
Modélisation des satisfactions personnelle et interactive d'agents situés coopératifs

Olivier Simonin - Jacques Ferber

LIRMM Université Montpellier II - CNRS
161, rue Ada - 34392 Montpellier Cedex 5 - France
{simonin,ferber}@lirmm.fr

RÉSUMÉ. Cet article aborde le problème de l'évaluation, par l'agent, de ses actions et interactions. Nous proposons un modèle des satisfactions différenciant actions individuelles et interactions avec les agents voisins. La satisfaction personnelle est calculée incrémentalement dans le temps suivant la perception de la progression de la tâche en cours. La satisfaction interactive est une évaluation de la présence des agents voisins et de leurs actions: gêne, coopération, indifférence. Cette évaluation est transformée en signaux d'intentions (attractions ou répulsions) agissant comme des champs de potentiels dynamiques. Nous proposons alors un module de sélection d'actions basé sur la compétition continue entre la satisfaction personnelle de l'agent et les signaux qu'il perçoit. Nous montrons comment l'application de cette architecture permet de résoudre des conflits spatiaux entre agents mobiles autonomes.

ABSTRACT. This paper deals with the problem of valuation by the agent of its actions and interactions. The agent satisfaction has been modeled into two aspects: the individual tasks performance and the quality of interaction with its neighborings agents. The personal satisfaction is computed by the perception of the current task progression. The interactive satisfaction is an estimation of the neighbors actions effects: hindrance, cooperation, indifference. This perception is transformed into intentional signals (attractions or repulsions) operating like dynamical potentials fields. Thus, we propose an action selection module based on the continuous competition between the personal satisfaction and the perceived signals. We show how the application of this architecture can solve spatial conflicts between autonomous mobile agents.

MOTS-CLÉS : agents situés, coopération/coordination réactive, sélection d'actions, modèle des satisfactions

KEYWORDS: situated agents, reactive cooperation/coordination, action selection, satisfactions model

1. Introduction

La résolution d'un problème distribué consiste souvent à le ramener à un problème multi-agent strict, c'est-à-dire qui peut s'exprimer par un ensemble de fonctions locales à satisfaire (c'est par exemple le principe de l'Eco-résolution [BUR 91]). Ces fonctions sont rattachées à des agents qui interagissent localement. Ainsi, chaque agent est un être qui cherche à se satisfaire, ou inversement qui cherche à éviter les frustrations. Le problème consiste alors à définir les comportements des agents pour que l'ensemble des fonctions de satisfaction locales soient maximisées.

Pour des agents situés, le problème revient alors à correctement définir les interactions entre agents. Il faut en effet que chaque agent puisse agir/communiquer en fonction des besoins et des actions de ses voisins. Ceci permet de définir des processus de coordination, de coopération ou de compétition. Généralement, ce problème est abordé au niveau de la sélection d'actions [FER 95]. Les architectures développées depuis une dizaine d'années se concentrent sur l'optimisation de la tâche/action à déclencher suivant les perceptions de l'agent et ses états/expériences (cf. [PIR 99]). Cependant, une fois la tâche/action déclenchée, sa progression, au sens *qualitatif*, est difficile à modéliser et donc à prendre en compte pour le contrôle des actions et des interactions de l'agent.

Nous abordons dans cet article le problème de la modélisation des satisfactions de l'agent et de son utilité dans la définition et la gestion des comportements. Notre approche consiste à considérer d'abord l'agent -individuellement-, puis avec -son voisinage- et enfin dans -le système-. Cette démarche est motivée par la conception d'agents coopératifs, qui sont dirigés par deux tendances principales: satisfaire leurs objectifs propres et rechercher des interactions optimales avec les autres agents.

L'article présente dans un premier temps une définition de la satisfaction individuelle, dont nous proposons en section 2 une mesure incrémentale. Puis nous montrons en section 3 que la coordination et la coopération entre agents nécessitent l'introduction d'une satisfaction complémentaire, qui tient compte uniquement des actions des autres agents. La section 4 présente le principe de l'architecture réactive développée sur ces deux types de satisfactions, puis les résultats obtenus sur des problèmes de SMA.

