# Authorization Decisions*

Itzhak Gilboa† and David Schmeidler‡

March 2012

**Abstract**

An *authorization decision* is a binary decision taken by an institution, determining whether certain economic transactions are allowed to take place. They may involve granting a status to an individual, approving a new product, and so forth. Institutions seek to be consistent with their past decisions, as well as with their regulations. Consistency with past decisions is axiomatized: it is shown that certain coherence notions between the decisions made across different histories imply that the institution would seek to be consistent with its own decisions within a given history. In this context, regulations are modeled as constraints on the institution's decisions, and the complexity of finding regulations that enforce specific decisions is studied.

# 1 Introduction

## 1.1 Motivation

An *authorization decision* is a binary decision taken by an institution, determining whether certain economic transactions are allowed to take place. Often the problem is whether to grant the agent a status that permits the

*We thank Antoine Billot, Eddie Dekel, Orna Harari, Elhanan Helpman, Ron Siegel, and Yoram Weiss for conversations and references. ISF Grant 396/10 (Gilboa and Schmeidler) and ERC Grant 269754 (Gilboa) are gratefully acknowledged.

†Tel-Aviv University, and HEC, Paris. tzachigilboa@gmail.com

‡Interdisciplinary Center (IDC), Herzliya, Tel-Aviv University, and the Ohio State University. schmeid@tau.ac.il

agent's participation in a market. For example, a professional guild has to decide whether to license a new member, thereby determining whether to allow the person to practice the profession, such as law or medicine; the immigration services have to determine whether to permit a foreigner to naturalize, and thus be active in the economy; and so forth.

Another type of authorization decisions involve the existence of the market itself. For instance, the government has to decide whether to approve a new medication, or whether to legalize a drug. In these cases it is not the agent who is the subject of authorization, but the product.

There are also situations where the product is not standardized enough to have a market, and specific transactions need to be approved. For example, when the government sells state property, when a municipality approves a construction project, or when an equal opportunity employer hires a professional, specific transactions need to be approved, and are, indeed, often challenged in court.

Finally, there are authorization decisions that involve granting property rights to agents, such as a tenure decision at a university. Such decisions determine whether an agent can have a certain status, where the status already implies certain economic transactions, as opposed to merely allowing participation in a market.

The institutions that make authorization decisions are typically not-for-profit organizations, such as government agencies at various levels, professional guilds, courts of law, and so forth. But they may also be firms making internal decisions such as whether to approve a request for a leave of absence. For simplicity, we focus on binary decisions, though there are cases where economic activity can be authorized at different levels.

How do institutions make authorization decisions? The rational choice paradigm offers two main approaches to modeling such decisions. First, one can conceive of an institution as a single decision maker. Under the standard assumptions of rationality, the institution can be thought of as if it were maximizing a utility function, or even the expectation of such a function relative to a certain probability (see Debreu, 1959, von Neumann and Morgenstern, 1944, Savage, 1954). Second, one may delve into the structure of the institution, and analyze the sub-units or even the individuals who function within it. Applying rational choice principles to each of these, and recognizing that they have different goals and potentially different beliefs, one may study the behavior of the institution as an equilibrium in a game whose players are these agents.

Both views of the institution, as a utility maximizing entity and as an equilibrium among utility maximizing agents, are insightful and useful. There are many problems for which these views suffice. But these two views also have their drawbacks, as the following example illustrates.

Consider a decision by the immigration services, whether to grant citizenship to a potential immigrant. The immigration services are a branch of the state. Viewing the state as a monolithic decision maker, one might assume that its utility function involves the well-being of its citizens, and perhaps the promotion of humanistic values. These goals tend to be rather vague, and the trade-offs between them may not be clearly sorted out. Worse still, these goals are only remotely related to the decision at hand. Consequently, if we consider the totality of the state's decisions, it is not obvious that they are satisfactorily captured by a complete and transitive preference relation that can be represented by utility maximization.

Alternatively, one may decompose the state into many sub-institutions and split these further to the individuals who constitute them. It is almost tautological that what the state does is a result of the decisions made by these individuals, starting from the head of the state issuing instructions and culminating in a clerk stamping a form. Indeed, the standard approach in economic theory would be to assume that each of these individual agents has a well-defined utility function and a subjective probability measure, and that they maximize their expected utility relative to their probability. But these assumptions do not suffice to explain and predict the institution's behavior. The agents' goals and beliefs are not directly observable, and even if they were, they would often fail to pin down a unique equilibrium outcome.

Clearly, similar problems arise when we consider a university's tenure decision, or the licensing of an expert by a professional guild. The rational choice paradigm is undoubtedly insightful, and at some level almost tautologically true. But it does not always seem to provide a useful account of the behavior of an institution. Hence one may seek other conceptualizations that will enrich our image of institutions, and perhaps deepen our understanding of the way that institutions function and interact in economic environments.

## 1.2 Background

Organization theory offer a variety of ways of conceptualizing organizations in general, and institutions in particular. The view of an organization as a rational entity is perhaps most natural in the analysis of economic firms, with an emphasis on efficient production and the division of labor (Smith, 1776, Marx, 1867, and Durkheim, 1893). The view of the organization as a well-tuned machine became more salient with the rise of mass production in the early 20th century (Taylor, 1911, Follett, 1918, Fayol, 1919).

Weber (1921, 1924) studied authority in general, and argued for the superiority of bureaucracy as a way of establishing legitimate authority based on rational grounds, and of achieving the maximal efficiency. While viewing bureaucracy is a much more positive light than is common today, Weber was a pioneer in decomposing organizations to their constituent agents.

In very bold strokes, these early studies in sociology and economics correspond to the two rational-choice conceptualizations mentioned above: one views the organization as a rational agent, and the other – as a game among rational agents (to use modern language). However, many other views have also been suggested in sociology and related fields. March and Simon (1958) highlighted the role of organizations as decision-making entities, but they pointed out the bounded rationality that characterizes organizational decision making. In particular, they viewed organizations as satisficing, rather than optimizing entities. Burns and Stalker (1961) suggested that mechanistic bureaucracies are ill-adapted to deal with changing environments, and that organizations are often adapting as organic systems.[1] In this sense they went a step further, as their model of an organization did not seem to have a well-defined objective function, apart from, perhaps, the organism's objective to survive. Moreover, Kanter (1977) argued that power inside an organization may not be easy to define, and suggested that the seemingly powerful are often powerless.[2]

Morgan (2006) identifies different images, or metaphors, that have been used to describe organizations. Among them are machines (Taylor, 1911, Fayol, 1919, Weber, 1924), organisms (Parsons, 1951, Burns and Stalker, 1961), brains (Sandelands and Stablein, 1987, Walsh and Ungson, 1991, March, 1999), cultures (Ouchi and Wilkins, 1985), and political systems (Burns, 1961, March, 1962).

Within economics, much attention has been devoted to the firm, its boundaries and its internal operations. Coase (1937) pointed out the absence of an economic theory of the size of the firm. Williamson (1975, 1979, 1981) discussed the transaction costs between and within organizations, with implications for vertical integration and the boundaries of the firm. In this analysis, Williamson rejected the view of an economic firm as a monolithic entity and suggested to analyze it as an arena in which different agents operate. Jensen and Meckling (1976) studied modern corporation from a principal-agent point of view, highlighting the difficulties that are generated by the separation of ownership from control.

[1] Both "organization" and "organism" derive from the Greek "organon", which means a tool, or an instrument. The adjective "organicon" was used in the context of live bodies.

[2] For extensive introductions to organization theory, see Handel (2003), Hatch-Cunliffe (2006), and Scott and Davis (2007).

Economic models have also been used to study organizations more generally. In particular, Buchanan and Tullock (1962) viewed the state as comprising of rational agents with different goals. Niskanen (1971, 1975) analyzed bureaucracy as a production entity, and questioned its efficiency. Studies such as Bendor and Moe (1985) analyzed bureaucracies, while relaxing the rationality assumptions.

To sum, there is a wide variety of models and metaphors used to understand organizations. Of these, but a fraction have been formally modeled in a way that would allow their incorporation in economic theory. Most formal models that have been proposed in the literature are within the rational choice paradigm, defining their unit of analysis to be the organization as a whole or its constituents. Exceptions, such as March and Simon (1958) or Bendor and Moe (1985), relaxed rationality to bounded rationality, viewed as satisficing rather than optimizing behavior.

The goal of this paper is to suggest a formal model of an institution's authorization decisions that will highlight other aspects. Our focus is on the institution's quest for consistency, in the absence of any other goal that can be attributed to the institution as a whole. Thus, we focus on an image of an institution without an objective function, but we seek to retain the formal framework that may allow one to combine the analysis of institutions with the theoretical analysis of economic agents.

