Games with Communication: from Belief to Preference Change

Guillaume Aucher[1], Bastien Maubert[2], Sophie Pinchinat[3], and François Schwarzentruber[4]

[1] IRISA - INRIA / Université de Rennes 1, Rennes, France.
guillaume.aucher@inria.fr
[2] LORIA - CNRS / Université de Lorraine, Nancy, France.
bastien.maubert@gmail.com
[3] IRISA / Université de Rennes 1, Rennes, France.
sophie.pinchinat@irisa.fr
[4] IRISA - ENS Rennes, Rennes, France.
francois.schwarzentruber@ens-rennes.fr

Abstract. In this work we consider simple extensive-form games with two players, Player **A** and Player **B**, where Player **B** can make announcements about his strategy. Player **A** has then to revise her preferences about her strategies, so as to better respond to the strategy she believes Player **B** will play. We propose a generic framework that combines methods and techniques from belief revision theory and social choice theory to address this problem. Additionally, we design a logic that Player **A** can use to reason and decide how to play in such games.

1   Introduction

Communication between players is a notion that arises naturally in a variety of contexts in game theory, and that led to the theory of games where players can communicate [4,5,11]. We are interested in non-cooperative games with two players, say Player **A** and **B**, in which Player **B** makes announcements about his strategy, before the game starts. Just as the *cheap talks* in [4], this preliminary communication round does not directly affect the payoffs of the game.

We illustrate our research problem with a classic example from [11] in which communication between players improves the payoff of both players. The extensive form game is described in Figure 1. Player **A** can go left or right. If **A** goes left, she gets 1\$ and **B** gets 0\$. If **A** goes right, player **B** can in turn choose to go left or right. If **B** goes left, he gets 100\$ and **A** gets 0\$, if **B** goes right both get 99\$. The solution given by the classic backward induction algorithm, which relies on the hypothesis that players are rational, is the following: **A** thinks that if she goes right, **B** will go left to maximize his payoff, and **A** will get 0\$. Therefore, **A** prefers to move left, and gets 1\$.

On the other hand, let us assume that the players communicate and trust each other, and that **B** tells **A**: "If you move right, I will move right". As a
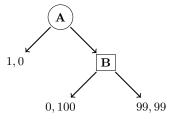
Fig. 1: Motivating example

consequence, **A** thinks she would better move right since she would collect 99\$ instead of 1\$: as such, **A** has revised her preferences about her own strategies.

Notice that in this example, **B**'s announcement could have been reflected by pruning the game, in the spirit of Public Announcement Logic [9]: we could have removed the moves (in the example, just one) of **B** that do not conform to his announcement, in this very case by ruling out his left move, and have recomputed a strategy of **A** by backward induction in the pruned game.

However, the pruning technique, although attractive in practice, has some serious limitations. First, we cannot guarantee that in any game, every announcement of **B** amounts to pruning the game, in particular those relying on conditional statements. Second, **B** can make a series of successive announcements, possibly conflicting each other. In that case, **A** will need to aggregate these announcements in order to revise her beliefs on what **B** will play. This phenomenon cannot be represented straightforwardly by means of a series of destructive prunings of the game, and we propose to work on the level of **B**'s strategies instead.

Preliminary announcements can be motivated by various reasons, such as trying to coordinate with the other player or to mislead him in order to get a better payoff. After these announcements, Player **A** needs to revise her strategy so as to better respond to what Player **B** pretends to play. Notice that depending on the context, the confidence Player **A** has on Player **B**'s commitment about his annoucements varies widely. In this work, like in belief revision theory [6], we assume that Player **A** always trusts Player **B**'s last announcement, which has also priority over the previous announcements.

The question we consider is the following:

> How can Player **A** take into account the announcements of Player **B** about his strategy in order to update her preferences on her strategies?

This question can be decomposed into:

> **Question 1:** How can Player **A** revise her beliefs about Player **B**'s preferences on his strategies?
> **Question 2:** How can Player **A** update her preference about her strategies on the basis of these beliefs?

