Social Influence Within Epistemic Logic

Vitaliy Dolgorukov
vdolgorukov@hse.ru

«Formal Philosophy», December 16, 2016

1 $\mathcal{L}_{EL} ::= p \mid \varphi \mid \neg\varphi \mid \varphi \wedge \psi \mid K_i\varphi \mid B_i\varphi$

2 $\mathcal{M} = (\mathcal{A}, W, \{\sim_i\}_{i\in\mathcal{A}}, \{\preceq_i\}_{i\in\mathcal{A}}, V)$
- $\mathcal{A}$ *is a finite non-empty set of agents*
- W *is a finite non-empty set of possible worlds (situations)*
- $\sim_i \subseteq W \times W$ *is an equivalence relation for each agent* $i \in \mathcal{A}$
- $\preceq_i \subseteq W \times W$ *– reflexive and transitive relation on* W *(s.t.* $\forall w' \forall w''(w' \sim_i w'' \equiv (w' \preceq_i w'' \vee w'' \preceq_i w'))$*);*
- $V : Var\mathcal{L} \mapsto \mathcal{P}(W)$*, where* Var *is a set of propositional variables*

3 $\mathcal{L}_{SI} ::= p \mid \varphi \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid K_i\varphi \mid B_i\varphi \mid SB_i\varphi \mid [\mathcal{I}, \circ\varphi]\psi$*, where*
- $\circ ::= \uparrow_i^n, \Uparrow_i^n, \uparrow_i^{sk}, \Uparrow_i^{sk}$
- p *ranges over* Var *(set of propositional variables)*
- $\mathcal{A}$ *is a finite set of agents,* $i \in \mathcal{A}$

4 $\mathcal{M}^{\mathcal{I}} = (\mathcal{A}, W, \{\sim_i\}_{i\in\mathcal{A}}, \{\preceq_i\}_{i\in\mathcal{A}}, V, Infl)$
- $\mathcal{M} = (\mathcal{A}, W, \{\sim_i\}_{i\in\mathcal{A}}, \{\preceq_i\}_{i\in\mathcal{A}}, V)$ *— epistemic plausibility model*
- $Infl : \mathcal{A} \times \mathcal{L}_{EL} \mapsto \Delta(\mathcal{A})$*, where* $\Delta(\mathcal{A})$ *is a probability distribution over* $\mathcal{A}$

5
1. $[\varphi]_{\mathcal{M}} := \{w \in W \mid \mathcal{M}, w \vDash \varphi\}$
2. $[w]_i := \{w' \in W \mid w \sim_i w'\}$
3. $[[\varphi]]_i^w := [\varphi]_{\mathcal{M}} \cap [w]_i$
4. $max_{\preceq_i}(X) := \{w \in X \mid \forall w' \in X : w' \preceq_i w\}$*, where* $X \subseteq W$
5. $best_i(\varphi, w) := max_{\preceq_i}([[\varphi]]_i^w)$

6 $Infl$ *is a function which maps a pair of agent* $i \in \mathcal{A}$ *and formula* $\varphi \in \mathcal{L}^{Pl}$ *into* $\ll_i^{\varphi} \subseteq \mathcal{A} \times \mathcal{A}$ *. The relation* $a \ll_i^{\varphi} b$ *interpreted as "for* i *b's opinion on* φ *is at least as important as a's ".*

7 $\mathcal{M}, w \vDash SB_i\varphi$ *iff* $\forall j \in max(\ll_i^{\varphi}) : \mathcal{M}, w \vDash B_iB_j\varphi$*,*
where $max(\ll_i^{\varphi}) := \{j \in \mathcal{A} \mid \forall k \in \mathcal{A} : k \preccurlyeq_i^{\varphi} j\}$

