Forcing in Strategic Belief Models

Fernando Tohmé   Gianluca Caterina   Rocco Gangle

Abstract

Forcing is a methodology for building models of Set Theory satisfying certain properties. Since its inception by Paul Cohen, in the early 1960s, it has been applied to several areas in Mathematical Logic, becoming a powerful tool in the analysis of axiomatic systems. In this paper we extend the applicability of forcing to game-theoretic strategic belief models. In particular, we propose a very general notion of solutions for such games by enlarging Brandenburger's $RmAR$ condition via extension through generic types.

1 Introduction

The methodology of forcing was introduced into Mathematics by Paul Cohen in order to show that Georg Cantor's famous Continuum Hypothesis is independent of the axioms of Zermelo-Frenkel Set Theory [10] [14]. This success prompted other set theorists to investigate other topics in the field with the aid of this powerful tool. Connections with other parts of Mathematical Logics were readily found and versions of forcing for Model Theory were developed at the end of the 1960s [3].

Forcing has remained in the realm of the foundations of Mathematics, without being adopted in applied fields. The reason can be found in Shoenfield's Theorem, from which it can be deduced that forcing yields results only in the non-absolute fragment of Mathematics, while most of applied science seems to be confined in the absolute realm [12]. Only two recent pieces of research dared to go beyond this limit, in Design Theory ([11]) and Abduction Theory ([9]). In the latter, forcing is seen as providing the formal basis for diagrammatic reasoning, embodied in Peirce's γ-graphs. The intuition behind them seems to extend to any belief formation process without defined boundaries.

A field in which the ideas of [9] might be applied is the characterization of *types* of players in games. While the conditions for the existence of complete types spaces are fairly well known, we are interested in providing definite features to the types that ensure the epistemic conditions for very general notions of solution in games. This can be accomplished, we claim, by means of a straightforward application of forcing.

In section 2 we present a conceptual discussion of Cohen's variety of forcing and how it allows to reason, from the point of view of a conceptual framework, about generic objects in it and to provide a characterization of them, even if they are indiscernable from inside the framework. In section 3 we make these ideas concrete by introducing the problem of defining *generic types* in games and apply forcing to define them.

2 Reasoning and Forcing

We make use of a mathematical technique called forcing. The method of forcing has become a standard tool in several branches of mathematics, notably model theory. Its primary purpose is to enact a specific form of mathematical reasoning, one that lends itself remarkably well, in particular, to proofs of independence results. It does this by providing a way to generate and at least partially control arbitrary models of set theory or other axiomatic systems with chosen properties, even when these models are non-constructible by normal means such as recursion or transfinite induction on ordinals.

Given a model M, we can define an extension of such a model by adjoining a set G to M, and denoting the new model by $M[G]$. The nature of set G, which we will call a *generic set* is such that, even being definable from within M, it is *indiscernable* from M. By this we mean that the language within M allows us only to *name* the elements of G, but not explicitly to describe its construction. In this way, we do not have access to the inner structure of G, which remains unknowable from the point of view of M and hence the use of the word *generic* (as referring to the expression of something so "mixed up" or "common" that it cannot be discerned). Once the generic G has been defined, the *extension* via G of the ground model M gives way to new and possibly surprising ways to satisfy the ground axioms, with profound epistemic consequences. Indeed, although *truths* in $M[G]$ are not directly accessible, we can define what we call a forcing relation between objects and relations at the level of the ground model. If one object forces a certain relation on M, then, if that relation belongs to G (and we might never be able to know that except as a modal claim across possible models), then we obtain "truth" in $M[G]$.

Cohen's original result with respect to the independence of the Generalized Continuum Hypothesis from the Zermelo-Fraenkel axioms of set theory is well known, and details may be found in [10] and [16]. In the wake of Cohen, a variety of other interesting results have been obtained. We apply the method to game-theoretic strategic belief models in the following section.

Smullyan and Fitting [16] elaborate an approach to forcing based in modal logic (specifically the standard system S4) and the Kripkean semantics of *frames*. A related approach, but one expressed in the diagrammatic logic of Peirce's EG-γ, may be found in [9]. These modal-logic based approaches to forcing emphasize the non-classical character of the reasoning forcing enacts, which is closely linked to the concept of the *generic* central to forcing. In what follows, this modal character of the rationality of forcing is introduced into game-theoretic strategic beliefs models by defining a generic set over the ordinal hierarchy of rational ascription beliefs of players' type-profiles. Models of games may then be forced which suggest new, non-classical rationalities. These motivating intuitions are formalized in the subsequent section.

