Logique dynamique pour le raisonnement stratégique dans les jeux extensifs

Cédric Dégremont⋆ Jonathan A. Zvesper⋆
degremont@hotmail.com jonathan@illc.uva.nl

⋆ Institute for Logic, Language and Computation
Plantage Muidergracht 24, 1018TV Amsterdam, Netherlands

Résumé :
Cet article poursuit l'analyse logique modale dynamique de la rationalité procédurale dans les jeux proposée par van Benthem [5]. Nous modélisons les jeux extensifs en utilisant une logique des préférences et proposons une analyse de processus analogue à l'induction à rebours. Ceci nous conduit à distinguer deux types de rationalité : la rationalité de la décision et celle des préférences. A ces deux types de rationalité correspondent des transformations de jeux, pour lesquels nous donnons des contreparties syntaxiques dans une logique modale. Dans le modèle final auquel nous parvenons par les transformations d'un jeu non générique, il peut subsister des chemins qui n'appartiennent à aucun équilibre parfait. Plus généralement la nature des solutions que notre approche peut induire est incompatible avec la nature retrospective des concepts de la théorie des jeux. Nous terminons par quelques remarques sur l'utilité d'une telle approche modale pour l'analyse des jeux en information imparfaite et en rationalité limitée.

Mots-clés : Jeux extensifs, rationalité, logique modale, logique dynamique, induction à rebours

Abstract:
This paper continues the dynamic modal logic analysis provided by van Benthem [5] of procedural rationality in games. Specifically we look at extensive games, and use preference logic to provide a closer analysis of backward induction type algorithms. This results in distinguishing two kinds of rationality : decision rationality and preference rationality. To these two kinds of rationality correspond game transformations, for which we give syntactic counterparts in a modal logic. In the final model arrived at through transformations of a non-generic game, there can be paths which are in no subgame-perfect equilibrium. More generally the nature of solutions that our approach can induce is incompatible with the retrospective nature of the usual concepts of game theory. We end the paper with some remarks on potential uses of such a modal logic analysis to the cases of imperfect information or where rationality is bounded.

Keywords: Extensive games, rationality, modal logic, dynamic logic, backward induction

Introduction

Les jeux extensifs représentent des situations d'interaction dans lesquelles les agents, ou joueurs, prennent des décisions de manière séquentielle. Leur raisonnement stratégique, notamment à propos de la rationalité des autres joueurs, est un aspect important de l'étude de tels jeux. Nous proposons une analyse logique qui prend en charge la description des jeux extensifs et la modélisation des *actions cognitives* effectuées par les joueurs lorsqu'ils raisonnent stratégiquement.

Une importante littérature s'intéresse à la modélisation des jeux en logique modale ([13] pour une vue d'ensemble) et en particulier des jeux extensifs [6]. Bonanno [10] est sans doute un des premiers à utiliser la logique temporelle pour analyser les concepts de solutions des jeux extensifs, à l'instar de la logique epistemique pour les jeux stratégiques ([24] ; pour une vue d'ensemble : [3]). Le concept d'équilibre partfait en sous-jeux a notamment fait l'objet d'une analyse modale par Bonanno [10], qui identifie une partie d'un jeu extensif *générique*[1] avec son seul équilibre parfait. Plus récemment [12] a montré comment on pouvait exprimer ce même concept, mais les outils utilisés ne sont spécifiquement modaux.

L'induction à rebours, introduite par [15], est probablement l'alogrithme de solution le plus important pour l'analyse des jeux extensifs. Elle en identifie les équi-

[1] voir définition 2.

libres parfaits en sous-jeux selon un processus itératif, consistant à éliminer les actions non-optimales des derniers joueurs à agir, puis à éliminer en fonction celle des joueurs agissant juste avant, et ainsi de suite. van Benthem [5] explore l'analogie entre "élimination itérée" de partie du jeu incompatible avec la rationalité des joueurs et restrictions de modèles épistémiques par des annonces publiques, se focalisant en particulier sur les jeux stratégiques. [5] a notamment montré comment, en utilisant un concept de rationalité globale, obtenir la solution de l'induction à rebours. Notre analyse poursuit l'exploration de l'analogie entre annonces publiques et concept de solution de jeux extensifs.

Nous modifions légèrement cette dernière perspective, en cherchant à définir un concept de rationalité locale, naturellement exprimable dans un langage modale relativement simple[2] capable d'exprimer les préférences (cf. [8]) puis recherchant quels profils de stratégies peuvent encore être construits à partir des seules arêtes épargnées par l'élimination itérée des états "irrationnels", c'est-à-dire supposant que le dernier coup joué ne respectait pas le critère de rationalité choisi. La contribution conceptuelle principale de ce papier est la distinction entre deux aspects du raisonnement stratégique effectué par les algorithmes de solution fonctionnant selon un processus itératif tel que celui de l'"induction à rebours". Nous distinguons entre la procédure d'élimination des états localement irrationnels et la procédure par laquelle les préférences sont étendues de manière causalement cohérente des noeuds terminaux vers la racine. La transformation de modèles que [5] considère prend la forme d'une annonce publique qui réduit le modèle à une de ses parties. Ce que [10] définit statiquement devient la partie du modèle résultant des annonces itérées. Nous analyserons l'élimination des états de la même façon, alors que la généralisation des préférences à l'arbre supposera de rendre certains points comparables qui ne l'étaient pas préalablement. De telles idées pour modéliser le changement de préférences ont été notamment proposées par [4].

Dans §1, nous présentons les jeux extensifs finis en information parfaite comme des structures sur lesquelles il est naturel d'interpréter un langage modal. Plus précisément, nous définirons un langage qui peut exprimer des notions de rationalité dans les jeux, comme nous le montrons dans §2. Nous définissons ensuite dans §3 une logique modale dynamique avec une expressivité suffisante. Nous ne chercherons pas ici à fournir un système de preuve complet pour une telle logique, mais plutôt d'illustrer le pouvoir expressif du langage relativement simple que nous présentons. Après avoir expliqué dans §4 quels types d'actions cognitives nous visons ici, nous montrons comment le langage modal dynamique permet d'exprimer celles-ci dans §5 (via la logique des annonces publiques) et §6. Dans §7 nous suggérerons quelques pistes pour étendre cette analyse aux jeux en information imparfaite et au cas des agents ayant des ressources cognitives limitées.

