ARTICLE IN PRESS YJETH:3606

Available online at www.sciencedirect.com
ScienceDirect

Journal of Economic Theory ••• (••••) •••–•••

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Aggregation of binary evaluations ☆

Elad Dokow, Ron Holzman *,[1]

Department of Mathematics, Technion-Israel Institute of Technology, 32000 Haifa, Israel

Received 24 February 2006; accepted 29 October 2007

Abstract

We study a general aggregation problem in which a society has to determine its position (yes/no) on each of several issues, based on the positions of the members of the society on those issues. There is a prescribed set of feasible evaluations, i.e., permissible combinations of positions on the issues. This framework for the theory of aggregation was introduced by Wilson and further developed by Rubinstein and Fishburn. Among other things, it admits the modeling of preference aggregation (where the issues are pairwise comparisons and feasibility reflects rationality), and of judgment aggregation (where the issues are propositions and feasibility reflects logical consistency). We characterize those sets of feasible evaluations for which the natural analogue of Arrow's impossibility theorem holds true in this framework.

JEL classification: D71

Keywords: Aggregation; Arrow's impossibility; Judgment; Logic; Social choice

1. Introduction

Various problems of aggregation may be cast in the following framework. A society has to determine its positions on each of several issues. There are two possible positions (say, 0 or 1) on

☆ This paper was presented, under the title "An Arrovian impossibility theorem for social truth functions," at the Second World Congress of the Game Theory Society, Marseille, July 2004. The first write-up, which contained more material than the current version, was completed in June 2005. The detailed comments of two referees are gratefully acknowledged.

* Corresponding author. Fax: +972 4 8293388.
E-mail address: holzman@techunix.technion.ac.il (R. Holzman).

[1] Part of this author's work was done while he was a Fellow of the Institute for Advanced Studies at the Hebrew University of Jerusalem.

0022-0531/$ – see front matter
doi:10.1016/j.jet.2007.10.004

Please cite this article in press as: E. Dokow, R. Holzman, Aggregation of binary evaluations, Journal of Economic Theory (2008), doi:10.1016/j.jet.2007.10.004

each issue, but the issues are interrelated and therefore not all combinations of 0–1 positions are feasible. Some set X of 0–1 vectors (of length equal to the number of issues, which we assume to be finite) is given, representing the feasible combinations of positions for each individual in the society as well as for the society as a whole. An aggregator is a function that assigns to every possible profile of individual evaluations in the set X, a social evaluation in the set X. The question is, how do well-behaved aggregators (i.e., that satisfy certain natural conditions) look like, and in particular, under what conditions are we forced to use dictatorial aggregators.

The first to propose such a framework for aggregation theory was Wilson [17]. His motivation was to show that Arrow's [1] impossibility theorem on aggregation of preferences extends to the aggregation of attributes other than preferences. To see how the aggregation of preferences fits into the above framework, consider the case of strict preferences over three alternatives a, b, and c. In this case there are three issues: whether a is preferred to b, whether b is preferred to c, and whether a is preferred to c. Any strict preference over $\{a, b, c\}$ can be encoded as a triple of 0–1 (no/yes) answers to these three questions, but not every such triple is allowed. Transitivity of preferences rules out the triples $(0, 0, 1)$ and $(1, 1, 0)$. The set X consists of the remaining six triples, and its members correspond to the six possible strict orderings of the set $\{a, b, c\}$. An aggregator mapping profiles of triples in X to triples in X thus corresponds to a social welfare function (in which both individual and social preferences are assumed to be strict).[2]

Following Wilson, we adapt Arrow's conditions for social welfare functions to the general framework. We say that an aggregator is independent of irrelevant alternatives (abbreviated IIA) if the society's position on any given issue depends only on the individuals' positions on that same issue. We say that an aggregator is Paretian if the society adopts any unanimously held position. In Arrow's context, his impossibility theorem asserts that when there are at least three alternatives, any IIA and Paretian aggregator must be dictatorial. The main question that we study here is: in the general framework, for which sets X is it the case that every IIA and Paretian aggregator mapping profiles of evaluations in X to evaluations in X must be dictatorial. In other words, we want to identify which limitations on feasibility have the same negative implications for IIA and Paretian aggregation that transitivity has in the context of aggregating preferences.

Rubinstein and Fishburn [8,16] pursued the study of Wilson's framework, and introduced an algebraic point of view.[3] They suggested several examples of aggregation problems that fit into this framework. One such example is the aggregation of equivalence relations, as a classification tool. There is a population of items, say plants, that is to be partitioned into families of similar items according to some criteria. In this application, every pair of items forms an issue, with the entry 1 meaning that they are equivalent and 0 meaning that they are not. The set X consists of those 0–1 vectors that represent equivalence relations. The individual equivalence relations that are to be aggregated may correspond to different experts, or to different criteria of classification.

Our main result is a characterization of those subsets X of $\{0, 1\}^m$ having the property that every IIA and Paretian aggregator over X (for a society of any size) must be dictatorial. There are

[2] The original setting of Arrow's theorem allows for indifference between alternatives. This does not fit into Wilson's framework described here, and therefore we do not recover Arrow's theorem as such. Rather, we recover and generalize the variant of Arrow's theorem (often referred to in the literature also as Arrow's theorem) in which only strict preferences are allowed. In a subsequent paper [6] we extend the framework to allow for abstentions on some of the issues, and thereby recover Arrow's theorem for weak preferences.

[3] They took X to be a subset of a finite-dimensional vector space over some field. Wilson's framework (and ours) corresponds to the case when the field is the two-element field $\{0, 1\}$. Some of Rubinstein and Fishburn's treatment was also specific to this case.

Please cite this article in press as: E. Dokow, R. Holzman, Aggregation of binary evaluations, Journal of Economic Theory (2008), doi:10.1016/j.jet.2007.10.004

two independent conditions on X which together are necessary and sufficient for this property to hold true. One condition was introduced earlier by Nehring and Puppe [13] and called "total blockedness." Roughly speaking, it requires that the limitations on feasibility embodied in the set X make it possible to deduce any position on any issue from any position on any issue, via a chain of deductions. So, intuitively, total blockedness expresses a strong form of cyclicity of deductions. Nehring and Puppe showed that total blockedness of X is a necessary and sufficient condition for every *monotone* IIA and Paretian aggregator over X to be dictatorial. Monotonicity means that changing an individual's position on some issue never results in a change of the society's position on that issue in the opposite direction. This is a plausible property of an aggregator, but it has not been postulated as a requirement in most of the literature on Arrovian aggregation. In not assuming monotonicity, we follow this tradition. It turns out that in the absence of the monotonicity assumption total blockedness is no longer a sufficient condition, and this led us to introduce our second condition. Stated in linear algebraic terms, it requires that the set X *not* be an affine subspace of $\{0, 1\}^m$. An equivalent way to state this is that the limitations on feasibility should not be entirely in the form of parity prescriptions. This condition on X is quite weak, as it rules out only sets with a very specific structure (affine subspaces) that are rare among all subsets of $\{0, 1\}^m$.

A major application of our result is to the problem of judgment aggregation, which has received a significant amount of attention recently. List and Pettit [10] were the first to offer a formal axiomatic treatment of this problem. There is a panel of n judges that faces a set $\mathcal{P}$ of m logical propositions, whose truth or falsehood has to be determined. The propositions are interrelated, and so only a certain subset of all 2^m such determinations are logically consistent. The problem is to aggregate the n individual evaluations, each of which is assumed to be logically consistent, into a joint evaluation that needs to be logically consistent. The literature has identified various combinations of conditions on the agenda $\mathcal{P}$ and requirements from the aggregator that force the aggregator to be dictatorial.

We note that the problem of judgment aggregation may naturally be cast in the framework studied in this paper. To do this, we consider the propositions in $\mathcal{P}$ as the issues, and take X to be the set of logically consistent evaluations.[4] Thus, our main result may be viewed as an impossibility theorem for judgment aggregation. What distinguishes it from the many recent such theorems is that (a) our requirements from the aggregator are the precise analogues of Arrow's requirements in preference aggregation, and subject to this (b) our conditions on the agenda are the weakest possible—they are necessary and sufficient.

