Graz Economics Papers – GEP

GEP 2016–04

The birth of social choice theory from the spirit of mathematical logic: Arrow's theorem in the framework of model theory

Daniel Eckert, Frederik Herzberg

March 2016

Department of Economics
Department of Public Economics
University of Graz

The birth of social choice theory from the spirit of mathematical logic: Arrow's theorem in the framework of model theory

Daniel Eckert* and Frederik Herzberg†

Abstract

Arrow's axiomatic foundation of social choice theory can be understood as an application of Tarski's methodology of the deductive sciences which is closely related to the latter's foundational contribution to model theory. In this note we show in a model-theoretic framework how his use of von Neumann and Morgenstern's concept of winning coalitions allows to exploit the algebraic structures involved in preference aggregation and provides an alternative indirect ultrafilter proof for Arrow's dictatorship result. This link also connects Arrows seminal result to key developments and concepts in the history of model theory, notably ultraproducts and model-preservation results.

Keywords: Arrow's theorem; model theory; winning coalitions; ultrafilter; ultraproducts; Boolean algebra; homomorphism

2010 Mathematics Subject Classification:
03C98, 91B14, 01A60, 03C20, 03G05

1 Introduction

By generalizing the classical Arrovian problem of preference aggregation, the recent literature on judgment aggregation (surveyed in List, Puppe and Polak (15) and (16)) has established so close a relation between logic and Arrovian aggregation theory that it has even been labelled "logical aggregation" in another recent survey (20).

A model of judgment aggregation typically consists in a set of logically interconnected propositions (the *agenda*) expressed in a formal language and on which a group of individuals has to reach a decision with the help of a rule that aggregates the individual evaluations of these propositions, subject to the satisfaction of normatively desirable properties inspired from social choice theory, as the Pareto property and some independence condition — the Pareto property being a particularly good example of a condition whose normative appeal carries over from a welfare economic to an epistemic context.

*Institute of Public Economics, University of Graz

†Center for Mathematical Economics, Bielefeld University and Munich Center for Mathematical Philosophy, Ludwig Maximilian University of Munich

Thus, the reformulation of Arrow's theorem in such a framework is often used as a benchmark for the model of judgment aggregation at hand.

This historical and expository note is written in the spirit of a rational-reconstruction approach to the history of economic thought (cf. e.g. (31)). We provide a reconstruction of a version of Arrow's theorem in a model-theoretic framework, in order to argue that this logical framework is not only consistent with Arrow's original research program, but even congenial: One can make the far-reaching claim that it was the methodology of the founding father of model theory, Alfred Tarski, that allowed Arrow to appropriate the concept of winning coalitions introduced in von Neumann and Morgenstern's seminal work (23) in a way that made his famous impossibility result a model-theoretic preservation result *avant la lettre*, i.e. before model theory started to grow to its maturity in the later 50es.

Roughly speaking, this significance consists in the formulation of the problem of preference aggregation as a preservation problem that is typical for a certain branch of model theory: On the one hand, the problem which Arrow's theorem answers in the negative is whether it is possible to preserve, at the "aggregate" or product level, those properties of the individual factor models that qualify them as "rational" preferences. And on the other hand, the preservation of first-order properties of the factor models under product formation is a core problem in the subsequent literature on model theory in the 60s and 70s (cf. e.g. (6))). In particular, Arrow's use of families of decisive coalitions anticipates the model-theoretic ultraproduct construction, as was already noted by Hodges (11) in his historical account of model theory.

Recently, Herzberg and Eckert (cf. e.g. (10)) have proposed a unified framework for aggregation theory (including judgment aggregation) based on the aggregation of model-theoretic structures, thus extending Lauwers and Van Liedekerke's (14) model-theoretic analysis of preference aggregation. This model-theoretic framework for aggregation theory conceives of an aggregation rule as a map $f : \mathrm{dom}(f) \to \Omega$ with $\mathrm{dom}(f) \subseteq \Omega^I$, wherein I is a set of individuals, the *electorate*, and Ω is the collection of all models of some fixed universal theory T (in a first-order language $\mathcal{L}$) with a fixed domain A. This map thus assigns to any direct product, or *profile*, of models of T an $\mathcal{L}$-structure that is also a model of T. Thus, in model-theoretic terms, an aggregation rule is equivalent to an operation on a product of models of some theory T that solves a so-called preservation problem by guaranteeing that the outcome of this operation is again a model of T, i.e. that all the properties of the factor models described by the theory T are preserved.

The fundamental observation in the model-theoretic analysis of aggregation is that the preservation of certain properties of the individual factor models together with a) some logical richness of the agenda in terms of the logical interconnections of the propositions involved and b) some normative conditions on the aggregation rule requires that this operation be some reduction of the direct product taken over a family of subsets of the electorate. Once this observation has been made, the proof of characterisations of aggregation rules (in the guise of (im)possibility theorems) only requires relatively basic facts from model theory, such as the construction of reduced products, ultraproducts, Łoś's theorem, and the properties of filters and ultrafilters on finite sets. Dictatorship then immediately follows in the finite case, if the aforementioned family of *coalitions* (subsets of the electorate) can be shown to be an ultrafilter, because an ultrafilter over a finite set is the collection of all supersets of some singleton – viz. the singleton containing the dictator.

While the significance of ultrafilters for proofs of Arrow's impossibility result has been well understood since the 70es (cf. (13) in particular, following (8) and (9)), the important role of model theory in understanding Arrovian aggregation problems has only been exploited in the 90es by Lauwers and van Liedekerke (14) – antedated by a working paper by Brown (4), which unfortunately remained unpublished and largely unknown.

Thus, the purpose of our reconstruction of the Arrovian problem of preference aggregation in the framework of model theory is to show how Tarski's methodological influence might have helped Arrow in appropriating von Neumann and Morgensterns concept of winning coalitions in view of the use of their algebraic properties. Moreover, this reconstruction reveals the algebraic structures which underlie this problem and thereby allows the derivation of an algebraic factorization result — which in turn can be used for a further indirect ultrafilter proof of Arrow's famous impossibility theorem.