Regarding Question 1, we propose to apply classical belief-revision techniques[5] to represent what **A** believes about **B**'s strategy and update these beliefs when **B** makes announcements. There exist several ways to perform this update/revision, but our approach aims at remaining as general as possible by not selecting a particular one, and by leaving the choice to peak the update mechanism that reflects how trustworthy **B**'s announcements are considered.

The main originality of our contribution lies in the solution we offer for Question 2, by combining techniques and methods from game theory and from social choice theory [2]: informally, each possible strategy of **B** is seen as a voter, who votes for strategies of **A** according to the payoff **A** would obtain in the play defined by both strategies. Individual votes are then aggregated to define the new preferred strategy of **A**. Here again we do not choose a particular type of ballot nor a precise aggregation method, but rather leave it open and free to be set according to the kind of strategy one wants to obtain: for instance, one that has best average payoff against **B**'s most plausible strategies, or one that is most often a best response.

The paper is organized as follows. In Section 2, we set up the mathematical framework we use to model games and communication/announcements. In Section 3, we develop the solution to the revision of beliefs, and in Section 4 we expose our proposal for the revision of preferences. Based on the developed setting, we propose in Section 5 a logic that Player **A** can use to reason and decide how to play. Section 6 illustrates our framework on a more complex example.

2 Games and announcements

We consider two-player extensive-form games in which at each decision node two distinct moves are available. A *finite rooted binary tree* (simply called *tree* from now on) is a prefix-closed finite set $T \subset \{0,1\}^*$. Elements of T are called *nodes*, ϵ is the *root*, if $w \cdot a \in T$, with $a \in \{0,1\}$, then w is called the *parent* of $w \cdot a$ and $w \cdot a$ is called the *left* (resp. *right*) *child* of w if $a = 0$ (resp. $a = 1$). If a node has no child, it is a *leaf*, otherwise it is an *interior node*. A tree is called *complete* if every interior node has exactly two children. If T, T' are trees such that $T \subseteq T'$, we say that T is a *subtree* of T'.

A *game* between **A** and **B** is a tuple $G = (T, \nu_\mathbf{A}, \nu_\mathbf{B})$ where T is a complete tree, and if we note $L \subseteq T$ the set of leaves of T, then $\nu_\mathbf{A} : L \to \mathbb{N}$ is the *utility function* for **A**, $\nu_\mathbf{B} : L \to \mathbb{N}$ is the utility function for **B**. Interior nodes are partitioned between nodes of **A** ($N_\mathbf{A}$) and those of **B** ($N_\mathbf{B}$), such that $T = N_\mathbf{A} \uplus N_\mathbf{B} \uplus L$.

Given a game $G = (T, \nu_\mathbf{A}, \nu_\mathbf{B})$, a *strategy*[6] for **A** (resp. **B**) is a subtree $\sigma_\mathbf{A}$ (resp. $\sigma_\mathbf{B}$) of T such that every node in $\sigma_\mathbf{A} \cap N_\mathbf{A}$ (resp. $\sigma_\mathbf{B} \cap N_\mathbf{B}$) has exactly

---

[5] Typically, **A** initially believes that **B** will play one of the strategies given by the classical backward-induction algorithm. Then **B** may announce a piece of information that is in contradiction with this belief, which thus needs to be revised.

[6] To be precise these are reduced strategies, but they are sufficient for what we present here.

one child, and every node in $\sigma_\mathbf{A} \cap N_\mathbf{B}$ (resp. $\sigma_\mathbf{B} \cap N_\mathbf{A}$) has exactly two children. Two strategies $\sigma_\mathbf{A}$ and $\sigma_\mathbf{B}$ define a unique path, hence a unique leaf in the tree T, that we shall write $\sigma_\mathbf{A}\hat{\ }\sigma_\mathbf{B}$. We note $\Sigma_\mathbf{A}$ and $\Sigma_\mathbf{B}$ the set of all strategies for $\mathbf{A}$ and $\mathbf{B}$, respectively.

For a strategy $\sigma_\mathbf{A} \in \Sigma_\mathbf{A}$, we define its *value* $\mathrm{val}(\sigma_\mathbf{A})$ as the minimum utility it can bring about for $\mathbf{A}$: $\mathrm{val}(\sigma_\mathbf{A}) := \min_{w \in L} \nu_\mathbf{A}(w)$. The value of a strategy for Player $\mathbf{B}$ is defined likewise.

