Implementation under Complete Information

Dirk Bergemann

10th February 2009
The goal of implementation theory is to study the relationship between outcomes in a society and the institutions under which those outcomes arise.

- The outcomes may include the production and allocation of public and private goods.
- By institutions, we mean the rules according to which the allocation is determined and enforced (property right, constitutions, contract law, voting rules, etc.)

In game theory, we study, analyze and predict the actions of individuals and the resulting outcomes (equilibrium) while taking game modeling interactions as given.

In implementation theory, our focus is on how to design the game itself. A typical question is: how can we construct a game form for which the strategic properties induce individuals to choose actions that lead to the desired outcomes.
A General Economic Environment

- $N$: set of agents: $i = 1, ..., n$
- $A$: set of outcomes (alternatives): $a \in A$
  - In a voting game, $N$ is the set of voters and $A$ is the set of finite candidates
  - In a market exchange model, $N$ is the set of economic agents and $A$ represents the final allocation of goods which are feasible given the endowments and production technology.
- $\mathcal{R}_A$: the classes of all orderings of the elements of $A$
- $R_i \in \mathcal{R}_i$: individual $i$’s preference (binary relation) over $A$, where $\mathcal{R}_i$ is a sub-class of $\mathcal{R}_A$
  - $aR_ib$, where $a, b \in A$, means that $i$ weakly prefers alternative $a$ to $b$.
  - Denote the preference profile by $R = (R_i, R_{-i}) \in \mathcal{R} \equiv \times_{j=1}^n \mathcal{R}_j$. 

Dirk Bergemann
Implementation under Complete Information
$F : \mathcal{R} \rightrightarrows A$ — social choice correspondence

- For any $R \in \mathcal{R}$, $F(R) \subseteq A$ represents the set of **socially desirable** (or $F$-optimal, in Maskin’s word) **alternatives** given the preference profile $R$.

- If $F$ is single-valued, we call it social choice function (SCF) and denote it by $f : \mathcal{R} \rightarrow A$

  - For any $R \in \mathcal{R}$, $f(R) \in A$. 

Dirk Bergemann

Implementation under Complete Information
A mechanism is denoted by a pair \( \langle M, g \rangle \)

\[ M = \times_{j=1}^{n} M_j, \text{ where } M_i \text{ is individual } i\text{’s message space} \]

\( g : M \rightarrow A \) is an outcome function

\( g (m) \in A \) represents the resulting outcome or allocation given the message profile \( m \).

\( s_i : R_i \rightarrow M_i \) is a pure strategy for individual \( i \)

Denote the strategy profile by \( s = (s_i, s_{-i}) \).

\( \mu_i : R_i \rightarrow \Delta (M_i) \) is a mixed strategy for individual \( i \) which assigns a probability \( \mu_i(s_i) \) to each pure strategy \( s_i \)

Denote the strategy profile by \( \mu = (\mu_i, \mu_{-i}) \).
Nash Implementation

Definition

A game form $g : M \rightarrow A$ implements the social choice correspondence $F$ in (pure strategy) Nash Equilibrium iff

(i) $\forall R \in \mathcal{R}, \forall a \in F(R)$, there exists a pure strategy $s$ such that $g(s) = a$ and $g(s) R_i g(s_i', s_{-i})$, $\forall s_i'$, $\forall i$ and

(ii) $\forall R \in \mathcal{R}$, if $\mu$ is a mixed strategy Nash equilibrium of $g$ with respect to $R$, then $g(s) \in F(R)$ for all realizations $s$ in the support of $\mu$. 
Part (i) states that any $F$-optimal alternatives can arise as a (pure strategy) Nash equilibrium.

Part (ii) is essentially the converse of part (i). In general, there can be multiple Nash equilibria of $g$. Without any reason to delete one or some of them, we cannot predict which of these will ultimately arise, part (ii) is necessary to ensure $F$-optimality of the outcome.
What can be derived from Nash Implementation? Or, what is necessary for Nash Implementation?

