferences. In this model an agent's power is modeled based on liberties and preferences. Liberties express freedom/inhibition forces regarding an agent decision and there are two types of liberties: endogenous (e.g. values and emotions) and exogenous (e.g. physical context limitations). The preferences affect the liberties forces according to an agent's characteristics: individual rationality (prefers individual welfare) or social rationality (prefers social welfare).

Many works are deeply rooted in the seminal work on Dependence Theory by Sichman and Conte [18] where a taxonomy of dependence and the fundamental concepts for agent reasoning in groups is supported through the formation of dependence networks. In [6] Castelfranchi relates the previous definitions of Dependence Theory to the concept of autonomy, and supports it on two subtypes, practical and deontic autonomy. Based on these two constituents of autonomy, in [3] personal power is grounded on an agent's individual capabilities and social power is formalized as originating from dependence relations of an individual regarding another's individual powers. In [2] dependence is related to utility based Decision Theory in order to reason about power in multi-agent plans. Power is conceptualized based on the possible costs and damages expected from a given dependence situation to the involved agents. These values are calculated as differences of utility between an intended joint plan and a possible alternative plan

from one of the agents. In this work power is defined as the capacity to harm another agent, i.e. perform a plan that can lower the utility of the other.

Another approach is normative and emphasizes the part that norms and roles play into a group's power structures and individual powers. The focus is on the restrictions an individual experiences in a given group, and also the benefits it collects from it. The work of López [12] is central to this approach and addresses the impact that power can have on agent behavior regarding the social processes of agent membership to a normative society, individual norm adoption and inter-agent goal delegation. At the core of these processes are the powers agents have due to their roles and personal capabilities. In [12] a taxonomy of powers is presented and agent powers are categorized under "Circumstancial Powers" (facilitation, illegal coercive, exchange, reciprocation, support) and "Institutional Powers" (legal, legal benefit, legal preventive, legal punishment, legal reward).

In [8] power is regarded as an exchangeable natural resource in the agents' environment where it replenishes/decays automatically. For an agent to influence another, it must transfer a given amount of power to it. In such a situation the influenced agent can resist to the influencer by transferring a higher amount of power back.

2.1   Discussion

The autonomy perspective heavily constraints the behavioral diversity of agents by abstracting all the influencing factors of an agent's power to four values. Especially the liberties value abstract components of agent's decisions which should be taken into account separately, for example different norms. The perspective of power as an exchangeable resource also largely disregards the behavioral aspects of power in which we are interested in. For example the perspective of power as resource that is spent or depleted upon usage when in fact a power (e.g. expert knowledge) might be used repeatedly without depletion.

The Dependence Theory approach relates more to agent reasoning about its individual ability or capacity of performing actions independently. However, social power is based on more than dependence. There is power grounded on relational factors such as attraction or credibility. Consider a situation where a music band fan dresses like its members because it wants to be like them. This situation cannot be represented in a context of pure dependence.

Finally, the normative approach connects power with norms and agent membership to a normative society so tightly that it is hard to model some situations. For example in many cases an agent society is defined a priory, e.g. the culture where we are born, no deliberation needed. Another difficulty relates to the taxonomy of powers which seems to limit power to specifically two types of context, circumstantial and institutional. An indication of this strict separation are the powers "illegal coercive power" and "legal punishment power". We argue that social powers are present in all contexts and they can be identified, used or manipulated in a given situation.

3   Power Based Agents

Our main goal is to create an agent architecture with social intelligence capabilities regarding social power. In this work our focus is on the core concepts which make such an architecture possible. Our contribution is two-fold, first a psychology inspired conceptualization of the different bases of power to enhance agents with the capability of identifying powers and perform situational analysis regarding power. Second, the description of the fundamental mechanisms of a power-based decision and the factors which underlie each base of power in that decision.

3.1   Elements of Social Power

But what is social power? The subject has been researched from different perspectives and there are many definitions. However, the one we follow is an adapted version from Lewin's [11] and Cartwright's [4] definitions of social power: "**Social Power** of A over T regarding a possible change in T is the resultant potential force that A can induce on T towards that change.". This definition captures the essence of power as a potential force that results from the accumulation of a variety of social power components with different sources.

