A Logic for Social Influence through Communication

Zoé Christoff

Institute for Logic, Language and Computation
University of Amsterdam

11th European Workshop on Multi-Agent Systems (EUMAS)
Logical Aspects of Multi-Agent System (LAMAS)
Toulouse, December 13 2013
Outline

1) Seligman, Girard & Liu (2011, 2014)
   ▶ social network
   ▶ peer pressure effects, influence inbetween “friends”
Introduction: the two approaches to combine
A two dimensional social network plausibility framework
Social influence through communication
Further research

Outline

1) Seligman, Girard & Liu (2011, 2014)
   ▶ social network
   ▶ peer pressure effects, influence inbetween “friends”

2) Baltag & Smets (2009, 2013)
   ▶ plausibility
   ▶ effects of group members sharing information with the rest of the group
Outline

1) Seligman, Girard & Liu (2011, 2014)
   - social network
   - peer pressure effects, influence inbetween “friends”

2) Baltag & Smets (2009, 2013)
   - plausibility
   - effects of group members sharing information with the rest of the group

3) Aim: a unified social network plausibility framework
   - model social influence on beliefs through communication among agents in a social network
   - define some particular communication protocols (in the new framework) inspired by 2) to represent some level of influence as defined in 1)
1) Social influence à la Girard, Liu & Seligman

The framework
Static hybrid logic to represent who is friend with whom and who believes what + an (external) influence operator

The main ideas
- Agents are influenced by their friends and only by their friends.
- Simple “peer pressure principle”: I tend to align with my friends.
- “Being influenced” is defined as “aligning my beliefs to the ones of my friends”.
- No communication is (at least explicitly) involved. (transparency?)
Friends network

Social network frame:
Friends network

Social network frame:

3 possible belief states (with respect to \( p \))

- \( Bp \)
- \( B \neg p \)
- \( Up := \neg Bp \text{ and } \neg B \neg p \)
Belief revision induced by (direct) social influence

1) Strong influence
When all of my friends believe that $p$, I (successfully) revise with $p$. When all of my friends believe that $\neg p$, I (successfully) revise with $\neg p$. 

\[
B \neg p \quad \cdots \quad B p
\]

\[
\cdots
\]

\[
B p \quad B \neg p
\]
Belief revision induced by (direct) social influence

1) Strong influence

When all of my friends believe that \( p \), I (successfully) \textit{revise} with \( p \). When all of my friends believe that \( \neg p \), I (successfully) \textit{revise} with \( \neg p \).
Belief revision induced by (direct) social influence

1) Strong influence
When all of my friends believe that \( p \), I (successfully) revise with \( p \). When all of my friends believe that \( \neg p \), I (successfully) revise with \( \neg p \).
Introduction: the two approaches to combine
A two dimensional social network plausibility framework
Social influence through communication
Further research

1) Social influence à la Girard, Liu & Seligman
2) Communication protocols à la Baltag & Smets
Comparison

Belief contraction induced by social influence

2) Weak influence
None of my friends supports my belief in $p$ and some believe that $\neg p$.
I (successfully) contract it.
(And similarly for $\neg p$)

\[
\begin{array}{c}
B\neg p \\
\vdots \\
Bp \\
\vdots \\
Up \\
\vdots \\
B\neg p
\end{array}
\]
Belief contraction induced by social influence

2) Weak influence

None of my friends supports my belief in $p$ and some believe that $\neg p$. I (successfully) contract it.

(And similarly for $\neg p$)
Stabilization

- Stable state: applying the social influence operator doesn’t change the state of any agent.
- Stabilization: some configurations will reach a stable state after a finite number of applications of the influence operator (see example of weak influence above) and some won’t (see example of strong influence).
- Sufficient condition for stability: all friends are in the same state.
2) Communication protocols à la Baltag & Smets

The framework
DEL type: plausibility modeling of (several) doxastic attitudes + communication events

The main ideas

- Agents communicate via public announcements.
- Assuming that they trust each other enough, agents all revise their beliefs with each of the announced formula, sequentially.
- In this sense, each announcement influences everybody (else) into belief revision.
Plausibility model
Plausibility model

\[ q \to a, b, d \to p \]
\[ v \to c \to w \]
Plausibility model
Reaching a stable state of agreement

How to communicate?

- Agents speak in turn (given expertise rank).
- An agent announces all and only (non-equivalent) sentences that she believes (exhaustivity + honesty).
- After a finite number of announcements (and corresponding revisions), everybody holds the same beliefs.
- This is a stable state: nothing which could be announced by any agent would change anything anymore.
Reaching a stable state of agreement

How to communicate?

- Agents speak in turn (given expertise rank).
- An agent announces all and only (non-equivalent) sentences that she believes (exhaustivity + honesty).
- After a finite number of announcements (and corresponding revisions), everybody holds the same beliefs.
- This is a stable state: nothing which could be announced by any agent would change anything anymore.