Fix a preference profile $R$ such that an alternative $a \in F(R)$. Suppose that $g$ Nash implements $F$, so there exists a $m \in M$ such that $m$ is a Nash Equilibrium at $R$ and $g(m) = a$.

Now consider another preference profile $\bar{R} \in \mathcal{R}$ such that $a \notin F(\bar{R})$.

$m$ cannot be a Nash equilibrium at $\bar{R}$, so someone prefers to deviate $\Rightarrow$ there must $\exists$ agent $i$ and a deviation $\bar{m}_i$ that $g(\bar{m}_i, m_{-i}) \mathcal{P}_i g(m)$.

Recall that $m$ is a Nash equilibrium at $R$ $\Rightarrow$ $g(m) R_i g(\bar{m}_i, m_{-i})$. 

Dirk Bergemann
Implementation under Complete Information
Letting \( b = g(\bar{m}_i, m_{-i}) \), we thus have:

**Definition**

A social choice correspondence \( F \) is **monotonic** if for \( \forall R, \bar{R} \in \mathcal{R} \), and \( a \in F(R) \) but \( a \notin F(\bar{R}) \), then there exists an agent \( i \) and \( b \in A \) such that \( aR_i b \) and \( b \mathcal{P}_i a \).

Conversely, if we cannot find an agent \( i \) and a strictly preferred alternative \( b \) under \( \bar{R} \), it must be the case that \( a \) is still socially desirable under \( \bar{R} \): \( a \in F(\bar{R}) \).

More precisely, suppose for \( \forall i \), we have \( g(m) R_i g(\bar{m}_i, m_{-i}) \) and \( g(m) \bar{R}_i g(\bar{m}_i, m_{-i}) \), then \( m \) is also a Nash equilibrium at \( \bar{R} \) and thus \( a \in F(\bar{R}) \).
Due to this logic, Maskin (1979, 1999) provides the contrapositive version of Monotonicity condition shown before:

Definition

A social choice correspondence $F$ is **monotonic** provided that $\forall a \in A, \forall R, \bar{R} \in \mathcal{R}$, if (i) $a \in F(R)$ and (ii) $\forall i \in N, \forall b \in A$, $aR_ib \Rightarrow a\bar{R}_ib$, then $a \in F(\bar{R})$.

In word, Monotonicity condition states that: if alternative $a$ is socially desirable w.r.t. some preference and if the profile is altered so that, in each individual’s ordering, $a$ does not fall below any alternative that it was not below before, then $a$ should still be socially desirable to the new profile.
Also, we have shown that monotonicity condition is necessary for Nash implementation:

**Theorem**

*If a social choice correspondence $F$ is Nash implementable, then $F$ is monotonic.*

Monotonicity alone is not sufficient to assure Nash implementation. Maskin (1999) constructs an example, well known as Borda Count Rule, that violates monotonicity.
In many applied economic environments, principal-agent literature on adverse selection in particular, a single-crossing property is often assumed.

To formulate, consider a standard one-principle-one-agent model. Let $t \in T \subset \mathbb{R}$ be the transfer from the principal to the agent, positive or negative, and $x \in X \subset \mathbb{R}$ denote the allocation, say, amount of good delivered from seller to buyer, probability of obtaining the item for an auctioner, etc.

Agent’s private information (type) is $\theta \in \Theta$, where his preference is characterized by a utility function $u : \Theta \times X \times T \rightarrow \mathbb{R}$, where $\partial u(\theta, x, t) / \partial t > 0$ and $\text{sgn} \{\partial u(\theta, x, t) / \partial x\}$ is fixed — positive everywhere or negative everywhere.
Now we introduce:

**Definition**

A utility function \( u(\theta, x, t) \) is called to satisfy the **single-crossing property** if for all \((x, t) \in X \times T\), taking \(\theta_1, \theta_2 \in \Theta\) such that \(\theta_1 < \theta_2\), then

\[
\frac{\partial u(\theta_1, x, t)}{\partial x} < \frac{\partial u(\theta_2, x, t)}{\partial x} < \frac{\partial u(\theta_1, x, t)}{\partial t} < \frac{\partial u(\theta_2, x, t)}{\partial t},
\]

or more generally, \(\frac{\partial u(\theta, x, t)}{\partial x} > \frac{\partial u(\theta, x, t)}{\partial t}\) is monotone in \(\theta\).

