Social Influence Within Dynamic Epistemic Logic: Plausibility Models

Vitaliy Dolgorukov
vdolgorukov@hse.ru

Research Seminar «Formal Philosophy», October 28, 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*
- $Infl : \mathcal{A} \mapsto \Delta(\mathcal{A})$*, where* $\Delta(\mathcal{A})$ *is a probability distribution over* $\mathcal{A}$

5 *Some definitions*

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 *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* $\forall w' \in max_{\preceq_i}([w]_i) : \mathcal{M}^{\mathcal{I}}, w' \vDash \varphi$

7 $\mathcal{A}_a^w(\varphi) := \{i \in \mathcal{A} \mid \mathcal{M}, w \vDash B_a B_i \varphi\}$

8 $\mathcal{M}^{\mathcal{I}}, w \vDash SB_a\varphi$ *iff* $\sum_{i\in\mathcal{A}_a^w(\varphi)} Infl(i,a) > \sum_{i\in\mathcal{A}_a^w(\neg\varphi)} Infl(i,a)$

9 *Some Properties*

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. $\nvDash (SB_a\varphi \wedge SB_a(\varphi \to \psi)) \to SB_a\psi$

10 *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*
  - $\forall x\forall y((x \in best_a(\varphi, w) \wedge y \in [w]_a) \to y \preceq_a^{\uparrow_a\varphi} x)$
  - $\forall x\forall y((x, y \in [w]_a - best_a(\varphi, w) \wedge x \preceq_a y) \to x \preceq_a^{\uparrow_a\varphi} y)$

11 *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]
  - $\forall x\forall y((x, y \in [[\varphi]]_a^w \wedge x \preceq_a y) \to x \preceq_a^{\Uparrow_a\varphi} y)$
  - $\forall x\forall y((x, y \in [[\neg\varphi]]_a^w \wedge x \preceq_a y) \to x \preceq_a^{\Uparrow_a\varphi} y)$

12 *Conservative update (for naive agents)*
$\mathcal{M}^{\mathcal{I}}, w \vDash [\mathcal{I}, \uparrow_a^n \varphi]\psi$ *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 $\hat{K}_a\varphi ::= \neg K_a\neg\varphi$

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

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

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

---

[1] $w' \prec_i w'' := w' \preceq_i w'' \wedge w'' \npreceq_i w'$.