2. Satisfaction personnelle: progression des tâches

2.1. Tâche et satisfaction

La plupart des travaux, théoriques ou pratiques, menés sur la conception d'agents situés, cherchent à maximiser la qualité de réalisation des tâches et à minimiser leur coût en temps. Les solutions optimales en temps peuvent être obtenues par des approches théoriques mais elles ne sont pas toujours applicables. En effet, par nature, ces solutions ne prennent pas en compte la totalité des paramètres des situations réelles et de leur caractère dynamique (cf. [PIR 99]). Par exemple, le calcul d'un chemin op-

timal entre deux points de l'environnement peut se révéler à l'exécution inadapté à la nature du terrain, obsolète face à la réalité des obstacles, etc.

Aussi, une bonne solution à un problème n'est pas nécessairement la plus optimale en temps (ou en énergie) mais celle qui garantit un résultat. Si l'on considère des agents autonomes évoluant dans un environnement *dynamique et inconnu*, la notion de temps perd de son importance au profit des capacités de réalisation et d'adaptation des agents. Dans ce type de système, un agent risque plus souvent d'être confronté à une situation inconnue qu'à une situation idéale où l'exécution des opérations peut être considérée comme *déterministe*.

A partir de ce constat, nous définissons une première satisfaction, liée uniquement à la tâche courante de l'agent, nommée satisfaction personnelle.

Définition (*Satisfaction personnelle*). La satisfaction personnelle d'un agent, à un instant donné, est un état fonction de la progression de sa tâche courante.

Cette définition, volontairement générale, s'appuie sur la notion assez floue de *progression de tâche*. En effet, celle-ci peut exprimer

- soit une progression instantanée de la tâche, prenant en compte l'évolution de la tâche entre deux instants *proches*,
- soit une progression plus globale dans le *temps* liée à la réalisation de la tâche depuis son commencement.

Nous pensons que l'état de satisfaction de l'agent doit représenter ces deux notions à la fois. Ainsi, nous proposons ci-dessous une définition du calcul de la satisfaction personnelle basée sur une mesure incrémentale de la progression de la tâche courante.

2.2. Mesure de la satisfaction personnelle

Généralement, un agent évalue sa satisfaction par la perception de la réussite, de l'échec ou de la durée relative de sa tâche en cours. Au contraire, notre objectif est de déterminer un niveau de satisfaction pendant l'exécution de la tâche.

Ce problème est souvent traité par la définition d'une satisfaction ou d'une motivation *fonction de la distance au but* (exemple [HOR 95]) ou de la *force d'un gradient* [ARK 98]. Ces méthodes ont l'inconvénient d'exprimer seulement une *proximité* au but, sans référence aux actions passées. La figure 1.a donne des exemples de fonctions de satisfaction pour un agent devant atteindre un lieu particulier de l'espace. Ainsi, un agent obtient un niveau de satisfaction très élevé lorsqu'il est proche du but (distance d_p) alors qu'il peut être bloqué ou obligé de "reculer" temporairement. Pour palier ce problème, des mesures *d'impatience* ont été proposées (exemple [PAR 94]). Elles sont efficaces lorsque la durée est le critère principal de satisfaction, mais elles ne permettent pas d'évaluer qualitativement la tâche en cours.

Notre approche consiste à augmenter ou à diminuer la valeur de la satisfaction en fonction d'une mesure régulière de l'avancement de la sous-tâche courante.

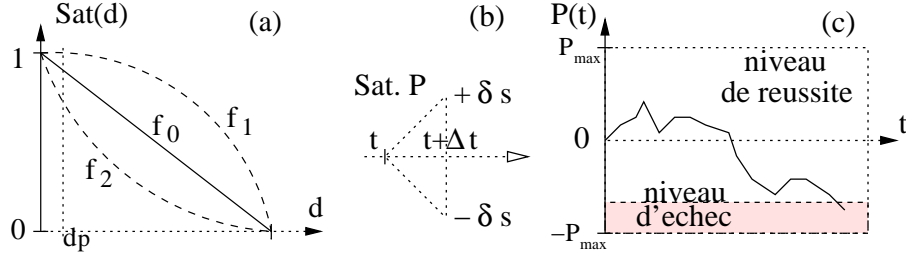


Figure 1. (a) Courbes de satisfaction fonctions de la distance au but (b) Evolution de P pour un pas de temps (c) Exemple de courbe de satisfaction personnelle

Définition (*Valeur de satisfaction personnelle*). Pour un agent qui exécute une sous-tâche, on définit $P(t) \in \mathbb{R}$ sa valeur de satisfaction personnelle à l'instant t par

$$\begin{aligned} - P(t) &= P(t - \Delta t) + v, \\ v \in \mathbb{R} &\text{ est une mesure de la progression de la sous-tâche sur } \Delta t \text{ et } |v| \leq \delta s \\ - \forall t \geq 0, |P(t)| &\leq P_{max}, P_{max} \text{ et } \delta s \text{ sont des constantes de } \mathbb{R}^+ \end{aligned}$$

La première condition précise que le niveau de satisfaction de l'agent est une variation, positive ou négative, de la *précédente* valeur de satisfaction (cf. schéma fig. 1.b).