3 Generic Types

Game Theory is the field that studies the strategic interaction among self-interested parties. That is, situations in which the outcome depends on the decisions made by several agents, who intend to maximize their respective payoffs [15]. These situations are defined as follows:

Definition **3.1** Let $\mathcal{G} = \langle I, \{S_i\}_{i\in I}, \{U_i\}_{i\in I}\rangle$ be a *game*, where $I = \{1, \ldots, n\}$ is a set of *players* and $S_i, i \in I$ is a finite set of *strategies* for each player. A *profile* of strategies, $s = (s_1, \ldots, s_n)$ is an element of $S = \prod_{i\in I} S_i$. In turn, $U_i : S \to \mathbb{R}$ is player i's payoff.

The goal is to assess the *solutions* of the game, i.e. the family of profiles $\mathcal{S} \subseteq S$ that might be expected to be chosen by the players. These profiles capture the rationality of players, seeking to maximize their payoffs. Given any $i \in I$, we denote by (s_i, s_{-i}), where $s_{-i} \in \prod_{j\neq i} S_j$, a joint profile of actions. All the aspects that contribute to the coordination among agents, which must remain implicit since no communication is allowed between players, have an *epistemic* nature. That is, they ensue solely from the beliefs and knowledge of the individuals.

A *Strategic Beliefs Model* captures the epistemic aspects involved in the choice of strategies [6]:

$$\mathbf{M} = (\{S_i\}_{i\in I}, \{T_i\}_{i\in I}, \{P_i\}_{i\in I})$$

where for each $i \in I$, S_i and T_i are i's sets of strategies and *types*, respectively. The structure intends to model a game $\mathcal{G}$ and each strategy-type pair is a *state* for a player, and each type of a player has beliefs about the states of the other players. These beliefs are captured by the relations P_i that satisfy:

1. $P_i : T_i \to S_{-i} \times T_{-i}$ is a correspondence.
2. For all $t_i \in T_i$, $P_i[t_i] \neq \emptyset$.

That is, $P_i[t_i]$ captures the strategies and types of the other players that i thinks are possible, and each t_i involves a non-empty set of beliefs.

The analysis on the rationality of players and the epistemic conditions of solutions to the game can be fully disclosed up from the *states of the game*, i.e. profiles of states of the players. The following example shows the expressive power of $\mathbf{M}$ [4]:

Example **3.1** *Let $\mathcal{G}$ be a two-player game, with $S_1 = \{A, B\}$ and $S_2 = \{I, D\}$:*

	I	D
A	$2,2$	$0,0$
B	$0,0$	$1,1$

To analyze this game we add:

- *A class of types T_i for each i, $T_1 = \{t^a, u^a\}$ and $T_2 = \{t^b, u^b\}$.*
- *A correspondence $P_i[\cdot] : T_i \to S_{-i} \times T_{-i}$, for each i.*

Suppose that $P_1[\cdot]$ and $P_2[\cdot]$ are as follows (the Xs indicate which pairs (s_{-i}, t_{-i}) belong to the range of $P_i[\cdot]$):

$P_1[t^a]$:

T_2/S_2	I	D
$\mathbf{u^b}$	0	X
$\mathbf{t^b}$	0	X

$P_1[u^a]$:

T_2/S_2	I	D
$\mathbf{u^b}$	X	0
$\mathbf{t^b}$	0	X

$P_2[t^b]$:

T_1/S_1	A	B
$\mathbf{u^a}$	0	X
$\mathbf{t^a}$	0	X

$P_2[u^b]$:

T_1/S_1	A	B
$\mathbf{u^a}$	X	0
$\mathbf{t^a}$	0	X

A state *of the game is* $(s_i, t_i; s_{-i}, t_{-i})$*. Let us consider state* $(B, t^a; D, t^b)$*:*

- *The response of* 1 *is "correct", since she considers* D *is the choice of* 2*. That is, the best she can do is choose* B*.*
- *The response of* 2 *is also right, since he considers that* B *is the choice of* 1 *and then his best response is* D*.*