## 1.3 Outline

Our starting point is that an institution, making authorization decisions, seeks to be *consistent*, both with its own past decisions, as well as with its rules and regulations.

There are several reasons for which consistency might be the institution's apparent goal. The first one, related to inertia, is that consistency with past choices is the path of least resistance. Doing the same thing that the institution has done in the past is a simple choice: everyone within the institution already knows what they need to do, and no additional energy needs to be expended to change the standard practice. Second, consistency of choice endows the institution with a sense of identity. Having a well-defined policy defines what the institution is. Indeed, employees might proudly say "we always do..." or "we would never do such a thing" suggesting that consistency is viewed as a merit. Yet another reason to seek consistency has to do with a sense of fairness. For example, it would feel wrong to admit a candidate to the university's graduate program while denying admittance from another candidate who has the same qualifications. Indeed, such discrimination might result in a lawsuit. But irrespective of legal complications, it appears to run against our natural notion of justice, of "equal treatment of equals".

We mention in passing that the legal system, especially when common law is concerned, exhibits certain features that remind one of an institution seeking consistency. However, the legal system has specific features not modeled here, and it is not the primary application of the model.

How does an institution trade off its past decisions in determining what decision, in a given problem, would be the "most consistent" with past decisions? Since there are many possible notions of consistency, we adopt an axiomatic approach to this problem. For clarity of exposition, we begin, in Section 2, by presenting an additive model of consistency with past cases: there exist relevance functions, such that the decision at each problem, given history, can be described as the decision that maximizes the sum of relevance values to past cases in which the same decision has been made.

We then proceed, in Section 3 to derive this additive decision rule. For the axiomatic derivation, we consider the institution's decisions not only in a given history of cases, but also given other, hypothetical histories. We assume that these decisions obey certain coherence axioms. The central of these is a so-called combination axiom, stating that a decision that would have been made given two disjoint histories would also be made given their union. As a result, we show that the institution can be described *as if* it were seeking consistency with past cases using the additive rule.

Importantly, our axioms do not pre-suppose that there are relevance functions or that new decision problems are analyzed in light of the relevance of past cases. The role played by past cases, the values of the relevance functions, and the additive nature of aggregation over past cases are all derived from the axioms. Thus, coherence of choices across different, hypothetical histories implies that, given a specific history, the institution would seek consistent choice within it.

Our axioms on the institution's coherence (across different histories) can be viewed as axioms on a "black box" that makes decisions. Indeed, these axioms are compatible with many decision models.[3] The representation of choice that results can be interpreted in two ways. First, it may be viewed as a mere representation, along the "as if" interpretation of utility maximization in consumer choice theory. According to this interpretation, institutions might be described as if they sought consistency with their own decisions in past cases, even if their decision making process makes no explicit reference to consistency.

Second, the representation can also be viewed as a description of an actual decision process. Again in analogy to consumer choice, where utility maximization can also be interpreted as an actual mental process the consumer goes through, in our case the quest for consistency with past choices can be viewed as part of the decision making process in the institution. Viewed thus, our theorem might provide another reason for the preference for consistency: consistency with past cases guarantees – and is, in fact, the only way to guarantee – coherence of choice across different hypothetical histories.

Institutions do not only seek consistency with past cases. Rather, they have rules such as implicit norms and explicit regulations, which partly gov-

[3] In particular, the relevance functions that are derived in our model can be identically zero, suggesting that the institution makes decisions in a history-independent way.

ern the institution's behavior. How do such rules interact with past cases? How do regulations affect the institution's decisions, and how are they introduced? We devote Section 4 to these questions. We first propose (in sub-section 4.1) a simple model in which explicit regulations constrain the institution's decision, while consistency with past cases guides decisions if existing regulations do not determine it. We consider a specific model of regulations that are conjunctions of conditions stated in a simple language, and show that it is a computationally simple problem to determine whether a set of regulations is consistent, and whether it determines the decision in a given problem. However, we show that if an agent wishes to find the most general regulation that would enforce a particular decision in a given problem, they face a computationally hard problem and may therefore not be able to find an appropriate regulation. As a result, when agents promote regulations with specific decisions in mind, the set of regulations may become inconsistent. We also discuss other ways of modeling regulations, which allow regulations to have an effect even if the set of regulations is inconsistent.

We then proceed, in sub-section 4.3 to describe an alternative approach to modeling regulations, according to which a regulation is not a clear-cut constraint on the institution's decision, but only a consideration favoring one decision over another. In this model there is a trade-off among past cases and existing regulations. If the weight of a regulation is high enough, it can overwhelm history and act as a constraint, dictating the decision. However, the regulation may also be weaker than common practices. Further, the weight of a regulation may be endogenous, so that the degree to which it has been followed in the past changes the respects it commands in the present. This model can explain why certain regulations are being enforced while others are ignored and become irrelevant.

Finally, Section 5 concludes.

# 2 Model

## 2.1 Set-up

An institution is viewed as an entity that processes cases. Each *problem* $p \in P$ presents the institution with a binary *decision* $d \in \{0,1\}$, such as whether to accept a candidate to a school. A problem and a decision made in it jointly generate a *case* $c = (p, d)$. The set of all cases is $C = P \times \{0,1\}$.

The institution has a *history* of cases, which is a finite subset $H \subset C$, in which every problem $q$ may appear at most once, that is, $c = (q, d)\,, c' = (q', d') \in H$ imply $q \neq q'$. The set of all histories is denoted $\mathcal{H}$. If new problems are considered identical to past ones, we will model them as formally distinct elements of $P$ though they will be equivalent in terms of their impact on future decisions.

We denote by $H_P$ the projection of $H$ on the first coordinate, that is

$$H_P = \{\, p \in P \mid \exists d \in \{0,1\}, (p, d) \in H \,\}\,.$$

Given a history $H$ and a new problem, $p \notin H_P$, the institution attempts to make a decision $d \in \{0,1\}$. Assume that the institution has a *decision correspondence*

$$f : \{\, (H, p) \mid p \in P, H \in \mathcal{H}, p \notin H_P \,\} \twoheadrightarrow \{0,1\}$$

that is never empty-valued. Thus, if $f(H, p) = \{d\}$ the institution's decision is $d \in \{0,1\}$, and if $.f(H, p) = \{0,1\}$ the institution may make either choice. Observe that the correspondence $f$ is defined given each and every history, even those that are incompatible with the correspondence $f$ itself. This is in line with the standard definition of a strategy in a game, and allows us to deal with random choices, errors, hypothetical choices, and so forth.

## 2.2 Consistency: Similarity and Relevance

We first model the institution decisions as seeking consistency with past decisions. The benchmark model of consistency consists of evaluation of the similarity of the problem at hand to past problems, and making the decision that has been chosen most often in similar cases in the past. To be precise, assume that there is a function

$$s : P \times P \to \mathbb{R}_+$$

that measures the degree of similarity of one problem to another. Consistency with similar past cases is modeled as follows. For history $H$ and problem $p$, and for $d \in \{0,1\}$, let

$$S_s(H,p,d) = \sum_{c=(q,d)\in H} s(q,p)$$

thus, $S_s(H,p,1)$ is the total similarity weight of all past problems, in which the decision was 1, to the present problem, and $S_s(H,p,0)$ is the corresponding sum for the decision 0.

Next define $f_s(H,p)$ to be the arg max of $S_s(H,p,\cdot)$, or, explicitly,

$$f_s(H,p) = \begin{cases} \{1\} & S_s(H,p,1) > S_s(H,p,0) \\ \{0,1\} & S_s(H,p,1) = S_s(H,p,0) \\ \{0\} & S_s(H,p,1) < S_s(H,p,0) \end{cases} .$$

We interpret $f_s$ as the decision most consistent with history, according to the similarity function $s$.

Often, the impact of a decision in a past case is asymmetric: it may happen that a decision 0 is the past is less relevant than a decision 1. For example, assume again that the institution is a university that makes admissions decisions based on GPA scores alone. Assume that a past case $(q,d)$ featured a candidate with a GPA score of 3.5, where the present candidate $p$ has a GPA score of 3.7. If the past decision was admittance ($d = 1$), it makes a strong case for admitting the present candidate, who has a higher GPA score. If, however, the candidate $q$ ended up in a negative decision ($d = 0$), this decision has lesser relevance to the case at hand than it would have had $d$ been 1. Indeed, the university may choose a cutoff that is between

3.5 and 3.7, and admit $p$ without being inconsistent with its past decision not to admit $q$.