The application to judgment aggregation raises several additional questions. The main one is whether the two conditions on X in our result, total blockedness and not being an affine subspace, may be expressed directly in terms of the agenda $\mathcal{P}$ (instead of the set X of logically consistent evaluations), and how easy it is to verify the conditions by looking at $\mathcal{P}$. We can answer these questions in the affirmative when the propositions in $\mathcal{P}$ are expressed in the propositional calculus, and every atomic proposition that appears in some member of $\mathcal{P}$ is itself a member of $\mathcal{P}$. A detailed treatment of these and other related questions will appear in a companion paper [7].

The role of affine subspaces in the aggregation problem is clarified by several additional results. We show that in the application to judgment aggregation, they arise exactly when the logical

[4] The first papers on judgment aggregation used the propositional calculus. More recently other logics were considered, and results were obtained that are valid for any logic satisfying certain criteria. From our point of view, the choice of logic is immaterial. As long as the number of propositions to be decided upon is finite, the logic is two-valued (true/false), and the concept of logical consistency is well-defined, our result applies.

Please cite this article in press as: E. Dokow, R. Holzman, Aggregation of binary evaluations, Journal of Economic Theory (2008), doi:10.1016/j.jet.2007.10.004

propositions can be expressed using negation and equivalence as the only connectives. We indicate that the escape from dictatorship afforded by affine subspaces is narrow and unattractive: if X is totally blocked and an affine subspace then every IIA and Paretian aggregator over X must be a parity rule (basing the society's position on each issue on the parity of the number of supporters within a certain odd-cardinality subset of the society). We prove that, along with the median spaces studied by Nehring and Puppe [13], affine subspaces are the only sets X that admit non-dictatorial aggregators which are IIA, Paretian and neutral (that is, they treat equally all issues and their negations).

In Section 2 we give the necessary definitions and formulate the main result. We also apply it to some examples (motivated by the problem of judgment aggregation), and show that Arrow's impossibility theorem on aggregation of strict preferences is an easy corollary of our result. We prove the main result in Section 3. We present the additional results about the role of affine subspaces in Section 4. We conclude in Section 5 with a comparison of our work to related results in the literature.

2. Formulation of the main result

We consider a finite, non-empty set of issues J. For convenience, if there are m issues in J, we identify J with the set $\{1, \dots, m\}$ of coordinates of vectors of length m. A vector $x = (x_1, \dots, x_m) \in \{0, 1\}^m$ is an *evaluation*. We shall also speak of partial evaluations: if K is a subset of J, a vector $x = (x_j)_{j \in K} \in \{0, 1\}^K$ with entries for issues in K only is a K-evaluation.

We assume that some non-empty subset X of $\{0, 1\}^m$ is given. The evaluations in X are called *feasible*, the others are infeasible. We shall also use this terminology for partial evaluations: a K-evaluation is feasible if it lies in the projection of X on the coordinates in K, and is infeasible otherwise.

A *society* is a finite, non-empty set N. For convenience, if there are n individuals in N, we identify N with the set $\{1, \dots, n\}$. If we specify a feasible evaluation $x^i = (x^i_1, \dots, x^i_m) \in X$ for each individual $i \in N$, we obtain a *profile* of feasible evaluations $\mathbf{x} = (x^i_j) \in X^n$. We may view a profile as an $n \times m$ matrix all of whose rows lie in X. We use superscripts to indicate individuals (rows) and subscripts to indicate issues (columns).

An *aggregator* for N over X is a mapping $f : X^n \to X$. It assigns to every possible profile of individual feasible evaluations, a social evaluation which is also feasible. Any aggregator f may be written in the form $f = (f_1, \dots, f_m)$ where f_j is the jth component of f. That is, $f_j : X^n \to \{0, 1\}$ assigns to every profile the social position on the jth issue.

An aggregator $f : X^n \to X$ is *independent of irrelevant alternatives* (abbreviated IIA) if for every $j \in J$ and any two profiles $\mathbf{x}$ and $\mathbf{y}$ satisfying $x^i_j = y^i_j$ for all $i \in N$, we have $f_j(\mathbf{x}) = f_j(\mathbf{y})$. This means that the social position on a given issue is determined entirely by the individual positions on that same issue. Viewing profiles as matrices, this means that the aggregation is done column-by-column. As we shall deal with IIA aggregators, we will slightly abuse notation and write also expressions of the form $f_j(x_j)$, where $x_j = (x^1_j, \dots, x^n_j)$ is the column vector of individual positions on the jth issue.

An aggregator $f : X^n \to X$ is *Paretian* if we have $f(\mathbf{x}) = x$ whenever the profile $\mathbf{x}$ is such that $x^i = x$ for all $i \in N$. Note that in the presence of IIA, this is equivalent to demanding that whenever all individuals agree on any one issue, the society adopts this position on that issue.

An aggregator $f : X^n \to X$ is *dictatorial* if there exists an individual $d \in N$ such that $f(\mathbf{x}) = x^d$ for every $\mathbf{x} \in X^n$. That is to say, the society always adopts the dictator's evaluation. A dictatorial aggregator is trivially IIA and Paretian.

We say that X is an *impossibility domain* if for every society N, every IIA and Paretian aggregator for N over X is dictatorial. Otherwise we say that X is a possibility domain.

Our aim is to characterize impossibility domains. In doing so, there is no loss of generality in assuming non-degeneracy in the following sense. We say that X is *non-degenerate* if for every issue $j \in J$ and every $u \in \{0,1\}$ there exists $x \in X$ with $x_j = u$. To see why we may assume this, suppose that some issues admit only one value in the feasible set X. If all issues are like that, then X has just one member and is trivially an impossibility domain. If some, but not all issues are like that, we may delete them from the set of issues and consider the projection X' of X on the remaining coordinates. It is easy to check that X is an impossibility domain if and only if X' is. We will henceforth assume non-degeneracy.

We turn now to the presentation of the first condition that appears in our characterization. The condition, named total blockedness, was introduced by Nehring and Puppe [13]. Let X be a non-degenerate subset of $\{0,1\}^m$. A *minimally infeasible partial evaluation* (abbreviated MIPE) is a K-evaluation $x = (x_j)_{j \in K}$ for some $K \subseteq J$ which is infeasible, but such that every restriction of x to a proper subset of K is feasible. By non-degeneracy, the length of any MIPE (i.e., the size of K) is at least two. We use the MIPEs to construct a directed graph associated with X, denoted by G_X. It has $2m$ vertices, labeled $0_1, 1_1, 0_2, 1_2, \ldots, 0_m, 1_m$. The vertex u_j is to be interpreted as holding the position u on issue j. There is an arc in G_X from vertex u_k to vertex v_ℓ (written $u_k \to v_\ell$) if and only if $k \neq \ell$ and there exists a MIPE $x = (x_j)_{j \in K}$ such that $\{k, \ell\} \subseteq K$ and $x_k = u$, $x_\ell = \bar{v}$ (where $\bar{v}$ denotes $1 - v$). The interpretation of $u_k \to v_\ell$ is that u_k conditionally entails v_ℓ in the following sense: conditional on holding the positions prescribed in the MIPE x on all issues in $K \setminus \{k, \ell\}$, holding position u on issue k entails holding position v on issue ℓ (since x is infeasible). If $u_k \to v_\ell$ by virtue of a MIPE of length two, then holding position u on issue k entails holding position v on issue ℓ, without conditions. Note that the arcs obey the logical law of contrapositives: $u_k \to v_\ell$ if and only if $\bar{v}_\ell \to \bar{u}_k$. We write $u_k \to\to v_\ell$ if there exists a directed path in G_X from u_k to v_ℓ. Finally, we say that X is *totally blocked* if G_X is strongly connected, that is, for any two vertices u_k and v_ℓ we have $u_k \to\to v_\ell$ (note that this relation is required to hold not only when $k \neq \ell$ but also when $k = \ell$). The following example illustrates these concepts.