But:

- 1 *considers possible that* 2 *may be mistaken about her choice:* 1 *thinks that* 2 *might be of type* u^b *while* u^b *considers possible that* 1 *may choose* A *instead of* B*.*
- 2 *thinks that* 1 *can be mistaken about his choice. Since* 1 *considers possible* u^a*, which implies that* 2 *thinks that* 1 *may play* I *instead of* D*.*

This means that 1 *and* 2 *are* rational*, since they maximize their payoffs given their beliefs. But* 1 *thinks that* 2 could *be irrational since she considers possible* (D, u^b)*, while if* 2 *were rational and had type* u^b*, he would get a higher payoff with* I *than with* D*. Analogously,* 2 *thinks that* 1 *might be irrational since he considers possible* (B, u^a)*, and at* u^a*,* 1 *fares better with* A *than with* B*.*

This example shows interesting features of $\mathbf{M}$. The first one is that each t_i can be "unfolded" in terms of the types of the other players, which in turn lead to beliefs about the type of i, etc.

To make this notion more precise, let us define for any t_i of i, the *unfolding* of t_i:

- $t_j \in P_i^1[t_i]$ if there exists $s_j \in S_j$ with[1] $\langle (s_{-ij}, t_{-ij}), (s_j, t_j) \rangle \in P_i[t_i]$.
- $t_j \in P_i^m[t_i]$, for any natural number m, if there exists $t_k \in P_i^{(m-1)}[t_i]$ such that $t_j \in P_k^1[t_k]$.

This means that, if $t_j \in P_i^m[t_i]$, t_i can be unfolded in m steps to t_j, i.e. t_j is believed by t_i by considering m steps of belief.

Another important feature of $\mathbf{M}$ is that it provides a powerful framework for describing the epistemic aspects involved in a game [1]. The fundamental concept here is that of *assumption* as defined over *events* of $\mathbf{M}$, i.e. on sets of states of the game. For any $E \subseteq \prod_i (S_i \times T_i)$, the types of i that assume E are denoted as $\mathcal{AS}_i[E]$ with:[2]

$$\mathcal{AS}_i[E] = \{t_i \in T_i : P_i[t_i] = E_{|\prod_{j\neq i}(S_j \times T_j)}\}$$

[1] Here s_{-ij} (t_{-ij}) denotes an element in $\prod_{k\neq i, k\neq j} S_k$ ($\prod_{k\neq i, k\neq j} T_k$).

[2] The notation $E_{|\prod_{j\neq i}(S_j \times T_j)\}}$ indicates the projection of E, defined over $\prod_i (S_i \times T_i)$ over $\prod_{j\neq i}(S_j \times T_j)$.

In particular, we say that at t_i, i assumes that $j \neq i$ is rational if $t_i \in \mathcal{AS}_i[j \text{ is rational}]$, where the event "$j$ is rational" is

$$\{\langle (s_j, t_j), (s_{-j}, t_{-j})\rangle \in \prod_i (S_i \times T_i) : U_j(s_j, s_{-j}) \geq U_j(s, s_{-j}) \text{ for any } s \in S_j\}$$

We may then define inductively the condition denoted $RmAR$ (for *Rationality and m-Assumption of Rationality*):[3]

- $R_i^0 = [i \text{ is rational}]_{|S_i \times T_i}$.
- $R_i^m = R_i^{m-1} \cap (S_i \times \mathcal{AS}_i[\bigcap_{j \neq i} R_j^{m-1}])$.

With all these elements at hand, it is known that there are certain properties that a particular $\mathbf{M}$ might fail to satisfy, in particular completeness [6]. Whether such properties hold is related to whether enough types exist to ensure the existence of solutions to a game. The question is if these properties can be imposed on $\mathbf{M}$ in a general way, without finiteness or countability restrictions on m. Here is where the technique of *forcing* can be applied.