2 Historical and methodological motivation

From a methodological point of view, Arrow's seminal 1951 monograph *Social Choice and Individual Values* (2) is rightly famous for its introduction of the axiomatic analysis of binary relations into economics – and welfare economics in particular (28). While the context of justification of this approach to the modelling of social welfare is the so-called ordinalist revolution of the 1930s, which put into question the measurability and, a fortiori, the interpersonal comparison of utilities ((2), p. 9)[1], its context of discovery is Arrow's exposure as a student to the work of the famous logician Alfred Tarski, in particular to the algebra of relations in the 1940s (for details to this aspect of Arrow's intellectual biography, cf. (12), pp. 43-47). Tarski's inspiration fell on fruitful grounds, as Arrow admits having been "fascinated by mathematical logic (...) as an undergraduate and even while in high school" ((1), p. 1).

Thus, Arrow explicitly motivates the formal framework of binary relations used for the representation of preferences by its familiarity "in mathematics and particularly in symbolic logic" ((2), p. 13), referring to Tarski's famous *Introduction to Logic and the Methodology of the Deductive Sciences* (30), which he had proofread as a student.

More generally, Arrow's analysis of the problem of preference aggregation can be read as a straightforward application of the deductive method exposed in Tarski's textbook. Central to Tarski's concept of a deductive theory is not only its derivation from a set of axioms, but the concept of a model of a theory obtained by an interpretation of its terms that makes all the axioms (and thus the theory derived

[1] For the relation of Arrovian social choice theory to the tradition of welfare economics, cf. (29).

from them) true.[2]

Arrow's representation of preferences by binary relations follows exactly the format prescribed in the section on the "Model and interpretation of a deductive theory" of Tarski's textbook: "Rational" preferences are introduced as consisting in binary relations on a set of alternatives that satisfy the two axioms of weak orderings, completeness and transitivity (Axiom I and II in (2), p. 13), i.e. in Tarski's terminology as models or realizations of the axiom system of the theory of weak orderings ((2), p. 13f.).

Thus, the central problem of Arrovian social choice theory – which is the characterization of a social welfare function which aggregates individual into social preferences subject to the satisfaction of normatively desirable conditions (such as the Pareto property and independence of irrelevant alternatives) – becomes the problem of preserving these first-order properties of weak orderings in the construction of "an ordering relation for society as a whole" that "may also be assumed to satisfy Axiom I and II" ((2), p. 19).

While the construction of various types of products with the help of families of sets on some index set would later play a central role in model theory (in particular in Łoś's (17) fundamental theorem on ultraproducts), Arrow's use of families of "decisive sets of individuals" for the formalization of this problem is inspired by von Neumann and Morgenstern's concept of winning coalitions in their foundational *Theory of Games and Economic Behavior* (23), to which Arrow often refers (cf. e.g. (2), p. 59, fn. 1).

However, this approach is perfectly consistent with the semantic spirit of model theory as it has its roots in another early semantical approach to logics. It was the mathematician Karl Menger (18), (19) who first introduced families of subsets of individuals into the logical analysis of norms, extensionally conceiving a norm as the set of individuals accepting it (for a modern reconstruction of Menger's deontic logic, cf. (24)). Precisely this semantical approach was then explicitly propagated by Morgenstern in his programmatic paper *Logistik und Sozialwissenschaften* (21) as a model for the application of mathematical logic to the social sciences in general and to economics in particular. This semantical approach can thus be even considered a major source of inspiration for the *Theory of Games and Economic Behavior*, in which winning coalitions, i.e. sets of individuals play a central role.

As we will see, it is the analysis of the algebraic properties of these families of winning coalitions that will make the construction of the social ordering relation (or social welfare function) equivalent to the model-theoretic ultraproduct construction. Thus, from an interpretational point of view, our reconstruction highlights the close connection between the logical structure and the social structure involved in problems of social choice.

[2]This semantical approach can be seen not only as the conceptual intuition underlying the further development of model theory (cf. e.g. (25)). It also accounts for the significance of model theory for the epistemological analysis of those social sciences that can be subsumed under the formal sciences, like theoretical economics (27).

3 A model-theoretic view on Arrow's theorem

In model-theoretic terms a binary relation constitutes the simplest relational structure, viz. an ordered pair $\mathfrak{A} = \langle A, R\rangle$ consisting of a fixed domain A and the binary relation R on it. In general model-theoretic terms a relational structure is defined by fixing an arbitrary set A, and letting $\mathcal{L}$ be a language consisting of constant symbols for all elements a of A as well as (at most countably many) predicate symbols P_n, $n \in \mathbf{N}$, where, for all $n \in \mathbf{N}$, the arity of P_n is denoted by $\delta(n)$. A relational structure $\mathfrak{A} = \langle A, \langle R_n : n \in \mathbf{N}\rangle\rangle$ is then called a *realisation of* $\mathcal{L}$ *with domain* A or simply an $\mathcal{L}$*-structure with domain* A if and only if the arities of the relations R_n correspond to the arities of the predicate symbols P_n and the relations are evaluated in A, that is, if $R_n \subseteq A^{\delta(n)}$ for each n.

An $\mathcal{L}$-structure $\mathfrak{A}$ is a *model* of the theory T if $\mathfrak{A} \models \varphi$ for all $\varphi \in T$, i.e. if all sentences of the theory hold true in $\mathfrak{A}$ (with the usual Tarskian definition of truth).

In the framework of Arrovian preference aggregation, the language $\mathcal{L}$ has just a single, binary relation symbol R, A is interpreted as the set of alternatives, and T is the theory of weak orderings, which Arrow expresses (as a student of Tarski in several respects) by the following universal sentences:

(i) $\forall x \forall y\ R(x,y) \vee R(y,x)$ (completeness, Axiom I in (2)) and

(ii) $\forall x \forall y \forall z\ R(x,y) \wedge R(y,z) \rightarrow R(x,z)$ (transitivity, Axiom II in (2)).

