
Agents BDI et simulations sociales

Unis pour le meilleur et pour le pire

Carole Adam* — **Benoit Gaudou****
Sarah Hickmott*** — **David Scerri*****

* Laboratoire d'Informatique de Grenoble, Université Joseph Fourier, Grenoble, France
carole.adam@imag.fr

** Université de Toulouse, IRIT, UMR 5505 (CNRS), Toulouse, France
benoit.gaudou@univ-tlse1.fr

*** RMIT University, School of CS & IT, Melbourne, Australie
sarah.hickmott@rmit.edu.au, david.scerri@rmit.edu.au

RÉSUMÉ. La modélisation et simulation ont longtemps été dominées par les approches basées sur les équations, jusqu'à l'avènement récent des approches orientées agents. Pour freiner l'augmentation de complexité des modèles que peut entraîner l'utilisation de cette nouvelle approche, la tendance est à la sursimplification des modèles. Des modèles plus descriptifs ont cependant été développés pour une variété de phénomènes, mais la cognition des agents est encore trop souvent négligée alors qu'elle a une grande importance dans certains domaines, en particulier en sciences humaines et sociales. La solution que nous proposons dans cet article est d'utiliser des agents BDI. Nous montrons qu'il s'agit d'un paradigme expressif, réaliste et simple qui apporte de nombreux bénéfices à la simulation à base d'agents.

ABSTRACT. Modeling and simulation have long been dominated by equation-based approaches, until the recent advent of agent-based approaches. To curb the increasing complexity of models resulting from this new approach, the trend has been to oversimplify the models. Some more descriptive models have still been developed for various phenomenons, but the cognition of agents is too often neglected, despite its great importance in some fields, such as Social and Human Sciences. The solution that we put forward in this paper is to use BDI agents. We will show that this is an expressive, realistic yet simple paradigm that thus offers numerous benefits to agent-based simulation.

MOTS-CLÉS : modélisation et simulation multi-agent, agents BDI, sciences sociales.

KEYWORDS: agent-based modeling and simulation, BDI agents, social sciences.

1. Introduction

Le domaine de la modélisation et de la simulation a été dominé pendant des années par les approches à base d'équations, qui ne décrivent le comportement du système ou du phénomène étudié que d'un point de vue global. Plus récemment, l'utilisation des systèmes multi-agents dans ce domaine a donné naissance à un nouveau paradigme, la modélisation et simulation à base d'agents (*Agent-based Modeling and simulation*, ABMS) qui permet aux modélisateurs de raisonner et de représenter le phénomène étudié au niveau individuel, et de prendre en compte l'hétérogénéité et la complexité à la fois au niveau de l'agent et au niveau de l'environnement. Ce paradigme a été utilisé avec succès par exemple en écologie (DeAngelis *et al.*, 1992), économie (Tesfatsion, 2002) ou sciences sociales (Gilbert *et al.*, 2005). Cependant un débat existe entre les partisans de différents types d'agents dans ces simulations, des agents les plus simples (types particules) aux plus complexes (architectures cognitives sophistiquées).

Ainsi dans le contexte de modèles à base d'agents destinés à "enrichir notre compréhension de processus fondamentaux", (Axelrod, 1997) introduit le principe KISS : "*keep it simple, stupid*". Son idée est qu'"à la fois le chercheur et l'audience ont des facultés cognitives limitées", ce qui réduit leur capacité à comprendre le contenu du modèle et par là même leur niveau de confiance dans l'émergence d'un "résultat surprenant". Il recommande donc de garder les hypothèses sous-jacentes simples, afin de faciliter la compréhension de la complexité résultant de la simulation. Un modèle simple se caractérise par : un petit nombre de paramètres, un environnement simple, un faible nombre d'entités et un comportement très simple (réactif) des agents.

Mais la question s'est posée dans la communauté agent (*p.ex.* (Drogoul *et al.*, 2002)) de savoir si la simplicité est *toujours* appropriée (et sinon, quand). En effet il a été suggéré qu'il y a beaucoup à gagner à simuler des modèles plus riches, plus réalistes ; un tel point de vue est capturé par le principe KIDS, introduit par (Edmonds *et al.*, 2005), qui recommande l'utilisation d'un formalisme aussi descriptif et direct que possible lors de la modélisation, même si cela implique une plus grande complexité. L'idée est qu'il vaut mieux partir d'un formalisme assez descriptif et bien adapté aux besoins de la simulation et le simplifier éventuellement plus tard, que de partir d'un formalisme simpliste inadapté et de devoir l'étendre ensuite de manière *ad hoc*. L'adoption de l'approche KIDS a donné lieu à de nombreux travaux sur la gestion de l'environnement (avec les systèmes d'information géographique), à des simulations qui considèrent un grand nombre de paramètres, des entités hétérogènes, une représentation précise du temps et à des simulations multi-échelles ; cependant le composant comportemental est en général resté très simple.

Le but de cet article est de défendre l'idée que l'utilisation d'agents ayant une architecture plus complexe et donc étant plus descriptifs pourrait apporter de nombreux bénéfices à la simulation basée sur les agents. Cependant, (Edmonds *et al.*, 2005) admettent que le choix entre descriptivité et simplicité, dépend énormément du contexte, du but de la simulation et de divers paramètres. C'est pourquoi cet article se concentre sur un domaine particulier, celui des sciences sociales, qui s'intéressent au comporte-

ment social d'entités complexes, à savoir les êtres humains. Intuitivement, il est plus réaliste de les représenter par des agents dotés de capacités cognitives évoluées que par des particules, utilisées dans d'autres domaines s'intéressant à des entités plus simples. Un formalisme plus descriptif offre aussi de nombreux bénéfices, comme de rendre la programmation plus intuitive et de permettre aux agents de s'adapter à l'environnement et d'expliquer leur comportement.

Dans la section 2, nous justifions l'utilisation d'agents complexes et plus particulièrement d'agents intelligents en simulation sociale (2.1 et 2.2) et notre choix de nous concentrer sur l'architecture BDI (2.3 et 2.4). La contrepartie de cette meilleure descriptivité des agents est l'augmentation de la complexité des modèles ; il est donc nécessaire d'identifier les types de simulations où les bénéfices justifient cet ajout de complexité : c'est l'objet de la section 3. Nous y discutons la littérature et comparons les simulations existantes sur différents axes : les caractéristiques d'agents nécessaires (*p.ex.* apprentissage, émotions, 3.1) ; le but de la simulation (*p.ex.* entraînement, découverte de connaissances, 3.2) ; et divers autres paramètres tels que le niveau de granularité ou le domaine d'application (3.3). Nous discutons aussi divers aspects qui peuvent décourager l'utilisation d'agents BDI dans les simulations (3.4).

2. Motivations pour l'utilisation d'agents BDI dans les simulations sociales

2.1. Exemples de simulations sociales basées sur différents types d'agents

2.1.1. Simulation de panique lors d'une évacuation (Helbing et al., 2000)

(Helbing *et al.*, 2000) proposent un modèle de comportement de foule dans le cas d'évacuations d'urgence. Leur but est de modéliser les situations de panique dans lesquelles la pression de la foule peut être dangereuse, et de trouver des solutions pour l'évacuation d'endroits dangereux (*p.ex.* une pièce remplie de fumée). Les auteurs ont choisi de représenter les agents comme des particules dans un champ de force : toutes les interactions physiques et sociales avec les autres agents et avec l'environnement sont représentées comme des forces. Le mouvement des particules représentant les êtres humains est décrit par des équations électromagnétiques.

2.1.2. DamasRescue (Paquet et al., 2004)

Le projet DamasRescue vise à simuler le comportement d'une équipe de sauvetage comprenant des policiers, médecins et pompiers dans le contexte d'une catastrophe naturelle. Le simulateur est intégré dans l'environnement RobocupRescue (Kitano, 2000) qui fournit des simulateurs pour le feu, la circulation et les passants. Les pompiers peuvent exécuter diverses actions (éteindre le feu, entrer dans un bâtiment pour secourir des victimes...). Les agents les représentant utilisent un arbre de décision (préconstruit avec des techniques d'apprentissage) pour sélectionner l'action la plus appropriée parmi les actions disponibles, en fonction de leur perception de l'environnement (distance par rapport aux bâtiments en feu, nombre de feux, nombre de victimes dans un bâtiment...).

2.1.3. *Simulation de la survie dans un groupe (Cecconi et al., 1998)*

(Cecconi *et al.*, 1998) proposent d'utiliser une simulation pour évaluer différentes stratégies de survie des individus dans un groupe social. La survie (et reproduction) ou la mort d'un individu dépend des ressources dont il dispose. Ces ressources peuvent être des ressources physiques (telles que l'argent, la nourriture), d'autres agents (tels qu'un partenaire de reproduction) ou des aptitudes personnelles (telles que la capacité de se battre ou de travailler). Les agents peuvent utiliser deux types de stratégies de survie : individuelles ou sociales. Une stratégie sociale particulièrement intéressante consiste à donner une partie de ce que l'agent récolte à un fonds commun de ressources. Les auteurs s'intéressent non seulement à l'évolution des aptitudes sur plusieurs générations, mais aussi à la taille idéale des groupes sociaux. Chaque agent est représenté comme un réseau de neurones. L'évolution est modélisée par un algorithme génétique qui vise à améliorer la capacité à trouver et produire de la nourriture.

2.1.4. *Simulation de foule pour la réponse aux situations d'urgence (Shendarkar et al., 2008)*

Cette simulation représente l'évacuation d'une foule pendant des attaques terroristes à la bombe dans une zone publique. Les auteurs étudient l'influence de divers paramètres (*p.ex.* densité ou taille de la foule, nombre de policiers gérant l'évacuation, nombre de sorties dans la pièce) sur la qualité de l'évacuation (*p.ex.* durée, nombre de victimes). Ils ont choisi d'implémenter les agents en utilisant une architecture BDI étendue, qui inclut un composant émotionnel et un planificateur en temps réel. Le modèle de comportement des agents est obtenu en capturant le comportement d'un humain placé dans la même situation dans un laboratoire de réalité virtuelle.

2.1.5. *Amélioration de la simulation de survie dans un groupe (Sun et al., 2007)*

(Sun *et al.*, 2007) proposent de donner d'authentiques capacités cognitives aux agents impliqués dans la simulation de (Cecconi *et al.*, 1998), afin de la rendre plus réaliste d'un point de vue cognitif. Par ailleurs, ils introduisent des institutions sociales qui imposent des obligations aux agents et des sanctions en cas de violation ; par exemple, un agent peut être puni s'il refuse de contribuer au fonds commun de nourriture quand cela est requis. Grâce à ces additions au modèle initial, les auteurs veulent étudier les interactions entre la cognition individuelle et les institutions sociales. Les processus cognitifs des agents sont implémentés en utilisant CLARION, une architecture cognitive principalement basée sur les travaux de Sun en sciences cognitives (Sun *et al.*, 2001).

2.2. *Justification de l'utilisation d'agents intelligents dans les simulations sociales*

2.2.1. *Synthèse des exemples précédents*

Ces quelques exemples montrent une grande variabilité dans les architectures d'agents utilisées. En effet, ils illustrent respectivement : une architecture de type

particules ; une architecture à base d'arbre de décision ; une architecture à base de réseaux de neurones ; une architecture BDI ; et une architecture cognitive. On peut faire une dichotomie entre ces cinq architectures par rapport aux concepts manipulés pour implémenter les comportements et à leur niveau de complexité.

Les trois premières architectures (particule, arbre de décision et réseau de neurones) n'impliquent aucun concept cognitif dans les processus de raisonnement et de décision. Même si les agents représentent des êtres humains, leurs comportements sont principalement fonction des stimuli externes (*p.ex.* les forces extérieures dans le premier exemple, la perception de l'environnement dans les deuxième et troisième). Ces architectures sont traditionnellement qualifiées d'*agents réactifs* (Wooldridge, 2009). Elles sont très bien adaptées pour représenter des entités simples sans capacité de raisonnement telles que des arbres, fourmis ou bâtiments, mais ne peuvent pas représenter les mécanismes de raisonnement d'entités complexes telles que des humains.