   Before specifying the different types of social power and their characteristics we must first introduce the three central elements to such a situation:

Actor (A) the agent which exerts power over the actions of another;
Target (T) the agent whose actions are affected by the Actor's power;
Action (C) the action evaluated by the Target in a given interaction.

3.2   Fundamental Notions

Our *MultiagentSystem* is formalized as the 5-tuple $\langle$*Agents, Actions, Roles, Rights, Relationships*$\rangle$. *Agents* is the set of agents in the multi-agent system, where each contains its own set of beliefs and goals. *Actions* is the set of actions available for the Agents to perform. *Roles* is the set of roles available in the society of Agents. The *Rights* is the set of right / obligation contract relationships. The predicate $Right\_to\_Influence(r_1, r_2, a)$ will be used to specify a mapping, belonging to *Rights*, between the right of role r_1 to influence role r_2 regarding action a. Finally the *Relationships* is the set of social relationships between the agents which will be used to represent information regarding how an agent regards another and can be of different types in the set $RelTypes = \{liking, friendliness, attraction\}$. The predicate $Relationship(a_1, a_2, t)$ will be used to represent that agent a_1 has a relationship of type t with a_2 (the unidirectionality is intentional in order to mimic human relations).

   An important notion for the conceptualization of power is context, since every power has an associated context. For example, an university teacher might easily prescribe actions to his master or doctoral students in an academic context due to his superior skill difference, but in a context of personal relationships he

cannot use the same power to direct behaviors of others in the same way. The context has, however, different meanings depending on the power base under consideration. For example, for a coercive power, the context is the resources and skills needed to perform that coercive action. When making an implementation of the defined predicates and functions, those contexts must be also implemented, as they are necessary for the predicates and functions to be well-defined.

3.3   Identifying Different Bases of Power

Many social power studies propose a set of bases of power. However, most can be represented by one of the first sets introduced by French and Raven [7]. Their work introduced a differentiation and dynamics of social power grounded on five bases of power: reward, coercive, legitimate, referent and expert. Our work is inspired on these bases of power due to their simplicity, behavior expressive potential and repeated validation over the years. We will formalize the five bases of power with four categories.

Welfare Power. This power is based on the ability of the Actor to mediate some welfare (reward or coercion) to the Target. It is formalized in definition (1): if there are two agents A and T, where A can do an action a which T values (positively or negatively), then we are in a situation where A has *Welfare_Power* over T. In this definition the predicate $Values(T, a)$ represents that action a brings about some benefit or harm to T. Additionally, the predicate $Can\_Do(A, a)$ represents the ability of agent A to perform action a. The force of an instance from this power increases with the value of a and also with the T's believed probability that it will perform a.

$$\exists a \in Actions \ \wedge \ a \neq C \ \wedge \ \{A, T\} \subseteq Agents$$

$$Welfare\_Power(A, T, a) := Values(T, a) \wedge Can\_Do(A, a) \qquad (1)$$

This power results from the abstraction of French and Raven's reward and coercive power bases under the *Values* predicate. Regarding the identification of the power an agent has over another the two underlying bases of power simply have symmetric valuations regarding a for T. In the reward case a is desired and in the coercive case a is undesired or avoided. However, even though these power bases have symmetric dynamics regarding their identification they have different effects when the interplay between power bases is considered [7].

A crucial example is that in a case of illegitimate coercion (e.g. bullying) the negative effects of using coercion are exacerbated, while in the illegitimate reward case (e.g. a bribe) it depends upon the personality characteristics (e.g. the relation of the conscientiousness trait with dutifulness) of the Target.

Legitimate Power. The power based on internalized beliefs in the Target regarding the right for the Actor to influence the Target and its obligation to accept that influence. It is formalized in definition (2): if there are two agents A and T with a *Rights* relationship confering role r_1 the right to influence role r_2

towards the compliance of C, where A plays role r_1 and T plays role r_2, then A has *Legitimate_Power* over T regarding C. In this definition the predicate $Plays(X, r)$ represents that the agent X assumes the role r. The force of an instance from this power increases with T's degree of adherence to the role he plays and also with T's believed probability that A will legitimately enforce C.

$$\exists r_1, r_2 \in Roles \ \wedge \ \{A, T\} \subseteq Agents$$

$$Legitimate\_Power(A, T, C) := Plays(A, r_1) \wedge Plays(T, r_2) \wedge$$
$$\wedge \ Right\_to\_Influence(r_1, r_2, C) \quad (2)$$

In multi-agent systems there are several representations which fit this generic description of internalized values. As presented in the work of Carabella *et al.*[3] there are norms (formal or informal), contracts and commitments.