This phenomenon calls for a slightly more general structure. Rather than using a similarity function $s(q,p)$ that is independent of the decision in the past case $q$, we use two *relevance functions*,

$$w_0, w_1 : P \times P \to \mathbb{R}_+$$

such that $w_d(q,p)$ measures the relevance of the decision $d$ in the past case $q$ to making the same decision in the present case $p$. As in the case of decision-independent similarity, we aggregate the relevance values across history: for history $H$ and problem $p$, and for $d \in \{0,1\}$, let

$$W_{w_0,w_1}(H,p,d) = \sum_{c=(q,d)\in H} w_d(q,p). \tag{1}$$

(Note that $W_{w_0,w_1}(H,p,1)$ does not depend on $w_0$, and similarly $W_{w_0,w_1}(H,p,0)$ does not depend on $w_1$.) Given the relevance functions $w_0, w_1$, we define the decision correspondence

$$f_{w_0,w_1}(H,p) = \arg\max_{d\in\{0,1\}} W_{w_0,w_1}(H,p,d)$$

and assume that such a correspondence guides the institution's decisions.

# 3 Axioms and Representation

In this section we show that the decision correspondence postulated above follows from rather simple assumptions about coherence of the institution's decisions given different histories. We begin by a richness assumption that is needed for the axiomatic derivation.

In the sequel, the notation $f(H,p)$ will be understood to assume that $p \notin H_P$.

Let there be given a decision correspondence $f$. Two problems $q, q' \in P$ are *$f$-equivalent*, denoted $q \sim_f q'$, if, for every history $H$ and every $p \in P$ such that $q, q' \notin H_P$, and for every $d \in \{0,1\}$,

$$f\left(H \cup \{(q,d)\}, p\right) = f\left(H \cup \{(q',d)\}, p\right).$$

In this section we assume that $P$ is infinite, and we will consider only decision correspondences $f$ for which the following holds:

**Richness Assumption:** For every $q \in P$ the set $\{\, q' \in P \mid q \sim_f q' \,\}$ is infinite.

It will also be convenient to define a corresponding equivalence relation on cases and on histories: two cases, $c = (q,d), c' = (q',d') \in C$, are *$f$-equivalent*, denoted $c \sim_f c'$, if $q \sim_f q'$ and $d = d'$. Next, two histories, $H, H' \in \mathcal{H}$, are *$f$-equivalent*, denoted $H \approx_f H'$, if there exists a bijection between them, $\pi : H \to H'$, such that, for every $c \in H$, $c \sim_f \pi(c)$. In other words, two histories are equivalent if they consist of the pairwise equivalent cases.

We now state our axioms.

**Axiom 1 (Combination):** Assume that $H, H' \in \mathcal{H}$ and $p \in P$ are such that $H_P \cap H'_P = \varnothing$. If $f\left(H,p\right) \cap f\left(H',p\right) \neq \varnothing$, then $f\left(H \cup H', p\right) = f\left(H,p\right) \cap f\left(H',p\right)$.

The combination axiom basically states that, if a certain decision is justifiable given each of two histories, it will also be justifiable given their union. To make this union a meaningful concept, the histories are assumed to be compatible in the sense that no problem appears in both of them ($H_P \cap H'_P = \varnothing$), and it is also implicitly assumed that the problem under consideration does not appear in either of the two. To illustrate this axiom, consider the admission problem: if, for instance, the admission decision $d = 1$ is acceptable given each of two records of past decisions (for a given new problem $p$), it should also be acceptable given their union. Further, the combination axiom states that if, given $H$, both decisions are possible $(f(H,p) = \{0,1\})$ but, given $H'$, only one is (say, $f(H,p) = \{1\}$), then the latter is the only decision possible given the union of the two histories. This reflects the notion that indifference between the two decisions is not robust: if one history fails to provide reasons for making a decision, but the other does, the institution would make the decision dictated by the latter. Observe that if the two histories suggest different decisions (say, $f(H,p) = \{1\}$ but $f(H',p) = \{0\}$), the combination axiom does not restrict the decision given the union of the two histories.

The combination axiom appears to be rather reasonable in a variety of situations. If we take a "black box" interpretation of the decision correspondence, an outside observer who notices that the institution made a certain decision in two histories may well risk a guess that the same decision will be made given the union of these histories. If we consider the actual decision making process of the institution, the axiom can be part of the reasoning that supports the decision given the longer history.

Clearly, there are reasonable counter-examples to the axiom. Consider, for example, an academic department that makes offers to candidates, and seeks to retain a healthy balance between fields. It is possible that such a department would make an offer to a candidate if the department already has only one similar researcher, but not if it has two such researchers. This would be a violation of the combination axiom. Generally, when there are goals having to do with collections, or patterns of choices, the combination axiom may be unreasonable.

Another class of counter-examples to the combination axiom involves an institution's reaction to a trend. Consider, for example, the decision to grant an immigrant status to a foreigner. It is possible that such a status would be granted to an individual, but will not be granted to millions of identical individuals. This might be the case if the state is concerned about masses of immigrations, and decides that the trend should be stopped. This phenomenon will be discussed in the next section, where we introduce regulations into the model: regulations are often introduced in order to change such trends. In this section, however, we are only interested in the trends themselves, in which one case follows another, and for that purpose the axiom appears to be a reasonable first approximation.[4]

For the next axiom we need an additional definition. Given the notion of equivalent histories, we can define the notion of a replica of a history: history $H'$ is a $k \geq 1$ $f$*-replica* of history $H$ if $H'$ is the disjoint union of $k$ histories $H_1, ..., H_k$, each of which is $f$-equivalent to $H$. We denote this relation by $H' \approx_f kH$.

**Axiom 2 (Archimedeanity):** Assume that $H \in \mathcal{H}$, $p \in P$, and $d \in \{0, 1\}$ are such that $f(H, p) = \{d\}$. For every $H' \in \mathcal{H}$ there exists a $k \geq 1$ and $H'' \in \mathcal{H}$ such that $H'' \approx_f kH$ and $f(H' \cup H'', p) = \{d\}$.

The Archimedean axiom states that, if history $H$, considered in isolation, gives reason to favor a decision $d$, then sufficiently long repetition of this history would overwhelm any other given history $H'$. The axiom means that the reasons for or against certain decisions, as provided by histories, are comparable: suppose that $f(H, p) = \{d\}$ but $f(H', p) = \{1 - d\}$. (Observe that if $d \in f(H', p)$, the conclusion would already follow from Axiom 1. Hence the interesting case is that the two histories induce different decisions.) Intuitively, Axiom 2 states that the reasons provided by $H'$ to choose $1 - d$ cannot be infinitely more powerful than the reasons provided by $H$ to choose $d$. The axiom has a flavor of continuity, and it is very useful in simplifying the representation. However, one may choose to drop it and obtain conceptually similar, though mathematically more cumbersome representations (using vector-valued functions applied lexicographically, or non-standard numbers).

**Axiom 3 (Monotonicity):** Assume that $H \in \mathcal{H}$, $q, p \in P \backslash H_P$, and

[4]Similar axioms appeared in Young (1975) and Myerson (1995), in the context of social choice. In their models, the union was taken over sets of ballots cast in elections, and the intersection was applied to the sets of acceptable candidates.

$d \in \{0, 1\}$ are such that $d \in f(H, p)$. Then $d \in f(H \cup \{(q, d)\}, p)$.

The monotonicity axiom implies that if $d$ is an acceptable decision given history $H$, where $q$ has not been encountered, then $d$ should certainly be acceptable if we were to add to $H$ a case in which $d$ was chosen at $q$. This axiom is the closest we get to assuming consistency with past decisions, and it is indeed the only axiom that makes reference to specific decisions made in the past. The axiom says that having made the decision $d$ more in the past cannot make it less acceptable in the present.

To see an example where this axiom is violated, consider an academic journal. Assume that the journal accepts a paper for publication. Next assume that an identical paper is submitted again to the same journal. The monotonicity axiom implies that the new paper be accepted, which is obviously a counter-intuitive prediction. One may point out, however, that the new paper is not truly identical to the first: an important feature of a research paper is its novelty, and this attribute distinguishes the second paper from the first. In other words, the two cases appear to be identical only because we left out an important attribute of the description of the case. Alternatively, one may hold that originality and novelty are among the features that can only be defined for sets of cases, and that preferences for such set-defined concepts may indeed violate both the combination and the monotonicity axioms.