8 *Truth for* $\mathcal{M}^{\mathcal{I}}$
$\mathcal{M}^{\mathcal{I}}, w \vDash p$ *iff* $w \in V(p)$
$\mathcal{M}^{\mathcal{I}}, w \vDash \neg\varphi$ *iff* $\mathcal{M}^{\mathcal{I}}, w \nvDash \varphi$
$\mathcal{M}^{\mathcal{I}}, w \vDash \varphi \wedge \psi$ *iff* $\mathcal{M}^{\mathcal{I}}, w \vDash \varphi$ *and* $\mathcal{M}^{\mathcal{I}}, w \vDash \psi$
$\mathcal{M}^{\mathcal{I}}, w \vDash K_i\varphi$ *iff* $\forall w' \in [w]_i : \mathcal{M}^{\mathcal{I}}, w' \vDash \varphi$
$\mathcal{M}^{\mathcal{I}}, w \vDash B_i\varphi$ *iff* $\forall w' \in max_{\preceq_i}([w]_i) : \mathcal{M}^{\mathcal{I}}, w' \vDash \varphi$

9 $\mathcal{A}_a^w(\varphi) := \{i \in \mathcal{A} \mid \mathcal{M}, w \vDash B_aB_i\varphi\}$

10 $\mathcal{A}_a^w(?\varphi) := \{i \in \mathcal{A} \mid \mathcal{M}, w \vDash B_a(\neg B_i\varphi \wedge \neg B_i\neg\varphi)\}$

11 *Social belief*
$\mathcal{M}^{\mathcal{I}}, w \vDash SB_a\varphi$ *iff*
1) $\sum\limits_{i\in\mathcal{A}_a^w(\varphi)} Infl(i,a) > \sum\limits_{i\in\mathcal{A}_a^w(\neg\varphi)} Infl(i,a)$ *and*
2) $\sum\limits_{i\in\mathcal{A}_a^w(\varphi)} Infl(i,a) + \sum\limits_{i\in\mathcal{A}_a^w(\neg\varphi)} Infl(i,a) > \sum\limits_{i\in\mathcal{A}_a^w(?\varphi)} Infl(i,a)$

12
1. $\nvDash B_a\varphi \to SB_a\varphi$
2. $\nvDash SB_a\varphi \to B_a\varphi$
3. $\vDash SB_a\varphi \to K_a(SB_a\varphi)$
4. $\vDash \neg SB_a\varphi \to K_a(\neg SB_a\varphi)$
5. $\nvDash (SB_a\varphi \wedge SB_a(\varphi \to \psi)) \to SB_a\psi$

13 *Set of credible agents (from agent a's point of view)*
$cred_a^w(X) := \{i \in X \mid \neg\exists\varphi : \mathcal{M}, w \vDash K_a\varphi \wedge K_aB_i\neg\varphi\}$, *where* $X \subseteq \mathcal{A}$

14 *Credible social influence*
$$Infl^*(j,i) = \frac{Infl(j,i)}{\sum\limits_{i\in cred_a^w(\mathcal{A})} Infl(j,i)}$$

15 *Credible social belief*
$\mathcal{M}^{\mathcal{I}}, w \vDash SB_a^*\varphi$ *iff*
1) $\sum\limits_{i\in cred_a^w(\mathcal{A}_a^w(\varphi))} Infl^*(i,a) > \sum\limits_{i\in cred_a^w(\mathcal{A}_a^w(\neg\varphi))} Infl^*(i,a)$ *and*
2) $\sum\limits_{i\in cred_a^w(\mathcal{A}_a^w(\varphi))} Infl^*(i,a) + \sum\limits_{i\in cred_a^w(\mathcal{A}_a^w(\neg\varphi))} Infl^*(i,a) > \sum\limits_{i\in cred_a^w(\mathcal{A}_a^w(?\varphi))} Infl^*(i,a)$