However, Arrow is primarily interested in relations that express the intuition of strict preference. It is in terms of strict preference that the central property of non-dictatorship is formulated, a dictator being defined as an individual whose strict preference determines the social ranking of any two alternatives. Consequently Arrow explicitly reserves the term 'preference relation' to the linear ordering which constitutes the asymmetric part of R and which he denotes by P.

Thus, we will use the equivalent axiomatization in terms of the asymmetric part P of R:

(i') $\forall x \forall y\ P(x,y) \rightarrow \neg P(y,x)$ (asymmetry) and

(ii') $\forall x \forall y \forall z\ \neg P(x,y) \wedge \neg P(y,z) \rightarrow \neg P(x,z)$ (negative transitivity)

A relation R satisfying properties (i) and (ii) can be defined by

$\forall x \forall y\ R(x,y) :\Leftrightarrow \neg P(y,x)$ as Arrow clarifies together with some further implications, as in particular the transitivity of P (Lemma 1 ((2), p. 14).

For our expository purposes and without great loss of generality, we will restrict our attention to linear orders, i.e. to weak orders that satisfy the additional property

(iii') $\forall x \forall y\ P(x,y) \vee P(y,x) \vee x = y$.

Thus, in our model-theoretic reconstruction the (unique) binary predicate symbol P denotes strict preference.[3]

Let $\mathcal{S}$ be the set of atomic formulae in $\mathcal{L}$, and let $\mathcal{T}$ be the *boolean closure* of $\mathcal{S}$, i.e. the closure of $\mathcal{S}$ under the logical connectives $\neg$, $\wedge$, $\vee$.

[3] Assumption (iii') is not only customarily used for convenience. In its absence, Arrow's dictatorship result no longer provides a characterization of social welfare functions satisfying the normatively desirable core conditions of Pareto and independence of irrelevant alternatives (as dictatorship is only defined in terms of the prevalence of the dictator's strict preferences, there exist dictatorial social welfare functions which do not satisfy independence of irrelevant alternatives, cf. e.g. (22)).

Obviously, in our case, $\mathcal{S} = \{P(x,y) : x, y \in A\}$ and T is the set of $\mathcal{L}$-sentences which axiomatize the class of linear orders, i.e. the set of sentences (i')–(iii').

Denote by Ω the set of all models $\mathfrak{A} = \langle A, P\rangle$ of the theory of linear orderings T with domain A and by I a (possibly infinite) set of at least two individuals. Elements of I will be called *individuals*, the direct (cartesian) products $\prod_{i\in I}\mathfrak{A}_i$ which are the elements of Ω^I will be called *profiles* and will be denoted by $\underline{\mathfrak{A}}$. Thus, Ω^I represents the set of all logically possible profiles of preferences.

Having fixed A and $\mathcal{L}$ once and for all, we shall sometimes identify an $\mathcal{L}$-structure $\langle A, R\rangle$ with the asymmetric part P of the binary relation R.

A *social welfare function* is then a map $f : \Omega^I \to \Omega$ which assigns to every profile of individual preferences a social preference. (Observe that this definition already implies the usual assumption of *universal domain.*)

As the concept of "decisive coalitions", which is central to Arrow's analysis of social welfare functions, is so clearly inspired by von Neumann and Morgenstern's concept of "winning coalitions", the latter will be consistently used for definitional purposes.

Definition 1 *A coalition $C \in \mathcal{P}(I)$ is said to be* ***winning*** *for the pair of distinct alternatives $x, y \in A$ under the social welfare function $f : \Omega^I \to \Omega$, if for all profiles $\underline{\mathfrak{A}} \in \Omega^I$*

$$\{i \in I : \mathfrak{A}_i \models P(x,y)\} = C \Rightarrow f(\underline{\mathfrak{A}}) \models P(x,y).$$

Winning coalitions can be used to express other properties of social welfare functions. In particular, it is well known that the property of independence of irrelevant alternatives plays a central role in Arrow's theorem (cf. e.g. (5)).

Definition 2 *A social welfare function $f : \Omega^I \to \Omega$ satisfies* ***independence of irrelevant alternatives*** *if for any pair of distinct alternatives $x, y \in A$ and all profiles $\underline{\mathfrak{A}}, \underline{\mathfrak{A}}' \in \Omega^I$,*

$$\{i \in I : \mathfrak{A}_i \models P(x,y)\} = \{i \in I : \mathfrak{A}'_i \models P(x,y)\}$$
$$\Rightarrow \left[f(\underline{\mathfrak{A}}) \models P(x,y) \Leftrightarrow f(\underline{\mathfrak{A}}') \models P(x,y)\right].$$

Obviously, there exists a close connection between the families of winning coalitions and the condition of independence of irrelevant alternatives, which is captured in the following proposition:

Proposition 3 *A social welfare function $f : \Omega^I \to \Omega$ satisfies independence of irrelevant alternatives if and only if for any pair of distinct alternatives $x, y \in A$ there exists a family of winning coalitions $\mathcal{W}^f_{(x,y)} \in \mathcal{P}(I)$ such that for any profile $\underline{\mathfrak{A}} \in \Omega^I$*

$$f(\underline{\mathfrak{A}}) \models P(x,y) \Leftrightarrow \{i \in I : \mathfrak{A}_i \models P(x,y)\} \in \mathcal{W}^f_{(x,y)}$$

Proof. Suppose f satisfies independence of irrelevant alternatives and let x, y be a pair of distinct alternatives. For all $\underline{\mathfrak{A}} \in \Omega^I$, define

$$C\left(\underline{\mathfrak{A}}, x, y\right) = \{i \in I : \mathfrak{A}_i \models P(x,y)\},$$

and let

$$\mathcal{C}^f_{(x,y)} = \left\{ C\left(\underline{\mathfrak{A}}, x, y\right) : \underline{\mathfrak{A}} \in \Omega^I \right\}.$$

Next let $g^f_{(x,y)} : \mathcal{C}^f_{(x,y)} \to \{0,1\}$ be such that

$$g^f_{(x,y)} : C\left(\underline{\mathfrak{A}}, x, y\right) \mapsto \begin{cases} 1, & f(\underline{\mathfrak{A}}) \models P(x,y) \\ 0, & f(\underline{\mathfrak{A}}) \models \neg P(x,y). \end{cases}$$

Since f satisfies independence of irrelevant alternatives, $g^f_{(x,y)}$ is well-defined. If we now put $\mathcal{W}^f_{(x,y)} = {g^f_{(x,y)}}^{-1}\{1\}$, we have found a family of winning coalitions as postulated in the Proposition. ■

Given the informational equivalence of such an independent social welfare function with its families of winnings coalitions it comes as no wonder, that with non-dictatorship and the Pareto property, the two other relevant properties can be expressed in terms of winning coalitions.

