Concepts of Implementation

Thomson, William

Working Paper No. 396
December 1994

University of Rochester
Concepts of Implementation

William Thomson

Rochester Center for Economic Research
Working Paper No. 396
CONCEPTS OF IMPLEMENTATION

William Thomson*

April 1993
Revised November 1994

*University of Rochester, Rochester, NY 14627. I would like to thank R. Calvert and G. McDonald for comments that motivated me to write this note. I am also grateful to L. Corchón for his suggestions.
Contents

1 Introduction 1

2 Full versus partial implementation. The case for full implementation 3
   2.1 The case for full implementation based on the definition of the notion of a social choice correspondence understood in its full generality 3
   2.1.1 Single-profile considerations 3
   2.1.2 Multi-profile considerations 5
   2.2 Characterizing full implementability is mathematically identical to characterizing partial implementability 6
   2.3 The relevance of implementation theory to non-cooperative games. Another case for full implementation 6

3 What to do when full implementation is not possible 7
   3.1 Overimplementation and minimal overimplementation 7
   3.2 Underimplementation and maximal underimplementation 9
   3.3 Approximate implementation 9

4 When partial implementation is a valid objective 9
   4.1 Minimal partial implementation 10
   4.2 Single-valued partial implementation 11

5 Two other issues: natural implementation and multiple implementation 12

6 Conclusion 13

7 References 15
1 Introduction

The purpose of this note is to present and discuss various notions of implementation that have been used in the literature. Its focus is on the considerations that should guide the choice of the required relationship between the set of equilibrium outcomes of the game form and the set of outcomes chosen by the correspondence that it is supposed to implement. Informal statements of the central definitions are as follows.

A social choice problem, or simply a problem, is given by specifying a set of feasible alternatives \( A \), and the preferences of a group of agents over \( A \). A social choice correspondence, or simply a correspondence, is a mapping \( F \) which associates with each profile of preferences \( R = (R_1, ..., R_n) \) in some admissible domain a non-empty subset of \( A \), denoted by \( F(R) \). The alternatives in \( F(R) \) are the \textbf{F-optimal alternatives} for the profile \( R \). The theory of implementation is concerned with the fact that preferences are typically not public information. When agents are asked to announce them, they might behave strategically, this resulting in the wrong decision being made.

However, instead of letting strategic behavior get in the way of reaching the social objectives, one can direct it to one’s advantage, by constructing well chosen “game forms”: a \textbf{game form} specifies for each agent a set of possible actions that he may choose, and a function, the \textbf{outcome function}, associating an alternative with each profile of actions. Each agent is free to choose his action so as to influence the outcome in his favor, that is, to choose strategically from his action space, which we will therefore call a \textbf{strategy space}. Naturally, which outcomes are obtained at equilibrium depends on the profile of true preferences. We say that the \textbf{game form implements the correspondence} if for each admissible profile of preferences, the set of equilibrium outcomes of the game form for that profile coincides with the set of \( F \)-optimal outcomes for it, that is, all \( F \)-optimal outcomes, and only \( F \)-optimal outcomes, are obtained as equilibrium outcomes.

The standard goals of the theory of implementation are (i) the characterization of which correspondences are implementable, and (ii) the construction of game forms achieving the implementation. Recently, much attention has been given to the goal of (iii) characterizing which correspondences can be implemented by “simple” game forms, and again, constructing such game forms.
A weaker implementation requirement is that the set of equilibrium outcomes be a non-empty subset of the set of $F$-optimal outcomes. Although this notion of partial implementation is sometimes mentioned in the literature, it is almost always the stronger one, which we will call full implementation, that is studied, no justification being given for the more stringent requirement.

Yet, the question should be asked: why not be satisfied with partial implementation? Central to our presentation in the next few pages is an attempt at answering this question. In so doing, we will find it useful to present yet other notions of implementation.

The points we will make are independent of which equilibrium concept is being used. However, in the illustrations, we mostly limit ourselves to implementation in Nash equilibrium.