1 Structures modales

Dans cette section nous présentons les jeux extensifs en information parfaite comme des structures sur lesquelles il est naturel d'interpréter un langage modal. Nous définissons un *jeu* $\mathcal{G}$ comme un tuplet de la forme

$$\langle W, N, A, (\stackrel{i\ a}{\rightarrow})_{a \in A, i \in N}, (\succ_i^z)_{i \in N} \rangle,$$

où W est un ensemble fini non-vide d'états, N est un ensemble fini de joueurs, et A est un ensemble fini d'actions. $\stackrel{i\ a}{\rightarrow} \subseteq W \times W$ pour chaque $a \in A$ et $i \in N$ définit les transitions possibles entre états :

[2] PDL avec intersection et converse mais sans test ni itération.

$w \xrightarrow{i\,a} v$ signifie qu'à l'état w, le joueur i peut prendre l'action a pour arriver à l'état v.

Soit $\xrightarrow{a} = \cup_{i \in N} \xrightarrow{i\,a}$, $\xrightarrow{i} = \cup_{a \in A} \xrightarrow{i\,a}$ et $\rightarrow = \cup_{i \in N} \xrightarrow{i}$ pour chaque $i \in N$ et $a \in A$. Soit $\rho(w) := \{i \in N | \exists v \in W : w \xrightarrow{i} v\}$ les contrôleurs de l'état w. Nous supposerons que les actions sont déterministes et qu'à chaque état un seul joueur joue, c'est-à-dire que si $a \neq b$ ou $i \neq j$ alors $(\xrightarrow{i\,a} \cap \xrightarrow{j\,b}) = \emptyset$. Ceci implique que chaque état a au plus un contrôleur.

$\succ_i^z$ exprime les préférences du joueur i. Suivant la tradition en théorie des jeux nous restreindrons dans un premier temps les préférences sur les issues possibles, qui correspondent dans notre présentation aux états terminaux $W_z := \{w \in W | \rho(w) = \emptyset\}$, c'est-à-dire aux états sans contrôleur. Nous stipulons pour chaque $i \in N$ que $\succ_i^z \subseteq W_z \times W_z$ est une relation bien fondée, transitive et irréflexive. $w \succ_i^z v$ signifiant que i préfère strictement l'état terminal w à l'état terminal v.

Il est immédiat de voir comment de telles structures correspondent à des jeux extensifs en information parfaite. Dans l'exemple de jeu de la figure 1 (cf. [17] p. 96), nous avons étiqueté chaque noeud avec une lettre. Ces lettres r, s, t, u, v sont les états de nos modèles. Le jeu $\mathcal{G}$ correspondant à la figure 1 est le suivant : $W = \{r, s, t, u, v\}$; $N = \{I, II\}$; $A = \{A, B, L, R\}$; $r \xrightarrow{I\,A} s$, $r \xrightarrow{I\,B} t$, $s \xrightarrow{II\,L} u$, $s \xrightarrow{II\,R} v$; $v \succ_I^z t \succ_I^z u$; $t \succ_{II}^z v \succ_{II}^z u$.

2 Aspects de la rationalité

Nous considérons que dans un jeu extensif, quatre aspects distincts de la rationalité des joueurs importent au raisonnement stratégique. Plus précisément, faire l'hypothèse de la rationalité des joueurs, c'est supposer que :

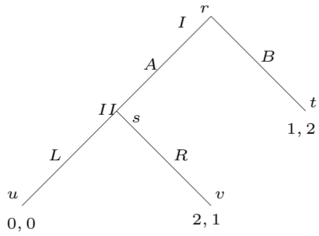

FIG. 1 – Un jeu extensif

1. les préférences des joueurs sont cohérentes, à la fois de manière statique habituelle (à savoir elles sont supposées être transitives et, dans leur version stricte, irreflexives.) mais aussi de façon “causale” : les situations conduisant à des situations préférées sont également préférées.
2. les joueurs choisissent toujours une des options conduisant à une de leurs issues préférées.
3. le raisonnement des joueurs est correct - ils ne font pas d'erreurs - et complet au sens où ils n'ont de limites computationnelles en temps ou en espace.
4. les joueurs mettent à jour leurs croyances de façon rationnelle.

Dans l'interprétation déductive de la théorie des jeux et de ses concepts de solution, les “joueurs cherchent à déduire les actions rationnelles de leurs opposants à partir des préférences de leurs opposants et de l'analyse du raisonnement de leurs opposants à propos de leurs propres actions rationnelles.”([16], p. 377). On pourrait défendre l'idée que la procédure par laquelle les joueurs éliminent de façon itérée les parties du jeu incompatibles avec l'hypothèse de la rationalité des autres joueurs peut être vu comme un processus de révision des croyances. Durant le processus de réduction du jeu, un joueur *révise* sa représentation du jeu. A la suite de [5] nous défendons précisément l'idée que ce processus peut être capturé dans le style *dynamique epistemique* (au sens de

[11]). Néanmoins nous réservons la question de la révision des croyances aux jeux en information imparfaite dans lesquels les joueurs peuvent avoir des ensembles d'information différents et peuvent ainsi ne pas partager la même information au début du jeu.

Nous défendons l'idée que la rationalité des agents couvre ces quatre aspects. La dernière sorte de rationalité sera l'objet de recherche ultérieure. Pour le moment nous nous concentrerons sur les jeux en information parfaite. Nous considérerons également pour commencer que les capacités cognitives ou computationnelles des agents sont illimitées. Enfin appelons les deux autre sortes de rationalité : "rationalité de la décision" et "rationalité des préférences". Les agents prennent des décision de façon cohérente avec des préférences cohérentes.

