Towards an axiomatized logic of Social Welfare

Matthijs Melissen

June 9, 2007

Abstract

In computational social choice, it is hard to reason about social welfare functions. In this paper, a logic is presented to do this job. The logic is strong enough for representing higher order properties, like Arrows theorem. We present the syntax and semantics of this logic. For two sublanguages, $\mathcal{L}_3$ and $\mathcal{L}_2$, an axiomatization is given. We also make a start with giving an axiomatization for the system itself.

1 Introduction

Suppose a school class want to choose an activity for their yearly excursion. They might go to the zoo, have a diner, go to cinema, or go to a theme park. Every child has a favourite option, and a second, third and forth preference. How do the children produce a common ranking of the options?

This is a typical example of a problem in the field of computational social choice. In this field several methods exists to aggregrate the individual **preference orders** of n agents over the **alternatives** A into a collective preference order. Preference orders are total linear orders on A. The set of all preference relations is denoted as $L(A)$.

At the moment it is hard to reason about **social welfare functions** (**SWF**s). An SWF F is a function from a **preference profile**, which consists of one preference order for each agent, to a preference order for society:

$$F : L(A)^n \to L(A)$$

To facilitate reasoning about SWFs, a logical system is designed (Ågotnes et al., 2006). However, no axiomatization is given. In this paper, we make a beginning with finding an axiomatisation for this system.

First we present the syntax of the system. After that, the semantics of the logic is given. Then we show some examples. In the next section, we present an axiomatization of $\mathcal{L}_2$ and $\mathcal{L}_3$, and give a starting point for $\mathcal{L}$. Finally, an overview of the paper is presented, and some suggestions for future work are made.

2 Syntax

The logical system is parametrized by the number of agents n, as well as by a set of symbols $\Pi = \{r, s, \ldots\}$. Furthermore Σ is defined to be the set $\{1, \ldots, n\}$. The system consists of three sublanguages, which are defined as follows:

$\mathcal{L} : \phi ::= \Box\psi \mid \neg\phi \mid \phi_1 \wedge \phi_2$
$\mathcal{L}_2 : \psi ::= \boxdot\gamma \mid \neg\psi \mid \psi_1 \wedge \psi_2$
$\mathcal{L}_3 : \gamma ::= r_i \mid \neg\gamma \mid \gamma_1 \wedge \gamma_2$ where $i \in \Sigma$ and $r \in \Pi$

Furthermore we define the following operators:

$$\begin{aligned} \Diamond\psi &\equiv \neg\Box\neg\psi \\ \Diamonddot\gamma &\equiv \neg \boxdot \neg\gamma \\ \phi \vee \psi &\equiv \neg(\neg\phi \wedge \neg\psi) \\ \phi \rightarrow \psi &\equiv \neg\phi \vee \psi \\ \phi \leftrightarrow \psi &\equiv (\phi \rightarrow \psi) \wedge (\phi \leftarrow \psi) \end{aligned}$$

3 Semantics

3.1 Semantics of $\mathcal{L}_3$

Informally, $\mathcal{L}_3$ gives information about a pair of distinct alternatives a and b, given a set of alternatives, an SWF F, and a profile function $\delta : \Pi \rightarrow L(A)^n$ (the profile function assigns a preference pofile to every letter in Π). A formula r_i is true on a pair (a, b) if the preference profile where the symbol r is mapped to let agent i rank a above b. If we do not talk about a single agent but about society, we simply write r instead of r_i. We can formalize the semantics of this language as follows:

$$\begin{aligned} (A, F, \delta, (a, b)) \models r_i &\Leftrightarrow (a, b) \in \delta_i(r) \\ (A, F, \delta, (a, b)) \models r &\Leftrightarrow (a, b) \in F(\delta(r)) \\ (A, F, \delta, (a, b)) \models \neg\gamma &\Leftrightarrow (A, F, \delta, (a, b)) \not\models \gamma \\ (A, F, \delta, (a, b)) \models \gamma_1 \wedge \gamma_2 &\Leftrightarrow (A, F, \delta, (a, b)) \models \gamma_1 \text{ and } (A, F, \delta, (a, b)) \models \gamma_2 \end{aligned}$$