La deuxième condition permet de borner la valeur de satisfaction dans l'intervalle $[-P_{max}, P_{max}]$ (cf. schéma fig. 1.c), par conséquent $\delta s < P_{max}$. Généralement, on prendra δs très petit devant P_{max} .

La discrétisation du temps Δt correspond au cycle perception-délibération-action de l'agent.

Définition (*Hédographe*). On appelle Hédographe¹ d'un agent la courbe de satisfaction personnelle construite à partir des valeurs $P(t)$ obtenues successivement dans le temps.

2.3. Progression instantanée d'une sous-tâche

L'évolution de P dépend de v qui est une mesure de la *progression instantanée* de la sous-tâche en cours, celle-ci peut être positive, négative ou nulle. Formellement,

$$v = \begin{cases} m & \text{si progression vers but} \\ n & \text{si éloignement du but} \\ f & \text{si agent immobilisé} \\ 0 & \text{sinon} \end{cases}$$

$$\text{avec } -\delta s < f < n < 0 \leq m < \delta s$$

1. Cette terminologie est dû à la concaténation des termes *hedos* (du Grec plaisir) et *graphe*.

Cette formule est une extension des *progress estimator* proposés par M.J. Mataric [MAT 94], car nous introduisons une pénalisation lorsque l'agent est contraint à l'immobilité. Cette pénalisation, plus importante que celle d'un éloignement du but, est introduite pour provoquer un changement rapide dans le comportement de l'agent lorsque celui-ci ne peut réaliser sa tâche.

Nous avons appliqué ce type de mesure à la simulation et à la réalisation de mini-robots mobiles explorant un environnement (cf. [LIE 01]). Ainsi, l'intensité de perception des attracteurs, respectivement des obstacles, permet de calculer la satisfaction de s'en approcher et respectivement de s'en éloigner. Mais un blocage spatial est plus pénalisé qu'un quelconque déplacement.

2.4. Satisfaction et comportement

La valeur $P(t)$ comme nous l'avons définie est un *niveau de satisfaction calculé dans le temps*, qui intègre les états passés. L'intérêt d'une telle approche est de cumuler de l'information pour définir des stratégies comportementales (cf. applications section 4) :

- suivant le niveau de satisfaction obtenu, un agent peut être renforcé dans sa motivation à continuer la tâche courante (valeur élevée) ou au contraire il peut décider de l'abandonner (valeur négative) (cf. fig. 1.c).
- un agent qui a régulièrement progressé vers son but atteint un niveau de satisfaction élevé, s'il rencontre momentanément des difficultés sa satisfaction chute mais reste positive. Ainsi, avec un niveau de satisfaction élevé, un agent peut accepter plus facilement et plus longtemps des contraintes qui l'éloignent de son but (c'est le principe de recherche d'une meilleure solution en acceptant de sortir temporairement d'un minima local), cf. illustration plus bas.

Remarquons que ce calcul progressif de la satisfaction des sous-tâches permet aussi d'absorber les erreurs éventuelles sur les mesures des capteurs.

Pour résumer, la mesure de la satisfaction personnelle permet d'évaluer:

- la *progression courante* des sous-tâches par la variation de P entre deux instants: $P(t) - P(t - \Delta t)$ (la dérivée),
- la *qualité de réalisation* des sous-tâches par l'intégration des valeurs $P(t)$ sur sa période d'exécution,
 - les réalisations rapides et efficaces, notées $P \nearrow^+$,
 - les comportements/situations d'échecs, notés $P \searrow_-$ (ex. fig. 1.c),
 - les actions qui rencontrent des difficultés dans leur réalisation mais qui aboutissent, notées $P \sim^+$.

