Logica epistemica dinamica per sistemi multi-agente in *liquid democracy*

Simone Cuconato* e Gianluigi Greco**

*Dottorando di Ricerca in ICT presso il Dipartimento di Ingegneria Informatica, Modellistica, Elettronica e Sistemistica - DIMES dell'Università della Calabria
cuconato.simone@gmail.com

**Professore Ordinario di Informatica Teorica e Direttore del Dipartimento di Matematica e Informatica - Demacs dell'Università della Calabria
gianluigi.greco@unical.it

1. Introduzione

Una parte rilevante delle moderne teorie economiche e sociali ha come obiettivo la costruzione di modelli logico-matematici in cui un insieme di agenti razionali prende decisioni individualmente (teoria della decisione), interattivamente (teoria dei giochi) o collettivamente (teoria della scelta sociale). I criteri che definiscono la razionalità di una decisione sono fortemente radicati nel paradigma della razionalità come coerenza che unisce i fondamenti della teoria della decisione e della teoria della probabilità a quelli della logica. La logica contemporanea occupa un posto sempre più ampio nell'analisi del comportamento razionale. Il seguente articolo elabora un modello logico in grado di analizzare il metodo di voto noto come *liquid democracy*. La democrazia liquida [1] è un approccio moderno al voto in cui gli elettori possono votare direttamente o delegare il proprio voto ad altri elettori. Una delle ragioni del successo della democrazia liquida è il fatto di essere vista come un compromesso pratico tra democrazia diretta e democrazia rappresentativa. Obiettivo del lavoro è definire un modello di logica epistemica dinamica[4,5] in grado di investigare, nel modo più rigoroso possibile, le decisioni interattive di voto nei sistemi multi-agente.

2. Logica epistemica multi-modale e multi-agente

Sintassi. La logica epistemica [2,6] è una estensione della logica classica che ha come oggetto di studio gli enunciati di credenza e di sapere. Esaminiamo il linguaggio logico del primo ordine $\mathcal{L}$. L'*alfabeto* di $\mathcal{L}$ è composto dai seguenti insiemi di simboli:

- lettere proposizionali: p_1, p_2, p_3, ...
- variabili individuali: x, y, z, ...
- costanti individuali: 1,2,3, ...
- costanti predicative: P_k^n
- connettivi vero-funzionali: $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$
- operatori modali: $\Box$, $\Diamond$
- operatori epistemici: B, K, I
- quantificatori esistenziali: $\forall, \exists$
- parentesi: (,)

L'*insieme delle formule* φ di $\mathcal{L}$ è definito come segue:

$$\varphi ::= p|\neg p|(\varphi \wedge \varphi)|(\varphi \vee \varphi)|\Box\varphi|K_i\varphi|B_i\varphi|I_i\varphi$$ [1]

Semantica. Le formule associate agli operatori modali ed epistemici verranno valutate vere o false relativamente a un particolare modello epistemico $\mathcal{M}$. Nello specifico il modello sarà la $n+2-$upla ordinata $\mathcal{M} = \langle W, R_{K_{1-n}}, R_{B_{1-n}}, R_{I_{1-n}}, I\rangle$[2], dove W sta per un insieme non vuoto di *mondi possibili*, R_K , R_B e R_I[3] per una relazione binaria in W e I per la funzione che, dato un mondo $w \in W$, associa uno e un solo valore di verità (T o F) ad ogni formula atomica di $\mathcal{L}$. Chiameremo la quadrupla ordinata $\langle W, R_K, R_B, R_I\rangle$ *frame epistemico*.

---

[1] Nello specifico, data una formula φ, $K_i\varphi$ sta per «L'agente cognitivo i sa che φ», $B_i\varphi$ sta per «L'agente cognitivo i crede che φ», mentre $I_i\varphi$ sta per «L'agente cognitivo *i* ignora che φ». Sulla logica dell'ignoranza si veda [7].

[2] Poiché la credenza è una condizione necessaria perché si abbia conoscenza, assumeremo che R_K sia un sottoinsieme di R_B:$R_K \subseteq R_B$.

[3] In generale, la relazione di accessibilità R nei calcoli modali è così definita: R è $seriale$ iff $(om\ w)(ex\ w')(wRw')$; R è $riflessiva$ iff $(om\ w)(wRw)$; R è $simmetrica$ iff $(om\ w)(om\ w')(wRw'$ iff $wRw')$;
R è $transitiva$ iff $(om\ w)(om\ w')(om\ w'')(wRw'$ and $w'Rw''$then $wRw'')$.

