Planification multi-agent et diagnostic stratégique

Ramzi Ben Larbi — benlarbi@cril.fr
Sébastien Konieczny — konieczny@cril.fr
Pierre Marquis — marquis@cril.fr

CRIL - CNRS, Université d'Artois, Lens

Résumé :
Quand plusieurs agents opèrent dans un environnement commun, leurs plans peuvent interférer. Le résultat de chaque plan peut être altéré par cette interaction et la notion de plan valide de la planification classique (mono-agent) ne convient plus. Dans cet article, nous étendons ce cadre à un cadre multi-agent. Nous montrons comment les "meilleurs" plans pour un agent rationnel peuvent être caractérisés en utilisant des notions de théorie des jeux, en particulier celle d'équilibre de Nash. Nous identifions par ailleurs les scénarios pour lesquels une coopération entre agents devrait s'effectuer et montrons que nombre d'informations stratégiques peuvent être dérivées du jeu.

Mots-clés : Planification, interaction, systèmes multi-agents

Abstract:
When several agents act in a common environment their plans may interfere. The predicted outcome of each plan may be altered and the usual notion of valid plan of classical (monoagent) planning is not adequate. In this paper we extend this framework to the multi-agent case. We show how the "best" plans of rational agents can be characterized using game-theoretic notions, especially Nash equilibrium. We also identify the scenarios for which a cooperation between agents is likely to occur and show that many strategic information can be derived from the game.

Keywords: Planning, interaction, multiagent systems

1 Introduction

La modélisation de l'interaction entre agents est un domaine de recherche qui a été exploré depuis des années en économie, psychologie mais aussi intelligence artificielle. En planification classique, on calcule des plans qui, une fois exécutés, permettent à l'agent qui les a formés d'atteindre son but. Parmi les hypothèses standard de planification classique figurent le fait que l'agent connaît l'état initial du monde, chaque action possible est déterministe et son résulat peut être parfaitement prédit quel que soit l'état où elle est exécutée, les buts sont binaires (i.e. un état du monde est soit complètement satisfaisant soit complètement insatisfaisant), et le monde est statique dans le sens où la seule manière de le modifier est d'exécuter l'une des actions de l'agent (ainsi, non seulement il n'y a pas d'évènement exogène mais aussi le monde n'a pas de dynamique intrinsèque).

Dans cet article, nous étendons le cadre de la planification classique à un cadre de planification multi-agent, i.e., nous considérons un groupe d'agents. Chaque agent possède ses propres actions et buts. Les agents agissent dans un environnement commun. Dans ce cadre, les hypothèses standard de planification classique sont faites (excepté le fait que les buts ne sont pas forcément binaires). Quand plus d'un agent est considéré, de telles hypothèses de planification (en particulier, le monde statique et les actions déterministes) ne sont pas suffisantes pour permettre de prédire comment le monde va évoluer après l'exécution du plan. En effet, même si les actions restent déterministes, l'interaction entre les plans des agents introduit un surplus de complexité. Chaque agent ignore généralement quels plans les autres agents vont finalement choisir et comment ses propres actions s'intercaleront avec les leurs. Nous suggérons de pallier cela en utilisant des concepts de théorie des jeux qui permettront à l'agent de construire un diagnostic stratégique exprimant ses chances d'atteindre ses buts étant données les interactions possibles avec les autres agents. Nous supposons que l'agent connaît les buts de chaque agent du groupe, ainsi que les plans que chaque agent peut proposer. Par ailleurs, les agents peuvent aussi se coordonner, ce qui veut dire qu'ils peuvent décider de bâtir un plan commun. Dans ce cas l'incertitude causée par l'interaction est dissipée. Mais il n'est pas toujours dans l'intérêt de l'agent de se coordonner.

Exemple 1 *Deux agents, un peintre et un électricien, agissent dans une même pièce. L'ampoule doit être changée (ce qui est le but de l'électricien) et le plafond doit être peint (ce qui est le but du peintre). L'électricien a une nouvelle ampoule et le peintre a le matériel nécessaire à la peinture. Il y a une seule échelle dans la pièce (l'échelle est donc une ressource critique). De plus, le peintre a besoin de lumière pour peindre. L'électricien possède trois actions Prendre-Echelle-Electricien (PEE), Changer-Ampoule (CA), Reposer-Echelle-Electricien (REE), et le peintre trois actions : Prendre-Echelle-Peintre (PEP), Peindre (P), Reposer-Echelle-Peintre (REP). Peindre réussit seulement si Changer-Ampoule a déjà été exécuté. Prendre-Echelle-Peindre réussit seulement si l'échelle est disponible (i.e., elle a été reposée auparavant). Les interactions suivantes peuvent être facilement envisagées :*

- *si le peintre prend l'échelle en premier, il ne sera pas capable d'atteindre son but (l'ampoule doit être changée avant) ; s'il ne repose pas l'échelle, l'électricien ne sera pas capable d'atteindre son but.*
- *si l'électricien prend l'échelle en premier, il sera capable d'atteindre son but ; alors, le peintre sera capable d'atteindre son but si et seulement si l'éléctricien repose l'échelle. En conséquence, si les deux agents peuvent se coordonner pour exécuter le plan joint PEE.CA.REE.PEP.P, alors les deux agents seront satisfaits.*

L'idée de se concentrer sur des plans linéaires peut être justifiée dans ce cadre lorsque l'agent en charge de l'exécution (qui peut être différent de l'agent qui construit le plan) ne peut observer l'environnement et ainsi ne peut adapter son plan aux événements extérieurs (i.e. actions des autres agents), ou lorsqu'il peut observer l'environnement mais ne peut replanifier dynamiquement à cause d'un manque de ressources calculatoires ou la présence de contraintes temps réel (c'est le cas par exemple d'agents autonomes et mobiles comme des drones volant à grande vitesse, ou des infobots - agents logiciels - devant agir sur des marchés hautement volatiles).