Another violation of monotonicity may be exhibited by the immigration example discussed above in the context of the combination axiom: if there is a critical mass of immigrants at which the state changes its immigration policy, there will be a history $H$, at which the last immigrant is still let in, but, after that decision, no new immigrants are allowed. Thus, for some $H$, the decision $d$ is acceptable for $H$ but not for $H \cup \{(q, d)\}$. However, as discussed above, in this section we ignore the regulatory activity that is designed to control trends, and focus on the trends themselves. In this context, the monotonicity axiom appears more reasonable.

We can now state a theorem that is an (almost) immediate implication of the main theorem in Gilboa and Schmeidler (2003):

**Theorem 1** *A decision correspondence* $f$ *satisfies Axioms 1-3 if and only if there are relevance functions* $w_0, w_1 : P \times P \to \mathbb{R}_+$ *such that* $f = f_{w_0,w_1}$.

The meaning of the theorem is that institutions that can be described, to an outside observer, as satisfying the coherence requirements posed by the axioms can also be thought of as seeking consistency with their own past decisions. It is important to note that the coherence requirements stated in Axioms 1-3 do not directly state that decisions are made so as to be consistent with previous ones. Of the three axioms, only the last one refers to the actual decisions made in past cases. In fact, one can drop Axiom 3 and obtain the same result, in which the relevance functions need not be non-negative. In such a formulation, past cases may be completely abstract pieces of information, for which "consistency with past choices" is meaningless.

The theorem obtains the quest for consistency mostly from the very formulation of the problem, as a choice $f(H, p)$ in a problem $p$ *given* history $H$. Once this formulation of the problem is agreed upon, Axioms 1-2, and, in particular, Axiom 1 (the combination axiom), provide the additive representation.[5] In this context, Axiom 3 is needed to make sure that past decisions make similar future decisions more likely rather than less likely. However, the coherence of choice across different histories, as stated in Axiom 1, is quite different from the consistency of the present choice with previous ones, as implied by the theorem.

[5] As mentioned above, Axiom 2 is needed to guarantee that this additive representation involves standard, rather than non-standard numbers.

# 4 Rules and Regulations

## 4.1 Regulations as Constraints

Institutions have rules and regulations that constrain, or otherwise affect their decisions. Some rules are formally stated regulations, while others are norms that have been established in the institution. To the extent that a norm has indeed been established, it can be represented by the totality of cases in which it has been followed, and it will therefore be captured by the tendency to make decisions consistent with past cases. Thus, we focus here on formally stated regulations, especially when these are at odds with, and designed to change common practices.

A *rule* to be a pair $r = (D, d)$ where $D \subset P$ is the *domain* in which the rule applies, and $d \in \{0, 1\}$ is interpreted as the decision that should be made in this domain (according to the rule). For example, when admitting candidates to a school, a rule might be that all candidates above a certain GPA should be admitted. Observe that the domain $D$ can be restricted by time, and, for example, to apply only to problems that are encountered after a certain date. We denote the set of all rules by $\mathcal{R} = 2^P \times \{0, 1\}$.

We first assume that rules are constraints on the institution's decisions. That is, if a rule applies in a case, it will gain precedence over whatever common practices might exist in the institution.[6] Let there be given a set of rules $R \subset \mathcal{R}$. For $d \in \{0, 1\}$, let $R(d)$ be the union of all the domains of the rules in $R$ that dictate the choice $d$:

$$R(d) = \cup_{(D,d) \in R} D$$

The set of rules $R$ is said to be *consistent* if $R(0) \cap R(1) = \varnothing$, that is, if there does not exist a problem $p$ for which there exists a rule that implies that the decision at $p$ be 0 and another rule that implies that this decisions be 1. Clearly, $R$ is consistent iff for every $(D, 0), (D', 1) \in R$, $D \cap D' = \varnothing$.

[6] Sub-section **??** briefly discusses alternative approaches.

Let there be given relevance functions $w_0, w_1 : P \times P \to \mathbb{R}_+$. We assume that these functions describe people's perception of the relevance of past cases, and are fixed for the time being. By contrast, the set of rules may change as new rules are introduced. Given a history $H \in \mathcal{H}$, a set of rules $R \subset \mathcal{R}$, and a problem $p \in P \backslash H_P$, define

$$f_{w_0,w_1}(H, R, p) = \begin{cases} d & if \quad p \in R(d) \backslash R(1-d) \\ \arg\max_{d \in \{0,1\}} W_{w_0,w_1}(H, p, d) & otherwise \end{cases}$$

Thus, if there is a decision $d \in \{0, 1\}$ and a rule $r = (D, d) \in R$ that dictates that $d$ be taken at $p$ (that is $p \in D$), but there is no other rule that dictates that the opposite decision be taken, then the institution follows the rule and makes the decision $d$. If, however, no rule applies at $p$ (that is, $p \notin R(0), R(1)$) or if there is a contradiction between different rules ($p \in R(0) \cap R(1)$), the rules do not determine the decision at $p$ and the institution follows common practices, as captured by history $H$ interpreted via the relevance functions $w_0, w_1$.

Observe that a set of rules $R$ may be inconsistent, but lead to no contradictions at $p$. Inconsistency means that there exists $p'$ in $R(0) \cap R(1)$, but this $p'$ may differ from $p$. It is possible that a set of rules that is found to be inconsistent (at some $p'$) is unlikely to be respected even in problems that do not lead to inconsistency (such as $p$). In the next sub-section we focus on sets of rules that are consistent, and therefore can be expected to respected.

## 4.2 The Structure of Regulations

For given relevance functions, $w_0, w_1 : P \times P \to \mathbb{R}_+$, a history $H \in \mathcal{H}$, a set of existing rules $R \subset \mathcal{R}$, and a problem $p \in P \backslash H_P$, suppose that $p \notin R(0), R(1)$, so that the existing rules do not constrain the institution's decision. Assume also that $f_{w_0,w_1}(H, R, p)$ suggests that a decision $1-d$ be made. However, an agent (or a group of agents) wishes to make the opposite decision, $d$. This may happen if the decision $1 - d$ is perceived as the "wrong" decision, say, one that is at odds with the institution's stated goals. For example, assume that, due to consistency with past decisions, a university is about to admit a clearly weak candidate. Alternatively, it may be the case that an agent wishes to promote some personal interest, say, to admit a relative of the university's president, who would not be admitted given past decisions. In short, we assume that an agent wishes to introduce a regulation that would constrain a decision $d$ in problem $p$.

Not every rule $r = (D, d)$ can be introduced as a regulation. For example, regulations cannot use proper names, as in "John Smith should be admitted". Moreover, in many institutions regulations may not refer to certain predicates such as gender, race, or even age. Thus, we first need to define the *language of regulations*. Let there be a class of binary *attributes* $a_1, ..., a_m$, where $a_j : P \to \{0, 1\}$ denotes, for each problem $p \in P$, whether or not it has the attribute $j$. We assume that the set of problems is rich enough so that every combination of attribute values is obtained, namely, $range\,((a_1, ..., a_m)) = \{0, 1\}^m$.

A regulation will be a rule whose domain is defined by setting some attributes to take particular values. Formally, let $J \subset \{1, ..., m\}$, $J \neq \varnothing$, $b : J \to \{0, 1\}$ and $d \in \{0, 1\}$. The *regulation* defined by $(J, b, d)$ is the rule $r = (D(J, b), d)$ where

$$D(J, b) = \{\, p \in P \mid a_j(p) = b(j) \quad \forall j \in J \,\}\,.$$

Thus, regulations are restricted to be definable in the language of the legitimate attributes, and, furthermore, they have to be cylindric, that is, the intersection of sets, each of which restricts the value of one attribute only.