Example 1. Let $X = \{(0,0,0), (0,1,0), (1,0,0), (1,1,1)\}$. Note that this set of feasible evaluations is characterized by the requirement that the third coordinate be 1 if and only if the first two are both 1. In judgment aggregation terminology, this corresponds to the case when the third proposition is the logical conjunction of the first two. This is the example that underlies the well-known "doctrinal paradox" (see e.g. [10]). In this example there are three MIPEs, namely (using vectors with $\cdot$ entries to denote partial evaluations): $(0, \cdot, 1)$, $(\cdot, 0, 1)$, and $(1, 1, 0)$. The graph G_X has vertices $0_1, 1_1, 0_2, 1_2, 0_3, 1_3$. The first of the above MIPEs gives rise to the arcs $0_1 \to 0_3$ and $1_3 \to 1_1$. Similarly, the second MIPE gives rise to $0_2 \to 0_3$ and $1_3 \to 1_2$. The MIPE of length three gives rise to six arcs: $1_1 \to 0_2$, $1_1 \to 1_3$, $1_2 \to 0_1$, $1_2 \to 1_3$, $0_3 \to 0_1$, $0_3 \to 0_2$. The arcs of G_X are the ten arcs listed above. Note that these arcs never go from a 0_k to a 1_ℓ. We conclude that G_X is not strongly connected, and hence X is not totally blocked.

The second condition that appears in our characterization comes from linear algebra. The set $\{0,1\}^m$ may be viewed as a vector space over the field $\{0,1\}$. In this space, addition is performed modulo 2, and subtraction is the same as addition. A linear subspace is a non-empty subset closed under addition (note that closure under scalar multiplication is not an issue here, the only scalars

Please cite this article in press as: E. Dokow, R. Holzman, Aggregation of binary evaluations, Journal of Economic Theory (2008), doi:10.1016/j.jet.2007.10.004

being 0 and 1). An *affine subspace* is a subset obtained from a linear subspace by adding a fixed vector to each of its elements. The following is a typical example of an affine subspace.

Example 2. Let X be the set of all $x = (x_1, \dots, x_5) \in \{0,1\}^5$ that satisfy the following two constraints:

(i) $x_1 \neq x_2$,
(ii) among x_2, x_3, x_4 there is an even number of 1's.

Note that the value of x_5 is unconstrained. It is easy to see that each constraint is satisfied by half of the vectors in $\{0,1\}^5$, and that one quarter of the vectors satisfy both constraints, that is, $|X| = 8$. Using addition modulo 2, the two constraints may equivalently be expressed as linear equations:

(i) $x_1 + x_2 = 1$,
(ii) $x_2 + x_3 + x_4 = 0$.

We see that X is the set of solutions of a system of two linear equations. If we replaced the right hand side of equation (i) by 0, the set of solutions X' of the resulting system would actually be a linear subspace. Now, X is obtained from X' by adding the fixed vector $(1, 0, 0, 0, 0)$ to each of its elements. Hence X is an affine subspace.

The following proposition lists several equivalent characterizations of an affine subspace of $\{0,1\}^m$.

Proposition 2.1. *Let X be a non-empty subset of* $\{0,1\}^m$. *The following are equivalent*:

1. *X is an affine subspace.*
2. *X is the set of solutions of a system of linear equations in m unknowns over* $\{0,1\}$, *that is, there exist a $k \times m$ matrix A over* $\{0,1\}$ *for some k and a column vector $b \in \{0,1\}^k$ so that* $X = \{x \mid Ax = b\}$.
3. *X is closed under addition of odd-tuples, that is, $x^1, \dots, x^k \in X,\ k$ odd $\Rightarrow x^1 + \dots + x^k \in X$.*
4. *X is closed under addition of triples, that is, $x, y, z \in X \Rightarrow x + y + z \in X$.*

Proof. The equivalence of clauses 1 and 2 is known from elementary linear algebra (and holds over any field). The other characterizations are specific to the binary field. To see that 2 implies 3, note that if k is odd and $Ax^i = b$ for $i = 1, \dots, k$, then

$$A(x^1 + \dots + x^k) = \underbrace{b + \dots + b}_{k} = b.$$

Trivially 3 implies 4. To see that 4 implies 1, let w be an arbitrary fixed element of X. We show that, assuming 4, the set $X + w = \{x + w \mid x \in X\}$ is closed under addition. Indeed, if $x, y \in X$ then $(x + w) + (y + w) = (x + w + y) + w \in X + w$. This shows that X is an affine subspace. □

Note that clause 2 characterizes an affine subspace as resulting from parity requirements: each linear equation requires a certain parity (even or odd) of the number of 1's within a certain subset of coordinates. This makes affine subspaces very special and rare objects among all subsets of $\{0,1\}^m$. It is easy to see that a necessary (but far from sufficient) condition on X in order to be an affine subspace is that its cardinality be a power of 2.

Clauses 3 and 4 in the proposition suggest non-dictatorial aggregators that are IIA and Paretian. For this reason, our characterization of impossibility domains will include the condition of *not* being an affine subspace.

We are now ready for the main result.

Theorem 2.2. *Let X be a non-degenerate subset of $\{0,1\}^m$. Then X is an impossibility domain if and only if X is totally blocked and is not an affine subspace.*

We recall that Nehring and Puppe [13] studied a weaker notion of an impossibility domain, namely the requirement that every monotone IIA and Paretian aggregator over X must be dictatorial. Their result was that this is equivalent to X being totally blocked. We note that the two results are logically incomparable, though one half of the 'only if' direction of our result (an impossibility domain must be totally blocked) follows from the corresponding direction of their result. Regarding the 'if' direction, our result requires an extra condition on the domain (not being an affine subspace) in exchange for waiving the monotonicity assumption on the aggregator.

The proof of the theorem will be given in the following section. It will actually show a little more. By definition, if X is a possibility domain, then there *exists* some n for which there exists a non-dictatorial aggregator $f: X^n \to X$ which is IIA and Paretian. The proof will show that if one of the conditions of the theorem is violated then such an aggregator in fact exists for *every* $n \geqslant 3$. It is not the case, though, that for every $n \geqslant 3$ we can come up with such an aggregator in which all n individuals actually take part in the decision making. Let us call an individual inessential (for a given aggregator) if the aggregator ignores his evaluation, and essential otherwise. If X is not totally blocked then, depending on the structure of G_X, we may be forced to have as few as two essential individuals. If X is totally blocked and is an affine subspace, then the number of essential individuals must be odd. Regarding the case $n = 2$, it is easy to see that in this case every IIA and Paretian aggregator is monotone. Therefore it follows from [13] that every IIA and Paretian aggregator $f: X^2 \to X$ is dictatorial if and only if X is totally blocked.

We proceed now to show how the theorem works in a number of examples.

Example 1 *(continued).* We saw that the set $X = \{(0,0,0), (0,1,0), (1,0,0), (1,1,1)\}$ is not totally blocked. Thus, although this set does satisfy the second condition of Theorem 2.2 (it is not an affine subspace, e.g., $(0,0,0) + (0,1,0) + (1,0,0) = (1,1,0) \notin X$), it is a possibility domain. We conclude that this set does admit non-dictatorial aggregation which is IIA and Paretian. How do such aggregators look like? The unanimity rule $f = (f_1, f_2, f_3)$ in which for every $j \in \{1,2,3\}$ we have $f_j(x_j) = 1$ if and only if $x_j^i = 1$ for every individual $i \in N$ works. In fact, all IIA and Paretian aggregators for this example must be of this form, except that instead of requiring unanimity in the entire society N, we may require it in some prescribed subset of N, ignoring the other individuals. The uniqueness of these oligarchic rules will be proved, in a more general setup, in a companion paper [7]. A similar characterization of oligarchic rules was independently obtained by Nehring and Puppe [14]; their characterization, though, uses the additional property of monotonicity.

