Outline:

1. Nash Equilibria in Games in normal form

- Normal Form games
- Multiplayer game for synthesis

2. Nash equilibria in infinite games

- (Motivating example)
- Definition of Multiplayer games on automata
- Reduction to a two player game
- Some extensions of Nash equilibria

3. Admissible strategies

- In normal form games
- Simple safety games
- Parity games
- Iterated eliminiation

4. Related Works

Todo:

- Example of Multiplayer game on automata (2.1)
- Example of admissibility in normal form game (3.1)
- Definition iterated elimination (3.4)
- Property of fixpoint for iterated elimination (3.5)
- Section about other results (4)

Introduction

In two player games seen so far, players had objectives that are opposite to each other's, so we were able to define them giving only Eve's objective. Adam was seen as a purely adversarial agent. Such games are called *zero-sum* games since, in a quantitative setting, the sum of the payoffs of the two players would sum up to zero in any outcome. However, the objectives of the players are not entirely conflictual in all games. In particular, multiplayer games, that is, games with more than two players, cannot be zero-sum by definition; there are also interesting examples of non-zero sum games with only two players (we will see one below). In this setting, winning strategies are no longer suitable to describe rational behaviors since the opponents should no longer be seen as purely adversarial. In fact, when the objectives of the players are not opposite, some cooperation becomes often possible. Then, rather than assuming opponents as purely adversarial, it is interesting to study the possible outcomes when they are simply *rational*, that is, follow the best strategy for their own objectives.

The notion of equilibria we will study in this chapter aims at describing such rational behaviors.

If one is expecting for sure a specific strategy to be played by the opponent, then the most rational response is to choose the *best response*, that is, the strategy that is optimal for the player against the specific strategy of the adversary. Thus, if we assign strategies to players, and if the players are all aware of the strategies of the other players, then each player will be willing to change their strategy if theirs does not turn out to be a best response. Such a situation is seen as unstable and is undesirable in many applications of game theory. *Nash equilibrium* is defined simply as a stable situation in such a setting: a strategy profile in which the strategy of each player is a best response to the rest of the strategies. Thus, no player has any incentive to change their strategy.

We will see the formal definition of a Nash equilibrium in the next section. Let us first consider the following example.

The following Hawk-Dove game was first presented by the biologists Smith and Price, and shown in Table ... *(add citation in the biblio section?). (We could also concider directly medium-access control here).* Here, two animals are fighting over some prey and can choose to either act as a hawk or as a dove. If a player chooses hawk then the best payoff for the opponent is obtained by choosing dove. In fact, ... (pourquoi? il faut expliquer ici l'idee du jeu) So, dove is the best response to hawk. Reciprocally, the best response to dove is to play hawk. There are two "equilibria": (Hawk, Dove) and (Dove, Hawk), where no player has an interest in changing their strategy. Note however that none of the players has a winning strategy.

Nash showed the existence of such equilibria in any normal-form game (citation?), which may require randomized strategies. This result revolutionized the field of economics, where it is used to analyze competitions between firms or government economic policies for example. Game theory and the concept of Nash equilibrium are now applied to diverse fields: in finance to analyze the evolution of market prices, in biology to understand the evolution of some species, in political sciences to explain public choices made by parties.

Table 1: The Hawk-Dove game. Each column corresponds to a strategy of P_1 and each line to a strategy of P_2.

	Hawk	Dove
Hawk	0 , 0	1 , 4
Dove	4 , 1	3 , 3

In this chapter, we are going to study the computation of Nash equilibria in multiplayer concurrent games with omega-regular objectives. The algorithms we present here differ from those that were given for normal-form games since ours are infinite-duration with omega-regular objectives. We will then present extensions of this notion such as secure and robust equilibria. The second result we develop is the notion of admissibility: this is a different approach to the study of rational behaviors and consist in eliminating for each player irrational choices of strategies.

Nash Equilibria in Games in Normal Form

Let us introduce the notions we will study in normal-form games.

Pas mal de définitions vont être héritées du chapitre concurrent

Normal form games

Definition: A normal-form game is given by a set of players $Players$, a set of strategies $Strat_P$ for each player $P \in Players$ and a payoff function $\texttt{payoff}_P : \prod_{P \in Players} Strat_P \to \mathbb{R}$.

The goal of each player is to maximize their payoff, and the payoff of each player can depend on other players' strategies. Note that we are in a concurrent setting, so each player chooses their strategy independently.

A *strategy profile* is a map assigning a strategy to each player. We simply denote it as a tuple $(\sigma_1, \sigma_2, \dots, \sigma_n) \in \prod_{p \in P} Strat_p$. Such a tuple will often be written σ_P, and for a subset of players $A \subseteq P$, σ_A denotes the strategy profile for the players. The pair~$(\sigma_A, \sigma_{P \setminus A})$ then refers to the full strategy profile.

Example: In the Hawk-Dove game, there are two players P_1 and P_2. Possible strategies are $Strat_{P_1} = Strat_{P_2} = \{Hawk, Dove\}$. The payoffs corresponding to different strategy profiles are given in Table *(reference to table)*, $\texttt{payoff}_{P_1}(Hawk, Hawk) = 0$, $\texttt{payoff}_{P_1}(Hawk, Dove) = 4$, etc…

Nash equilibrium and Existence Theorem:

A *Nash equilibrium* is a stable situation in the sense that no player has an interest in changing its strategy. Nash proved that when player are allowed to randomise among all there strategies, there always exists a Nash equilibrium.

A *randomised* strategy is a probability distributing over the actions, i.e. $\forall a \in Act.\ \sigma_P(a) \in [0,1]$ and $\sum_{a \in Act} \sigma_P(a) = 1$. The resulting payoff of a strategy profile is then $\sum_{a_{Agt} \in Act^{Agt}} \texttt{payoff}(a_{Agt}) \cdot \prod_{A \in Agt} \sigma_A(a_A)$.

probablement sera déjà défini

A strategy is said to be *pure* if the support of its distribution is a singleton. A pure strategy profile is a profile consisting of pure strategies. A randomised strategy profile is a general strategy profile possibly containing randomised strategies.

