Strategic Disclosure of Opinions on a Social Network

(Extended Abstract)

Umberto Grandi
IRIT, University of Toulouse
France
Umberto.Grandi@irit.fr

Emiliano Lorini
IRIT-CNRS
University of Toulouse, France
Emiliano.Lorini@irit.fr

Laurent Perrussel
IRIT, University of Toulouse
France
Laurent.Perrussel@irit.fr

ABSTRACT

We study the strategic aspects of social influence in a society of agents linked by a trust network, introducing a new class of games called games of influence. A game of influence is an infinite repeated game with incomplete information in which, at each stage of interaction, an agent can make her opinions visible (public) or invisible (private) in order to influence other agents' opinions. The influence process is mediated by a trust network, as we assume that the opinion of a given agent is only affected by the opinions of those agents that she considers trustworthy (i.e., the agents in the trust network that are directly linked to her). Each agent is endowed with a goal, expressed in a suitable temporal language inspired from linear temporal logic (LTL). We show that games of influence provide a simple abstraction to explore the effects of the trust network structure on the agents' behaviour, by considering solution concepts from game-theory such as Nash equilibrium, weak dominance and winning strategies.

1. INTRODUCTION

At the micro-level, social influence can be conceived as a process where an agent forms her opinion on the basis of the opinions expressed by other agents in the society. Social influence depends on trust since an agent can be influenced by another agent, so that her opinions are affected by the expressed opinions of the other, only if she trusts her. At the macro-level, social influence is the basic mechanism driving the diffusion of opinions in human societies: certain agents in the society influence other agents in the society towards a given view, and these agents, in turn, influence other agents to acquire the same view, and so on. In other words, social influence can be seen as the driving force of opinion diffusion in human and human-like agent societies. This view is resonant of existing studies in social sciences and social psychology which emphasize the role of interpersonal processes in how people construe and form their perceptions, judgments, and impressions.

Recent work in multi-agent systems [3, 1] proposed a formal model of opinion diffusion that combined methods and techniques from social network analysis with methods and techniques from belief merging and judgment aggregation. The two models aim at studying how opinions of agents on a given set of issues evolve over time due to the influence of other agents in the population. The basic component of these models is the trust network, as it is assumed that the opinions of a certain agent are affected only by the opinions of the agents that she trusts (i.e., the agents in the trust network that are directly linked to her). Specifically, the opinions of a certain agent at a given time are the result of aggregating the opinions of the trustworthy agents at the previous time.

In this work we build on these models to look at social influence from a strategic perspective. We do so by introducing a new class of games, called games of influence. Games of influence provide a simple abstraction to explore the effects of the trust network structure on the agents' behaviour. Specifically, a game of influence is an infinite repeated game with incomplete information in which, at each stage of interaction, an agent can make her opinions visible (public) or invisible (private) to the other agents. Incompleteness of information is determined by the fact that an agent has uncertainty about the private opinions of the other agents, as she cannot see them. At each stage of the game, every agent is influenced by the *public* opinions of the agents she trusts (i.e., her neighbors in the trust network) and changes her opinions on the basis of the aggregation criterion she uses.

Following the representation of agents' motivations given in [2], in a game of influence each agent is identified with the goal that she wants to achieve. This goal is represented by a formula of a variant of linear temporal logic (LTL), in which we can express properties about agents' present and future opinions. For example, an agent might have the achievement goal that at some point in the future there will be consensus about a certain proposition p (i.e., either everybody has the opinion that p is true or everybody has the opinion that p is false), or the maintenance goal that two different agents will always the same opinion about p.

2. BASIC DEFINITIONS

Let $\mathcal{I} = \{p_1, \dots, p_m\}$ be a finite set of propositions or issues and let $\mathcal{N} = \{1, \dots, n\}$ be a finite set of individuals or agents. In this section we provide some of the basic definitions of our setting, starting from the modelling of private and public opinions of individuals in $\mathcal{N}$ about the issues in $\mathcal{I}$, presenting then a model of opinion diffusion based on aggregation, together with a suitable logical language to express temporal goals of the individuals, and finally giving a general definition of influence games.

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

2.1 Private and public opinions

Agents have opinions about all issues in $\mathcal{I}$ in the form of a propositional evaluations:

DEFINITION 1 (PRIVATE OPINION). *The private opinion of agent i is a function $B_i : \mathcal{I} \to \{1, 0\}$.*

We also assume that each agent has the possibility of declaring or hiding her private opinion on each of the issues.

DEFINITION 2 (VISIBILITY FUNCTION). *The visibility function of agent i is a map $V_i : \mathcal{I} \to \{1, 0\}$.*

By combining the private opinion with the visibility function of an agent we can build her public opinion as a three-valued function on the set of issues.

DEFINITION 3 (PUBLIC OPINION). *Let B_i be agent i's opinion and V_i her visibility function. The public opinion induced by B_i and V_i is a function $P_i : \mathcal{I} \to \{1, 0, ?\}$ s.t.:*

$$P_i(p) = \begin{cases} B_i(p) & \text{if } V_i(p) = 1 \\ ? & \text{if } V_i(p) = 0 \end{cases}$$

Observe that an agent can only hide or declare her opinion about a given issue, but is not allowed to lie.