Au contraire, les deux dernières architectures représentent deux types d'*agents intelligents* (Wooldridge, 2009) : les agents BDI et les agents à architecture cognitive¹. Les architectures d'agents intelligents permettent aux agents d'être plus autonomes au lieu de seulement réagir aux stimuli environnementaux, et ainsi de ressembler plus aux humains. Elles permettent aussi la modélisation de problèmes plus complexes, impossibles à aborder avec des agents purement réactifs (comme cela est illustré par la progression entre 2.1.3 et 2.1.5). Elles ont enfin une forte base théorique qui peut fournir des indications lors de la modélisation (mais aussi de nombreuses contraintes).

2.2.2. Pourquoi utiliser des agents intelligents en simulation

Les cinq exemples ci-dessus montrent un continuum d'architectures, de la plus simple (agents particules) à la plus complexe (agents à architecture cognitive), et une augmentation parallèle des capacités des agents.

(Sun, 2007) remarque qu'"un important défaut dans le domaine de la simulation sociale en informatique est que la plupart des travaux supposent une cognition très rudimentaire de la part des agents. [...] Non seulement cela limite le réalisme, et donc l'applicabilité, des simulations sociales, mais cela exclut aussi la possibilité de répondre à la question du lien micro-macro en termes d'interaction cognitive-sociale"². L'utilisation d'agents intelligents en sciences sociales permettrait de corriger ces limitations, et offrirait aussi aux agents la possibilité de raisonner sur les autres agents, sur les normes ou sur les règles de la société ou d'un groupe d'agents.

1. L'expression "architecture cognitive" utilisée ici renvoie aux architectures basées sur des théories cognitives, comme (Sun, 2002; Newell, 1990; Anderson, 1983). Elle ne doit pas être confondue avec le terme "agent cognitif" qui est généralement un synonyme d'"agent intelligent" dans la littérature SMA, et englobe donc aussi les agents BDI.

2. "a significant shortcoming of current computational social simulation is that most of the work assumes very rudimentary cognition on the part of agents. [...] It not only limits the realism, and hence the applicability, of social simulations, but also precludes the possibility of tackling the question of the micro-macro link in terms of cognitive-social interaction" (notre traduction)

En effet, le comportement d'un être humain en société n'est pas seulement fonction de ce qu'il veut faire, mais aussi de ce qu'il doit faire. L'agent ne peut pas suivre aveuglément ses plans sans tenir compte de l'évolution de ses relations avec les autres agents. De même que des phénomènes spatiaux émergent dans le cas de simulations avec des agents très simples (*p.ex.* la ségrégation de (Schelling, 1971)), des phénomènes "cognitifs" intéressants peuvent émerger suite aux interactions entre agents intelligents, comme l'émergence de normes, de croyances et d'intentions de groupe. Il est cependant nécessaire de disposer d'un modèle d'agent suffisamment descriptif pour permettre de simuler de tels phénomènes ; c'est le cas des architectures d'agents intelligents (architectures BDI ou cognitives).

2.3. *Choix d'un paradigme d'agent intelligent*

Aussi bien les architectures BDI que les architectures cognitives permettent aux modélisateurs de représenter les processus cognitifs et le raisonnement des agents, et fournissent donc la descriptivité dont nous revendiquons l'utilité dans les simulations sociales (en accord avec (Edmonds *et al.*, 2005)). Cependant ces deux types d'architectures sont très différentes. Nous détaillons dans la suite les principales raisons qui nous ont poussés à choisir l'architecture BDI.

2.3.1. *Multiplicité des théories sous-jacentes vs théorie unique*

Tout d'abord, les architectures cognitives se basent sur les sciences cognitives mais chaque implémentation est basée sur sa propre théorie. Par exemple, l'architecture CLARION (Sun *et al.*, 2001) est basée sur (Sun, 2002) ; SOAR (Rosenbloom *et al.*, 1993) est basée sur (Newell, 1990) ; et ACT-R (Anderson *et al.*, 1998) est basée sur (Anderson, 1983 ; Anderson, 1993). De plus elles n'ont pas de centre d'intérêt commun : CLARION s'intéresse à la dichotomie entre implicite et explicite, en particulier dans le processus d'apprentissage, alors qu'ACT-R se concentre sur le processus d'activation. Il n'y a pas non plus de consensus sur les composants mémoire de l'architecture, que ce soit le nombre de mémoires ou leur nature.

Au contraire, l'architecture BDI est basée sur un courant philosophique (par exemple les travaux de (Bratman, 1987)), formalisé principalement par les logiques modales ((Rao *et al.*, 1991) ou (Cohen *et al.*, 1990)). Ces logiques décrivent les trois principaux états intentionnels (croyances, désirs, intentions) qui régissent le comportement d'un agent, le menant de ses croyances sur le monde au choix d'actions à exécuter, et à l'exécution de ces actions. L'architecture BDI est largement adoptée dans la communauté SMA et ces trois concepts centraux sont une constante dans le nombre considérable de formalisations et d'implémentations existantes de cette architecture. Elle peut aussi être étendue avec d'autres concepts déontiques ou sociaux tels que les émotions (Pereira *et al.*, 2005) ou les normes (Jones *et al.*, 2009).

2.3.2. Niveau d'abstraction et simplicité de compréhension

Un autre différence essentielle entre les architectures cognitives et BDI est le niveau d'abstraction des concepts manipulés. Le but des architectures cognitives étant de décrire les processus cognitifs humains aussi précisément que possible, elles en nécessitent une compréhension en profondeur (par exemple quelle partie du cerveau exécute quelle fonction) et manipulent des concepts de bas niveau (*p.ex.* les *chunks* de mémoire dans ACT-R). Au contraire les concepts BDI sont des concepts de haut niveau (attitudes mentales) qui résultent des concepts de bas niveau et des principes décrits dans les architectures cognitives. D'après Sun, les architectures BDI peuvent être considérées comme la partie explicite et symbolique des architectures cognitives.

Les architectures BDI se concentrent en effet sur le niveau conscient de l'esprit, le niveau auquel les humains peuvent consciemment manipuler les concepts et les communiquer aux autres. Ainsi (Norling, 2004, p.1) affirme que "les agents BDI ont été utilisés avec grand succès pour modéliser les humains et créer des personnages humanoïdes dans les environnements virtuels. Une raison essentielle de ce succès est que le paradigme BDI est basé sur la psychologie naïve (*folk psychology*), ce qui signifie que les concepts centraux de l'architecture agent correspondent directement au langage utilisé par les humains pour décrire leur raisonnement et leurs actions dans les conversations de tous les jours."³. Cette formalisation directe du raisonnement humain en termes de concepts intuitifs (croyances, désirs et intentions) est exactement ce que le principe KIDS recommande. Elle procure en effet de nombreux avantages, tels qu'identifiés par (Norling, 2004, p.1) : "la correspondance directe entre les concepts essentiels BDI et la terminologie naturellement utilisée par les experts modélisés pour décrire leur raisonnement facilite la capture, la représentation et la correction des connaissances."⁴. L'utilisation de ces concepts BDI de haut niveau facilite également la compréhension des comportements des agents par les non-informaticiens.

2.3.3. Conclusion

Ainsi les architectures BDI semblent être à la fois assez descriptives pour représenter les processus cognitifs influençant le comportement et assez intuitives pour être facilement compréhensibles par les modélisateurs et les non-informaticiens en général. Dans la suite de cet article nous nous concentrons donc sur les agents BDI et commençons par détailler leurs caractéristiques.

3. Citation originale : "BDI agents have been used with considerable success to model humans and create human-like characters in simulated environments. A key reason for this success is that the BDI paradigm is based in folk psychology, which means that the core concepts of the agent framework map easily to the language people use to describe their reasoning and actions in everyday conversations"; notre traduction.

4. Citation originale : "the close match between the core BDI concepts and the natural terminology that the experts being modelled used to describe their reasoning facilitated knowledge capture, representation and debugging"; notre traduction.

2.4. Description d'un agent BDI

2.4.1. Théorie BDI

Le modèle BDI a ses fondements théoriques dans les travaux en philosophie de (Bratman, 1987) et (Dennett, 1989) qui définissent les trois composants de base permettant de capturer l'interprétation courante du raisonnement humain : les *croyances* représentant le savoir de l'agent sur son environnement et son propre état interne ; les *désirs* ou plus spécifiquement les *buts* (des désirs non conflictuels que l'agent a décidé qu'il veut réaliser) ; et les *intentions* qui sont l'ensemble des plans ou séquences d'actions que l'agent veut suivre afin de réaliser ses buts.

(Cohen *et al.*, 1990) affirment qu'en permettant un niveau d'engagement à un ensemble d'intentions pour réaliser un but à long terme, le modèle BDI se distingue des modèles qui permettent uniquement un comportement réactif. Cette notion d'engagement est donc un composant essentiel du modèle BDI. (Rao *et al.*, 1991) montrent qu'il doit aussi y avoir une certaine forme de *processus rationnel* pour décider quelles intentions sélectionner selon les circonstances. Bien que la manière spécifique de réaliser cela varie d'une implémentation à l'autre, elle est généralement guidée par les principes de psychologie naïve (Norling, 2004).

2.4.2. Systèmes implémentés

Les fonctionnalités centrales que tout système BDI doit implémenter sont donc : une représentation des croyances, désirs et intentions ; un processus rationnel et logique de sélection des intentions ; et un mécanisme flexible et adaptable d'engagement sur l'ensemble actuel d'intentions. Il existe un nombre substantiel de systèmes BDI implémentés (*p.ex.* JACK (Busetta *et al.*, 1999), Jason (Bordini *et al.*, 2007), 3APL (Hindriks *et al.*, 1999)) qui supportent la création d'agents définis en utilisant ces trois concepts et proposent chacun leur propre un moteur d'exécution BDI, ainsi que plusieurs systèmes qui les étendent pour fournir d'autres capacités de raisonnement humain.

Un point commun à toutes ces implémentations est le concept d'une *bibliothèque de plans* qui est une collection de plans que l'agent peut exécuter. Chacun de ces plans se compose d'un corps, d'un but que le plan permet de réaliser, et d'un contexte dans lequel le plan est applicable. Le corps du plan est constitué à la fois d'actions qui peuvent agir directement sur l'environnement, et de sous-buts, qui seront développés en d'autres plans quand nécessaire. Cette imbrication de buts dans les plans permet de construire une hiérarchie de plans et buts, où les branches partant d'un but mènent vers les plans permettant de le réaliser, et les branches partant d'un plan mènent vers les buts qui doivent être accomplis pour compléter ce plan.

2.4.3. Caractéristiques centrales et bénéfices

L'architecture BDI offre les caractéristiques de base suivantes :

– Adaptabilité - *Les plans ne sont pas complètement spécifiés et peuvent être constitués de sous-buts imbriqués, ce qui rend l'agent flexible et lui permet de s'adapter à un environnement changeant.*

– Robustesse - *L'utilisation d'une hiérarchie de plans signifie que si un plan échoue, potentiellement à cause de changements dans l'environnement pendant son exécution, l'agent est capable de se ressaisir en exécutant un autre plan applicable s'il en a un disponible.*

– Programmation abstraite - *Le programmeur spécifie les croyances, désirs et intentions (concept de haut niveau d'abstraction), ce qui permet de créer des systèmes complexes tout en maintenant un code transparent et facile à comprendre.*

– Orientation vers les buts - *L'utilisation d'une approche basée sur les buts par rapport à une approche basée sur la tâche signifie qu'un agent sait **pourquoi** il exécute une tâche, ce qui lui permet d'expliquer son comportement d'une manière intuitive.*

2.4.4. Extensions de l'architecture BDI

L'architecture BDI est ouverte aux extensions et de nombreux travaux se sont donc intéressés à l'extension des concepts de base pour supporter d'autres facteurs qui influencent le processus de raisonnement humain. (Dignum *et al.*, 2000) introduisent le concept de *normes* dans les systèmes BDI. De la même manière que le comportement humain est très influencé par les normes sociales, le comportement des agents peut lui aussi être influencé par les normes, permettant aux agents de mieux comprendre ce que les autres agents sont susceptibles de faire. (Pereira *et al.*, 2005) présentent une extension de l'architecture BDI pour inclure des émotions artificielles, qui affectent le raisonnement de l'agent de la même manière que les émotions réelles affectent le raisonnement d'un être humain. Des travaux ont aussi intégré des planificateurs temps réel dans l'architecture pour supporter la création de plans lors de l'exécution (de Silva *et al.*, 2009). Les émotions, les normes et la planification, et la manière dont elles sont intégrées dans les simulations, seront discutées plus loin.