Example 3. Let $X = \{(0,0,1), (0,1,0), (1,0,0), (1,1,1)\}$. In judgment aggregation terminology, this set corresponds to the case when the third proposition asserts the logical equivalence of the first two. This X is an affine subspace, defined by the requirement that the total number of 1's be odd. Thus, although this set is totally blocked (as can be easily checked), it is a possibility domain. How do non-dictatorial, IIA and Paretian aggregators over this set look like? Proposition 2.1 indicates that if the number of individuals is odd, we can use addition modulo 2 to aggregate evaluations. More generally, we can prescribe a subset of N of odd cardinality and add their evaluations, ignoring the other individuals. In Section 4 we will show, in a more general setup, the uniqueness of these parity rules. We will also show that this example is representative of a general family of examples in judgment aggregation, for which the set of feasible evaluations is an affine subspace. We will see that this occurs if and only if the propositions to be evaluated can be expressed using only logical negation and equivalence.

We note that the standard presentations of the well-known "doctrinal paradox" use the setting of Example 1, or some variant of it in which $m = 3$, the first two coordinates may assume any combination of values, and the value of the third coordinate is a function of those of the first two. It may be checked that all such examples are actually possibility domains. This remarkable fact may explain why List and Pettit [10] resorted to stronger assumptions on the aggregator in order to fit the paradox into the framework of an impossibility theorem. Our next example shows that in a similar setting, but with a fourth coordinate added, we may reach impossibility.

Example 4. Let $X = \{(0,0,0,0), (0,1,0,1), (1,0,0,1), (1,1,1,1)\}$. In judgment aggregation terminology, this set corresponds to the case when the third proposition is the logical conjunction of the first two, and the fourth one is the logical disjunction of the first two. This X is totally blocked, as may be checked with some effort. Moreover, X is not an affine subspace, e.g., $(0,0,0,0) + (0,1,0,1) + (1,0,0,1) = (1,1,0,0) \notin X$. We conclude from Theorem 2.2 that X is an impossibility domain: no non-dictatorial aggregation which is IIA and Paretian is possible.

Example 5 *(Arrow's impossibility theorem for strict preferences).* We explained in the introduction how the problem of (strict) preference aggregation can be expressed in our terminology. Let us do this explicitly for any number $k \geqslant 3$ of alternatives. We enumerate the alternatives as $a_1, \ldots, a_k$. We have $\binom{k}{2}$ issues, corresponding to the pairs of alternatives. Each issue is indexed by a pair rs such that $1 \leqslant r < s \leqslant k$. By convention, $x_{rs} = 1$ means that a_r is preferred to a_s and $x_{rs} = 0$ means the opposite. The set of feasible evaluations is

$$X = \left\{x \;\middle|\; \forall 1 \leqslant r < s < t \leqslant k \ (x_{rs}, x_{st}, x_{rt}) \neq (0,0,1), (1,1,0)\right\}.$$

Arrow's impossibility theorem for strict preferences is equivalent to the assertion that this X is an impossibility domain. Let us check that X satisfies the conditions of Theorem 2.2. By the definition of X, for every $r < s < t$ we have two MIPEs on the issues rs, st, rt: $(0,0,1)$ and $(1,1,0)$. The arcs generated in G_X by these two MIPEs between the vertices $0_{rs}, 1_{rs}, 0_{st}, 1_{st}, 0_{rt}, 1_{rt}$ suffice to connect any two of them by a path. This in turn implies that any two vertices in G_X are connected by a path, because they either belong together to such a block of six vertices, or belong to two such blocks that intersect. This shows that X is totally blocked. Also, X is not an affine subspace, for example because the cardinality of an affine subspace must be a power of 2, whereas $|X| = k!$. Thus, Arrow's impossibility theorem for strict preferences follows from Theorem 2.2. We note that Nehring [12] expressed the problem of preference aggregation in a similar way, and checked total blockedness. This enabled him to obtain a monotone version of Arrow's theorem as a corollary of his result with Puppe.

Please cite this article in press as: E. Dokow, R. Holzman, Aggregation of binary evaluations, Journal of Economic Theory (2008), doi:10.1016/j.jet.2007.10.004

Table 1
Construction for Claim 3.1

	1	2	3	$\cdots$	r
S	u	v	x_3	$\cdots$	x_r
$N \setminus S$	$\bar{u}$	$\bar{v}$	x_3	$\cdots$	x_r
	u	$\bar{v}$	x_3	$\cdots$	x_r

3. Proof of the main result

Certain collections of winning coalitions associated with an IIA and Paretian aggregator will play a central role in our proof. Suppose that X is a non-degenerate subset of $\{0,1\}^m$. If $f : X^n \to X$ is an IIA and Paretian aggregator, then f may be written in the form $f = (f_1, \dots, f_m)$ where f_j maps columns of positions on the jth issue (of the form $x_j = (x_j^1, \dots, x_j^n) \in \{0,1\}^n$) into $\{0,1\}$. For each issue j and each position $u \in \{0,1\}$, we say that a subset S of N is a *u_j-winning coalition* if

$$x_j^i = \begin{cases} u & \text{if } i \in S, \\ \bar{u} & \text{if } i \in N \setminus S, \end{cases} \quad \Rightarrow \quad f_j(x_j) = u.$$

Thus, S is u_j-winning if it prevails on issue j when its members, and only they, hold the position u. We denote by $\mathcal{W}_j^u$ the collection of all u_j-winning coalitions. The Pareto property implies that $N \in \mathcal{W}_j^u$ and $\emptyset \notin \mathcal{W}_j^u$ for every j and u. The collections $\mathcal{W}_j^u$ may be thought of as simple games, though not necessarily monotone ones. It follows from the definition that for each j the two collections $\mathcal{W}_j^0$ and $\mathcal{W}_j^1$ are *dual* to each other, in the sense that $S \in \mathcal{W}_j^0 \Leftrightarrow N \setminus S \notin \mathcal{W}_j^1$. Note that, conversely, if we arbitrarily specify collections of coalitions $\mathcal{W}_j^u$ for every j and u, such that $N \in \mathcal{W}_j^u$ and $\mathcal{W}_j^0$ and $\mathcal{W}_j^1$ are dual to each other, then we have implicitly defined the components $f_1, \dots, f_m$. The resulting function $f = (f_1, \dots, f_m)$ may not map X^n into X, but if it does then it is an IIA and Paretian aggregator. Clearly, f is dictatorial if and only if there exists $d \in N$ so that for every j and u we have $S \in \mathcal{W}_j^u \Leftrightarrow d \in S$.

In the first part of the proof we show that the conditions of Theorem 2.2 are sufficient for X to be an impossibility domain. So we consider some IIA and Paretian aggregator $f : X^n \to X$ and, using the conditions of the theorem, we gradually establish properties of the associated collections of winning coalitions.

Claim 3.1. *If $u_k \to v_\ell$ in the graph G_X then $\mathcal{W}_k^u \subseteq \mathcal{W}_\ell^v$.*

Proof. Assume, for the sake of contradiction, that the coalition S is in $\mathcal{W}_k^u$ but not in $\mathcal{W}_\ell^v$. By the definition of G_X, there exists a MIPE $x = (x_j)_{j \in K}$ such that $\{k, \ell\} \subseteq K$ and $x_k = u$, $x_\ell = \bar{v}$. In Table 1 we construct a profile of feasible evaluations and the resulting social evaluation, all restricted to issues in K (for ease of exposition, we assume that $K = \{1, \dots, r\}$ and $k = 1$, $\ell = 2$).