Since the marginal rate of substitution between \(x\) and \(t\), \(\frac{\partial u(\theta, x, t)}{\partial x} / \frac{\partial u(\theta, x, t)}{\partial t}\), specifies the slope of the indifference curve of utility function \(u(\cdot)\) on the \(x-t\) plane, the single-crossing property ensures that the two indifference curves associated with two different types could only cross once.
The connection between the single-crossing property and Maskin monotonicity can be summarized as

**Claim.** The single-crossing property implies that any SCF satisfies Maskin Monotonicity.

**Proof.**
A SCF $f$ satisfies Maskin monotonicity if for any $\theta_1, \theta_2 \in \Theta$, and $(x_1, t_1) = f(\theta_1) \neq f(\theta_2)$ (note that $f$ is univariant), there $\exists (x_2, t_2)$ such that $u(\theta_2, x_2, t_2) > u(\theta_2, x_1, t_1)$ and $u(\theta_1, x_1, t_1) \geq u(\theta_1, x_2, t_2)$ which is directly implied by the single-crossing property.
Monotonicity plus the following No Veto Power condition is sufficient for Nash implementation:

**Definition**

A social choice correspondence $F$ satisfies no veto power if for all $R \in \mathcal{R}$ and $a \in A$, whenever there exists $i \in N$ such that, for all $j \neq i$ and all $b \in A$, $aR_j b$, then $a \in F(R)$.

No Veto Power Condition says that if an alternative is at the top of $n - 1$ individuals’ preference orderings, the remaining agent cannot "veto" it, so this alternative should be socially desirable.
We now present the main theorem of this lecture

**Theorem**

If \( n \geq 3 \), a social choice correspondence \( F \) satisfies monotonicity and no veto power, then it is Nash implementable.
Consider the following mechanism.

Let \( M_i = R \times A \times \mathbb{N} \) where \( \mathbb{N} \) is the set of positive integers. Therefore, each individual announces a preference profile in \( R \) (not necessarily the true one), an alternative from \( A \) and an integer in \( \mathbb{N} \) (in order to break tie).

Now we construct a game form \( g : M \rightarrow A \) that implements SCC \( F \) in three steps.

(★) If \( \forall i, m_i = (R, a, k) \) and \( a \in F(R) \), take \( g (m_1, \ldots, m_n) = a \).

**Interpretation:** If all individuals' announcements are identical, then their mutually agreed alternative is chosen as the outcome.
(★★) If \( \exists i \), such that \( m_j = (R, a, k) \) for all \( j \neq i \), where \( a \in F(R) \), and \( m_i = (\cdot, b, \cdot) \neq m_j \), then take

\[
g(m_1, \ldots, m_n) = \begin{cases} 
  b & \text{if } aR_i b \\
  a & \text{if } bP_i a
\end{cases}
\]

Notice: (1) \( i \)'s preference that determines \( a \) or \( b \) is the one proposed by all the others instead of by individual \( i \); (2) \( a \) is chosen only when individual \( i \) strictly prefers \( b \) to \( a \) as proposed by all others.

**Interpretation**: If all but one individual makes exactly the same announcement (where alternative \( a \) is chosen by them), and the other announces differently (who chooses alternative \( b \)), the outcome is \( b \) if \( b \) is no better than \( a \) for \( i \) in all others' perspective of \( i \)'s preference. Otherwise, \( a \) is the outcome.
For any other $m$ where $m_i = (\cdot, a_i, k_i)$, let $i^*$ be the lowest indexed $i$ such that $k_i \geq k_j$ for all $j \neq i$, and then $g(m) = a_{i^*}$.

Notice: $i^* = \min \{i | k_i = \max_j k_j \}$. This is used for breaking ties when multiple individuals proposed the maximal number.