Referent Power. The power based on the identification of the Actor with the Target. It is formalized in definition (3): if there are two agents A and T, where T acknowledges a relationship of type t with A, then A has referent power over T. In this definition the predicate $Relationship(T, A, t)$ acknowledges the relationship of type $t \in RelTypes$ and from T with A. The predicate $Identifies(T, A, t)$ represents the recognition of the relationship factors of type t between T and A. The valuation is based on factors underlying each type of relationship and it can be positive (e.g. attract, like) or negative (e.g. repulse, dislike). Additionally, notice that the relationship does not need to be bidirectional, a person can acknowledge and establish a "relationship" with another and the later not even know of the first's existence. The force of an instance from this power increases with T's magnitude of identification with A (based to the value of the liking/friendliness/attraction) and also with T's probability of identification with A.

$$\{A, T\} \subseteq Agents \ \wedge \ Relationship(T, A, t)$$
$$Referent\_Power(A, T) := Identifies(T, A, t) \quad (3)$$

Expert Power. The power based on the perceived skill difference between the Actor and the Target. It is formalized in definition (4): if there are two agents A and T, where A has a higher skill than T regarding the knowledge domain of the interaction C between the agents, then A has *Expert_Power* over T. In this definition the predicate $Higher(X, Y)$ represents that X is higher than Y. The function $topic(X)$ identifies the knowledge domain of an interaction. Additionally, function $skill(X, Y)$ quantifies the skill of X regarding the knowledge domain Y. The force of an instance from this power increases with T's believed skill difference to A regarding $topic(C)$ and also with the T's believed credibility (a probability value) of A regarding $topic(C)$.

$$\{A, T\} \subseteq Agents$$
$$Expert\_Power(A, T, C) := Higher(skill(A, topic(C)), skill(T, topic(C))) \quad (4)$$

3.4 Social Power Decision Mechanism

The social power bases identified in the previous section impact intelligent agents' decisions in the environment. As such we must be able to operationalize them and to do so we consider an environment representation as presented in 3.2 and the basic elements of a social power interaction: Actor(A), Target(T), Action(C). To operationalize an agent decision in an influence attempt situation by A over T regarding C, we define the possible environment outcome for the case in which it decides to perform C as $S_{f,C}$ in (5) and for the case it decides not to perform it as $S_{f,\neg C}$ in (6). Notice that S_i represents the environment state before the decision.

$$S_{f,C} = Do(T, C, S_i) \qquad (5)$$

$$S_{f,\neg C} = \neg Do(T, C, S_i) \qquad (6)$$

In a given social power interaction, an agent may identify several powers according to the different power bases. This set of powers may contain powers from only one base (e.g. have several punishments) or be composed by powers of several distinct bases. We define the set of identified power bases as $IdentifiedPowers$ (IP). Each of these powers is a $Force$ (F) exerted by the Actor on the Target, influencing its decision. To operationalize each force we quantify its strength based on a probability and a magnitude according to equation (7). The probability represents the perception of mediation capability (e.g. from history of interaction) and the magnitude captures the strength of the power underlying factor (e.g. value of a coercion or degree of liking).

$$p \in IdentifiedPowers$$

$$Force_p = probability_p * magnitude_p \qquad (7)$$

The probability component of the $Force$ assumes values between $[0, 1]$. The higher the probability value the stronger the force. The magnitude component of the $Force$ assumes values in $\mathbb{R}$ in which positive values represent a positive influence towards the decision $S_{f,C}$ and negative ones towards $S_{f,\neg C}$ (conceptualizing the concept of negative power[17], e.g. from disliking). The total social power force exerted by the Actor over the Target in a given situation is the sum of all the individual forces in the IP, as presented in equation (8).

$$social\_power\_force(T, A, C) = \sum_{p \in IP} Force_p \qquad (8)$$

Besides social power there is another indisputable force in agent decisions: utility. We can better expose this parallel force by comparing two situations. First, if a person asks a friend to lend him a cellphone temporarily. It is reasonable that the friend does so given the friendship relation and the low loss of utility he experiences by lending the cellphone temporarily. In a second situation the person now asks the friend to buy him a cellphone. In this case he is also

reasonable that the friend refuses given the high loss of utility he would experience by buying the cellphone. Generically utility measures the usefulness of world states regarding the agent's goals. Notice that this can incorporate many other social concepts such as emotions.