Dans un premier temps la rationalité des préférences se définira par une façon canonique d'induire une relation de préférence pour chaque joueur qui s'étend à l'ensemble des états contenus dans l'arbre du jeu, à partir des préférences sur les états terminaux. Une fois les préférences étendues à tous les états (plutôt qu'aux seuls états terminaux) il devient facile de définir un langage modal dans lequel nous pouvons exprimer certains aspects du raisonnement stratégique, et notamment le fait que la dernière action prise par un joueur ait été une décision optimale. Nous décrirons cette méthode canonique et rechercherons comment nous pouvons définir une notion de rationalité de l'action, composée des deux aspects mentionnés. Plus précisément nous montrerons comment la rationalité de l'action réduit un jeu à son 'noyau rationnel', de façon comparable à l'induction à rebours, bien que ne coïncidant pas avec cette dernière.

Ainsi suivant [5] nous proposons un cadre dans lequel les actions cognitives constituant le raisonnement stratégique des joueurs sont des citoyens de première classe. A la suite de [5], nous explicitons ces actions dans le style des actions changeant le modèle de la logique des annonces publiques [19]. Dans un premier temps l'étape de raisonnement que nous considérerons est l'élimination des états qui ne peuvent être atteints que si un joueur a agi de façon irrationnelle.

La façon canonique et rationnelle d'induire les préférences, comme nous le notons dans §6 prend en charge une partie du raisonnement constituant la procédure de réduction du jeu. Puisque que nous cherchons à mettre au premier plan le raisonnement stratégique des joueurs, nous chercherons à rendre compte de l'action par laquelle les préférences sont induites des noeuds terminaux vers les autres dans un style dynamique de changement de modèles.

3 Langage modal dynamique

Nous définissons un une logique dynamique dans l'esprit de [18]. Ainsi fixant un ensemble A d'actions et N de joueurs, nous définissons récursivement l'ensemble des actions ACT :

$$ACT ::= \ A \mid N \mid \succ_N \mid \ \rightarrow \ \mid ACT \cup ACT \mid ACT \cap ACT \mid ACT^c \mid ACT; ACT,$$

A chaque action correspond une modalité dans le langage $\mathcal{L}$. Nous supposons un ensemble dénombrable de lettres propositionnelles Φ :

$$\mathcal{L} ::= \ \mid \Phi \mid \neg\mathcal{L} \mid \mathcal{L} \wedge \mathcal{L} \mid \langle ACT \rangle \mathcal{L}.$$

Nous utilisons les abréviations habituelles. Nous interprétons ce langage sur un jeu $\mathcal{G}$ de la forme

$$\langle W, N, A, (\overset{i\ a}{\rightarrow})_{a \in A, i \in N}, (\succ_i^z)_{i \in N} \rangle$$

accompagné d'une valuation $V : \Phi \rightarrow \wp(W)$. Comme nous l'avons mentionné, nous étendons la relation de préférences au-delà des noeuds terminaux. Nous définissons pour chaque joueur i la nouvelle relation de préférence étendue à $W \times W$ inductivement comme la plus petite relation contenant $\succ_i^z$ et clos sous les règles suivantes :

1. $((\exists x \ : w \rightarrow x) \ \& \ \forall x \ (w \rightarrow x \Rightarrow x \succ_i v)) \Rightarrow w \succ_i v$
2. $((\exists x \ : w \rightarrow x) \ \& \ \forall x \ (w \rightarrow x \Rightarrow v \succ_i x)) \Rightarrow v \succ_i w$
3. $(\exists x \ : (w \xrightarrow{i} x \ \& \ x \succ_i v)) \Rightarrow w \succ_i v$

Expliquons le processus d'induction des préférences. Fixons un joueur i. Nous commençons par les états terminaux, c'est-à-dire avec $\succ_i \ := \ \succ_i^z$. Puis nous étendons la relation aux états non-terminaux. Dans la précondition de chacune des règles de clôture nous spécifions $\exists x \ : w \rightarrow x$, de telle sorte qu'elles ne s'appliquent à un état w que si ce *n*'est *pas* un point terminal. L'idée est de sélectionner deux états arbitraires w and v et de voir s'ils peuvent être comparés étant données les préférences déjà explicites. La règle 3 est la plus courte et sans doute la plus simple, elle énonce que si à un état w c'est à i de jouer et qu'une des actions que i peut accomplir en w le conduit à un état qu'il (i) préfère à v, alors i préfère w à v. Les deux autre conditions sont très similaires, nous n'expliquons donc que la première. Elles s'appliquent même si ce n'est pas à i de jouer en w. Dans ce cas, la règle énonce que i préférera w à v si *tous* les successeurs possibles de w sont préférées à v par i. Nous reviendrons à cette opération d'induction des préférences §6.

Nous pouvons donc maintenant assigner à chaque action $\alpha \in ACT$ une relation R_α :

$$
\begin{array}{lcl}
R_{\rightarrow} & = & \rightarrow \\
R_a & = & \xrightarrow{a} \\
R_i & = & \xrightarrow{i} \\
R_{\succ_i} & = & \succ_i \\
R_{\beta\cup\gamma} & = & R_\beta \cup R_\gamma \\
R_{\beta\cap\gamma} & = & R_\beta \cap R_\gamma \\
R_{\beta^c} & = & \{(v,w) | wR_\beta v\} \\
R_{\beta;\gamma} & = & \{(v,w) | \exists x : vRxRw\}
\end{array}
$$

Supposant une fonction de valuation $V : \Phi \rightarrow \wp(W)$, l'interprétation de $\mathcal{L}$ est standard pour la logique modale :

$$
\begin{array}{lll}
w \vDash \top & & \\
w \vDash p & \text{ssi} & w \in V(p) \\
w \vDash \phi \wedge \psi & \text{ssi} & w \vDash \phi \text{ et } w \vDash \psi \\
w \vDash \neg\phi & \text{ssi} & w \nvDash \phi \\
w \vDash \langle\alpha\rangle\phi & \text{ssi} & \exists v : wR_\alpha v \text{ et } v \vDash \phi
\end{array}
$$