Les questions clés que nous posons dans ce papier sont les deux suivantes : pour chaque agent du groupe, quels sont ses "meilleurs" plans ? Est-ce qu'un plan donné requiert une coordination afin d'être exécuté d'une manière satisfaisante pour les deux agents ? En nous concentrant principalement sur le cas de deux agents et en considérant seulement des buts binaires, nous montrons comment un jeu peut être associé à n'importe quel problème de planification multi-agent ; en conséquence, les "meilleurs" plans pour un agent rationnel peuvent être caractérisés en utilisant des notions de théorie des jeux, spécialement l'équilibre de Nash. Nous identifions aussi les scénarios pour lesquels une coopération entre agents est opportune et montrons comment plusieurs informations stratégiques peuvent être dérivées du jeu sous forme stratégique. Finalement, nous montrons que plusieurs cadres formels dans lesquels on considère l'interaction entre agents peuvent être intégrés au nôtre, incluant ceux de la planification robuste et des jeux booléens.

2 Un cadre formel pour la planification multi-agent

On considère un groupe d'agents N= $\{1, 2, \cdots, k\}$, où chaque agent est identifié par un entier. Soit S un ensemble fini non vide d'états abstraits. Notons s_0 l'état initial, supposé être l'état actuel du monde. s_0 est connu par chacun des agents de N. Chaque agent est associé à un ensemble d'actions :

Définition 1 (action) *Une action α est une application de S dans S. L'ensemble des actions de l'agent i est noté A^i.*

Dans la suite, une action sera notée par une lettre grecque $(\alpha, \beta, \cdots)$. La définition précédente impose que les actions soient déterministes et toujours exécutables. Cette dernière hypothèse n'est pas excessive. En effet, si l'on veut modéliser le fait qu'une action n'est pas exécutable si l'état du monde est s, on peut typiquement la représenter par une action qui ne change pas l'état du monde dans cet état, i.e. $\alpha(s) = s$, ou qui conduit à un état "puits", i.e., $\alpha(s) = s_\perp$, avec $s_\perp$ un état qui a la pire évaluation par rapport aux buts de l'agent et tel que $\beta(s_\perp) = s_\perp$ pour toute action β. A partir de son ensemble d'actions, chaque agent peut construire des plans :

Définition 2 ***(plan)*** *Soit A un ensemble d'actions. Un plan p sur A est une suite (possiblement vide) d'actions de A, i.e. $p = \alpha_1.\alpha_2.\cdots.\alpha_n$, où chaque $\alpha_i \in A$. Sémantiquement, c'est une application de S dans S, définie à partir de la composition de ses actions, i.e., pour toute action $\alpha \in S$, p(s) = s si p = ϵ (la séquence vide), et $p(s) = \alpha_n(\cdots(\alpha_1(s))\cdots)$ autrement. L'ensemble des plans sur A est noté A^*.*

Soit un plan $p = \alpha_1.\alpha_2.\cdots.\alpha_n$. Un sous-plan de p est une sous-suite de ses actions, i.e., $p' = \alpha_1'.\cdots.\alpha_m'$ est un sous-plan de p si et seulement si il existe une application strictement croissante t de $\{1, \cdots, m\}$ dans $\{1, \cdots, n\}$ telle que $\forall q \in \{1, \cdots, m\}$, $\alpha_a' = \alpha_{t(q)}$. Soit un autre plan $p' = \beta_1.\cdots.\beta_k$, $p.p'$ dénote la concaténation de p et p', i.e., $p.p'$= $\alpha_1.\cdots.\alpha_n.\beta_1.\cdots.\beta_k$. Enfin, si p est un plan sur A, $A(p)$ dénote le sous-ensemble de A formé par les actions de p.

Les buts d'un agent sont exprimés d'une manière qualitative au moyen d'une relation de préférence (un pré-ordre) sur l'ensemble des états : ${\preceq_G}^i \subseteq S \times S$. Ainsi, pour $s, s' \in S, s{\preceq_G}^i s'$ signifie que pour chaque agent i, l'état s' est au moins aussi préféré que s. Quand S est fini, chaque préordre ${\preceq_G}^i$ sur S peut être représenté par une fonction réelle G^i telle que pour tout $s, s' \in S$, $s{\preceq_G}^i s'$ si et seulement si $G^i(s) \leq G^i(s')$. Dans la suite, nous allons souvent nous concentrer sur le cas binaire dans lequel les états sont divisés entre états buts et états non-buts :

Définition 3 ***(but binaire)*** *On dit qu'un agent i a des buts binaires $G^i \subseteq S$ si et seulement si sa relation de préférence ${\preceq_G}^i$ est telle que $s{\preceq_G}^i s'$ si et seulement si $s' \in G^i$ ou $s, s' \notin G^i$. Nous utiliserons la notation $G^i(s) = 1$ si $s \in G^i$ et $G^i(s) = 0$ si $s \notin G^i$.*

Assez naturellement, toute relation de préférence sur les états induit une relation de préférence sur les plans :

Définition 4 ***(préférence sur les plans)*** *Soit i un agent, A^* un ensemble de plans, s_0 un état et ${\preceq_G}^i$ une relation de préférence sur les états. La relation de préférence $\leq^i$ sur A^* est définie comme suit : pour tout $p, p' \in A^*$, $p \leq^i p'$ si et seulement si $p(s_0){\preceq_G}^i p'(s_0)$.*

La qualité d'un plan est donnée par la qualité de l'état atteint ; comme en planification classique, des critères additionnels (e.g., le coût du plan) peuvent être utilisés pour discriminer les meilleurs plans ainsi définis. Dans plusieurs cas, il est raisonnable de supposer que seulement un sous-ensemble Π^i de A^{i^*} est envisagé par l'agent i ; en particulier, à cause de capacités de calcul limitées, les plans dont la longueur excède un seuil donné peuvent être éliminés. Toutefois, cela a du sens de considérer que Π^i est clos pour les sous-plans, i.e., quand un plan p appartient à Π^i, alors tout sous-plan de p appartient aussi à Π^i. En particulier, le plan vide ϵ appartient toujours à Π^i. Nous sommes maintenant prêts à définir la notion de représentation d'un agent et celle de problème de planification multi-agent :

