Journal of Mathematical Economics 46 (2010) 356–363

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

The Computational Complexity of Rationalizing Behavior[☆]

Jose Apesteguia[a,*], Miguel A. Ballester[b]

[a] *Universitat Pompeu Fabra*
[b] *Universitat Autónoma de Barcelona*

ARTICLE INFO

Article history:
Received 23 October 2008
Received in revised form 11 January 2010
Accepted 2 February 2010
Available online 10 February 2010

JEL classification:
D00

Keywords:
Rationalization
Computational complexity
NP-complete
Arbitrary choice domains

ABSTRACT

We study the computational complexity of rationalizing choice behavior. We do so by analyzing two polar cases, and a number of intermediate ones. In our most structured case, that is where choice behavior is defined in universal choice domains and satisfies the "weak axiom of revealed preference," finding the complete preorder rationalizing choice behavior is a simple matter. In the polar case, where no restriction whatsoever is imposed, either on choice behavior or on choice domain, finding a collection of complete preorders that rationalizes behavior turns out to be intractable. We also show that the task of finding the rationalizing complete preorders is equivalent to a graph problem. This allows the search for existing algorithms in the graph theory literature, for the rationalization of choice.

1. Introduction

The theory of individual decision making is typically formulated on the basis of two different approaches. A revealed preference approach that directly studies individual choice behavior. Or a preference relation approach, where tastes are summarized through binary relations. Clearly, the relation between these two distinct formal approaches to individual behavior is a fundamental question in economics. It is said that choice behavior is rationalized whenever the two approaches yield the same results, in every possible choice problem that may arise.

Our objective is to study the problem of the rationalization of different classes of choice patterns that may or may not be consistent. Therefore, we need a rationalizing method applicable to any possible choice pattern. Kalai, Rubinstein, and Spiegler (2002; hereafter KRS) provide such a method. In KRS choice behavior is rationalized by a collection of preference relations (rationales), such that for every choice problem A in the domain of feasible choice problems, the choice $c(A)$ is maximal in A for some rationale in the collection. That is, it is as if the decision-maker partitions the set of choice problems into different categories and applies one rationale to each category in the partition. Of importance for our purposes, it is immediate that every possible choice pattern admits a rationalization a la KRS. Further, there are multiple collections of rationales that rationalize a given choice rule. KRS naturally propose to focus on those collections that use the minimal number of rationales. Consequently, when choice behavior is fully consistent, i.e. it satisfies the property known

☆ This paper supercedes a previous paper with the title "Minimal Books of Rationales." We thank Jon Benito, Ricard Gavalda, Ignacio Palacios-Huerta, Xavier Vila and especially Coralio Ballester for his key contribution to the proof of Theorem 2. Financial support by the Spanish Commission of Science and Technology CICYT (ECO2009-12836, ECO2008-04756), the Barcelona GSE research network, and the Government of Catalonia is gratefully acknowledged.

* Corresponding author.
E-mail address: jose.apesteguia@upf.edu (J. Apesteguia).

0304-4068/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jmateco.2010.02.001

as the "weak axiom of revealed preference" (WARP), the minimal number of rationales that rationalizes choice is one, and we are back in the classical world. If behavior is inconsistent, then the minimal number of rationales will be above one.

Drawing on the tools of theoretical computer science, we study the question of how complex it is to find a minimal collection of rationales that rationalizes choice behavior. We analyze two polar cases and some intermediate ones. In the first place we study what we call the rational procedure, i.e. when WARP holds. We show that finding the unique rationale in such a case is a simple matter, as there are algorithms that can easily construct it.

We then turn the analysis to the polar case where no restriction whatsoever is imposed, neither on choice behavior nor on the choice domain. In this case, we show that, contrary to the previous one, finding a minimal collection of rationales is a difficult computational problem (see Theorem 2). That is, there is little hope for the existence of an algorithm that for every possible choice rule finds a minimal collection in a reasonable time frame. Hence, in this extremely unstructured case, the task of finding the collection of preferences that rationalizes behavior is in general intractable.

Now, the question arises whether it is the conjunction of (i) unstructured choice behavior and (ii) unrestricted choice domain, that leads to the computational hardness of the problem of rationalization. In principle, it could be that the real difficulty in finding a minimal collection of rationales is completely triggered by choice behavior per se, or it could be that it is the interplay of behavior and domain that drives the result. It turns out that the answer to this question seems to depend on the single or multi-valued nature of the choice correspondence. Theorem 4 suggests that in the case of single-valued choice correspondences, the essence of the intractability of rationalization is triggered by the interplay of both unstructured behavior and unrestricted domain. However, we argue that in the case of multi-valued choice correspondences the practical difficulty of rationalizing behavior may be due to the nature of choice behavior per se.

The challenge is then to understand the driving force of the computational complexity of rationalization. If we are able to understand the roots of the complexity in rationalizing choice behavior, we may use this to search for specific algorithms that behave well under certain circumstances. We start this challenge by defining two binary relations on the space of choice problems, that capture two fundamental properties on the structural relation of choice problems. By doing so we will be able to draw a connection with a graph theory problem (see Theorem 7). This is especially useful since there is a wealth of algorithms for graph problems that may be used to solve the problem of rationalization of certain choice structures.