16 *If* $\mathcal{M} = (\mathcal{A}, W, \{\sim_i\}_{i\in\mathcal{A}}, \{\preceq_i\}_{i\in\mathcal{A}}, V)$ *then* $\mathcal{M}^{\uparrow a\varphi} = (\mathcal{A}^{\uparrow a\varphi}, W^{\uparrow a\varphi}, \{\sim_i^{\uparrow a\varphi}\}_{i\in\mathcal{A}}, \{\preceq_i^{\uparrow a\varphi}\}_{i\in\mathcal{A}}, V^{\uparrow a\varphi})$, *where*
- $\mathcal{A}^{\uparrow a\varphi} = \mathcal{A}$
- $W^{\uparrow a\varphi} = W$
- $\{\sim_i^{\uparrow a\varphi}\}_{i\in\mathcal{A}} = \{\sim_i\}_{i\in\mathcal{A}}$
- $V^{\uparrow a\varphi} = V$
- $\preceq_i^{\uparrow a\varphi} = \preceq_i$, *for* $i \neq a$
- *For a* $\preceq_a^{\uparrow a\varphi}$ *is the smallest relation satisfying*
  - $\forall x\forall y((x \in best_a(\varphi, w) \wedge y \in [w]_a) \to y \preceq_a^{\uparrow a\varphi} x)$
  - $\forall x\forall y((x, y \in [w]_a - best_a(\varphi, w) \wedge x \preceq_a y) \to x \preceq_a^{\uparrow a\varphi} y)$

17 *If* $\mathcal{M} = (\mathcal{A}, W, \{\sim_i\}_{i\in\mathcal{A}}, \{\preceq_i\}_{i\in\mathcal{A}}, V)$ *then* $\mathcal{M}^{\Uparrow a\varphi} = (\mathcal{A}^{\Uparrow a\varphi}, W^{\Uparrow a\varphi}, \{\sim_i^{\Uparrow a\varphi}\}_{i\in\mathcal{A}}, \{\preceq_i^{\Uparrow a\varphi}\}_{i\in\mathcal{A}}, V^{\Uparrow a\varphi})$, *where*
- $\mathcal{A}^{\Uparrow a\varphi} = \mathcal{A}$
- $W^{\Uparrow a\varphi} = W$
- $\{\sim_i^{\Uparrow a\varphi}\}_{i\in\mathcal{A}} = \{\sim_i\}_{i\in\mathcal{A}}$
- $V^{\Uparrow a\varphi} = V$
- $\preceq_i^{\Uparrow a\varphi} = \preceq_i$, *for* $i \neq a$
- *For a* $\preceq_a^{\Uparrow a\varphi}$ *is the smallest relation satisfying*
  - $\forall x\forall y((x \in [[\varphi]]_a^w \wedge y \in [[\neg\varphi]]_a^w) \to y \prec_a^{\Uparrow a\varphi} x)$[1]
  - $\forall x\forall y((x, y \in [[\varphi]]_a^w \wedge x \preceq_a y) \to x \preceq_a^{\Uparrow a\varphi} y)$
  - $\forall x\forall y((x, y \in [[\neg\varphi]]_a^w \wedge x \preceq_a y) \to x \preceq_a^{\Uparrow a\varphi} y)$

[1] $w' \prec_i w'' := w' \preceq_i w'' \wedge w'' \not\preceq_i w'$.

18 *Conservative update (for naive agents)*
$\mathcal{M}^{\mathcal{I}}, w \vDash [\mathcal{I}, \uparrow^n_a \varphi]\psi$ *iff* $\mathcal{M}^{\mathcal{I}}, w \vDash \hat{K}_a\varphi \wedge SB_a\varphi$ *implies* $\mathcal{M}^{\uparrow_a\varphi}, w \vDash \psi$,
where $\hat{K}_a\varphi ::= \neg K_a \neg\varphi$

19 *Radical update (for naive agents)*
$\mathcal{M}^{\mathcal{I}}, w \vDash [\mathcal{I}, \Uparrow^n_a \varphi]\psi$ *iff* $\mathcal{M}^{\mathcal{I}}, w \vDash \hat{K}_a\varphi \wedge SB_a\varphi$ *implies* $\mathcal{M}^{\Uparrow_a\varphi}, w \vDash \psi$

20 *Conservative update (for cautious agents)*
$\mathcal{M}^{\mathcal{I}}, w \vDash [\mathcal{I}, \uparrow^c_a \varphi]\psi$ *iff* $\mathcal{M}^{\mathcal{I}}, w \vDash \neg(B_a\varphi \vee B_a\neg\varphi) \wedge SB_a\varphi$ *implies* $\mathcal{M}^{\uparrow_a\varphi}, w \vDash \psi$