3. Caractéristiques des simulations influençant l'utilité des agents BDI

Dans la section précédente nous avons justifié pourquoi nous nous focalisons sur les agents BDI dans cet article. Nous nous attachons maintenant à identifier les critères qui permettent d'utiliser cette architecture mais aussi ceux qui empêchent son utilisation (ou diminuent ses bénéfices).

3.1. Influence des capacités qu'on souhaite donner aux agents

3.1.1. Capacités émotionnelles

D'après les théories de psychologies cognitives (Ortony *et al.*, 1988; Lazarus, 1991), les émotions peuvent être directement exprimées en termes des attitudes

mentales décrites par le formalisme BDI (croyances, désirs, intentions). Par exemple la joie est la conjonction d'un désir et d'une croyance que ce désir est satisfait. En effet, (Nair *et al.*, 2005, p.12) affirment que "l'un des intérêts des modèles de l'évaluation cognitive comme base des modèles informatiques est la facilité avec laquelle l'évaluation peut être liée au cadre BDI souvent utilisé dans les systèmes agent"⁵ ; inversement, ce lien facile entre concepts BDI et émotions (tel qu'il est supposé par les théories de l'évaluation cognitive) constitue aussi un argument en faveur de l'utilisation d'agents BDI dans les simulations où des capacités émotionnelles sont requises.

Le cadre BDI standard nécessite une extension pour intégrer les émotions, mais il permet ensuite de raisonner à leur sujet et de comprendre leur influence sur le comportement des agents. Plusieurs modèles BDI des émotions ont donc été développés (Steunebrink *et al.*, 2008; Adam, 2007) et un certain nombre d'agents BDI ont été dotés d'émotions pour les rendre plus crédibles (*p.ex.* agents conversationnels animés (de Rosis *et al.*, 2003)) ou plus efficaces (*p.ex.* tuteurs virtuels (Gratch *et al.*, 2005)). De tels agents émotionnels sont aussi applicables dans différents types de simulations.

(Jones *et al.*, 2009) ont développé une extension appelée PEP→BDI qui enrichit le moteur BDI avec des concepts de personnalité, d'émotions, ainsi que des facteurs physiologiques ; ils appliquent cette extension à une simulation visant à entraîner les acteurs de la sécurité à réagir lors de crises à grande échelle, comme des attaques terroristes. (Bosse *et al.*, 2009) étudient l'influence de facteurs psychologiques (émotions), sociaux et biologiques sur le comportement criminel. Ils utilisent leur propre langage, LEADSTO, qui permet de modéliser à la fois les aspects qualitatifs cognitifs (BDI) et les aspects quantitatifs biologiques.

(Gratch *et al.*, 2004) ont développé un cadre général de planification BDI, intégré avec les processus d'*appraisal* et de *coping* pour déclencher des émotions chez un agent selon son interprétation du monde et pour modéliser leur influence sur son comportement. Ce cadre nommé EMA a été appliqué à des simulations pour l'entraînement de militaires : des officiers humains interagissent avec des agents dans un monde virtuel afin d'améliorer leurs compétences dans des situations stressantes réelles. Les agents doivent avoir des comportements complexes (en particulier émotionnels) pour ajouter du réalisme et favoriser l'immersion, ce qui améliore l'entraînement.

Les agents BDI ont aussi été dotés d'émotions dans des simulations visant à explorer les relations entre les émotions et divers autres concepts. (Staller *et al.*, 2001) montrent que les émotions ont une relation bidirectionnelle avec les normes : les émotions peuvent inciter à respecter les normes (*p.ex.* honte ou culpabilité) ou à les imposer aux autres (*p.ex.* mépris) ; et les normes influencent le déclenchement et la régulation des émotions. (Nair *et al.*, 2005) ont étudié comment les capacités émotionnelles peuvent améliorer le travail d'équipe dans des équipes composées soit uniquement d'agents soit à la fois d'agents et d'humains.

5. "one of the appeals of cognitive appraisal models as the basis of computational model is the ease with which appraisal can be tied to the belief, desire and intention (BDI) framework often used in agent systems" ; notre traduction.

Enfin les agents BDI dotés d'émotions peuvent aussi être plus performants dans la réalisation de leur tâche. Ainsi (Nair *et al.*, 2005) construisent des agents émotionnels en utilisant BDI combiné avec POMDP pour gérer l'incertitude. Ils les intègrent ensuite dans divers types d'équipes : dans des équipes purement agents, le comportement est plus réaliste, car les êtres humains utilisent leurs émotions pour ordonner leurs buts et choisir l'action à réaliser ; dans des équipes mixtes humains-agents, leur conscience des émotions de leurs partenaires humains facilite la communication avec eux et l'expression de leurs émotions améliore leurs relations.

Notons que les émotions peuvent aussi être intégrées dans des agents ayant un autre type d'architecture. Par exemple (Silverman *et al.*, 2005) ont intégré les émotions dans leur système MACES de simulation de foule en utilisant PFMserv, une architecture cognitive dont ils prétendent qu'elle permet un comportement plus réaliste des agents, en leur procurant deux types de mémoires utilisées dans leur cycle de décision : une mémoire de travail similaire à la mémoire à court terme des humains, ainsi qu'une mémoire à long-terme. Cependant pour les raisons exposées en section 2.3 nous ne nous intéressons ici qu'aux architectures BDI.

3.1.2. *Raisonnement sur les normes*

Les normes sont un ensemble de faits déontiques (obligations, permissions ou interdictions) spécifiant comment les agents sont censés se comporter ; elles sont souvent représentées en logique déontique (von Wright, 1963; Conte *et al.*, 1995). Les normes sociales sont des normes partagées par un groupe d'agents ; elles constituent un outil très puissant pour décrire les règles dans un groupe d'agents. Il est important de considérer le fait que ces normes *peuvent être* violées ; il est donc nécessaire d'avoir un mécanisme pour gérer ces violations, tel que l'application de sanctions aux agents coupables. Chaque agent doit considérer les normes dans son processus de décision, *i.e.* ses désirs devraient être pondérés par les normes. La version basique du cadre BDI n'inclut pas les normes, mais certaines extensions les intègrent, *p.ex.* l'architecture BOID (Broersen *et al.*, 2001; Neumann, 2010) qui met particulièrement l'accent sur les conflits entre les quatre modalités.

Bien que les normes soient le plus souvent utilisées dans des applications comme le commerce électronique ou le B2B (*business to business*), certaines expériences ont aussi montré leur pertinence dans les simulations. (Savarimuthu *et al.*, 2009) donnent un aperçu assez complet de l'utilisation des normes en simulation et catégorisent les simulations par rapport aux mécanismes normatifs. Dans les simulations intéressées par les normes et leur influence sur le comportement des agents, il est important d'avoir un modèle précis de la cognition des agents qui leur permet de raisonner explicitement sur les normes (voir section 3.3.2). Par exemple certaines simulations visent à expliquer le comportement criminel d'un point de vue cognitif (Bosse *et al.*, 2009) et utilisent donc des agents BDI. (Burgemeestre *et al.*, 2009) montrent que dans les SMA normatifs, différents types de normes nécessitent différentes architectures d'agents, et discutent en particulier comment diverses architectures cognitives permettent de raisonner sur la conformité aux normes.

Dans (Conte *et al.*, 1995), la fonction des normes sociales est de contrôler les agressions dans la société. Les auteurs utilisent délibérément des agents très simples pour étudier le rôle des normes au niveau macroscopique. Ces agents peuvent exécuter quelques routines simples (se déplacer, manger, attaquer) pour survivre dans un environnement pauvre en nourriture ; ils ont une force qui augmente en mangeant et diminue lors des déplacements et attaques. Les auteurs comparent deux stratégies pour contrôler les agressions dans cette société virtuelle : avec la stratégie normative, les agents respectent une norme (*p.ex.* “le premier qui trouve quelque chose le garde”), alors qu’avec la stratégie utilitaire ils attaquent pour sa nourriture tout agent qui n’est pas plus fort qu’eux. La simulation a montré que la stratégie normative était bien meilleure pour réduire le nombre d’agressions.

(Staller *et al.*, 2001) étendent (Conte *et al.*, 1995) mais mettent l’accent sur l’étude de la relation bi directionnelle entre les émotions et les normes sociales : les normes sociales influencent le déclenchement des émotions (*p.ex.* être honteux d’avoir violé une norme) ainsi que leur régulation et leur expression (*p.ex.* il ne faut pas rire à un enterrement) ; et les émotions influencent à leur tour le comportement normatif (*p.ex.* la honte pousse les gens à respecter les normes, alors que le mépris les pousse à appliquer des sanctions pour les faire respecter). À cause de leur focus sur le niveau microscopique et sur ses interactions avec le niveau macroscopique, les auteurs ne peuvent pas se contenter du modèle d’agent très simple utilisé par (Conte *et al.*, 1995). Ils le remplacent donc par l’architecture d’agent à trois couches TABASCO, ainsi qu’une architecture BDI appelée JAM pour la partie raisonnement de leurs agents.

Les architectures BDI semblent très adaptées pour représenter deux mécanismes : l’internalisation des normes, et l’influence des normes sur le processus de décision. En effet la représentation explicite des normes permet aux agents de raisonner à leur sujet et de prendre des décisions autonomes à propos de leur internalisation (au lieu d’accepter automatiquement toutes les normes) ; selon ses désirs et croyances, un agent peut donc choisir de violer des normes (comme dans (Castelfranchi *et al.*, 2000)).

Une des principales critiques à l’encontre des architectures BDI concernant les normes est que ces architectures utilisent des normes existantes *a priori* (implémentées avant l’exécution). L’émergence et la reconnaissance des normes en cours d’exécution peut être gérée avec d’autres architectures comme EMIL (Campenn *et al.*, 2009) ; mais ce problème n’a été abordé que très récemment et marginalement dans la littérature BDI. (Criado *et al.*, 2010) étendent une architecture BDI valués pour permettre l’acceptation autonome de normes (indépendamment pour chaque agent, en fonction de ses préférences, contrairement à EMIL).

Pour conclure, les architectures BDI offrent un outil puissant pour raisonner sur les normes et les intégrer dans le processus de décision, mais elles sont plus adaptées aux simulations où les normes sont établies initialement. L’émergence de normes sociales informelles dans des groupes d’agents serait plus difficile à aborder avec une architecture BDI. Cependant, cette restriction n’est pas trop contraignante puisque les architectures BDI sont tout de même applicables dans un grand nombre de simulations où les lois et les régulations sont établies par une institution supérieure.

3.1.3. *Agents auto-explicatifs*

Le problème de l'explication du comportement remonte au développement des systèmes experts (Swartout *et al.*, 1991). En effet, il est très important pour un système expert utilisé dans un environnement d'apprentissage (comme pour l'entraînement des militaires (Gomboc *et al.*, 2005)), de fournir à ses utilisateurs des explications sur son raisonnement et son comportement. Plus généralement, (Ye *et al.*, 1995) ont montré qu'un système expert capable d'expliquer ses réponses était mieux accepté par les utilisateurs finaux car ces justifications augmentent leur confiance.