Definition 4 *A social welfare function $f : \Omega^I \to \Omega$ is* ***non-dictatorial****, if and only if there does not exist an individual $k \in I$ such that for any pair of distinct alternatives $x, y \in A$,*

$$\mathcal{W}^f_{(x,y)} = \{S \subseteq I : k \in S\}.$$

(Otherwise, such an individual k is called ***dictator****.)*

Similarly, the familiar Pareto property can be formulated in the following way:

Definition 5 *A social welfare function $f : \Omega^I \to \Omega$ is* ***paretian*** *if and only if for any pair of distinct alternatives $x, y \in A$*

$$I \in \mathcal{W}^f_{(x,y)}$$

Interestingly, it is the largely uncontroversial Pareto property that will strengthen independence to a neutrality property knwon in the literature on judgement aggregation as systematicity:

Definition 6 *A social welfare function $f : \Omega^I \to \Omega$ is* ***systematic*** *if and only if for any pair of distinct alternatives $x, y \in A$ there exists a single family of winning coalitions $\mathcal{W}^f \in \mathcal{P}(I)$ such that for any profile $\underline{\mathfrak{A}} \in \Omega^I$*

$$f(\underline{\mathfrak{A}}) \models P(x,y) \Leftrightarrow \{i \in I : \mathfrak{A}_i \models P(x,y)\} \in \mathcal{W}^f$$

Proposition 7 *If a paretian social welfare function $f : \Omega^I \to \Omega$ satisfies independence of irrelevant alternatives, it is already systematic.*

This proposition follows, as is well known, from the so-called Contagion Lemma in social choice theory.

Lemma 8 ***(Contagion Lemma)*** *Let $f : \Omega^I \to \Omega$ be a paretian social welfare function which satisfies independence of irrelevant alternatives. Then for any two pairs of distinct distinct alternatives $a, b \in A$ and $x, y \in A$*

$$\mathcal{W}^f_{(a,b)} = \mathcal{W}^f_{(x,y)}.$$

Proof. Let a, b, x, y as above. It is sufficient to prove the inclusion $\mathcal{W}^f_{(x,y)} \subseteq \mathcal{W}^f_{(a,b)}$ as the converse inclusion will then follow from interchanging a, b, x, y.

Let hence $C \in \mathcal{W}^f_{(x,y)}$. Since the domain of f is the whole of Ω^I, we can construct a profile $\underline{\mathfrak{A}} \in \Omega^I$ such that
a) $\{i \in I : \mathfrak{A}_i \models P(x, y)\} = C = \{i \in I : \mathfrak{A}_i \models P(a, b)\}$ and
b) $\{i \in I : \mathfrak{A}_i \models P(a, x)\} = \{i \in I : \mathfrak{A}_i \models P(b, y)\} = I$, whence by the Pareto property and C being a winning coalition $f(\underline{\mathfrak{A}}) \models P(a, x) \wedge P(x, y) \wedge P(y, b)$. Since $f(\underline{\mathfrak{A}})$ is $\in \Omega$ and thus complete and transitive, we may deduce $f(\underline{\mathfrak{A}}) \models P(a, b)$, which in turn proves $C \in \mathcal{W}^f_{(a,b)}$. ■

Independence of irrelevant alternatives is thus technically one of the most crucial properties of Arrovian social welfare functions, but also interpretationally controversial. Interestingly, Arrow interprets this property by reference to voting, which in turn is presented as a natural valuation of the preference judgment on any pair of alternatives:

"The condition of the independence of irrelevant alternatives implies that in a generalized sense all methods of social choice are of the type of voting. If S is the set $[x, y]$ consisting of the two alternatives x and y, (it) tells us that the choice between x and y is determined solely by the preferences of the members of the community as between x and y. That is, if we know which members of the community prefer x to y, which are indifferent, and which prefer y to x, then we know what choice the community makes" ((2), p. 27f.).

In our framework this means that each profile $\underline{\mathfrak{A}} \in \Omega^I$ (considered as a mapping $\pi : I \to \Omega$) induces a mapping $t : \mathcal{T} \to \mathcal{P}(I)$ which maps each formula λ in the closure of $\mathcal{S}$ under the logical connectives $\neg$, $\wedge$, $\vee$ to the set of individuals $\{i \in I : \mathfrak{A}_i \models \lambda\}$ in the models of which it holds true. This mapping then determines the outcome of the social welfare function. In this way, the power set algebra of the set of individuals $P(I) = (\mathcal{P}(I); \cup, \cap, \complement, \varnothing, I)$, which we will call *coalition algebra*, becomes a natural algebraic valuation of $\mathcal{T}$, as it clearly preserves the logical connectives $\neg$, $\wedge$, $\vee$.[4]

However, as the Boolean closure of a set of atomic formulae is not in itself a Boolean algebra, the pivotal role of the coalition algebra is best seen with the help of an algebraic factorization result for social welfare functions. For this factorisation it is helpful to use the Lindenbaum algebra of the theory T.

Definition 9 *Let $\equiv$ denote provable equivalence given a theory T (i.e. $\phi \equiv \psi$ if and only if $T \cup \{\phi\} \vdash \psi$ and $T \cup \{\psi\} \vdash \phi$. The set of equivalence classes of $\mathcal{L}$-formulae under T is known as the Lindenbaum algebra and is denoted by $\mathcal{L}/\equiv$.*

[4] That the use of the coaliton algebra as an algebra of truth values leads to a (trivial) possibility result was first shown by (26). For the significance of boolean-valued aggregation rules, cf. (7).