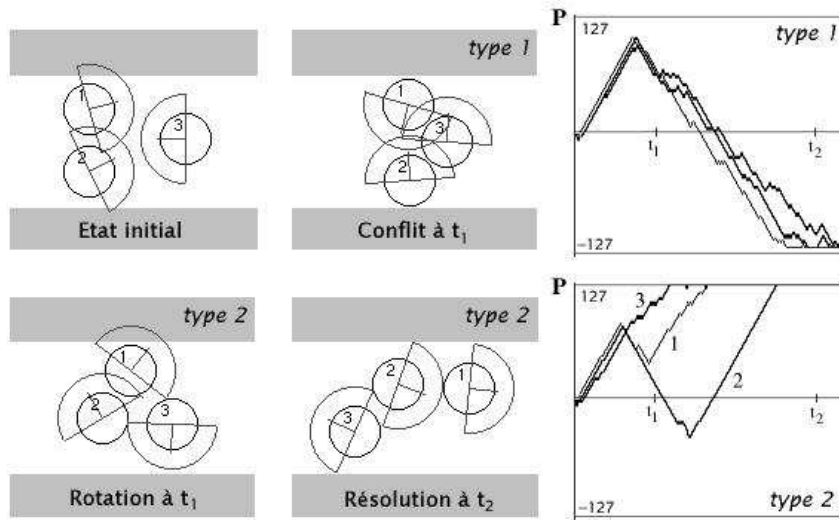


Figure 2. Snapshots de simulations avec hédographes des agents de type 1: évitement simple des obstacles et de type 2: intègre le contournement des agents par la gauche.

2.5. Illustration: action de coordination spatiale

Considérons des agents mobiles capables de se déplacer en évitant les obstacles qu'ils perçoivent devant eux. Ces évitements sont effectués en contournant les obstacles par le côté qu'ils perçoivent comme le plus "libre" (agents notés *type 1*). Cette technique est efficace pour contourner des obstacles fixes, mais elle peut mener à des blocages ou des ralentissements lorsque les agents se font obstacles [ARK 98].

Prenons une situation simple impliquant trois agents de *type 1* qui évoluent à vitesse égale dans un couloir. Ces agents ont pour seule tâche de traverser le couloir. Deux d'entre eux avancent de front vers le troisième qui arrive dans le sens opposé (cf. fig. 2 état initial). Pour étudier le comportement de ces "robots", nous avons réalisé des simulations² de cette situation et tracé progressivement l'hédographe de chaque agent. La variation de satisfaction personnelle dépend de l'angle α entre la direction de déplacement de l'agent et la direction de son but: $v = 0.5 + \cos(\alpha)$ (forte *progression* quand l'agent *avance vers son but*, faible *régression* quand l'agent *s'éloigne de son but*), si l'agent se trouve *immobilisé/bloqué* $v = -1.5$ (forte *régression*) ($P_{max} = 127$).

Lorsque les trois agents se rencontrent, le couloir est trop étroit pour qu'ils puissent se croiser facilement. Les agents de *types 1* sont incapables de se coordonner

2. Nous avons développé un simulateur d'agents mobiles à l'aide de la plate-forme MadKit [GUT 00], cf. [LIE 01].

pour parvenir à se croiser. Les tracés d'hédographes montrent clairement une chute régulière des satisfactions personnelles jusqu'à la valeur minimum (cf. fig. 2).

Modifions le module d'évitement pour que les agents puissent détecter la présence de gêneurs, et qu'ils les contournent toujours par le même côté (par exemple à gauche). Nous avons vérifié que cette action était profitable en réalisant de nouvelles simulations (agents notés *type 2*). Les tracés d'hédographes montrent qu'avec l'action de coordination, la satisfaction ne chute que temporairement lorsque les agents opèrent un croisement par "rotation à gauche". Plus précisément, cela nous permet d'évaluer la résolution: l'agent 3 trouve rapidement un passage (sa satisfaction continue à progresser), l'agent 1 subit une baisse légère avant d'exploiter l'espace libéré par l'agent 3. Enfin, l'agent 2 est plus longuement gêné par ce processus de rotation, sa satisfaction chute jusqu'à des valeurs légèrement négatives avant de remonter vers le maximum dès que l'agent 1 peut progresser.

3. Satisfaction interactive

Nous avons vu précédemment que la satisfaction personnelle modélise un état dépendant de la progression des actions menées par un agent. Si nous définissons un agent à partir de cette seule satisfaction il aura un comportement très "égoïstes", car incapable d'actions intentionnellement coopératives ou altruistes.