Dans le cadre des simulations à base d'agents, les bénéfices sont assez similaires. Ainsi (Helman *et al.*, 1986) ont montré que les explications étaient utiles pour aider les modélisateurs ou les experts à bien comprendre une simulation informatique. L'introduction d'agents BDI dans la simulation est particulièrement intéressante de ce point de vue, car ceux-ci peuvent expliquer leur comportement en termes de leurs buts (pourquoi ils ont agi ainsi) et leurs croyances (quelles étaient les raisons de leur choix d'action), des concepts très intuitifs pour les êtres humains car ils correspondent directement à leur propre langage naturel et à leur mode de pensée (Norling, 2004). Par exemple (Harbers *et al.*, 2010) ont étudié l'utilité des agents BDI pour fournir des explications à l'étudiant dans une simulation pour l'entraînement. De plus ces explications peuvent devenir bien plus précises en ajoutant des concepts supplémentaires (*p.ex.* normes...) au modèle BDI basique.

Les concepts BDI facilitent donc la création d'explications plus efficaces, mais ils impliquent aussi un processus de raisonnement beaucoup plus complexe. Cela n'est pas nécessairement rhédibitoire à l'utilisation des agents BDI, à condition que ces agents soient systématiquement équipés d'une capacité d'auto-explication. En effet (Harbers *et al.*, 2010) soulignent que sans cette capacité, leur comportement et donc toute la simulation deviennent presque impossible à comprendre, aussi bien par les informaticiens qui l'implémentent que par les modélisateurs ou experts qui essaient de l'utiliser. Pour résumer, l'introduction d'agents BDI dans la simulation facilite la génération d'explications par les agents, mais rend aussi cette capacité indispensable à la bonne compréhension de la simulation.

3.1.4. *Planification*

Les plans sont des séquences d'actions qu'un agent peut exécuter pour atteindre son(ses) but(s). Dans certaines simulations, un modélisateur peut vouloir que les agents reconsidèrent leur comportement en permanence, en fonction de leur situation et de leur(s) but(s), par des mécanismes qui ne peuvent pas être réduits à de simples règles si-alors-sinon. Cela peut induire, par exemple, de modéliser des processus de décisions complexes, de fournir à l'agent une librairie de plans compréhensibles ou de l'équiper des capacités essentielles à la planification. Si on considère l'application de la simulation multi-agent dans le seul domaine de la gestion de crise, il est déjà possible de trouver des agents dotés d'un large panel de capacités de planification.

PLAN C⁶ utilise des agents réactifs dotés de simples plans scriptés pour modéliser les unités médicales d'intervention dans des simulations d'accident à grande échelle, afin d'explorer les conséquences sur le plan médical de différentes stratégies de réponse. DrillSim (Balasubramanian *et al.*, 2006) a pour but d'évaluer l'apport de la collecte et de la diffusion d'informations dans un scénario d'évacuation incendie ; cette simulation comprend des agents représentant les secouristes et les victimes évacuées, modélisés à l'aide d'un réseau de neurones artificiels récurrents permettant de choisir le cours de leurs actions, et utilisent un algorithme A* pour la planification de leurs mouvements de bas niveau. D-AESOP (Buford *et al.*, 2006) modélise les opérations de secours médicaux en situation de crise, dans le but de favoriser une meilleure organisation, direction et utilisation des ressources dans une situation de réponse à une crise. Les développeurs de ce modèle prennent en compte l'influence d'un nombre beaucoup plus important d'informations (sociales, médicales, géographiques, psychologiques, politiques et technologiques) sur le comportement de l'agent. Par conséquent, ils affirment avoir besoin d'un modèle d'agent riche et choisissent le paradigme BDI dans lequel ils remplacent les buts par des plans de "situation" qui capturent une vision plus large de la situation de catastrophe.

Dans la version standard du paradigme BDI (cf section 2.4), la planification s'appuie sur une bibliothèque prédéfinie de plans spécifiant leur contexte d'application et le(s) but(s) qu'ils cherchent à atteindre. Mais il existe aussi plusieurs extensions du paradigme de base permettant de doter les agents de capacités essentielles de planification : par exemple de planification classique (*p.ex.* (de Silva *et al.*, 2009)), qui leur permettent de construire des plans en tenant compte des connaissances du domaine disponibles ; ou bien planification hiérarchique (*p.ex.* (Sardina *et al.*, 2006)), leur permettant de générer des plans de haut niveau en se basant sur un raisonnement hypothétique sur la valeur relative des plans.

De telles extensions ont été utilisées dans le cadre de simulations multi-agent. Par exemple (Shendarkar *et al.*, 2008) utilisent des agents dotés de capacités de planification en temps réel dans des simulations de foule. JAMES (Schattenberg *et al.*, 2001) est un environnement Java de modélisation d'agents pour la simulation, qui supporte le développement d'agents BDI ayant des capacités de planification. Les auteurs affirment qu' "ajouter des capacités de planification à l'architecture d'un agent offre l'avantage d'une représentation explicite des connaissances combinée avec une procédure de décision stable et robuste pour le comportement stratégique de l'agent"⁷.

3.1.5. Interactions complexes via communication

Dans la communauté programmation orientée-agent, en parallèle du développement des architectures BDI, de nombreux travaux se sont intéressés au développement et à la formalisation de Langages de Communication entre Agents (LCA, *p.ex.* KQML

6. <http://www.nyu.edu/ccpr/laser/plancinfo.html>

7. "adding planning capabilities to an agent architecture offers the advantage of an explicit knowledge representation combined with a stable and robust decision procedure for the strategic behavior of the agent".

(Finin *et al.*, 1994) ou FIPA-ACL (FIPA, 2002a) pour ne citer que les plus influents) et de Protocoles d'Interaction (PI, *p.ex.* la formalisation du protocole d'appel d'offres Contract-Net par le consortium FIPA (FIPA, 2002b)). Ces LCA et PI ont deux avantages principaux pour la simulation.

Tout d'abord, ils permettent aux agents de communiquer d'une manière réaliste (proche de la communication humaine). En effet les sémantiques de ces langages explicitent les liens étroits entre la communication et les états mentaux des agents. De plus ces formalismes facilitent la modélisation de l'effet de la communication sur le comportement des agents : les messages reçus modifient leurs états mentaux qui à leur tour ont une influence sur leur comportement. Par exemple dans le contexte des systèmes de transport intelligents, (Fernandes *et al.*, 2008) proposent un simulateur de trafic routier dans lequel les véhicules peuvent communiquer entre eux et doivent tenir compte de l'influence de la communication sur les conducteurs. Ces auteurs suivent la méthodologie de modélisation Prometheus (Padgham *et al.*, 2002) et utilisent donc des agents BDI communiquant par le langage et les protocoles de FIPA.

Le second intérêt concerne les PI, qui peuvent être particulièrement utiles pour coordonner des essais d'agents comme ceux manipulés par les simulations de sauvetage. Un certain nombre de protocoles de coordination de tâches ont donc été formalisés et standardisés, en particulier par le consortium FIPA et certaines plate-formes intègrent la gestion des LCA dans le but d'utiliser ces PI. Par exemple, (Sakellariou *et al.*, 2008) ont étendu la plate-forme NetLogo (Wilensky, 1999) en proposant des bibliothèques facilitant le développement d'agents BDI échangeant des messages dans le langage FIPA-ACL et vérifiant les protocoles FIPA. Ils ont développé un exemple de simulation de feu de forêt où les agents pompiers coopèrent et se coordonnent en utilisant le protocole Contract-Net. Un autre exemple est le modèle présenté dans (Chu *et al.*, 2009). Ce modèle représente, sur la plate-forme GAMA (Amouroux *et al.*, 2007), une opération de sauvetage dans laquelle la coordination entre pompiers, policiers et ambulances est basée sur la communication à l'aide des primitives FIPA.⁸

Les mécanismes de communication évoluée entre des agents BDI sont également utilisés pour modéliser des comportements de négociation. Par exemple (Traum *et al.*, 2003) présentent une simulation visant à améliorer les qualités de leader des humains en les faisant interagir avec des partenaires virtuels aussi crédibles que possible, dans des mondes virtuels mêlant agents artificiels et humains.

Pour conclure, lorsque l'on utilise des agents BDI, l'emploi de mécanismes de communication complexe avec des LCA et des PI est assez aisée, car ceux-ci sont basés sur les mêmes fondations philosophiques et les mêmes concepts. De plus leur emploi combiné peut fournir de très puissants mécanismes de coordination, coopération ou négociation via la communication, lorsque c'est utile pour le modèle.

8. Néanmoins ce dernier exemple n'intègre pas d'agents BDI pour des raisons développées dans la section 3.4.1.

3.2. *Influence du but de la simulation*

(Axelrod, 2005) admet que selon le but du modèle l'accent doit être mis soit sur sa simplicité soit sur sa précision. Globalement, quand la simulation vise à enrichir notre compréhension d'un phénomène, le modèle doit rester simple pour faciliter l'analyse des résultats ; cela facilite également sa réplication ou son extension. D'un autre côté, quand on souhaite reproduire fidèlement un phénomène et que le réalisme et la précision sont importants, il peut être nécessaire d'avoir un modèle plus complexe.

Pour être plus précis, on considère la typologie des buts de la simulation proposée par (Axelrod, 2005). Dans les paragraphes suivants nous détaillons chacun d'entre eux, fournissons des exemples de simulations ayant ce but, et discutons si les agents BDI sont appropriés ou non pour atteindre ce but, et pourquoi.

3.2.1. *Prédiction*

Les simulations peuvent être utilisées pour prévoir les sorties possibles à partir de données complexes en entrée et d'hypothèses sur les mécanismes sous-jacents. Par exemple (Gunderson *et al.*, 2000) appliquent des techniques de fouille de données à des statistiques pour créer des simulations qu'ils utilisent pour prévoir le comportement de criminels. Ces techniques sont parfaitement adaptées à la manipulation des grandes quantités de données statistiques disponibles. Au contraire, les auteurs n'étant pas intéressés ici par les détails de la cognition des agents mais plutôt par les modèles globaux de crimes à l'échelle de la ville, le modèle BDI ne serait pas adapté.

(Shendarkar *et al.*, 2008) proposent un modèle de foule pour simuler l'évacuation en cas d'attaque terroriste ; leur but est de prédire où placer les policiers pour améliorer l'évacuation. La nouveauté de leur approche est que les agents ont une architecture BDI et que les paramètres de leur comportement sont capturés sur des humains placés dans un environnement de réalité virtuelle.

Ces deux modèles mettent en lumière l'importance pour de bonnes prédictions de la phase de collecte et de traitement de données précises, et donc la nécessité de techniques et d'outils dédiés et efficaces. L'architecture BDI est un bon candidat quand l'échelle choisie pour le modèle est au niveau de l'individu ; en effet le comportement des agents doit être aussi réaliste que possible, et doit pouvoir être adapté à de nouvelles données. Cependant il est indispensable que les données disponibles soient aussi au niveau de l'individu pour informer la construction d'un tel modèle.

3.2.2. *Exécution d'une tâche*

La simulation peut également être utilisée pour réaliser une tâche, dans le sens où les techniques d'intelligence artificielle traditionnelle (reconnaissance de la parole ou de l'image, prise de décision, raisonnement...) peuvent être vues comme la simulation des processus cognitifs humains. Contrairement aux autres buts proposés par Axelrod, il s'agit ici de simuler un individu plutôt qu'une communauté et les relations sociales en son sein. Par exemple, la communauté des Agents Conversationnels Animés (ACA)

a pour but de créer des humains artificiels dotés d’une grande palette de capacités inspirées de celles des humains (*p.ex.* dialogue, raisonnement, interactions multimodales, émotions, personnalité...) (Gratch *et al.*, 2002; Swartout *et al.*, 2006). De tels agents sont très utiles comme agents interface, compagnons virtuels (Wilks, 2010) ou équipiers de travail (Nair *et al.*, 2005).

(Maatman *et al.*, 2005) proposent, par exemple, un agent capable d’écouter son interlocuteur humain et de réagir en temps réel de manière réaliste, en se basant uniquement sur des indices de bas niveau comme le ton et le volume de la voix (à cause des difficultés actuelles de compréhension sémantique en temps réel). Par ailleurs, d’autres ACA sont dotés d’une architecture BDI dans le but d’obtenir un comportement plus complexe et réaliste. Par exemple (Traum *et al.*, 2003) dotent des agents BDI de capacités de négociation humanoïdes puis les utilisent comme tuteurs dans des jeux de rôles visant à entraîner des humains pour améliorer leurs capacités de négociation en situation de stress. Dans le cadre de scénarios issus de l’intelligence ambiante, un agent BDI doté de capacités émotionnelles peut exprimer des émotions réalistes ou détecter et prendre en compte dans son raisonnement les émotions des utilisateurs (Adam, 2007). (Nair *et al.*, 2005) ont montré que des agents BDI dotés de capacités émotionnelles calquées sur les humains réalisent un meilleur travail en équipe.