Definiamo la *verità di una formula nel mondo w del modello $\mathcal{M}$* attraverso le seguenti clausole:

- $\mathcal{M}, w \vDash p$ iff $w \in \mathrm{I}(p)$
- $\mathcal{M}, w \vDash \neg p$ iff $w \notin \mathrm{I}(p)$
- $\mathcal{M}, w \vDash (\varphi_1 \wedge \varphi_2)$ iff $\mathcal{M}, w \vDash \varphi_1$ and $\mathcal{M}, w \vDash \varphi_2$
- $\mathcal{M}, w \vDash (\varphi_1 \vee \varphi_2)$ iff $\mathcal{M}, w \vDash \varphi_1$ or $\mathcal{M}, w \vDash \varphi_2$
- $\mathcal{M}, w \vDash \Box\varphi$ iff $(om\ w')(w\mathrm{R}w'$ then $\mathcal{M}, w' \vDash \varphi)$[4]
- $\mathcal{M}, w \vDash K_i\varphi$ iff $(om\ w')(w\mathrm{R}_K w'$ then $\mathcal{M}, w' \vDash \varphi)$
- $\mathcal{M}, w \vDash B_i\varphi$ iff $(om\ w')(w\mathrm{R}_B w'$ then $\mathcal{M}, w' \vDash \varphi)$
- $\mathcal{M}, w \vDash I_i\varphi$ iff $(ex\ w', w'')(w\mathrm{R}_I w'$ and $w\mathrm{R}_I w''$ then $\mathcal{M}, w' \vDash \varphi$ and $\mathcal{M}, w'' \vDash \neg\varphi$ $)$

3. Liquid democracy

Sintassi. Per trattare in modo adeguato la nozione di *liquid democracy* estendiamo il linguaggio $\mathcal{L}$ per ottenere un linguaggio $\mathcal{L}^+$. L'*alfabeto* di $\mathcal{L}^+$ contiene due nuovi predicati a due posti:

- $\mathcal{V}$, da leggersi "vota"
- $\mathcal{D}$, da leggersi "delega"

L'*insieme delle formule ψ* di $\mathcal{L}^+$ è definito come segue:

$$\psi ::= \Gamma_i^{K \oplus B \oplus I} p_{w_i} \big| \neg\psi \big| (\psi \wedge \psi) \big| (\psi \vee \psi) \big|$$

Dove $\Gamma_i^{K \oplus B} p_{w_i}$ indica che «L'agente cognitivo i sa (K), crede (B) o ignora (I) p_{w_i}», con p_{w_i} avente la seguente forma:

$$p_{w_i} =_{df} \mathcal{V}_{i \oplus j}^{yes \oplus no} \text{ oppure } \mathcal{D}_{i \oplus j}^{l}$$

Con $\mathcal{V}_{i \oplus j}^{yes \oplus no}$ intendiamo che i o j *$vota\ yes$* o no, mentre con $\mathcal{D}_{i \oplus j}^{l}$ intendiamo che i o j $delega$ il voto a l. Di conseguenza, in $\mathcal{L}^+$ l'insieme delle formule ψ è sempre costituito da *formule epistemiche*.

Specifichiamo inoltre l'irriflessività di $\mathcal{D}$:

- $\mathcal{D}$ è irriflessiva iff not $\mathcal{D}_i^i$.

---

[4] $\mathcal{M}, w \vDash \Diamond\, \varphi$ iff $(ex\ w')(w\mathrm{R}w'$ and $\mathcal{M}, w' \vDash \varphi)$.