21 *Radical update (for cautious agents)*
$\mathcal{M}, w \vDash [\mathcal{I}, \Uparrow^c_a \varphi]\psi$ *iff* $\mathcal{M}^{\mathcal{I}}, w \vDash \neg(B_a\varphi \vee B_a\neg\varphi) \wedge SB_a\varphi$ *implies* $\mathcal{M}^{\Uparrow_a\varphi}, w \vDash \psi$

The formula $B^d_i\varphi$ is read as "i's believes that φ is true to degree d"

22 *An epistemic probability model is a tuple* $\mathcal{M} = (\mathcal{A}, W, \{\sim_i\}_{i\in\mathcal{A}}, \{P_i\}_{i\in\mathcal{A}}, V)$, *where*
- $\mathcal{A}$ *is a finite non-empty set of agents*
- W *is a finite non-empty set of possible worlds (situations)*
- $\sim_i \subseteq W \times W$ *is an equivalence relation for each agent* $i \in \mathcal{A}$ *interpreted as agent* i*'s epistemic indistinguishability.*
- $P_i : W \mapsto (W \mapsto [0,1])$ *such that*
  - $\forall w' \sum_{w''\in[w']_i} P_i(w')(w'') = 1$
  - $\forall w' \forall w'' (w'' \notin [w']_i \to P_i(w')(w'') = 0)$
  - $\forall w' \forall w'' \forall w''' (w'' \in [w']_i \to P_i(w')(w''') = P_i(w'')(w'''))$
- $V : Var \mapsto \mathcal{P}(W)$ *is a valuation map.*

23 *Suppose that* $\mathcal{M} = (\mathcal{A}, W, \{\sim_i\}_{i\in\mathcal{A}}, \{P_i\}_{i\in\mathcal{A}}, V)$ *is an epistemic probability model. The belief operator is defined as follows:*
- $\mathcal{M}, w \vDash B^d_i\varphi$ *iff* $d = \sum_{w'\in[\varphi]_{\mathcal{M}}} P_i(w)(w')$

24 *Let* Var *be a set of propositional variables and* $\mathcal{A}$ *a finite set of agents. The language of Static Social Influence Plausibility Model* $\mathcal{L}^{Pr}_S$ *is defined by the following BNF:*

$$\varphi_0 ::= \varphi_o \in \mathcal{L}^{Pr}$$
$$\varphi_1 ::= \varphi_o \mid SB^d_i\varphi_o \mid \neg\varphi_1 \mid \varphi_1 \wedge \psi_1 \mid \varphi_1 \vee \psi_1 \mid \varphi_1 \to \psi_1$$

where $p \in Var$, $i \in \mathcal{A}$, $d \in \mathbb{Q}$.

25 $d^w_a(b, \varphi) = \sum_{(x',y')\in[{}^a_b]^w_\varphi} x' \cdot y'$
where $[{}^a_b]^w_\varphi := \{(x,y) \mid \mathcal{M}, w \vDash B^x_a B^y_b \varphi\}$

26 $\mathcal{M}, w \vDash SB^d_a\varphi$ *iff* $\sum_{i\in\mathcal{A}} Infl(\varphi, i, a) \cdot d^w_a(i, \varphi)$

27 *Let* Var *be a set of propositional variables and* $\mathcal{A}$ *a finite set of agents. The language of Dynamic Social Influence Plausibility Model* $\mathcal{L}^{Pr}_D$ *is defined by the following BNF:*

$$\varphi ::= p \mid \neg\varphi \mid \varphi \wedge \psi \mid \varphi \vee \psi \mid \varphi \to \psi \mid K_i\varphi \mid B_i\varphi$$
$$[\mathcal{I}, {}_x\uparrow^n_a \varphi]\psi \mid [\mathcal{I}, {}_x\Uparrow^n_a \varphi]\psi \mid [\mathcal{I}, {}_x\uparrow^c_a \varphi]\psi \mid [\mathcal{I}, {}_x\Uparrow^c_a \varphi]\psi$$

where $p \in Var$, $i \in \mathcal{A}$, $\mathcal{I}$ *is a model of social influence.*