Example 1. *For example, if we want to evaluate $(A, F, \delta, (c, d)) \models r_3$, we first find out to which preference profile r is mapped, by calculating $\delta(r)$. From this profile we take the preference relation for agent 3. Now we return true if in this relation c is ranked over d, and false otherwise.*

3.2 Semantics of $\mathcal{L}_2$

The language $\mathcal{L}_2$ gives information about preference profiles. It is interpreted in a set of alternatives A, an SWF F and a profile function δ. Intuitively, the symbol $\boxdot$ has the meaning 'for all pairs', and the symbol $\diamonddot$ means 'for some pair'.

$$
\begin{array}{rcl}
(A,F,\delta) \models \boxdot\gamma & \Leftrightarrow & \forall_{(a,b)\in A\times A} a \neq b \Rightarrow (A,F,\delta,(a,b)) \models \gamma \\
(A,F,\delta) \models \neg\psi & \Leftrightarrow & (A,F,\delta) \not\models \psi \\
(A,F,\delta) \models \psi_1 \wedge \psi_2 & \Leftrightarrow & (A,F,\delta) \models \psi_1 \text{ and } (A,F,\delta) \models \psi_2
\end{array}
$$

Example 2. *If we want to check whether* $(A,F,\delta) \models \diamonddot(r_3 \leftrightarrow r)$*, we first rewrite it to* $(A,F,\delta) \models \neg \boxdot \neg(r_3 \leftrightarrow r)$*. Now we have to check whether* $\neg(r_3 \leftrightarrow r)$ *is not in all pairs true, in other words, whether there is a pair* (a,b) *in which in* $\mathcal{L}3$ $(A,F,\delta,(a,b)) \models (r_3 \leftrightarrow r)$ *holds. To do this, we look up to which preference profile* r *is mapped to, and check whether there is a pair* (a,b) *in which agent 3 agrees with society in this preference profile.*

3.3 Semantics of $\mathcal{L}$

Finally, the language $\mathcal{L}$ is interpreted in a social welfare function. We define Δ to be the set of all profile functions. The symbol $\Box$ can be read as 'in all preference profiles', and the symbol $\Diamond$ can be read as 'for some preference profiles'. Formally the semantics is as follows:

$$
\begin{array}{rcl}
(A,F) \models \Box\psi & \Leftrightarrow & \forall_{\delta\in\Delta}(A,F,\delta) \models \psi \\
(A,F) \models \neg\phi & \Leftrightarrow & (A,F) \not\models \phi \\
(A,F) \models \phi_1 \wedge \phi_2 & \Leftrightarrow & (A,F) \models \phi_1 \text{ and } (A,F) \models \phi_2
\end{array}
$$

4 Examples

Now we can express some properties from social choice in this logic. Pareto optimality can be expressed as follows:

$$PO = \Box \boxdot ((r_1 \wedge \ldots \wedge r_n) \rightarrow r)$$

Remember that this intuitively means that for all mappings from r to a preference profile - or all possible preference profiles r - and all pairs (a,b) holds that if agents $1\ldots n$ rank a over b, society does the same.

Non-dictatorship can be expressed in this way:

$$ND = \wedge_{i\in\Sigma}\Diamond\diamonddot\neg(r \leftrightarrow r_i)$$

This says that for all agents it holds that there is a preference profile r in which there is a pair (a, b) such that the agent disagrees with society about the ranking of that pair.

Finally, we show how to represent independence of irrelevant alternatives:

$$IIA = \Box \boxdot (((r_1 \leftrightarrow s_1) \wedge \ldots \wedge (r_n \leftrightarrow s_n)) \rightarrow (r \leftrightarrow s)$$

This formula says that for all mappings from r and s to preference profiles - so for all preference profiles r and s - and for all pairs it holds that if all agents have the same opinion about that pair in two preference profiles, society also agrees about that pair in both profiles.