We then end the analysis by exploring a specific case, inspired by the work of Manzini and Mariotti (2007; hereafter MM). MM study the nature of choice behavior that can be rationalized by sequentially applying a fixed set of asymmetric binary relations. Among other results, they provide a characterization for the case when choice is sequentially rationalizable by two rationales. Such a choice rule is called a Rational Shortlist Method (RSM).[1] Using the tools derived for establishing the connection with graph theory, we show that the rationalization of RSMs through multiple rationales turns out to be a computationally tractable problem.

There is very little work on the relationship between complexity and rationality. We will now discuss some relevant exceptions. An early proposal for the measurement of the complexity of decision-making is Futia (1977). Futia proposes two methods, one algebraic in nature and one topological in nature, in order to take into account the requirements different rules of thumb make in terms of computational resources, time, attention and imagination. Both methods suggest that rules of thumb are procedures of low complexity. Campbell (1978a,b) studies the class of choice correspondences for which the chosen element can be computed by procedural methods, with good adaptability properties, that do not waste time, and that yield satisfactory intermediate alternatives. He shows that the more structured choice behavior is, the better properties the associated procedural method has. Salant (2003; see also Salant, forthcoming) studies two computational aspects of choice: the amount of memory choice behavior requires, and the computational power needed for the computation of choice. He shows that the rational procedure is favored by these considerations. Johnson (2006; see also Johnson 1990, 1995) uses a semiautomaton model to investigate algebraic complexity and computational complexity issues inherent in specific classes of choice domains.[2] With regard to algebraic complexity, he asks how powerful the semiautomaton that implements the choice correspondence has to be. Johnson shows that the more structure on the class of choice correspondences, the environments in which the semiautomaton operates become simpler. With regard to computational complexity, the question Johnson addresses is how hard is to construct the semiautomaton that will implement a particular choice correspondence. Interestingly, here the conclusions reverse. That is, the more structured choice behavior is, the more complex it is to construct the semiautomaton. Johnson (2006) shows, therefore, that the notion of complexity is critical. Finally, Mandler (2009) studies a procedure where the decision-maker chooses by proceeding sequentially through a list of criteria.[3] He shows that a rational decision-maker needs the minimum number of criteria. Any deviation from rationality needs at least as many criteria, and on some domains must use strictly more. Therefore, if one measures computational burden by the minimum number of required criteria, then rationality is less computationally demanding than irrationality.

---

[1] Apesteguia and Ballester (2009) offer a property that is equivalent to the notion of choice by sequential procedures, regardless of the number of rationales required.

[2] Concretely, he investigates those rationalized by a linear order, those rationalized by a preorder, those rationalized by a quasi-transitive relation, and path independent choice correspondences.

[3] See also, Mandler et al. (2009).

The rest of the paper is organized as follows. Section 2 gives the definitions on choice behavior and binary relations, and makes precise the notion of rationalization we will use throughout the paper. Furthermore, it contains a brief introduction to the theory of computational complexity. Section 3 contains the complexity results. In Section 4 we draw on the connection between the problem of rationalization and the literature on graph theory. Finally, Section 5 concludes.

2. Preliminaries

2.1. Choice behavior and preference relations

Let X be a finite set of n objects. We denote by $\mathcal{U}$ the set of all non-empty subsets of X. We will often refer to $\mathcal{U}$ as the universal choice domain. We also consider the general case of arbitrary domains $\mathcal{D} \subseteq \mathcal{U}$, with $\mathcal{D} \neq \varnothing$. A choice correspondence c on $\mathcal{D}$ assigns to every $A \in \mathcal{D}$ a non-empty set of elements $c(A) \subseteq A$. Whenever for every $A \in \mathcal{D}$, $|c(A)| = 1$, we say that the choice correspondence c is single-valued. Define $H(A) = A \setminus c(A), A \in \mathcal{D}$. That is, $H(A)$ denotes the unchosen elements of A. For the computational treatment, let that c is described explicitly by listing for all $A \in \mathcal{D}$, the elements $c(A)$.

In order to elaborate on the connection between choice behavior and preference relations we first introduce some notation. Denote by $\succcurlyeq$ a binary relation on X, $\succcurlyeq \subseteq X \times X$. Binary relations $\succ$ and $\sim$ are the asymmetric and symmetric parts of $\succcurlyeq$, respectively. Hence, $x \succ y$ if and only if $x \succcurlyeq y$ and $\neg(y \succcurlyeq x)$, and $x \sim y$ if and only if $x \succcurlyeq y$ and $y \succcurlyeq x$. We will say that a binary relation is a preorder if it is reflexive and transitive. We will often refer to complete preorders by 'rationales.' We will say that a binary relation is a linear order if it is antisymmetric, transitive, and complete. Finally, we will say that it is a partial order if it is reflexive, antisymmetric, and transitive. For any $A \in \mathcal{D}$, $M(A, \succcurlyeq)$ denotes the set of maximal elements in A with respect to $\succcurlyeq$, that is, $M(A, \succcurlyeq) = \{x \in A : y \succ x \text{ for no} y \in A\}$. Let $c(A) \succcurlyeq c(B), A, B \in \mathcal{D}$, denote the case when for every $x \in c(A)$ and $y \in c(B)$, we have $x \succcurlyeq y$.

KRS propose the rationalization of any possible choice pattern through collections of rationales. The decision-maker has in mind a partition of the set of choice problems and applies one rationale to each category in the partition. A category may represent a state of the world, that once internalized triggers the application of a particular rationale. The interest is naturally directed to the minimal number of rationales that rationalizes choice behavior. Here we use this general notion of rationalization. To this end we extend the original definition to include multi-valued choice correspondences and arbitrary choice domains $\mathcal{D}$. Accordingly, we also substitute linear orders by complete preorders.