28 ($P^{x\uparrow\varphi}$) *Given* $\mathcal{M} = (\mathcal{A}, W, \{\sim_i\}_{i\in\mathcal{A}}, \{P_i\}_{i\in\mathcal{A}}, V)$ *update model is* $\mathcal{M} = (\mathcal{A}, W, \{\sim_i\}_{i\in\mathcal{A}}, \{P_i\}_{i\in\mathcal{A}}, V)$, *where*
$W' = W$;
$\sim_i = \sim_i$

$$P^{x\uparrow\varphi}(w)(w') = \begin{cases} \dfrac{P(w)(w')\cdot x}{\sum\limits_{w''\in[[\varphi]]_i^w} P(w)(w'')\cdot x + \sum\limits_{w''\in[[\neg\varphi]]_i^w} P(w)(w'')\cdot(1-x)}, & \text{if } w' \in [[\varphi]]_i^w \\ \dfrac{P(w)(w')\cdot x}{\sum\limits_{w''\in[[\varphi]]_i^w} P(w)(w'')\cdot x + \sum\limits_{w''\in[[\neg\varphi]]_i^w} P(w)(w'')\cdot(1-x)}, & \text{otherwise} \end{cases}$$

$V' = V$

29 ($P^{x\Uparrow\varphi}$) *Given* $\mathcal{M} = (\mathcal{A}, W, \{\sim_i\}_{i\in\mathcal{A}}, \{P_i\}_{i\in\mathcal{A}}, V)$ *update model is* $\mathcal{M} = (\mathcal{A}, W, \{\sim_i\}_{i\in\mathcal{A}}, \{P_i\}_{i\in\mathcal{A}}, V)$, *where*

- $W' = W$;
- $\sim_i = \sim_i$
- $P^{x\Uparrow\varphi}(w)(w') = \begin{cases} \dfrac{P(w)(w')\cdot x}{\sum\limits_{w''\in[[\varphi]]_i^w} P(w)(w'')}, & \text{if } w' \in [[\varphi]]_i^w \\ \dfrac{P(w)(w')\cdot(1-x)}{\sum\limits_{w''\in[[\varphi]]_i^w} P(w)(w'')}, & \text{otherwise} \end{cases}$
- $V' = V$

30 *Naive agent*
$\mathcal{M}, w \vDash [\mathcal{I},\ {}_x\uparrow_a^n \varphi]\psi$ *iff* $\mathcal{M}^{\mathcal{I}}, w \vDash \hat{K}_a\varphi \wedge SB_a^x\varphi$ *implies* $\mathcal{M}^{x\uparrow_a\varphi}, w \vDash \psi$

31 *Naive agent*
$\mathcal{M}, w \vDash [\mathcal{I},\ {}_x\Uparrow_a^n \varphi]\psi$ *iff* $\mathcal{M}^{\mathcal{I}}, w \vDash \hat{K}_a\varphi \wedge SB_a^x\varphi$ *implies* $\mathcal{M}^{x\uparrow_a\varphi}, w \vDash \psi$

32 *Cautious agent*
$\mathcal{M}, w \vDash [\mathcal{I},\ {}_x\uparrow_a^c \varphi]\psi$ *iff* $\mathcal{M}^{\mathcal{I}}, w \vDash B_a^{.5}\varphi \wedge SB_a^x\varphi$ *implies* $\mathcal{M}^{x\uparrow_a\varphi}, w \vDash \psi$

33 *Cautious agent*
$\mathcal{M}, w \vDash [\mathcal{I},\ {}_x\Uparrow_a^c \varphi]\psi$ *iff* $\mathcal{M}^{\mathcal{I}}, w \vDash B_a^{.5}\varphi \wedge SB_a^x\varphi$ *implies* $\mathcal{M}^{x\uparrow_a\varphi}, w \vDash \psi$