A *pure (resp. randomised) Nash equilibrium* is pure (resp. randomised) strategy profile σ_P such that for all players $A_i \in P$, and all strategies σ'_{A_i}, $\texttt{payoff}(\sigma'_{A_i}, \sigma_{P \setminus A_i}) \leq \texttt{payoff}(\sigma_{A_i}, \sigma_{P \setminus A_i})$.

Theorem: In every normal-form game with a finite number of players, each having a finite number of pure strategies, there exists a randomised Nash equilibrium.

Note that not all games contain pure Nash equilibria. For example, ...

Algorithm: *Should we present an algorithm in polynomial time to find pure Nash equilibria?*

Why not. Is it simple?

or present a Reduction to linear programming for mixed strategies

Add small paragraph on LP techniques to find mixed equilibria

Action-graph games

Action-graphs are succinct representation of matrix games. Indeed, representing games with matrices can be costly when the number of players increases. The size of the matrix is in fact exponential in the number of players: when each player has two strategies there are $2^{|Players|}$ cells in the table. In this compact representation, the algorithm is no longer polynomial because the representation can be exponentially smaller. Computer scientist wanted to show that this was indeed a difficult problem. Usally we do this by showing that the corresponding decision problems are NP-hard. However existence of Nash equilibrium cannot be made into a decision problem, because they always exist (the answer is always true so the problem is trivial). To characterize the complexity of this problem, the PPAD complexity class was created, and it can be shown that finding a Nash equilibrium in an action-graph game is a PPAD complete problem.

J'ai l'impression que ce paragrahphe est un peu trop rapide et n'apporte pas grand chose au lecteur. Ca n'a

Multiplayer Games for Synthesis

This part can be dropped as this is a theoretical book, and the example is a "fakely" applied one O: I like this example. It's simple enough to look like a fake application

Medium Access Control: Consider a medium access control problem, where several users share access to a wireless channel. A communication over the channel is successful if thera are no collisions, that is, if a single user is transmitting its message only. During each slot, each user chooses either to transmit or to idle. Intuitively, the probability that a user is successful in its transmission decreases with the number of users emitting in the same slot. (Attention, on n'a pas de proba ici) Furthermore each attempt at transmitting has a cost or reward depending on whether the tranmission is succesful. The expected reward for one slot and two players, is represented in Table(reference) assuming a cost of 2 for each transmission, a reward of 4 for a successful transmission, a probability 1 to be successful if only one player emit, and of 1 if they both transmit at the same time. (Pourquoi est-ce qu'on parle de proba?)

Table 2: A game of medium access.

	Emit	Wait
Emit	-1, -1	2, 0
Wait	0 , 2	0, 0

The game described above corresponds to a single slot of this system. In reality, there would be a succession of slots and the payoff would be the sum of payoffs for each slot. Normal-form games are thus not sufficient to represent games with repetitions and to study the evolution of the behaviors as the game evolves.

One possibilty to model repetition is to use *games in extensive form* which are games played on finite trees. However such games only model a fixed number of repetitions unlike infinite or arbitrary duration games as studied in this book. We thus study, in the rest of this chapter, algorithms for games played on graphs.

Exercice: What are/is the Nash equilibrium/a in the medium access games?

Nash Equilibria in Omega-Regular Games

Definition

To be adapted later according to chapter 8

A multiplayer arena $\mathcal{A}$ is a tuple $V, Agt, Act, Tab, (c_A)_{A \in Agt_i}$, where:

- V is a finite set of vertices;
- Agt is a finite set of players;
- Act is a finite set of actions, a tuple $(a_A)_{A \in Agt}$ containing one action a_A for each player $A \in Agt$ is called a move, thus Act^{Agt} is the set of possible moves;
- $Tab : V \times Act^{Agt} \to V$ is the transition function, it associates with a given vertex and a given move, the resulting state;
- $(c_A)_{A \in Agt}$ is a tuple of coloring function, one for each player which define its objectives.

Thus, a multiplayer arena is simply a concurrent arena with multiple players.

Example *INSERT simple and abstract example*

Note: Add figure

History and plays: A history of the multiplayer arena $\mathcal{A}$ is a finite sequence of states and moves ending with a state, i.e. an word in $(V \cdot Act^{Agt}) \cdot V$. Note that unlike for two player games we include actions in the history, because knowing the source and target vertices does not mean you know which player chose what actions.

We write h_i the i-th vertex of h, starting from 0, and $move_i(h)$ its i-th move, thus $h = h_0 \cdot move_0(h) \cdot h_1 \cdots move_{n-1}(h) \cdot h_n$. The length |h| of such a history is n + 1. We write last(h) the last vertex of h, i.e. $h_{|h|-1}$. A play ρ is an infinite sequence of vertices and moves, i.e. an element of $(V \cdot Act^{Agt})^{\omega}$.

Defs: Define coalition, notation C, -C. Notation $(move)_A$.

Outcomes: Let C be a coalition, i.e. a subset of the players in Agt, and a strategy σ_C for C is a function which associates a strategy σ_A to each player $A \in C$. Given a strategy A, when it is clear from the context, we simply write σ_A for $\sigma_C(A)$ A history h is compatible with the strategy σ_C if, for all $k < |h|-1$ and all $A \in C$, $(move_k(h))_A = \sigma_A(h {\leq}\, k)$, and $Tab(h_k, move_k(h)) = h_{k+1}$. A play ρ is compatible with the strategy σ_C if all its prefixes are. We write $Out_A(v_0, \sigma_C)$ for the set of plays in $\mathcal{A}$ that are compatible with strategy σ_C and have initial vertex v_0. These paths are called *outcomes* of σ_C from v_0. We write $Out_{\mathcal{A}}$ the set of plays that are compatible with some strategy σ_{Agt} of Agt. Note that when the coalition C is composed of all the players (and the strategies are deterministic *Note: I don't know whether this is assumed on not*) the outcome is unique.