Minimal Rationalization by Multiple Rationales (RMR): A K-tuple of complete preorders $\{\succcurlyeq_k\}_{k=1,\dots,K}$ on X is a rationalization by multiple rationales (RMR) of c if for every $A \in \mathcal{D}$, there is a $k \in \{1, \dots, K\}$, such that $c(A) = M(A, \succcurlyeq_k)$. It is said to be minimal if any other RMR of c has at least K complete preorders.

Note that a minimal RMR is an extension of the classic idea of rationalization by one rationale. In fact, if c satisfies WARP,[4] then the minimal RMR is composed of one complete preorder.

We highlight an especially important class of choice sets; the collection of c-maximal sets. These are the sets that must be explained in order to rationalize behavior. A set A is c-maximal if any addition to A of elements (consistent with the domain of choice problems $\mathcal{D}$) leads to a change in choice behavior with respect to the original elements of A. Formally,

c-Maximal Sets: A subset $A \in \mathcal{D}$ is said to be c-maximal if for all $B \in \mathcal{D}$, with $A \subset B$, it is the case that $c(A) \neq c(B) \cap A$. Denote the family of c-maximal sets under the choice domain $\mathcal{D}$ by $M_c^{\mathcal{D}}$.

2.2. Computational complexity

There are different notions of complexity of relevance to economics. We already referred to the notion of algebraic (or implementation) complexity. Another prominent example is the notion of descriptive (or Kolmogorov) complexity, that measures the complexity of a string by the length of the shortest description of the string in some descriptive language. In this paper we adopt the notion of computational complexity, and accordingly make use of the theory of NP-completeness. We now present an informal introduction to the notion of NP-completenesss. For an excellent, detailed and formal account see Garey and Johnson (1979).[5]

In theoretical computer science the computational complexity of a problem is measured by the relation between the size of the problem and the time required to solve it. There are different classes of problems. The *class P* encompasses problems that can be solved in polynomial time. That is, a problem belongs to the class P if there is an algorithm that solves the problem, and the time required is upper bounded by a polynomial function of the size of the problem. More precisely, a time algorithm $f(t)$ is polynomial if there is a polynomial function p such that $|f(t)| \leq |p(t)|$ for every $t \geq 0$, with t denoting the size (or the input length) of the problem. In this sense we write that the complexity function of $f(t)$ is $O(p(t))$.

The *class NP* is the family of problems that may be difficult to solve (it may take exponential time), although it is relatively easy (it takes polynomial time) to *verify* a particular instance of the problem (that is, a concrete case where the parameters have been realized).

---

[4] **Weak Axiom of Revealed Preference (WARP)**: Let $A, B \in \mathcal{D}$ and assume $x, y \in A \cap B$; if $x \in c(A)$ and $y \in c(B)$, then we must also have $x \in c(B)$.

[5] See also Cormen et al. (2001). Ballester (2004) and Aragones et al. (2005) provide introductions in the context of economics.

Within the class NP there is one class of problems of particular interest. This is the class of *NP-complete* problems. The NP-complete problems are regarded as the most difficult problems in NP. This is because no polynomial time algorithm is known to solve any of them, but if a polynomial algorithm were found for one problem, then such an algorithm could be translated into polynomial time algorithms for all other problems in the NP class. In this sense only exponential time algorithms are known for NP-complete problems, and these problems are regarded as *intractable*. It is interesting to note that, although there is no known polynomial time algorithm solving an NP-complete problem, a proof showing that no such algorithm exists is still awaited. In fact, this is regarded as one of the major unsolved problems in mathematics.

We signify a final type of time bounds that is super-polynomial and sub-exponential: the quasi-polynomial type. A quasi-polynomial time algorithm has time complexity $O(t^{\text{polylog}t})$, where polylog$t$ is a polynomial in $\log t$. An example of a quasi-polynomial function is $t^{\log t}$.

The theory of NP-completeness is conventionally centered around *decision problems*. These are problems formulated with a yes-or-no answer. Consequently, we define a decision problem that is the binary analog of the problem of finding a minimal RMR of a choice correspondence c on $\mathcal{D}$ as follows.

Rationalization (RAT): Given a choice correspondence c on $\mathcal{D}$, and an integer K, can we find $k \leq K$ complete preorders that constitute a rationalization by multiple rationales of c?

In occasions, we will refer to the especial cases of RAT where $\mathcal{D} = \mathcal{U}$, or where the choice correspondence c is single-valued, or where c satisfies some consistency property.

3. Computational complexity of rationalization

3.1. Structured choice behavior: The rational procedure

We start by considering the case where a choice correspondence defined over the universal choice domain $\mathcal{U}$ satisfies WARP. We have already mentioned that this represents a very structured case, as it is well-known that there exists a unique complete preorder relation $\succcurlyeq$ rationalizing c.

Now the question arises of how difficult it is to find the rationalizing binary relation $\succcurlyeq$. Intuitively, it seems that the high degree of structure of the rational procedure implies that finding the rationale is not a difficult task. This is in fact the case. It is easy to show that the set $M_c^{\mathcal{U}}$ is small, as it contains at most n elements. The simplicity of the family of maximal sets allows us to consider very simple algorithms to obtain the rationalization of choice behavior in polynomial time. Consider for instance the following trivial one.