2.2 Opinion diffusion through aggregation

First, we assume that individuals are connected by an *influence network* which we model as a directed graph $E \subseteq \mathcal{N} \times \mathcal{N}$. We interpret $(i, j) \in E$ as "agent j is influenced by agent i". Given a profile of public opinions and an influence network E, we model the process of opinion diffusion by means of an aggregation function, which shapes the private opinion of an agent by taking into consideration the public opinions of her influencers.

DEFINITION 4 (AGGREGATION PROCEDURE). *An aggregation procedure for agent i is a class of functions*

$$F_i : \mathcal{B} \times \mathcal{P}_J \longrightarrow \mathcal{B} \text{ for each } J \subseteq \mathcal{N} \setminus \{i\}$$

that maps agent i's individual opinion and the public opinions of a set of agents J to agent i's individual opinion.

2.3 A language for goals

We introduce a logical language based on a combination of simple version of multi-agent epistemic logic and linear temporal logic (LTL) that can be interpreted over histories of influenced opinions. In line with our framework, term *epistemic state* should be interpreted as *private opinion.* Goals in our perspective consists of targeting an epistemic state: typically ''agent i wants that agent j has private opinion φ in the future".

We call ELTL–I the following logic, from epistemic linear temporal logic of influence. Its language, denoted by $\mathcal{L}_{\mathsf{ELTL-I}}$, is defined by the following BNF:

$$\begin{array}{lll} \alpha & ::= & \mathsf{op}(i,p) \mid \mathsf{vis}(i,p) \mid \neg\alpha \mid \alpha_1 \wedge \alpha_2 \mid \mathsf{K}_i\alpha \\ \varphi & ::= & \alpha \mid \neg\varphi \mid \varphi_1 \wedge \varphi_2 \mid \mathsf{X}\varphi \mid \varphi_1 \mathsf{U} \varphi_2 \end{array}$$

where i ranges over $\mathcal{N}$ and p ranges over $\mathcal{I}$. $\mathsf{op}(i,p)$ has to be read "agent i's opinion is that p is true" while$\neg\mathsf{op}(i,p)$ has to be read "agent i's opinion is that p is true" (since we assume that agents have binary opinions). $\mathsf{vis}(i,p)$ has to be read "agent i's opinion about p is visible". Finally, $\mathsf{K}_i\alpha$ has to be read "agent i knows that α is true". $\mathsf{X}\varphi$ and U are the standard LTL operators 'next' and 'until'.

The operator K_i is rather peculiar, and should not be interpreted as a classical individual epistemic operator. It mixes public and private opinions of our model. Operator K_i reading is rather "agent i is uncertain about other agents private opinion as this opinion is not visible" and $\mathsf{K}_i\alpha$ stands for agent i knows α despite this uncertainty.

2.4 Influence games

We are now ready to give the following definition:

DEFINITION 5 (INFLUENCE GAME). *An influence game is a tuple $IG = (\mathcal{N}, \mathcal{I}, E, F_i, S_0, \gamma_1, \ldots, \gamma_n)$ where $\mathcal{N}$, $\mathcal{I}$, E and S_0 are, respectively, a set of agents, a set of issues, an influence network, and an initial state, F_i are aggregation procedures, and $\gamma_i \in \mathcal{L}_{\mathsf{ELTL-I}}$ is agent i's goal.*

Strategies, best-responses and Nash equilibria, can be defined according to the relevant literature.

3. SUMMARY OF RESULTS

We inquire into the multiple aspects of the relation between the structure of the influence network, and the existence of well-known game-theoretic solution concepts. such as Nash equilibrium and weak dominance. For instance, we show that if the trust network is fully connected and every agent wants to reach a consensus about a certain proposition p, then there always exists a least one Nash equilibrium.

Moreover, we study how the relative position of an agent in the trust network determines her influencing power, that is, her capacity to influence opinions of other agents, no matter what the others decide to do (which corresponds to the concept of uniform strategy).

We are also able to show that model checking for our language, as well as the problem of checking whether a given profile is a Nash equilibrium, is in PSPACE, hence no harder than the linear temporal logic on which our language is based.

Acknowledgments

This work was partially supported by the project "Social Choice on Networks" of Labex CIMI ANR-11-LABX-0040-CIMI within the program ANR-11-IDEX-0002-02.

REFERENCES

[1] U. Grandi, E. Lorini, and L. Perrussel. Propositional opinion diffusion. In *Proceedings of the 14th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-2015)*, 2015.

[2] J. Gutierrez, P. Harrenstein, and M. Wooldridge. Iterated boolean games. *Information and Computation*, 242:53–79, 2015.

[3] N. Schwind, K. Inoue, G. Bourgne, S. Konieczny, and P. Marquis. Belief revision games. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI-2015)*, 2015.