Algorithm for Finding Nash Equilibria

We will now present an algorithm to compute Nash equilibria in multiplayer games.

The problem we are interested in is to decide the existence of a Nash equilibrium in which the objectives of a given set of players are satisfied.

Problem: Given a multiplayer game $(G, (\Omega_p)_{p \in Agt})$, and $A \subset Agt$ decide if there exists a Nash equilibrium σ_{Agt} such that for all $p \in A$, ...

Is this the problem we solve? Do we want to use quantitative payoffs?

The algorithm is based on a reduction to zero-sum two-players games, which allows us to use algorithms presented in the previous chapters of this book. More precisely, we present the *deviator game*, which is a transformation of a concurrent multiplayer game into a turn-based zero-sum game, such that there are strong links between equilibria in the first one and winning strategies in the second one. The proofs of this section are independent of the type of objectives we consider.

Deviators

A central notion we use is that of *deviators*. These are the players who have played different moves from those prescribed in a given profile, thus causing a deviation from the expected outcome. Formally, a deviator from move a_{Agt} to $(a')_{Agt}$ is a player $D \in Agt$ such that $a_D \neq a'_D$. We denote the set of deviators by

$$Dev(a_{Agt}, a'_{Agt}) = \{A \in Agt \mid a_A \neq a'_A\}.$$

We extend the definition to pairs of histories and strategies by taking the union of deviator sets of each step along the history. Formally,

$$Dev(h, \sigma_{Agt}) = \bigcup_{0 \leq i < |h|} Dev(move_i(h), \sigma_{Agt}(h_{\leq i})).$$

For a play ρ, we define $Dev(\rho, \sigma_{Agt}) = \bigcup_{i \in \mathbb{N}} Dev(move_i(\rho), \sigma_{Agt}(\rho_{\leq i}))$. Intuitively, having chosen a strategy profile σ_{Agt} and observed a play ρ, deviators represent the agents that must have changed their strategies from σ_{Agt} in order to generate ρ.

Note that Nash equilibria are defined only with respect to deviations by single players, so the case of deviators exceeding size 2 (two players simultaneously changing their strategies) should be of no matter to us. %Indeed, we will define the objective of Eve so that it is satisfied in that case.

Lemma Given a play ρ, coalition C contains $Dev(\rho)$, if and only if, there exists a strategy σ'_C such that $Out(\sigma_{-C}, \sigma'_C) = \rho$.

Proof Assume that coalition C contains $Dev(\rho, \sigma_{Agt})$. We define the strategy σ_C to be such that for all $i \in \mathbb{N}$, $\sigma_C(\rho_{\leq i}) = (move_i(\rho))_C$. By hypothesis, we have, for all indices i, $Dev(move_i(\rho), \sigma_{Agt}(\rho_{\leq i})) \subseteq C$, so for all agents $A \notin C$, $\sigma_A(\rho_{\leq i}) = (move_i(\rho))_A$. Then $Tab(\rho_i, \sigma'_C(\rho_{\leq i}), \sigma_{-C}(\rho_{\leq i})) = \rho_{i+1}$. Hence ρ is the outcome of the profile (σ_{-C}, σ'_C).

For the other direction, let σ_{Agt} be a strategy profile, σ'_C a strategy for coalition C, and $\rho \in Out_G(\rho_0, \sigma_{-C}, \sigma'_C)$. We have for all indices i that $move_i(\rho) = (\sigma_{-C}(\rho_{\leq i}), \sigma'_C(\rho_{\leq i}))$. Therefore for all agents $A \notin C$, $(move_i(\rho))_A = \sigma_A(\rho_{\leq i})$. Then $Dev(move_i(\rho), \sigma_{Agt}(\rho_{\leq i})) \subseteq C$. Hence $Dev(\rho, \sigma_{Agt}) \subseteq C$.

Deviator Game

We now use the notion of deviators to draw a link between multiplayer games and a two-player game. Given a *game* $G = (\mathcal{A}, v_0)$, we define the deviator game $D(G)$. $\boxed{\text{Game is arena + objective}}$ Intuitively, Eve needs to play according to an equilibrium, while Adam tries to find a profitable deviation for any player. The vertices are $V' = V \times 2^{Agt}$, where the second component, a subset of Agt, records the deviators of the current history.

At each step, Eve chooses an action profile, and Adam chooses the move that will apply. Adam can either respect Eve's choice, or pick a different action profile in which case the deviators will be added to the second component of the vertex. The game begins in $(v_0, \varnothing)$ and then proceeds as follows: from a vertex (s, D), Eve chooses an action profile a_{Agt}, and Adam chooses another one a'_{Agt} which can be equal. The next vertex is $(Tab(s, a'_{Agt}), D \cup Dev(a_{Agt}, a'_{Agt}))$.

We define projections π_V and π_{Dev} from from V' to V and from V' to 2^{Agt} respectively. As well as π_{Act} from $Act^{Agt} \times Act^{Agt}$ to Act^{Agt} which maps to the second component of the product, that is, Adam's action.

For a history or play ρ, define $\pi_{Out}(\rho)$ as a play ρ' such that for all index i, $\rho'_i = \pi_V(\rho_i)$ and $move_i(\rho') = \pi_{Act}(move_i(\rho))$. This is thus the play induced by Adam's actions.

We can associate a strategy of Eve in to each strategy profile σ_{Agt} such that she chooses the moves prescribed by σ_{Agt} at each history of $D(G)$. Formally, we write $\kappa(\sigma_{Agt})$ for the strategy defined by $\kappa(\sigma_{Agt})(h) = \sigma_{Agt}(\pi_{Out}(h))$

The following lemma states the correctness of the construction of the deviator game $D(G)$, in the sense that it records the set of deviators in the strategy profile suggested by Adam with respect to the strategy profile suggested by Eve.