Take $X_0 = X$ and iteratively define $X_k = X_{k-1} \setminus c(X_{k-1})$ if $X_{k-1} \setminus c(X_{k-1}) \neq \varnothing$ holds. Then, the set of maximal elements is the collection of sets $\{X_k\}$, with $|\{X_k\}| \leq n$. Clearly, every element is chosen from exactly one set X_k. Then, the rationale can be defined by stating that $x \succcurlyeq y$ if and only if $x \in X_l$, $y \in X_j$, with X_l, X_j in the family of c-maximal sets $\{X_k\}$ and $l \leq j$.

We summarize the above in the following observation.

Observation 1. Let the choice correspondence c satisfy WARP and $\mathcal{D} = \mathcal{U}$. Then $|M_c^{\mathcal{U}}| \leq n$ and RAT is polynomial.

There are two important remarks to Observation 1. First, as a corollary to the above, finding the rationale rationalizing a single-valued choice correspondence defined on $\mathcal{U}$ that satisfies WARP is also polynomial. This follows from the fact that the single-valued case is but a special case of the multi-valued choice correspondence.

Second, RAT remains polynomial when c is defined in arbitrary choice domains $\mathcal{D}$. To find the rationale rationalizing c when $\mathcal{D} \neq \mathcal{U}$, simply write $x \succcurlyeq y$ if and only if $x \in c(A)$ and $y \in A$, $A \in \mathcal{D}$. It is easy to see that WARP guarantees that such a rationale is a complete preorder.

3.2. Unstructured behavior and unrestricted domain

We now turn to the polar case of the rational procedure where we neither impose structure on choice behavior, nor on the domain of choice sets. This means that, in general, there is not a single rationale rationalizing choice behavior, but a set of them in the sense of KRS. It is clear that the problem of finding the rationales in this case is a much more demanding task than the one we faced for the rational procedure in the previous section. In fact we show below that the task is demanding to the point of being intractable. That is, we show that finding a minimal RMR in this setting belongs to the class of NP-complete problems, and hence unless P=NP, there is no hope of finding an efficient algorithm that for every choice correspondence c gives a minimal RMR in a reasonable time frame.

We can now state our first NP-completeness result.

Theorem 2. *RAT is NP-complete.*

Proof of Theorem 2: It is easy to show that RAT is in NP. A candidate for a solution consists of $k \leq K$ complete preorders that can be easily checked to rationalize the choice corrrespondence c. Now consider the following problem that is known to be NP-complete.[6]

[6] See Karp (1972). See also Garey and Johnson (1979).

Partition into Cliques (PIC): Given a graph $G = (V, E)$, and an integer K, can the vertices of G be partitioned into $k \leq K$ disjoint sets $V_1, V_2, \ldots, V_k$ such that for $1 \leq i \leq k$ the subgraph induced by V_i is a complete graph?

We now find a restriction of RAT that makes it identical to PIC. First, define the auxiliary set $H = \{h_1,\ldots,h_{|V|}\}$. Now, let X be the union of the set of vertices V and the set H. For each vertex $i \in V$ define the set $S_i = \{i\} \cup \{j : (i, j) \notin E\} \cup \{h_i\}$ and denote $\mathcal{S} = \{S_i : i \in V\}$. Choice behavior is defined as follows. For every $S_i \in \mathcal{S}$ let $c(S_i) = \{i\}$. Clearly, this defines a single-valued choice correspondence in $\mathcal{D}$. Note that the construction of X, $\mathcal{D}$, and c can be done in polynomial time in the size of $|V|$.

We now prove that G can be partitioned into less than K cliques if and only if c can be rationalized by less than K rationales. First, consider a collection $\{\succcurlyeq_p\}_{p=1,\ldots,k}$ with $k \leq K$ complete preorders rationalizing c. Consider a single-valued mapping (there may be several) $f : \mathcal{S} \to \{1, \ldots, k\}$ such that $f(S_i) = p$ if and only if $\succcurlyeq_p$ rationalizes the choice in S_i. Define $V_p = \{i : f(S_i) = p\}$. Clearly, the union of all sets V_p is a partition of the vertex set V. Now, we show that V_p is a clique. Otherwise, let $i, j \in V_p$ such that $(i, j) \notin E$. Then, $i, j \in S_i \cap S_j$ and hence, given the definition of c over the class $\mathcal{S}$, for the rationalization of S_i and S_j it must hold that $i \succ_p j$ and $j \succ_p i$. But this contradicts that $\succcurlyeq_p$ is a complete preorder. We conclude that V_p is a clique, as desired. Therefore the partition of G is composed of at most k cliques, and hence there is obviously a PIC of G with at most K cliques.

In the other direction, let $\{V_p\}_{p=1,\ldots,k}$ be a partition of G into $k \leq K$ cliques. Define for every $p = 1, \ldots, k$, the partial order $\succcurlyeq'_p$ by: $x \succcurlyeq'_p y$ if and only if $[x \in V_p$ or $y \in X \setminus V_p]$. Take any linear extension of $\succcurlyeq'_p$ and denote it by $\succcurlyeq_p$. We now have to show that $\{\succcurlyeq_p\}_{p=1,\ldots,k}$ rationalizes all the selections. We claim that the collection of sets S_i, $i \in V_p$, is rationalized by $\succcurlyeq_p$. Otherwise, there would be a set S_i with $i \in V_p$, and an element $w \in S_i$, $w \neq i$ with $w \succ_p i$. This implies that given the definition of $\succcurlyeq_p$, since $i \in V_p$, it must be $w \in V_p$. Since V_p is a clique it must be that $(i, w) \in E$, but this contradicts that $w \in S_i$. This shows that there is a collection of at most K complete preorders that rationalizes c. □

Note that the c used in the proof is single-valued and the rationales defined from the partition of the graph are linear orders. Hence, the following corollary to (the proof of) Theorem 2 is immediate.