Les agents BDI semblent fournir un bon paradigme pour modéliser de manière réaliste le comportement humain dans le but de réaliser des tâches de manière plus efficace en reproduisant la façon dont les humains les réalisent.

3.2.3. Entraînement

Selon (Axelrod, 2005), certaines simulations ont pour but d’“entraîner des humains en leur proposant une représentation interactive assez réaliste et dynamique d’un environnement donné”⁹. Il remarque qu’elles ont besoin de précision et de réalisme plutôt que de simplicité. On constate, comme on l’avait mentionné en introduction de cette section, que beaucoup de ces simulations utilisent des agents complexes de type BDI, souvent augmentés de capacités supplémentaires comme les émotions ou la négociation.

Par exemple, un domaine de recherche très récent porte sur l’utilisation d’humains virtuels comme patients virtuels pour entraîner les personnels médicaux. (Kenny *et al.*, 2007; Kenny *et al.*, 2008; Kenny *et al.*, 2009) ont utilisé un humain virtuel doté d’une architecture BDI développée antérieurement dans (Gratch *et al.*, 2002), pour créer différents patients virtuels ; ceux-ci présentent de manière réaliste des maladies mentales devant être diagnostiquées ou discutées, afin d’améliorer par la pratique les compétences des médecins dans ces activités. Par exemple, Justina est une jeune femme souffrant de troubles de stress post-traumatique suite à une agression sexuelle (Kenny *et al.*, 2008) et Justin est un garçon de 16 ans ayant des troubles du comporte-

9. “train people by providing a reasonably accurate and dynamic interactive representation of a given environment”.

ment (Kenny *et al.*, 2007). Les auteurs affirment que leur approche peut être appliquée à la modélisation d’une grande variété de maladies psychiatriques.

Cette architecture d’humain virtuel a également été utilisée dans une grande variété de simulations à but militaire, par exemple dans la plate-forme *Mission Rehearsal Exercise*, simulant des scénarios de guerre afin d’entraîner les capacités de négociation (Traum *et al.*, 2003) et de prise de décision (Gratch *et al.*, 2004) des militaires américains. Les auteurs remarquent que la pratique, généralement dans le cadre d’un jeu de rôle avec un professeur, est nécessaire pour négocier efficacement ; (Traum *et al.*, 2005) ont donc enrichi cet humain virtuel avec des stratégies de conversation pour en faire un véritable tuteur de négociation. Les auteurs constatent aussi le rôle important des émotions dans la négociation, ce qui rend les rend indispensables à l’efficacité du négociateur virtuel. Celui-ci a été intégré dans un agent médecin dans un scénario de guerre, afin de former l’armée à la négociation en situation de crise.

(McIlroy *et al.*, 1996) se sont intéressés à la modélisation de la recherche opérationnelle et combinent un modèle physique du combat et un modèle du raisonnement du pilote utilisant l’architecture (D’Inverno *et al.*, 2004), une architecture basée sur le cadre BDI. (Lui *et al.*, 2002) utilisent Jack Teams, un cadre BDI de modélisation d’équipes, pour modéliser les commandants militaires dans des scénarios d’opérations terrestres. L’approche BDI est supposée permettre aux agents d’utiliser un processus de prise de décision proche de celui de l’humain.

3.2.4. Divertissement

Les simulations à but de divertissement sont très similaires à celles pour l’entraînement, à la différence que les mondes virtuels peuvent cette fois être imaginaires. Des agents autonomes (ou *bots*) utilisant diverses architectures sont intégrés dans différents jeux. Par exemple, (Small, 2008) intègre au jeu *Unreal Tournament* un simple agent à base de règles appelé Steve qui peut décider de ses actions (*p.ex.* se déplacer vers un endroit pour se soigner) rapidement en se basant sur des conditions de haut niveau (*p.ex.* un faible niveau d’énergie) dans un environnement dynamique ; pour concurrencer des joueurs humains Steve imite leur processus d’apprentissage grâce à un algorithme évolutionniste, *i.e.* ses stratégies s’améliorent au fil du temps ; cependant il est encore incapable de battre systématiquement ses opposants humains.

(Hindriks *et al.*, 2010) ont tenté de proposer une alternative aux *bots* réactifs traditionnellement utilisés dans les jeux, convaincus que des agents BDI pourraient avoir un comportement plus humanoïde et ainsi rendre le jeu plus réaliste. Ils ajoutent qu’il est aussi plus facile de spécifier les données d’observation de joueurs réels en termes d’attitudes mentales d’un agent BDI que de les traduire en termes d’une machine à états finis. Ils ont donc essayé d’intégrer le langage agent GOAL dans Unreal Tournament 2004, ce qui a soulevé deux problèmes principaux : trouver le bon niveau d’abstraction du raisonnement et raccourcir les temps de calcul du module cognitif pour permettre une réponse en temps réel. En effet, ils ont constaté que le passage à l’échelle restait un problème pour l’architecture BDI, les performances devenant mauvaises avec un plus grand nombre d’agents. (Urlings *et al.*, 2006) s’intéressent aux

équipes humains-agents dans le domaine des véhicules sans équipage. Ils ont proposé un modèle de telles équipes, l'ont implémenté dans des agents BDI sous Jack et ont utilisé le jeu Unreal Tournament comme environnement de simulation en temps réel pour démontrer les capacités de leur cadre. Les agents BDI ont aussi été utilisés dans d'autres jeux comme *Quake* (Norling, 2003) ou *Black and White* (Molyneux, 2001).

Ces exemples montrent que les agents BDI peuvent être appliqués avec succès pour créer des personnages crédibles et ainsi rendre les jeux plus réalistes. Malheureusement leur utilisation dans les jeux est freinée par des problèmes de performance lorsque le nombre d'agents croît.

3.2.5. Éducation

Les simulations à but éducatif sont aussi similaires aux simulations pour l'entraînement, à la différence qu'elles n'ont pas besoin d'un environnement aussi précis. Le point important est ailleurs : comme le dit (Axelrod, 2005) "le principal usage de la simulation en éducation est de permettre aux utilisateurs d'apprendre par eux-mêmes des principes et des relations"¹⁰.

Il existe ainsi de nombreux travaux sur les tuteurs virtuels, des agents qui simulent le comportement d'un professeur humain afin d'améliorer l'apprentissage. Par exemple, (Elliott *et al.*, 1999) étudient l'intégration des émotions dans un tuteur virtuel appelé Steve qui évolue dans un environnement en 3D et enseigne à des étudiants comment utiliser un compresseur d'air haute pression. Steve a aussi été récemment utilisé dans des simulations pour l'entraînement (scénario de maintien armé de la paix pour préparer les officiers à faire face aux situations difficiles (Rickel *et al.*, 2001)).

De plus, dans les simulations éducatives, il est utile que les agents puissent fournir des explications de leur comportement. (Harbers *et al.*, 2010) utilisent donc des agents BDI qui du fait de leur représentation explicite de leurs buts et de leurs croyances, peuvent fournir des explications en des termes facilement compréhensibles par les humains ; ces explications intuitives améliorent l'apprentissage.

3.2.6. Preuve

Un autre but possible des simulations est de fournir une preuve d'existence, comme dans le jeu de la vie de Conway (Gardner, 1970), un jeu basé sur un ensemble de règles simples gouvernant la naissance et la mort de cellules, qui peut être utilisé pour prouver ou infirmer des conjectures telles que "une population finie ne peut pas grandir indéfiniment".

La littérature fournit peu d'exemples de modèles à base d'agents ayant ce but. Un exemple très connu est le modèle de ségrégation de (Schelling, 1971) : des règles très simples régissent le mouvement des agents d'un quartier à l'autre en fonction de leur niveau de bonheur avec leurs voisins. Ce modèle montre que, même avec un assez

10. "The main use of simulation in education is to allow the users to learn relationships and principles for themselves".

fort niveau de tolérance envers les personnes différentes, on observe l'apparition de groupes de quartiers homogènes. Ce modèle est censé apporter une preuve de l'existence de ségrégation sans que les individus concernés soient racistes.

La limite entre exploration et preuve est difficile à fixer précisément. Ainsi (Axtell *et al.*, 2002) fournissent un modèle à base d'agents réactifs relativement simples de l'exode des Kayenta Anasazi afin d'explorer ses causes possibles. Bien qu'il suggère fortement que le manque de nourriture ne soit pas la seule cause d'exode, il est loin d'être considéré comme une preuve que les facteurs socioculturels ont eu une certaine influence. Il n'est cependant pas clair si c'est parce que ce modèle est trop complexe ou au contraire trop basique pour soutenir un tel raisonnement.

(Sun *et al.*, 2007) soutiennent que la simulation socio-cognitive est essentielle pour vérifier par l'expérimentation les hypothèses fonctionnelles (justifiant un phénomène social par sa fonction dans la société), généralement difficiles à prouver. Cependant les capacités cognitives rudimentaires des agents utilisés dans ces simulations réduisent leur réalisme et donc leur applicabilité dans des preuves, et empêchent aussi de prouver des relations entre les niveaux micro (cognitif) et macro (social) de la simulation. Les auteurs plaident donc pour l'utilisation d'une architecture cognitive appelée CLARION ; nous pensons que des agents BDI permettraient d'atteindre le même but.

Ainsi, selon l'hypothèse à prouver, la complexité nécessaire des agents mis en jeu n'est pas la même. Les hypothèses relatives à la survie d'une population de cellules, ou aux mouvements géographiques d'une population humaine, peuvent être prouvées avec des agents très simples ; en revanche, des hypothèses plus complexes sur les relations entre la cognition individuelle et l'organisation sociale nécessitent des agents plus complexes, et le modèle BDI pourrait être une option.

3.2.7. Découverte

Le dernier but de la simulation identifié par (Axelrod, 2005) est la découverte de connaissances. Axelrod soutient que des sociologues ont réussi à “découvrir des relations et des principes importants à partir de modèles très simples”¹¹ et défend l'idée que “plus le modèle est simple, plus il est facile de découvrir et de comprendre les effets subtils des mécanismes choisis comme hypothèses”¹².

(Sun, 2004) ont tenté de découvrir des relations et des interactions entre les niveaux micro et macro de la simulation, par exemple dans une simulation de groupes sociaux dépendant de la stratégie de gestion de la nourriture (Sun *et al.*, 2007). Ce modèle nécessitant un modèle fin d'agents, les auteurs utilisent l'architecture cognitive CLARION. (Heinze *et al.*, 2001) ont développé un modèle BDI des patrouilles maritimes, utilisé dans une simulation pour l'exploration, l'évaluation et le développement de tactiques et de procédures, afin de conseiller les forces de défense australiennes.

11. “discover[ed] important relationships and principles from very simple models”.

12. “the simpler the model, the easier it may be to discover and understand the subtle effects of its hypothesised mechanisms”.

Les simulations peuvent aussi être utilisées pour tenter de reproduire et donc de comprendre le comportement humain. Par exemple (Kamphorst *et al.*, 2009), dans le but de comprendre les processus cognitifs impliqués dans la prise de décision humaine, ont utilisé une plateforme de modélisation de décision existante, Colored Trails ; ils ont ensuite dû l'étendre avec des agents BDI programmés en 2APL afin de gagner le pouvoir explicatif nécessaire.