Observe that each of the rows corresponding to S and to $N \setminus S$ in the table differs from the MIPE x in exactly one place, and therefore by the minimality of a MIPE these rows are feasible (i.e., can be extended to a feasible evaluation on J). The resulting social positions are determined in the first two columns by our assumptions on S, and in the remaining columns by the Pareto property. Thus, the social K-evaluation equals x, which is a contradiction since x is infeasible. □

Please cite this article in press as: E. Dokow, R. Holzman, Aggregation of binary evaluations, Journal of Economic Theory (2008), doi:10.1016/j.jet.2007.10.004

Table 2
Construction for Claim 3.3

	1	2	3	4	$\cdots$	r
S	$\overline{x_1}$	x_2	x_3	x_4	$\cdots$	x_r
T	x_1	$\overline{x_2}$	x_3	x_4	$\cdots$	x_r
$N \setminus U$	x_1	x_2	$\overline{x_3}$	x_4	$\cdots$	x_r
	x_1	x_2	x_3	x_4	$\cdots$	x_r

By repeated applications of Claim 3.1, it follows that $u_k \to\to v_\ell$ implies $\mathcal{W}_k^u \subseteq \mathcal{W}_\ell^v$. Therefore, if X is totally blocked then $\mathcal{W}_k^u = \mathcal{W}_\ell^v$ for any u_k and v_ℓ. Thus, there exists one common collection, that we denote by $\mathcal{W}$, of winning coalitions. In this case we say that f is *neutral*, that is, it treats equally all issues and their negations. Note that the collection $\mathcal{W}$ is *self-dual*, in the sense that $S \in \mathcal{W} \Leftrightarrow N \setminus S \notin \mathcal{W}$.

Claim 3.2. *If X is totally blocked then there exists a MIPE of length at least three.*

Proof. As explained in the previous section, if the arc $u_k \to v_\ell$ exists on account of a MIPE of length two, then any feasible evaluation that specifies position u on issue k must specify position v on issue ℓ. So, if X is totally blocked and every MIPE has length two, then we can iterate and deduce that any feasible evaluation that specifies position u on issue k must specify position $\bar{u}$ on issue k. This is a contradiction. $\square$

Claim 3.3. *If there exists a MIPE of length at least three, and f is neutral, then $\mathcal{W}$ is decomposable in the following sense: if $U \in \mathcal{W}$ and (S, T) is any partition of U then either $S \in \mathcal{W}$ or $T \in \mathcal{W}$.*

Proof. Suppose, for the sake of contradiction, that $S, T \notin \mathcal{W}$, $S \cap T = \emptyset$, and $U = S \cup T \in \mathcal{W}$. Let $x = (x_j)_{j \in K}$ be a MIPE with $|K| \geqslant 3$. Now, consider the construction in Table 2 (where for ease of exposition $K = \{1, \ldots, r\}$). By similar arguments to those presented for Table 1, this construction is justified and leads to a contradiction. $\square$

Starting from $N \in \mathcal{W}$ and repeatedly using decomposability, we conclude that there exists $d \in N$ so that $\{d\} \in \mathcal{W}$. By self-duality of $\mathcal{W}$, in order to prove that d is a dictator it suffices to show that $\mathcal{W}$ is *monotone*, in the following sense: if $S \in \mathcal{W}$ and $S \subset T$ then $T \in \mathcal{W}$. We do this in the next two claims. We will say about two partial evaluations that they are at *distance* 2 if they are both K-evaluations for the same K and they differ in exactly two places.

Claim 3.4. *If there exist a MIPE and a feasible partial evaluation that are at distance* 2*, and f is neutral, then $\mathcal{W}$ is monotone.*

Proof. Assume, for the sake of contradiction, that $S \subset T$, $S \in \mathcal{W}$, and $T \notin \mathcal{W}$. Let $x = (x_j)_{j \in K}$ be a MIPE and let $y = (y_j)_{j \in K}$ be a feasible K-evaluation at distance 2 from x. Then the construction in Table 3 is legitimate and leads to a contradiction (again, for ease of exposition, $K = \{1, \ldots, r\}$ and the two issues where y differs from x are 1 and 2). $\square$

The next claim shows that the premise of Claim 3.4 must hold true, unless X is an affine subspace. This will complete the proof of sufficiency in Theorem 2.2.

Table 3
Construction for Claim 3.4

	1	2	3	$\cdots$	r
S	x_1	$\overline{x_2}$	x_3	$\cdots$	x_r
$T \setminus S$	$\overline{x_1}$	$\overline{x_2}$	x_3	$\cdots$	x_r
$N \setminus T$	$\overline{x_1}$	x_2	x_3	$\cdots$	x_r
	x_1	x_2	x_3	$\cdots$	x_r

Claim 3.5. *If for every MIPE, every partial evaluation at distance* 2 *from it is infeasible, then X is an affine subspace.*

Proof. We first show that, under the assumption of our claim, more is true: if $x = (x_j)_{j \in K}$ is a MIPE and $y = (y_j)_{j \in K}$ is at distance 2 from x, then y is actually a MIPE. Suppose that y is infeasible, but not minimal. Then the restriction of y to some $K' \subsetneqq K$, which we denote by y', is a MIPE. Supposing w.l.o.g. that 1 and 2 are the places where x, y differ, we must have $\{1, 2\} \subseteq K'$ for otherwise y' would be feasible (as it would be contained in either the first or the third evaluation in Table 3). Now consider x', the restriction of x to K'. It is at distance 2 from y', which is a MIPE, so by our assumption x' is infeasible. This contradicts the minimality of x.

We will call a K-evaluation even (respectively odd) if it has an even (respectively odd) number of 1's. Clearly, any two K-evaluations (for the same K) of the same parity can be linked by a sequence of K-evaluations in which every two successive members are at distance 2. Therefore, what we proved above implies that the MIPEs can be partitioned into blocks, each block consisting of all K-evaluations for a given K that have a given parity. Consider such a block. If it consists of the even (respectively odd) K-evaluations, then $x \in \{0, 1\}^m$ does not contain any MIPE in this block if and only if $\sum_{j \in K} x_j = 1$ (respectively $\sum_{j \in K} x_j = 0$) modulo 2. Obviously $x \in X$ if and only if x does not contain any MIPE in any block, which is equivalent to x satisfying the system of linear equations corresponding as above to the blocks. Thus X is the set of solutions of this system, and hence an affine subspace. □

The next two claims show that each of the conditions of Theorem 2.2 is necessary for X to be an impossibility domain.

Claim 3.6. *If X is not totally blocked then for every $n \geqslant 2$ there exists a non-dictatorial, IIA and Paretian aggregator $f : X^n \to X$.*

Proof. As X is not totally blocked, there exists a partition of the vertices of G_X into two non-empty parts V_1 and V_2 so that there is no arc in G_X from a vertex in V_1 to a vertex in V_2. Let N be a society, $|N| = n \geqslant 2$. We define f by specifying the collections of coalitions $\mathcal{W}_j^u$ as follows:

$$\mathcal{W}_j^u = \begin{cases} \{S \subseteq N \mid 1 \in S\} & \text{if } \{u_j, \bar{u}_j\} \subseteq V_1, \\ \{S \subseteq N \mid 2 \in S\} & \text{if } \{u_j, \bar{u}_j\} \subseteq V_2, \\ \{S \subseteq N \mid S \neq \emptyset\} & \text{if } u_j \in V_1,\ \bar{u}_j \in V_2, \\ \{N\} & \text{if } u_j \in V_2,\ \bar{u}_j \in V_1. \end{cases}$$

All we need to show is that the resulting f maps X^n into X (the other required properties of f are obvious). Suppose, for the sake of contradiction, that $f(\mathbf{x}) \notin X$ for some $\mathbf{x} \in X^n$. Then some restriction of $f(\mathbf{x})$, say $y = (y_j)_{j \in K}$, is a MIPE. For each $j \in K$, let $S_j = \{i \in N \mid x_j^i = y_j\}$. Then we must have $S_j \in \mathcal{W}_j^{y_j}$ for all $j \in K$, and $\bigcap_{j \in K} S_j = \emptyset$. By the definition of the collections $\mathcal{W}_j^u$, this requires the existence of some $k, \ell \in K$, $k \neq \ell$, so that $\mathcal{W}_k^{y_k}$ is either $\{S \subseteq N \mid 1 \in S\}$ or $\{S \subseteq N \mid S \neq \emptyset\}$, and $\mathcal{W}_\ell^{y_\ell}$ is either $\{S \subseteq N \mid 2 \in S\}$ or $\{S \subseteq N \mid S \neq \emptyset\}$. But then, letting $u = y_k$ and $v = y_\ell$, we have $u_k \in V_1$, $\bar{v}_\ell \in V_2$, and $u_k \to \bar{v}_\ell$ due to the MIPE y. This contradicts our assumption about V_1 and V_2. □