Semantica. Il modello per $\mathcal{L}^+$ è sempre la $n+2-$upla ordinata $\mathcal{M} = \langle$W,R$_{K_{1-n}}$, R$_{B_{1-n}}$, R$_{I_{1-n}}$,I$\rangle$. Naturalmente, in questo caso l'interpretazione I associa uno dei due valori di verità (T/F) come segue:

- $K_i\left(\mathcal{V}_{i\oplus j}^{yes\oplus no}\right) = \text{T/F}$
- $B_i\left(\mathcal{V}_{i\oplus j}^{yes\oplus no}\right) = \text{T/F}$
- $I_i\left(\mathcal{V}_{i\oplus j}^{yes\oplus no}\right) = \text{T/F}$
- $K_i\left(\mathcal{D}_{i\oplus j}^{l}\right) = \text{T/F}$
- $B_i\left(\mathcal{D}_{i\oplus j}^{l}\right) = \text{T/F}$
- $I_i\left(\mathcal{D}_{i\oplus j}^{l}\right) = \text{T/F}$

Definiamo la *verità di una formula nel mondo w del modello* $\mathcal{M}$ attraverso le seguenti clausole:

- $\mathcal{M}, w \vDash \Gamma_i^{K\oplus B\oplus I} P_{w_i}$ iff $K_i\left(\mathcal{V}_{i\oplus j}^{yes\oplus no}\right)$ or $B_i\left(\mathcal{V}_{i\oplus j}^{yes\oplus no}\right)$ or $I_i\left(\mathcal{V}_{i\oplus j}^{yes\oplus no}\right)$ or $K_i\left(\mathcal{D}_{i\oplus j}^{l}\right)$ or $B_i\left(\mathcal{D}_{i\oplus j}^{l}\right)$ or $I_i\left(\mathcal{D}_{i\oplus j}^{l}\right)$
- $\mathcal{M}, w \vDash \neg\psi$ iff $\mathcal{M}, w \nvDash \psi$
- $\mathcal{M}, w \vDash (\psi_1 \wedge \psi_2)$ iff $\mathcal{M}, w \vDash \psi_1$and $\mathcal{M}, w \vDash \psi_2$
- $\mathcal{M}, w \vDash (\psi_1 \vee \psi_2)$ iff $\mathcal{M}, w \vDash \psi_1$or $\mathcal{M}, w \vDash \psi_2$

Struttura. Consideriamo la seguente struttura $\mathcal{S} = \langle A, W, P_{w_i}, S\rangle$, dove:

- $A = \{1,2,3, \dots\}$ è un insieme non vuoto di agenti;
- $W = \{w_1, \dots, w_m\}$ è un insieme non vuoto di mondi epistemici ($|W| = m \in \mathbb{N}$);
- $P_{w_1} = \left\{p!, p_{1_{w_1}}, \dots, p_{m_{w_1}}\right\}$ è un insieme non vuoto di proposizioni $\left(\left|P_{w_1}\right| = m \in \mathbb{N}\right)$;
- $S = \{s_1, \dots, s_m\}$ è un insieme di scenari epistemici $\big(S \in P_{w_1}$ and $|S| = |A|\big)$.

$\mathcal{S}$ è una struttura dinamica, una struttura nella quale si verificano *mondi possibili epistemici* W. In ciascun mondo epistemico ogni agente possiede uno *scenario epistemico* S, ossia una serie di *credenze dinamiche* su se stesso e sugli altri agenti. Naturalmente, gli operatori epistemici *dinamici* sono B e

I, mentre l'operatore del sapere K è un operatore *statico*. A è l'insieme degli agenti della struttura, mentre P_{w_1} è l'insieme delle proposizioni epistemiche *iniziali* possedute da ciascun individuo appartenente ad A. Nello specifico, ogni agente ha delle *conoscenze di base* $p!$ e delle credenze o conoscenze precise su se stesso e gli altri agenti $p_{i_{w_i}}$. Mentre $p_{i_{w_i}}$ sappiamo già essere uguale per definizione a $\mathcal{V}_{i \oplus j}^{yes \oplus no}$ oppure $\mathcal{D}_{i \oplus j}^{l}$, al contrario per $p!$ intendiamo:

$$p! =_{df} \bigwedge_{i \in A} K_i(\Sigma)$$

Ogni individuo i che appartiene ad A conosce la *condizione epistemica* Σ. Dove a sua volta per Σ intendiamo quattro *sottocondizioni* $\alpha, \beta, \gamma, \delta$:

- α: in $\mathcal{S}$ si ha una *maggioranza assoluta* quando il voto *yes* o il voto *no* prendono il $50\% + 1$ di voti;
- β: se non si verifica in w_1 una maggioranza assoluta si passerà ad un mondo epistemico successivo w_2 ed eventualmente a w_3, w_4, ecc.;
- γ: ogni qual volta si effettua il passaggio ad un mondo epistemico successivo si verifica una situazione che chiameremo "*propagazione di opinione*". Per propagazione di opinione intendiamo il fatto che ogni agente all'interno della struttura può modificare il proprio voto *iff* la credenza relativa al voto appena dato è diversa da quella della *maggioranza relativa* degli agenti con cui è in contatto (*rete sociale*). Formalmente:
  - $B_i\left(\mathcal{V}_i^{yes}\right)_{yes}^{no}$;
  - $B_i(\mathcal{V}_i^{no})_{no}^{yes}$;
  - $B_i\left(\mathcal{D}_i^{j}\right)_{j}^{l}$

- δ: indicheremo il passaggio da credenza a conoscenza come segue:

$$[\uparrow_B]K_i\left(\mathcal{V}_{i \oplus j}^{yes \oplus no}\right) =_{df} \text{da } B_i\left(\mathcal{V}_{i \oplus j}^{yes \oplus no}\right) \text{ a } K_i\left(\mathcal{V}_{i \oplus j}^{yes \oplus no}\right)$$

  dove il cambiamento da credenza a conoscenza $[\uparrow_B]K_i$ può verificarsi *iff* almeno il 75% degli agenti appartenenti alla stessa rete sociale sanno di votare *yes* o *no*. Formalmente la $verit$à di $[\uparrow_B]K_i\left(\mathcal{V}_{i \oplus j}^{yes \oplus no}\right)$ $nel\ mondo\ w\ del\ modello\ \mathcal{M}$ sarà:

$\mathcal{M}, w \vDash [\uparrow_B]K_i(\mathcal{V}_{i\oplus j}^{yes\oplus no})$ iff $\forall i$
$\wedge (om\ w^*)\left(i \in A \text{ and } wR_K w^* \text{then } \mathcal{M}, w^* \vDash \left(\mathcal{V}_{i\oplus j}^{yes\oplus no}\right)\right)$

Inoltre, indicheremo con:

- $[\uparrow_B]K_i(\mathcal{V}_{i\oplus j}^{yes})_{yes}^{no}$;
- $[\uparrow_B]K_i(\mathcal{V}_{i\oplus j}^{no})_{no}^{yes}$

il passaggio da credenza a conoscenza e il relativo cambiamento di opinione da parte di un agente di A

- ε: gli agenti che ignorano se votare o delegare hanno delle condizioni di propagazione di opinione *soggettive*: devono essere scelte individualmente per ogni agente di A appartenente alla struttura $\mathcal{S}$. Formalmente indicheremo il passaggio da ignoranza a credenza o conoscenza come segue:
  - ε_1:$[\uparrow_I]B_i(\mathcal{V}_i^{yes\oplus no} \oplus \mathcal{D}_i^l) =_{df}$ da $I_i(\mathcal{V}_i^{yes\oplus no} \oplus \mathcal{D}_i^l)$ a $B_i(\mathcal{V}_i^{yes\oplus no} \oplus \mathcal{D}_i^l)$
  - ε_2:$[\uparrow_I]K_i(\mathcal{V}_i^{yes\oplus no} \oplus \mathcal{D}_i^l) =_{df}$ da $I_i(\mathcal{V}_i^{yes\oplus no} \oplus \mathcal{D}_i^l)$ a $K_i(\mathcal{V}_i^{yes\oplus no} \oplus \mathcal{D}_i^l)$

Posta la condizione epistemica Σ possiamo ora definire l'*operatore di transizione di opinione* μ:

Definizione 1 (operatore di transizione di opinione) Data una struttura $\mathcal{S}$ e una proposizione $p_{i_{w_i}} \in P_{w_i}$, l'operatore di transizione di opinione μ definirà la transizione di opinione come segue:

$$\mu p'_{1_{w_i}}, p'_{2_{w_i}}, \dots, p'_{m_{w_i}} \longleftarrow p_{1_{w_i}}, p_{2_{w_i}}, \dots, p_{m_{w_i}}$$

Nello specifico, μ definirà il passaggio secondo le sottocondizioni $\gamma, \delta, \varepsilon$:

1. $\gamma.\ p'_{i_{w_i}} = [\gamma]p_{i_{w_i}}$
2. $\delta.\ p'_{i_{w_i}} = [\delta]p_{i_{w_i}}$
3. $\gamma + \delta.\ p'_{i_{w_i}} = [\gamma + \delta]p_{i_{w_i}}$
4. not γ and not $\delta.\ p'_{i_{w_i}} = p_{i_{w_i}}$