Corollary 3. *RAT is NP-complete in the subclass of single-valued choice correspondences.*

Theorem 2 (and Corollary 3) show that the conjunction of (i) unstructured choice behavior and (ii) unrestricted choice domain lead to the NP-completeness of the problem of rationalization. But are the two conditions required to get the intractability result? In principle, it could be that the real difficulty in finding a minimal RMR is completely triggered by choice behavior per se, or it could be that it is the interplay of behavior and domain that drives the result. Theorem 4 below suggests that in the case of single-valued choice correspondences it is the interplay of both, unstructured behavior and unrestricted domain, that triggers the intractability of finding the rationales.

Theorem 4. *RAT is quasi-polynomially bounded in the subclass of single-valued choice correspondences defined on the universal choice domain $\mathcal{U}$.*

Proof of Theorem 4: Let c be a single-valued choice correspondence defined on $\mathcal{U}$, and consider the following naive algorithm. The naive algorithm would examine all the k-tuples of order relations, with $1 \leq k \leq n-2$. That is, it will start by checking whether there is one rationale that rationalizes behavior. If yes, the algorithm stops. Otherwise, it checks whether there are two rationales rationalizing behavior, and so on. Propositions 1 and 2 in KRS show that the minimal collection is composed of at most $n-1$ rationales, and that in fact this bound is tight. We now proceed to compute the order of growth of the algorithm's runtime. The main components of the input size and the number of operations are the number of choice sets and the number of collections of rationales to check. That is, since the choice domain is the universal one, the input size x of the algorithm is roughly 2^n. The naive algorithm requires a number of operations t which is roughly $n!^{n-1}$.[7]

Now, to check the order of growth of the algorithm's runtime we relate t to x^α. This gives an α that is $O(\log n!)$ or equivalently $O(\log x \log\log x)$. Hence, it corresponds to the case where the naive algorithm has a *quasi-polynomial* order of magnitude of $O(x^{\log x \log\log x})$. □

Theorem 4 suggests that the problem RAT is not NP-complete in the subclass of single-valued choice correspondences defined on the universal choice domain $\mathcal{U}$.[8] Otherwise, Theorem 4 would imply that *all* NP-complete problems are quasi-polynomially bounded. However, in spite of continuous efforts such a bound has never been found for NP-complete problems. Indeed, there is the strong conviction this will never happen.

The question with regard to multi-valued choice correspondences defined on $\mathcal{U}$, however, remains open. The naive algorithm used in Theorem 4 seems not to be quasi-polynomially bounded in this case. It is not difficult to see that the number of possible complete preorders is upper bounded by $n!^2$. The problem arises because, in the case of complete preorders, the $n-1$ bound on the maximum number of rationales to check for rationalization does not hold. In fact, the bound may be of the order of 2^n. This would imply that the naive algorithm for choice correspondences cannot be quasi-polynomially

[7] These values take into account only the number of choice problems and the set of possible collections of $n-1$ linear orders. Note that a more precise formulation would include other considerations such as the number of elements in each choice set, the chosen element, a checking protocol for the rationalization of choice, or the length of the data expressed in bits. It is not difficult to see, however, that these considerations have no influence on the final conclusion. Hence, we avoid these details here.

[8] The naive algorithm of Theorem 4 does not necessarily behave well in unrestricted choice domains $\mathcal{D}$. The reason being that in this case the input size need not be as high as 2^n, but for example, it could be n. This gives an exponential order of magnitude for the naive algorithm.

bounded, having an exponential order of magnitude. Therefore, it may well be the case that with choice correspondences, the practical difficulty in finding a minimal RMR is triggered by choice behavior per se. Interestingly, this is the first case where the single-valued or multi-valued nature of the choice correspondence may play a significant role. We showed that in the rational case finding the preference relation that rationalizes choice behavior is a simple matter in both cases, and Theorem 2 and Corollary 3 show that in the polar case the problem of rationalization is NP-complete both with single-valued and with multi-valued choice correspondences.

The challenge for future research is then to understand the driving force of the complexity of rationalization, and use this understanding to find specific algorithms that behave well under certain circumstances. In the next section we start this task by drawing a connection with graph theory.

4. On the structure of the computational complexity of rationalization

4.1. Rationalization and graph theory

The following binary relations capture two fundamental properties on the structural relation of maximal sets.[9]

Definition 5. Let $A, B \in M_c^{\mathcal{D}}$, ARB if and only if $c(A) \cap B \notin \{\varnothing, c(B) \cap A\}$.

This first binary relation R states the conditions under which a set A *blocks* another set B, in the sense that A and B cannot be rationalized by a binary relation $\succcurlyeq$ writing $c(A) \succcurlyeq c(B)$. Then, when c is single-valued, R reduces to: ARB if and only if $c(A) \subseteq H(B)$.

Definition 6. *Let $A, B \in M_c^{\mathcal{D}}$, AQB if and only if $c(A) \cap c(B) \neq \varnothing$.*

AQB means that sets A and B are linked in the rationalization problem. That is, whenever AQB, if A and B are to be rationalized by the same rationale $\succcurlyeq$, then it must be that $c(A) \sim c(B)$.

