Arrow's Theorem in Modal Logic

Giovanni Ciná
Joint work with Ulle Endriss

20/03/2015

Institute for Logic, Language and Computation

Logics for Social Choice Theory

Quite a few logics for Social Choice: [1, 2, 4, 5, 6, 7, 9, 11, 12, 13].

Logics for Social Choice Theory

Quite a few logics for Social Choice: [1, 2, 4, 5, 6, 7, 9, 11, 12, 13].
What is logic useful for? (list borrowed from [8])

- formal representation and retrieval
- makes hidden assumptions explicit
- confirms existing results
- cleans up proofs
- suggests new proof strategies
- helps find new results (inc. new types of results)
- helps review work

Logics for Social Choice Theory

Quite a few logics for Social Choice: [1, 2, 4, 5, 6, 7, 9, 11, 12, 13].
What is logic useful for? (list borrowed from [8])

- formal representation and retrieval
- makes hidden assumptions explicit
- confirms existing results
- cleans up proofs
- suggests new proof strategies
- helps find new results (inc. new types of results)
- helps review work

To test the expressive power of the *modal logic of social choice functions* proposed by Troquard et al. [12], Ulle Endriss and I gave a syntactic proof Arrow's Theorem.

Outline

1. Arrow's Theorem
2. A proof
3. A logic
4. Encoding the proof

Outline

The setting

Given a set of alternatives X, we suppose each agent has a preference over these alternatives, namely a reflexive, antisymmetric, complete, and transitive relation over X.
Question: given a set of agents N, how do we aggregate the preferences of individuals into a unique collective preference?

The setting

Given a set of alternatives X, we suppose each agent has a preference over these alternatives, namely a reflexive, antisymmetric, complete, and transitive relation over X.
Question: given a set of agents N, how do we aggregate the preferences of individuals into a unique collective preference?
Let $\mathcal{L}(X)$ denote the set of all such linear orders. Call $\succcurlyeq_i$ the *ballot* provided by agent i. A *profile* is an n-tuple $(\succcurlyeq_1, \ldots, \succcurlyeq_n) \in \mathcal{L}(X)^n$ of such ballots. Indicate with $N^w_{x \succcurlyeq y}$ the set of agents preferring x over y in profile w.

Definition

A resolute social choice function is a function $F : \mathcal{L}(X)^n \to X$ mapping any given profile of ballots to a single winning alternative.

The properties of a SCF

Three properties are mentioned in the statement of Arrow's Theorem: Independence of Irrelevant Alternatives (IIA), Pareto efficiency and Dictatorship.

The properties of a SCF

Three properties are mentioned in the statement of Arrow's Theorem: Independence of Irrelevant Alternatives (IIA), Pareto efficiency and Dictatorship.

Definition

A SCF F satisfies IIA if, for every pair of profiles $w, w' \in \mathcal{L}(X)^n$ and every pair of distinct alternatives $x, y \in X$ with $N^{w}_{x \succcurlyeq y} = N^{w'}_{x \succcurlyeq y}$, $F(w) = x$ implies $F(w') \neq y$.

The properties of a SCF

Three properties are mentioned in the statement of Arrow's Theorem: Independence of Irrelevant Alternatives (IIA), Pareto efficiency and Dictatorship.

Definition

A SCF F satisfies IIA if, for every pair of profiles $w, w' \in \mathcal{L}(X)^n$ and every pair of distinct alternatives $x, y \in X$ with $N^{w}_{x \succcurlyeq y} = N^{w'}_{x \succcurlyeq y}$, $F(w) = x$ implies $F(w') \neq y$.

Definition

A SCF F is Pareto efficient if, for every profile $w \in \mathcal{L}(X)^n$ and every pair of distinct alternatives $x, y \in X$ with $N^{w}_{x \succcurlyeq y} = N$, we obtain $F(w) \neq y$.

The properties of a SCF

Three properties are mentioned in the statement of Arrow's Theorem: Independence of Irrelevant Alternatives (IIA), Pareto efficiency and Dictatorship.

Definition

A SCF F satisfies IIA if, for every pair of profiles $w, w' \in \mathcal{L}(X)^n$ and every pair of distinct alternatives $x, y \in X$ with $N^w_{x \succcurlyeq y} = N^{w'}_{x \succcurlyeq y}$, $F(w) = x$ implies $F(w') \neq y$.

Definition

A SCF F is Pareto efficient if, for every profile $w \in \mathcal{L}(X)^n$ and every pair of distinct alternatives $x, y \in X$ with $N^w_{x \succcurlyeq y} = N$, we obtain $F(w) \neq y$.

Definition

A SCF F is a dictatorship if there exists an agent $i \in N$ (the dictator) such that, for every profile $w \in \mathcal{L}(X)^n$, we obtain $F(w) = top^w_i$.