5. $\varepsilon_1.\ p'_{i_{w_i}} = [\varepsilon_1]p_{i_{w_i}}$
6. $\varepsilon_2.\ p'_{i_{w_i}} = [\varepsilon_2]p_{i_{w_i}}$
7. not $\varepsilon.\ p'_{i_{w_i}} = p_{i_{w_i}}$

Ricapitolando, la definizione formale dell'operatore μ è la seguente:

$$\mu p'_{i_{w_i}} \longleftarrow p_{i_{w_i}} \equiv \left\{ \begin{array}{c} p'_{i_{w_i}} = [\gamma]p_{i_{w_i}} \text{ se vale } \gamma \\ p'_{i_{w_i}} = [\delta]p_{i_{w_i}} \text{ se vale } \delta \\ p'_{i_{w_i}} = [\gamma+\delta]p_{i_{w_i}} \text{ se vale } \gamma+\delta \\ p'_{i_{w_i}} = p_{i_{w_i}} \text{ se vale not } \gamma \text{ and not } \delta \\ p'_{i_{w_i}} = [\varepsilon_1]p_{i_{w_i}} \text{se vale } \varepsilon_1 \\ p'_{i_{w_i}} = [\varepsilon_2]p_{i_{w_i}} \text{se vale } \varepsilon_2 \\ p'_{i_{w_i}} = p_{i_{w_i}} \text{se vale not } \varepsilon \end{array} \right\}$$

4. Osservazioni conclusive

La creazione di modelli logico-matematici delle capacità deduttive degli agenti razionali è da alcuni decenni percepita come un problema urgente nell'ambito della teoria della razionalità. In questo articolo abbiamo presentato un approccio logico innovativo in grado di investigare le decisioni interattive di voto nei sistemi multi-agente. In particolare sono stati sintatticamente e semanticamente definiti due nuovi predicati a due posti "vota" e "delega", le nozioni di "struttura dinamica", "scenario epistemico", "condizione epistemica" e, infine, la situazione di "propagazione di opinione" con la definizione di uno specifico "operatore di transizione di opinione". Fornire un modello espresso in un linguaggio logico governato da precise regole formali garantisce uno strumento indispensabile alla creazione di una rigorosa teoria della scelta razionale. Tuttavia, l'obiettivo abituale di fornire, attraverso la formalizzazione, un'esplicitazione dei concetti fondamentali della teoria che ne favoriscano una più profonda comprensione non può essere considerato l'unico obiettivo. Un secondo obiettivo, non meno importante, è quello di aprire la strada all'automazione. L'idea che la logica possa giocare un ruolo più diretto nella specificazione formale e nel controllo delle procedure di scelta sociale, così come è stato a lungo usata nell'informatica per specificare e verificare automaticamente le proprietà del *software*, è insita nel programma delineato da Rohit Parihk e suggestivamente chiamato *Social Software*[3]. In conclusione, l'idea alla base di questo lavoro è che per caratterizzare adeguatamente le decisioni interattive di voto nei sistemi multi-agente occorra tenere conto dei suoi aspetti "dinamici": i soggetti epistemici, infatti, tendono a rivedere costantemente il proprio corpus doxastico alla luce delle informazioni di volta in volta disponibili.

5. **Riferimenti bibliografici**

[1] Behrens, J., Kistner, A., Nitsche A., Swierczek B., 2014, "Principles of Liquid Feedback", Interaktieve Demokratie.

[2] Meyer, Ch., van der Hoek, W., 1995, *Epistemic Logic for AI and Computer Science*, Cambridge University Press.

[3] Parikh, R., 2002, "Social software", Synthese (3): 187-211.

[4] van Benthem, J., 2007, "Dynamic logic for belief revision", Journal of Applied Non- Classical Logics, 17(2): 129–155.

[5] van Ditmarsch, H., van der Hoek, W., Kooi, B., 2007, *Dynamic Epistemic Logic*, Volume 337 of Synthese Library, Netherlands: Springer.

[6] van Ditmarsch, H., Halpern, J., Van Der Hoek, Kooi, B., 2015, *Handbook of Epistemic Logic*, College Publications.

[7] van der Hoek, W., Lomuscio, A., 2004, "A Logic for Ignorance", Electronic Notes in Theoretical Computer Science, 85 No. 2.