Claim 3.7. *If X is an affine subspace then for every $n \geqslant 3$ there exists a non-dictatorial, IIA and Paretian aggregator $f : X^n \to X$.*

Proof. Let N be a society, $|N| = n \geqslant 3$. We choose some subset R of N of odd cardinality $k \geqslant 3$, and define $f(\mathbf{x}) = \sum_{i \in R} x^i$. By Proposition 2.1 f maps X^n into X, and all other required properties are obvious. □

We conclude this section with a discussion of the relation between our proof and that of Nehring and Puppe's [13] corresponding result for monotone aggregators. The general structure of our proof of the sufficiency part above fits that of the corresponding part in [13]. However, the proof there relies on a characterization of the consistency of the aggregator (i.e., its image being contained in X) by a simple "intersection property" of the associated collections of winning coalitions. Without monotonicity, the intersection property need not hold, and there seems to be no similar simple property that characterizes consistency. Thus, we had to redo the proof of Claim 3.1. We deduced from it, using total blockedness, that f is neutral, in the same way as it was done in [13]. Our Claims 3.4 and 3.5, which have no counterpart in [13], show that given neutrality, the assumption that X is not an affine subspace implies monotonicity of the aggregator. Thus, we could have proceeded directly to Claims 3.4 and 3.5 (skipping Claims 3.2 and 3.3), and once we established monotonicity we could have invoked the result of [13] to deduce the existence of a dictator. We preferred to carry out the proof as we did for two reasons. First, to keep the proof self-contained. Secondly, to emphasize that the decomposability property (Claim 3.3) and the conclusion that $\{d\} \in \mathcal{W}$ for some $d \in N$ are valid regardless of monotonicity (and hence do not require the assumption that X is not an affine subspace). The corresponding conclusion was proved in [13] using again the intersection property, that hinges on monotonicity. We note, finally, that Claim 3.2, as well as Claim 3.6 in the necessity part of our proof, appeared already in [13] in essentially the same form.

4. More about the affine case

In this section we give additional characterizations of the case when X is an affine subspace, and further results about aggregation in this case.

First, we consider the judgment aggregation interpretation of our framework, where issues are represented by logical propositions, and X is the set of all logically consistent evaluations. It turns out that X is an affine subspace exactly when the propositions can be expressed using only logical negation and equivalence. Formally, consider a logical language that contains the atomic propositions $p_1, \ldots, p_k$ and the connectives $\neg$ and $\leftrightarrow$. Let $\mathcal{P}_{\neg,\leftrightarrow}(p_1, \ldots, p_k)$ be the set of all propositions in this language, that is, the smallest set $\mathcal{P}$ of expressions that includes $p_1, \ldots, p_k$ and satisfies: $\varphi \in \mathcal{P}$ implies $(\neg\varphi) \in \mathcal{P}$, and $\varphi, \psi \in \mathcal{P}$ implies $(\varphi \leftrightarrow \psi) \in \mathcal{P}$. Given an assignment $t = (t_1, \ldots, t_k) \in \{0, 1\}^k$ of truth values to the atomic propositions (where 1 means 'true' and 0 means 'false'), define inductively the truth value $T_t(\varphi) \in \{0, 1\}$ of $\varphi \in \mathcal{P}_{\neg,\leftrightarrow}(p_1, \ldots, p_k)$ by the following rules: (a) $T_t(p_i) = t_i$, (b) $T_t(\neg\varphi) = \overline{T_t(\varphi)}$, and (c) $T_t(\varphi \leftrightarrow \psi) = \overline{T_t(\varphi) + T_t(\psi)}$

Please cite this article in press as: E. Dokow, R. Holzman, Aggregation of binary evaluations, Journal of Economic Theory (2008), doi:10.1016/j.jet.2007.10.004

(using addition modulo 2). For a list of m propositions $\varphi_1, \dots, \varphi_m \in \mathcal{P}_{\neg,\leftrightarrow}(p_1, \dots, p_k)$ that represent the m issues in our framework, let

$$X(\varphi_1, \dots, \varphi_m) = \big\{\big(T_t(\varphi_1), \dots, T_t(\varphi_m)\big) \bigm| t \in \{0,1\}^k\big\}.$$

Thus $X(\varphi_1, \dots, \varphi_m)$ is the subset of $\{0,1\}^m$ that consists of all logically consistent evaluations of $\varphi_1, \dots, \varphi_m$.

Proposition 4.1. *For any $\varphi_1, \dots, \varphi_m \in \mathcal{P}_{\neg,\leftrightarrow}(p_1, \dots, p_k)$ the set $X(\varphi_1, \dots, \varphi_m)$ is an affine subspace of $\{0,1\}^m$. Conversely, given any affine subspace X of $\{0,1\}^m$, one can find $\varphi_1, \dots, \varphi_m \in \mathcal{P}_{\neg,\leftrightarrow}(p_1, \dots, p_k)$ so that $X(\varphi_1, \dots, \varphi_m) = X$ (and one can achieve such a representation with $k \leqslant m$).*

Proof. Recall that a mapping from the vector space $\{0,1\}^k$ to the vector space $\{0,1\}^\ell$ is said to be linear if it preserves addition, and is said to be affine if it is of the form $g(t) = h(t) + y$ where h is linear and y is a fixed vector in $\{0,1\}^\ell$. It is easy to check, by induction on the structure of φ, that for every fixed $\varphi \in \mathcal{P}_{\neg,\leftrightarrow}(p_1, \dots, p_k)$ the mapping $t \mapsto T_t(\varphi)$ from $\{0,1\}^k$ to $\{0,1\}$ is affine. It follows that for any $\varphi_1, \dots, \varphi_m \in \mathcal{P}_{\neg,\leftrightarrow}(p_1, \dots, p_k)$ the mapping $t \mapsto (T_t(\varphi_1), \dots, T_t(\varphi_m))$ from $\{0,1\}^k$ to $\{0,1\}^m$ is affine. By a well-known fact from linear algebra, this implies that the image of the mapping, namely $X(\varphi_1, \dots, \varphi_m)$, is an affine subspace of $\{0,1\}^m$.

Conversely, let X be an affine subspace of $\{0,1\}^m$. By Proposition 2.1, X is the set of solutions of a system of linear equations in m unknowns over $\{0,1\}$. By performing Gaussian elimination the system can be brought to a canonical form, which in the binary case can be described as follows. There exists a subset $K \subseteq J = \{1, \dots, m\}$ so that the variables x_j, $j \in K$, are free, and the values of the remaining variables are determined by the equations

$$x_i = \sum_{j \in K_i} x_j + u_i, \quad i \in J \setminus K,$$

where $K_i \subseteq K$ and $u_i \in \{0,1\}$ for each $i \in J \setminus K$. Suppose first that $K \neq \emptyset$. Assume, for notational convenience, that $K = \{1, \dots, k\}$. We proceed to define $\varphi_1, \dots, \varphi_m$. For $j = 1, \dots, k$, we let $\varphi_j = p_j$ (these are the k atomic propositions). For $i = k+1, \dots, m$ we define first an auxiliary proposition ψ_i as follows. If $K_i \neq \emptyset$ then each p_j, $j \in K_i$, appears in ψ_i exactly once, and the only connective used in ψ_i is $\leftrightarrow$ (e.g., if $K_i = \{1\}$ then $\psi_i = p_1$, if $K_i = \{1,2\}$ then $\psi_i = (p_1 \leftrightarrow p_2)$, if $K_i = \{1,3,5\}$ then $\psi_i = ((p_1 \leftrightarrow p_3) \leftrightarrow p_5)$, and so on). If $K_i = \emptyset$ then let $\psi_i = (p_1 \leftrightarrow p_1)$. Note that ψ_i is true if and only if the number of false p_j, $j \in K_i$, is even, or equivalently, the number of true p_j, $j \in K_i$, is congruent modulo 2 to $|K_i|$. Now we define φ_i, $i = k+1, \dots, m$, as follows. If $|K_i| \equiv u_i \pmod 2$ then $\varphi_i = (\neg\psi_i)$, otherwise $\varphi_i = \psi_i$. It follows from the above that the truth values of the φ_i are determined by those of the p_j exactly according to the equations $x_i = \sum_{j \in K_i} x_j + u_i$, and hence $X(\varphi_1, \dots, \varphi_m) = X$. Finally, if $K = \emptyset$ then we let $k = 1$ and define φ_i, $i = 1, \dots, m$, as in the case $K_i = \emptyset$ above. □

Our next characterization of the affine case is more technical. It shows that the property that appeared in the course of the proof of the main result, in Claim 3.5, characterizes affine subspaces. Recall that two partial evaluations are said to be at distance 2 if they are both K-evaluations for the same K and they differ in exactly two places.