The theorem

We are ready to state Arrow's Theorem itself:

Theorem (Arrow)

Any SCF for $\geqslant 3$ *alternatives that satisfies IIA and the Pareto condition is a dictatorship.*

Outline

A proof

We present a well known proof of the theorem [5, 10], exploiting the notion of decisive coalition.

A proof

We present a well known proof of the theorem [5, 10], exploiting the notion of decisive coalition.

Definition

A coalition $C \subseteq N$ is *decisive* over a pair of alternatives $(x, y) \in X^2$ if $C \subseteq N^w_{x \succcurlyeq y}$ entails $F(w) \neq y$.
A coalition $C \subseteq N$ is *weakly decisive* over $(x, y) \in X^2$ if $C = N^w_{x \succcurlyeq y}$ entails $F(w) \neq y$.

A proof

The general strategy of the proof is the following.

1. If a coalition is weakly decisive over one pair then it is decisive over any pair.

A proof

The general strategy of the proof is the following.

1. If a coalition is weakly decisive over one pair then it is decisive over any pair.
2. By 1, if a coalition C is decisive over any pair and C is partitioned into two disjoint sets C_1 and C_2 then one of the two latter sets must be decisive over any pair (Contraction Lemma).

A proof

The general strategy of the proof is the following.

1. If a coalition is weakly decisive over one pair then it is decisive over any pair.
2. By 1, if a coalition C is decisive over any pair and C is partitioned into two disjoint sets C_1 and C_2 then one of the two latter sets must be decisive over any pair (Contraction Lemma).
3. By Pareto the whole set N is decisive over all pairs; by repeated application of Contraction Lemma we infer that there is a singleton coalition that is decisive over any pair, i.e. a dictator.

Outline

Syntax

Troquard et al. [12] introduced a modal logic, called $\Lambda^{\mathrm{scf}}[N, X]$, to reason about resolute SCF's as well as the agents' truthful preferences. We use a fragment of this logic, called here $L[N, X]$.

Syntax

Troquard et al. [12] introduced a modal logic, called $\Lambda^{\mathrm{scf}}[N, X]$, to reason about resolute SCF's as well as the agents' truthful preferences. We use a fragment of this logic, called here $L[N, X]$.

Definition

The language of $L[N, X]$ is the following:

$$\varphi \quad ::= \quad p \,|\, x \,|\, \neg\varphi \,|\, \varphi \vee \psi \,|\, \Diamond_C \varphi$$

where $p \in \{p^i_{x \succcurlyeq y} \mid i \in N \text{ and } x, y \in X\}$, $x \in X$ and $C \subseteq N$.

Semantics

Definition

A model is a triple $M = \langle N, X, F \rangle$, consisting of a finite set of agents N with $n = |N|$, a finite set of alternatives X, and a SCF $F : \mathcal{L}(X)^n \rightarrow X$.

Semantics

Definition

A model is a triple $M = \langle N, X, F \rangle$, consisting of a finite set of agents N with $n = |N|$, a finite set of alternatives X, and a SCF $F : \mathcal{L}(X)^n \to X$.

Definition

Let M be a model. We write $M, w \models \varphi$ to express that the formula φ is true at the world $w = (\succcurlyeq_1, \ldots, \succcurlyeq_n) \in \mathcal{L}(X)^n$ in M. Define:

- $M, w \models p^i_{x \succcurlyeq y}$ iff $x \succcurlyeq_i y$
- $M, w \models x$ iff $F(w) = x$
- $M, w \models \neg\varphi$ iff $M, w \not\models \varphi$
- $M, w \models \varphi \vee \psi$ iff $M, w \models \varphi$ or $M, w \models \psi$
- $M, w \models \Diamond_C \varphi$ iff $M, w' \models \varphi$ for some world $w' = (\succcurlyeq'_1, \ldots, \succcurlyeq'_n) \in \mathcal{L}(X)^n$ with $\succcurlyeq_i = \succcurlyeq'_i$ for all $i \in N \setminus C$.

Notation

We can encode some semantic notions into formulas:

$$ballot_i(w) \quad := \quad p^i_{x_1 \succcurlyeq x_2} \wedge p^i_{x_2 \succcurlyeq x_3} \wedge \cdots \wedge p^i_{x_{m-1} \succcurlyeq x_m}$$

encodes the ballot of agent i.

Notation

We can encode some semantic notions into formulas:

$$ballot_i(w) \quad := \quad p^i_{x_1 \succcurlyeq x_2} \wedge p^i_{x_2 \succcurlyeq x_3} \wedge \cdots \wedge p^i_{x_{m-1} \succcurlyeq x_m}$$

encodes the ballot of agent i.

$$profile(w) \quad := \quad ballot_1(w) \wedge ballot_2(w) \wedge \cdots \wedge ballot_n(w)$$

$profile(w)$ is true at world w, and only there; hence *nominals*, i.e., formulas uniquely identifying worlds [3], are definable within this logic at no extra cost.

