Arrow's Theorem as a Corollary

Klaus Nehring
University of California, Davis

October 2002

Abstract

Arrow's Impossibility Theorem is derived from a general theorem on social aggregation in "property spaces". In the present proof, the weak-order structure of the domain and co-domain of the aggregation plays a purely combinatorial role.

Over the years, Arrow's (1963) celebrated result on the impossibility of ordinal preference aggregation has received a number of attempts at improved, more transparent and concise direct proofs; see, among others, Barbera (1980), Geanakoplos (1996) and Reny (2001). While some of these efforts have been strikingly successful in achieving what they set out to do, they still leave a tangible gap. Precisely by exploiting the specific assumptions of Arrow's Theorem to their fullest, they let it stand as an intriguing, isolated result, and do not reveal the existence of a more general underlying mathematical logic at work. By contrast, the appeal of many of the most compelling results in mathematical economics lies precisely in the basic and general nature of the underlying mathematical structure, as when for example the existence of equilibrium boils down to the application of a fixed-point theorem, or the existence of efficiency prices to a separating hyperplane theorem. The goal of this note is to contribute towards closing this gap by obtaining Arrow's Theorem as a corollary of a substantially more general, "abstract" result on the existence of non-degenerate aggregation mappings, i.e. functions from the n-fold Cartesian product of a space into itself. This abstract aggregation result has been obtained in the context of characterizing domains that admit non-dictatorial strategy-proof social choice functions on generalized single-peaked domains (Nehring-Puppe 2002, henceforth NP).[1]

Let X be a (finite) abstract set of "objects" referred to here as "views". In NP, the elements of X are most-preferred social alternatives; here, X will be interpreted as the set of all weak orders $\mathcal{R}(A)$ on the set of social alternatives A. A *property space* is a pair $(X, \mathcal{H})$, where $\mathcal{H}$ is a family of subsets of X that is closed under complementation and that separates points (i.e., for any $x \neq y \in X$, there exists $H \in \mathcal{H}$ such that $x \in H$ and $y \notin H$); a member $H \in \mathcal{H}$ can be interpreted as the set of objects possessing a particular "property" (here: as a set of "views" agreeing in some "aspect"). Property spaces are closely related to the so-called S3-convexities in abstract convexity theory; see, e.g. Van de Vel (1993).

Let $I = \{1, ..., n\}$ be the finite set of "individuals". An aggregation rule $f : X^I \rightarrow X$ maps profiles of individual views to a social view. The rule f is *unanimous* if $f(x, ..., x) = x$ for all x; f is *dictatorial* if, there exists some $i \in I$ such that $f(x_1, ..., x_n) = f(y_1, ..., y_n)$ whenever $x_i = y_i$. Finally,

[1] In a somewhat congenial spirit, recent work by Sethuraman-Teo-Vohra (2001) derives Arrow's Theorem and a number of related results using the technique of integer prgramming.

f is *monotone* if, for all $H \in \mathcal{H}$, all profiles $(x_1, .., x_i, .., x_n)$, individuals $i \in I$ and views $y_i \in X$ such that $y_i \in H$ if $x_i \in H$, $f(x_1, ..., x_n) \in H$ implies $f(x_1, .., y_i, .., x_n) \in H$. In words: if some of i's views change, but if i continues to support some aspect H of the social view, then the social view does not change in this aspect.

To characterize the property spaces that admit non-dictatorial, unanimous and monotone aggregation rules, some auxiliary terminology is needed. A *critical family* is a minimal subset $\mathcal{G} = \{H_1, ..., H_m\} \subseteq \mathcal{H}$ with empty intersection. Let Γ denote the set of all critical families in $(X, \mathcal{H})$. Let $H \geq G$ iff there exists a critical family $\mathcal{G}$ containing H and G^c; let $\geq^*$ denote the transitive closure of $\geq$.[2] The property space $(X, \mathcal{H})$ is *totally blocked* if $\geq^*$ is total, i.e. if, for all $G, H \in \mathcal{H}, G \geq^* H$. The following result is proved in NP (Theorem 5).