Définition 5 ***(représentation d'un agent)*** *Chaque agent $i \in N$ est caractérisé par un triplet $\mathscr{A}^i = \langle A^i, \Pi^i, \preceq_G{}^i \rangle$ formé par un ensemble d'actions A^i, un ensemble de plans $\Pi^i \subseteq A^i$ et une relation de préférence $\preceq_G{}^i$.*

Définition 6 ***(problème de planification multi-agent)*** *Un problème de planification multi-agent (MAPP) pour un ensemble N d'agents est un triplet $\langle S, s_0, \{\mathscr{A}^i | i \in N\} \rangle$ formé par un ensemble d'états S, un état initial $s_0 \in S$ et un ensemble de représentations d'agents $\mathscr{A}_i$. Un MAPP avec buts binaires est tel que chacun des agents possède une structure de buts binaire.*

Lorsque chaque agent a choisi un plan, la suite d'évènements correspondant à leur exécution jointe est l'un de leurs mélanges, sauf si une coordination est réalisée. Nous notons $\oplus$ l'application de $A^* \times A^*$ dans 2^{A^*} qui associe à chaque paire de plans p_i et p_j, l'ensemble contenant leurs mélanges :

Définition 7 ***(mélange, ensemble de mélanges)*** *Soit $p_i = \alpha_1^i . \ldots . \alpha_n^i \in A^i, p_j = \alpha_1^j . \ldots . \alpha_p^j \in A^j$. Alors $p_i \oplus p_j$ est l'ensemble de plans p qui sont des permutations de $p_i . p_j$ pour lesquel p_i et p_j sont des sous-plans. Chaque p est appelé un* mélange *de p_i et p_j et $p_i \oplus p_j$ est appelé l'* ensemble de mélanges *de p_i et p_j.*

Exemple 2 *Reprenons l'exemple 1 avec p_1 le plan de l'éléctricien :* PEe.CA *et p_2 le plan du peintre :* PEp.P. *Alors $p_1 \oplus p_2 =$* {PEe.CA.PEp.P, PEe.PEp.CA.P, PEe.PEp.P.CA, PEp.PEe.P.CA, PEp.P.PEe.CA, PEp.PEe.CA.P}.

Observons que $\oplus$ est une fonction permutative (i.e., commutative et associative). Il s'en suit que les définitions précédentes de mélange et d'ensemble de mélanges peuvent être facilement étendues au cas où $n > 2$. Observons aussi que ϵ (la suite vide) est un élément neutre pour $\oplus$. Dans le cas déterministe avec un seul agent, évaluer un plan est une tâche facile. L'état prédit est le résultat de l'exécution du plan. Caractériser un meilleur plan est aussi facile pour l'agent considéré : le plan est d'autant meilleur que l'état atteint l'est. Dans le cas non déterministe, l'agent doit considérer tous les états possiblement atteints et agréger leurs scores afin d'évaluer le plan (plusieurs fonctions d'agrégation peuvent être utilisées, e.g. min (critère de Wald) pour traduire le comportement d'un agent pessimiste, ou un critère d'utilité espérée quand les scores sont quantitatifs et les actions non déterministes sont données par des ensembles de distributions de probabilité).

Dans le cas multi-agent (quoique déterministe), qui est le cas étudié dans cet article, la situation est similaire à celle du cas non-déterministe à un agent dans le sens où chaque agent doit considérer tous les états possiblement atteints pour évaluer ses plans. La différence principale vient de la nature de l'incertitude : dans notre cadre, l'incertitude vient de l'interaction avec les plans fournis par les autres agents. En conséquence, chaque agent doit exploiter le fait qu'il connaît les représentations des autres agents (il connaît les buts des agents ainsi que leurs plans) afin de déduire quel est son "meilleur" plan. Il diffère en cela du cas non déterministe où l'incertitude vient de l'impossibilité de prédire précisement le résultat de certaines actions, comme "tirer à pile ou face". Dans plusieurs cas, une telle impossibilité résulte d'évènements extérieurs (sur lesquels notre connaissance est imparfaite), qui ne peuvent être totalement observés ou prédits et qui ont un certain effet sur le monde. Par exemple, dans le cas de planification des mouvements d'un robot, l'effet normal de l'action "avancer(1m)" est d'avancer le robot d'un mètre ; toutefois, il se peut que cet effet normal ne se produise pas : si le sol est mouillé (et que cela ne puisse pas être observé), un effet exceptionnel de "avancer(1m) sera d'avancer le robot de 0.5 mètre, seulement. Toutefois, dans la section 5, nous expliquerons comment la planification robuste, qui traite le problème de trouver un plan robuste dans un cadre non déterministe, peut être exprimée dans notre cadre.

Exemple 3 *Si le peintre dans l'exemple 1 propose le plan* p = PEp.P.REp *, il est seulement assuré que les actions de* p *seront exécutés dans l'ordre désiré. Alors qu'il connaît la représentation de l'électricien, il ne sait pas quel plan l'électricien va proposer (en effet, l'ensemble des plans possibles n'est en général pas un singleton). Même si cet ensemble est un singleton, le peintre ignore encore l'ordre d'exécution, i.e., comment son plan va s'intercaler avec celui de l'électricien. Supposons que l'électricien propose le plan* p' = PEe.CA.REe*, le plan joint qui va être finalement exécuté peut être n'importe quel plan de* $p \oplus p'$*. L'incertitude résultante disparait dès que les deux agents se coordonnent pour exécuter un plan commun comme* p'' = PEp.P.REp.PEe.CA.PEe.

Dans notre approche, une tâche capitale pour chaque agent est celle d'évaluer l'interaction de ses plans avec ceux des autres agents. Formellement, cela requiert l'évaluation de chaque ensemble de mélanges. A cette fin, nous associons à chaque ensemble de mélanges son profil de satisfaction (SP), qui est une vue résumée et abstraite de l'évaluation des mélanges pour tous les agents du groupe. Expliquons comment construire un profil de satisfaction pour un groupe de deux agents ayant des buts binaires. Etant donnée un couple de plans, $p_i \in \Pi^i$ et $p_j \in \Pi^j$, chaque mélange de l'ensemble de mélanges $p_i \oplus p_j$ est un plan construit à partir des actions des deux agents ; l'exécution d'un tel plan conduit à un état spécifique qui est plus ou moins satisfaisant pour chaque agent. L'évaluation d'un plan dépend de l'état résultant de son exécution. On peut représenter l'évaluation d'un ensemble de mélanges par les agents en utilisant une représentation sur 2 axes (chaque axe exprime la satisfaction de l'agent correspondant) qui associe un point de coordonnées (x,y) à un mélange p ssi $G^i(p(s_0)) = x$ et $G^j(p(s_0)) = y$. Notons qu'une telle représentation peut être facilement généralisée au cas de n agents.