First, and in view of the central role of the coalition algebra, the set Ω^I of all possible profiles (considered as mappings $\pi : I \to \Omega$) induces for any given profile $\underline{\mathfrak{A}} \in \Omega^I$ a map

$$H(\cdot, \underline{\mathfrak{A}}) : \mathcal{L}/\equiv \to P(I), \quad \langle [\lambda]_\equiv, \underline{\mathfrak{A}} \rangle \mapsto \{i \in I : \mathfrak{A}_i \models \lambda\},$$

which assigns to each class of formulae $[\lambda]_\equiv$ the coalition $\{i \in I : \mathfrak{A}_i \models \lambda\}$ of all individuals in whose models λ is satisfied. By the soundness of $\vdash$, this map is well-defined. Moreover, it is a homomorphism, as can be seen from Tarski's definition of truth.

On the other hand, it is also obvious that any social welfare function f induces for any given profile $\underline{\mathfrak{A}} \in \Omega^I$ a truth valuation as another homomorphism

$$V(\cdot, \underline{\mathfrak{A}}) : \mathcal{L}/\equiv \to \{0,1\}, \quad \langle [\lambda]_\equiv, \underline{\mathfrak{A}} \rangle \mapsto \|\lambda\|_f^{\mathfrak{A}} = \begin{cases} 1, & f(\underline{\mathfrak{A}}) \models \lambda \\ 0, & f(\underline{\mathfrak{A}}) \not\models \lambda \end{cases},$$

between the Lindenbaum algebra and the algebra of (classical) truth values $\{0,1\}$.

Consider now the binary relation h_f defined by

$$h_f(C, v) \Leftrightarrow \exists [\lambda]_\equiv \in \mathcal{L}/\equiv \quad \exists \underline{\mathfrak{A}} \in \Omega^I \quad \left(\{i \in I : \mathfrak{A}_i \models \lambda\} = C \;\&\; \|\lambda\|_f^{\mathfrak{A}} = v \right)$$

for all $C \in P(I)$ and $v \in \{0,1\}$.

If the social welfare function f is systematic, this relation will be a function $h_f : P(I) \to \{0,1\}$, and it will admit the following factorization for any $\underline{\mathfrak{A}} \in \Omega^I$:

$$\begin{array}{ccc} \mathcal{L}/\equiv & \overset{H(\cdot,\underline{\mathfrak{A}})}{\longrightarrow} & P(I) \\ {\scriptstyle V(\cdot,\underline{\mathfrak{A}})}\downarrow & \swarrow {\scriptstyle h_f} & \\ \{0,1\} & & \end{array}$$

An algebraic understanding of the coalition algebra, enables us to complete the proof of Arrow's theorem: We shall first verify that h_f is a homomorphism of Boolean algebras. We will then be able to deduce that the family of winning coalitions forms an ultrafilter[5] This is most easily done with the help of the following well-known result from Boolean algebra:

Lemma 10 *Consider the mapping $g : B_1 \to B_2$, where B_1 and B_2 are Boolean algebras. If g preserves complements, then it preserves meets if and only if it preserves joins. In that case, g is a homomorphism and the set $F = \{x \in B_1 : g(x) = 1\}$ is a filter, and even an ultrafilter if $B_2 = \{0,1\}$.*

---

[5]Recall that a *filter* on the set I is a family $\mathcal{W} \in \mathcal{P}(I)$ such that
(F1) $\mathcal{W} \neq \varnothing$ and $\varnothing \notin \mathcal{W}$ (non-triviality)
(F2) $U \cap V \in \mathcal{W}$ for all $U, V \in \mathcal{W}$ (finite intersection closure)
(F3) $V \in \mathcal{W}$ whenever $V \supseteq U$ for some $U \in \mathcal{W}$ (superset closure).

A filter is *principal* if it is generated by some non-empty set $X \subset I$, i.e. if it is of the form $\mathcal{W} = \{U \subseteq I : X \subseteq U\}$.

An *ultrafilter* $\mathcal{W}$ on the set I is a filter on I such that for any $U \subseteq I$ either $U \in \mathcal{W}$ or $I \backslash U \in \mathcal{W}$ (but not both).

The so-called Ultrafilter Theorem asserts that an ultrafilter on I is a "finest possible" filter on I, i.e. a maximal filter when ordering filters on I by set-theoretic inclusion. It is not difficult to deduce that an ultrafilter is principal if and only if there exists some $k \in I$ such that $\mathcal{W} = \{U \subseteq I : k \in I\}$.

Combining this lemma with the above factorization we thus obtain an alternative indirect proof of the ultrafilter structure of the family of winning coalitions for our social welfare function.

Theorem 11 *Let $f : \Omega^I \to \Omega$ be a paretian social welfare function which satisfies independence of irrelevant alternatives. Then the corresponding family of winning coalitions $\mathcal{W}^f$ is an ultrafilter on I.*

Proof. We first verify that h_f preserves complements:

For any $C \in P(I)$ there will be $\lambda \in \mathcal{S}$ such that $C = H([\lambda]_\equiv, \underline{\mathfrak{A}})$. This means
$h(\complement C) = h\left(\complement H([\lambda]_\equiv, \underline{\mathfrak{A}})\right)$
$= h\left(H([\neg\lambda]_\equiv, \underline{\mathfrak{A}})\right)$
$= V([\neg\lambda]_\equiv, \underline{\mathfrak{A}})$
$= V([\lambda]_\equiv, \underline{\mathfrak{A}})^*$
$= h(H([\lambda]_\equiv, \underline{\mathfrak{A}}))^*$
$= h(C)^*$

It remains to verify that h_f preserves meets:

For any $C, C' \in P(I)$, there will be $\lambda, \mu \in \mathcal{S}$ such that $C = H([\lambda]_\equiv, \underline{\mathfrak{A}})$ and $C' = H([\mu]_\equiv, \underline{\mathfrak{A}})$. This means
$h(C \cap C') = h\left(H([\lambda]_\equiv, \underline{\mathfrak{A}}) \cap H([\mu]_\equiv, \underline{\mathfrak{A}})\right)$
$= h\left(H\left([\lambda]_\equiv \wedge [\mu]_\equiv, \underline{\mathfrak{A}}\right)\right)$
$= V\left([\lambda]_\equiv \wedge [\mu]_\equiv, \underline{\mathfrak{A}}\right)$
$= V\left([\lambda]_\equiv, \underline{\mathfrak{A}}\right) \sqcap V\left([\mu]_\equiv, \underline{\mathfrak{A}}\right)$
$= h(H([\lambda]_\equiv, \underline{\mathfrak{A}})) \sqcap h\left(H\left([\mu]_\equiv, \underline{\mathfrak{A}}\right)\right)$
$= h(C) \sqcap h(C')$

By the first half of the previous lemma, we conclude that h_f preserves joins, too, and thus is a homomorphism. But $\mathcal{W}^f = {h_f}^{-1}(\{1\})$. Hence, by the second part of the previous lemma, $\mathcal{W}^f$ is indeed an ultrafilter. ■

In light of the fact that ultrafilters on a finite set are always principal, one obtains Arrow's theorem as the following simple corollary:

Corollary 12 *Let $f : \Omega^I \to \Omega$ be a paretian social welfare function which satisfies independence of irrelevant alternatives. If I is finite, then f is dictatorial.*

Now, a major model-theoretic significance of ultrafilters consists in their role in the construction of *restricted ultraproducts* $\underline{\mathfrak{A}}/\mathcal{W}$ by the reduction of a direct product $\underline{\mathfrak{A}}$ over an ultrafilter $\mathcal{W}$. Herein, reduction means that two profiles are identified in the ultraproduct if the coalition of all individuals to which both profiles assign the same preferences is an element of the ultrafilter.[6]

---

[6]More formally, for all $\underline{\mathfrak{A}}, \underline{\mathfrak{A}'}$, define a relation $\sim_\mathcal{W}$ by

$$\underline{\mathfrak{A}} \sim_\mathcal{W} \underline{\mathfrak{A}'} \Leftrightarrow \left\{i \in I \ : \ \mathfrak{A}_i = \mathfrak{A}'_i\right\} \in \mathcal{W}.$$

Since $\mathcal{W}$ is an ultrafilter, one can verify that $\sim_\mathcal{W}$ is an equivalence relation. By $\underline{\mathfrak{A}}/\mathcal{W}$ we denote the equivalence class of $\underline{\mathfrak{A}}$ under $\sim_\mathcal{W}$.

Theorem 13 *Let $f : \Omega^I \to \Omega$ be a paretian social welfare function which satisfies independence of irrelevant alternatives. Then there exists some ultrafilter $\mathcal{W}$ on I such that for any profile $\underline{\mathfrak{A}} \in \Omega^I$, $f(\underline{\mathfrak{A}}) = \underline{\mathfrak{A}}/\mathcal{W}$.*

Proof. Let $\mathcal{W}$ be the coalition of winning coalition under f. We have previously shown it to be a filter. Combining the definition of systematicity with the definition of an ultrafilter, we obtain for every $\underline{\mathfrak{A}} \in \Omega^I$ and all $x, y \in A$,

$$f(\underline{\mathfrak{A}}) \models P(x,y) \Leftrightarrow \{i \in I : \mathfrak{A}_i \models P(x,y)\} \in \mathcal{W}^f \Leftrightarrow \underline{\mathfrak{A}}/\mathcal{W}^f \models P(x,y),$$

which completes the proof. ■

Now, an important aspect of the use of ultraproducts in model theory is that they yield, through Łoś's theorem, so-called preservation results, i.e. results about the preservation of first-order properties of relational structures (through Łoś's theorem). Arrow's theorem is itself a result about the possibility of preserving such properties in the presence of other conditions: One of its key assumptions is that the range of f consists of structures which satisfy the theory T. This is reflected by our derivation of Arrow's theorem as well: Its pivotal part was the Contagion Lemma, whose proof crucially rested on the assumption that $f(\underline{\mathfrak{A}})$ is a model of T for every $\underline{\mathfrak{A}} \in \Omega^I$.

One can compare Arrow's impossibility theorem with Fishburn's possibility theorem (8) by noting that the former is a negative model preservation result, while the latter is a positive one. In model-theoretic terms, Fishburn's possibility theorem becomes a converse of the previous theorem:

Theorem 14 *Let $\mathcal{W}$ be an ultrafilter on I and define a mapping $f : \Omega^I \to A \times A$ by $f(\underline{\mathfrak{A}}) = \underline{\mathfrak{A}}/\mathcal{W}$ for all $\underline{\mathfrak{A}} \in \Omega^I$. Then $f : \Omega^I \to \Omega$ is a paretian social welfare function which satisfies independence of irrelevant alternatives. If I is infinite, f can be chosen non-dictatorial.*

Proof. Let $\underline{\mathfrak{A}} \in \Omega^I$. Since each $\mathfrak{A}_i$ is a model of T and T is universal, $f(\underline{\mathfrak{A}}) = \underline{\mathfrak{A}}/\mathcal{W}$ must be a model of T, too. Therefore, $f(\underline{\mathfrak{A}}) \in \Omega$ for every $\underline{\mathfrak{A}} \in \Omega^I$. Verifying the Pareto property and independence of irrelevant alternatives is straightforward – as is the demonstration of non-dictatorship for non-principal $\mathcal{W}$. ■

4 Conclusion

Our reconstruction of the Arrovian problem of preference aggregation in the framework of model theory shows how Tarski's methodological influence might have helped Arrow in appropriating von Neumann and Morgenstern's concept of winning coalitions in a way that allows to exploit the algebraic structures involved in this problem. The algebraic factorization result that can thus be obtained provides another, indirect ultrafilter proof of his famous dictatorship result. Moreover, the equivalence of an Arrovian social welfare function with an ultraproduct construction gives his result the flavour of a model-theoretic preservation result *avant la lettre.*

References

[1] K. Arrow. *Collected Papers of K.J. Arrow: Social Choice and Justice.* The Belknap Press of Harvard University Press, 1983.