Theorem 1 *The property space $(X, \mathcal{H})$ admits a non-dictatorial, unanimous and monotone aggregation rule if and only if it is not totally blocked.*

To derive from this result a (version of) Arrow's Theorem, the set of all weak orders on A must be endowed with the structure of a property space. Let R denote the generic weak order on A, with asymmetric component P. "Aspects" of weak orders are naturally defined in terms of preferences over given pairs. Thus, for $a, b \in A$, let $H_{ab} := \{R \in \mathcal{R}(A) : aRb\}$. The canonical property space of weak orders $(\mathcal{R}(A), \mathcal{H}(\mathcal{R}(A)))$ is given by

$$\mathcal{H}(\mathcal{R}(A)) := \{H_{ab}\}_{a,b \in A} \cup \{H_{ab}^c\}_{a,b \in A}. \tag{1}$$

In this context, aggregation rules f can be interpreted as social welfare functions. Unanimity and dictatoriality retain their usual meaning; f is "monotone" if and only if its satisifies "*monotone independence of irrelevant alternatives*":

> For all $a, b \in A$ and all profiles $(R_1, .., R_n)$ and $(R_1', ..., R_n')$ such that $aR_i'b$ whenever aR_ib, $af(R_1, .., R_n)b$ implies $af(R_1', .., R_n')b$.[3]

[2]For an interpretation of the relation $\geq$ in terms of "conditinional entailement", see NP, section 5.1.

[3]Note that on a universal domain of weak orders, monotone IIA is just slightly stronger than IIA, for it is equivalent to IIA plus the following restricted version of monotone IIA:

For all $a, b \in Z$ and all profiles $(R_1, .., R_n)$ and $(R_1', ..., R_n')$ such that $aR_i'b$ whenever aR_ib *and such that, for all* $c \neq a, b$ *and all* d, aP_ic, bP_ic, *and* cR_id *if and only if* $cR_i'd$ *for all* i, $af(R_1, .., R_n)b$ implies $af(R_1', .., R_n')b$.

We are not aware of any motivated social choice rule violating this condition.

We shall show that the property space $(\mathcal{R}(A), \mathcal{H}(\mathcal{R}(A)))$ is totally blocked whenever $\#A \geq 3$, thereby obtaining Arrow's Theorem as the following corrolary to Theorem 1.

Corollary 2 *If an aggregation rule $f : \mathcal{R}(A)^I \to \mathcal{R}(A)$ satisfies unanimity and monotone IIA and if $\#A \geq 3$, f must be dictatorial.*

Proof. In view of Theorem 1, it remains to verify the total blockedness of $(\mathcal{R}(A), \mathcal{H}(\mathcal{R}(A)))$ given that $\#A \geq 3$.

By the transitivity of weak orders, for any three distinct $a, b, c \in A$, the set $\{H_{ab}, H_{bc}, H^c_{ac}\}$ is a critical family. Hence, for all distinct $a, b, c \in A$,

$$H_{ab} \geq H_{ac} \text{ and } H_{bc} \geq H_{ac}, \tag{2}$$

$$H^c_{ac} \geq H^c_{ab} \text{ and } H^c_{ac} \geq H^c_{bc}, \text{ and} \tag{3}$$

$$H_{ab} \geq H^c_{bc} \text{ and } H_{bc} \geq H^c_{ab}. \tag{4}$$

Let $\mathcal{H}^+ := \{H_{ab}\}_{a \neq b \in A}$ and $\mathcal{H}^- := \{H^c_{ab}\}_{a \neq b \in A}$.

From (2), one easily infers that

$$\text{for all } H, H' \in \mathcal{H}^+\text{: } H \geq^* H'. \tag{5}$$

In particular, to obtain $H_{ab} \geq^* H_{ba}$, take any c different from a and b and note that by (2) $H_{ab} \geq H_{ac}$, $H_{ac} \geq H_{bc}$ and $H_{bc} \geq H_{ba}$. From this and (2), (5) is immediate.