Let us start with a given $\mathbf{M}_0$ intended as a family of events of the game plus their underlying states of the game. It will constitute our ground model, on the basis of which a new model $\mathbf{M}$ with the desired properties will be built. We then define a family of *forcing conditions* $\mathcal{P}$ with a partial order $\preceq$ defined as follows:

- $\mathcal{P}$= $\{\pi = \langle (s_1, t_1), \ldots, (s_n, t_n)\rangle$: every i is rational and there exists a natural number $m(\pi)$ such that for any i, j, $t_j \in P_i^{m(\pi)}[t_i]\}$.
- For any $\pi, \pi^{'} \in \mathcal{P}$, each one defined by a natural number ($m(\pi)$ and $m(\pi^{'})$), $\pi^{'} \preceq \pi$ iff $m(\pi) \geq m(\pi^{'})$.

We say that if $\pi^{'} \preceq \pi$, then π *dominates* $\pi^{'}$.

Let us define now a *correct set* δ of forcing conditions. A set δ is said to be *correct* if and only if it satisfies the properties of a *filter* in $(\mathcal{P}, \preceq)$:

- If $\pi^{'} \in \delta$ and $\pi^{'} \preceq \pi$ then $\pi \in \delta$.
- If $\pi^{'}, \pi^{''} \in \delta$ there exists $\pi \in \delta$ such that $\pi^{'} \preceq \pi$ and $\pi^{''} \preceq \pi$.

Our candidate is $\delta = \{\phi = \langle (s_1, t_1), \ldots, (s_n, t_n)\rangle$: there exists a natural number $m(\phi)$ such that for any i, $(s_i, t_i) \in R_i^{m(\phi)}\}$. We have that:

Proposition 3.1 δ is a correct set in $\mathcal{P}$.

[3]If $E^i \subseteq S_i \times T_i$, $\mathcal{AS}_i[E^i]$ is a shorthand for $\mathcal{AS}_i[\bigcap_{E: E_{|S_i \times T_i} = E^i} E]$.

Proof:*Let us see first that* $\delta \subseteq \mathcal{P}$. *That is, that for every* $\phi \in \delta$ *there exists* $\pi \in \mathcal{P}$ *such that* $\phi = \pi$. *We know that* $\phi=\langle (s_1, t_1), \ldots, (s_n, t_n)\rangle$ *where for some* $m \geq 0$ *every* i *is such that,* $(s_i, t_i) \in R_i^m$. *From this condition follows that each* i *is rational. We have to see now that for every* i, j, $t_j \in P_i^m[t_i]$. *Suppose not. Then, there exists a* t_k *such that* $t_k \notin P_i^{m'}[t_i]$, *with* $m' \leq m$. *This means that there exists* $t_l \in P_i^{(m'-1)}[t_i]$ *such that* $(s_k, t_k) \notin P_l[t_l]$. *But, on the other hand,* $(s_l, t_l) \in R_l^m$. *We have by definition that* $(s_l, t_l) \in S_l \times \mathcal{AS}_l[R_k^0]$ *which means that* t_l *is such that* $(s_k, t_k) \in P_l[t_l]$. *Contradiction.*

The converse is also true: given a state $\langle (s_1, t_1), \ldots, (s_n, t_n)\rangle \in \mathcal{P}$, *it follows that there exists* m *such that each* $(s_i, t_i) \in R_i^m$. *Suppose not. Then, for a pair* i, j, $t_j \notin \mathcal{AS}_j[R_i^{m'}]$ *for some* $m' \leq m$. *Then* $P_j[t_j]_{|S_i \times T_i} \neq R_i^{m'}$. *In particular, we have that* $t_i \notin P_j^{m'}[t_j]$. *Contradiction.*

From this last implication it follows that if $\pi' \in \delta$ *and* $\pi' \preceq \pi$ *then* $\pi \in \delta$. *This is because* $m(\pi') \leq m(\pi)$ *and* π *is such that every* i *is rational and for every pair* i, j, $t_j \in P_i^{m(\pi)}$, *which in turn implies that every* $(s_i, t_i) \in R_i^{m(\pi)}$ *and therefore,* $\pi \in \delta$.

Finally, given $\pi', \pi'' \in \delta$, *just take* m *as the maximum of* $m(\pi')$ *and* $m(\pi'')$. *Without loss of generality let us assume that* $m = m(\pi')$. *Then, we take* $\pi = \pi'$ *and it is easy to see that* $\pi' \preceq \pi$ *and* $\pi'' \preceq \pi$.