Pour développer des comportements coopératifs, il faut donc que chaque agent tienne compte de ses voisins, de leurs actions et de leur potentiel à coopérer. Nous avons ainsi posé une deuxième satisfaction, dite *interactive*, qui évalue uniquement les interactions des agents.

La satisfaction interactive est une estimation par l'agent des *effets* des actions externes sur sa tâche courante: gêne, indifférence ou coopération (potentielle). Ce sont essentiellement des processus de *perception* de présences ou de gênes occasionnées.

L'évaluation de cette deuxième satisfaction peut alors permettre à l'agent d'émettre des influences ou des requêtes vers ses voisins, dans le but de faire évoluer une situation insatisfaisante ou de créer une coopération.

Pour développer une architecture réactive, nous avons exprimé ces influences sous forme de signaux très simples:

Définition (*Signal de satisfaction interactive*). Un signal de satisfaction interactive d'un agent A vers un agent B est une valeur $I \in [-P_{max}, P_{max}]$ qui exprime une demande de A:

- $I > 0$: pour attirer l'agent B (recrutement, partage d'une ressource, aide),
- $I < 0$: pour repousser l'agent B (conflit, gêne).

La section suivante présente comment ces signaux, ainsi que les satisfactions personnelle et interactive, sont exploités pour définir une architecture d'agents mobiles coopératifs.

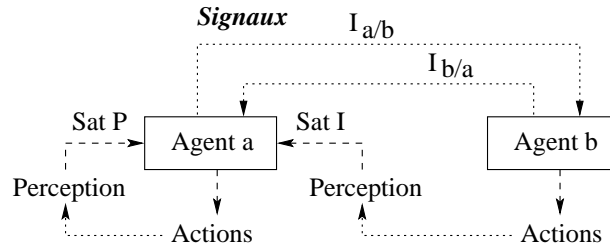


Figure 3. *Modèle des satisfactions personnelle (Sat P) et interactive (Sat I).*

4. Applications

4.1. Des agents altruistes

Le modèle des satisfactions proposé (cf. fig. 3) a été utilisé pour concevoir des agents capables de décider, en temps réel, d'actions de coordination et de coopération. Ce comportement est basé sur une architecture réactive déclenchant des sous-tâches par perceptions de stimuli et d'états internes (type behavior-based [DRO 93] [PAR 94] [ARK 98]). Le principal état interne pris en compte est la satisfaction personnelle. Les stimuli externes sont complétés par des signaux/messages de satisfaction interactive échangés localement.

Le principe général du comportement de l'agent est *la recherche et le contrôle en continue des actions optimales à exécuter vis à vis de lui même et de son voisinage* :

- Tâches personnelles: L'agent utilise la satisfaction personnelle pour décider de la continuation d'une sous-tâche ou pour diminuer la motivation à la déclencher en cas d'insatisfaction. De plus, il emploie l'émission de signaux de satisfaction interactive pour signifier aux autres agents les gênes ou les besoins d'aides.

- Tâches extérieures: L'agent écoute les signaux de satisfaction interactive pour répondre aux besoins exprimés par ses voisins, et y réagir en conséquence.

Lorsqu'un agent reçoit des requêtes qui sont incompatibles avec ses propres objectifs, il ne peut agir simultanément pour lui-même et pour les autres. Nous avons adopté une solution plaçant toujours l'agent dans l'une des deux alternatives. Le choix est déterminé par une comparaison entre le niveau de satisfaction personnelle de l'agent et les signaux de satisfactions interactives qu'il perçoit. Ce modèle, nommé Satisfaction-Altruisme [SIM 00a], fonctionne sur le principe suivant:

- l'agent effectue une sous-tâche tant que la satisfaction personnelle de l'agent n'atteint pas un seuil d'échec (très négatif) et qu'aucun signal extérieur de satisfaction interactive n'est plus intense que sa satisfaction personnelle $P(t)$ (cf. test ci-dessous),

- Soit $I_{ext}(t)$ le signal perçu le plus intense, si $\alpha \cdot |I_{ext}(t)| > (1 - \alpha) \cdot P(t)$ alors l'agent effectue une *action altruiste*: il abandonne sa tâche et réalise une action liée à la valeur de satisfaction interactive perçue.

Le coefficient α permet de pondérer la force du signal extérieur pour définir des agents plus ou moins "sociables". Si $\alpha > 0.5$ on parlera d'agents très coopératifs ou altruistes, pour $\alpha < 0.5$ d'agents peu coopératifs. Enfin, α peut aussi dépendre de l'intensité de réception du signal, ainsi le poids de $|I_{ext}(t)|$ peut être diminué par la distance à l'émetteur.