The language Player $\mathbf{B}$ uses to make the announcements about his strategies is the bimodal language $\mathcal{L}_2$, the syntax of which is:

$$\psi ::= p \mid \neg\psi \mid \psi \wedge \psi \mid \Diamond_i \psi$$

where $p \in \{turn_\mathbf{A}, turn_\mathbf{B}\}$ and $i \in \{0, 1\}$.

For $i \in \{0, 1\}$, we write $\top$ for $\neg(p \wedge \neg p)$, $\Box_i \psi$ for $\neg\Diamond_i\neg\psi$, $\Box\varphi$ for $\Box_0\varphi \wedge \Box_1\varphi$, and $move_i$ for $\Diamond_i\top$, meaning that the strategy at this point chooses direction i.

Example 1. For instance, in the example of Figure 1, the strategy of $\mathbf{B}$ consisting in playing the action leading to 99, 99 is $\Diamond_1\Diamond_1\top$.

Given a game $G = (T, \nu_\mathbf{A}, \nu_\mathbf{B})$, a strategy σ can be seen as a Kripke structure with two relations (one for left child, one for right child). The valuations of propositions $turn_\mathbf{A}$ and $turn_\mathbf{B}$ are given by the partition between positions of Player $\mathbf{A}$ and Player $\mathbf{B}$. Formally, the truth conditions are defined inductively as follows:

$$
\begin{aligned}
\sigma, w &\models turn_a & &\text{if } w \in N_a,\ a \in \{\mathbf{A}, \mathbf{B}\} \\
\sigma, w &\models \neg\psi & &\text{if } \sigma, w \not\models \psi \\
\sigma, w &\models \psi \wedge \psi' & &\text{if } \sigma, w \models \psi \text{ and } \sigma, w \models \psi' \\
\sigma, w &\models \Diamond_i\psi & &\text{if } w \cdot i \in \sigma \text{ and } \sigma, w \cdot i \models \psi
\end{aligned}
$$

3 Belief revision: from announcements to beliefs

We now represent the beliefs $\mathbf{A}$ has about what $\mathbf{B}$ is more likely to play, and how these beliefs evolve as $\mathbf{B}$ makes new announcements.

From a purely semantic point of view, the framework of belief revision theory [1,7] can be roughly described as follows. Given a universe $\mathcal{U}$ of possible worlds, a player ranks each possible world via a *ranking function* $\kappa : \mathcal{U} \to \mathbb{N}$, also called *belief state*, such that $\kappa^{-1}(0) \neq \emptyset$. This ranking induces a plausibility preorder between possible worlds: among two possible worlds, the one with the lowest rank is considered to be more plausible than the other by the player. Given a ranking function κ, the set of *most plausible worlds* for the player is the set $\kappa^{-1}(0)$.

The impact of a new piece of information on these beliefs is modelled by a *revision function* which takes a ranking function together with the new information, and returns the revised ranking function that induces the new belief state of the player. Many such revision functions exist in the literature, that correspond amongst other things to various degrees in the trust put in the received information, the reluctance to modify one's beliefs, etc (see *e.g.* [10]). Formally, if one

chooses say formulas of propositional logic PL to represent new pieces of information, a revision function is a binary function $* : (\mathcal{U} \to \mathbb{N}) \times PL \to (\mathcal{U} \to \mathbb{N})$, and given $F \in PL$, a belief state κ is changed into $\kappa * F$.