Lemma . Let G be a multiplayer game and σ_{Agt} be a strategy profile and $\sigma_\exists = \kappa(\sigma_{Agt})$ the associated strategy in the deviator game.

1. If $\rho \in Out_{D(G)}(\sigma_\exists)$, then $Dev(\pi_{Out}(\rho), \sigma_{Agt}) = Dev(\rho)$.
2. If $\rho \in Out_G$ and for all index i, $\rho'_i = (\rho_i, Dev(\rho_{\leq i}, \sigma_{Agt}))$, $move_i(\rho') = (\sigma_{Agt}(\rho_{\leq i}), move_i(\rho))$ then $\rho' \in Out_{D(G)}(\sigma_\exists)$.

Proof We prove that for all i, $Dev(\pi_{Out}(\rho_{\leq i}, \sigma_{Agt}) = \pi_{Dev}(\rho_{\leq i})$, which implies the property. The property holds for i = 0, since initially both sets are empty. Assume now that it holds for $i \geq 0$. $Dev(\pi_{Out}(\rho_{\leq i+1}, \sigma_{Agt}) =$

- $Dev(\pi_{Out}(\rho_{\leq i}, \sigma_{Agt}) \cup Dev(\sigma_{Agt}(\pi_{Out}(\rho_{\leq i}), \pi_{Act}(move_{i+1}(\rho))))))$ (by definition of deviators)
- $\pi_{Dev}(\rho_{\leq i}) \cup Dev(\sigma_{Agt}(\pi_{Act}(\rho)_{\leq i}), \pi_{Act}(move_{i+1}(\rho)))$ (by induction hypothesis)
- $\pi_{Dev}(\rho_{\leq i}) \cup Dev(\sigma_\exists(\rho_{\leq i}), \pi_{Act}(move_{i+1}(\rho)))$ (by definition of $\sigma_\exists$)
- $\pi_{Dev}(\rho_{\leq i}) \cup Dev(move_{i+1}(\rho))$ (by assumption $\rho \in Out_{D(G)}(\sigma_\exists)$)
- $\pi_{Dev}(\rho_{\leq i+1})$ (by construction of $D(G)$)

Which concludes the induction.

We now prove the second part. The property is shown by induction. It holds for v_0. Assume it is true up to index i, then $Tab'(\rho'_i, \sigma_\exists(\rho'_{\leq i}), move_i(\rho)) =$

- $Tab'((\rho_i, Dev(\rho_{\leq i}, \sigma_{Agt})), \sigma_\exists(\rho'_{\leq i}), move_i(\rho))$ (by definition of ρ')
- $(Tab(\rho_i, move_i(\rho)), Dev(\rho_{\leq i}, \sigma_{Agt}) \cup Dev(\sigma_\exists(\rho'_{\leq i}), \rho_{i+1}))$ (by construction of Tab')

- $(\rho_{i+1}, Dev(\rho_{\leq i}, \sigma_{Agt}) \cup Dev(\sigma_\exists(\rho'_{\leq i}), \rho_{i+1}))$ (since ρ is an outcome of the game)
- $(\rho_{i+1}, Dev(\rho_{\leq i}, \sigma_{Agt}) \cup Dev(\sigma_{Agt}(\rho_{\leq i}), \rho_{i+1}))$ (by construction of $\sigma_\exists$)
- $(\rho_{i+1}, Dev(\rho_{\leq i+1}, \sigma_{Agt}))$ (by definition of deviators)
- ρ'_{i+1}

Objectives in the Deviator Game

The objective of Eve in the deviator game is defined so that winning strategies correspond to equilibria of the original game. First, as an intermediary step, for a coalition C, a player A and an objective g in the game G $\boxed{\text{for the arena G?}}$, consider the following objective in $D(G)$: $\Omega(C, A, g) = \{\rho \mid Dev(\rho) \subseteq C \Rightarrow \texttt{payoff}_A(\pi_{Out}(\rho)) \in g\}$. It is such that a profile which ensures some goal g against coalition C corresponds to a winning strategy for $\Omega(C, A, g)$ in the deviator game, which is what we formalise now. $\boxed{\text{On ne comprend pas bien l'intuition de cette def}}$

Lemma . Let $C \subseteq Agt$ be a coalition, σ_{Agt} be a strategy profile, $g \subset V^\omega$ an objective and A a player. For all strategies σ'_C for coalition C, $\texttt{payoff}_A(\sigma_{-C}, \sigma'_C) \in g$ if, and only if, $\kappa(\sigma_{Agt})$ is winning in $D(G)$ for objective $\Omega(C, A, g)$.

Proof Let ρ be an outcome of $\sigma_\exists = \kappa(\sigma_{Agt})$. By Lemma **??**, we have that $Dev(\rho) = Dev(\kappa_V(\rho), \sigma_{Agt})$. By Lemma , $\kappa_V(\rho)$ is the outcome of $(\sigma_{-Dev(\rho)}, \sigma'_{Dev(\rho)})$ for some $\sigma'_{Dev(\rho)}$. If $Dev(\rho) \subseteq C$, then $\texttt{payoff}_A(\kappa_V(\rho)) = \texttt{payoff}_A(\sigma_{-C}, \sigma_{C \setminus Dev(\rho)}, \sigma'_{Dev(\rho)}) = \texttt{payoff}_A(\sigma_{-C}, \sigma''_C)$ where $\sigma''_A = \sigma'_A$ if $A \in Dev(\rho)$ and σ_A otherwise. By hypothesis, this payoff belongs to g. This holds for all outcomes ρ of $\sigma_\exists$, thus $\sigma_\exists$ is a winning strategy for $\Omega(C, A, g)$.