4 Actions cognitives dans le raisonnement stratégique

Nous passons maintenant à la représentation du raisonnement stratégique des joueurs, amenant au premier plan les étapes de ce raisonnement, que nous appelons des "actions cognitives" (suivant plus ou moins le sens de [7]). Reprenant l'idée proposée pour la première fois par [5], nous utiliserons les ressources de la logique des annonces publiques (PAL) pour modéliser les actions cognitives. Une annonce publique est une opération changeant le modèle, le restreignant à l'un de ses sous-modèles. Une annonce publique d'une formule ϕ assigne à un modèle $\mathcal{M}$ l'un de ses sous-modèles $\mathcal{M} \restriction \phi$, défini comme le modèle dont le domaine est l'ensemble des états $w \in \mathcal{M}$ tels que $\mathcal{M}, w \vDash \phi$, et dont les relations et valuations sont les restrictions correspondantes à celle de $\mathcal{M}$ à ce domaine. Syntactiquement, dans PAL il existe une modalité $[!\phi]$ pour chaque formule du langage en question, dont la sémantique est la suivante :

$$\mathcal{M}, w \vDash [!\phi]\psi \text{ ssi}$$

$$(\text{si } \mathcal{M}, w \vDash \phi \text{ alors } \mathcal{M} \restriction \phi, w \vDash \psi).$$

Selon l'interprétation déductive des concepts de solution dans les jeux extensifs, le raisonnement stratégique des joueurs est le processus par lequel – dans un jeu joué une seule fois – les joueurs raisonnent au sujet des choix des autres joueurs sous l'hypothèse de leur rationalité. Ce raisonnement peut-être analysé et réduit à une série de plus petites actions cognitives, qui peuvent être conçues comme des annonces par lesquelles les noeuds du jeu qui sont incompatibles avec la rationalité des joueurs sont éliminés. Plus précisément, et telle est l'idée cruciale de [5], si I conclut que II ne choisira pas une certaine action parce qu'elle est rationnelle, ceci peut être représenté comme une annonce publique véridique de la forme, "la joueuse II est rationnelle". Puisque que nous supposons que le processus de raisonnement de chaque joueur fait partie de la rationalité même d'un joueur, nous supposons en fait que ce raisonnement est connaissance commune entre les joueurs. En effet si le joueur I était le seul à faire ce raisonnement, nous devrions considérer la représentation que chaque joueur se fait du jeu et le modèle résultant ne serait pas approprié pour présenter le raisonnement *commun* des joueurs au sujet du jeu. Pour résumer, éliminer un point du modèle revient à dire qu'aucun des joueurs ne le considère plus. Il faudrait donc parler d'une *action cognitive commune*, dont la légitimité repose sur la connaissance commune de la rationalité des joueurs (dans les quatre sens mentionnés dans §2).

L'abandon de l'hypothèse d'un raisonnement commun doit avoir lieu lorsque nous acceptons non seulement l'information imparfaite (des ensembles d'information différents), mais si nous acceptons que les joueurs puissent ne pas savoir ce que les autres savent et ignorent. Dans ce cadre plus général, une analyse plus subtile des actions cognitives pourrait être proposée, permettant par exemple de modéliser le cas où un joueur accomplit une certaine étape dans son raisonnement mais n'est pas sêr que les autres l'ont accomplie aussi. Prendre en compte des types d'actions plus complexes pourrait par exemple permettre de proposer une nouvelle analyse des jeux de dé-synchronisation [1]. Mais nous laissons ces considérations à des recherches futures, pour lesquelles notre analyse des jeux en information parfaite pourra fournir une fondation.

5 Rationalité de la décision et induction à rebours

Nous nous intéressons maintenant à la partie du jeu stable sous l'annonce itérée de rationalité. Plus précisément, nous nous intéressons aux chemins terminaux survivant à l'annonce itérée de rationalité. On pourrait apprécier une définition de la rationalité tels que les noeuds terminaux accessibles après l'itération de l'annonce lui correspondant serait précisément les équilibres parfait en sous-jeux du jeu. Mais plutôt que de chercher la notion modale de rationalité correspondante, inversons la perspective et demandons : qu'est-ce qu'une notion modale naturelle de rationalité, et (*via* son annonce itérée) quel concept de solution induit-elle ?

5.1 Quelle rationalité ?

De nombreux concepts de rationalité peuvent être envisagés. Comme nous sommes intéressés par les perspectives que la logique modale peut apporter à l'analyse du raisonnement stratégique dans les jeux extensifs, nous sommes particulièrement intéressés par la question suivante :

> "Y a-t-il une manière naturellement modale d'exprimer la rationalité d'une décision ?"

Pour inspirer et motiver notre réponse à cette question, nous considérerons deux slogans tirés de deux manuels, l'un de logique modale, l'autre de théorie des jeux.

SLOGAN 1 (LOGIQUE MODALE) Les langages modaux fournissent une perspective interne et locale sur les structures relationnelles.[9]

Le second est proposé pour caractériser l'hypothèse crucial de la théorie du choix rationnel :

SLOGAN 2 (CHOIX RATIONNEL) Lors qu'il prend une décision, un agent choisit une action qui est au moins aussi bonne, d'après ses préférences, que n'importe quelle autre action disponible.[16], p. 6

Essayons de rendre le second slogan dans l'esprit du premier, c'est-à-dire en prenant la rationalité dans un sens "local", en restreignant l'attention de la formule exprimant la rationalité au dernier coup d'un joueur. L'avantage de cette notion locale de rationalité est qu'elle autorise une caractérisation modale simple. our la clarté de l'exposition, nous définissons la rationalité de la décision comme la négation de l'*ir*rationalité de la décision. Un joueur i est considéré comme *venant de prendre une décision irrationnelle* au point w si i contrôle l'unique prédécesseur de w, disons v, et qu'en v, i pouvait prendre une action qu'il aurait emmené (en un seul pas) dans un état u qu'il (i) préfère à w. C'est-à-dire si i vient de faire quelque chose qu'on pourrait lui faire regretter à l'aide d'une explication très simple. C'est un concept extrêmement locale d'irratationalité : dans le jeu présentée dans la figure 2, le joueur I n'est irrationnel à aucun des points terminaux. Néanmoins, si pour commencer il va à gauche, alors au noeud intermédiaire, il *est* irrationnel.

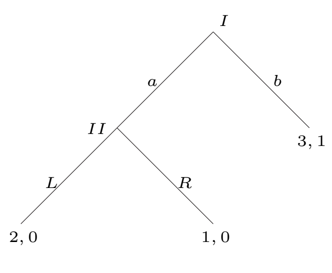

FIG. 2 – Un jeu illustrant la *localité* de notre notion d'irrationalité de la décision.