Proposition 4.2. *Let X be a non-degenerate subset of $\{0,1\}^m$. The following are equivalent:*

Please cite this article in press as: E. Dokow, R. Holzman, Aggregation of binary evaluations, Journal of Economic Theory (2008), doi:10.1016/j.jet.2007.10.004

Table 4
Construction for Proposition 4.3

	1	2	3	4	$\cdots$	r
$S \setminus T$	$\overline{x_1}$	x_2	x_3	x_4	$\cdots$	x_r
$T \setminus S$	x_1	$\overline{x_2}$	x_3	x_4	$\cdots$	x_r
$N \setminus (S \cup T)$	x_1	x_2	$\overline{x_3}$	x_4	$\cdots$	x_r
$S \cap T$	$\overline{x_1}$	$\overline{x_2}$	$\overline{x_3}$	x_4	$\cdots$	x_r
	x_1	x_2	x_3	x_4	$\cdots$	x_r

1. *X is an affine subspace.*
2. *For every MIPE, every partial evaluation at distance 2 from it is infeasible.*
3. *For every MIPE, every partial evaluation at distance 2 from it is a MIPE.*

Proof. In the proof of Claim 3.5 we showed that 2 implies 3, which in turn implies 1. Let us show that 1 implies 2. Let $x = (x_j)_{j \in K}$ be a MIPE, and let y be at distance 2 from x. Suppose, for the sake of contradiction, that y is feasible. Then each of the three rows of individual K-evaluations in Table 3 is feasible (again, assuming that $K = \{1, \dots, r\}$ and the two issues where y differs from x are 1 and 2). As X is an affine subspace, it is closed under addition of triples, and therefore the sum of these three rows, which is x, is also feasible. This contradicts the fact that x is a MIPE. □

Our main result showed that among all sets X that are totally blocked, the only ones that admit any non-dictatorial aggregator which is IIA and Paretian are the affine subspaces. It is natural to ask what form do the aggregators take in this case. The proof exhibited one family of aggregators of the form $f(\mathbf{x}) = \sum_{i \in R} x^i$, where R is a subset of N of odd cardinality. We call this an *R-parity* rule, because the decision on any issue is that which is supported by an odd number of individuals in R. Our next result shows that these are all the IIA and Paretian aggregators that are enabled in the case under consideration.

Proposition 4.3. *Let X be a non-degenerate subset of $\{0, 1\}^m$. Assume that X is totally blocked and is an affine subspace. Then for any society N, the family of IIA and Paretian aggregators for N over X coincides with the family of R-parity rules, where R ranges over the odd-cardinality subsets of N.*

Proof. We already know that R-parity rules constitute IIA and Paretian aggregators when X is an affine subspace.

Conversely, suppose that $X \subseteq \{0, 1\}^m$ has the properties stated in the proposition, and $N = \{1, \dots, n\}$ is a society. Let $f : X^n \to X$ be an IIA and Paretian aggregator. From the proof of the main result, we know that f is neutral, and the corresponding collection $\mathcal{W}$ of winning coalitions is self-dual. We also know that there exists a MIPE of length at least three.

Let $\mathcal{L}$ be the collection of losing coalitions, that is, the subsets of N that are not in $\mathcal{W}$. We show now that $\mathcal{L}$ is closed under symmetric difference: if $S, T \in \mathcal{L}$ then $S \triangle T \in \mathcal{L}$, where $S \triangle T = (S \setminus T) \cup (T \setminus S)$. Suppose, for the sake of contradiction, that $S, T \in \mathcal{L}$ but $S \triangle T \in \mathcal{W}$. Let $x = (x_j)_{j \in K}$ be a MIPE of length at least three. Consider the construction in Table 4 (where for ease of exposition $K = \{1, \dots, r\}$).

Observe that each of the first three evaluations in the table differs from the MIPE x in exactly one place, and is therefore feasible. The fourth one is the sum of the first three, and since X is an affine subspace, it is also feasible. The resulting social positions are determined in the first three columns by our assumptions on S, T, and $S \triangle T$, respectively, and in the remaining columns by the Pareto property. Thus, the social K-evaluation equals x, which is a contradiction since x is infeasible.

In terms of the characteristic vectors in $\{0,1\}^n$ that are associated with coalitions, symmetric difference corresponds to addition modulo 2. So what we showed above means that the characteristic vectors of the losing coalitions form a linear subspace of $\{0,1\}^n$. By self-duality, the cardinality of this subspace is 2^{n-1}, and therefore it is the solution space of one homogeneous linear equation, say $\sum_{i=1}^{n} a^i x^i = 0$ (we use superscripts to be consistent with our earlier conventions). Taking R to be the set of $i \in N$ for which $a^i = 1$, this means that $\mathcal{W} = \{S \subseteq N \mid |S \cap R| \text{ is odd}\}$. As $N \in \mathcal{W}$, the cardinality of R is odd. Hence f has the required form. □

We end this section with a characterization of the sets X that admit a non-dictatorial aggregator which is not only IIA and Paretian but also neutral. The monotone version of this characterization (that is, when monotonicity is added to the list of properties required of the aggregator) was obtained by Nehring and Puppe [13]. They proved that such an aggregator exists if and only if X is a median space. By definition, X is a *median space* if for every three evaluations in X, their coordinate-wise median (adopting on each issue the majority view) is also in X. This is equivalent, as shown in [13], to the property that all MIPEs are of length two. The result below (suggested to us by Nehring and Puppe) shows that affine subspaces are the only type of spaces, other than median spaces, that admit IIA and Paretian aggregators which are neutral and non-dictatorial.

Proposition 4.4. *Let X be a non-degenerate subset of $\{0,1\}^m$. Then there exists an aggregator $f : X^n \to X$, for some n, which is IIA, Paretian, neutral and non-dictatorial, if and only if X is a median space or an affine subspace.*

Proof. If X is a median space we can use majority rule as an aggregator with the required properties (this works, by the definition of a median space, for $n = 3$, but in fact it works for any odd $n \geqslant 3$). If X is an affine subspace we can use parity rule as an aggregator with the required properties (this, too, works for any odd $n \geqslant 3$).

