Strategic Disclosure of Opinions on a Social Network

Umberto Grandi, Emiliano Lorini, Arianna Novaro, Laurent Perrussel
IRIT, CNRS and University of Toulouse

Agents, opinions, network and unanimous aggregation

Agents in $\mathcal{N}$ have opinions over the issues in $\mathcal{I}$.

- $B_i : \mathcal{I} \to \{0,1\}$ is the opinion of agent i
  $B_i(p) = 1$ iff agent i believes that p
- $V_i : \mathcal{I} \to \{0,1\}$ is the influence power (visibility) of agent i
  $V_i(p) = 1$ iff agent i makes her opinion on p "visible"

A state consists of a specification of B_i and V_i for all $i \in \mathcal{N}$.

Influence network
A directed irreflexive graph $E \subseteq N \times N$, such that $(i,j) \in E$ iff agent i influences agent j.

We let $Inf(i) = \{k \in \mathcal{N} \mid (k,i) \in E\}$ be i's influencers in E.

Unanimous aggregation F
Agent i updates her opinion on p iff all visible opinions of i's influencers are unanimous (else, she keeps her opinion).

Strategic actions and opinion update

Actions available to players are the following:
$\mathcal{A} = \{(\mathsf{reveal}(J), \mathsf{hide}(J')) \mid J, J' \subseteq \mathcal{I} \text{ and } J \cap J' = \emptyset\}$

Opinion update process
1. Agents decide how to use their influence power on issues
2. Agents update opinions via (unanimous) aggregation

Given $\mathcal{I}$, $\mathcal{N}$, E and F_i for $i \in \mathcal{N}$, an history is an infinite sequence of states $H = (H_0, H_1, \dots)$ such that for all $t \in \mathbb{N}$ there exists a joint action in $\mathcal{A}^{|\mathcal{N}|}$ leading from H_t to H_{t+1}.

Example: the two actions $\boldsymbol{a}_0 = (\mathsf{skip}, \mathsf{skip}, \mathsf{reveal}(p))$ and $\boldsymbol{a}_1 = (\mathsf{skip}, \mathsf{hide}(p), \mathsf{skip})$ generate the following history:

$$\underset{H_0}{((0,1,1),(1,1,0))} \xrightarrow[\boldsymbol{a}_0]{} \underset{H_1}{((1,1,1),(1,1,1))} \xrightarrow[\boldsymbol{a}_1]{} \underset{H_2}{((1,1,1),(1,0,1))}$$

Individual goals

We express agents' goals in Linear Temporal Logic (LTL):

$$\varphi \quad ::= \quad \mathsf{op}(i,p) \mid \mathsf{vis}(i,p) \mid \neg\varphi \mid \varphi \wedge \varphi \mid \bigcirc\varphi \mid \varphi\mathcal{U}\varphi$$

where $\bigcirc$ is the 'next' temporal operator and $\mathcal{U}$ indicates 'until'. 'Eventually' ($\Diamond$) and 'henceforth' ($\Box$) are definable from them. We interpret goals over histories.

Examples of influence goals

$$\mathsf{Positive-Cons}(C,J) := \bigwedge_{i \in C} \bigwedge_{p \in J} \mathsf{op}(i,p) \qquad [\mathsf{pcons}]$$

$$\mathsf{Consensus}(C,J) := \Diamond\Box(\mathsf{pcons}(C,J) \vee \mathsf{ncons}(C,J))$$

$$\mathsf{Influence}(i,C,J) := \Diamond\Box \bigwedge_{p \in J} \big((\mathsf{op}(i,p) \to \bigcirc\mathsf{pcons}(C,p)) \wedge (\neg\mathsf{op}(i,p) \to \bigcirc\mathsf{ncons}(C,p))\big).$$

Games of influence

An influence game is a tuple $IG = \langle \mathcal{N}, \mathcal{I}, E, \{F_i\}_{i \in \mathcal{N}}, S_0, \{\gamma_i\}_{i \in \mathcal{N}} \rangle$, where:

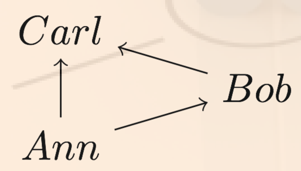
- $\mathcal{N}$ is the set of agents and $\mathcal{I}$ is the set of issues
- E is the network and F_i is the (unanimous) aggregator for agent i
- S_0 is the initial state and γ_i is the individual goal of agent i

Strategies of the agents
Memory-less: associate an action to each state
Perfect-recall: associate an action to each finite sequence of states

We are interested in studying the following classical game-theoretic solution concepts:

- Winning strategy: the agent gets her goal, regardless of the strategies of the others
- Weakly dominant strategy: the agent is not better off by playing a different strategy
- Nash equilibrium: no agent gains by unilaterally deviating from her current strategy

Example: Let $B_{Ann}(p) = 1$ and $B_{Bob}(p) = B_{Carl}(p) = 0$ on the following network:



If $\gamma_{Ann} = \Diamond\Box\mathsf{op}(Carl, p)$, Ann's winning memory-less strategy is to play $\mathsf{reveal}(p)$ in all states. Bob will be influenced to believe p at the second stage in the process, and Carl at the third — since his influencers are unanimous even if Bob plays $\mathsf{hide}(p)$.

Game-theoretic results

Weak-dominance
The memory-less strategy associating action **reveal** to all states is not necessarily weakly dominant for goals of type **Influence**.

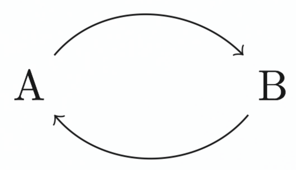
Network: A $\longrightarrow$ B $\longrightarrow$ C $\longrightarrow$ D

	A	B	C	D
$S_0(p)$:	0	1	0	1

- $\gamma_B = \mathsf{Influence}(B, \{D\}, \{p\})$
- B's strategy is to use her influence power over p in all states
- C uses her influence power on p unless A, B and C agree on p

$\Rightarrow$ What if B does *not* use her influence power over p in S_0?

Nash Equilibrium
For a cycle with two agents, if $\gamma_A = \gamma_B = \mathsf{Consensus}(\{A,B\}, C)$ where $C \subseteq \mathcal{I}$, then there is a Nash equilibrium for any S_0.



- Agents A and B have to *anti*-coordinate over the issues
- Given agent A's strategy, build the "inverse" of agent B's strategy

Computational complexity problems

M-Nash
INPUT: an influence game IG, a strategy profile
QUESTION: is this profile a Nash Equilibrium of IG?

E-Nash
INPUT: an influence game IG
QUESTION: is there a Nash Equilibrium of IG?

U-Nash
INPUT: an influence game IG
QUESTION: is there a unique Nash Equilibrium of IG?

E-Winning
INPUT: an influence game IG, an agent i
QUESTION: does i have a winning strategy in IG?

Logics for multi-agent systems

The language of Alternating-time Temporal Logic (ATL):

$$\varphi \;::=\; q \mid \neg\varphi \mid \varphi \wedge \varphi \mid \langle\!\langle C \rangle\!\rangle \bigcirc \varphi \mid \langle\!\langle C \rangle\!\rangle (\varphi\mathcal{U}\varphi)$$

The language of Graded Strategy Logic ($\mathsf{G-SL}$):

$$\varphi \;::=\; q \mid \neg\varphi \mid \varphi \wedge \varphi \mid \bigcirc\varphi \mid \varphi\mathcal{U}\varphi \mid \langle\!\langle x_1, \dots, x_\ell \rangle\!\rangle^{\geq k}\varphi \mid [i \mapsto x]\varphi$$

Both logics are interpreted over concurrent game structures:

Concurrent Game Structure
A tuple $\mathbf{G} = (\mathcal{W}, \mathcal{M}, R, T, Val)$ where $\mathcal{W}$ are worlds, $\mathcal{M}$ are moves, $R : \mathcal{N} \times \mathcal{W} \longrightarrow 2^{\mathcal{M}} \backslash \emptyset$ defines available moves for an agent at a world, $T : \mathcal{W} \times \mathcal{M}^n \longrightarrow \mathcal{W}$ maps a world and a move profile to a successor world, and Val is a valuation.