Cette localité suggère une limitation au raisonnement des joueurs, une question que nous aborderons rapidement dans §7.

Nous pouvons définir la notion d'irrationalité de la décision de i avec la $\mathcal{L}$-formule suivante :

$$\langle (i^c; i) \cap \succ_i{}^c \rangle \top$$

La négation nous donne notre concept de rationalité de la décision rat_i :

$$rat_i :=_{df} [(i^c; i) \cap \succ_i{}^c] \bot$$

rat_i peut être lu de la façon suivante il n'est pas possible ($[]\bot$) que i prenne une action (i) au point précédent (i^c) et que cela le conduise ($\cap$) à un état qu'il préfère au point actuel ($\succ_i{}^c$).

5.2 Quel concept de solution ?

Une annonce de la rationalité d'un joueur réduira (potentiellement) la taille du jeu, éliminant les états dans lequel un joueur a joué de façon irrationnelle. Etant donné la notion (c'est-à-dire l'ensemble d'opérations) d'induction des préférences que nous avons stipulées pour nos modèles ; l'itération des annonces de la rationalité de tous les joueurs peut conduire à une série de réduction.

Par exemple, prenons le jeu donné par la figure 1. Observons qu'en u, rat_{II} n'est pas satisfait : il y a un état, à savoir v, tel que $uR_{(II^c;II) \cap \succ_{II}{}^c} v$. Après une annonce de rat_{II}, nous obtenons le sous-jeu décrit par la figure 3.

Etant donné la façon dont les préférences sont stipulées dans nos modèles, il suit que le joueur I préfère *désormais* s à t. Dès lors une annonce de rationalité additionnelle, à savoir $[!rat_I]$, et nous obtenons le sous-jeu décrit dans la figure 4.

Dans *ce* cas nous avons obtenu, via une séquence d'annonces publiques, un jeu

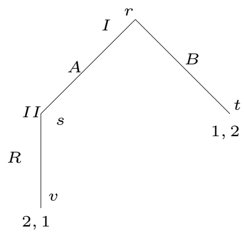

FIG. 3 – Le jeu de la figure 1 après une étape de réduction.

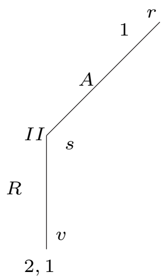

FIG. 4 – Le jeu de la figure 1 après deux réductions.

contenant comme unique chemin terminal le seul chemin terminal qui peut être construit en utilisant les actions contenues dans l'*unique* équilibre parfait en sous-jeux de jeu original. En fait cette coïncidence ne se vérifie pas dans le cas *général*. Mais nous avons le résultat suivant :

PROPOSITION 1 . *Un chemin maximal dans le jeu $\mathcal{G}$ survit à l'itération de l'annonce de rationalité commune des joueurs ($\bigwedge_{i \in N} rat_i$) ssi il est composé d'arêtes contenues dans l'union des équilibres parfaits en sous-jeux du jeu $\mathcal{G}$.*

Preuve. De gauche à droite, la preuve est par induction sur la longueur du sous-jeu dans lequel une action qui n'est contenue dans aucun équilibre parfait en sous-jeux apparaît et par réduction à l'absurde. ■

DÉFINITION 2 Un jeu en information parfaite est générique si aucun joueur n'est indifférent entre deux noeuds terminaux, c'est-à-dire si $\forall i \in N$,$\forall w, w' \in W_z$ $w \neq w' \rightarrow (w \succ_i^z w' \vee w' \succ_i^z w)$.

FAIT 3 *Un jeu générique ne contient qu'un seul équilibre parfait en sous-jeux.*

DÉFINITION 4 ([5]) L'assertion de rationalité momentanée (MR) énonce qu'à chaque étape d'une branche dans le modèle actuel, celui qui doit jouer, n'a pas choisi une action dont toutes les continuations finissent plus mal pour lui que toutes celles suivant une autre action.

Nous obtenons comme corollaire le résultat suivant prouvé par [5].

COROLLAIRE 5 ([5]) *Sur les arbres de jeux extensifs (en information parfaite) génériques finis, l'annonce itérée de MR conduit exactement à la solution de l'induction à rebours.*

Mais revenons au concept de solution que nous obtenons. On pourrait opposer qu'il s'agit là d'un concept de solution faible. En effet dans l'exemple de la figure 2, $\{\{a\}, \{R\}\}$ n'est pas un équilibre parfait en sous-jeux. Néanmoins ce profile de stratégie demeure dans notre procédure. Nous défendons l'idée que le concept d'équilibre parfait en sous-jeux (comme plus généralement celui d'équilibre de Nash) est *rétrospectif* par nature. Un tel concept nous indique que si la stratégie de II en s est de jouer R, alors le joueur I *s'en serait* mieux sorti s'il *avait* choisi b. Le concept de solution que nous obtenons fonctionne lui de façon *prospective* et *prescriptive*. En ce sens une *prescription* répond à certaines contraintes de cohérence. Vous pouvez également dire à I qu'il doit jouer b et ne pas jouer a, vous obtiendriez un concept de solution tout aussi *prescriptif* mais plus fort, à savoir la maximisation du minimum. Mais si vous admettez l'action a comme une chose que I pourrait rationnellement faire, alors la cohérence de la prescription ne permet de traiter l'issue t comme irrationnelle.

Il est à noter que cette situation est plus générale. Dès que nous sortons des jeux génériques pour traiter le cas des jeux extensifs en information parfaite, aucun concept de solution qui n'est pas clos sous les futurs rationnels ne peut être modélisé par une annonce itérée de rationalité. Plus précisément si un concept de solution qui n'admet pas tous les profils de stratégies qui peuvent être construits sur la base des chemins maximaux survivant à l'itération de l'annonce de rationalité, alors celui-ci n'est pas exprimable avec l'appareil modale dynamique que nous proposons. En particulier nous ne pouvons pas obtenir dans le cas général la solution de l'induction à rebours.