5 Towards an axiomatization

5.1 Axiomatization of $\mathcal{L}_3$

We define ϕ to be valid in $\mathcal{L}_3$, written as $\models \phi$, if and only if $(A, F, \delta, (a, b)) \models \phi$ for all sets of alternatives A, social welfare functions F, preference functions δ and pairs (a, b).

Theorem 1. *ϕ is valid in $\mathcal{L}_3$ if and only if ϕ is valid in proposition logic.*

Proof. It holds that all valuations of r and r_i with $r \in \Pi$ and $i \in \Sigma$ are independent of each other, if we do not know in advance the set of alternatives, social welfare function, preference function and pair. Therefore we can treat those formulas as atomic propositions in proposition logic. □

5.2 Axiomatization of $\mathcal{L}_2$

We define ϕ to be valid in $\mathcal{L}_2$, written as $\models \phi$, if and only if $(A, F, \delta) \models \phi$ for all sets of alternatives A, social welfare functions F and preference functions δ.

We propose the following axiomatization for $\mathcal{L}_2$:

Axioms

$$\vdash \boxdot\phi \text{ , where } \phi \text{ is a tautology in proposition logic.} \tag{A1}$$

$$\vdash \neg \boxdot \phi\text{, where } \neg\phi \text{ is satisfiable in proposition logic.} \tag{A2}$$

Inference rules

$$\frac{\vdash \phi \quad \vdash \psi}{\vdash \phi \wedge \psi} \tag{R1}$$

$$\frac{\vdash \neg\phi}{\vdash \neg(\phi \wedge \psi)} \tag{R2}$$

$$\frac{\vdash \neg\psi}{\vdash \neg(\phi \wedge \psi)} \tag{R3}$$

$$\frac{\vdash \phi}{\vdash \neg\neg\phi} \tag{R4}$$

Theorem 1. *This axiomatization is sound and complete.*

Proof. This follows from lemmas 2 and 3. □

Lemma 2. *The axiomatization is sound (if $\vdash \phi$ then $\models \phi$).*

Proof. Assume $\vdash \boxdot\phi$ where ϕ is a tautology in proposition logic. Because ϕ is a tautology in proposition logic, it holds that $\vdash \phi$ in $\mathcal{L}_3$ by the axiomatization of $\mathcal{L}_3$. This means that ϕ holds in all pairs, which in turn means that $\models \boxdot\phi$. Therefore $A1$ is sound.

Now assume that $\vdash \neg\boxdot\phi$ where $\neg\phi$ is satisfiable in proposition logic. The satisfiability means that there is a pair in which $\models \neg\phi$ in $\mathcal{L}_3$, because of the independence of atomic variables in $\mathcal{L}_3$. Now we can conclude that ϕ does not hold in all pairs: $\not\models \boxdot\phi$. This is by definition equivalent to $\models \neg \boxdot \phi$, thus we know that $A2$ is sound.

Assume $\models \phi$ and $\models \psi$. Now it holds by definition that $\models (\phi \wedge \psi)$. Therefore also $R1$ is sound.

Assume $\models \neg\phi$. Then it holds that $\not\models \phi$. This means that not both $\models \phi$ and $\models \psi$, or not $\models (\phi \wedge \psi)$. This is equivalent to $\not\models (\phi \wedge \psi)$, thus we proved soundness of $R2$. Soundness for $R3$ can be proven in the same way.

Finally assume $\models \psi$. Then we can conclude that $\not\models \neg\psi$, and $\models \neg\neg\psi$, thereby proving that $R4$ is sound. □

Lemma 3. *The axiomatization is complete (if $\models \phi$ then $\vdash \phi$).*

Proof. We prove this by induction over the structure of the formula. Assume $\models \phi$. The formula ϕ can either be in the form $\boxdot\psi$, $\psi_1 \wedge \psi_2$ or $\neg\psi$. We can subdivide the last form in $\neg \boxdot \psi$, $\neg(\chi_1 \wedge \chi_2)$ or $\neg\neg\chi$. Therefore we have two base cases $\boxdot\psi$ and $\neg \boxdot \psi$, and three induction step cases $\psi_1 \wedge \psi_2$, $\neg(\chi_1 \wedge \chi_2)$ and $\neg\neg\chi$.