Finally, we use the above two binary relations to define an oriented cycle.

Oriented Cycle: The collection $\{A_t\}_{t=1}^n \subseteq M_c^{\mathcal{D}}$, $n \geq 2$, is an oriented cycle if

(1) $A_1 = A_n$,
(2) for every $i \in \{1, \ldots, n-1\}$, either $A_i R A_{i+1}$, or $A_i Q A_{i+1}$, and
(3) there is $j \in \{1, \ldots, n-1\}$ such that $A_j R A_{j+1}$.

We are now in a position to introduce a graph theory problem over the space of c-maximal sets.

Minimal Non-Oriented Partition (NOP): A partition of $M_c^{\mathcal{D}}$, $\{V_p\}_{p=1,\ldots,P}$, is said to be a non-oriented partition if for every class V_p there is no oriented cycle. It is said to be minimal if any other NOP has at least P classes.

A (minimal) NOP is constructed over the set $M_c^{\mathcal{D}}$ according to the binary relations R and Q. Hence, a class V_p in the partition has a clear interpretation: all the sets in the class V_p can be rationalized through a complete preorder $\succcurlyeq_p$. Then, an NOP gives information on which choice problems can be rationalized together.

The following theorem establishes that finding a minimal RMR is equivalent to finding a minimal NOP. This opens the possibility of drawing upon the established algorithm knowledge on graph theory problems.

Theorem 7. Let c be a choice correspondence:

-If $\{\succcurlyeq_p\}_{p=1,\ldots,P}$ is a minimal RMR, then there is a minimal NOP $\{V_p\}_{p=1,\ldots,P}$.
-If $\{V_p\}_{p=1,\ldots,P}$ is a minimal NOP, then there is a minimal RMR $\{\succcurlyeq_p\}_{p=1,\ldots,P}$.

Proof of Theorem 7: Let $\{\succcurlyeq_p\}_{p=1,\ldots,P}$ be an RMR of c. Consider a single-valued mapping (there may be several) $f : M_c^{\mathcal{D}} \to \{1, \ldots, P\}$ such that $f(A) = j$ if and only if $\succcurlyeq_j$ rationalizes the choice in A. The mapping f naturally induces a partition of $M_c^{\mathcal{D}}$ containing at most P classes. Let $V_p = \{A_p^s\}_{s=1,2,\ldots,S_p}$ be the collection of c-maximal sets associated with class p, $1 \leq p \leq P$ (some of these collections may be empty). We now show that there is no oriented cycle in V_p. Assume the contrary is the case. That is let $\{A_t\}_{t=1}^n \subseteq M_c^{\mathcal{D}}$, $n \geq 2$ be an oriented cycle in V_p. Then, by definition, there is $j \in \{1,\ldots,n-1\}$ such that $A_j R A_{j+1}$, and hence $c(A_j) \cap A_{j+1} \notin \{\varnothing, c(A_{j+1}) \cap A_j\}$.

Assume that $c(A_{j+1}) \cap A_j \neq \varnothing$. Then $A_{j+1} R A_j$ and there must be an element $x \in X$ such that either $x \in c(A_j) \cap H(A_{j+1})$ or $x \in c(A_{j+1}) \cap H(A_j)$. Without loss of generality, assume the first case. Then for $\succcurlyeq_p$ to rationalize A_{j+1} it must be $c(A_{j+1}) \succ_p x$. Since $c(A_{j+1}) \cap A_j \neq \varnothing$, there is $y \in A_j$ such that $y \succ_p x$. But this contradicts the rationalization of A_j since $x \in c(A_j)$. Hence, it must be that $c(A_{j+1}) \cap A_j = \varnothing$, and therefore, for $\succcurlyeq_p$ to rationalize A_j and A_{j+1} it must hold $c(A_{j+1}) \succ_p c(A_j)$.

Now, consider A_{j+2} (A_2 if $A_{j+1} = A_n = A_1$). Since, by assumption, it is an oriented cycle, it must be either that $A_{j+1} R A_{j+2}$ or $A_{j+1} Q A_{j+2}$. In the first case, the argument in the previous paragraph shows that $c(A_{j+2}) \succ_p c(A_{j+1})$. In the second case, it must

[9] Interestingly, Johnson and Dean (2001; see also Koshevoy 1999) study lattice representations of choice sets generated by path independent choice correspondences. Their approach and ours share the view of ordering choice sets, but have different aims. Ours is to help in the finding of a minimal set of rationales that rationalizes choice behavior. It is unclear how their lattice representations may be informative in this task.

obviously hold that $c(A_{j+2}) \sim_p c(A_{j+1})$. In both cases, by the transitivity of $\succcurlyeq_p$ it must be that $c(A_{j+2}) \succ_p c(A_j)$. Applying this reasoning iteratively we conclude that $c(A_j) \succ_p c(A_j)$, which is of course absurd. This shows that there is no oriented cycle.

Now take an NOP $\{V_p\}_{p=1,\ldots,P}$. For every class V_p we construct a rationale $\succcurlyeq_p$ that rationalizes all the elements in the class V_p. We start by defining an equivalence relation E between the elements of V_p: AEB if and only if there is a chain of elements in V_p such that $A = A_1 Q A_2 Q \cdots Q A_q = B$. The equivalence relation E partitions V_p into disjoint classes $\{L_p^i\}_{i=1}^{I_p}$. We write $L_p^i R L_p^j$ when there are $A \in L_p^i$ and $B \in L_p^j$ with ARB.