Définition 8 ***(profil de satisfaction)*** *Soit un MAPP avec buts binaires pour un ensemble* $N = \{1, \ldots, m\}$ *d'agents, avec un état intial* s_0*. Un* profil de satisfaction (SP) *pour l'ensemble de mélanges* $p_1 \oplus p_2 \oplus \ldots \oplus p_k$ *(*$p_i \in \Pi^i$ *pour* $i \in \{1, \ldots k\}$*) est un ensemble* $SP(p_1 \oplus p_2 \oplus \ldots \oplus p_k)$ *de vecteurs* $(x_1, \ldots, x_k)$ *vérifiant* $(x_1, \ldots, x_k) \in SP(p_1 \oplus p_2 \oplus \ldots \oplus p_k)$ *si et seulement si* $\exists p \in p_1 \oplus p_2 \oplus \ldots \oplus p_k$ *tel que* $\forall i \in \{1, \ldots k\}$*,* $G^i(p(s_0)) = x_i$*.*

Quand nous considérons seulement deux agents ayant des buts binaires, les profils de satisfaction possibles est décrit dans la figure 1.

De nombreuses conclusions peuvent être tirées à partir de tels SPs. Ainsi, quelques

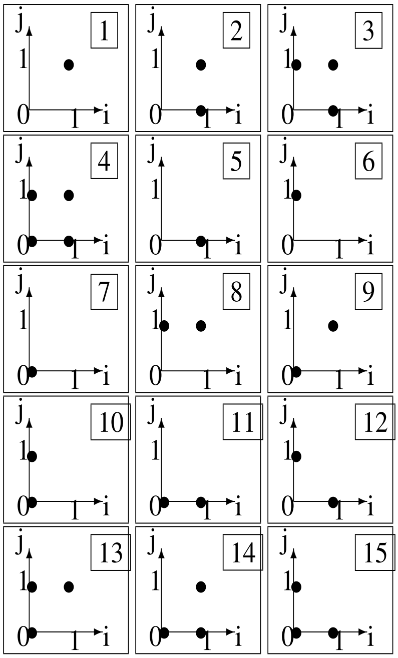

FIG. 1 – SPs possibles dans le cas binaire avec deux agents

SPs sont clairement meilleurs pour un agent donné que d'autres. Clairement, SP 2 dans lequel les mélanges conduisent seulement à des états que l'agent i évalue à 1, est plus intéressant pour lui que SP 10, dans lequel les mélanges conduisent toujours à des états non buts pour cet agent. De plus, considérons SP 3 : pour chacun des deux agents, au moins l'un des mélanges conduit à un mauvais état (i.e., un état non but), et au moins l'un des mélanges conduit à un état but. Cet SP montre aussi l'exsistence d'au moins un mélange gagnant-gagnant (évalué comme le vecteur (1,1)). Dans un tel cas, si les deux agents sont rationnels (i.e., ils agissent pour changer le monde vers un état but), alors il devraient se coordonner pour exécuter un tel mélange. En effet, la coordination est un moyen d'éliminer l'incertitude. Si les deux agents i et j proposent deux plans $p_i \in A^{i*}$ et $p_j \in A^{j*}$ de manière indépendante, ils courent le risque que l'exécution jointe de $p_i \oplus p_j$ conduise à un état évalué à (0,1) ou (1,0), auquel cas, l'un des agents sera insatisfait. A l'inverse, s'ils se coordonnent et proposent conjointement un plan correspondant à un mélange gagnant-gagnant, ils auront la garantie d'être tous deux satisfaits. Dans une situation correspondant au SP 3, les deux agents ont intérêt à offrir (et accepter) une coordination. En l'absence de plus d'information (comme une distribution de probabilité sur l'ensemble des mélanges), cela a un sens de classer les SPs sur une échelle ordinale. Prenons pour cela le point de vue de l'agent i et montrons comment les SPs peuvent être rassemblés et ordonnés :

– **Toujours Satisfait** SP 1,2,5. Pour ces SPs, l'agent i est assuré d'atteindre ses buts même si l'agent j n'accepte aucune coordination. C'est le cas le plus favorable pour i.
– **Intérêt Mutuel** SP 3,4,9,13,14. Pour chacun de ces SPs, une certaine exécution jointe est bénéfique et d'autres non (pour les deux agents), mais ils partagent tous le vecteur (1,1), ce qui signifie que si les deux agents se coordonnent, il peuvent tous deux atteindre leurs buts.
– **Dépendance** SP 8,11. Pour ces SPs, l'évaluation des mélanges par l'autre agent est constante. Cela signifie que, a priori, il n'y a aucune raison pour l'autre agent d'accepter une coordination afin d'aider l'agent i à atteindre son but.
– **Antagonisme** SP 12,15. Ces SPs reflètent des situations plus problématiques étant donné que les intérêts des deux agents sont clairement distincts. Cela signifie que si l'un est satisfait, alors l'autre ne l'est pas (i.e. la coordination (1,1) n'est pas une option). Dans de tels cas, l'agent i peut juste espérer que l'exécution jointe lui sera favorable.
– **Toujours Insatisfait** SP 6,7,10. Quelle que soit la suite des évènements, l'agent i sera insatisfait (aucune exécution jointe ne permet à l'agent d'atteindre son but). De tels SPs sont clairement les pires pour l'agent i.