Complexity results for memory-less strategies

Theorem
M-Nash, E-Nash and U-Nash are in PSPACE for memory-less strategies (and unanimous rule).

Proof idea. An action $a = (\mathsf{reveal}(J), \mathsf{hide}(J'))$ can be encoded as a LTL formula $\beta_i(a)$ as follows:

$$\beta_i(a) = \bigwedge_{p \in J} \bigcirc\mathsf{vis}(i,p) \wedge \bigwedge_{q \in J'} \bigcirc\neg\mathsf{vis}(i,q).$$

A memory-less strategy Q_i can also be encoded in a formula $\tau_i(Q_i)$ (for S a state and being $\alpha(S)$ its specification):

$$\tau_i(Q_i) = \bigwedge_{S} \alpha(S) \to \beta_i(Q_i(S)).$$

We can thus write as a formula the profile of agents' strategies.

Let $unan(i,p)$ be the following formula:

$$\bigcirc\,\mathsf{op}(i,p) \leftrightarrow \Big(\Big[\bigwedge_{j \in Inf(i)} \bigcirc\neg\mathsf{vis}(j,p) \wedge \mathsf{op}(i,p)\Big] \vee$$
$$\Big[\bigvee_{j \in Inf(i)} \bigcirc\,\mathsf{vis}(j,p) \wedge \bigwedge_{j \in Inf(i)} (\bigcirc\,\mathsf{vis}(j,p) \to \mathsf{op}(j,p))\Big] \vee$$
$$\Big[\bigvee_{j,z \in Inf(i):} (\bigcirc\,\mathsf{vis}(j,p) \wedge \bigcirc\,\mathsf{vis}(z,p) \wedge$$
$$\mathsf{op}(j,p) \wedge \neg\mathsf{op}(z,p)) \wedge \mathsf{op}(i,p)\Big]\Big).$$

An analogous formula for $\neg p$ allows us to express the unanimous aggregation rule.

Having a translation of strategies and unanimity, we can then use the LTL validity-checking problem (known to be in PSPACE).

Complexity results for perfect-recall strategies

Theorem
E-Winning is in EXPTIME for perfect-recall strategies (restricted goals and unanimous rule).

Proof idea. Construct a corresponding CGS from IG and use ATL model checking for formula $\langle\!\langle i \rangle\!\rangle\gamma_i$.

Theorem
E-Nash and U-Nash are in 3EXPTIME for perfect-recall strategies (and unanimous rule).

Proof idea. Construct a corresponding CGS from IG, translate the problems as $\mathsf{G-SL}$ formulas and then use the model-checking problem of $\mathsf{G-SL}$.

This poster in a nutshell

We study strategic aspects of opinion diffusion on a network:

- Agents can use or retain their influence
- Opinions are updated with a unanimous aggregator
- Goals are expressed as temporal logic formulas

We focus on game-theoretic concepts: winning strategy, weak dominance and Nash equilibrium. We show that:

- The interplay between the network structure, goals and solution concepts is non-trivial
- Solution concepts with memory-less strategies can be expressed as LTL satisfiability (PSPACE)
- Perfect-recall strategies require ATL or Graded Strategy Logic (EXPTIME/3EXPTIME)

Next step: full-blown strategic opinion diffusion with lies.

AAMAS-2017, São Paolo, 8–12 May 2017

{umberto.grandi, emiliano.lorini, arianna.novaro, laurent.perrussel}@irit.fr
Acknowledgments: Partially supported by LabEx CIMI project "Social choice on networks"