Base cases: If ϕ is of the form $\boxdot\psi$, then ψ holds in $\mathcal{L}_3$ for all pairs ($\models \psi$), so by the axiomatization of this language, ψ is a tautology in proposition logic. Now we prove by $A1$ that $\vdash \boxdot\phi$.

If ϕ is of the form $\neg \boxdot \psi$, then it holds that $\not\models \boxdot\psi$, so not every pair satisfies ψ, so there is a pair which does satisfy $\neg\psi$ in $\mathcal{L}_3$. Because atomic variables in $\mathcal{L}_3$ are independent, $\neg\psi$ is also satisfiable in proposition logic. Now we can conclude by $A2$ that $\neg \boxdot \psi$ holds.

Induction step: Assume that lemma 3 holds for all subformulas of ϕ.

If ϕ is of the form $\psi_1 \wedge \psi_2$, then it holds that $\models \psi_1$ and $\models \psi_2$. By the induction hypothesis, we can conclude that $\vdash \psi_1$ and $\vdash \psi_2$ and now by $R1$ it holds that $\vdash \psi_1 \wedge \vdash \psi_2$.

If ϕ is of the form $\neg(\psi_1 \wedge \psi_2)$, then it holds that $\not\models (\psi_1 \wedge \psi_2)$ and $\models \psi_1$ and $\models \psi_2$. By the induction hypothesis, we can conclude that $\vdash \psi_1$ and $\vdash \psi_2$ and now by $R1$ it holds that $\vdash \psi_1 \wedge \psi_2$.

Finally if ϕ is of the form $\neg\neg\psi$, then it holds by definition that $\not\models \neg\psi$ and $\models \psi$. Now it holds by induction hypothesis that $\vdash \psi$. From this we can derive by $A4$ that $\vdash \neg\neg\psi$.

$\square$

5.3 Axiomatization of $\mathcal{L}$

We say that a formula ϕ is *valid* on A if $A, F \models \phi$ for all sets of alternatives A and possible social welfare functions F. A formula ϕ is *valid*, written $\models \phi$, if $A \models \phi$ for all A. Now we can ask which $\mathcal{L}$ formulas are valid.

For this language, no axiomatization is yet known. The following properties, which are proved in (Ågotnes et al., 2006), might be useful starting points in finding an axiomatization:

$\models \square \boxdot \phi \Leftrightarrow \models \phi$
$\models \square \diamonddot q$ if q is a literal
$\not\models \square \diamonddot (q_1 \wedge q_2)$ if q_1 and q_2 are different literals

6 Conclusion

In this paper, we reviewed a logic for reasoning about social welfare functions (SWFs). This logic has three sublanguages: $\mathcal{L}$ for reasoning about SWFs, $\mathcal{L}_2$ for expressing properties about profile functions and $\mathcal{L}_1$ for reasoning about pairs of alternatives. The logic is strong enough for representing higher order properties like Arrow's theorem.

We tried to find an axiomatization for the logic. We saw an axiomatization for $\mathcal{L}_3$ and $\mathcal{L}_2$, but could not come up with one for $\mathcal{L}_3$.

Although it seems hard to find an axiomatization, I do recommend others to continue on this problem. I believe an axiomatization for this system will be a breakthrough in Computational Social Choice, because it makes many propositions in the field much easier to prove. However, without an axiomatization the system does not seem to have much added value. It would also be interesting to extend the system in such a way that it also works when two alternatives are ranked equal. Finally, we will get more insight in the system if we find some properties of SWFs which are not expressible in the language.

References

[1] T. Ågotnes, W. van der Hoek, and M. Wooldridge. Towards a Logic of Social Welfare. Proc. COMSOC-2006, University of Amsterdam, 2006.