Essentially, our conclusions are as follows. (ia) When the concept of a choice correspondence is understood in its full generality, as a mapping from a domain of preference profiles to some outcome space embodying desirable social norms, it is full implementation that is the correct objective. (ib) Moreover, the theoretical problem of characterizing the class of partially implementable correspondences is mathematically identical to the problem of characterizing the class of fully implementable correspondences. (ic) Third, we point out a useful connection between implementation theory and the theory of non-cooperative games. Here too, full implementation is the relevant notion.

(iiia) It is only when full implementation of a correspondence is not possible that one may want to consider its partial implementation; but then one should refine this goal and have what we will call its “maximal partial implementation” as second-best objective. (iib) Alternatively, “overimplementation” of the correspondence can be considered, and here the right second-best objective is what we will discuss under the phrase of “minimal overimplementation”. We show that these concepts are all well-defined whenever the conditions for implementation are closed under intersection and union, as is the case for the conditions that play a central role for Nash implementation. (iic) We will also discuss a notion of approximate implementation.

(iiiia) In spite of our claim that full implementation is in most cases the correct objective, we admit that practical considerations may sometimes justify that attention be paid to partial implementation. One such consideration is that having small equilibrium sets may increase the likelihood of
equilibrium being reached. Another reason is that it may be mathematically
difficult to obtain a characterization of full implementability. Then, some
useful results may be available about partial implementability.

2 Full versus partial implementation. The
case for full implementation

In this section we develop several arguments for focusing on full implement-
tation.

2.1 The case for full implementation based on the def-
inition of the notion of a social choice correspond-
dence understood in its full generality

The argument that is made in favor of partial implementation is based on the
interpretation of $F$ as representing the “preferences” of a social planner over
alternatives. According to that interpretation, the fact that for the profile $R$,
the set $F(R)$ may contain several alternatives, (perhaps many, perhaps an
infinity), is simply a reflection of the planner’s indifference. So, the argument
goes, Why should he care if only some of his preferred outcomes are obtained
as equilibrium outcomes?

2.1.1 Single-profile considerations

A partial answer to this criticism is that $F$ may alternatively be interpreted
as providing a first “narrowing down” of the set of feasible outcomes, the
final choice being decided at a later stage, on the basis of considerations that
are not explicitly specified at the time the game form is being played.\footnote{In
order to have that freedom in the future, and this “ex ante” justification to be
valid, it is necessary that all the points in $F(R)$ be identified. It is of course
difficult to imagine how this identification will occur. It certainly will not if the
game form is played only once.}

\footnote{Note that, whether one is interested in implementation or in other issues, the position
could be taken that one should only care about social choice functions since eventually
no more than one outcome can ever be selected. Then the concept of a social choice correspondence has
never any relevance.}
That such considerations may have to be brought in is clear since otherwise it would for example be satisfactory to partially implement a given correspondence by the subcorrespondence that systematically picks, for each preference profile in its domain, the one among the outcomes that the correspondence would have selected that the first agent prefers. But such partial implementation is unlikely to be acceptable, and certainly will not be acceptable if the correspondence embodies some minimal concerns about fairness in distribution, as almost all do.

To show that this is more than a theoretical possibility, consider the Divide and Choose game, a game designed to generate a “fair” division of a vector of resources on which two agents are assumed to have equal rights. It has the following rules: one agent selected in advance, and called the “divider”, divides the resources into two shares; the other, called the “chooser”, chooses one of the two shares. At equilibrium, this game produces an envy-free allocation that is unique up to Pareto-indifference, and therefore, it provides a partial implementation of the no-envy solution. However, this envy-free allocation is the one that the divider prefers (Crawford, 1979); the chooser would benefit from being the chooser instead.\(^3\) Such partial implementation of the no-envy correspondence seems in contradiction with the basic objective of fairness.

The above discussion suggests that one should make sure that the complete list of desired properties of correspondences have been identified - in the example, such a list would include the “symmetric” treatment of agents - and then identify the class of correspondences satisfying them. If all the properties are satisfied only by \(F\) (that is, if \(F\) is characterized by these properties), then full implementation of \(F\) is indeed what we should be after. If, on the other hand, they are satisfied by correspondences other than \(F\), then the question of implementation should be phrased more generally thus: among all correspondences in the class, is there any that is fully implementable? But again, the objective is full implementation.