In our mechanism we model the utility force according to this comparison of the world states as presented in equation (9). The *utility_force* assumes values in $\mathbb{R}$ and increases with the increase of $utility(T, S_{f,C})$ or decrease of $utility(T, S_{f,\neg C})$ and vice versa. Positive values represent beneficial situations (favoring final state $S_{f,C}$), negative values harmful ones, and when 0 indicates indifference.

$$utility\_force(T, C) = utility(T, S_{f,C}) - utility(T, S_{f,\neg C}) \qquad (9)$$

Finally we model the agent's decision to either perform C or not ($\neg C$) as a combination of the two major forces identified: social power and utility. To do so we assume a simple resultant force approach in accordance with our definition of power (see 3.1) and represented in equation (10).

$$res\_force(T, A, C) = utility\_force(T, C) \ + social\_power\_force(T, A, C)$$
$$Decision = \begin{cases} Do(T, C, S_i), & \text{if } res\_force(T, A, C) > 0 \\ \neg Do(T, C, S_i), & \text{if } res\_force(T, A, C) \leq 0 \end{cases} \qquad (10)$$

If the value is positive then the agent choses C, if not then it choses $\neg C$. Notice that this decision formalization takes into account the resistance that an agent can offer to a given prescription of behavior from another agent [7]. This is present at two distinct levels. First, if utility is negative it represents an opposing force to the social power being exerted. Second, for each *Force* (from any base of power) the magnitude can reflect negative power, which is also another form of opposing force to the influence attempt. For instance when a person dislikes another this will be represented as a negative force (in case of any influence attempt) from the referent power base.

4   Example Scenario

Consider a situation where a boy is going out with his friends, but before leaving home he is instructed by his father to be at home before midnight. At a certain point after leaving home, and before midnight, he will be faced with the decision to either do as told by his father or defy his wishes. What are the forces at play for the boy's decision? How do we model it under our framework? In this situation the boy is the Target (T), his father the Actor (A) and the Action (C) upon which he must decide is "be at home before midnight".

4.1   Initial Situation

The agent set for this scenario is $Agents = \{boy, father\}$ where $A = father$ and $T = son$. The *Roles* and *Rights* for the situation depend on the existing norms for a specific family which can vary a lot. However, for this example we can at least consider one (informal) norm which is frequently adopted

in families: children should always obey to their parents. Based on this we have $Roles = \{parent, children\}$ and the rights relationship for the situation $Rights = \{\{parent, children\}\}$ meaning that agents in the role of parent have the $Right\_to\_Influence$ agents in the role children. As for the relationships they are represented by the following set $Relationships = \{\{boy, father, like\},$ $\{father, boy, like\}\}$ meaning that there is a bi-directional relationship of type "liking" between the boy and father. Finally as for actions we consider the following set of possible actions $Actions = \{WithdrawAllowance, GroundChild,$ $AllowReturnLate, ReturnOnTime\}$ where the action $C = ReturnOnTime$. Finally, the moment before the boy makes his decision is S_i and the moment after he decides is either $S_{f,ReturnOnTime}$ or $S_{f,\neg ReturnOnTime}$.

4.2 Modeling Boy's Decision

Following our conceptual framework, the *boy* starts by analyzing what power categories are at play by using the definitions (1) to (4):

1. $Welfare\_Power(father, boy, WithdrawAllowance)$
2. $Welfare\_Power(father, boy, GroundChild)$
3. $Welfare\_Power(father, boy, AllowReturnLate)$
4. $Legitimate\_Power(father, boy, ReturnOnTime)$
5. $Referent\_Power(father, boy)$

The welfare powers represent the $father$'s ability to punish the boy by withdrawing his allowance (1) or ground him (2) on future opportunities to go out with its friends. As for benefits, the $father$ can allow the boy to stay out longer (3) next time if it behaves properly this time. The legitimate power (4) represents the $father$'s right to influence the boy regarding $ReturnOnTime$, since boy plays the role of child and $father$ that of parent. The referent power (5) is based on the son-father relationship between the two agents. Expert Power does not exert any force because there is no relevant skill in this situation.