Notre thèse est que, en l'absence d'information supplémentaire, une telle classification est la plus rationnelle. Par conséquent, nous considérons que chaque agent possède les préférences suivantes sur les évaluations des ensembles de mélanges :

Toujours Satisfait> Intérêt Mutuel > Dépendance > Antagonisme> Toujours Insatisfait

X>Y signifie que les SPs de la classe X sont strictement préférés à ceux de la classe Y. Tous les SPs d'une même classe sont indifférents. On peut facilement encoder un tel préordre total en utilisant des nombres. Ainsi, nous écrivons $e_i(p_i \oplus p_j) = 4$ si et seulement si $SP(p_i \oplus p_j) \in$ **Toujours Satisfait(i)**, $\cdots$, $e_i(p_i \oplus p_j) = 0$ si et seulement si $SP(p_i \oplus p_j) \in$ **Toujours Insatisfait(i)** (voir table1).

Classe	Evaluation
Toujours Satisfait	4
Intérêt Mutuel	3
Dépendance	2
Antagonisme	1
Toujours Insatisfait	0

TAB. 1 – Evaluation des SPs

De telles évaluations $e_i(p_i \oplus p_j)$ peuvent grossièrement être vues comme des utilités, mais elles ne dépendent pas seulement des buts de l'agent i. Notons aussi que les nombres utilisés importent peu, seul l'ordre compte (notre cadre n'est pas quantitatif). Notons finalement que, alors que les définitions à venir vont utiliser des évaluations $e_i(p_i \oplus p_j)$ et $e_j(p_i \oplus p_j)$, elles ont encore du sens quand d'autres évaluations sont utilisées. Ainsi, si quelqu'un est en désaccord avec l'échelle proposée, les définitons suivantes s'appliquent toujours (tant que l'on utilise une évaluation qui permet de comparer tous les couples de plans).

3 Résolution du jeu et génération de diagnostic stratégique

A partir de la construction précédente, nous sommes maintenant capables d'associer à chaque mélange une évaluation pour chaque agent. Ceci nous permet de modéliser l'interaction entre les plans des agents comme un jeu (non-coopératif) sous forme stratégique. En faisant cela, on peut utiliser deux concepts de solutions pour ces jeux : ceux de niveau de sécurité et d'équilibre de Nash. En effet, à chaque MAPP à buts binaires pour un ensemble de deux agents N=$\{i, j\}$, on peut associer un jeu sous forme stratégique, défini par l'ensemble N de joueurs, l'ensemble de stratégies pour chaque joueur (les ensembles Π^i et Π^j de plans dans notre cas), et par une fonction d'évaluation pour chaque joueur qui associe une évaluation à chaque profil de stratégies (les évaluations $e_i(p_i \oplus p_j)$ et $e_j(p_i \oplus p_j)$ pour chaque ensemble de mélanges $p_i \oplus p_j$ dans notre cas).

Exemple 4 *Considérons le MAPP suivant :* $\langle S, s_0, \{\mathscr{A}^i \mid i \in 1, 2\}\rangle$. $\mathscr{A}^1 = \langle A^1, \Pi^1 = \{p_1, p'_1\}, \preceq^1_G\rangle$. $\mathscr{A}^2 = \langle A^2, \Pi^2 = \{p_2, p'_2\}, \preceq^2_G\rangle$. *Supposons qu'il en résulte le SP de la figure 2 :*

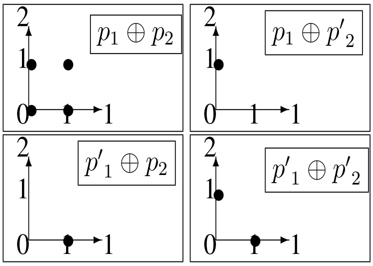

FIG. 2 – Exemple de SPs

On peut maintenant associer un MAPP avec le jeu suivant sous forme stratégique de la table 2

Une première analyse qu'un agent peut faire est basée sur la notion de niveau de

	p_2	p'_2
p_1	(3,3)	(0,4)
p'_1	(4,0)	(1,1)

TAB. 2 – Jeu associé

sécurité de ses plans.

Définition 9 ***(niveau de sécurité d'un plan)*** *Etant donné un MAPP avec buts binaires pour $N = \{1, 2\}$, le* niveau de sécurité *d'un plan p_i d'un agent i ($i \in N$) face à un ensemble Π^j de plan de l'agent j ($j \neq i$) est défini comme l'évaluation minimum de l'ensemble de mélanges entre le plan p_i et un plan du joueur j, i.e.,*

$$S_{\Pi^j}(p_i) = \min_{p_j \in \Pi^j} e_i(p_i \oplus p_j).$$

A partir des niveaux de sécurité des plans d'un agent on peut définir le niveau de sécurité de l'agent :

Définition 10 ***(niveau de sécurité d'un agent)*** *Etant donné un MAPP avec buts binaires $N = \{1, 2\}$, le* niveau de sécurité *de l'agent i face à l'ensemble Π^j de plans de l'agent j est le plus grand niveau de sécurité des plans de l'agent i, i.e.,*

$$S_{\Pi^j}(i) = \max_{p_i \in \Pi^i} S_{\Pi^j}(p_i).$$

Une solution au jeu associé à un MAPP donné peut être définie comme un couple de plans $\langle p_1 \in \Pi^1, p_2 \in \Pi^2\rangle$ telle que p_1 (resp. p_2) maximise le niveau de sécurité de l'agent 1 (resp 2) face à Π^2 (resp. Π^1). Une telle solution a du sens dans notre cadre étant donné qu'elle peut être vue comme une analyse au pire cas de l'interaction stratégique. En effet, les SPs sont des ensembles de vecteurs de satisfaction possibles, et comme la classification des SPs que nous avons fournie repose sur une analyse au pire cas, il semble raisonnable d'utiliser les niveaux de sécurité pour comparer des mélanges. Toutefois, les niveaux de sécurité ne prennent pas en compte toutes les opportunités offertes aux agents. Une notion de solution beaucoup plus largement acceptée est basée sur la notion d'équilibre de Nash.