In the other direction, assume $\sigma_\exists = \kappa(\sigma_{Agt})$ is a winning strategy in $D(G)$ for $\Omega(C, A, g)$. Let σ'_C be a strategy for C and ρ the outcome of (σ'_C, σ_{-C}). By Lem.~**??**, $Dev(\rho, \sigma_{Agt}) \subseteq C$. By Lem.~**??**, $\rho' = (\rho_j, Dev(\rho_{\leq j}, \sigma_{Agt}))_{j \in \mathbb{N}}$ is an outcome of $\sigma_\exists$. We have that $Dev(\rho') = Dev(\rho, \sigma_{Agt}) \subseteq C$. Since $\sigma_\exists$ is winning, ρ is such that $\texttt{payoff}_A(\kappa(\rho)) \in g$. Since $\texttt{payoff}_A(\kappa_V(\rho')) = \texttt{payoff}_A(\rho)$, this shows that for all strategy σ'_C, $\texttt{payoff}_A(\sigma_{-C}, \sigma'_C) \in g$

Now if Eve wants to prove there is a Nash equilibria, if there is one deviator, she has to prove it does not gain anything, and if there is more than one she has nothing to prove.

$\boxed{\text{Attention il n'y a pas de fonction p sauf erreur de ma part}}$ **Theorem .** Let G be a game, σ_{Agt} a strategy profile in G, $p = \texttt{payoff}(Out(\sigma_{Agt}))$ the payoff profile of σ_{Agt}. The strategy profile σ_{Agt} is a Nash equilibrium if, and only if, strategy $\kappa(\sigma_{Agt})$ is winning in $D(G)$ for the objective $N(p)$ defined by:

$$N(p) = \{\rho \mid |Dev(\rho)| \neq 1\} \cup \bigcup_{A \in Agt} \{\rho \mid Dev(\rho) = \{A\} \wedge \texttt{payoff}_A(\pi_{Out}(\rho)) \leq p(A)\}.$$

Proof By *previous Lemma*, σ_{Agt} is a Nash equilibrium if, and only if, for each player A, $\pi(\sigma_{Agt})$ is winning for $\Omega(\{A\}, A, (-\infty, \texttt{payoff}_A(\sigma_{Agt})])$. So it is enough to show that for each player A, $\pi(\sigma_{Agt})$ is winning for $\Omega(\{A\}, A, (-\infty, \texttt{payoff}_A(\sigma_{Agt})])$ if, and only if, $\pi(\sigma_{Agt})$ is winning for $N(k,p)$.

Implication Let ρ be an outcome of $\pi(\sigma_{Agt})$.

- If $|Dev(\rho)| \neq 1$, then ρ is in $N(k,p)$ by definition.
- If $|Dev(\rho)| = 1$, then for $\{A\} = Dev(\rho)$, $\texttt{payoff}_A(\pi(\rho)) \in (-\infty, p(A)]$ because $\pi(\sigma_{Agt})$ is winning for $\Omega(Dev(\rho), A, (-\infty, p(A)])$. Therefore ρ is in $N(k,p)$.

This holds for all outcome ρ of $\pi(\sigma_{Agt})$ and shows that $\pi(\sigma_{Agt})$ is winning for $N(k,p)$.

Reverse implication Let p be such that strategy $\pi(\sigma_{Agt})$ is winning for $N(k,p)$. We now show that $\pi(\sigma_{Agt})$ is winning for $\Omega(\{A\},]-\infty, p(A)])$ for each player A. Let ρ be an outcome of $\pi(\sigma_{Agt})$, we have $\rho \in N(k,p)$. We show that ρ belongs to $\Omega(\{A\}, A,]-\infty, p(A)])$:

- If $Dev(\rho) = \varnothing$ then $\rho = Out(\sigma_{Agt})$ and $\texttt{payoff}(\rho) = p$, so ρ in $\Omega(\{A\}, A, (-\infty, p(A)])$
- If $Dev(\rho) \not\subseteq \{A\}$, then $\rho \in \Omega(C, A,]-\infty, p(A)])$ by definition.
- Otherwise $Dev(\rho) = \{A\}$. Since $\rho \in N(k,p)$, $\texttt{payoff}_A(\rho) \leq p(A)$ and therefore $\texttt{payoff}_A(\rho) \in (-\infty, p(A)]$. Hence $\rho \in \Omega(C, A, (-\infty, p(A)])$.

This holds for all outcome ρ of $\pi(\sigma_{Agt})$ and shows it is winning for $\Omega(\{A\}, A, (-\infty, p(A)])$ for each player A in Agt, which shows that σ_{Agt} is a Nash equilibrium. $\square$

Algorithm for Parity Objectives

Let us now give an explicit algorithm for parity objectives. Given a payoff p we can deduce from the previous theorem an algorithm that constructs a Nash equilibrium if there exists one. We construct the deviator game and note that we can reduce the number of vertices as follows: when $Dev(\rho_{\leq k})$ reaches a size greater than~1 it never decreases again so we know that Eve necessarily wins. During construction these states can be replaced by a sink vertex that is winning for Eve. This means that the constructed game has at most $|V| \times (|Agt| + 1) + 1$ states.

The objective can be expressed as a Parity condition in the following way:

- for each vertex $v' = (v, \{A\})$, $c'(v') = c_A(v) + 1$ if $p(A) = 0$ and 0 otherwise;
- for each vertex $v' = (v, D)$ with $|D| \neq 1$, $c'(v') = 0$ i.e. it is winning for Eve.

what is p(A)?

Note: do we have to define the payoff for a parity game here?

Lemma $\texttt{maxinf}(c'(\rho_i)) \in 2\mathbb{N}$ if, and only if, $\rho \in N(p)$.

Proof For the implication, we will prove the contrapositive. let ρ that is not in $N(p)$, then since the deviators can only increase along a play, we have that $\delta(\rho) = \{A\}$ for some player A and $\texttt{payoff}_A(\rho) > p(A)$. This means $p(A) = 0$ and $\texttt{maxinf}(c_A(\rho_i)) \in 2\mathbb{N}$. By definition of c' this implies that $\texttt{maxinf}(c'(\rho_i)) \in 2\mathbb{N}+1$ This prooves the implication.