Comme nous l'avons déjà mentionné, une bonne partie de ce résultat est secrètement encodé dans notre définition de la façon dont les préférences sont induites des noeuds terminaux vers les autres. Etant donné que notre objectif est d'expliciter les actions cognitives composant le raisonnement stratégique des agents, il est raisonnable de demander que ce type d'action cognitive, rendant nos préférences causalement cohérente, soit modélisé par une action dans une logique dynamique, à l'instar de ce que nous avons fait pour la rationalité de la décision. Ce que nous faisons dans la section suivante.

6 Préférences

Notre procédure de réduction du jeu à son noyau rationnel, telle que nous l'avons décrite, tire profit de ce que nous avons décrit comme la rationalité des préférences. Ainsi un aspect du processus de raisonnement est effectuée en silence par les contraintes qui définissent les relations de préférences généralisées à l'arbre. Ceci ne rend pas justice à ce processus autonome comme étant d'égale importance pour isoler la solution du jeu. Nous altérons donc la sémantique et prenons une perspective plus générale, abandonnant l'hypothèse selon laquelle les préférences sont automatiquement induites de façon canonique à chaque fois que nous passons d'un modèle à une de ses parties.

Nous n'irons pas ici jusqu'à considérer des préférences qui ne sont pas "nettes" (cf. [17] p. 4) : nous n'autorisons pas l'imprécision des préférences, ni préférences intransitives ou symétriques. Nous supposons des joueurs dont les préférences sont "statiquement" cohérentes. Mais nous abandonnons ici l'idée que les préférences sont automatiquement spécifiées sur tout l'arbre afin de mettre en avant le processus de raisonnement par lequel les joueurs - en analysant la structure du jeu - peuvent décider quelles options intermédiaires sont préférables, étant donné leurs préférences sur des noeuds plus proche de l'issue du jeu. Ce processus doit conduire à des préférences "causalement" cohérentes. Il s'agit par exemple d'exclure comme irrationnel qu'un joueur d'échecs préfère prendre la reine de son adversaire (à un autre coup), alors que cela le conduit à un noeud duquel son adversaire possède une stratégie de victoire.

Un modèle sera dotée d'une relation $\succ_i$ qui avant toute action cognitive ne sera pas différente de $\succ_i^z$, exactement comme les modèles traditionnels de la théorie des jeux.

Pour chaque propriété $\mathcal{P}$ que nous utilisons pour définir $\succ_i$ et chaque joueur $i \in N$ nous considérons une action "rationalisant les préférences" du joueur i d'une certaine façon, comme une opération $\mathcal{O}_i^{\mathcal{P}}$ de $(W \times W)^n$ vers $W \times W$. Soit $\mathcal{O}_I^{\mathcal{Q}}$ l'opération correspondant à $((\exists x \; : \; w \rightarrow x) \wedge \forall x \; (w \rightarrow x \Rightarrow x \succ_I v)) \Rightarrow w \succ_I' v$. Par abus de langage nous écrirons $\mathcal{O}_I^{\mathcal{Q}}(\mathcal{M})$ pour $\langle W, N, A, (\overset{i\ a}{\rightarrow})_{a \in A, i \in N}, \succ_I \cup \{(w,v) \in W \times W : \exists x \in W \; ((w \rightarrow x) \wedge \forall x \in$

$W\ (w \rightarrow x \Rightarrow x \succ_I v))\}, (\succ_i)_{i \in N-I}\rangle$. Intuitivement une telle opération rendra comparables de noeuds de l'arbre entre lesquels un joueur était préalablement indifférent, selon des maximes du type “préfère les moyens qui te mènent à tes fins”.

Dans notre cas, l'ordre dans lequel nous appelons les différentes opérations de rationalisation des préférences est sans importance. Pour simplifier définissons donc une opération unique pour toutes les opérations et tous les joueurs que nous notons $\mathcal{RP}$.

REMARQUE 6 L'axiome de réduction suivant donne la clé de n'analyse compositionnelle dans la syntaxe de l'opération $\mathcal{RP}$:

$$[ratPref_i]\langle\succ_i\rangle\phi \equiv$$
$$(\langle\succ_i\rangle[ratPref_i]\phi \quad \vee$$
$$(\langle\rightarrow\rangle\top \wedge [\rightarrow]\langle\succ_i\rangle[ratPref_i]\phi) \quad \vee$$
$$\langle\succ_i\rangle([\succ_i{}^c \cap (\rightarrow^c; \rightarrow)]\bot \wedge \langle\rightarrow^c\rangle[ratPref_i]\phi)$$
$$\vee \quad \langle i; \succ_i\rangle[ratPref_i]\phi)$$

La formule est correcte par rapport à la clause sémantique suivante :

$$\mathcal{G}, V, w \vDash [ratPref]\phi \Leftrightarrow \mathcal{RP}(\mathcal{G}), V, w \vDash \phi$$

où $[ratPref]\phi := \bigwedge_{i \in N}[ratPref_i]\phi$

Cette remarque est rendue naturelle par le fait que le premier disjoint corresponde à la relation initiale alors que chacune des autres clauses correspond à une des règles que nous utilisons pour définir $\succ_i$ à partir de $\succ_i^z$ dans §3.

Afin de voir que ces deux actions cognitives (l'une rationalisant les préférences, l'autre les décisions) peuvent cohabiter, l'ordre d'application de ces deux actions ne doit pas importer (ce qui pourrait être le cas si nous avions choisi les versions “faibles” de nos opérations). De ce fait suit la version explicite de notre précédente proposition. Définissant $\mathcal{RD} := \mathcal{M} \mapsto \mathcal{M} \upharpoonright \bigwedge_{i \in N} rat_i$, nous obtenons :

PROPOSITION 7 . *Un chemin maximal dans le jeu $\mathcal{G}$ survit à l'application itérée des opérateurs $\mathcal{RP}$ et $\mathcal{RD}$ ssi il est composé d'arêtes contenues dans l'union des équilibres parfaits en sous-jeux du jeu $\mathcal{G}$.*

Enfin, il serait conceptuellement plus correct de voir ces opérations comme la mise en oeuvre de normes de rationalité plutôt - comme le parallèle avec les annonces publiques pourrait le suggérer - que comme l'action de révéler que les joueurs sont rationnels. Elles conseillent aux joueurs d'éviter certaines décisions, de même ces normes indiquent quels états intermédiaires les joueurs *devraient* préférer atteindre. La mise en oeuvre de ces normes de rationalité est précisément ce qu'effectuent les actions cognitives qui constituent le raisonnement stratégique.