Définition 11 ***(équilibre de Nash)*** *Etant donné un MAPP avec buts binaires pour $N = \{i, j\}$, un couple de plans $\langle p_i \in \Pi^i, p_j \in \Pi^j\rangle$ est un équilibre de Nash si aucun des agents ne peut avoir une meilleure évaluation en choisissant un autre plan, i.e., $\langle p_i, p_j\rangle$ est un équilibre de Nash si et seulement si $\nexists p \in \Pi^i$ s.t $e_i(p \oplus p_j) > e_i(p_i \oplus p_j)$ et $\nexists p \in \Pi^j$ s.t. $e_j(p_i \oplus p) > e_j(p_i \oplus p_j)$.*

Exemple 5 *Revenons au jeu donné à la table 2. Considérons le couple $\langle p'_1, p'_2\rangle$. L'agent 1 n'a aucun intérêt à dévier seul de ce couple. En effet, $\langle p_1, p'_2\rangle$ le conduit à une situation moins favorable ($e_1(p_1 \oplus p'_2) < e_1(p'_1 \oplus p'_2)$). De même, $\langle p'_1, p_2\rangle$ est clairement moins favorable à l'agent 2 que $\langle p'_1, p'_2\rangle$. Ainsi, on peut conclure que $\langle p'_1, p'_2\rangle$ est un équilibre de Nash. Il est facile de vérifier que c'est le seul du jeu.*

Comme dans le cas général en théorie des jeux, il se peut dans notre cadre qu'un jeu n'ait pas d'équilibre de Nash en stratégie pure [**?**], ou qu'il y en ait plusieurs. Quand il y a plusieurs équilibres de Nash, d'autres critères, comme la Pareto optimalité [1], sont souvent utilisés pour les différentier. Les propositions suivantes donnent deux conditions suffisantes à l'existence d'un équilibre de Nash.

Proposition 1 *Considérons un MAPP avec buts binaires et deux agents 1 et 2 tel que $G^1 = G^2$. Alors le jeu associé exhibe un équilibre de Nash.*

[1] Un vecteur Pareto domine un autre si chacune des composantes du premier est supérieure ou égale à la composante correspondante du second

Autrement dit, si les deux agents partagent les mêmes buts et s'il existe un plan formé sur l'ensemble de leurs actions qui permette d'y parvenir, alors le modèle présenté retient ce plan comme solution.

Proposition 2 *Considérons un MAPP avec buts binaires pour deux agents* 1 *et* 2. *Notons* $G^{1,+}$ *(resp.* $G^{2,+}$*) le sous-ensemble* G^1 *(resp.* G^2*) des états atteignables en utilisant des plans sur* A^1 *(resp.* A^2*) et* $G^{1,2,+}$ *(resp.* $G^{2,1,+}$*) le sous-ensemble de* G^1 *(resp.* G^2*) des états atteignables en utilisant des plans sur* $A^1 \cup A^2$. *Si* $G^{1,+} = G^{2,+} = \emptyset$ *et* $G^{1,2,+} = G^{2,1,+} \neq \emptyset$, *alors le jeu associé au MAPP exhibe un équilibre de Nash.*

Notons que, dans notre cadre, le dilemne des prisonniers, un jeu particulier largement étudié[**?**, **?**], peut aussi être atteint.

Exemple 6 *Considérons encore une fois l' exemple 4. Le jeu associé (table 2) exhibe une situation de "dilemne du prisonnier".* $\langle p'_1, p'_2\rangle$ *est un équilibre de Nash. Le couple* $\langle p_1, p_2\rangle$ *qui est plus profitable que* $\langle p'_1, p'_2\rangle$ *pour les deux agents n'est pas un équilibre de Nash (chaque agent est tenté d'utiliser un autre plan).*

Au delà de la notion de solution, chacun des deux agents i et j considérés dans le MAPP peut dériver beaucoup d'informations stratégiques à partir du jeu sous forme stratégique associé. Concentrons nous sur les notions de plans robustes, d'effets de synergie, et d'indépendance. Un plan p_i pour l'agent i est *robuste* par rapport à l'agent j si et seulement si son exécution jointe avec n'importe quel plan de l'agent j lui assure d'atteindre son but. Dans le jeu sous forme stratégique, un tel plan correspond à une ligne (ou colonne) pour laquelle toutes les évaluations pour cet agent sont égales à 4 : $\forall p_j \in \Pi^j, e_i(p_i \oplus p_j) = 4$. Assez clairement, un tel plan maximise le niveau de sécurité de l'agent i. Si un plan robuste existe pour un agent i, alors aucune coordination n'est nécéssaire avec l'agent j. L'existence d'une synergie entre deux agents peut assez facilement être déduite du jeu sous forme stratégique. En effet, un *effet synergétique* pour les agents i et j est possible si et seulement si il existe $p_i \in \Pi^i$ et $p_j \in \Pi^j$ tel que $e_i(p_i \oplus p_j) > max_{p\in\Pi^i} e_i(p)$ et $e_j(p_i \oplus p_j) > max_{p\in\Pi^j} e_j(p)$. Assez clairement, aucun effet synergétique n'est possible quand au moins l'un des agents possède un plan robuste. La proposition suivante donne une condition suffisante pour assurer qu'un couple de plans $\langle p_1, p_2\rangle$ exhibant un effet synergétique pour les deux agents 1 et 2 soit aussi une solution du jeu :

Proposition 3 *Considérons un MAPP avec buts binaires et deux agents* 1 *et* 2. *Supposons que* $\exists p \in (A^1 \cup A^2)^*$ *satisfaisant* $G^1(p(s_0)) = 1$ *et* $G^2(p(s_0)) = 1$ *et* $\forall p' \in (A^1 \cup A^2)^*$, $p' \neq p \Rightarrow (G^1(p'(s_0)) = 0$ *et* $G^2(p'(s_0)) = 0)$. *Soient* $p_1 \in \Pi^1, p_2 \in \Pi^2$. *Si* $p \in p_1 \oplus p_2$ *alors* $\langle p_1, p_2\rangle$ *est un équilibre de Nash du jeu associé.*

Une notion d'indépendance entre agents, reflétant le fait qu'il n'y a pas d'interaction entre leurs plans, peut aussi être facilement dérivée du jeu sous forme stratégique. En effet, les deux agents sont indépendants si et seulement si $\forall p_i \in \Pi^i, \forall p_j \in \Pi^j, e_i(p_i \oplus p_j) = e_i(p_i \oplus \epsilon)$ et $e_j(p_i \oplus p_j) = e_j(\epsilon \oplus p_j)$.