For the other implication, let ρ be such that $\texttt{maxinf}(c'(\rho_i)) \in 2\mathbb{N} + 1$. By definition of c' this means ρ contains infinitely many states of the form $(v, \{A\})$ with $p(A) = 0$. Since the deviators only increase along the run, there is a player A such that ρ stays in the component $V \times \{A\}$ after some index k. Then for i after index k, $c'(\rho_i) = c_A(\rho_i)$, hence $\texttt{maxinf}(c'(\rho_i)) = \texttt{maxinf}(c_A(\rho_i)) + 1$. Therefore $\texttt{maxinf}(c_A(\rho_i)) \in 2\mathbb{N}$, which means $\texttt{payoff}_A(\rho) = 1 > p(A)$. By definition of $N(p)$, $\rho \notin N(p)$. $\square$

Given that the size of the game is polynomial and that parity games can be decided in nondeterministic polynomial time, the preceeding lemma implies the following theorem.

Theorem There is an NP algorithm to decide if there is a Nash equilibrium with a particular payoff.

Extensions of Nash equilibria

Subgame perfect equilibria

Nash equilibria present the disavantage that once a player has deviated, the others will try to punish him, forgetting everything about their own objectives. If we were to observe the game after this point of deviation, it would not look like the players are playing rationaly and in fact it does not look like a Nash equilibrium. The concept of *subgame perfect equilibria* refines the concept of Nash equilibrium by imposing that at each step of the history, the strategy behave like Nash equilibrium if we were to start the game now. Formally, let us write $\sigma_A \circ h$ the strategy which maps all histories h' to $\sigma_A(h \cdot h')$, that is the strategy that behave like σ_A after h. Then $(\sigma_A)_{A\in Agt}$ is a *subgame perfect equilibrium* if for all history h, $(\sigma_A \circ h)_{A\in Agt}$ is a Nash equilibrium.

Imposing such a strong restriction is justified by the fact that subgame perfect Nash equilibria exist for a large class of games. In particular subgame perfect equilibria always exist in turn-based games with reachability objectives.

Note: This is a result of Brihaye Bruyère, De Pril, Gimber, I don't know if there is a similar result for parity games

Example ....

Note: should we talk about Secure equilibria since it is not really multiplayer

Robust equilibria.

The notion of robust equilibria refines Nash equilibria in two ways:

- a robust equilibrium is *resilient*, i.e. when a small coalition of player change its strategy, it can not improve the payoff of one of its players;
- it is *immune*, i.e. when a small coalition changes its strategy, it will not lower the payoff of the non-deviating players.

The size of small coalitions is determined by some parameters k for resilience and t for immunity. When a strategy is both k-resilient and t-immune, it is called a (k,t)-robust equilibrium.

The motivation behind this concept is to address these two weaknesses of Nash equilibra:

- There is no guarantee when two (or more) users deviate together. It can happen on a network that the same person controls several devices (a laptop and a phone for instance) and can then coordinate there behavior. In that case, the devices would be considered as different agents and Nash equilibria offers no guarantee.
- When a deviation occurs, the strategies of the equilibrium can punish the deviating user without any regard for payoffs of the others. This can result in a situation where, because of a faulty device, nobody can use the protocol anymore.

By comparison, finding resilient equilibria with k greater than 1, ensures that clients have no interest in forming coalitions (up to size k), and finding immune equilibria with t greater than 0 ensures that other clients will not suffer from some agents (up to t) behaving differently from what was expected.

The deviator construction can be reused for finding such equilibria. We only need to adapt the objectives need to change. Given a game G, a strategy profile σ_{Agt}, and k, t integers:

- The strategy profile σ_{Agt} is k-resilient if, and only if, strategy $\pi(\sigma_{Agt})$ is winning in Dev for the *resilience objective* $\mathcal{R}e(k,p)$ where $p = \texttt{payoff}(Out(\sigma_{Agt}))$ is the payoff profile of σ_{Agt} and $\mathcal{R}e(k,p)$ is defined by: $\mathcal{R}e(k,p) = \{\rho \mid |Dev(\rho)| > k\} \cup \{\rho \mid |Dev(\rho)| = k \land \forall A \in Dev(\rho).\ \texttt{payoff}_A(\pi(\rho)) \leq p(A)\} \cup \{\rho \mid |Dev(\rho)| < k \land \forall A \in Agt.\ \texttt{payoff}_A(\pi(\rho)) \leq p(A)\}$
- The strategy profile σ_{Agt} is t-immune if, and only if, strategy $\pi(\sigma_{Agt})$ is winning for the *immunity objective* $\mathcal{I}(t,p)$ where $p = \texttt{payoff}(Out(\sigma_{Agt}))$ is the payoff profile of σ_{Agt} and $\mathcal{I}(t,p)$ is defined by: $\mathcal{I}(t,p) = \{\rho \mid |Dev(\rho)| > t\} \cup \{\rho \mid \forall A \in Agt \setminus Dev(\rho).\ p(A) \leq \texttt{payoff}_A(\pi(\rho))\}$
- The strategy profile σ_{Agt} is a (k,t)-robust profile in G if, and only if, $\pi(\sigma_{Agt})$ is winning for the *robustness objective* $\mathcal{R}(k,t,p) = \mathcal{R}e(k,p) \cap \mathcal{I}(t,p)$ where $p = \texttt{payoff}(Out(\sigma_{Agt}))$ is the payoff profile of σ_{Agt}.

The proof can be done as an exercise.

Extension to games with hidden actions

To represent network interaction, it makes sense to model the fact that players can have different information about the network. Unfortunately, in games with imperfect information over the states, the existence of Nash equilibria is undecidable. It is because of "forks" of information between the different players. However a restriction that is decidable is the case where only the actions are invisible. To model this, we concider strategies that only observe states of the system instead of edges, hence a strategy σ is now a function from V^* to Act.