We start by noting that there must be at least one class, denoted by L_p^1, such that $L_p^1 R L_p^i$ holds for no L_p^i, $i > 1$. Otherwise, due to the finiteness of X and hence to the number of classes L_p^i in V_p, there would be a cycle. Using the same argument there must be another class L_p^2 such that $L_p^2 R L_p^i$ holds for no L_p^i, $i > 2$. Clearly, this argument extends to all classes $\{L_p^i\}_{i=1}^{I_p}$. We write $c(L_p^i) = \cup_{A \in L_p^i} c(A)$ and $c(V_p) = \cup_{A \in V_p} c(A)$. Now, we show that the complete preorder $\succcurlyeq_p$ with asymmetric part $\succ_p$, where

$$c(L_p^1) \succ_p c(L_p^2) \succ_p \cdots \succ_p c(L_p^{I_p}) \succ_p X \setminus c(V_p),$$

rationalizes V_p. Assume the contrary is the case. Let $j \in \{1, \ldots, I_p\}$ and $A \in L_p^j$ with $y \in H(A)$ but $y \succcurlyeq_p c(A)$. By construction of $\succcurlyeq_p$, $y \in c(B)$ for some $B \in L_p^k$, $k \le j$. From $y \in H(A) \cap c(B)$, obviously $c(B) \cap A \notin \{\varnothing, c(A) \cap B\}$ and, hence, BRA. By construction of the equivalence classes, it can only be $k = j$. But, since A and B belong to the same equivalence class there is a chain of elements $A = A_1 Q A_2 Q \ldots Q A_q = B$ and, therefore, together with BRA, it implies that there is an oriented cycle. This is absurd and hence $\succcurlyeq_p$ rationalizes V_p.

Finally, to end the claims of the theorem we have to show the conditions on minimality. Then, for the first claim suppose that $\{\succcurlyeq_p\}_{p=1,\ldots,P}$ is minimal while $\{V_p\}_{p=1,\ldots,P}$ is not. Then there is an NOP with T number of classes, $T < P$. But then, we have shown that there is an RMR with T classes, contradicting the minimality of $\{\succcurlyeq_p\}_{p=1,\ldots,P}$. The reverse direction of the proof is analogous. This concludes the proof. □

The significance of Theorem 7 is best appreciated in the case of single-valued choice correspondences. Recall that when c is single-valued, ARB if and only if $c(A) \subseteq H(B)$, $A, B \in M_c^{\mathcal{D}}$. That is, A blocks B if and only if the chosen element in A belongs to B and, at the same time this element is not chosen in B. It is clear that such a case is inconsistent with a linear order rationalizing B and writing $c(A) \succ c(B)$. Also, note that AQB if and only if $c(A) = c(B)$. Then the relation between minimal RMRs and minimal NOPs simplifies considerably. This is because for any three sets $A_1, A_2, B \in M_c^{\mathcal{D}}$, whenever $c(A_1) = c(A_2)$, then $A_1 RB$ if and only if $A_2 RB$. This implies that, whenever there is an oriented cycle, there is a cycle composed of elements related only through R. Hence the analysis can obviate the equivalence relation Q, and focus on $(M_c^{\mathcal{D}}, R)$. The latter is simply a standard directed graph, and the structure defined as an oriented cycle reduces to the standard notion of a directed cycle. There is an immense literature on graphs without directed cycles, typically known as directed acyclic graphs (DAGs).[10] This is especially important since Theorem 7 guarantees that there is much to gain from the results in this literature. Hence, our problem reduces to find a minimal partition into DAGs. In passing we not that this discussion shows that Theorem 7 represents another instance where the single-valued or multi-valued nature of the choice correspondence makes a difference.

4.2. An example

The study of concrete choice procedures appears particularly appealing. It is likely that the structure inherent to specific choice procedures allows for the tractability of rationalization. Here we provide an example that draws on the previous subsection.

Manzini and Mariotti (2007) study the nature of single-valued choice correspondences defined on $\mathcal{U}$ that can be rationalized by sequentially applying a set of asymmetric binary relations in a fixed order. Among other results, they provide a full characterization of the case when a single-valued choice correspondence is sequentially rationalized by two asymmetric binary relations. Such a correspondence is called a Rational Shortlist Method (RSM). The collection of RSMs encompasses the collection of single-valued choice correspondences that satisfy WARP, and it is a strict subset of the collection of all possible single-valued choice correspondences. In fact, MM's characterization makes use of the classical property of expansion.

Expansion: If $\{x\} = c(A)$ and $\{x\} = c(B)$, $A, B \in \mathcal{U}$, then $\{x\} = c(A \cup B)$.

It is not difficult to observe that the set $M_c^{\mathcal{U}}$ shrinks considerably whenever this property holds. Clearly, for each element x in X there is at most one element in $M_c^{\mathcal{U}}$ for which x is the chosen element, namely, the union of all subsets A for which x is chosen. Denote the latter by $M(x)$. Then the problem of finding a minimal RMR here reduces to finding a minimal partition into DAGs over the universal set of alternatives X according to N, where for every $x \neq y$, xNy if and only if $M(x)RM(y)$ if and only if $x \in M(y)$.

Hence the rationalization of RSMs through multiple rationales turns out to be a computationally tractable problem.

[10] See, e.g., Cormen et al. (2001).

5. Final remarks

We have used the tools of theoretical computer science to study the complexity of finding the rationales that rationalize choice behavior. The question of rationalizability of choice behavior has played a central role in economics. However, surprisingly enough, virtually no attention has been given to the practical problem of computing the rationales.