Une approche plus générale consisterait à *libéraliser* les préférences des joueurs, autorisant des incohérences “causales”. Autorisant par exemple qu'un joueur i préfère touts les successeurs de s à ceux de t mais préfère t à s. Néanmoins le processus de rationalisation ne serait plus une simple expansion et nous ferions face à différentes façons de rationaliser ces préférences. Donner la priorité aux préférences entre noeuds plus proches des noeuds terminaux pourrait permettre d'obtenir un concept de solution ayant un comportement régulier, mais nous laissons cette question à d'autres recherches.

7 Information imparfaite et rationalité limitée

Nous nous sommes concentrés sur les actions cognitives, car il est clair qu'elles transforment la représentation que les joueurs se font du jeu. Parfois toute l'information n'est pas disponible au début du jeu pour tous les joueurs. Une lecture (équivalente à la lecture classique) des jeux en information imparfaite repose dans l'idée qu'il faut parfois attendre que le jeu soit joué pour obtenir certaines informations précieuses. Prenons l'exemple de la figure 5. Dans ce jeu, le joueur II peut élimi-

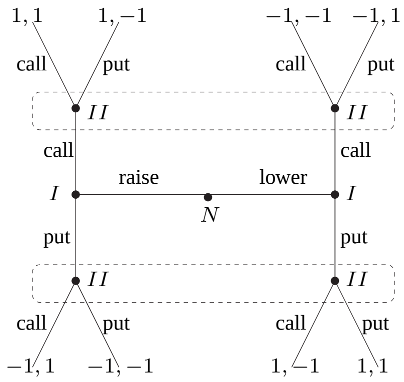

FIG. 5 – Un jeu en information imparfaite.

ner certains états irrationnels, voire figure 6. Néanmoins jusqu'à ce que I joue, le

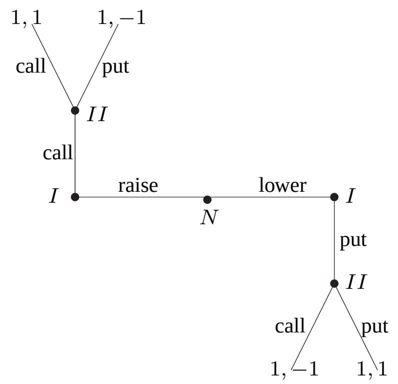

FIG. 6 – Le jeu de la figure 5 après une annonce de rationalité

joueur II ne peut plus réduire l'arbre du jeu et fixer une stratégie. Nous interprétons les jeux en information imparfaite et incomplète comme des jeux dans lesquels l'information fournie par les actions effectives des joueurs n'est pas redondante. Nous pouvons étendre l'interprétation des actions cognitives constituant le raisonnement stratégique en termes d'actions dynamiques changeant le modèle. Généralisant le processus par lequel l'information est pris en compte par les joueurs dans la représentation du jeu, de nouvelles questions apparaissent : quelle information peut révéler les préférences d'un joueur, le joueur i a-t-il une stratégie rationnelle qui ne révèle pas l'information qu'il possède, etc. Néanmoins cette lecture d'un jeu en information imparfaite est équivalente à celle utilisant des stratégies conditionnelles. Son principale intérêt est d'introduire l'idée que la temporalité des actions cognitives et des actions en général est la même, même si, ayant à faire à des agents idéaux, on peut aussi bien considérer les stratégies comme définies préalablement à toute action effective.

Or l'économie et la théorie des jeux classiques supposent habituellement des agents ou des joueurs ayant des capacités cognitives (ou computationnelles) illimitées. Il serait d'ailleurs plus exact de dire que ces théories ne supposent rien du tout, puisque le calcul n'y est même pas considéré comme un processus consommant des ressources. Sur le plan descriptif, les théories prédisent des résultats contredits par l'expérience[3]. Sur le plan normatif, les prescriptions correspondantes peuvent très bien se révéler largement sous-optimales. Suivant le travail fondateur de Simon [22] nous proposons de chercher une théorie plus exacte de "la façon dont les vrais êtres humains prennent de vraies décisions dans un monde qui leur fournit rarement les données et les *ressources cognitives* qui seraient exigées pour appliquer, littéralement, les théories des manuels" [23], c'est-à-dire les théories qui ne prennent pas le processus de raisonnement au sérieux ou le fait que les agents aient des capacités cognitives limitées.

Analysant les jeux en information parfaite, nous avons nous aussi supposé que le rai-

[3]Ainsi l'unique équilibre parfait en sous-jeux du dilemme des prisonniers répété de façon finie, à savoir trahir l'autre à chaque tour, est fortement invalidée par le comportement des sujets expérimentaux (cf. e.g.[2]).

sonnement était mené à bien une fois pour toutes, isolant les profiles de stratégies respectant certaines contraintes de rationalité. Une fois ce raisonnement achevé les agents agiraient en suivant un des chemins terminaux restants. Mais il est clair que les agents réels ne procèdent pas ainsi (cf. [20]) : déjà dans l'Antiquité, Aristote soulignait le fait que la délibération ne doit pas durer indéfiniment. Plus concrètement, imaginez que vous êtes devant une carte du métro parisien à *Châtelet* essayant de calculer la façon la plus rapide de parvenir à *La rue de Tanger*. Comme vous êtes déjà en retard, votre fonction d'utilité est donnée par $u(t) = -(t_{tanger} - t_0)^2$. Il est hautement improbable[4] que vous vous en sortiriez mieux en commençant par calculer l'itinéraire optimal et commenciez *seulement ensuite* à vous déplacer. Tout ceci pour indiquer que le temps de la réflexion, pendant lequel vous exécutez des actions cognitives, et le temps de l'action, pendant lequel vous agissez est bien le même.