In this version the deviators are not has obvious as before, as it may not always be possible to identify one unique deviator responsible for a deviation. The construction for decidability use a notion of suspect. Suspects for a transition (v, v') with respect to a move $(a_A)_{A \in Agt}$ are players A such that there is a'_A and $Tab(a'_A, a_{-A}) = (v, v')$. Along a history instead of taking the union of deviators, we take the intersection of suspects: note that if there is no deviation everyone is suspect, and we assume only one player will deviate (this is enough for Nash equilibria).

A construction similar that deviators, is done replacing the deviator component by a suspect component. Then the objective for Eve is that no suspect player improve its payoff. The reason is that we know the deviator is among them but we don't know which one.

Example ...

Link between determinacy and existence

Note: After having a look at the results of Stephane Le Roux it seems to me that the hypothesis of the theorem would be to complicated to expose and it is not really worth it if we are not going to give the proof or other details

Admissible strategies

Normal form games

TODO: give an example in a normal form game

Definition

Dominance

Remark: I give here the notion in a quantitative setting although it may not be needed..

Let $S \subseteq \mathcal{S}^{Agt}$ be a set of the form $S = S_1 \times S_2 \times \cdots \times S_n$ which we will call a rectangular set. Let $\sigma, \sigma' \in S_i$. Strategy σ *very weakly dominates* strategy σ' with respect to S, written $\sigma_i \geq_S \sigma'_i$, if from all vertices v_0:

- $\forall \sigma_{-i} \in S_{-i}, \mathtt{payoff}_i(Out_{v_0}(\sigma'_i, \sigma_{-i})) \geq \mathtt{payoff}_i(Out_{v_0}(\sigma_i, \sigma_{-i}))$.

Strategy σ_i *weakly dominates* strategy σ'_i with respect to S, written $\sigma >_S \sigma'$, if $\sigma \geq_S \sigma'$ and $\neg(\sigma' \leq_S \sigma)$. A strategy $\sigma \in S_i$ is weakly dominated in S if there exists $\sigma' \in S_i$ such that $\sigma' >_S \sigma$. A strategy that is not weakly dominated in S is *admissible* in S. The subscripts on $\geq_S$ and $>_S$ is omitted when the sets of strategies are clear within the context.

Remark: I give here the notion of value that is used in quantitative games, because I find it easier to generalise and a bit more natural to define

Algorithms rely on the notion of *optimistic* and *pessimistic value* of a history. The pessimistic value is the maximum payoff that a player can secure, (restricting the strategies to the ones that have not been eliminated so far *does not make sense if we don't speak of elimination beforehand*). The optimistic value is the best the player can achieve if others player help him, with the same restriction on strategies.

Definition: Values The *pessimistic value* of a strategy σ_i for a history h with respect to a rectangular set of strategies S, is

- $pes_i(S, h, \sigma_i) = \inf_{\sigma_{-i} \in S_{-i}} \mathtt{payoff}_i(h \cdot Out_{last(h)}(\sigma_i, \sigma_{-i}))$.

The *pessimistic value of a history* h for A_i with respect to a rectangular set of strategies S is given by:

- $pes_i(S, h) = \sup_{\sigma_i \in S_i} pes_i(S, s, \sigma_i)$.

The *optimistic value* of a strategy σ_i for a history h with respect to a rectangular set of strategies S is given by:

- $opt_i(S, h, \sigma_i) = \sup_{\sigma_{-i} \in S_{-i}} \mathtt{payoff}_i(h_{\leq |h|-2} \cdot Out_{last(h)}(\sigma_i, \sigma_{-i}))$.

The *optimistic value* of a history h for A_i with respect to a rectangular set of strategies S is given by:

- $opt_i(S, h) = \sup_{\sigma_i \in S_i} \mathtt{payoff}_i(opt_i(S, h, \sigma_i))$

Simple Safety games

Simple safety games, are safety games in which there are no transition from losing vertices to non-losing one. Restricting to this particular class of game makes the problem simpler because the objective becomes prefix independent. The pessimistic and optimistic value do not depend on the full history but only on the last state: for all history $pes_i(h) = pes_i(last(h))$ and $opt_i(h) = opt_i(last(h))$.

Note that in safety games (and any qualitative game) values can be only 1 (for winning) and 0 (for losing) and since the pessimistic value is always less than the optimistic one, the pair (pes_i, opt_i) can only take three values: $(0,0)$, $(0,1)$ and $(1,1)$.

Intuitively, making this value pair decrease is bad, and this is formalized by showing admissible strategies correspond exactly to strategies that do not decrease their own value in their turn.

Definition We write D_i for the set of edges $v \to v' \in E$, such that v is controlled by player A_i and $pes_i(v) > pes_i(v')$ or $opt_i(v) > opt_i(v')$. These are called *dominated edges*.

Theorem: characterisation of addmissible startegies Admissible strategies for player A_i are the strategies that never take actions in D_i.

Proof We show that if A_i plays an admissible strategy σ_i then the value cannot decrease on a transition controled by A_i. Let $\rho \in Out(\sigma_i, \sigma_{-i})$, and k an index such that $\rho_k \in V_i$. Let $(v, v') = \sigma_i(\rho_{\leq k})$:

- If $pes_i(\rho_k) = 1$, then σ_i has to be winning against all strategies σ_{-i} of A_{-i}, otherwise it would be weakly dominated by such a strategy. Since there is no such strategy from a state with value $pes_i \leq 0$, $pes_i(s') = 1$.
- If $opt_i(s) = 1$, then there is a profile σ_{-i} such that $\rho = \texttt{payoff}(Out(\sigma_i, \sigma_{-i})) = 1$. Note that $h \cdot s \cdot s'$ is a prefix of ρ. If $opt_i(s') = 0$, there can be no such profile, thus $opt_i(s') = 1$.
- If $opt_i(s) = 0$, the value cannot decrease.

In the other direction, let σ_i, σ_i' be two strategies of player A_i and assume $\sigma_i' >_S \sigma_i$. We will prove σ_i takes a transition in D_i at some point.