To summarize, our first point is that the recommendations made by \(F\) for each profile of preferences should be evaluated “globally”, as a set, and that it is not meaningful to think of individual allocations in \(F(R)\) as being preferred by the planner to allocations not in \(F(R)\). Instead, it is the set

\(^3\)Removing this asymmetry is one of Crawford’s main objectives in his work on the subject.
that the planner has chosen for the profile \( R \), as opposed to other subsets of the feasible set.

2.1.2 Multi-profile considerations

This argument is incomplete however, since some of the properties of \( F \) that make it desirable may have to do with the way it responds to changes in the preference profile, an issue that the planner has had to consider since the implementation issue arises precisely because the economy may be any element in a certain class. The considerations underlying the formulation of most of the criteria that investigators have found useful in their search for good allocation rules are simply not compatible with the view that all subcorrespondences of a chosen correspondence are as desirable as that correspondence. The fairness property discussed in connection with the previous example pertains to the choices \( F \) makes for each profile separately, but \( F \) may have been selected because it responds in a particularly appealing way to changes in profiles. This desirable responsiveness may not hold for its subcorrespondences. The choice of \( F(R) \) for each profile \( R \) in the admissible domain should be understood in relation to the choice of \( F(R') \) for each other admissible profile \( R' \). If earlier, we made an argument for full implementation on the basis of “single-profile” considerations, our point here is based on “multi-profile” considerations.

To make it in another way, consider a change in the admissible domain but let \( R \) be a profile that is present in both cases. This change in domains may actually change the set of allocations that the planner would pick for that profile. Earlier, we said that one should not speak of an allocation being desirable for the profile \( R \); here we claim that one should not even speak of a set of allocations as being desirable for the profile \( R \); one should speak of a set of allocations being desirable for the profile \( R \) given the domain of possible profiles from which \( R \) is drawn.

To emphasize the point, suppose that one of the reasons why a certain correspondence \( F \) was selected is that it satisfies a certain property pertaining to changes in the profile. It might be that all proper subcorrespondences of \( F \) violate the property. Accepting a partial implementation of \( F \) would therefore imply giving it up. But if the domain of preferences were to change, subcorrespondences of \( F \) may very well become available that both satisfy the property and are fully implementable.
2.2 Characterizing full implementability is mathematically identical to characterizing partial implementability

Consider a correspondence $F$ that is fully implementable. Augment it by adding, for each profile in its domain, an arbitrary subset of the feasible set to whatever $F$ would have selected for that profile. This enlarged correspondence is obviously partially implementable. Moreover, since the enlargement was arbitrary, the correspondence has no property that makes it partially implementable beyond the fact that it has a fully implementable subcorrespondence, namely $F$. The task of characterizing the class of partially implementable correspondences is therefore identical to the task of characterizing the class of fully implementable correspondences. A theorem of the form: “A correspondence is fully implementable if and only if it satisfies conditions $C1 - Ck$” can automatically be converted into one that states: “A correspondence is partially implementable if and only if it contains a correspondence satisfying conditions $C1 - Ck$”.

Similarly, one could be looking for a game form whose equilibrium correspondence has a non-empty intersection with $F$, as proposed by Bandyopadhyay and Samuelson (1992). But since correspondences implementable in this way are obtained by arbitrary augmentations of a fully implementable subcorrespondence of $F$, characterizing the class they constitute is again identical to the task of characterizing the class of fully implementable correspondences.

2.3 The relevance of implementation theory to non-cooperative games. Another case for full implementation

We provide one last reason for being interested in understanding the implications of full implementability.

The theory of implementation is usually understood to pertain to the interface between welfare economics and social choice on the one hand, and the theory of non-cooperative games on the other. However, we claim that it has direct relevance to the theory of non-cooperative games: consider a game form, and let $F$ be the correspondence that associates with every pref-
erence profile for which this game form can meaningfully be played the set of equilibrium outcomes of the game form when played by agents having these preferences. This is its equilibrium correspondence. The game form obviously fully implements the correspondence. If we know what fully implementable correspondences look like and what properties they necessarily satisfy, that information will be useful in understanding the equilibrium correspondences of game forms.

3 What to do when full implementation is not possible

When a correspondence of interest is not fully implementable, it is natural to turn to second-best implementation objectives and to ask how "close it is to being implementable". In this section we present several such objectives.