Tout ceci était folklorique. Mais qu'est-ce que la logique modale peut avoir à dire au sujet de la prise de décision en rationalité limitée ? D'autres réponses visent les logiques épistémiques [21] ou les logiques de l'action [14]. Or, revenons à nos opérations $\mathcal{RD}$ et $\mathcal{RP}$; nous avons pour le moment admis que les agents pouvait itérer l'exécution de l'une des deux opérations sans limite avant d'agir. Autrement dit les agents peuvent exécuter $(\mathcal{RD} \cup \mathcal{RP})^n$ où $n \in \omega$, $\mathcal{R}^1 := \mathcal{R}$,$\mathcal{R}^{i+1} := \mathcal{R}^i; \mathcal{R}$, avant de prendre la moindre décision. Intégrer la limitation cognitive en termes de temps alloué au calcul entre chaque décision pourrait en première approche être capturé en disant que dans la clause précédente, n ne peux pas dépasser un k certain fini, indice des capacité cognitives de l'agent. N'importe quelle action survivant à toutes les compositions de k opérations de rationalisation serait acceptable selon le concept de solution correspondant.

[4]Même en faisant l'hypothèse que vous connaissez le temps que vous aurez encore à marcher à partir des différentes stations d'arrivée, le temps que le métro met pour parcourir une interstation etc.

Il ne s'agirait pas d'isoler un sous-ensemble ou un super-ensemble des équilibres classiques, mais à analyser l'impact de la limitation des ressources (en termes d'actions cognitives exécutables *entre* chaque tour) sur les concepts de solutions. Bien entendu la complexité des différentes opérations doit être analysée, pondérant ainsi le nombre de ressource que son exécution requiert. Il serait même encore mieux, d'intégrer, par exemple dans un langage épistémique, les pas effectifs que requiert l'opération (vérification, stockage, déduction). Par la suite il serait intéressant de comparer les performances de ces différentes sous-opérations cognitives sur différentes classes de jeux bien connues de la littérature.

8 Conclusion

Nous avons présenté une logique modale pour raisonner sur les jeux extensifs non-coopératifs. Nous utilisons les modalités correspondant aux relations de préférences et aux actions pour caractériser ce que nous appelons la rationalité de la décision. Nous avons également discuté la rationalité des préférences et avons montré comment ces deux notions sont liées. Nous avons également montré comment des opérations des logiques modales dynamiques peuvent être utilisés pour modéliser les actions cognitives constituant le raisonnement stratégique. Nous avons vu que cette approche conduit à des concepts de solution d'une nature plus "prospective" que ceux la littérature principale en théorie des jeux. Nous avons également suggéré qu'une notion moins globale d'action cognitive pourrait permettre d'aborder la question de concepts de solutions pour des agents cognitivement limitées, suggérant un nouveau rapprochement entre la logique modale et la théorie des jeux en rationalité limitée.

Références

[1] W Brian Arthur. Inductive reasoning and bounded rationality. *Amer. Econ. Rev.*, 84(2) :406–11, 1994.

[2] Robert Axelrod. *The Evolution of Cooperation*. Basic Books, 1985.

[3] Pierpaolo Battigalli and Giacomo Bonanno. Recent results on belief, knowledge and the epistemic foundations of game theory. *Research in Economics*, 53 :149–225, 1999.

[4] J. van Benthem and F. Liu. Dynamic logic of preference upgrade. To appear in the *Journal of Applied Non-Classical Logic*, 2007.

[5] Johan van Benthem. Rational dynamics and epistemic logic in games. to appear in International Journal of Game Theory.

[6] Johan van Benthem. Extensive games as process models. *JoLLI*, 11(3) :289–313, 2002.

[7] Johan van Benthem. Cognition as interaction. *ILLC*, PP-2005-10, 2005.

[8] Johan van Benthem, Sieuwert van Otterloo, and Olivier Roy. Preference logic, conditionals, and solution concepts in games. *ILLC*, PP-2005(28), 2005.

[9] Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic*. Cambridge, 2001.

[10] Giacomo Bonanno. Branching time, perfect information games, and backward induction. *Games and Economic Behavior*, 36 :57–73, 2001.

[11] H. P. van Ditmarsch, W. van der Hoek, and B. P. Kooi. Playing cards with Hintikka : An introduction to dynamic epistemic logic. *Austral. Journ. of Logic*, 3 :108–134, 2005.

[12] Paul Harrenstein, John-Jules Meyer, Wiebe van der Hoek, and Cees Witteveen. A modal characterization of Nash equilibrium. *Fundamenta Informaticae*, 57(2-4) :281–321, 2003.

[13] Wiebe van der Hoek and Marc Pauly. Modal logic for games and information. In Patrick Blackburn, Johan van Benthem, and Frank Wolter, editors, *The Handbook of Modal Logic*. Elsevier, 2006.

[14] Zhisheng Huang, Michael Masuch, and László Pólos. ALX, an action logic for agents with bounded rationality. *Artificial Intelligence*, 82 :75–127, 1996.

[15] John von Neumann and Oskar Morgenstern. *Theory of Games and Economic Behavior*. Princeton, 1944.

[16] Martin J. Osborne. *An Introduction to Game Theory*. Oxford, 2004.

[17] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. MIT Press, 1994.

[18] Solomon Passy and Tinko Tinchev. An essay in combinatory dynamic logic. *Information and Control*, 93(2) :263–332, 1991.

[19] Jan A. Plaza. Logics of public communications. In M. L. Emrich, M. S. Pfeifer, M. Hadzikadic, and Z. W. Ras, editors, *Proceedings of the 4th International Symposium on Methodologies for Intelligent Systems*, pages 201–216, 1989.

[20] John L. Pollock. Rational decision-making in resource-bounded agents.

[21] Ariel Rubinstein. *Modeling Bounded Rationality*. MIT Press, 1997.

[22] Herbert Simon. Theories of decision-making in economics and behavioral science. *Amer. Econ. Rev.*, 49 :253–283, 1959.

[23] Herbert Simon. *Theories of Bounded Rationality vol. 3 : Empirically Grounded Economy*. 1997.

[24] Robert Stalnaker. On the evaluation of solution concepts. *Theory and Decision*, 37 :49–73, 1994.