L'architecture d'agent mobile proposée calcule l'action altruiste comme un mouvement relatif à l'agent émetteur de $I_{ext}(t)$. C'est un *vecteur* déplacement dont le signe est déterminé par la valeur $I_{ext}(t)$: négatif=répulsion, positif=attraction. Cette réaction est combinée à d'autres vecteurs comme les répulsions ou glissements déduits des obstacles perçus localement.

Notons que ces signaux agissent comme des champs de potentiels créés dynamiquement par les agents, dans le but de maximiser leurs interactions. Pour augmenter l'efficacité du principe, les agents en état d'altruisme peuvent *relayer* les signaux perçus et ainsi créer des chemins d'attraction ou de répulsion (cf. [SIM 00b]).

Des robots fourrageurs: Nous avons appliqué ce modèle d'agent au problème classique des robots fourrageurs [STE 90]. Le principe réactif et adaptatif de l'architecture a été mis en évidence sur de nombreuses simulations. Ainsi, nous avons montré que les signaux de satisfaction permettaient *d'explorer* efficacement des environnements dynamiques très contraints, d'améliorer *la coordination* des trajectoires et d'effectuer des *recrutements* par formation de chemins attractifs (cf. détails [SIM 00a][SIM 00b]). Récemment, le modèle a été appliqué, en simulation, à un système hétérogène de robots découpeurs et de robots transporteurs de déchets toxiques ([LIE 01]).

Dans le cadre du développement de robots mobiles autonomes, nous présentons ci-dessous l'application du modèle à des problèmes d'encombrements en navigation collective.

4.2. Navigation collective - détection et résolution des blocages -

Nous abordons ici les problèmes de conflits d'accès et de blocages liés à la navigation d'agents mobiles autonomes en environnements très contraints. Nous avons étendu le modèle pour établir un algorithme distribué permettant aux agents de détecter et résoudre rapidement les situations bloquantes.

Principe: Lorsqu'ils sont en déplacement, les agents émettent des signaux de satisfaction interactive pour signifier à leurs voisins immédiats les gênes ou les blocages éventuels. La résolution de ces conflits est basée sur le fait que ce sont les agents les moins contraints qui doivent adopter un comportement altruiste (ici de recul) devant les autres. Ainsi, les agents les plus contraints (les plus insatisfaits) propagent des vagues d'insatisfactions autour d'eux jusqu'à faire s'écarter ou reculer les gêneurs.

Plus précisément, nous avons spécifié des règles sur l'émission des valeurs de satisfactions interactives. Celles-ci peuvent être résumées ainsi:

– Chaque agent répond prioritairement aux signaux les plus répulsifs (c'est-à-dire négatif et d'intensité maximale) en adoptant un comportement altruiste. L'agent sort de l'état d'altruisme lorsque le signal répulsif perçu est moins intense que sa propre insatisfaction, ou lorsque le signal n'est plus reçu,

– Un agent qui perçoit un signal répulsif plus intense que sa propre insatisfaction *relaie* l'information en émettant un signal d'intensité quasi-égale³. Ce processus de diffusion assure la propagation entre agents des insatisfactions les plus importantes vers les agents les moins contraints.

– Un agent insatisfait par un voisin (gêne, blocage) déclenche l'émission d'un signal répulsif I d'une valeur *égale* à son insatisfaction personnelle P . Il *persiste* dans cette répulsion en émettant la valeur d'insatisfaction *initiale* tant qu'il perçoit ce gêneur (même si les deux agents peuvent se déplacer).

L'application de ces règles locales permet d'obtenir un algorithme complètement distribué assurant la résolution rapide de conflits spatiaux entre de nombreux agents.

La figure 4 montre un exemple de ce type de résolution sur une situation difficile: les agents explorent un corridor très étroit et sans issue. Rapidement des conflits apparaissent: certains agents veulent sortir du corridor, d'autres veulent l'explorer, alors qu'il leur est impossible de se croiser.