Let there be given a set of regulations $R = \{(J_i, b_i, d_i)\}_{i=1}^n$, a problem $p$, and a decision $d$. An agent who wishes to impose the decision $d$ at $p$ may wish to know, first, whether the set of rules is consistent and whether it determines the decision at $p$. If the decision at $p$ is not determined by $R = \{(J_i, b_i, d_i)\}_{i=1}^n$, one may introduce a new regulation that would restrict the choice at $p$ to be the desired $d$. However, to do that one should verify that, if $R$ is consistent to begin with, then it also remains consistent after the addition of the new regulation. It turns out that, if the set of regulations is consistent but does not dictate the choice at $p$, there exists a regulation that can be added to $R$, dictating $d$ at $p$, while retaining consistency. Importantly, one may find such a regulation by an efficient algorithm:

**Proposition 2** *Let there be given a number of attributes $m$, a set of regulations $R = \{(J_i, b_i, d_i)\}_{i=1}^n$, a problem $p$ and a decision $d$. There exists a polynomial-time algorithm that finds out whether $R$ is consistent, whether $p \in R(0), R(1)$, and, if $R$ is consistent and $p \notin R(0), R(1)$, finds a regulation $(J_{n+1}, b_{n+1}, d_{n+1})$, such that $d_{n+1} = d$, $p \in D(J_{n+1}, b_{n+1})$ and $R' = \{(J_i, b_i, d_i)\}_{i=1}^{n+1}$ is consistent.*

Thus, the proposition states that one can find out whether a regulation can be added to a set of regulations, imposing a certain decision in a given problem, without violating consistency. However, this result might be misleading because it says nothing about the generality of the regulation that is used to impose the decision $d$ at $p$. A regulation that uses too many attributes, all of which happen to hold at $p$, would be considered ad hoc or even ad hominem. Thus, there is an interest in regulations that use a small set of attributes $J$. The smaller is the set $J$, the more does the regulation appear to be introduced in good faith, and the simpler it is to state and to enforce. Thus, given a consistent set of regulations and a desired decision in a given problem, one may seek to introduce the most general regulation that imposes the desired decision without rendering the set of regulations inconsistent. However, it turns out that this is a computationally difficult task:[7]

**Theorem 3** *Let there be given a number of attributes $m$, a set of regulations $R = \{(J_i, b_i, d_i)\}_{i=1}^n$, a problem $p$, a decision $d$, and a number $k \geq 1$ such that $R$ is consistent and $p \notin R(0), R(1)$. Finding whether there exists a regulation*

[7] See Appendix B for an informal introduction to a few concept in complexity theory.

$(J_{n+1}, b_{n+1}, d_{n+1})$ *such that* $d_{n+1} = d$, $p \in D(J_{n+1}, b_{n+1})$, $R' = \{(J_i, b_i, d_i)\}_{i=1}^{n+1}$ *is consistent, and* $|J_{n+1}| \leq k$ *is NP-Complete.*

The theorem suggests that, in general, agents will not be able to find the most general regulations that retain consistency, and enforce the decisions they are trying to impose. This computational complexity problem is added to several complications that are assumed away in Proposition 2: first, regulations might not all be cylindric, and some may involve disjunctions as well as conjunctions. In this case, finding out whether a collection of regulations is consistent will already be a computationally difficult task. Second, the vast number of regulations and attributes used in real institutions may make a task that is theoretically simple (that is, of polynomial complexity) practically complex. This problem may be especially acute if the relevant regulations are not organized in a single codex. For example, a large institution such as a state will typically have different collections of regulations issued by different authorities and appearing in different publications. Third, our model assumes that attributes are binary functions, so that it is immediately clear whether a certain problem has a certain attribute or not. In reality there is much more room for interpretation of the regulations in particular cases, as is evident in the legal domain. As a result, it may not be a simple task to find out whether a particular regulation applies in a given problem.

Theorem 3 should therefore be read as follows: even if one were to make several simplifying assumptions, namely that regulations are clearly stated and easily available, and that all of them are defined by simple conjunctions of binary attributes, it would still be a complicated task to find a reasonably general regulation that would enforce a given decision in a given problem.

What happens, then, when an agent copes with such a complex problem? One possibility is that the agent states a general regulation that is later on found to render the set of regulations inconsistent. Another possibility is that consistency is retained at the expense of generality, and the new regulation is perceived as ad hoc, or as clearly serving a certain agenda. In both cases, the institution's set of regulations loses of its appeal and authority. This possibility leads us to a discussion of other conceptualizations of regulations.

## 4.3 Case-Regulation Trade-offs

The assumption that regulations are taken to be strict constraints on decisions may be realistic for some institutions. Some cultures tend to respect regulations more than others, and if the set of regulations is relatively small and simple, one can expect them to be followed. But there are cultures in which rules are expected to be bent and in which some regulations end up being completely ignored. This suggests that one also needs to consider other models, in which regulations are not strict constraints.

One such approach suggests that regulations are "paradigmatic cases", describing a decision that could have, or should have been made. To save on notation, we use the model of Section 2 and assume that $P = P_C \cup P_R$ (with $P_C \cap P_R = \varnothing$) where $P_C$ is a set of "real" problems, which present themselves at a given time, while $P_R$ are "paradigmatic" problems, which were generalized from concrete ones to generate regulations. Similarity, the set of cases $C$ is split into $C_C = P_C \times \{0, 1\}$ and $C_R = P_R \times \{0, 1\}$, and every history $H \subset C$ is described as the disjoint union of $H_C \subset C_C$ and $H_R \subset C_R$.

The rest of the notation of Section 2 is unchanged. Thus, relevance functions,

$$w_0, w_1 : P \times P \to \mathbb{R}_+$$

define a total weight for each decision $d \in \{0, 1\}$ at problem $p \in P_C$ and given history $H$,

$$W_{w_0,w_1}(H, p, d) = \sum_{c=(q,d)\in H} w_d(q, p).$$

and the decision correspondence $f_{w_0,w_1}(H, p)$ is the argmax of $W_{w_0,w_1}$.[8]

[8] Introducing regulations as cases does not pose a problem for the definition of the function $f_{w_0,w_1}$, but its axiomatic derivation calls for a new model, as it does not make sense to assume that regulations are repeated in the same way that cases are.

Suppose that actual problems $q \in P_C$, which are index by time, fade with memory. For example, suppose that, for $p, q \in P_C$, $w_d(q,p)$ is an exponentially decreasing function of the time difference between $q$ and $p$. If at every period only one concrete case $(q,d) \in C_C$ is observed, the total weight of relevance of past concrete cases is bounded (uniformly across all histories). In this case, a single regulation $c \in C_R$ that has a high enough relevance function $w_d$ may outweigh all past concrete cases. Hence, even if a certain decision $d$ has been made throughout history, a regulation $c = (q, 1-d) \in C_R$ can change the common practice.

What determines the relevance of a regulation $c = (q, 1-d)$ to a problem $p \in P_C$? First, one would like to judge whether the regulation applies to the case. In the model of sub-section 4.1 this was a dichotomous question: a regulation was modeled as a rule $(D, d)$ with $D \subset P$, so that $p$ either did or did not belong to the rule's domain. This can easily be simulated by the relevance functions $w_d$, if they vanish outside the rule's domain, and are constant over it. However, the present model allows regulations to apply to concrete cases to a varying degree, depending on the interpretation of the regulation, the exceptionality of the case, and so forth.

Other important factors affecting the relevance of a regulation are the reputation of the authority introducing the regulation, and the regulation's own history. For example, a regulation that has been on the books for years but has been ignored is likely to continue to be ignored (though one may strategically choose to revive it). This suggests that the relevance functions for a regulations should also be a function of history, decreasing in the proportion of cases in which the regulation was violated.

Yet another way in which regulations can be captured in the model is by their inclusion in the description of a problem. A regulation can be thought of as an attribute of the problem, whose value indicates whether the regulation applies in the problem. Thus, regulations can change the perceived relevance of past cases to a present problem. Also, one may be interested in the relevance of one regulation to another: if the institution's culture is such that previous regulations have been ignored, new regulations might also be ignored, more so than in an institution that has a history of respecting regulations. Clearly, these effects are beyond the scope of the model presented here.

# 5 Concluding Comments

## 5.1 Decisions without Utility

We study an institution who makes an authorization decision. Our basic model of the decision making unit is that it seeks consistency with past decisions, perhaps constrained by regulations. Importantly, we do not ascribe a utility function to the institution. In this sense, our model differs from the standard approach to modeling institutional decisions, in which agents have utility functions. The agents may be entire institutions, sub-units thereof, or even individuals; they can be fully rational maximizing agents, or boundedly rational satisficing agents; but in all these different models there is a concept of a utility that one can meaningfully ascribe to the agent. By contrast, we suggest a formal model in which there is no useful concept of "utility".

Applying this approach to economic requires a combination of the utility-based and the inertia-based models of agents. For example, suppose that we wish to study competition in a market for medications. The firms that produce medications might well be modeled as standard profit-maximizing agents. By contrast, the government agency that approves medications might be described as an institution that seeks consistency with its past decision, given regulations. Clearly, there is a conceptual cost in having more than one type of agent in a model. Yet, the plurality of decision making models may provide insights that the utility-based model might miss.

## 5.2 Case-Based Decisions

Gilboa and Schmeidler (1995) suggest a theory of case-based decisions, where the decision maker faces uncertainty and tends to choose acts that performed well in similar past cases. Each case is a triple consisting of a problem, an act, and an outcome. In the basic model an act is evaluated based only on the cases in which it has been chosen. Thus, each past problem gives an indication about the quality of the act chosen in it, but not about that of the other acts: the decision maker is not assumed to engage in counterfactual reasoning about the outcomes that other acts would have resulted in.