We have shown that, in the classical case, when the weak axiom of revealed preference holds, finding the rationale rationalizing choice behavior is easy. There are polynomial time algorithms that compute the rationale quickly. On the other hand, when we neither impose any restriction on choice behavior nor on the domain, we have shown that the problem of rationalization is NP-complete. Therefore, there is little hope of finding an efficient algorithm bounded above by a a polynomial function. Furthermore, we have shown that in the case of single-valued choice correspondences, it is the conjunction of unstructured choice and unrestricted domain that drives the intractability result. Under the universal domain, the problem of finding a minimal collection of rationales is quasi-polynomially bounded. On the other hand, we argue that in the choice correspondences case, it may well be the case that the difficulty in finding a minimal collection of rationales is triggered by choice behavior per se.

We then turned to trying to better understand the complexity of rationalization. To this end we identified two binary relations over choice sets that capture part of the essence of rationalization. Furthermore, these binary relations define a problem in graph theory that is equivalent to the problem of rationalization. This is particularly interesting since the complexity issues have attracted a great deal of attention in graph theory. The equivalence result provided allows for the searching of existing algorithms in graph theory that can be used for the problem of rationalization.

Apart from the literature on rationalization, our results relate to two other strands in the literature. First, the problem of rationalization can be read as the problem of transmission of information (choice behavior in different situations), given a specific grammar (complete preorders a la KRS). Under this interpretation, our paper is related to the economics literature on language (see Rubinstein, 2000). Rubinstein stresses the importance of binary relations to natural language. In this sense, our results establish that there are practical limitations to the design of a grammar with the ability to transmit any kind of information. Several questions arise. Does the structure of natural speech imply the existence of a collection of complete preorders computable in polynomial time? What, if anything, is lost in the transmission of information, if the problem of constructing a grammar is upper bounded by a polynomial time algorithm?

Second, an immediate conclusion from our results is that the more structured behavior is, the easier it is to rationalize it in practice. That is, rationality makes things easier. This type of observation has been recognized in the bounded rationality literature from different perspectives. Tversky and Simonson (1993) note that the standard maximization problem is hard to beat in terms of its simplicity of formulation. It is most likely that any descriptive bounded rationality model is condemned to involve a more cumbersome formulation. Also, as we mention in the introduction, Campbell (1978a,b); Salant (2003), and Johnson (2006) reach analogous conclusions.

References

Apesteguia, J., Ballester, M.A., 2009. Choice by Sequential Procedures. Barcelona GSE, mimeo.
Aragones, E., Gilboa, I., Postlewaite, A., Schmeidler, D., 2005. Fact-Free Learning. American Economic Review 95, 1355–1368.
Ballester, C., 2004. NP-completeness in Hedonic Games. Games and Economic Behavior 49, 1–30.
Campbell, D.E., 1978a. Realization of Choice Functions. Econometrica 46, 171–180.
Campbell, D.E., 1978b. Rationality from a Computational Standpoint. Theory and Decision 9, 255–266.
Cormen, T.H., Leiserson, C.E., Rivest, R.L., Stein, C., 2001. Introduction to Algorithms. MIT Press.
Futia, C., 1977. The Complexity of Economic Decision Rules. Journal of Mathematical Economics 4, 289–299.
Garey, M.R., Johnson, D.S., 1979. Computers and Intractability: A Guide to the Theory of NP-Completeness. Freeman.
Johnson, M.R., 1990. Information, Associativity and Choice Requirements. Journal of Economic Theory 52, 440–452.
Johnson, M.R., 1995. Ideal Structure of Path Independent Choice Functions. Journal of Economic Theory 65, 468–504.
Johnson, M.R., 2006. Economic Choice Semiutomata, Structure, Complexities and Aggregations. Econometric Society World Congress, London.
Johnson, M.R., Dean, R.A., 2001. Locally Complete Path Independent Choice Functions and Their Lattices. Mathematical Social Sciences 42, 53–87.
Kalai, G., Rubinstein, A., Spiegler, R., 2002. Rationalizing Choice Functions by Multiple Rationales. Econometrica 70, 2481–2488.
Karp, R.M., 1972. Reducibility among Combinatorial Problems. In: Miller, R.E., Thatcher, J.W. (Eds.), Complexity of Computer Computations. Plenum Press, New York, pp. 85–103.
Koshevoy, G.A., 1999. Choice Functions and Abstract Convex Geometries. Mathematical Social Sciences 38, 35–44.
Mandler, M., 2009. Rational Agents are the Quickest. Royal Holloway College, University of London, mimeo.
Mandler, M., Manzini, P., Mariotti, M., 2009. A Million Answers to Twenty Questions: Choosing by Checklist. Royal Holloway College, University of London and University of St Andrews, mimeo.
Manzini, P., Mariotti, M., 2007. Sequentially Rationalizable Choice. American Economic Review 97, 1824–1839.
Rubinstein, A., 2000. Economics and Language. Cambridge University Press.
Salant, Y. (2003). Limited Computational Resources Favor Rationality, Discussion paper No. 320, The Center for the Study of Rationality, Hebrew University.
Salant, Y. (2010), Procedural Analysis of Choice Rules with Applications to Bounded Rationality, American Economic Review, forthcoming.
Tversky, A., Simonson, I., 1993. Context-dependent Preferences. Management Science 39, 1179–1189.