Notation

We can encode some semantic notions into formulas:

$$ballot_i(w) \quad := \quad p^i_{x_1 \succcurlyeq x_2} \wedge p^i_{x_2 \succcurlyeq x_3} \wedge \cdots \wedge p^i_{x_{m-1} \succcurlyeq x_m}$$

encodes the ballot of agent i.

$$profile(w) \quad := \quad ballot_1(w) \wedge ballot_2(w) \wedge \cdots \wedge ballot_n(w)$$

$profile(w)$ is true at world w, and only there; hence *nominals*, i.e., formulas uniquely identifying worlds [3], are definable within this logic at no extra cost.

$$profile(w)(x, y) \;:=\; \bigwedge_{i \in N} \{p^i_{x \succcurlyeq y} \mid x \succcurlyeq_i y\} \wedge \bigwedge_{i \in N} \{p^i_{y \succcurlyeq x} \mid y \succcurlyeq_i x\}$$

Axiomatization

1. all propositional tautologies
2. formulas $p^i_{x \succcurlyeq y}$ are arranged in a linear order
3. $\Box_i(\varphi \to \psi) \to (\Box_i \varphi \to \Box_i \psi)$ (K(i))
4. $\Box_i \varphi \to \varphi$ (T(i))
5. $\varphi \to \Box_i \Diamond_i \varphi$ (B(i))
6. $\Diamond_i \Box_j \varphi \leftrightarrow \Box_j \Diamond_i \varphi$ (confluence)
7. $\Box_{C_1} \Box_{C_2} \varphi \leftrightarrow \Box_{C_1 \cup C_2} \varphi$ (union)
8. $\Box_\emptyset \varphi \leftrightarrow \varphi$ (empty coalition)
9. $(\Diamond_i p \wedge \Diamond_i \neg p) \to (\Box_j p \vee \Box_j \neg p)$, where $i \neq j$ (exclusive)
10. $\Diamond_i ballot_i(w)$ (ballot)
11. $\Diamond_{C_1} \delta_1 \wedge \Diamond_{C_2} \delta_2 \to \Diamond_{C_1 \cup C_2}(\delta_1 \wedge \delta_2)$ (cooperation)
12. $\bigvee_{x \in X}(x \wedge \bigwedge_{y \in X \setminus \{x\}} \neg y)$ (resolute)
13. $(profile(w) \wedge \varphi) \to \Box_N(profile(w) \to \varphi)$ (functional)

Nice results

The logic $L[N, X]$ behaves well:

Lemma

Determining whether a formula in the language of $L[N, X]$ is valid is a decidable problem.

Theorem

The logic $L[N, X]$ is sound and complete w.r.t. the class of models of SCF's.

Outline

Here is how the aforementioned properties are coded in the logical language:

$$IIA := \bigwedge_{w \in \mathcal{L}(X)^n} \bigwedge_{x \in X} \bigwedge_{y \in X \setminus \{x\}} [\Diamond_N(profile(w) \wedge x) \rightarrow (profile(w)(x,y) \rightarrow \neg y)]$$

$$P := \bigwedge_{x \in X} \bigwedge_{y \in X \setminus \{x\}} \left[\left(\bigwedge_{i \in N} p^i_{x \succcurlyeq y} \right) \rightarrow \neg y \right]$$

$$D := \bigvee_{i \in N} \bigwedge_{x \in X} \bigwedge_{y \in X \setminus \{x\}} (p^i_{x \succcurlyeq y} \rightarrow \neg y)$$

Proof

We use the following formula to encode decisiveness of C over (x, y):

$$Cdec(x, y) \quad := \quad \left(\bigwedge_{i \in C} p^i_{x \succcurlyeq y} \right) \rightarrow \neg y$$

If C is decisive on every pair, we will simply write $Cdec$.

Proof

We use the following formula to encode decisiveness of C over (x, y):

$$Cdec(x, y) \quad := \quad \left(\bigwedge_{i \in C} p^i_{x \succcurlyeq y} \right) \rightarrow \neg y$$

If C is decisive on every pair, we will simply write $Cdec$.
We define a *weakly decisive* coalition C for (x, y) as a coalition that can bar y from winning if *exactly* the agents in C prefer x to y:

$$Cwdec(x, y) \quad := \quad \left(\bigwedge_{i \in C} p^i_{x \succcurlyeq y} \wedge \bigwedge_{i \notin C} p^i_{y \succcurlyeq x} \right) \rightarrow \neg y$$

Proof

We first prove that every possible profile exists in the semantics:

Lemma (Universal domain)

For every possible profile $w \in \mathcal{L}(X)^n$, we have $\vdash \Diamond_N \mathit{profile}(w)$.