We can define now a class of conditions called *dominations*. D is a domination if and only if $D \subseteq \mathcal{P}$ is *dense* in $\mathcal{P}$:

$$\forall \pi' \in \mathcal{P}\ \exists \pi \in D \text{ such that } \pi' \preceq \pi$$

Then, a correct set G is said to be *generic* if $G \subseteq \delta$ and $G \cap D \neq \emptyset$ for any domination D. We have that:

THEOREM **3.1** $G = \{\phi \in \delta : \phi = \langle (s_1, t_1), \ldots, (s_n, t_n)\rangle$ with for every m and every i, $(s_i, t_i) \in R_i^m\}$ is a generic set.

Proof:*By Proposition 3.1,* δ *is a correct set and so is* $G \subset \delta$. *To see that it is generic, just consider any* $\pi \in \mathcal{P}$, *which is identified by a finite natural number* $m(\pi)$. *Then, by definition,* $\phi \in G$ *is such that for every* i, $(s_i, t_i) \in R_i^m$ *for every* m, *in particular with* $m \geq m(\pi)$. *Then,* $\pi \preceq \phi$.

We can say that G defines a set of types $\{t_i^*\}_{i \in I}$ such that each one, joint with the corresponding s_i^*, satisfies that $(s_i^*, t_i^*) \in R_i^m$ for *every* $m \geq 0$. That is, each i is, with her type and the correponding strategy, rational and assumes rationality at all levels. In other words, it satisfies the condition called $R\infty AR$. These *generic types* cannot be defined in the language of $\mathbf{M_0}$. That is, there is no property λ expressible in $\mathbf{M_0}$ such that:[4]

$$\forall t_i \in T_i\ \lambda(t_i) \Leftrightarrow \exists \pi \in G \text{ such that } \exists s_i \in S_i\ (s_i, t_i) \in R_i^{m(\pi)}$$

[4]The generic types are *indiscernable* in $\mathbf{M_0}$ [9].

This realization is quite important since, as shown in [6], $\mathbf{M_0}$ is *not definable complete*, i.e. there exists some event $E \in \mathbf{M_0}$, definable by a property λ_E (i.e. $\langle(s_1,t_1),\ldots,(s_n,t_n)\rangle \in E \Leftrightarrow \lambda_E(\langle\langle(s_1,t_1),\ldots,(s_n,t_n)\rangle\rangle)$) such that there exists an i for whom no $t_i \in T_i$ satisfies $P_i[t_i] = E_{|\prod_{j\neq i}(S_j\times T_j)}$.[5]

However forcing shows that $\mathbf{M_0}$ can be extended to $\mathbf{M_0}[G]$, in which the class of generic types defined by G is included. To define $\mathbf{M_0}[G]$ consider the *names* of objects in G. The G-names are recursively defined sets of the form $\{(\mu,\pi) : \mu \text{ is a } G-\text{name and } \pi \in G\}$. They can be ordered in terms of their rank. A name μ of rank 0 is the set of pairs $(\emptyset,\pi)$ with $m(\pi)=0$. Recursively, we say that μ is of rank m, if it includes all the pairs $(\mu^{'},\pi)$ such that $m(\pi)=m$ and the rank of $\mu^{'}$, $m^{'}$, verifies $m^{'} < m$.

The *referential value* of a name μ, $r_G(\mu)$ is also defined recursively:

- If the rank of μ is 0, $r_G(\mu) = \{\langle(s_1,t_1),\ldots,(s_n,t_n)\rangle \in \prod_i(S_i\times T_i) : (s_i,t_i)\in R_i^0\}$ iff there exists $(\emptyset,\pi)\in\mu$. It is $r_G(\mu) = \mathbf{M_0}[G]$, otherwise.
- If the rank of μ is m, $r_G(\mu) = \{r_G(\mu^{'}) : \exists(\mu^{'},\pi)\in\mu\}$.

It is easy to see that names of rank 0 yield all the states in which all players are rational, while for any $m > 0$ the names have as referential values all the states in which the players are rational and assume up to level m the rationality of all the others.

Then, $\mathbf{M_0}[G] = \{r_G(\mu) : \mu \text{ is a name in } \mathbf{M_0}\}$. We have:

Proposition **3.2** $\mathbf{M_0} \subset \mathbf{M_0}[G]$.