The *boy*'s decision between $S_{f,ReturnOnTime} = Do(boy, ReturnOnTime, S_i)$ or $S_{f,\neg ReturnOnTime} = \neg Do(boy, ReturnOnTime, S_i)$ then only depends on the actual values of each of these forces. From the perspective of what happens in real life both cases are believable given the appropriate personal characteristics, relationships and beliefs. We will present both cases accordingly.

4.3 Case 1: A Well-behaved Child

If we assume a well-behaved son all the social power forces work towards C, and even though the utility offers resistance, he will probably return home before midnight. In order to illustrate this case let's consider the values in table 1 and the utility values $Utility_{S_{f,C}} = -100$ and $Utility_{S_{f,\neg C}} = 20$. Notice that for space reasons we abbreviated the names of the components. For example, for the power situation $Referent\_Power(father, boy)$, in the table its probability component and value component are represented by P_{RP1} and V_{RP1} respectively.

Table 1. Example values for the identified social powers

Property	P_{WP1}	V_{WP1}	P_{WP2}	V_{WP2}	P_{WP3}	V_{WP3}	P_{LP1}	V_{LP1}	P_{RP1}	V_{RP1}
Value	0.2	100	0.7	20	0.15	50	0.95	100	0.6	40

Based on table 1 we can then calculate the *utility_force* using definition (9):

$$utility\_force(boy, ReturnOnTime) = -100 - 20 = -120$$

From the previous subsection (see 4.2) we can also determine the set *IdentifiedPowers* = {*WP1, WP2, WP3, LP1, RP1*}. Now according to definitions (7) and (8) it is possible to calculate the social power force:

$$social\_power\_force(father, boy, ReturnOnTime) =$$
$$= F_{WP1} + F_{WP2} + F_{WP3} + F_{LP1} + F_{RP1} =$$
$$= 0.2 * 100 + 0.7 * 20 + 0.15 * 50 + 0.95 * 100 + 0.6 * 40 = 160.5$$

Once we have both the *social_power_force* and the *utility_force* we can then calculate the *res_force* and know the decision of the agent *boy* according to (10). Therefore the resultant force value is the following:

$$res\_force(boy, father, ReturnOnTime) = -120 + 160.5 = 40.5$$

Based on the value of the *res_force* and the condition *res_force(boy, father, ReturnOnTime)* > 0 from definition (10) the *boy* agent decides for $Do(boy, ReturnOnTime, S_i)$ meaning that it will return home on time, before midnight.

4.4   Case 2: Disobedience in a Well-behaved Child

Now imagine the *boy* is having more fun than he has ever had? In this case it might occur the situation where a well behaved son actually disobeys to the father's command since the utility force can surpass the social power force exerted by his father. This means the utility evaluation of the situation changes drastically. For example, consider the values $Utility_{S_{f,C}} = -100$ and $Utility_{S_{f,\neg C}} = 100$ and table 1 with the same values. Since the values associated with the *social_power_force* did not change, its value remains the same. However, the new utility force is the following:

$$utility\_force(boy, ReturnOnTime) = -100 - 100 = -200$$

The impact of this change in the *res_force* is given by:

$$res\_force(boy, father, ReturnOnTime) = -200 + 160.5 = -39.5$$

Based on the value of the *res_force* and the condition *res_force(boy, father, ReturnOnTime)* ≤ 0 from definition (10) the *boy* agent decides for $Do(boy, \neg ReturnOnTime, S_i)$ meaning that it will not return home on time and will stay out late.

4.5 Other Complex Cases

An additional interesting case would be that of the rebellious son, where the legitimate component of the social power force now exerts negative influence since the command from the father actually works as a resistance force to $ReturnOnTime$. In this case, the rebellious son is much more probable to stay out with its friends (do $\neg C$) even in the first case. Notice that this does not mean that he "does not like" his father, a son might like his father and simultaneously exhibit rebellious behavior. Our model enables the simulation of these complex situations.