Analogously, one infers from (3) that

$$\text{for all } H, H' \in \mathcal{H}^-\text{: } H \geq^* H'. \tag{6}$$

Moreover, (4) implies that, for all $H \in \mathcal{H}^+$ there exists $H' \in \mathcal{H}^-$ such that $H \geq H'$. Hence by (6),

$$\text{for all } H \in \mathcal{H}^+, H' \in \mathcal{H}^-\text{: } H \geq^* H'. \tag{7}$$

Finally, by the completeness of weak orders, for all $a \neq b$, the set $\{H^c_{ab}, H^c_{ba}\}$ is a critical family, implying that $H^c_{ab} \geq H_{ba}$. Hence by (5),

$$\text{for all } H \in \mathcal{H}^-, H' \in \mathcal{H}^+\text{: } H \geq^* H'. \tag{8}$$

The assertions (5), (6), (7) and (8) establish the total blockedness of $(\mathcal{R}(A), \mathcal{H}(\mathcal{R}(A)))$.

A beauty of the present proof of Arrow's Theorem is its clean separation of the role of its two key assumptions, the monotone IIA condition and the weak-order structure of the *aggregendum.* Only the former is used in the general aggregation theorem, Theorem 1, and only the latter in the derivation of the corollary. In that derivation, the characteristic properties of weak orders, transitivity and completeness, play a purely combinatorial role.

Viewing Arrow's Theorem as a special case of aggregation in property spaces has the advantage of suggesting variations in the domain and co-domain by modifications of the rationality assumptions on individual and social preferences, or by material domain restrictions. For example, if one allows the domain and co-domain of the aggregation rule to be a preorder rather than a weak order by dropping completeness, and if one uses an analogous definition of the property space of preorders in terms of (1), the transitivity-based families $\{H_{ab}, H_{bc}, H_{ac}^c\}$ remain critical, hence the implications (5), (6), and (7) remain valid. Thus, while the property space of preorders is no longer totally blocked, it is nearly so, and a version of Gibbard's (1969) generalization of Arrow's Theorem can be derived easily. This line of research will be explored further in Nehring-Puppe (2003).

Finally, it is worth noting that Theorem 1 also entails the Gibbard-Satterthwaite Theorem as a corollary, as shown in NP (Corollary to Theorem 5).[4] Here, the set X is given by the set of social alternatives interpreted as individual's preference peaks. Appealing to a result of Barbera-Masso-Neme (1997), we show there that strategy-proof social choice gives rise to a monotone aggregation rule over peaks. We also show there that universality of the preference domain is captured by the totally blocked property space $(X, \mathcal{H})$ with $\mathcal{H}$ given as the set of all singletons and their complements. Equipped with these two auxiliary results, the Gibbard-Satterthwaite Theorem follows from Theorem 1.

[4] For other connections between the Arrow and Gibbard-Satterthwaite Theorems, see Satterthwaite (1975) and Reny (2001).

References

[1] Arrow, K. (1963): *Social Chioce and Individual Values,* 2nd Edition, Wiley, New York.

[2] Barbera, S. (1980): "Pivotal Voters: A New Proof of Arrow's Theorem," *Economics Letters* 6, 13-16.

[3] Barbera, S., J. Masso, and A. Neme (1997): "Voting under Constraints," *Journal of Economic Theory* 76, 298-321.

[4] Geanakoplos, J. (1996): "Three Brief Proofs of Arrow's Impossibility Theorem," mimeo, Cowles Foundation.

[5] Gibbard, A. (1969): "Social Choice and the Arrow Condition," mimeo.

[6] Nehring, K. and C. Puppe: "Strategy-proof Social Choice on Single-peaked Domains: Possibility, Impossibility, and the Space Between", mimeo. Available at http://www.econ.ucdavis.edu/faculty/nehring/

[7] Nehring, K. and C. Puppe (2003): "Non-Binary Arrowian Aggregation," in preparation.

[8] Reny, P. (2001): "Arrow's Theorem and the Gibbard-Satterthwaite Theorem: A Unified Approach," *Economics Letters* 70, 99-105.

[9] Satterthwaite (1975): "Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions, *Journal of Economic Theory* 10, 187-217.

[10] Sethuraman, J., C. Teo and R. Vohra (2001): "Integer Programming and Arrowian Social Welfare Functions," mimeo Northwestern.