[2] K.J. Arrow. *Social choice and individual values. 2nd ed.* Cowles Commission Monograph. 12. New York, NY: Wiley, 1963.

[3] J.L. Bell and A.B. Slomson. *Models and ultraproducts. An introduction.* North Holland, Amsterdam, 1969.

[4] D.J. Brown. Collective rationality. Technical report, Cowles Foundation for Research in Economics, 1975.

[5] D. Campbell and J. Kelly. Impossibiiity theorems in the arrovian framework. In K.J. Arrow, A. Sen, and K. Suzumura, editors, *Handbook of Social Choice and Welfare*, volume 1, chapter 1, pages 35–94. Amsterdam: North-Holland, 2002.

[6] C.C. Chang and H. Jerome Keisler. *Model Theory.* Amsterdam: North-Holland, 1990.

[7] D. Eckert and F. Herzberg. The problem of judgment aggregation in the framework of boolean-valued models. In N. Bulling, L. van der Torre, S. Villata, W. Jamroga, and W. Vasconcelos, editors, *Computational Logic in Multi-Agent Systems. CLIMA XV, August 2014, Proceedings*, volume 8624 of *Lecture Notes in Computer Science*, pages 138–147, 2014.

[8] P.C. Fishburn. Arrow's impossibility theorem: concise proof and infinite voters. *Journal of Economic Theory*, 2(1):103–106, 1970.

[9] B. Hansson. The existence of group preference functions. *Public Choice*, 38(1):98–98, 1976.

[10] F. Herzberg and D. Eckert. The model-theoretic approach to aggregation: Impossibility results for finite and infinite electorates. *Mathematical Social Sciences*, 64:41–47, 2012.

[11] W. Hodges. Model theory. Technical report, Queen Mary, University of London, July 2000. http://wilfridhodges.co.uk/history07.pdf.

[12] J.S. Kelly. An interview with Kenneth J. Arrow. *Social Choice and Welfare*, 4:43–62, 1987.

[13] A.P. Kirman and D. Sondermann. Arrow's theorem, many agents, and invisible dictators. *Journal of Economic Theory*, 5(2):267–277, 1972.

[14] L. Lauwers and L. Van Liedekerke. Ultraproducts and aggregation. *Journal of Mathematical Economics*, 24(3):217–237, 1995.

[15] C. List and B. Polak. Introduction to judgment aggregation. *Journal of Economic Theory*, 145(2):441–466, 2010.

[16] C. List and C. Puppe. Judgment aggregation: A survey. In P. Anand, P.K. Pattanaik, and C. Puppe, editors, *The Handbook of Rational and Social*

Choice: An Overview of New Foundations and Applications, pages 457–482. Oxford, Oxford University Press, 2009.

[17] J. Łoś. Quelques remarques, théorèmes et problèmes sur les classes définissables d'algèbres. In Th. Skolem, G. Hasenjaeger, G. Kreisel, A. Robinson, H. Wang, L. Henkin, and J. Łoś, editors, *Mathematical interpretation of formal systems*, volume 16 of *Studies in Logic and the Foundations of Mathematics*, pages 98–113. Amsterdam: North-Holland, Amsterdam, 1955.

[18] K. Menger. *Moral, Wille und Weltgestaltung. Grundlegung zur Logik der Sitten.* Springer, 1934.

[19] K. Menger. Ein Satz über endliche Mengen mit Anwendungen auf die formale Ethik (77. Kolloquium, 30.V.1934). In E. Dierker and K. Sigmund, editors, *Ergebnisse eines Mathematischen Kolloquiums*, pages 23–26. Springer: Wien New York, 1998.

[20] Mongin. The doctrinal paradox, the discursive dilemma, and logical aggregation theory. *Theory and Decision*, 73(3):315–355, 2012.

[21] O. Morgenstern. Logistik und sozialwissenschaften. *Zeitschrift für Nationalökonomie*, 7(1):1–24, 1936.

[22] A. Mossel and O. Tamuz. Complete characterization of functions satisfying the conditions of arrow's theorem. *Social Choice and Welfare*, 39(1):127–140, 2012.

[23] J. v. Neumann and O. Morgenstern. *Theory of Games and Economic Behavior.* Princeton University Press, 1944.

[24] A. Siegesleitner and H. Leitgeb. Mengers Logik für Ethik und Moral: Nichts von Sollen, nichts von Güte, nichts von Sinnlosigkeit. In *Logischer Empirismus, Werte und Moral*, volume 15 of *Veröffentlichungen des Instituts Wiener Kreis*, pages 197–218. Wien: Springer, 2010.

[25] K. Simmons. Tarski's logic. In D.M. Gabbay and J. Woods, editors, *Handbook of the History of Logic*, volume 5, pages 511–616. Elsevier, 2009.

[26] H.J. Skala. Arrow's impossibility theorem: Some new aspects. In H. Gottinger and W. Leinfellner, editors, *Decision theory and social ethics.* Reidel, Dordrecht, 1978.

[27] B.P. Stigum. *Toward a Formal Science of Economics: The Axiomatic Method in Economics and Econometrics.* The MIT Press, 1990.

[28] P. Suppes. The pre-history of Kenneth Arrow's social choice and individual values. *Social Choice and Welfare*, 25(2-3):319–326, 2005.

[29] K. Suzumura. Introduction. In K. Arrow, A. Sen, and K. Suzumura, editors, *Handbook of Social Choice and Welfare*, volume 1-32. Elsevier, 2002.

[30] A. Tarski. *Introduction to Logic and the Methodology of the Deductive Sciences.* Oxford University Press, 1941.

[31] A.M.C. Waterman. Mathematical modelling as an exegetical tool: Rational reconstruction. In W.J. Samuels, J.E. Biddle, and J.B. Davis, editors, *A companion to the History of Economic Thought*, pages 553–570. Blackwell, 2007.