3.1 Overimplementation and minimal overimplementation

If the correspondence $F$ is not fully implementable, one appealing requirement on a game form is that the equilibrium set always contains the set of $F$-optimal outcomes. Guaranteeing that all of the desirable allocations be included, which could be called "overimplementation", might be a reasonable compromise. Of course, one should be careful not to include too many additional allocations. For instance, the correspondence that associates with each economy its feasible set, the feasibility correspondence, is trivially implementable, and of course it overimplements any correspondence that one may be interested in. But it does not provide a very satisfactory resolution of the problem posed by the fact that the correspondence may not be implementable. A more appealing resolution would keep the enlargement needed to obtain implementability at a minimum.

Luckily, there is a formal sense in which this can be done. Indeed, many of the conditions on correspondences that have been found relevant for implementation are closed under arbitrary intersections. For instance, for implementation in Nash equilibrium, the central conditions of Maskin's monotonicity and no veto power conditions are. The class of all correspondences
overimplementing a given correspondence is therefore non-empty, since it
contains the feasibility correspondence, and it has a minimal element, which
is the intersection of all of its members. This minimal implementable
extension of the correspondence that is implementable is of course fully
implementable.

In exchange economies in which agents have strictly monotonic preferences
with respect to a common good, monotonicity is a necessary and suffi-
cient condition for implementability, and the minimal implementable exten-
sion of a correspondence is therefore its minimal monotonic extension.

The notion of minimal monotonic extension was proposed and system-
atically studied in the context of abstract social choice by Sen (1987). In
concretely specified economic models, it has been the object of very little
attention although the following observations, taken from Thomson (1992)
might be useful. The first one pertains to “the constrained Walrasian corre-
spondence”, defined by Hurwicz, Maskin and Postlewaite (1983). It differs
from the usual Walrasian correspondence in that agents are required to max-
imize their preferences in budget sets “truncated” by the aggregate feasibility
constraint. The Walrasian and constrained Walrasian correspondences
coincide in the interior of the feasible set, but boundary allocations can be
constrained Walrasian allocations without being Walrasian allocations. Now,
it is easy to show that the constrained Walrasian correspondence is the mini-
mal monotonic extension of the Walrasian correspondence. The “difference”
between the two correspondences is small, and for some interesting classes
of economies (economies for which boundary allocations can be eliminated
from considerations), the two actually coincide. However, for other corre-
spondences that are not implementable, a considerable enlargement may be
required in order to obtain implementability. This is illustrated on classical
domains by means of several examples discussed in the paper just cited
(Thomson, 1992). There, the minimal monotonic extensions of a number of
widely used correspondences are explicitly calculated, and general results re-
lating the minimal monotonic extensions of the union and the intersection of
two correspondences to the minimal monotonic extensions of the components
are established.4

4Cobb-Douglas preferences are an example.

5It should be noted that not all of the implementability conditions that have been
discovered are closed under intersection. For instance, the Moore and Repullo (1988)
condition, which is a necessary and sufficient condition on arbitrary domains, is not. This
3.2 Underimplementation and maximal underimplementation

Conversely, implementability conditions are often closed under union, so that in order to get as close as possible to full implementation, when full implementation cannot be achieved, it suffices to verify that there be at least one subcorrespondence of the correspondence of interest that is fully implementable. Then, the correspondence defined by taking the union of all the subcorrespondences that are implementable is the closest subcorrespondence of the correspondence aimed at. This correspondence can be said to maximally underimplement the desired correspondence. Here too, monotonicity can be used as an example since the property is closed under union, and on the domain of exchange economies in which preferences are strictly monotonic in a common good, the maximal implementable subcorrespondence of a given correspondence is its maximal monotonic subcorrespondence. This notion remains to be analyzed in detail.

3.3 Approximate implementation

A different definition is given by Matsushima (1988), and Abreu and Sen (1991). They propose looking for game forms such that for all $\epsilon > 0$, and for all profiles, the set of equilibrium outcomes of the induced game and the set of desired outcomes be $\epsilon$-close to each other.\(^6\) This notion of approximate implementation - they use the term of "virtual implementation" - is an appealing alternative to the concepts that we have discussed so far. Note that the proposal here is to approximate full implementation.