Description de la résolution (figure 4): Les agents 3 et 1 sont en conflit (directions recherchées opposées) (1), ce qui déclenche des chutes de satisfactions, puis un signal répulsif de l'agent 1⁴ (2). L'agent 3 perçoit ce signal plus intense que sa propre insatisfaction, il passe en état d'altruisme (3). Il recule jusqu'à être bloqué au fond du corridor (4). L'insatisfaction personnelle de 3 chute alors plus vite que celle de 1 (5), permettant à l'agent 3 d'émettre une répulsion plus intense. Ainsi, l'agent 3 repousse l'agent 1 (6), qui par propagation du signal repousse l'agent 2. L'agent 2 s'éloigne rapidement puis tente de revenir, tandis que les agents 3 et 1 forment un début de "chaîne" qui progresse vers la sortie (7). L'agent 2, à l'approche de l'agent 1, passe à nouveau en état d'altruisme et agrandit la "chaîne", idem pour l'agent 4 au contact de l'agent 2 (8). Finalement, l'agent 3 parvient à repousser les agents jusqu'à sortir du corridor (9).

Ce processus de coordination décentralisé est une illustration de la formation d'un agent de niveau supérieur (la chaîne d'agents) par auto-organisation de plusieurs agents. Le *mouvement* de cette "chaîne" résulte d'interactions qui s'apparentent à celles de l'éco-résolution pour des problèmes du type Taquin [DRO 93]: le but d'un agent est de prendre la place d'un autre (avec priorité au plus contraint), l'agression est réalisée par l'émission d'un signal répulsif et la fuite obligée est effectuée par altruisme.

3. Le signal relayé est légèrement moins intense que celui qui est capté. Ainsi, chaque agent peut déterminer le sens de propagation des répulsions. Ceci est très utile lorsque des erreurs de transmissions surviennent (ce qui est généralement le cas avec des robots réels), et qu'un agent doit *agir* dans le sens de propagation, cf. détails dans [LIE 01].

4. Si l'agent 3 avait été le plus insatisfait, il aurait émis en premier une répulsion et la résolution serait passée directement à l'étape (6).

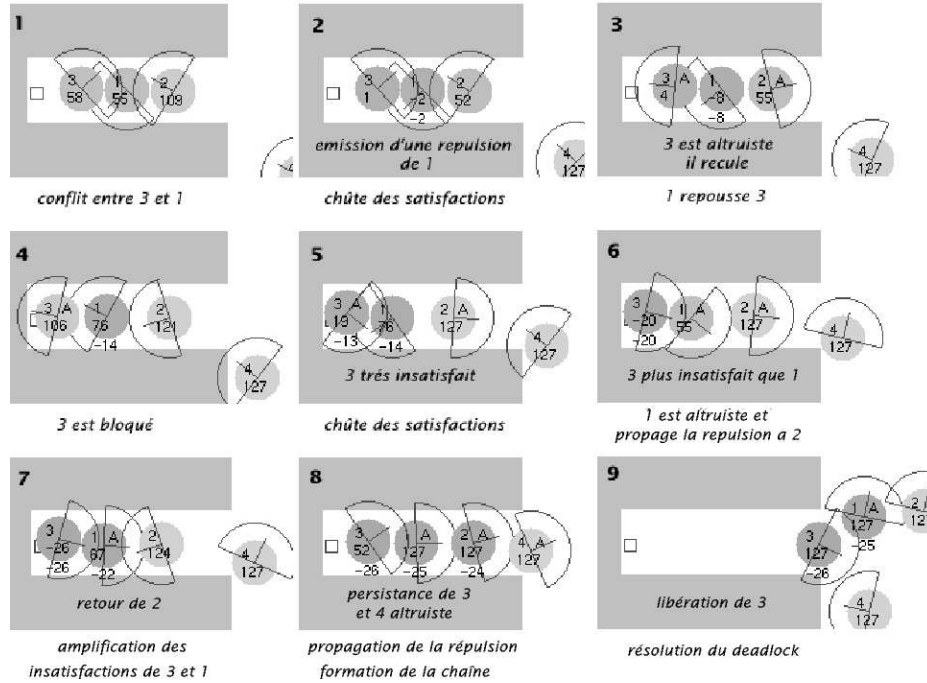


Figure 4. Simulation de robots mobiles autonomes - résolution de conflits spatiaux par le modèle satisfaction-altruisme - Chaque agent est représenté par un disque et sa zone de perception par un demi-cercle. La valeur de satisfaction personnelle (valeur entière dans $[-127,127]$) et l'état d'altruisme (lettre A) sont affichés dans le disque. Lorsqu'un signal de satisfaction interactive est émis (dans un disque de rayon égal à celui de la perception) sa valeur est affichée.