Conversely, suppose that $f : X^n \to X$ has the required properties, but X is neither a median space nor an affine subspace. We refer to the proof of the sufficiency part of our main result. There we used total blockedness (which we do not assume here) to deduce neutrality of f and the existence of a MIPE of length at least three. Here we have neutrality as an assumption, and a MIPE of length at least three exists because X is not a median space. The other steps of the proof there are valid here as well, and lead to the conclusion that f is dictatorial, contrary to our assumption. □

5. Relation to other works

Wilson [17] was the first to study the general aggregation problem treated in this paper. He required the same properties of aggregators as we do (IIA and Pareto). He proved that under a certain condition (being a “frame”) on the set of feasible evaluations, every such aggregator is a parity rule defined with respect to some odd-cardinality subset of the society. He went on to show that under some further condition such an aggregator must actually be dictatorial. His conditions on the set of feasible evaluations were sufficient to determine the form of aggregators,

Please cite this article in press as: E. Dokow, R. Holzman, Aggregation of binary evaluations, Journal of Economic Theory (2008), doi:10.1016/j.jet.2007.10.004

but were not necessary and sufficient as ours are. The fundamental concept used in his conditions, called a "single frame," amounts in our terminology to a pair of MIPEs of the form $x = (x_j)_{j \in K}$, $y = (\overline{x_j})_{j \in K}$ with $|K| \geqslant 3$. In the application to preference aggregation, MIPEs come in pairs like that (see our Example 5), and this enabled Wilson to deduce Arrow's theorem, as well as some extensions. But in general such antipodal pairs of MIPEs may not exist, and this limits the applicability of Wilson's results.

Rubinstein and Fishburn [16] introduced the algebraic framework for the theory of aggregation. Their model was more general than the one treated here, in that the evaluations were vectors over an arbitrary field. However, part of their work was specialized to the case of the binary field, and was devoted to conditions on the set X of feasible evaluations that force every IIA and Paretian aggregator to be dictatorial. This part of their work is directly comparable to our Theorem 2.2. They showed that each of two alternative conditions on the set X (being a "W_0-set" or a "W_1-set") suffices to force the aggregation to be dictatorial.[5] Like Wilson's conditions, the conditions found by Rubinstein and Fishburn sufficed in order to deduce Arrow's theorem, but were not necessary and sufficient for X to be an impossibility domain.

We owe the concept of total blockedness to Nehring and Puppe [13]. They introduced it in a different model, where individuals have single-peaked preferences over the evaluations in X, and the objects studied are strategy-proof social choice functions mapping profiles of such preferences into X. Though the models are different, one of their main results translates to the following result in the model considered here: if X is non-degenerate, a necessary and sufficient condition for every monotone IIA and Paretian aggregator $f : X^n \to X$ to be dictatorial is that X be totally blocked. The difference between this and our main result is that we do not assume monotonicity of the aggregator. Without the monotonicity assumption, total blockedness is no longer a sufficient condition, and this led us to add the condition of not being an affine subspace. In a subsequent paper, Nehring [12] applied their result to the problem of preference aggregation, in the same way as we did in Example 5. However, because he needed monotonicity as an assumption, he did not obtain a proof of Arrow's theorem, but rather a weaker version of it in which "monotone IIA" was substituted for IIA.

There is a growing body of recent literature on the problem of judgment aggregation. In this literature the problem is described in terms of a set $\mathcal{P}$ of propositions, in some logical language, that need to be evaluated subject to the constraint of logical consistency. As explained in the introduction, any such problem (with finite $\mathcal{P}$) may be cast in the framework of our paper by letting X be the set of logically consistent evaluations. In fact, this reformulation captures the essential aspects of the aggregation problem at hand. By now there exist quite a number of different impossibility theorems for judgment aggregation, and we do not survey them individually here. In some of these theorems, the list of properties of an aggregator for which impossibility is proved includes properties that we (following Arrow) do not assume. To be specific, List and Pettit [10] assumed anonymity and systematicity (the latter amounts to IIA and neutrality); Dietrich [4] and Dietrich and List [5, Theorem 1] assumed systematicity; Nehring and Puppe [14] assumed monotonicity. In another group of impossibility theorems, including Pauly and van Hees [15][6] and Dietrich [3], there are no extra properties of the aggregator, but the conditions on the agenda

[5] They defined a W_0-set as a set $X \subseteq \{0,1\}^m$ with $|X| \geqslant 3$ for which there exists $y \in \{0,1\}^m$ so that every $x \in X$ differs from y in exactly one coordinate. Their definition of a W_1-set was more involved and somewhat careless, as it does not prevent it from being an affine subspace. Therefore, their result is not valid as stated for W_1-sets.

[6] Pauly and van Hees extended their theorem to a model in which the evaluations may take more than two values and are subject to the rules of a certain many-valued logic.

Please cite this article in press as: E. Dokow, R. Holzman, Aggregation of binary evaluations, Journal of Economic Theory (2008), doi:10.1016/j.jet.2007.10.004

are stronger than ours. We note, though, that some of the above-mentioned theorems do not assume the Pareto property (in some cases weaker versions of it are assumed). The impossibility theorem of Mongin [11] stands out in not requiring the full IIA property.

Finally, we would like to mention the work of Beigman [2]. He addressed a problem of preference aggregation due to Kalai [9]. This problem generalizes the classical setup, by assuming that the domain of permissible preferences (both for the individuals and for the society) is given by an arbitrary class of tournaments on the set of alternatives A. In the classical setup, the domain is given by the class of transitive tournaments. Kalai expected that Arrow's theorem would extend to any proper subclass of the class of all tournaments on A that is closed under permutations of A, provided that $|A| \geqslant 4$. Beigman showed that this is not true in general, but becomes true if the aggregator is assumed to be monotone, or alternatively non-neutral. We point out that Beigman's results can be better understood and sharpened using our results. Indeed, as we did for the classical setup, any domain of preferences may be encoded as a set X of 0–1 vectors of length $\binom{k}{2}$, where $k = |A|$. It can be checked that any domain satisfying Kalai's conditions gives rise to a totally blocked X. However, X may be an affine subspace and then the analogue of Arrow's theorem fails. Moreover, we can characterize all domains for which this happens. In particular, it turns out (somewhat curiously) that the analogue of Arrow's theorem does hold true in general if $|A| \equiv 2 \pmod 4$.

References

[1] K.J. Arrow, Social Choice and Individual Values, Wiley, New York, 1951.
[2] E. Beigman, Extension of Arrow's theorem to symmetric sets of tournaments, Discrete Math. 307 (2007) 2074–2081.
[3] F. Dietrich, Judgment aggregation: (Im)possibility theorems, J. Econ. Theory 126 (2006) 286–298.
[4] F. Dietrich, A generalised model of judgment aggregation, Soc. Choice Welfare 28 (2007) 529–565.
[5] F. Dietrich, C. List, Arrow's theorem in judgment aggregation, Soc. Choice Welfare 29 (2007) 19–33.
[6] E. Dokow, R. Holzman, Aggregation of binary evaluations with abstentions, Working Paper, Technion-Israel Institute of Technology, 2006.
[7] E. Dokow, R. Holzman, Aggregation of binary evaluations for truth-functional agendas, Soc. Choice Welfare, in press.
[8] P.C. Fishburn, A. Rubinstein, Aggregation of equivalence relations, J. Classification 3 (1986) 61–65.
[9] G. Kalai, Social choice without rationality, reviewed in NAJ Econ. 3 (2001).
[10] C. List, P. Pettit, Aggregating sets of judgments: An impossibility result, Econ. Philos. 18 (2002) 89–110.
[11] P. Mongin, Factoring out the impossibility of logical aggregation, J. Econ. Theory, in press.
[12] K. Nehring, Arrow's theorem as a corollary, Econ. Letters 80 (2003) 379–382.
[13] K. Nehring, C. Puppe, Strategy-proof social choice on single-peaked domains: Possibility, impossibility and the space between, Working Paper, University of California at Davis, 2002.
[14] K. Nehring, C. Puppe, Consistent judgement aggregation: The truth-functional case, Soc. Choice Welfare 31 (2008) 41–57.
[15] M. Pauly, M. van Hees, Logical constraints on judgment aggregation, J. Philos. Logic 35 (2006) 569–585.
[16] A. Rubinstein, P.C. Fishburn, Algebraic aggregation theory, J. Econ. Theory 38 (1986) 63–77.
[17] R. Wilson, On the theory of aggregation, J. Econ. Theory 10 (1975) 89–99.

Please cite this article in press as: E. Dokow, R. Holzman, Aggregation of binary evaluations, Journal of Economic Theory (2008), doi:10.1016/j.jet.2007.10.004