**Interpretation:** this case applies if neither (★) or (★★★) applies. The outcome is the alternative proposed by the player with the minimum index among those whose proposed number is maximal.
Based on the construction of an augmented mechanism \( g \), we now show that it Nash implements \( F \).

Let \( R \) be the true preference profile.

**Claim:** for any \( a \in F(R) \), it is a Nash equilibrium for all individuals to announce \( m_i = (R, a, 1) \).

- Check by a unilateral deviation. If individual \( j \) announces \( m_j \neq (R, a, 1) \), then situation falls in case (★★★) and \( j \) only gets \( b \) when \( aR_jb \), which is no better off if he announces \( (R, a, 1) \).

This establishes part (i) in the definition of Nash Implementation: there exists a Nash equilibrium of \( g \) corresponding to each socially desirable alternative.
We now check that every Nash equilibrium results in some \( a \in F(R) \).

Consider a Nash equilibrium in which everyone announces \( m_j = (R', a, k), \forall j \) and \( a \in F(R') \) where \( R' \neq R \) and \( R \) is the true preference profile. We are in the scenario (★) and \( a \) is the equilibrium outcome.

Since everyone reporting \((R', a, k)\) is a Nash equilibrium with respect to \( R \), no one (say, agent \( i \)) wants to make a unilateral deviation to achieve any \( b \) such that \( aR'_ib \), then \( aR_ib \).

If, instead, \( \exists i \) such that \( bP_ia \), \( i \) would rather report \((\cdot, b, \cdot)\) and this falls into the scenario (★★). Since \( aR'_ib \) as confirmed by anyone else’s report other than \( i \)’s, \( b \) is the equilibrium outcome as \( i \) expected.

So we get: for all \( i \) and all \( b \in A \), \( aR'_ib \Rightarrow aR_ib \).

By monotonicity, we must have \( a \in F(R') \Rightarrow a \in F(R) \), i.e., \( a \) is \( F \)-optimal.
Suppose $m$ is a Nash equilibrium for the true preference profile $R$ that falls into (★★★), i.e., for all $j \neq i$, 
$m_j = (R', a, k)$ and $m_i \neq m_j$, where $a \in F (R')$. Suppose finally $a'$ is chosen as the equilibrium outcome: $g (m) = a'$.

Any player $j \neq i$ can induce any alternative $a$ he wants by reporting a high enough number.

- Suppose, instead, agent $j$ announces $m'_j = (\cdot, a, k_j)$, where $k_j > \max_{l \neq j} k_l$. Now we are in the scenario (★★★★). By construction, his proposal will be adopted: $a$ is the equilibrium outcome.

- To guarantee that $a'$ is the equilibrium outcome under $R$, we must have $a' R_j b$ for all $j \neq i$ and $\forall b \in A$.

- This together with No-Veto Power condition lead to $a \in F (R)$.
Suppose $m$ is a Nash equilibrium for the true preference profile $R$ that falls into (★★★★), then $g(m) R_i a$, for $\forall i, \forall a \in A$.

Otherwise, if $\exists j$ and $b \in A$ such that $b P_j g(m)$, he could induce $b$ by announcing a number higher than anyone else’s.

By No-Veto Power condition, it must be that $g(m) \in F(R)$.

We thus have checked part (ii) in the definition of Nash Implementation. This completes the whole proof. ■

We just show the proof for pure strategies. We relegate the proof for mixed strategies to an exercise.

The mechanism shown in the proof is usually called the canonical mechanism for Nash implementation.
Actually, no-veto power condition is not necessary and might be restrictive.

How can we characterize fully the necessary and sufficient conditions for Nash implementation?

Also, we require $n \geq 3$. However, the case of $n = 2$ is also important as bilateral interactions are a typical theoretical topic.

Further topics include

- Solution refinement
- Virtual implementation
- Implementation in a sequential game
- Bayesian implementation under incomplete information
  
  . . . .

*Now this is not the end. It is not even the beginning of the end. But it is, perhaps, the end of the beginning.*

  —Sir Winston Churchill