There are situations in which the problem of counterfactual reasoning does not arise. In a prediction problem, the predictor attempts to guess the outcome, but her decision, namely, the guess, has no impact on it. Thus, one can think of a case as a pair of a problem and an outcome, omitting the act. In a model such as this, case-based decision theory is conceptually simplified, and becomes closer to kernel classification methods (see Gilboa and Schmeidler, 2003).

The model presented here, while mathematically very similar to the kernel classification model, can be viewed as the other extreme simplification of the general case-based decision model: rather than suppressing the act, it suppresses the outcome. When an institution makes a permit decision in our model, there is no "outcome", no external source of evidence about the quality of the decision made. In the absence of such evidence, all that is left to base future decisions on is the decision itself. This is akin to a court decision, which, for the most part, is not externally evaluated as right or wrong, good or bad.[9] Our institution, like the court, can seek consistency with precedences, but it cannot ask which were the "right" decisions in them. In a sense, the decision that has been made in each case is, by definition, the correct one for that case.

[9] A court decision might be reversed by an appeal, but in the final analysis it is the legal system that defines the "correct" decision.

## 5.3 Randomness

A natural generalization of the basic model would introduce random factors into the institution's decisions. Thus, in section 2 one may assume that the actual decision that the institution makes, and that is added to history $H$, is $f(H,p)$ with probability $(1-\varepsilon)$ and the opposite decisions $(1-f(H,p))$ with probability $\varepsilon < 0.5$. The noise variable $\varepsilon$ can capture various factors that are left out of the model, or that are inherently random. In such a model one may find that, starting with a given history, the institution might evolve to develop different norms: with some probability it might converge to tend to make the decision 1 in certain cases, and with some probability – the decision 0, where switching between these decisions can also occur indefinitely. This type of indeterminacy appears to be natural whenever the institution is assumed to aim for consistency with its own past decisions, rather than with a well-defined external goal.

# 6 Appendix A: Proofs

## 6.1 Proof of Theorem 1

The proof of this theorem relies on that of the main theorem in Gilboa and Schmeidler (2003). The main differences between these theorems are the following: in this paper we deal with the simpler case in which there are only two decisions to be ranked (0 and 1). On the other hand, we do not assume here a "diversity" axiom, and this allows the possibility that, for some problems $p$, one or both of the sets $\{0\}$, $\{1\}$ is not in the domain of $f(\cdot, p)$. We now turn to the proof, relying on Gilboa and Schmeidler (2003) whenever possible.

It is straightforward to verify that, for any pair of relevance functions $w_0, w_1 : P \times P \to \mathbb{R}_+$, the decision correspondence $f_{w_0,w_1}$ satisfies Axioms 1-3. We now wish to prove the converse.

Let there be given a decision correspondence $f$ satisfying Axioms 1-3. Fix $p \in P$. We will show that, for this $p$, there are functions

$$w_0^p, w_1^p : P \to \mathbb{R}_+$$

such that, for every $H$ such that $p \notin H_P$,

$$f(H,p) = \arg\max_{d \in \{0,1\}} \sum_{c=(q,d) \in H} w_d^p(q)$$

denoting $w_0(q,p) = w_d^p(q)$ will complete the proof. Since the argument is carried out for each $p$ separately, we will omit $p$ from the notation for the rest of the proof. Thus, $f$ will be defined on histories $H$ such that $p \notin H_P$ and we seek $w_0, w_1 : P \to \mathbb{R}_+$ such that

$$f(H) = \arg\max_{d \in \{0,1\}} \sum_{c=(q,d) \in H} w_d(q). \tag{2}$$

Recall that $f$ is assumed to satisfy the richness assumption. Observe that $\sim_f$ is an equivalence relation on $P$ and on $C$, and that, by assumption, all their equivalence classes are infinite.

Denote by $\mathbb{T}$ the set of *case types*, namely, the set of the equivalence classes of $\sim_f$. For $H \subset C$, define $I_H : \mathbb{T} \to \mathbb{Z}_+ \equiv \{0, 1, 2, ...\}$ to be the counter vector defined by $H$, so that, for $T \in \mathbb{T}$, $I_H(T) = \#(H \cap T)$. Since $H$ is finite, $\sum I_H(T) < \infty$ for every $H$. Define the set of all possible counter vector to be

$$\mathcal{I} = \left\{ I : \mathbb{T} \to \mathbb{Z}_+ \;\middle|\; \sum I(T) < \infty \right\}.$$

Note that, indeed, for every $I \in \mathcal{I}$ there exists $H \in \mathcal{H}$ such that $I = I_H$. Further, for two histories $H, H' \in \mathcal{H}$, if $I_H = I_{H'}$, then $H \approx H'$. It follows that we can define $f$ on $\mathcal{I}$ rather than on $\mathcal{H}$: for $I \in \mathcal{I}$ let $f(I)$ be $f(H)$ for a history $H$ such that $I_H = I$.

For each vector $I \in \mathcal{I}$ the correspondence $f$ defines a binary relation $\succsim_I$ on $\{0, 1\}$: for $d \in \{0, 1\}$, $d \succsim_I d$ and $d \succsim_I (1 - d)$ iff $d \in f(I)$. It is straightforward to see that $\succsim_I$ satisfies axioms A1*, A2*, A3* of Theorem 2 in Gilboa and Schmeidler (2003, p. 16).

Next, assume that there exist $I_0, I_1 \in \mathcal{I}$ for which $f(I_d) = \{d\}$ for $d = 0, 1$. In this case, axiom A4* of that theorem also holds, and the theorem implies that there exists

$$v : \{0, 1\} \times \mathbb{T} \to \mathbb{R}$$

such that, for all $I$, and $d = 0, 1$,

$$d \in f(I) \qquad \Leftrightarrow \qquad \sum_{T \in \mathbb{T}} I(T)v(d, T) \geq \sum_{T \in \mathbb{T}} I(T)v(1 - d, T). \tag{3}$$

Clearly, the function $v(\cdot, T)$ may be shifted by a constant $c_T$ for each $T$ separately, without changing the representation (3).

Let there be given $q \neq p$. Consider the two cases $c_0 = (q, 0)$ and $c_1 = (q, 1)$ and let $T_0$ and $T_1$ be their case types (so that $c_d \in T_d$, $T_0, T_1 \in \mathbb{T}$). Shifting the function $v(\cdot, T)$ by a $c_T = -v(d, T_{1-d})$, we may assume, without loss of generality, that $v(d, T_{1-d}) = 0$. If, under this normalization, it were the case that $v(d, T_d) < 0$, monotonicity (Axiom 3) would be violated. Hence we find that (3) holds with $v(d, T_{1-d}) = 0$ and $v(d, T_d) \geq 0$. Defining $w_d(q) = v(d, T_d)$, the representation (3) yields (2).

We wish to establish that a representation (2) can be obtained also if A4* does not hold. Suppose, first, that $f(I) = \{0, 1\}$ for all $I$. In this case, $w_d \equiv 0$ (for $d = 0, 1$) satisfies (2). Next assume that for one $d \in \{0, 1\}$ there exists $I_d$ such that $f(I_d) = \{d\}$ but not for the other. Without loss of generality, assume that there exists $I_1 \in \mathcal{I}$ for which $f(I_1) = \{1\}$ but there does not exist $I_0 \in \mathcal{I}$ with $f(I_0) = \{0\}$. Let $\mathbb{T}_1 \subset \mathbb{T}$ be defined as the case types that lead to the choice of 1 alone, that is,

$$\mathbb{T}_1 = \{\, T \subset \mathbb{T} \mid f(I_T) = \{1\} \,\}$$

where $I_T$ is defined by $I_T(T) = 1$ and $I_T(T') = 0$ for $T' \neq T$. It is easy to verify that, due to Axiom 1 (Combination),

$$f(I) = \begin{cases} \{1\} & if \quad \exists T \in \mathbb{T}_1 \quad f(I_T) > 0 \\ \{0, 1\} & \qquad otherwise \end{cases}$$

In this case, set $w_0 \equiv 0$. As for $w_1$, let there be given $q \neq p$. Let $T$ be the case-type containing $(q, 1)$. If $T \in \mathbb{T}_1$, choose $w_1(q) > 0$, and set $w_1(q) = 0$ otherwise. It is easy to verify that with this definition, $w_0, w_1$ satisfy (2). $\square$

## 6.2 Proof of Proposition 2

Let there be given a number of attributes $m$, a set of regulations $R = \{(J_i, b_i, d_i)\}_{i=1}^n$, a problem $p$ and a decision $d$. Assume, without loss of generality, that $d_i = 0$ for $i \leq l$ and $d_i = 1$ for $i > l$, where $0 \leq l \leq n$. Then $R$ is consistent iff for every $i \leq l$ and $i' > l$, there exists $1 \leq j \leq m$, such that $j \in J_i \cap J_{i'}$ and $b_i(j) \neq b_{i'}(j)$. This can be verified in time complexity which is $\mathcal{O}(mn^2)$.