Les simulations à base d'agents ont également été utilisées récemment comme un outil d'analyse de la politique publique, car elles permettent de découvrir des sources possibles de résistance aux politiques, c'est-à-dire des effets négatifs imprévus. Par exemple, des transports publics plus économes en carburant peuvent réduire les émissions de carbone autour de la ville, mais l'argent économisé sur l'essence pourrait être consacré à des voyages aériens, résultant globalement en des émissions de carbone inchangées. Ces effets ne sont généralement pas intuitifs car le système est en constante évolution, avec un couplage étroit entre les acteurs et des effets non linéaires (Stermann, 2006). Ainsi (Cirillo *et al.*, 2006) ont développé une simulation multi-agent pour explorer la restructuration et la déréglementation du marché de l'énergie électrique dans l'Illinois, qui a fourni des informations sur les problèmes liés aux politiques mises en place. Il existe également de nombreux exemples de simulations agent utilisés pour étudier les politiques liées au climat, *p.ex.* (Moss *et al.*, 2001; Posada *et al.*, 2008).

Bien que l'élaboration de politiques soit un nouveau domaine d'application pour les modèles à base d'agents, le lien étroit entre l'impact des politiques et le comportement humain suggère déjà un besoin de modèles riches et réalistes, ce qu'offre le paradigme BDI. Dans le cadre d'une simulation à base d'agents pour explorer une politique économique, (Farmer *et al.*, 2009) suggèrent que "le défi majeur réside dans la spécification des comportements des agents, et en particulier dans le choix des règles qu'ils utilisent pour prendre leurs décisions"¹³.

3.3. Influence d'autres paramètres

3.3.1. Domaine d'application

(Macal *et al.*, 2005), dans leur tutoriel sur la modélisation et la simulation à base d'agents, listent divers domaines dans lesquels il en existe des applications pratiques. Nous examinons ici la pertinence de l'approche BDI dans chacun d'eux.

Dans les *simulations d'entreprises et d'organisations*, la plupart des applications industrielles impliquent des règles et des listes de tâches claires, de sorte qu'une approche basée sur les buts ne serait pas appropriée ; au contraire les modélisateurs des *marchés de consommation* pourraient traduire directement leur compréhension du processus de prise de décision d'achat en termes de concepts BDI ; dans les simulations de

13. "the major challenge lies in specifying how the agents behave and, in particular, in choosing the rules they use to make decisions".

chaînes d'approvisionnement et d'assurance, selon la complexité et l'incertitude dans les modèles, l'approche BDI pourrait également aider car elle permettrait un modèle plus complexe mais toujours transparent.

En économie, les *modèles de marchés financiers artificiels et de réseaux commerciaux* sont actuellement simplifiés à l'extrême en supposant des agents parfaitement rationnels. L'utilisation d'agents BDI gérant notamment des émotions et des normes permettrait de capturer des hypothèses plus raisonnables sur le raisonnement humain.

Dans les *simulations d'infrastructure*, les participants des *marchés de l'électricité ou des économies de l'hydrogène* doivent “effectuer des tâches diverses en utilisant des règles de décision spécialisées”¹⁴ (Macal *et al.*, 2005), ce qui pourrait être modélisé en spécifiant une bibliothèque de plans et en permettant au paradigme BDI de saisir la complexité des raisonnements. Dans des *simulations de transport*, les agents BDI seraient idéaux pour modéliser des humains et leurs diverses façons d'utiliser les systèmes de transport, chacune avec ses propres avantages ; leur nature auto-explicative permettrait aussi au modélisateur de comprendre les raisons des choix des agents.

Dans les simulations de foules et en particulier dans *la modélisation du déplacement humain et des évacuations*, l'approche BDI offre une modélisation plus précise du comportement humain et peut enrichir une simulation de foule.

Dans les simulations culturelles et sociales, par exemple de *civilisations anciennes*, l'approche BDI permet de définir des objectifs de haut niveau que les civilisations voulaient atteindre et les méthodes à leur disposition sous forme de plans. Cette approche devrait produire un modèle reflétant assez fidèlement la réalité et permettrait également aux agents d'auto-expliquer leur comportement ; le concept de norme, ainsi que l'hétérogénéité des comportements parmi les individus dans une organisation, rendent aussi l'approche BDI appropriée pour la modélisation des *réseaux organisationnels*.

Dans le *domaine militaire*, le paradigme BDI qui reflète bien la manière dont les humains pensent qu'ils pensent est parfait pour les *simulations participatives de commandement et de contrôle*.

En *écologie*, l'absence d'un processus de prise de décision dans les agents suggère que l'approche BDI n'est pas appropriée ; de même en *biologie cellulaire*, la nature simple des comportements des molécules rend inadaptée l'approche BDI. Par contre en *biologie animale*, l'approche BDI peut être appropriée pour l'étude du comportement d'un groupe d'animaux, en fonction de l'espèce et de ses capacités cognitives.

Intuitivement, les agents BDI sont mieux adaptés dans les domaines d'application où les agents représentent des humains, plutôt que dans ceux qui modélisent des entités plus simples, non humaines ou sans processus de prise de décision.

14. “perform diverse tasks using specialized decision rules”.

3.3.2. *Focus de la simulation*

Nous étudions ici l'influence du centre d'intérêt d'une simulation sur le choix de l'architecture des agents. D'abord, les simulations centrées sur la dynamique d'une population nombreuse n'ont pas besoin d'une granularité très fine et se contentent généralement d'agents assez simples, comme dans les simulations de foules où l'intérêt est d'observer un comportement global de flocage (Reynolds, 1987). Cependant pour simuler une foule plus réaliste, par exemple dans un environnement virtuel où un joueur humain est impliqué, les modélisateurs essayent d'utiliser l'architecture BDI (Shendarkar *et al.*, 2008; Silverman *et al.*, 2005).

Par ailleurs lorsque le centre d'observation se déplace du comportement global de la foule vers le comportement local des agents individuels, l'architecture BDI, permettant de doter l'agent d'un comportement plus réaliste, est plus susceptible d'être utilisée. Ceci est particulièrement vrai dans le cas d'ACA (*p.ex.* (de Rosis *et al.*, 2003)), où l'accent est mis sur un seul agent. De même dans des modèles de *bots* réalistes pour les jeux vidéos (Norling, 2003) ou des modèles des effets de la psychologie et de la physiologie sur la cognition (*p.ex.* comportement criminel (Bosse *et al.*, 2007)), il y a une prévalence des approches BDI qui permettent une description fine de la cognition des agents et de ses liens avec d'autres facteurs.

Nous pouvons en conclure que l'objet de l'observation de la simulation a une grande influence sur le choix d'une architecture d'agent : si l'accent est mis sur l'observation du comportement global du groupe d'agents, alors une architecture simple peut suffire ; mais lorsque l'observation porte sur les comportements individuels des agents, une approche plus fine et réaliste doit être favorisée.

3.3.3. *Disponibilité des données*

Pour construire un modèle aussi réaliste et utile que possible, les données disponibles sur le système réel ont une importance énorme et peuvent influencer le choix de l'architecture des agents. Par exemple, si les seules données disponibles décrivent le comportement de groupes sociaux (des statistiques sur les familles par exemple), le modélisateur ne dispose pas d'informations suffisamment précises pour donner une architecture très riche aux agents représentant les individus (Gil-Quijano *et al.*, 2007).

Il est généralement beaucoup plus facile de recueillir des données objectives et mesurables (comme l'âge, les revenus...), suffisantes pour un modèle simple à base d'agents réactifs, que de recueillir des informations subjectives expliquant le comportement en termes d'états mentaux (non observables), nécessaires pour un modèle à base d'agents complexes tels que des agents BDI. Cependant une solution est offerte par la modélisation participative : il s'agit d'immerger un homme dans un monde virtuel afin de capturer son comportement et de le reproduire dans les agents de la simulation (Shendarkar *et al.*, 2008).

Par ailleurs (Edmonds *et al.*, 2005) affirment que les modèles descriptifs (notamment à base d'agents) aident à exploiter une plus grande partie des données disponibles. En particulier les concepts du modèle BDI sont en correspondance directe

avec les informations qualitatives telles les preuves basées sur des anecdotes ou du sens commun ; ce modèle se combine donc bien avec la modélisation participative.

(Norling, 2004) remarque que “les personnes ont tendance à expliquer leurs actions à partir de leurs intentions, qui sont à leur tour expliquées en termes de leurs buts et de leurs croyances”¹⁵. Elle préconise donc l’utilisation d’une approche descriptive intuitive comme l’approche BDI car “le fait que la personne qui modélise et l’objet modélisé réfèrent aux mêmes concepts simplifie les choses”¹⁶.

Pour conclure, les modèles à base d’agents BDI peuvent d’abord sembler difficiles à concevoir à cause de l’absence de données statistiques suffisamment fines. Mais d’un autre côté ils permettent d’exploiter d’autres types de données, telles que les observations de la modélisation participative, en raison de la correspondance directe entre les concepts et la façon dont les gens expliquent instinctivement leur comportement.

3.4. Paramètres qui découragent l’utilisation d’agents BDI

3.4.1. Apprentissage

Une des principales critiques faite aux architectures BDI est exprimée ainsi par (Urlings *et al.*, 2006, p.9) : “une faiblesse que les modèles BDI d’agent partagent avec d’autres modèles à opérateur humain (à base d’agents ou non) est leur incapacité à supporter l’apprentissage chez les agents”¹⁷. C’est en particulier dû au fait que de puissants algorithmes d’apprentissage numérique existent (*p.ex.* réseaux de neurones) alors que l’apprentissage symbolique est un problème beaucoup plus complexe.

Néanmoins il existe des extensions aux architecture BDI visant à intégrer des capacités d’apprentissage dans les agents BDI. Dans le jeu vidéo Black and White (Molyneux, 2001), le personnage principal a une architecture BDI étendue avec des capacités d’apprentissage par renforcement, permettant au joueur de lui enseigner des choses. Par ailleurs, (Norling, 2004, p.4) a étendu l’architecture BDI avec la stratégie d’apprentissage RPD (*Recognition-Primed Decision*), utilisée par les experts d’un domaine : l’agent apprend progressivement à partir de ses erreurs à faire des distinctions très fines entre situations ce qui permet une sélection quasi automatique de plan.

Ces exemples montrent qu’à l’inverse des idées reçues, les agents BDI peuvent être utiles même dans des simulations nécessitant des capacités d’apprentissage.

15. “people have a tendency to explain their actions in terms of their intentions, which in turn are explained in terms of their goals and beliefs” (p.2).

16. “the fact that model builder and subject being modelled are referring to the same concepts does simplify matters”.

17. “one weakness that the BDI agent model shares with other human operator models (agent-based or otherwise) is its inability to support agent learning”.

3.4.2. *Passage à l'échelle*

Un autre frein à l'utilisation d'agents BDI dans la simulation est la lourdeur de calcul de cette approche. La simulation d'un seul agent BDI avec ses différents modules de raisonnement consomme déjà beaucoup de temps, aussi un tel outil semble peu applicable à des simulations à grande échelle contenant des milliers d'agents. Un certain nombre de travaux ont intégré des agents BDI dans des environnements temps réel comme des jeux (voir section 3.2.4) et ils tendent à confirmer cette intuition. Par exemple (Hindriks *et al.*, 2010) développent des *bots* à architecture BDI pour le jeu Unreal Tournament et concluent qu'il y a effectivement des problèmes de passage à l'échelle et de performances lorsque le nombre d'agents augmente.

Contre cet inconvénient technique, nous pouvons répondre que ce qui était impossible il y a dix ans n'est plus que difficile, pour deux raisons. La première est que la puissance de calcul des ordinateurs individuels étant en augmentation exponentielle, une simulation de calcul impossible aujourd'hui pourrait être possible dans un avenir proche. La deuxième raison est le développement des réseaux qui permettent à un modélisateur de distribuer sa simulation sur de nombreux ordinateurs pour accélérer son exécution. Par exemple (Zhao *et al.*, 2007) ont construit un agent BDI dans un environnement distribué, remplaçant partiellement le processus de décision des humains dans un système automatisé ; cette simulation ne comporte qu'un seul agent pour l'instant mais pourrait techniquement être étendue à un plus grand nombre d'agents. Cependant la distribution d'un simulateur sur un réseau d'unités de calcul est elle-même très difficile et loin d'être résolue.