Graz Economics Papers

For full list see:
http://ideas.repec.org/s/grz/wpaper.html
Address: Department of Economics, University of Graz,
Universitätsstraße 15/F4, A-8010 Graz

04–2016	**Daniel Eckert, Frederik Herzberg**: The birth of social choice theory from the spirit of mathematical logic: Arrow's theorem in the framework of model theory
03–2016	**Christoph Kuzmics, Jan-Henrik Steg**: On public good provision mechanisms with dominant strategies and balanced budget
02–2016	**Florian Gauer, Christoph Kuzmics**: Cognitive Empathy in Conflict Situations
01–2016	**Thomas Aronsson, Olof Johansson-Stenman, Ronald Wendner**: Redistribution Through Charity, and Optimal Taxation when People are Concerned with Social Status
09–2015	**Sofie R. Waltl**: Variation across price segments and locations: A comprehensive quantile regression analysis of the Sydney housing market
08–2015	**Karl Farmer, Bogdan Mihaiescu**: Pension systems and financial constraints in a three-country OLG model of intra-EMU and global trade imbalances
07–2015	**Javier López Prol, Karl W. Steininger**: Photovoltaic self-consumption regulation in Spain: profitability analysis and alternative regulation schemes.
06–2015	**Miriam Steurer, Caroline Bayr**: Quantifying Urban Sprawl using Land Use Data
05–2015	**Porfirio Guevara, Robert J. Hill, Michael Scholz**: Hedonic Indexes for Public and Private Housing in Costa Rica
04–2015	**Sofie R. Waltl**: A Hedonic House Price Index in Continuous Time
03–2015	**Ronald Wendner**: Do Positional Preferences for Wealth and Consumption Cause Inter-Temporal Distortions?
02–2015	**Maximilian Gödl, Christoph Zwick**: Stochastic Stability of Public Debt: The Case of Austria

01–2015 **Gabriel Bachner**: Land Transport Systems under Climate Change: A Macroeconomic Assessment of Adaptation Measures for the Case of Austria

10–2014 **Thomas Schinko, Judith Köberl, Birgit Bednar-Friedl, Franz Prettenthaler, Christoph Töglhofer, Georg Heinrich, Andreas Gobiet**: An Integrated Economic Analysis of Climate Change Impacts on Tourism in Austria Across Tourism Types and Seasons

09–2014 **Sugata Ghosh, Ronald Wendner**: Positional Preferences, Endogenous Growth, and Optimal Income- and Consumption Taxation

08–2014 **Stan Metcalfe**: George Shackle and The Schumpeterian Legacy.

07–2014 **Birgit Bednar–Friedl, Karl Farmer**: Existence and efficiency of stationary states in a renewable resource based OLG model with different harvest costs

06–2014 **Karl Farmer, Irina Ban**: Modeling financial integration, intra-EMU and Asian-US external imbalances

05–2014 **Robert J. Hill, Michael Scholz**: Incorporating Geospatial Data in House Price Indexes: A Hedonic Imputation Approach with Splines

04–2014 **Y. Hossein Farzin, Ronald Wendner**: The Time Path of the Saving Rate: Hyperbolic Discounting and Short-Term Planning

03–2014 **Robert J. Hill, Iqbal A. Syed**: Hedonic Price-Rent Ratios, User Cost, and Departures from Equilibrium in the Housing Market

02–2014 **Christian Gehrke**: Ricardo's Discovery of Comparative Advantage Revisited

01–2014 **Sabine Herrmann, Jörn Kleinert**: Lucas Paradox and Allocation Puzzle – Is the euro area different?

08–2013 **Christoph Zwick**: Current Account Adjustment in the Euro-Zone: Lessons from a Flexible-Price-Model

07–2013 **Karl Farmer**: Financial Integration and EMUs External Imbalances in a Two-Country OLG Model

06–2013 **Caroline Bayr, Miriam Steurer, Rose-Gerd Koboltschnig**: Scenario Planning for Cities using Cellular Automata Models: A Case Study

05–2013 **Y. Hossein Farzin, Ronald Wendner**: Saving Rate Dynamics in the Neoclassical Growth Model – Hyperbolic Discounting and Observational Equivalence

04–2013 **Maximilian Gödl, Jörn Kleinert**: Interest rate spreads in the Euro area: fundamentals or sentiments?

03–2013 **Christian Lininger**: Consumption-Based Approaches in International Climate Policy: An Analytical Evaluation of the Implications for Cost-Effectiveness, Carbon Leakage, and the International Income Distribution

02–2013 **Veronika Kulmer**: Promoting alternative, environmentally friendly passenger transport technologies: Directed technological change in a bottom-up/top-down CGE model

01–2013 **Paul Eckerstorfer, Ronald Wendner**: Asymmetric and Non-atmospheric Consumption Externalities, and Efficient Consumption Taxation

10–2012 **Michael Scholz, Stefan Sperlich, Jens Perch Nielsen**: Nonparametric prediction of stock returns with generated bond yields

09–2012 **Jörn Kleinert, Nico Zorell**: The export-magnification effect of offshoring

08–2012 **Robert J. Hill, Iqbal A. Syed**: Hedonic Price-Rent Ratios, User Cost, and Departures from Equilibrium in the Housing Market

07–2012 **Robert J. Hill, Iqbal A. Syed**: Accounting for Unrepresentative Products and Urban-Rural Price Differences in International Comparisons of Real Income: An Application to the Asia-Pacific Region

06–2012 **Karl Steininger, Christian Lininger, Susanne Droege, Dominic Roser, Luke Tomlinson**: Towards a Just and Cost-Effective Climate Policy: On the relevance and implications of deciding between a Production versus Consumption Based Approach

05–2012 **Miriam Steurer, Robert J. Hill, Markus Zahrnhofer, Christian Hartmann**: Modelling the Emergence of New Technologies using S-Curve Diffusion Models

04–2012 **Christian Groth, Ronald Wendner**: Embodied learning by investing and speed of convergence

03–2012 **Bettina Brüggemann, Jörn Kleinert, Esteban Prieto**: The Ideal Loan and the Patterns of Cross-Border Bank Lending

02–2012 **Michael Scholz, Jens Perch Nielsen, Stefan Sperlich**: Nonparametric prediction of stock returns guided by prior knowledge

01–2012 **Ronald Wendner**: Ramsey, Pigou, heterogenous agents, and non-atmospheric consumption externalities