Next, for a problem $p$ and a regulation $(J_i, b_i, d_i)$, it takes $\mathcal{O}(m)$ steps to verify whether $p \in D(J_i, b_i)$. Hence it takes no more than $\mathcal{O}(mn)$ steps to find out whether $p \in R(0)$ and whether $p \in R(1)$.

Finally, if $R$ is consistent and $p \notin R(0), R(1)$, one can find the regulation $(J_{n+1}, b_{n+1}, d_{n+1})$ defined by $J_{n+1} = \{1, ..., m\}$, $b_{n+1}(j) = a_j(p)$ and $d_{n+1} =$

$d$. Clearly, $D(J_{n+1}, b_{n+1})$ consists only of problems that are identical to $p$ in terms of all attributes $a_1, ..., a_m$. Hence the decision in none of them is restricted by $R = \{(J_i, b_i, d_i)\}_{i=1}^n$ and therefore $R' = \{(J_i, b_i, d_i)\}_{i=1}^{n+1}$ is consistent. $\square$

## 6.3 Proof of Theorem 3

That the problem is in NP is straightforward (and follows from Proposition 2). The proof that the problem is NP-Complete is by a straightforward reduction of the set cover problem (which is known to be NP-Complete, see Gary and Johnson, 1979):

**Problem COVER:** Given a natural number $r$, a set of $q$ subsets of $S \equiv \{1, ..., r\}$, $\mathfrak{S} = \{S_1, ..., S_q\}$, and a natural number $t \leq q$, are there $t$ subsets in $\mathfrak{S}$ whose union contains $S$?

(That is, are there indices $1 \leq j_1 \leq ... \leq j_t \leq q$ such that $\bigcup_{l \leq t} S_{j_l} = S$ ?)

Let there be given an instance of the COVER problem, $r$, $\mathfrak{S} = \{S_1, ..., S_q\}$, $t \leq q$. Assume without loss of generality that $\bigcup_{j \leq q} S_j = S$.[10]

Define $m = q$, $n = r$, and $k = t$. Next define the set of regulations $R = \{(J_i, b_i, d_i)\}_{i=1}^n$ as follows: for $i \leq n$, $J_i = \{\, j \mid i \in S_j \,\}$, for $j \in J_i$, $b_i(j) = 1$, and $d_i = 1$. Finally, choose $p$ such that $a_j(p) = 0$ for all $j \leq m$ and $d = 0$. Because $\bigcup_{j \leq q} S_j = S$, $J_i$ is nonempty.

The set of regulations is consistent because all regulations imply $d_i = 1$ (and none has $d_i = 0$). This also means that $R(0) = \varnothing$, hence, in particular, $p \notin R(0)$. Observe also that, for each $i \leq n$, $b_i(j) = 1$ for at least one $j$ (because $b_i(j) = 1$ for every $j$ in $J_i$, while we know that $J_i \neq \varnothing$). This means that $p \notin D(J_i, b_i)$ and therefore $p \notin R(1)$.

We now turn to show that there exists a regulation $(J_{n+1}, b_{n+1}, 0)$ such that $p \in D(J_{n+1}, b_{n+1})$, $R' = \{(J_i, b_i, d_i)\}_{i=1}^{n+1}$ is consistent and $|J_{n+1}| \leq k$ iff

[10] Clearly, COVER is also NP-Complete if we restrict the input to be such that $\bigcup_{j \leq q} S_j = S$. To see this, observe that the standard problem COVER can be reduced to the restricted one by adding, for each $i \in S \setminus \bigcup_{j \leq q} S_j$, a singleton $\{i\}$ to the collection $\mathfrak{S}$.

there is a cover of size $t = k$ for $S$ in the original problem.

A regulation $(J_{n+1}, b_{n+1}, 0)$ such that $p \in D(J_{n+1}, b_{n+1})$ has to satisfy $b_{n+1}(j) = 0$ for every $j \in J_{n+1}$. Hence there is a bijection between the regulations $(J_{n+1}, b_{n+1}, d_{n+1})$ we need to consider and subsets of $\{1, ..., m\}$: every subset $J \subset \{1, ..., m\}$ defines the regulation $J_{n+1} = J$, $b_{n+1}(j) = 0$ for $j \in J$, and $d_{n+1} = 0$, and vice versa.

Given such a subset $J_{n+1} \subset \{1, ..., m\}$, when will $R' = \{(J_i, b_i, d_i)\}_{i=1}^{n+1}$ be consistent? Because every regulation $(J_i, b_i, d_i)$ for $i \leq n$ has $d_i = 1$, it has to be the case that $D(J_i, b_i) \cap D(J_{n+1}, b_{n+1}) = \varnothing$. But this implies that for every $i \leq n$ there exists $j \in J_{n+1}$ such that $b_i(j) = 1$, namely that $i \in S_j$. In other words, $D(J_i, b_i) \cap D(J_{n+1}, b_{n+1}) = \varnothing$ for every $i \leq n$ iff $\{S_j\}_{j \in J_{n+1}}$ is a cover for $S$. Thus, there exists such a regulation with $|J_{n+1}| \leq k = t$ if and only if there exists a cover of size $t$ for $S$ in the original problem.

It remains to note that the construction is carried out in polynomial time.

$\square$

# 7 Appendix B: Computational Complexity

A **problem** can be thought of as a set of legitimate inputs, and a correspondence from it into a set of legitimate outputs. For instance, consider the problem "Given a graph, and two nodes in it, $s$ and $t$, find a minimal path from $s$ to $t$". An input would be a graph and two nodes in it. These are assumed to be appropriately encoded into finite strings over a given alphabet. The corresponding encoding of a shortest path between the two nodes would be an appropriate output.

An **algorithm** is a method of solution that specifies what the solver should do at each stage. **Church's thesis** maintains that algorithms are those methods of solution that can be implemented by **Turing machines**. It is neither a theorem nor a conjecture, because the term "algorithm" has no formal definition. In fact, Church's thesis may be viewed as defining an "algorithm" to be a Turing machine. It has been proved that Turing machines are equivalent, in terms of the algorithms they can implement, to various other computational models. In particular, a PASCAL program run on a modern computer with an infinite memory is also equivalent to a Turing machine and can therefore be viewed as a definition of an "algorithm".

It is convenient to restrict attention to **YES/NO problems**. Such problems are formally defined as subsets of the legitimate inputs, interpreted as the inputs for which the answer is YES. Many problems naturally define corresponding YES/NO problems. For instance, the previous problem may be represented as "Given a graph, two nodes in it $s$ and $t$, and a number $k$, is there a path of length $k$ between $s$ and $t$ in the graph?" It is usually the case that if one can solve all such YES/NO problems, one can solve the corresponding optimization problem. For example, an algorithm that can solve the YES/NO problem above for any given $k$ can find the minimal $k$ for which the answer is YES (it can also do so efficiently). Moreover, such an algorithm will typically also find a path that is no longer than the specified $k$.

Much of the literature on computational complexity focuses on **time complexity**: how many operations will an algorithm need to perform in order to obtain the solution and halt. It is customary to count input/output operations, as well as logical and algebraic operations as taking a single unit of time each. Taking into account the amount of time these operations actually take (for instance, the number of actual operations needed to add two numbers of, say, 10 digits) typically yields qualitatively similar results.

The literature focuses on **asymptotic** analysis: how does the number of operations grow with the size of the input. It is customary to conduct **worst-case** analyses, though attention is also given to average-case performance. Obviously, the latter requires some assumptions on the distribution of inputs, whereas worst-case analysis is free from distributional assumptions. Hence the complexity of an algorithm is generally defined as the order of magnitude of the number of operations it needs to perform, in the worst case, to obtain a solution, as a function of the input size. The complexity of a problem is the minimal complexity of an algorithm that solves it. Thus, a problem is **polynomial** if there exists an algorithm that always solves it correctly within a number of operations that is bounded by a polynomial of the input size. A problem is **exponential** if all the algorithms that solve it may require a number of operations that is exponential in the size of the input, and so forth.