Le passage à l'échelle semble donc demeurer un frein important à l'utilisation d'agents BDI dans les simulations à grande échelle. La puissance de calcul des ordinateurs personnels actuels permet toutefois l'utilisation d'agents BDI dans la simulation. La distribution de simulations offrirait une solution mais reste encore très complexe. La meilleure solution pour intégrer un grand nombre d'agents BDI dans une simulation est donc d'utiliser une plate-forme optimisée ou un cadre dédié.

3.4.3. *Manques conceptuels*

Comme souligné dans la section 2.4.4, l'architecture BDI a l'avantage d'être facilement extensible avec des concepts additionnels. Néanmoins, avant de l'intégrer dans un agent BDI, les propriétés d'un concept supplémentaire et ses liens avec les concepts existants (au moins les croyances, les désirs et les intentions, mais peut-être d'autres dans une architecture BDI déjà étendue) doivent être étudiés en profondeur.

Les programmeurs de la DSTO (la Défense australienne) ont énuméré un certain nombre de caractéristiques à ajouter au cadre BDI pour éviter aux développeurs de les re-implémenter dans tous les modèles, et rendre les applications moins contraintes. (Norling, 2004, p.3) cite comme exemple de telles caractéristiques les émotions, la fatigue, l'apprentissage, la mémoire... On pourrait aussi étendre cette liste avec des attitudes de groupe ou des engagements sociaux. En particulier, dans une société d'agents, l'architecture BDI peut être utilisée efficacement pour représenter le processus de raisonnement individuel, mais peu de formalismes ont été développés pour le processus

de raisonnement des groupes sociaux. En pratique, rien ne s’oppose à utiliser métaphoriquement une architecture BDI pour un agent représentant un groupe d’agents définis statiquement (avant le début de la simulation). Mais il y a peu de travaux sur la façon d’attribuer des états mentaux à des groupes sociaux d’agents émergeant au cours de la simulation, lorsque les attitudes du groupe doivent interagir dynamiquement avec les attitudes individuelles des agents qui le composent.

Seules quelques études théoriques ont été menées sur les concepts ci-dessus. Il existe donc un énorme manque conceptuel. Néanmoins nous restons optimistes et supposons que dans les années à venir de nombreux chercheurs vont aborder ces questions. Par exemple, des travaux récents ont commencé à analyser les liens entre l’architecture BDI et les émotions (Adam, 2007) ou les attitudes sociales (Gaudou, 2008). Ce type de travaux doivent être continués pour fournir un cadre BDI clair, bien fondé et prolongé dans un effort commun.

4. Conclusion

Dans cet article, nous avons montré les avantages associés à l’utilisation de l’architecture BDI pour la modélisation du comportement des agents dans les simulations sociales. Nous avons parlé des capacités fournies par l’architecture de base : adaptabilité, robustesse, programmation abstraite et capacité à expliquer leur comportement. Nous avons aussi discuté (et réfuté) les principaux freins à l’utilisation de cette architecture : le champ d’application – l’approche BDI est certes inadaptée pour la modélisation d’entités simples comme les bactéries, mais idéale pour modéliser l’humain ; le but de la simulation – toutes les simulations n’ont effectivement pas besoin de modèles très complexes, mais comme le reconnaît Axelrod, lorsque l’accent est mis sur la précision, la complexité est inévitable ; la difficulté du passage à l’échelle des architectures BDI – il existe pourtant des jeux vidéos en temps réel basés sur une architecture BDI ; ou les manques conceptuels de la logique BDI – qui existent, mais peuvent être corrigés grâce à des extensions telles que l’ajout des émotions ou des normes.

Ainsi l’approche BDI a des avantages évidents, et les arguments classiques contre son utilisation sont discutables, mais ce paradigme est pourtant peu utilisé dans les modèles et simulations à base d’agents. Nous suggérons que cela peut être dû au manque d’outils appropriés pour aider les concepteurs à intégrer les agents BDI dans leurs simulations. En particulier les plates-formes de simulation et de modélisation existantes n’offrent généralement pas la possibilité d’intégrer facilement des agents BDI dans les simulations, ou alors uniquement avec des capacités limitées et sans fournir de moteur d’exécution BDI (comme c’est le cas dans NetLogo (Sakellariou *et al.*, 2008)). Ainsi (Uhrmacher *et al.*, 2009) a montré “combien le comportement des agents est trivial dans le type de simulations sociales dont certains chercheurs ont dû se contenter, en raison de la limitation des outils de simulation multi-agent”.

Notre objectif est donc de remédier à ce problème dans nos travaux futurs, en proposant une extension à une plate-forme de simulation multi-agent, comme GAMA (Amouroux *et al.*, 2007), pour permettre la création et l’utilisation facile d’agents BDI.

5. Bibliographie

- Adam C., The emotions : from psychological theories to logical formalisation and implementation in a BDI agent, PhD thesis, IRIT, Toulouse, July, 2007.
- Amouroux E., Chu T.-Q., Boucher A., Drogoul A., « GAMA : An Environment for Implementing and Running Spatially Explicit Multi-agent Simulation », *PRIMA*, p. 359-371, 2007.
- Anderson J., *The architecture of cognition*, Cambridge, MA : Harvard University Press, 1983.
- Anderson J., *Rules of the mind*, Hillsdale, NJ : Erlbaum, 1993.
- Anderson J., Lebiere C., *The atomic components of thought*, Mahwah, NJ : Erlbaum, 1998.
- Axelrod R., *The Complexity of Cooperation : Agent-Based Models of Competition and Collaboration*, Princeton University Press, 1997.
- Axelrod R., « Advancing the art of simulation in the social sciences », in J.-P. Rennard (ed.), *Handbook of Research on Nature Inspired Computing for Economy and Management*, vol. 3, Hersey, PA : Idea Group, p. 16-22, 2005.
- Axtell R., *et al.* , « Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley », *Proc. of the Nat. Academy of Sciences of the U.S.A.*, 2002.
- Balasubramanian V., Massguer, Mehrotra S., « DrillSim : A Simulation Framework for Emergency Response Drills », *ISCRAM*, 2006.
- Bordini R., Hübner J., Wooldridge M., *Programming multi-agent systems in AgentSpeak using Jason*, Wiley-Interscience, 2007.
- Bosse T., Gerritsen C., Treur J., « Cognitive and social simulation of criminal behaviour : the intermittent explosive disorder case », *AAMAS*, 2007.
- Bosse T., Gerritsen C., Treur J., « Towards Integration of Biological, Psychological and Social Aspects in Agent-Based Simulation of Violent Offenders », *Simulation*, vol. 85, n° 10, p. 635-660, October, 2009.
- Bratman M., *Intentions, Plans, and Practical Reason*, Harvard University Press, 1987.
- Broersen J., Dastani M., Hulstijn J., Huang Z., van der Torre L., « The BOID architecture : conflicts between beliefs, obligations, intentions and desires », *AGENTS'01*, ACM, 2001.
- Buford J., Jakobson G., Lewis L., Parameswaran N., Ray P., « D-AESOP : a situation aware BDI agent system for disaster situation », *Agent Technology for Disaster Management*, 2006.
- Burgemeestre B., Hulstijn J., Tan Y.-H., « Agent architectures for compliance », *ESAW*, 2009.
- Busetta P., Ronnquist R., Hodgson A., Lucas A., « Jack intelligent agents-components for intelligent agents in Java », *AgentLink News Letter*, vol. 2, p. 2-5, 1999.
- Campenn M., Andrighetto G., Cecconi F., Conte R., « Normal = Normative ? The Role of Intelligent Agents in Norm Innovation », *Normative Multi-Agent Systems*, 2009.
- Castelfranchi C., Dignum F., Jonker C. M., Treur J., « Deliberative normative agents : Principles and architecture », *ATAL'99*, n° 1757 in *LNCS*, Springer-Verlag, 2000.
- Cecconi F., Parisi D., « Individual versus social survival strategies », *JASSS*, 1998.
- Chu T.-Q., Drogoul A., Boucher A., Zucker J.-D., « Learning of Independent Experts' Criteria for Rescue Simulations », *J. of Univ. Computer Science*, vol. 15, n° 13, p. 2701-2725, 2009.
- Cirillo R., Thimmapuram P., Veselka T., Koritarov V., Conzelmann G., Macal C., Boyd G., North M., Overbye T., Cheng X., Evaluating the Potential Impact of Transmission

- Constraints on the Operation of a Competitive Electricity Market in Illinois, Technical Report n° ANL-06/16, Illinois Commerce Commission, 2006.
- Cohen P. R., Levesque H. J., « Intention is Choice with Commitment », *Artificial Intelligence Journal*, vol. 42, n° 2–3, p. 213-261, 1990.
- Conte R., Castelfranchi C., « Understanding the functions of norms in social groups through simulation », in N. Gilbert, R. Conte (eds), *Artificial Societies : The Computer Simulation of Social Life*, vol. 1, Taylor and Francis Group, p. 252-267, June, 1995.
- Criado N., Argente E., Noriega P., Botti V., « Towards a Normative BDI Architecture for Norm Compliance », in N. Fornara, G. Vouros (eds), *COIN@MALLOW2010*, 2010.
- de Rosi F., *et al.* , « From Greta's mind to her face : modelling the dynamics of affective states in a conversational embodied agent », *Int. J. of Human-Computer Studies - Special issue on Application of affective computing in HCI*, vol. 59, n° 1-2, p. 81-118, July, 2003.
- de Silva L., Sardina S., Padgham L., « First principles planning in BDI systems », *AAMAS'09, IFAAMAS*, p. 1105-1112, 2009.
- DeAngelis D., Gross L., *Individual-Based Models and Approaches in Ecology*, Chapman and Hall, New York, 1992.
- Dennett D., *The intentional stance*, The MIT Press, 1989.
- Dignum F., Morley D., Sonenberg E., Cavendon L., « Towards socially sophisticated BDI agents », *ICCCN*, IEEE Computer Society, p. 01-11, 2000.
- D'Inverno M., Luck M., Georgeff M., Kinny D., Wooldridge M., « The dMARS Architecture : A Specification of the Distributed Multi-Agent Reasoning System », *AAMAS*, p. 5-53, 2004.
- Drogoul A., Vanbergue D., Meurisse T., « Multi-agent Based Simulation : Where Are the Agents ? », *MABS*, p. 1-15, 2002.
- Edmonds B., Moss S., « From KISS to KIDS : an anti-simplistic modelling approach », in P. Davidsson (ed.), *MABS*, vol. 3415 of *LNAI*, Springer, p. 130-144, 2005.
- Elliott C., Rickel J., Lester J., « Lifelike pedagogical agents and affective computing : An exploratory synthesis », in M. Wooldridge, M. Veloso (eds), *Artificial Intelligence Today*, vol. 1600 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, p. 195-212, 1999.
- Farmer J. D., Foley D., « The economy needs agent-based modelling », *Nature*, August, 2009.
- Fernandes P., Nunes U., « Multi-agent architecture for simulation of traffic with communications », *Int. Conf. on Informatics in Control, Automation and Robotics (ICINCO)*, 2008.
- Finin T., Fritzson R., McKay D., Robin M., « KQML as an agent communication language », *Proc. of the third int. conf. on Information and knowledge management*, 1994.
- FIPA, FIPA Communicative Act Library Specification, <http://www.fipa.org/specs/fipa00037/>, Foundation for Intelligent Physical Agents, 2002a.
- FIPA, FIPA Contract Net Interaction Protocol Specification, <http://www.fipa.org/specs/fipa00029/>, Foundation for Intelligent Physical Agents, 2002b.
- Gardner M., « Mathematical Games : The fantastic combinations of John Conway's new solitary game 'life' », *Scientific American*, 1970.
- Gaudou B., Formalizing social attitudes in modal logic, PhD thesis, IRIT, Toulouse, July, 2008.
- Gil-Quijano J., Piron M., Drogoul A., « Mechanisms of automated formation and evolution of social-groups : A multi-agent system to model the intra-urban mobilities of Bogota city », in