4 When partial implementation is a valid objective

In the previous section we have made a case for full implementation and discussed alternative notions for situations where full implementation is not

\(^6\)They derive conditions guaranteeing the existence of such game forms.
possible. In the final section, we will present some arguments for partial implementation.

4.1 Minimal partial implementation

Even when full implementation is thought to be the natural objective, it may be difficult to characterize the class of fully implementable correspondences. Then from the viewpoint of research strategy, turning to partial implementation makes good sense. In some contexts, it has been quite easy to show the possibility of partially implementing a correspondence, and in fact some general results have been obtained on this issue.

One of the most interesting theorems of the theory of implementation pertains to the class of upper-semicontinuous and fully implementable subcorrespondences of the correspondence that associates with each classical exchange economy its set of individually rational and efficient allocations: Hurwicz (1979) shows that any such correspondence contains the Walrasian correspondence. Under appropriate assumptions on the domain of preferences that are necessary to eliminate boundary allocations (see section 3.1), the Walrasian correspondence is fully implementable. This result is a theorem on what could be called minimal partial implementability. In situations where one may be concerned about agents being able to reach the equilibria, one could be interested in understanding which “small” subcorrespondences of the correspondence chosen by the planner are fully implementable. Hurwicz tells us that the class of fully implementable subcorrespondences of the individual rationality and Pareto correspondence has a smallest element, which is nothing other than the Walrasian correspondence.

Other results of this kind have been proved by Hurwicz (1979) for public good economies; there, the Lindahl correspondence plays the role played by the Walrasian correspondence in exchange economies. Thomson (1979) searched for fully implementable subcorrespondences of the no-envy and Pareto correspondence in exchange economies, and in one-commodity economies with single-peaked preferences preferences (1990). In these contexts, the equal income Walrasian correspondence and the uniform rule respectively are the counterparts of the Walrasian correspondence. The relevance of these last results is widened by Fleurbaey and Maniquet (1994)'s result that, under conditions on domains that are typically met, Maskin-monotonicity together with the very minimal condition of equal treatment of
equals imply no-envy.

4.2 Single-valued partial implementation

Instead of allowing partial implementation of the chosen correspondence, one could actually insist on it and go so far as to demand implementation of a single-valued selection. Indeed, as already noted, the actual success of the implementation will depend on agents actually reaching the desired equilibria. The coordination problems to be solved when equilibria are multiple, which are pervasive in game theory, are compounded here by the fact that, according to most equilibrium concepts that have been used in this literature, including Nash equilibrium and several of its refinements, it is hard to see how equilibrium will be reached unless agents know each others' preferences; only the planner is assumed to be in the dark. Some of the game forms that have been used actually require agents to announce whole preference profiles and at equilibrium, each agent is announcing the true profile! This may be too much to ask, as the critics of the implementation literature have repeatedly noted.\(^7\)

In defense of the theory, it should be pointed out that the theorems in which such announcements are required typically do much more than stating the existence of a game form implementing a correspondence if the correspondence satisfies a certain list of properties. They identify conditions guaranteeing implementability and they provide game forms achieving the implementation. In his pioneering paper, Maskin (1977) constructs a procedure that associates with each implementable correspondence a game form that implements it, and many of the subsequent contributions provide such constructions. It is not surprising that this enormous generality should have come at a price, both in terms of the complexity of the strategy spaces and of the informational demands imposed on agents in order to reach the equilibria. It is true that in subsequent work, these demands have been significantly decreased; for instance, McKelvey (1989) and Saijo (1988) have shown that game forms in which each agent announces an indifference set for himself and one for only one other agent would suffice. However, at equilibrium, agents are still truthfully announcing rather complex data. (This observation motivates the line of research discussed in the next section.)

\(^7\)See Postlewaite, 1985, for a discussion of these issues.
Single-valuedness of the correspondence certainly enhances the chance of equilibrium being reached and that is why this property may be desirable. Unfortunately, it is very restrictive. Indeed, if the set of alternatives is unstructured and the domain of possible preference includes all strict orderings, a social choice function whose range contains at least three alternative is implementable in Nash equilibrium only if it is in fact dictatorial (Muller and Satterthwaite, 1977). Under a certain assumption of richness of domains, a social choice function function is implementable in Nash equilibrium if and only if it is constant (Saijo, 1987). Fortunately these rather negative results have no counterpart on economic domains, where the set of feasible alternatives is endowed with additional structure and preferences are appropriately restricted. Moreover, as explained in the previous subsections, one can in some settings completely describe the “smallest implementable correspondences” satisfying some reasonable assumptions. This may be one reason why partial implementation may be not only an acceptable objective but in fact a desirable one.