Lets fix some objects before digging into the proof. There is a state s and strategy profile $\sigma_{-i} \in S_{-i}$ such that $\texttt{payoff}(Out_s(\sigma_i', \sigma_{-i})) = 1 \wedge \texttt{payoff}(Out_s(\sigma_i, \sigma_{-i}) = 0$. Let $\rho = Out_s(\sigma_i, \sigma_{-i})$ and $\rho' = Out_s(\sigma_i', \sigma_{-i})$. Consider the first position where these runs differ: write $\rho = w \cdot s' \cdot s_2 \cdot w'$ and $\rho' = w \cdot s' \cdot s_1 \cdot w''$.

There are some simple facts, we can note right away:

- s' belongs to A_i, because the strategy of the other players are identical in the two runs.
- $opt_i(s_1) = 1$ because $\texttt{payoff}(Out(\sigma_i', \sigma_{-i})) = 1$
- $pes_i(s_2) = 0$ because $\texttt{payoff}(Out(\sigma_i, \sigma_{-i})) = 0$

If $opt_i(s_2) = 0$ or $pes_i(s_1) = 1$ then $s' \to s_2 \in D_i$ so σ_i takes a transition of D_i. The remaining case to complete the proof is $opt_i(s_2) = 1$ and $pes_i(s_1) = 0$. We will show that this case leads to a contradiction.

We first construct a profile $\sigma_{-i}^2 \in S_{-i}$ such that $\texttt{payoff}(\sigma_i, \sigma_{-i}^2) = 1$ from s_2. Let h be a history such that $last(h) \notin V_i$, if for all $\sigma_{-i}^2 \in S_{-i}$, $opt_i(\sigma_{-i}^2(h)) = 0$

then $opt_i(last(h)) = 0$. Thus $\sigma^2_{-i} \in S_{-i}$ never decreases the optimistic value from 1 to 0. The strategy σ_i itself does not decrease the value of A_i because it does not take transitions of D_i. So the outcome of $(\sigma_i, \sigma^2_{-i})$ never reaches a state of optimistic value 0. Hence it never reaches a state in Bad_i and therefore it is winning for A_i.

We now show there is a profile $\sigma^1_{-i} \in S_{-i}$ such that $\texttt{payoff}(\sigma'_i, \sigma^1_{-i}) = 0$ from s_1. Since $pes_i(s_1) = 0$ so there is no winning strategy for A_i from s_1 against all strategies σ_{-i}. Then there exists a strategy profile σ^1_{-i} such that σ'_i loses from s_1.

Now consider strategy profile σ'_{-i} that plays like σ_{-i} if the play does not start with w, then σ^1_{-i} after s_1 and σ^2_{-i} after s_2. Formally, given a history h, $\sigma'_{-i}(h) =$

- $\sigma^1_{-i}(h')$ if $w \cdot s_1$ is a prefix of h and $w \cdot s_1 \cdot h' = h$
- $\sigma^2_{-i}(h')$ if $w \cdot s_2$ is a prefix of h and $w \cdot s_2 \cdot h' = h$
- $\sigma_{-i}(h)$ otherwise

Clearly we have $\texttt{payoff}_i(Out_s(\sigma_i, \sigma'_{-i})) = 1 \wedge \texttt{payoff}_i(Out_s(\sigma'_i, \sigma'_{-i})) = 0$, which contradicts $\sigma'_i \geq_S \sigma_i$.

□

Once we know how to solve simple safety games, we can solve safety games by converting the safety game to an equivalent simple safety game. This is done by encoding in the states which players have visited a losing state. Note that this translation can be exponential in the number of players.

Parity games

The caracterization given for simple safety game is not enough for parity objectives.

TODO: Illustrate on an example why it is not the case.

However the fact that an admissible strategy should not decrease its own value still holds. Assuming strategy σ_i of player A_i does not decrease its own value, we can classify its outcome in three categories according to their ultimate values.

- either ultimately $opt_i = 0$, in which case all strategies are losing, and thus any strategy is admissible
- or ultimately $pes_i = 1$, in which case admissible strategies are exactely the winning ones
- or ultimately $pes_i = 0$ and $opt_i = 1$, this case is more involved and we will now focus on that case.

From a state of value 0, an admissible strategy of A_i should allow a winning play for A_i with the help of other players.

We write H_i for set of vertices v controlled by a player $A_j \neq A_i$ that have at least two successors of optimistic value 1. Formally, the *help-states* of player A_i are defined as: $H_i = \bigcup_{A_j \in Agt \setminus \{i\}} \{s \in V_j \mid \exists s', s'',\ s' \neq s'' \wedge\ s \to s' \wedge s \to s'' \wedge\ opt_i(s') = 1 \wedge\ opt_i(s'') = 1\}$

These states have the following property.

Lemma Let $s \in V$, $A_i \in Agt$ and ρ a play be such that $\exists^\infty k.opt_i(\rho_k) = 1$. There exists σ_i admissible such that $\rho \in Out_s(\sigma_i)$ if, and only if, $\mathtt{payoff}_i(\rho) = 1$ or $\exists^\infty k.\rho_k \in H_i$.

Iterated elimination

Definition: iterated elimination

TODO: give a definition

Theorem: fixpoint exists and is non-empty (not proved)

TODO: give the theorem

Other results

- Indecidability of winning strategies for imperfect information games
- Remorse-free strategies (optionnel)
- Strategy logic
- Equilibria in Stochastic Games
- Indecidability for stochastic games
- Algorithm for memoryless strategis in stochastic games (use existential theory of reals)

Pointers

Reference for medium access control: A.B. MacKenzie, S. Wicker. Stability of multipacket slotted aloha with selfish users and perfect information. IEEE INFOCOM, 3:1583–1590, 2003.

Link between determinacy and existence of Nash equilibria: Extending Finite Memory Determinacy to Multiplayer Games Stéphane Le Roux (Université Libre de Bruxelles), Arno Pauly (Université Libre de Bruxelles)