Lexicographic belief merge protocol

\[ \rho_a := \prod \{ \uparrow \phi : \|\phi\| \subseteq S \text{ such that } M, w \models B_a \phi \} \]

\[ \rho_b := \prod \{ \uparrow \phi : \|\phi\| \subseteq S \text{ such that } M[\rho_a], w \models B_b \phi \} \]

etc for all \( c \in \mathcal{A} \)

where \( \prod \) is a sequential composition operator and \( M[\rho_a] \) is the new model after joint revision with each formula announced by \( a \).
Big picture

Common features

- Agents are influenced into revising their beliefs to make them closer to the ones of (some) others.
- A global agreement state is stable (both under honest communication and under social conformity pressure).

From 1)

- **Social network**
- Synchronous
- **Over friends only**
- Equal power (among friends)
- Direct
- **Tools**: nominals, @, F

From 2)

- **Plausibility**
- Sequential
- Over everybody
- Ranking
- **Via communication**
- **Tools**: B, ↑, ↑↑
3) A social network plausibility framework

plausibility model:

\[ V \quad \text{and} \quad W \]

\[ a, b, d \quad \text{and} \quad c \]
3) A social network plausibility framework

Social network plausibility model:
Social network plausibility model

\[ \mathcal{M} = (S, \mathcal{A}, \leq_{a \in \mathcal{A}}, \| \cdot \|, s_0, \simeq_{s \in S}) \]

- \( S \) is a (finite) set of possible states.
- \( \mathcal{A} \) is a (finite) set of agents.
- \( \leq_{a \subseteq S \times S \text{ is a locally connected preorder, interpreted as the subjective plausibility relation of agent } a, \text{ for each } a \in \mathcal{A}} \]
- \( s_0 \in S \) is a designated state, interpreted as the actual state
- \( \simeq_s \subseteq \mathcal{A} \times \mathcal{A} \) is an irreflexive and symmetric relation, interpreted as friendship, for each state \( s \in S \)
- \( \| \cdot \| : \Phi \cup \mathcal{N} \to \mathcal{P}(S \times \mathcal{A}) \) is a valuation, assigning:
  - a set \( \| p \| \subseteq S \times \mathcal{A} \) to every element \( p \) of some given set \( \Phi \) of “atomic propositions”
  - a set \( \| n \| = S \times \{ a \} \) for some \( a \in \mathcal{A} \) to every element \( n \) of some given set \( \mathcal{N} \) of “nominals”.
Syntax

$$\phi ::= p \mid n \mid \neg \phi \mid \phi \land \phi \mid F\phi \mid \Box n \phi \mid B\phi$$

where \( p \) belongs to a set of atomic propositions \( \Phi \) and \( n \) to a set of nominals \( N \).
Inherited indexicality

Formulas evaluated both at a state $w \in S$ and at an agent $a \in A$.

- $p$ : “I am blonde.”
- $BFp$ : “I believe that all my friends are blonde.”
- $FBp$ : “All of my friends believe that they are blonde”.
Semantics clauses

- If $M, w, a \models p$ iff $\langle w, a \rangle \in \lVert p \rVert$
- If $M, w, a \models n$ iff $\langle w, a \rangle \in \lVert n \rVert$ iff $a = n$
- If $M, w, a \models \neg \phi$ iff $M, w, a \not\models \phi$
- If $M, w, a \models \phi \land \psi$ iff $M, w, a \models \phi$ and $M, w, a \models \psi$
- If $M, w, a \models F\phi$ iff $M, w, b \models \phi$ for all $b$ such that $a \preceq b$
- If $M, w, a \models \Box b \phi$ iff $M, w, b \models \phi$
- If $M, w, a \models B\phi$ iff $M, v, a \models \phi$ for all $v \in S$ such that $v \in \text{best}_a w(a)$

Notation:
- $n$ the unique agent at which the nominal $n$ holds
- $s(a)$ the comparability class of state $s$ relative to agent $a$: $t \in s(a)$ iff $s \preceq_a t$ or $t \preceq_a s$
- $\text{best}_a s(a)$ the most plausible states in $s(a)$ according to $a$: $\text{best}_a s(a) := \{ s \in s(a) : t \preceq_a s$ for all $t \in s(a) \}$
Example

- $M, v, c \models p$
- $M, v, a \models Fp$
- $M, v, a \models \langle F \rangle b$

- $M, w, d \models FBp$
- $M, w, a \models BFp$
Example

\[
\begin{align*}
M, v, c &\models p \\
M, v, a &\models Fp \\
M, v, a &\models \langle F \rangle b
\end{align*}
\]

\[
\begin{align*}
M, w, d &\models FBp \\
M, w, a &\models BFp \\
M, w, c &\models B\@ b \langle F \rangle d
\end{align*}
\]
Example