5 Two other issues: natural implementation and multiple implementation

The recent work devoted to the characterization of correspondences that can be fully implemented by means of “simple” game forms constitutes an important development, and probably a necessary one if the theory is to be applied. In such games, strategies are points in finite dimensional euclidean spaces, interpretable as consumptions bundles, allocations and prices. The low dimensionality of strategy spaces should however not be an objective in itself since complex information can be “smuggled” in simple messages by means of cleverly defined outcome functions. On this point, see Hurwicz (1977), who was the first to recognize and analyze the possibility, and Chakravorti (1991) and Dutta, Sen and Vohra (1993) for additional discussion. These works show that the “simplicity” of a game form should be defined to also include requirements on the outcome function. In their work on the subject, Dutta, Sen and Vohra (1993), and Saijo, Tatamitani and Yamato (1993) have considered games in which agents announce consumptions, allocations and prices, and moreover, they have required that at equilibrium, the announced
consumptions be actually assigned to agents, and the announced prices be supporting prices at the resulting allocation. They identify conditions on correspondences that permit their implementation by such game forms. For early examples of simple game forms implementing the Walrasian and Lindahl correspondences, see Hurwicz (1979) and Walker (1981). For a recent example concerning the no-envy correspondence and several variants of it, see Thomson (1993), Maniquet (1994) and Suh (1994). A final contribution, concerned with the implementation of the core, is by Serrano and Vohra (1994).

Another recent literature has been concerned with the issue of multiple implementation: when the behavior of agents cannot be confidently described in terms of a single equilibrium concept, the search for a game form such that the set of outcomes chosen by the correspondence coincides with the sets of equilibria obtained for each one of several possible specifications of behavior is a natural objective. Most work on the subject pertains to implementation for two notions of equilibrium. This is referred to as double implementation, a concept introduced by Maskin (1977). The pairs that have been mainly considered are Nash and undominated Nash equilibria, and Nash and strong Nash equilibria. Simple game forms achieving double implementation of the so-called ratio correspondence in economies with public goods have been developed by Corchón and Wilkie (1991). These authors refer to their game forms as “market mechanisms”. Yamato (1990), Corchón and Ortúñor Ortín (1991), Suh (1994) and Maniquet (1994) are other contributors to the theory of double implementation.

The range of issues discussed earlier concerning the distinction between full and partial implementation are still relevant to the question of implementation by means of simple game forms, and to the question of implementation with respect to multiple equilibria.

6 Conclusion

We have investigated various notions of implementation and attempted to show that, unless one is concerned about the practical issue of achieving the equilibria (admittedly an important one), the notion of implementation that is worth studying is that of full, as opposed to partial, implementation. When a correspondence is not implementable, we have also suggested ways
of evaluating how close it is to having the property. Finally, we have briefly described new directions recently taken by the implementation literature.
7 References


—, E. Maskin and A. Postlewaite, "Implementation with unknown endowments", mimeo, 1983.

Maniquet, F., "Implementing the envy-free and the Pareto and envy-free correspondences with minimal symmetric strategy spaces", University of Namur mimeo, 1994.

Maskin, E. "Nash equilibrium and welfare optimality" M.I.T. mimeo 1977


McKelvey, R.D. "Game forms for Nash implementation of general social choice correspondences", *Social Choice and Welfare* 6 (1989), 139-


Suh, S-C., "Double implementation of the no-envy solution", University of Rochester mimeo, 1994.


——, "Manipulation and implementation in economies with single-peaked preferences", University of Rochester mimeo, 1990.

——, "Minimal monotonic extensions", University of Rochester mimeo, 1992.

——, "Divide and Permute" and the implementation of solutions to the problem of fair division", University of Rochester working paper no 360, 1993, forthcoming in *Games and Economic Behavior*. 

16