In our framework, the universe $\mathcal{U} = \Sigma_{\mathbf{B}}$ is the set of Player $\mathbf{B}$'s strategies, and the new pieces of information are modal formulas of $\mathcal{L}_2$, representing $\mathbf{B}$'s announcements about his strategy. For a belief state κ, $\kappa^{-1}(0)$ is then what $\mathbf{A}$ believes $\mathbf{B}$ is the most likely to play. Initially, we assume that $\mathbf{A}$ has an *a priori* belief, represented by κ_0, that may for example arise from the very values of the strategies:

$$\kappa_0(\sigma_{\mathbf{B}}) := \max_{\sigma'_{\mathbf{B}} \in \Sigma_{\mathbf{B}}} \mathrm{val}(\sigma'_{\mathbf{B}}) - \mathrm{val}(\sigma_{\mathbf{B}}) \tag{1}$$

The revision function signature is now $(\Sigma_{\mathbf{B}} \to \mathbb{N}) \times \mathcal{L}_2 \to (\Sigma_{\mathbf{B}} \to \mathbb{N})$, and we can use any kind of revision function. For example here, we present the classic *moderate revision* [8,10], written $*_m$, and defined by: for $\kappa, \psi \in \mathcal{L}_2$ and $\sigma \in \Sigma_{\mathbf{B}}$,

$$(\kappa *_m \psi)(\sigma) = \begin{cases} \kappa(\sigma) - \min_{\sigma' \models \psi} \kappa(\sigma') & \text{if } \sigma \models \psi \\ \max_{\sigma' \models \psi} \kappa(\sigma') + 1 + \kappa(\sigma) \\ - \min_{\sigma' \not\models \psi} \kappa(\sigma') & \text{if } \sigma \not\models \psi \end{cases}$$

The moderate revision makes all the possible worlds that verify the announcement ψ more believed than those which do not; it preserves the original order of preference otherwise.

4   Voting: from beliefs to preferences

The belief Player $\mathbf{A}$ has about $\mathbf{B}$'s strategy induces some preference over $\mathbf{A}$'s strategies. We describe a mechanism that, given a belief state κ, computes a *preference set* $\mathcal{P}_\kappa \subseteq \Sigma_{\mathbf{A}}$. This preference set is made of all the strategies that should be preferred by $\mathbf{A}$ if she believes that $\mathbf{B}$ will play a strategy in $\kappa^{-1}(0)$. This mechanism relies on voting systems.

A plethora of different voting systems have been proposed and studied [3], verifying different properties one may want a voting system to verify (majority criterion, Condorcet criterion etc). Since we are interested in quantitative outcomes, we argue that a relevant choice is to use a *cardinal voting system* [12]. In a cardinal voting system, a voter gives each candidate a *rating* from a set of grades; we take here grades in $\mathbb{N}$. Take a set of n candidates, $C = \{c_1, \ldots, c_n\}$, and a set of m voters, $V = \{v_1, \ldots, v_m\}$. A *ballot* is a mapping $b : C \to \mathbb{N}$ and a *voting correspondence* is a function $r^C : (C \to \mathbb{N})^m \to 2^C \backslash \{\emptyset\}$ that takes a vector $(b_1, b_2, \ldots, b_m)$ of ballots (one for each voter) and returns a nonempty set of *winning* candidates[7]. In this work we take as an example the *range voting system*, but the method is generic and any other cardinal voting system can be used. Range voting works as follows: for each candidate, we sum the grades obtained in the different ballots, and the set of winners is the set of candidates

---

[7] It is called a voting rule if there is a unique winner.

who share the highest overall score: if b_i is voter i's ballot, for $i \in \{1, \ldots, m\}$, r^C is defined by

$$r^C(b_1, \ldots, b_m) := \underset{c \in C}{\mathrm{argmax}} \sum_{i=1}^{m} b_i(c).$$

We aim at electing the strategies of Player $\mathbf{A}$ that she should prefer with regard to the most plausible strategies of Player $\mathbf{B}$. Therefore, the set of candidates consists in Player $\mathbf{A}$'s possible strategies ($C = \Sigma_\mathbf{A}$), and each of Player $\mathbf{B}$'s most plausible strategie is seen as a voter ($V = \kappa^{-1}(0)$). We assume that Player $\mathbf{A}$ prefers strategies that in average give her the best payoff, which leads us to define ballots as follows. For each strategy $\sigma_\mathbf{B} \in \kappa^{-1}(0)$, we let $b_{\sigma_\mathbf{B}}$ be the ballot that assigns to each $\sigma_\mathbf{A} \in \Sigma_\mathbf{A}$ the payoff of $\mathbf{A}$ in the play $\sigma_\mathbf{A} \hat{} \sigma_\mathbf{B}$, that is $b_{\sigma_\mathbf{B}}(\sigma_\mathbf{A}) = \nu_\mathbf{A}(\sigma_\mathbf{A} \hat{} \sigma_\mathbf{B})$. In other words, each voter ranks the candidates according to the corresponding payoff for Player $\mathbf{A}$. The voting system aggregates these "individual" preferences in order to obtain a "collective" preference $\mathcal{P}_\kappa$ against all strategies of $\kappa^{-1}(0)$, defined by:

$$\mathcal{P}_\kappa := r^C(b_{\sigma_\mathbf{B}^1}, \ldots, b_{\sigma_\mathbf{B}^m}), \text{ whenever } \kappa^{-1}(0) = \{\sigma_\mathbf{B}^1, \ldots, \sigma_\mathbf{B}^m\}.$$

Remark 1. Note that we could use more of the information we have by letting all strategies in $\Sigma_\mathbf{B}$ vote, and weigh their votes according to their respective plausibility.

5   A logic for strategies, announcements and preferences

We present the formal language $\mathcal{L}_{SAP}$, where SAP stands for "Strategies, Announcements and Preferences", to reason about Player $\mathbf{A}$'s preferences concerning her strategies, and how these evolve while Player $\mathbf{B}$ makes announcements about his strategy. The syntax of $\mathcal{L}_{SAP}$ is the following:

$$\varphi ::= \psi \mid \neg \varphi \mid \varphi \wedge \varphi \mid P_\mathbf{A} \varphi \mid [\psi!]\varphi$$

where $\psi \in \mathcal{L}_2$.

The formula $P_\mathbf{A}\varphi$ reads as 'φ holds in all the preferred strategies of Player $\mathbf{A}$'; $[\psi!]\varphi$ reads as 'φ holds after Player $\mathbf{B}$ announces that her strategy satisfies ψ'.

$\mathcal{L}_{SAP}$ formulas are evaluated in models of the form $(\kappa, \sigma_\mathbf{A})$, where κ is the *belief state* of Player $\mathbf{A}$ and $\sigma_\mathbf{A} \in \Sigma_\mathbf{A}$ is the strategy $\mathbf{A}$ is considering. The truth conditions are given inductively as follows:

$$
\begin{aligned}
(\kappa, \sigma_\mathbf{A}) &\models \psi && \text{if } (\sigma_\mathbf{A}, \epsilon) \models \psi \\
(\kappa, \sigma_\mathbf{A}) &\models \neg \varphi && \text{if } (\kappa, \sigma_\mathbf{A}) \not\models \varphi \\
(\kappa, \sigma_\mathbf{A}) &\models \varphi \wedge \varphi' && \text{if } (\kappa, \sigma_\mathbf{A}) \models \varphi \text{ and } (\kappa, \sigma_\mathbf{A}) \models \varphi' \\
(\kappa, \sigma_\mathbf{A}) &\models P_\mathbf{A}\varphi && \text{if for all } \sigma'_\mathbf{A} \in \mathcal{P}_\kappa, (\kappa, \sigma'_\mathbf{A}) \models \varphi \\
(\kappa, \sigma_\mathbf{A}) &\models [\psi!]\varphi && \text{if } (\kappa *_m \psi, \sigma_\mathbf{A}) \models \varphi
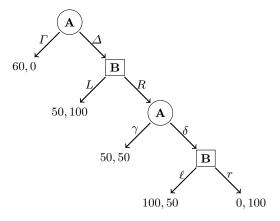\end{aligned}
$$

Fig. 2: Second example game

6 Example

Consider the game in Figure 2. By backward induction, we get that $\mathbf{B}$ chooses r, $\mathbf{A}$ thus chooses γ, $\mathbf{B}$ chooses L, and finally $\mathbf{A}$ chooses Γ, obtaining 60\$ while $\mathbf{B}$ gets nothing. $\mathbf{B}$ would therefore like $\mathbf{A}$ to change her mind and play Δ on the first move, so that he can play L and get 100. The problem is that if he announces that he will do so, then $\mathbf{A}$ will stick to her strategy, as she will know that changing it will give her a payoff of 50 instead of 60. So $\mathbf{B}$ announces, instead, that he commits to play either L, or R and then ℓ (we note this strategy $R\ell$), but not Rr. This announcement can be described by the following $\mathcal{L}_2$-formula:

$$\psi = \Box(turn_{\mathbf{B}} \rightarrow move_0) \vee \Box\Box\Box(turn_{\mathbf{B}} \rightarrow move_0)$$