Polynomial problems are generally considered relatively "easy", even though they may still be hard to solve in practice, especially if the degree of the polynomial is high. By contrast, exponential problems become intractable already for inputs of moderate sizes. To prove that a problem is polynomial, one typically points to a polynomial algorithm that solves it. Proving that a YES/NO problem is exponential, however, is a very hard task, because it is generally hard to show that there does *not* exist an algorithm that solves the problem in a number of steps that is, say, $O(n^{17})$ or even $O(2^{\sqrt{n}})$.

A **non-deterministic Turing machine** is a Turing machine that allows multiple transitions at each stage of the computation. It can be thought of as a parallel processing modern computer with an unbounded number of processors. It is assumed that these processors can work simultaneously, and, should one of them find a solution, the machine halts. Consider, for instance, the Hamiltonian path problem: given a graph, is there a path that visits each node precisely once? A straightforward algorithm for this problem would be exponential: given $n$ nodes, one needs to check all the $n!$ permutations to see if any of them defines a path in the graph. A non-deterministic Turing machine can solve this problem in linear time. Roughly, one can imagine that $n!$ processors work on this problem in parallel, each checking a different permutation. Each processor will therefore need no more than $O(n)$ operations. In a sense, the difficulty of the Hamiltonian path problem arises from the multitude of possible solutions, and not from the inherent complexity of each of them.

The class **NP** is the class of all YES/NO problems that can be solved in **P**olynomial time by a **N**on-deterministic Turing machine. Equivalently, it can be defined as the class of YES/NO problems for which the validity of a suggested solution can be verified in polynomial time (by a regular, deterministic algorithm). The class of problems that can be solved in polynomial time (by a deterministic Turing machine) is denoted **P** and it is obviously a subset of NP. Whether P=NP is considered to be the most important open problem in computer science. While the common belief is that the answer is negative, there is no proof of this fact.

A problem $A$ is **NP-Hard** if the following statement is true ("the conditional solution property"): if there were a polynomial algorithm for $A$, there would be a polynomial algorithm for any problem $B$ in NP. There may be many ways in which such a conditional statement can be proved. For instance, one may show that using the polynomial algorithm for $A$ a polynomial number of times would result in an algorithm for $B$. Alternatively, one may show a polynomial algorithm that translates an input for $B$ to an input for $A$, in such a way that the $B$-answer on its input is YES iff so is the $A$-answer of its own input. In this case we say that $B$ is **reduced** to $A$.

A problem is **NP-Complete** if it is in NP, and any other problem in NP can be reduced to it. It was shown that the **SATISFIABILITY** problem (whether a Boolean expression is not identically zero) is such a problem by a direct construction. That is, there exists an algorithm that accepts as input an NP problem $B$ and input for that problem, $z$, and generates in polynomial time a Boolean expression that can be satisfied iff the $B$-answer on $z$ is YES. With the help of one problem that is known to be NP-Complete (**NPC**), one may show that other problems, to which the NPC problem can be reduced, are also NPC. Indeed, it has been shown that many combinatorial problems are NPC.

NPC problem are NP-Hard, but the converse is false. First, NP-Hard problems need not be in NP themselves, and they may not be YES/NO problems. Second, NPC problems are also defined by a particular way in which the conditional solution property is proved, namely, by reduction.

There are by now hundreds of problems that are known to be NPC. Had we known one polynomial algorithm for one of them, we would have a polynomial algorithm for each problem in NP. As mentioned above, it is believed that no such algorithm exists.

# 8 References


Bendor, J. and T. M. Moe (1985), "An Adaptive Model of Bureaucratic Politics", *The American Political Science Review*, **79**: 755-774.

Buchanan, J. M. and G. Tullock (1962), *The Calculus of Consent: Logical Foundations of Constitutional Democracy.* Ann Arbor: The University of Michigan Press.

Burns, T. (1961), "Micropolitics: Mechanisms of Organizational Change", *Administrative Science Quarterly*, **6**: 257-281.

Burns, T. and G. M. Stalker (1961), *The Management of Innovation.* London: Tavistock.

Coase, R. H. (1937), "The Nature of the Firm", *Economica*, New Series, **IV**: 386-405.

Debreu , G. (1959). *The Theory of Value.* Cowles Foundation Monographs Series.

Durkheim, E. (1893), *The Division of Labor in Society.* English translation, 1949, Glencoe, Il: Free Press.

Fayol, H. (1919), *General and Industrial Management.* English translation, 1949, London: Pitman.

Follett, M. P. (1918), *The New State: Group Organization and the Solution of Popular Government.* 1923 edition, New York: Longman, Green, and Co.

Gary, M. and D. S. Johnson (1979), *Computers and Intractability: A Guide to the Theory of NP-Completeness.* San-Francisco, CA: W. Freeman and Co.

Gilboa, I. and D. Schmeidler (1995). "Case Based Decision Theory", *Quarterly Journal of Economics,* **110:** 605-639.

Gilboa, I. and D. Schmeidler (2003), "Inductive Inference: An Axiomatic Approach", *Econometrica*, 71: 1-26.

Handel, M. J. (2003), *The Sociology of Organizations.* Thousand Oaks, CA: Sage Publications.

Hatch, M.-J. and A. L. Cunliffe (2006), *Organization Theory.* 2nd Edition. Oxford: Oxford University Press.

Jensen, M. C. and W. H. Meckling (1976), "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure", *Journal of Financial Economics*, **3**: 305-360.

Kanter, R. M. (1977), *Men and Women of the Corporation.* New York: Basic Books.

March, J. G. (1962), "The Business Firm as a Political Coalition", *Journal of Politics*, **24**: 662-678.

March, J. G. (1999), *The Pursuit of Organizational Intelligence.* Malden, MA: Blackwell.

March J. G. and J. P. Olsen (1976), *Ambiguity and Choice in Organizations.* Bergen: Universitetsforlaget.

March J. G. and H. A. Simon (1958), *Organizations.* New York: John Wiley and Sons.

Marx, K. (1867), *Capital.* English Translation, 1954, Moscow: Foreign Languages Publishing House.

Morgan, G. (2006), *Images of Organization.* Thousand Oaks, CA: Sage Publications.

Myerson, R. B. (1995), "Axiomatic Derivation of Scoring Rules Without the Ordering Assumption", *Social Choice and Welfare*, 12, 59-74.

von Neumann, J. and O. Morgenstern (1944), *Theory of Games and Economic Behavior.* Princeton, NJ: Princeton University Press.

Niskanen, W. (1971), *Bureaucracy and Representative Government.* New Brunswick: Aldine Transaction.

Niskanen, W. (1975), “Bureaucrats and Politicians”, *Journal of Law and Economics*, **18**: 617-644.

Ouchi, W. and A. Wilkins (1985), “Organizational Culture”, *Annual Review of Sociology*, **11**: 457-483.

Parsons, T. (1951), *The Social System*. New York: The Free Press.

Sandelands, L. E. and R. E. Stablein (1987), “The Concept of Organization Mind”, *Research in the Sociology of Organizations*, **5**. Greenwich, CT: JAI Press.

Savage, L. J. (1954), *The Foundations of Statistics*. New York: John Wiley and Sons.

Scott, W. R. and G. F. Davis (2007), *Organizations and Organizing: Rational, Natural, and Open System Perspectives*. Upper Saddle River, NJ: Pearson Education.

Simon, H. A. (1947), *Administrative Behavior*. New York: McMillan.

Smith, A. (1776), *The Wealth of Nations*. 1957 edition, New York: Appleton Century Crofts.

Taylor, F. W. (1911), *The Principles of Scientific Management*. New York: Harper.

Thompson, J. D. (1967), *Organization in Action*. New York: McGraw-Hill.

Walsh, J. P. and G. R. Ungson (1991), “Organizational Memory”, *Academy of Management Review*, **16**: 57-91.

Weber, M. (1921), *Economy and Society: An Outline of Interpretive Sociology*. English translation, 1968, Berkeley, CA: University of California Press.

Weber, M. (1924), *The Theory of Social and Economic Organization*. English translation, 1947, Glencoe, Il: Free Press.

Williamson, O. E. (1975), *Markets and Hierarchies*. New York: Free Press.

Williamson, O. E. (1979), "Transaction Cost Economics: The Governance of Contractual Relations", *Journal of Law and Economics*, **22**: 233-261.

Williamson, O. E. (1981), "The Economics of Organization: The Transaction Cost Approach", *American Journal of Sociology*, **87**: 548-577.

Young, H. P. (1975), "Social Choice Scoring Functions", *SIAM Journal of Applied Mathematics*, 28: 824-838.