- B. Edmonds, C. Hernandez, K. Troitzsch (eds), *Social Simulation : Technologies, Advances and New Discoveries*, Idea Group Inc., chapter 12, p. 151-168, 2007.
- Gilbert N., Troitzsch K. G., *Simulation For The Social Scientist*, Open University Press, 2005.
- Gomboc D., Solomon S., Core M. G., Lane H. C., Lent M. V., « Design recommendations to support automated explanation and tutoring », *BRIMS05*, 2005.
- Gratch J., Marsella S., « A Domain-independent framework for modeling emotion », *Journal of Cognitive Systems Research*, vol. 5, n° 4, p. 269-306, 2004.
- Gratch J., Marsella S., « Some Lessons from Emotion Psychology for the Design of Lifelike Characters », *Journal of Applied Artificial Intelligence - special issue on Educational Agents - Beyond Virtual Tutors*, vol. 19, n° 3-4, p. 215-233, 2005.
- Gratch J., Rickel J., André E., Badler N., Cassell J., Petajan E., « Creating Interactive Virtual Humans : Some Assembly Required », *IEEE Intelligent Systems*, vol. 17, p. 54-63, 2002.
- Gunderson L., Brown D., « Using a multi-agent model to predict both physical and cyber criminal activity », *IEEE Int. Conf. on Systems, Man, and Cybernetics*, vol. 4, 2000.
- Harbers M., van den Bosch K., Meyer J.-J., « Explaining simulations through self-explaining agents », *Journal of Artificial Societies and social simulation*, 2010.
- Heinze C., Goss S., Josefsson T., Bennett K., Waugh S., Lloyd I., Murray G., Oldfield J., « Interchanging Agents and Humans in Military Simulation », *IAAI*, 2001.
- Helbing D., Farkas I., Vicsek T., « Simulating dynamical features of escape panic », *Nature*, vol. 407, n° 6803, p. 487-490, 2000.
- Helman D. H., Bahuguna A., « Explanation systems for computer simulations », *Proceedings of the 18th conference on Winter simulation*, ACM, p. 453-459, 1986.
- Hindriks K., De Boer F., Van der Hoek W., Meyer J., « Agent programming in 3APL », *AAMAS journal*, vol. 2, n° 4, p. 357-401, 1999.
- Hindriks K. V., van Riemsdijk M. B., Behrens T., *et al.* , « Unreal Goal Bots. Connecting Agents to Complex Dynamic Environments », *AGS*, 2010.
- Jones H., Saunier J., Lourdeaux D., « Personality, Emotions and Physiology in a BDI Agent Architecture : The PEP - BDI Model », *Web Intelligence and Intellig. Agent Techno.*, 2009.
- Kamphorst B., van Wissen A., Dignum V., « Incorporating BDI agents into human-agent decision making research », *ESAW*, 2009.
- Kenny P. G., Parsons T. D., Gratch J., Rizzo A. A., « Evaluation of Novice and Expert Interpersonal Interaction Skills with a Virtual Patient », *IVA*, vol. 5773 of *LNAI*, Springer, 2009.
- Kenny P., Parsons T. D., Gratch J., Leuski A., Rizzo A. A., « Virtual Patients for Clinical Therapist Skills Training », *7th IVA*, Paris, France, September, 2007.
- Kenny P., Parsons T., Gratch J., Rizzo A., « Evaluation of Justina : A Virtual Patient with PTSD », *8th IVA*, Tokyo, Japan, September, 2008.
- Kitano H., « Robocup rescue : A grand challenge for multi-agent systems », *ICMAS*, 2000.
- Lazarus R. S., *Emotions and adaptation*, Oxford University Press, 1991.
- Lui F., Connell R., Vaughan J., « An architecture to support autonomous command agents for onesaf testbed simulations », *SimTecT Conference*, 2002.
- Maatman R. M., Gratch J., Marsella S., « Natural Behavior of a Listening Agent », *IVA*, 2005.
- Macal C. M., North M. J., « Tutorial on agent-based modeling and simulation », *37th Winter Simulation Conference. Introductory tutorials : agent-based modeling*, p. 2-15, 2005.

- McIlroy D., Heinze C., « Air Combat Tactics Implementation in the Smart Whole AiR Mission Model », *First International SimTecT Conference*, 1996.
- Molyneux P., « Postmortem : Lionhead studios' Black and White », *Game Developer*, 2001.
- Moss S., Pahl-Wostl C., Downing T., « Agent-based integrated assessment modelling : the example of climate change », *Integrated Assessment*, vol. 2, n° 1, p. 17-30, 2001.
- Nair R., Tambe M., Marsella S., « The Role of Emotions in Multiagent Teamwork », in J.-M. Fellous, M. Arbib (eds), *Who needs emotions : the brain meets the robot*, OUP, 2005.
- Neumann M., « Norm Internalisation in Human and AI », *JASSS*, vol. 13, n° 1, p. 12, 2010.
- Newell A., *Unified theories of cognition*, Cambridge, MA : Harvard University Press, 1990.
- Norling E., « Capturing the quake player : using a BDI agent to model human behaviour », *AAMAS*, p. 1080-1081, 2003.
- Norling E., « Folk psychology for human modeling : extending the BDI paradigm », *AAMAS*, New York, 2004.
- Ortony A., Clore G., Collins A., *The cognitive structure of emotions*, Cmburg. Uni. Press, 1988.
- Padgham L., Winikoff M., « Prometheus : A Methodology for Developing Intelligent Agents », *AOSE @ AAMAS*, 2002.
- Paquet S., Bernier N., Chaib-draa B., « DAMAS-Rescue Description Paper », *Proceedings of RoboCup-2004 : Robot Soccer World Cup VIII*, Springer-Verlag, Berlin, p. 12, 2004.
- Pereira D., Oliveira E., Moreira N., Sarmento L., « Towards an architecture for emotional BDI agents », *Twelfth Portuguese Conf. on AI*, p. 40-47, 2005.
- Posada M., Hern'andez C., L'opez-Paredes A., « Emissions Permits Auctions : An Agent Based Model Analysis », in B. Edmonds, K. G. Troitzsch, C. H. Iglesias (eds), *SocialSimulation : Technologies, Advances and New Discoveries*, IGI Global, chapter 14, p. 180-191, 2008.
- Rao A. S., Georgeff M. P., « Modeling Rational Agents within a BDI-Architecture », in J. A. Allen, R. Fikes, E. Sandewall (eds), *KR'91*, Morgan Kaufmann, p. 473-484, 1991.
- Reynolds C., « Flocks, herds, and schools : A distributed behavior model », *SIGGRAPH*, 1987.
- Rickel J., Gratch J., Hill R., Marsella S., W. S., « Steve Goes to Bosnia : Towards a New Generation of Virtual Humans for Interactive Experiences », *AAAI Spring Symposium on Artificial Intelligence and Interactive Entertainment*, Stanford University, CA, March, 2001.
- Rosenbloom P. S., Laird J. E., Newell A., *The SOAR papers : research on integrated intelligence*, Cambridge, MA : MIT Press, 1993.
- Sakellariou I., Kefalas P., Stamatopoulou I., « Enhancing NetLogo to simulate BDI communicating agents », *AI : Theories, Models and Applications*, vol. 5138, p. 263-275, 2008.
- Sardina S., de Silva L., Padgham L., « Hierarchical planning in BDI agent programming languages : a formal approach », *AAMAS'06*, ACM, New York, NY, USA, p. 1001-1008, 2006.
- Savarimuthu B. T. R., Cranefield S., « A categorization of simulation works on norms », *Dagstuhl Seminar Proceedings 09121 : Normative Multi-Agent Systems*, p. 39-58, 2009.
- Schattenberg B., Uhrmacher A., « Planning agents in JAMES », *Proc. of the IEEE*, vol. 89, n° 2, p. 158-173, 2001.
- Schelling T., « Dynamic Models of Segregation », *Journal of Mathematical Sociology*, vol. 1, p. 143-186, 1971.

- Shendarkar A., Vasudevan K., Lee S., Son Y.-J., « Crowd simulation for emergency response using BDI agents based on immersive virtual reality », *Simulation Modelling Practice and Theory*, vol. 16, n° 9, p. 1415-1429, 2008.
- Silverman B., Badler N. I., Pelechano N., O'Brien K., « Crowd Simulation Incorporating Agent Psychological Models, Roles and Communication », 2005, Centre Human Mod. and Sim.
- Small R. K., « Agent Smith : a real-time game-playing agent for interactive dynamic games », *Genetic And Evolutionary Computation Conference*, p. 1839-1842, 2008.
- Staller A., Petta P., « Introducing Emotions into the Computational Study of Social Norms : A First Evaluation », *Journal of Artificial Societies and Social Simulation*, vol. 4, n° 1, p. 1-25, 2001.
- Sterman J. D., « Learning from Evidence in a Complex World », *J. of Public Health*, vol. 96, n° 3, p. 505-514, 2006.
- Steunebrink B., Dastani M., Meyer J.-J., « A Formal Model of Emotions : Integrating Qualitative and Quantitative Aspects », *ECAI'08*, IOS Press, p. 256-260, 2008.
- Sun R., *Duality of the Mind*, Lawrence Erlbaum Associates, Mah- wah, NJ, 2002.
- Sun R., « Desiderata for cognitive architectures », *Philosophical psychology*, vol. 17, n° 3, p. 341-373, December, 2004.
- Sun R., « Cognitive social simulation incorporating cognitive architectures », *IEEE Intelligent Systems*, vol. 22, n° 5, p. 33-39, 2007.
- Sun R., Merrill E., Peterson T., « From implicit skills to explicit knowledge : A bottom-up model of skill learning », *Cognitive Science*, vol. 25, n° 2, p. 203-244, 2001.
- Sun R., Naveh I., « Social institution, cognition, and survival : a cognitive-social simulation », *Mind & Society*, vol. 6, n° 2, p. 115-142, 2007.
- Swartout W., Gratch J., Hill R., Hovy E., Marsella S., Rickel J., Traum D., « Toward Virtual Humans », *AI Magazine*, 2006.
- Swartout W., Paris C., Moore J., « Explanations in knowledge systems : design for explainable expert systems », *IEEE Expert*, vol. 6, n° 3, p. 58-64, 1991.
- Tesfatsion L., « Agent-based computational economics : growing economies from the bottom up », *Artificial Life*, vol. 8, p. 55-82, 2002.
- Traum D., Rickel J., Gratch J., Marsella S., « Negotiation over Tasks in Hybrid Human-Agent Teams for Simulation-Based Training », *AAMAS*, Melbourne, Australia, July, 2003.
- Traum D., Swartout W., Marsella S., Gratch J., « Fight, Flight, or Negotiate : Believable Strategies for Conversing under Crisis », *Intelligent Virtual Agents*, Kos, Greece, 2005.
- Uhrmacher A., D. W., « Agent-Based Simulation Using BDI Programming in Jason », *Multi-Agent Systems : Simulation and Applications*, Taylor and Francis Group, p. 451-476, 2009.
- Urlings P., Sioutis C., Tweedale J., Ichalkaranje N., Jain L., « A future framework for interfacing BDI agents in a real-time teaming environment », *Journal of Network and Computer Applications*, vol. 29, n° 2 - Innovations in agent collaboration, p. 105-123, April, 2006.
- von Wright G. H., *Norm and Action*, Routledge and Kegan, 1963.
- Wilensky U., « NetLogo », 1999, <http://ccl.northwestern.edu/netlogo/>.
- Wilks Y. (ed.), *Close Engagements With Artificial Companions : Key Social, Psychological, Ethical and Design Issues*, Natural Language Processing, John Benjamins Pub Co, 2010.
- Wooldridge M., *An Introduction to MultiAgent Systems - 2nd Ed.*, John Wiley & Sons, 2009.

Ye L. R., Johnson P. E., « The impact of explanation facilities on user acceptance of expert systems advice », *MIS Quarterly*, vol. 19, n° 2, p. 157-172, 1995.

Zhao X., Son Y.-j., « BDI-based human decision-making model in automated manufacturing systems », *Int. Journal of Modeling and Simulation*, vol. 28, n° 3, p. 347-356, 2007.