Proof: *Trivial. Just take any name μ such that for every $\pi \in G$, $(\emptyset,\pi)\notin\mu$. It is easy to find an event in $\mathbf{M_0}$ satisfying this condition: take anyone in which the states are such that there exist i and j, with $(s_i,t_i)\in R_i^{m_i}$ and $(s_j,t_j)\in R_j^{m_j}$ and $m_i \neq m_j$.*

Instead, the object defined by G, namely the class of states in which each $(s_i,t_i)\in R\infty AR$, exists in $\mathbf{M_0}[G]$. If we consider a statement $\Gamma(\langle\langle(s_1,t_1),\ldots,(s_n,t_n)\rangle\rangle)$ which is true iff each $(s_i,t_i)\in R\infty AR$, we know that:

$$\mathbf{M}_0[G] \models \Gamma(\langle\langle(s_1,t_1),\ldots,(s_n,t_n)\rangle\rangle).$$

Using a well-known result proven in [2], [13] and [8] among others, we also know that if μ is a name in $\mathbf{M_0}$ such that $\langle(s_1,t_1),\ldots,(s_n,t_n)\rangle \in r_G(\mu)$ we have that $\pi\in G$ is such that:

$$\pi \Vdash \Gamma(\mu)$$

i.e. the generic types force $R\infty AR$.

In this sense, we have defined a broad class of types involved in a very general notion of solution for games. This class of types yields, when S_i is finite, those

[5] See also [17].

profiles of strategies that obtain by iterated elimination of weakly dominated strategies. This is because $R\infty AR$ has been shown to be the epistemic precondition for a *SAS* (self-admissible set) [7], [5].

References

[1] Aumann, R. and Brandenburger, A.: Epistemic Conditions for Nash Equilibrium, *Econometrica* **63**: 1161-1180, 1995.

[2] Burgess, J.: Forcing, Chap.B4 in J. Barwise (ed.) **Handbook of Mathematical Logic**, North-Holland, 1977.

[3] Barwise, J. and Robinson, A.: Completing Theories by Forcing, *Annals of Mathematical Logic* **2**: 119–142, 1970.

[4] Brandenburger, A.: The Power of Paradox: Some Recent Developments in Interactive Epistemology, *International Journal of Game Theory* **35**: 465–492, 2007.

[5] Brandenburger, A. and Friedenberg, A.: Self-Admissible Sets, *Journal of Economic Theory* **145**: 785–811, 2009.

[6] Brandenburger, A. and Keisler, J.: An Impossibility Theorem on Beliefs in Games, *Studia Logica* **84**: 211-240, 2006.

[7] Brandenburger, A. - Friedenberg, A. - Keisler, J.: Admissibility in Games, *Econometrica* **76**: 307–352, 2008.

[8] Chow, T.: A Beginners Guide to Forcing, *Contemporary Mathematics* **479**: 25–40, 2009.

[9] Caterina, G. and Gangle, R.: Consequences of a Diagrammatic Representation of Paul Cohens Forcing Technique Based on C.S. Peirces Existential Graphs , *Studies in Computational Intelligence* **314**: 429–443, 2010.

[10] Cohen, P.: **Set Theory and the Continuum Hypothesis**, Addison-Wesley, 1966.

[11] Hatchuel, A. and Weil, B.: Design as Forcing: Deepening the Foundations of CK Theory, *Proceedings of the 16th International Conference on Engineering Design.* (`http://www.cgs.ensmp.fr/design/docs/Hatchuel-Weil-ICED07paperVdef21.pdf.`), 2007.

[12] Jech, T.: What is Forcing?, *Notices of the AMS* **55**: 692-693, 2008.

[13] Jech, T.: **Set Theory**, Springer Verlag, 2002.

[14] Kanamori, A.: Cohen and Set Theory, *Bulletin of Symbolic Logic* **14**: 351–378, 2008.

[15] Osborne, M. and Rubinstein, A.: **A Course in Game Theory**, MIT Press, 1994.

[16] Smullyan, R. and Fitting, M.: **Set Theory and the Continuum Problem**, Clarendon Press, Oxford, 1996.

[17] Tohmé, F.: Existence and Definability of States of the World, *Mathematical Social Sciences* **49**: 81-100, 2005.