5. Conclusion

Le modèle proposé permet de rendre compte des différentes satisfactions intervenant dans les processus d'actions et d'interactions d'agents situés. La satisfaction liée aux actions individuelles est complétée d'une satisfaction interactive évaluant les situations favorables ou défavorables à la coopération.

Ces perceptions locales sont transformées en signaux d'influences qui permettent d'intégrer au modèle réactif des comportements coopératifs intentionnels. Ainsi, l'architecture proposée gère en temps réel, et suivant les influences extérieures, le choix entre action individuelle et coopérative. De plus, la diffusion des signaux propage ces influences locales à un niveau plus global, permettant la résolution de problèmes distribués tel que celui de l'encombrement spatial.

L'architecture proposée a été implémentée sur des mini-robots mobiles autonomes (réalisés au LIRMM [LIE 01]) et validée par la résolution de conflits spatiaux en envi-

ronnements contraints, dont le problème de l'exploration par trois robots d'un corridor étroit et sans issue. Des expériences avec plus de robots sont en préparation.

Enfin, nous travaillons à l'intégration d'un module d'apprentissage utilisant les signaux de satisfaction interactive pour développer des agents apprenant à coopérer. Les premiers résultats en simulation sont encourageants.

6. Bibliographie

- [ARK 98] ARKIN R., *Behavior-based Robotics*, The MIT Press, 1998.
- [BUR 91] BURA S., DROGOUL A., FERBER J., JACOPIN E., « Eco-résolution: un modèle de résolution par interactions », *Huitième congrès sur la reconnaissance des formes et l'intelligence artificielle (RFIA'91)*, Lyon, AFCET, 1991.
- [DRO 93] DROGOUL A., *De la simulation multi-agent à la résolution collective de problèmes. Une étude de l'émergence de structures d'organisation dans les systèmes multi-agents.*, Thèse de l'université Paris 6, 1993.
- [FER 95] FERBER J., *Les Systèmes Multi-Agents, Vers une intelligence collective*, InterEditions, 1995.
- [GUT 00] GUTKNECHT O., J.FERBER, MICHEL F., « MadKit: Une expérience d'architecture de plate-forme multi-agent générique », *JFIADSMA'00*, 2000, p. 281–284.
- [HOR 95] HORSWILL I., « Analysis of adaptation and environment », *Artificial Intelligence*, vol. 73, 1995, p. 1–30.
- [LIE 01] LIEGEOIS A., FERBER J., SIMONIN O., LUCIDARME P., « Le projet COROM: <http://www.lirmm.fr/~simonin/project.html>, article soumis: Implementation and Evaluation of a Satisfaction/Altruism Based Architecture for Multi-Robot Systems », 2001.
- [MAT 94] MATARIC M., « Reward Functions for Accelerated Learning », COHEN W. W., H. HIRSH E., Eds., *Machine Learning Proc. of the Eleventh International Conference*, Morgan Kaufmann, 1994, p. 181-189.
- [PAR 94] PARKER L., « ALLIANCE: An Architecture for Fault Tolerant, Cooperative Control of Heterogeneous Mobile Robots », *Proceedings of International Conference on Intelligent Robots and Systems (IROS'94)*, Munich, Germany, 1994, p. 776-783.
- [PIR 99] PIRJANIAN P., « Behavior coordination mechanisms – state-of-the-art », *Technical Report IRIS-99-375, Institute of Robotics and Intelligent Systems*, School of Engineering, University of Southern California. <http://iris.usc.edu/~irislib>, 1999.
- [SIM 00a] SIMONIN O., FERBER J., « Modeling Self Satisfaction and Altruism to handle Action Selection and Reactive Cooperation », *From Animals to Animats 6, SAB'2000 supplement volume*, Paris, France, 2000, p. 314-323.
- [SIM 00b] SIMONIN O., LIEGEOIS A., RONGIER P., « An Architecture for Reactive Cooperation of Mobile Distributed Robots », *DARS'2000 Distributed Autonomous Robotic Systems 4*, Springer, 2000, p. 35–44.
- [STE 90] STEELS L., « Cooperation between distributed agents Through self-organization », DEMAZEAU Y., MÜLLER J., Eds., *Decentralized AI– Proceedings of the First European Workshop on Modelling Autonomous Agents in a Multi-Agent World (MAAMAW-89)*, Elsevier Science B.V., Amsterdam, 1990, p. 175-196.