- $M, v, c \models p$
- $M, v, a \models Fp$
- $M, v, a \models \langle F \rangle b$
- $M, w, d \models FBp$
- $M, w, a \models BFp$
- $M, w, c \models B@b\langle F \rangle d$
Example

\[
\neg \bigcirc b \langle F \rangle d
\]

\[
\begin{align*}
M, v, c & \models p \\
M, v, a & \models Fp \\
M, v, a & \models \langle F \rangle b \\
M, w, d & \models FBp \\
M, w, a & \models BFp \\
M, w, c & \models B \bigcirc b \langle F \rangle d \\
\end{align*}
\]
Influence dynamics

Simplifying assumptions

- agents speak in turn (rank)
- only friends communicate
- agents revise with (all) sentences announced (trust)
Revision operator

Joint radical upgrade $\uparrow \phi$

- “Promote” all the $\|\phi\|$-worlds so that they become more plausible than all $\neg\|\phi\|$-worlds (in the same information cell), keeping everything else the same:
Revision operator

Joint radical upgrade $\uparrow \phi$

- “Promote” all the $\|\phi\|$-worlds so that they become more plausible than all $\neg\|\phi\|$-worlds (in the same information cell), keeping everything else the same:

- $\uparrow \phi$ is a model transformer which takes as input any model $\mathcal{M} = (S, A, \leq_{a \in A}, \|\cdot\|, s_0, \preceq s \in S)$ and outputs a new model $\mathcal{M}' = (S, A, \leq'_{a \in A}, \|\cdot\|, s_0, \preceq s \in S)$ such that:

  $s \leq'_{a} t$ iff either $(s, t \notin \|\phi\| \text{ and } s \leq_{a} t)$ or $(s, t \in \|\phi\| \text{ and } s \leq_{a} t)$ or $(t \in s(a) \text{ and } s \notin \|\phi\| \text{ and } t \in \|\phi\|)$. 


Belief merge

Baltag & Smets’ lexicographic belief merge protocol

$$\rho_a := \prod \{\uparrow \phi : \|\phi\| \subseteq S \text{ such that } \mathcal{M}, w \models B_a \phi\}$$

$$\rho_b := \prod \{\uparrow \phi : \|\phi\| \subseteq S \text{ such that } \mathcal{M}_{[\rho_a]}, w \models B_b \phi\}$$

etc for all $c \in \mathcal{A}$

where $\prod$ is a sequential composition operator and $\mathcal{M}_{[\rho_a]}$ is the new model after joint revision with each formula announced by $a$. 
Belief merge

Indexical lexicographic belief merge protocol

\[ \rho_a := \prod \{ \uparrow \otimes a \phi : \| \phi \| \subseteq S \times A \text{ such that } M, w, a \models B\phi \} \]

\[ \rho_b := \prod \{ \uparrow \otimes b \phi : \| \phi \| \subseteq S \times A \text{ such that } M[a], w, b \models B\phi \} \]

etc for all \( c \in A \)

where \( \prod \) is a sequential composition operator and \( M[a] \) is the new model after joint revision with each formula announced by \( a \).
A central friend

Assumptions

- $a$ is other agents’ only friend.
- $a$ speaks first.

One-to-others unilateral strong influence protocol

One step version of the indexical lexicographic belief merge protocol:

$$\rho_a := \prod \{ \uparrow \Theta_a \phi : \|\phi\| \subseteq S \times A \text{ such that } M, w, a \models B\phi \}$$
Everybody is friends with everybody else

Assumption

- Connectedness

Others-to-one unilateral strong influence protocol

\[
\rho_b := \prod \{ \uparrow \otimes_b B\phi : \|\phi\| \subseteq S \times A \text{ such that } M, w, b \models B\phi \}
\]

\[
\rho_c := \prod \{ \uparrow \otimes_c B\phi : \|\phi\| \subseteq S \times A \text{ such that } M, w, c \models B\phi \}
\]

etc, for all \( d \in A \) such that \( M, w, d \models \langle F \rangle a \)

\[
\rho_a := \prod \{ \uparrow \otimes_a \phi \text{ iff } M_{[\rho_b;\rho_c;\ldots]}, w, a \models BFB\phi \}
\]

where \( M_{[\rho_b;\rho_c;\ldots]} \) is the model resulting from the successive revisions (by all friends) with each of the formulas announced by each of them.
Summary

- Social network plausibility framework with communication events
- Indexical protocol to merge beliefs
- Unilateral strong influence *one-to-all-the-others* protocol
- Unilateral strong influence *all-the-others-to-one* protocol
To do next

- Private (and synchronous?) communication: *friends to friends* influence (level of privacy to determine)
- Different doxastic attitudes (conditional belief, strong belief, safe belief) + different levels of trust (dynamic attitudes) corresponding to different types of revision (minimal revision, update).
- Consider how to merge (as quickly as possible) knowledge and/or belief within a social network.
Thank you

zoe.christoff@gmail.com
References