5 Conclusions and Future Work

In this work we introduced a conceptual framework for agent social intelligence regarding social power awareness. The identification of situations where different sources of power emerge enable agents to take many social influences that an individual has to deal with when making decisions in a social context. The definitions presented are founded in a well established social psychology study and build upon a small but behaviorally expressive set of social powers that enable agents to participate in a wide range of power games as found in human societies. These are the fundamentals of an agent's power assessment. We also introduced the basic mechanisms for an agent's decision making processes including social power influences aiming for behavioral believability which can be employed in diverse agent-based applications.

In future work we will first conceptualize the different effects on power dynamics associated with different power bases, the mechanism used to replicate power strategies utilization and its integration in a theory of mind reasoning. Finally this will be applied to a test scenario where agents can reason and influence other agents or humans in a believable way, considering the distinct social power forces and its effects.

References

1. Boella, G., Sauro, L., van der Torre, L.: Power and dependence relations in groups of agents. In: Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT 2004, pp. 246–252 (2004)
2. Brainov, S., Sandholm, T.: Power, dependence and stability in multiagent plans. In: Proceedings of the Sixteenth National Conference on Artificial Intelligence, AAAI 1999/IAAI 1999, pp. 11–16. American Association for Artificial Intelligence, Menlo Park (1999)
3. Carabelea, C., Boissier, O., Castelfranchi, C.: Using social power to enable agents to reason about being part of a group. In: Gleizes, M.-P., Omicini, A., Zambonelli, F. (eds.) ESAW 2004. LNCS (LNAI), vol. 3451, pp. 166–177. Springer, Heidelberg (2005)

4. Cartwright, D.: Power: A neglected variable in social psychology. Studies in Social Power 6, 1–14 (1959)
5. Castelfranchi, C.: Modelling social action for AI agents. Artificial Intelligence 103(1-2), 157–182 (1998)
6. Castelfranchi, C.: Founding agents' "autonomy" on dependence theory. In: ECAI, vol. 1, pp. 353–357 (2000)
7. French Jr., J., Raven, B.: The bases of social power. Studies in Social Power, 150–167 (1959)
8. Hayes, D., Hexmoor, H.: Social power as an exchangeable resource for distributed multi-agent systems. In: Proceedings of the International Symposium on Collaborative Technologies and Systems, pp. 278–281. IEEE Computer Society, Washington, DC (2006)
9. Hexmoor, H.: Absolute model of autonomy and power: Toward group effects. Connection Science 14(4), 323–333 (2002)
10. Keltner, D., Van Kleef, G., Chen, S., Kraus, M.: A reciprocal influence model of social power: Emerging principles and lines of inquiry. Advances in Experimental Social Psychology 40, 151–192 (2008)
11. Lewin, K.: Field theory in social science: selected theoretical papers (Cartwright, D. (ed.)). Harpers (1951)
12. López, F.: Social Power and Norms: Impact on agent behaviour. PhD thesis, University of Southampton (2003)
13. Marsella, S., Pynadath, D.: Modeling influence and theory of mind. In: Artificial Intelligence and the Simulation of Behavior. Joint Symposium on Virtual Social Agents, pp. 199–206 (2005)
14. Milgram, S.: Obedience to authority: an experimental view. Harper & Row (1974)
15. Pfeffer, J.: Power in organizations. Pitman, Marshfield (1981)
16. Prada, R., Paiva, A.: Teaming up humans with autonomous synthetic characters. Artificial Intelligence 173(1), 80–103 (2009)
17. Raven, B.H.: The bases of power: Origins and recent developments. In: Annual Meeting of the American Psychological Association (100th), Washington, DC (August 1992); Raven, B.H., Department of Psychology, UCLA, Los Angeles, CA 90024-1563
18. Sichman, J.S., Conte, R., Demazeau, Y., Castelfranchi, C.: A social reasoning mechanism based on dependence networks. In: Proceedings of 11th European Conference on Artificial Intelligence, pp. 188–192 (1994)
19. Ward, D., Hexmoor, H.: Deception as a means for power among collaborative agents. In: Int. WS on Collaborative Agents: Autonomous Agents for Collaborative Environments, pp. 61–66 (2003)