4 Exemple : le pont

On considère deux agents 1 et 2. L'agent 1 est en position a et l'agent 2 est en position b. Afin d'aller de a à c, l'agent 1 doit traverser le pont, qui doit être ouvert auparavant (action B^1). Il en va de même pour 2 (action B^2). Si le pont est ouvert pour 1, il est fermé pour 2 et inversement. Chaque agent a une action C^i lui permettant de traverser le pont (mais requiert que le pont soit ouvert pour réussir), i.e., C^1 change la position de l'agent 1 de a à c. L'agent 1 a une action supplémentaire J^1 qui lui pemet de sauter par dessus le pont. En utilisant cette action, l'agent 1 n'a pas besoin d'ouvrir le pont. Cette action conduit directement l'agent 1 à la position c.

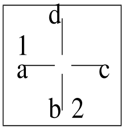

FIG. 3 – Traverser le pont

	ϵ	B^2C^2
ϵ	$0,0$	$0,4$
J^1	$4,0$	$4,4$
B^1C^1	$4,0$	$3,3$
B^1J^1	$4,0$	$4,2$
C^1J^1	$4,0$	$4,4$
J^1B^1	$4,0$	$4,2$
J^1C^1	$4,0$	$4,4$

TAB. 3 – Jeu sous forme stratégique (simplifié)

Clairement, l'agent 1 doit exécuter le plan $B^1.C^1$ ou le plan J^1 afin d'atteindre la position désirée ; l'agent 2 doit exécuter $B^2.C^2$. Etant donné qu'un agent ne peut ouvrir le pont pour un autre, les plans qui ne contiennent pas l'un de ces sous-plans ne peuvent conduire à un état but. Si chaque agent était tout seul dans cet environnement, le problème de planification serait facilement résolu étant donné qu'un agent serait alors sûr d'atteindre son but ($B^i.C^i$ permet d'atteindre le but de i). Ce n'est plus la même histoire lorsque les deux agents agissent conjointement. En effet, dans ce cas, une coordination est nécessaire : si l'exécution jointe est $B^1.B^2.C^1.C^2$ l'agent 1 ne pourra pas traverser le pont et ne pourra pas atteindre son but. Représentons le jeu sous forme stratégique associé à ce problème de planification multi-agent. Nous restreignons la longueur des plans examinés à deux actions (observons que les plans de longueur supérieure à 2 sont inutiles). Ce jeu peut être simplifié en supprimant les plans nuls (un plan nul est un plan qui conduit à une satisfaction de 0 quel que soit le mélange dans lequel il est impliqué). Toutefois, nous gardons le plan ϵ dans la version simplifiée, même lorsque c'est un plan nul (voir Table 3).

Avec n'importe lequel des plans $J^1.B^1$, $J^1.C^1$, J^1, $B^1.J^1$ ou $C^1.J^1$, l'agent 1 a un niveau de sécurité de 4. Comme J^1 est un sous-plan de tous ces plans, les autres plans incluent des actions inutiles. L'agent 1 va probablement choisir le plan J^1. Pour l'agent 2, le seul plan dont le niveau de sécurité est non nul est $B^2.C^2$ (tous les autres plans sont des plans nuls, ils ne peuvent donc conduire à un état but). Ainsi, dans cette situation, le résultat probable du jeu sera le couple de plans $\langle J^1, B^2.C^2\rangle$ qui est évaluée à 4 par chaque agent, ce qui signifie que les deux agents vont sûrement atteindre leurs buts et que cette situation stratégique ne requiert aucune coordination. L'agent 1 peut aussi choisir $B^1.J^1$ au lieu de J^1. Ces deux plans sont pareillement évalués par l'agent 1. Cependant, avec $B^1.J^1$ l'agent 1 peut s'assurer que le plan de l'agent 2 obtiendra une plus faible évaluation (2 au lieu de 4) face à J^1. Si l'agent 1 choisit $B^1.J^1$, il exhibe un comportement agressif par rapport à l'agent 2. On ne développera pas ce point dans la suite, mais il est intéressant d'observer que de telles attitudes peuvent être modélisées dans notre cadre. Les équilibres de Nash de ce MAPP correspondent ici exactement aux solutions obtenues en utilisant la notion de niveau de sécurité.

5 Généralité du cadre

Nous allons voir dans ce paragraphe que plusieurs cadres formels dans lesquels on considère l'interaction entre agents peuvent être facilement vus comme des cas particuliers du nôtre. Il s'agit de la planification robuste et des jeux booléens.

5.1 Planification robuste

En planification robuste (see e.g. [**?**]), le but est de déterminer si une suite d'actions (i.e., un plan) est robuste, i.e., s'il permet d'atteindre le but pour toutes les contingences possibles.

Définition 12 ***(planification robuste)***
- *Une* action non déterministe α *sur un ensemble fini et non vide S d'états est une application de S dans $2^S \setminus \{\emptyset\}$.*
- *Un* plan non déterministe π *sur un ensemble A d'actions non déterministes est une suite finie d'éléments de A.*
- *Une* trajectoire *pour un plan non déterministe $\pi = \alpha_1.\cdots.\alpha_n$ étant donné un état initial $s_0 \in S$ est une suite d'états $s_O, \ldots, s_{n+1}$ telle que pour tout $i \in 0 \ldots n$, $s_{i+1} \in \alpha_i(s_i)$.*
- *Un plan non déterministe $\pi = \alpha_1.\cdots.\alpha_n$ sur A est* robuste *pour un but $G \subseteq S$ étant donné un état initial $s_0 \in S$ si et seulement si pour chaque trajectoire $s_O, \ldots, s_{n+1}$ pour π, $s_{n+1} \in G$.*

Le problème de la planification robuste peut être facilement exprimé dans notre cadre. Le codage est assez technique, donnons-en simplement le principe : l'idée est de considérer chaque trajectoire du plan non déterministe considéré π comme le résultat d'un mélange avec le plan d'un second agent qui joue le rôle de la nature ; considérons la première action α de π et supposons qu'elle possède au plus k effets ; dans ce cas, le plan du second agent va débuter par le sous-plan $\alpha'_1, ..., \alpha'_k$, où α'_j est l'action vide si α n'a pas été exécutée lorsque α'_j est rencontrée (information que l'on mémorise dans les états via un fluent supplémentaire), et produit le $k$ème effet de α sinon ; il reste essentiellement à répéter ce traitement pour les actions suivantes de π en mettant à jour le plan du second agent par concaténation avec les sous-plans produits à chaque étape.