Consider now the following $\mathcal{L}_{SAP}$-formula:

$$\varphi = turn_{\mathbf{A}} \wedge P_{\mathbf{A}}move_0 \wedge [\psi!]P_{\mathbf{A}}move_1$$

φ expresses that it is Player $\mathbf{A}$'s turn to play, and that in all her preferred strategies she goes left (*i.e.* she plays Γ), but in case Player $\mathbf{B}$ announces ψ, Player $\mathbf{A}$ prefers to play differently, namely moving right.

Now, considering this game, moderate revision, range voting, with the initial belief ranking κ_0 of Equation (1) on Page 5, and any strategy $\sigma_{\mathbf{A}} \in \Sigma_{\mathbf{A}}$, one can check that indeed we have:

$$(\kappa_0, \sigma_{\mathbf{A}}) \models \varphi$$

This is because going right ensures $\mathbf{A}$ a better mean-payoff against $\mathbf{B}$'s most plausible strategies after the announcement ψ, which are L and Rl. However, consider now the classic plurality voting system, where each voter only gives one voice to its preffered candidate (here, the one that ensures $\mathbf{A}$ the best outcome), and where the winner is the one with most votes for him. This amounts to

electing **A**'s strategy that is most often a best response against **B**'s most plausible strategies. Using this instead of range voting system, one can verify that after the announcement, the vote results into a tie, with strategy Γ of **A** obtaining one vote (from **B**'s strategy L), and strategy $\Delta\delta$ receiving the other one (from strategy Rl). Therefore, $P_{\mathbf{A}} move_1$ does not hold in the state resulting from the announcement, so that we have:

$$(\kappa_0, \sigma_{\mathbf{A}}) \not\models \varphi$$

7 Conclusion

Our work contributes to the study of games with communication. We have defined a generic framework that uses belief revision techniques to take into account communication, and voting for choosing strategies to play. A specific revision function and voting system may characterize the behavior of Player **A** (trustful, optimistic, etc), and the kind of strategies she wants (best mean payoff, most often best-response...). Investigating the theoretical properties of the agent's behavior in terms of combinations of revision and voting mechanisms is left for future work.

References

1. Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *J. Symb. Log.*, 50(2):510–530, 1985.
2. K.J. Arrow and A.K. Sen. *Handbook of social choice and welfare*, volume 19. North Holland, 2002.
3. S.J. Brams and P.C. Fishburn. Voting procedures. *Handbook of social choice and welfare*, 1:173–236, 2002.
4. Vincent P Crawford and Joel Sobel. Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pages 1431–1451, 1982.
5. Joseph Farrell and Matthew Rabin. Cheap talk. *The Journal of Economic Perspectives*, pages 103–118, 1996.
6. Peter Gärdenfors. *Knowledge in Flux (Modeling the Dynamics of Epistemic States)*. Bradford/MIT Press, Cambridge, Massachusetts, 1988.
7. Peter Gärdenfors and Hans Rott. Belief revision. In *Handbook of Logic in Artificial Intelligence and Logic Programming*, volume 4, pages 35–132. 1995.
8. A.C. Nayak. Iterated belief change based on epistemic entrenchment. *Erkenntnis*, 41(3):353–390, 1994.
9. Jan Plaza. Logics of public communications. *Synthese*, 158(2):165–179, 2007.
10. H. Rott. Shifting priorities: Simple representations for twenty-seven iterated theory change operators. *Towards mathematical philosophy*, pages 269–296, 2009.
11. J. van Benthem. *Logical dynamics of information and interaction*. Cambridge University Press, 2011.
12. S. Vasiljev. Cardinal voting: the way to escape the social choice impossibility. *Available at SSRN 1116545*, 2008.