Proof

We first prove that every possible profile exists in the semantics:

Lemma (Universal domain)

For every possible profile $w \in \mathcal{L}(X)^n$, we have $\vdash \Diamond_N profile(w)$.

Proof.

Take any w. Then $ballot_1(w)$ encodes the preferences of the first agent. By axiom (10) we have $\Diamond_1 ballot_1(w)$, and similarly we get $\Diamond_2 ballot_2(w)$. Because $ballot_1(w)$ and $ballot_2(w)$ contain different atoms, we can apply axiom (11) and obtain $\Diamond_{\{1,2\}}(ballot_1(w) \wedge ballot_2(w))$. We repeat this reasoning for all the finitely many agents in N to prove $\Diamond_N profile(w)$. □

Proof

Lemma (1)

Consider a language parametrised by X such that $|X| \geqslant 3$. Then for any coalition $C \subseteq N$ and any two distinct alternatives $x, y \in X$, we have that:

$$\vdash P \wedge IIA \wedge Cwdec(x, y) \to Cdec$$

Proof

Lemma (1)

Consider a language parametrised by X such that $|X| \geqslant 3$. Then for any coalition $C \subseteq N$ and any two distinct alternatives $x, y \in X$, we have that:

$$\vdash P \wedge IIA \wedge Cwdec(x, y) \rightarrow Cdec$$

Lemma (2, Contraction Lemma)

Consider a language parametrised by X such that $|X| \geqslant 3$. Then for any coalition $C \subseteq N$ with and any two coalitions C_1 and C_2 that form a partition of C, we have that:

$$\vdash P \wedge IIA \wedge Cdec \rightarrow (C_1dec \vee C_2dec)$$

Proof

Theorem

Consider a language parametrised by X such that $|X| \geqslant 3$. Then we have:

$$\vdash P \wedge IIA \rightarrow D$$

Proof.

We know P is equivalent to $Ndec$. Exploiting the premise $P \wedge IIA$, we can apply the Contraction Lemma and prove that one of two disjoint subsets of N is decisive. Repeating the process finitely many times (we have finitely many agents), we can show that one of the singletons that form N is decisive. But this is tantamount to deriving D, i.e. saying that there exist a dictator. □

Further work

The plan for the near future:

- Encode more commonly studied notions of voting theory in the logic considered here and prove other results such as May's Theorem or Sen's approach to rights.
- Exploit the computational feasibility of modal logic by working on an optimised implementation.

References I

Thomas Ågotnes, Wiebe van der Hoek, and Michael Wooldridge.
On the logic of preference and judgment aggregation.
Autonomous Agents and Multiagent Systems, 22(1):4–30, 2011.

Bernhard Beckert, Rajeev Goré, Carsten Schürmann, Thorsten Bormer, and Jian Wang.
Verifying voting schemes.
Journal of Information Security and Applications, 19(2):115–129, 2014.

Patrick Blackburn, Maarten de Rijke, and Yde Venema.
Modal Logic.
Cambridge University Press, 2001.

Felix Brandt and Christian Geist.
Finding strategyproof social choice functions via SAT solving.
In *Proc. 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2014)*, 2014.

References II

Ulle Endriss.
Logic and social choice theory.
In A. Gupta and J. van Benthem, editors, *Logic and Philosophy Today*, volume 2, pages 333–377. College Publications, 2011.

Christian Geist and Ulle Endriss.
Automated search for impossibility theorems in social choice theory: Ranking sets of objects.
Journal of Artificial Intelligence Research, 40:143–174, 2011.

Umberto Grandi and Ulle Endriss.
First-order logic formalisation of impossibility theorems in preference aggregation.
Journal of Philosophical Logic, 42(4):595–618, 2013.

References III

Christoph Lange, Colin Rowat, and Manfred Kerber.
The ForMaRE Project: Formal mathematical reasoning in economics.
In *Intelligent Computer Mathematics*, pages 330–334. Springer-Verlag, 2013.

Tobias Nipkow.
Social choice theory in HOL: Arrow and Gibbard-Satterthwaite.
Journal of Automated Reasoning, 43(3):289–304, 2009.

A. K. Sen.
Social choice theory.
In K. J. Arrow and M. D. Intriligator, editors, *Handbook of Mathematical Economics*, volume 3. North-Holland, 1986.

P. Tang and F. Lin.
Computer-aided proofs of Arrow's and other impossibility theorems.
Artificial Intelligence, 173(11):1041–1053, 2009.

References IV

Nicolas Troquard, Wiebe van der Hoek, and Michael Wooldridge.
Reasoning about social choice functions.
Journal of Philosophical Logic, 40(4):473–498, 2011.

F. Wiedijk.
Arrow's Impossibility Theorem.
Formalized Mathematics, 15(4):171–174, 2007.