5.2 Jeux booléens

Les jeux booléens (see e.g. [**?**, **?**]) traitent le cas d'agents contrôlant un ensemble de variables (binaires) propositionnelles. Plus précisément, ce sont des jeux où les utilités des agents sont binaires et les buts sont spécifiés par des formules propositionnelles.

Définition 13 ***(jeu booléen)*** *Un* jeu booléen *est un quadruplet $\langle A, V, \Pi, \Phi\rangle$ où $A = \{1 \cdots n\}$ est un ensemble d'agents, V est un ensemble de variables propositionnelles (variables de décision), $\Pi : A \to 2^V$ une fonction d'assignation qui induit une partition $\{\pi_1, \cdots, \pi_n\}$ de V où π_i est l'ensemble de variables contrôlées par l'agent i, $\Phi = \{\phi_1 \cdots \phi_n\}$ un ensemble de formules propositionnelles.*

Pour un joueur $i \in A$, une stratégie est une instanciation des variables qu'il contrôle (i.e., une application de $\Pi(i) = \pi_i$ dans $\{0, 1\}$). Un profil de stratégies P consiste en l'instanciation de toutes les variables considérées et peut être vu comme une application de V dans $\{0, 1\}$). Un agent i est satisfait par un profil de stratégies P si et seulement si P est un modèle de Φ_i. On peut exprimer ce cadre dans le nôtre en associant à chaque variable $v \in V$ une action v^+ qui affecte la variable v à 1. A chaque jeu booléen $G = \langle A, V, \Pi, \Phi\rangle$ nous associons un MAPP $\langle S, s_0, \{\mathscr{A}^i \mid i \in 1 \cdots n\}\rangle$ où P est l'ensemble de toutes les affectations de V, s_0 est l'affectation telle que $s_0(v) = 0$ pour tout $v \in V$. Pour chaque agent i, $A^i = \{v^+ \mid v \in \pi_i\}$, Π^i est le sous-ensemble de plans de A^{i^*} tels que chaque action possède au plus une seule occurence dans chaque plan et G^i est l'ensemble des modèles de ϕ_i.

6 Conclusion

Dans ce travail, nous avons proposé un cadre pour modéliser des problèmes de planification multi-agents. Ce cadre nous permet de former diverses conclusions stratégiques à propos d'interactions spécifiques et nous permet de "résoudre" de nombreuses situations. Ce travail ouvre de nombreuses perspectives. L'une d'elles consiste à ajouter des coûts aux actions, comme dans certains problèmes de planification. Dans ce cas, l'objectif principal de chaque agent est d'atteindre un état but et un objectif auxiliaire est de dériver un plan de coût minimal. Une autre extension est de considérer plus en profondeur le cas de n agents (n>2), et de rechercher les coalitions possibles dans ce cadre.

Si de nombreux travaux ont été consacrés à la planification multi-agent, on y suppose souvent que les agents partagent un certain nombre de buts. Relâcher cette hypothèse a un impact majeur sur les approches possibles du problème et appelle à des notions provenant de la théorie des jeux.

Une approche comparable à la nôtre est décrite dans [**?**]. Dans ce papier, les politiques sont évaluées au niveau du groupe par rapport à chaque agent et les "meilleures" sont caractérisés comme des équilibres de Nash, comme c'est le cas dans notre travail. Cette approche est néanmoins différente de la nôtre par de nombreux aspects :

- le cadre formel considéré est celui de la planification sous incertitude et observabilité totale et non celui de la planification classique. Des actions non déterministes sont considérées et un ensemble d'états initiaux possibles (et non un seul état) est connu par chaque agent. Les politiques sont des applications associant des actions à des états et non des plans linéaires (suites d'actions), et la qualité d'un plan n'est pas binaire par essence (à l'inverse de ce qui se passe dans le cadre classique).
- les politiques au niveau du groupe font partie de l'entrée du problème mais les politiques au niveau des agents ne le sont pas (alors que les plans possibles au niveau du groupe sont caractérisés comme des mélanges de plans au niveau des agents dans notre travail).
- enfin, aucune notion de diagnostic stratégique n'est abordée (en particulier, le besoin de coordination ne peut être déduit de l'entrée considérée).

Remerciements

Merci aux relecteurs pour leurs remarques avisées. Les auteurs ont bénéficié du soutien de la Région Nord/Pas-de-Calais, de l'IRCICA et du programme FEDER de la Communauté Européenne.

Références

[1] R. Axelrod. *The Evolution of Cooperation*. Basic Books, New York, USA, 1984.

[2] B. Beaufils, J.-P. Delahaye, and Ph. Mathieu. Complete classes of strategies for the classical iterated prisoner's dilemma. In *Proc. of EP'98*, pages 33–41, 1998.

[3] E. Bonzon, M.-C. Lagasquie-Schiex, J. Lang, and B. Zanuttini. Boolean games revisited. In *Proc. of ECAI'06*, pages 265–269, 2006.

[4] M.H. Bowling, R.M. Jensen, and M.M. Veloso. A formalization of equilibria for multiagent planning. In *Proc. of IJCAI'03*, pages 1460–1462, 2003.

[5] A. Cimatti and M. Roveri. Conformant planning via model checking. In *Proc. of ECP'99*, pages 21–34, 1999.

[6] P. Harrenstein. *Logic in Conflict*. PhD thesis, Utrecht University, 2004.

[7] J.F Nash. Equilibrium points in n-person games. *Proc. of the National Academy of Sciences of the USA*, 36(1) :48–49, 1950.