Une approche facilitant la couverture et l'intelligibilité des influences des traits de personnalité sur le raisonnement rationnel des agents

F. Bouchet[a] francois.bouchet@mcgill.ca
J. P. Sansonnet[b] jps@limsi.fr

[a]SMART Laboratory, McGill University, 3700 McTavish, Montreal, Canada

[b]LIMSI-CNRS, BP 133, 91403 Orsay Cedex, France

Résumé

Nous présentons une approche systématique de l'implémentation du principe stipulant que les traits de personnalité ont une influence potentielle et effective sur le processus de décision rationnelle d'agents cognitifs. L'apport de ce travail se situe au niveau de la couverture du domaine psychologique traité, de sa généricité par rapport aux modèles d'agents rationnels utilisés. Surtout, par sa nature déclarative il facilite l'intelligibilité des relations associant les phénomènes psychologiques aux influences sur le cycle de délibération des agents.

Mots-clés : *Modélisation cognitive, Psychologie computationnelle, Traits de personnalité, Agents cognitifs.*

Abstract

We present a generic approach to the implementation of a principle stating that personality traits have an actual influence over the rational decision making process of cognitive agents. The contribution of this work concerns the coverage of the domain of psychology that is treated, the genericity of the proposal with regard to the models of rational agents used and above all, the improvement of the comprehensiveness of the relationships that linking psychological phenomena with computational operators influencing the deliberation cycle of the agents.

Keywords: *Cognitive modeling, Computational psychology, Personality traits, Cognitive agents.*

1 Introduction

1.1 Proposition

Le principe selon lequel les traits de personnalité d'un humain [5, 11] influencent son état mental (en particulier l'activation de ses émotions [19]) a été transposé à la modélisation d'agents artificiels de manière extensive, par exemple dans la communauté étudiant les personnages virtuels. L'idée que ce même principe s'applique aussi au processus de *raisonnement rationnel* des humains a été débattue en psychologie [7], mais a été encore peu étudiée au niveau de la modélisation cognitive d'agents artificiels. En effet, depuis les travaux fondateurs de Rousseau et Hayes-Roth [23, 24] portant sur l'implémentation informatique de phénomènes de nature psychologique sur le raisonnement rationnel d'agents, les études qui ont suivi (détaillées à la Section 1.2) souffrent de deux problèmes que nous désignerons par les termes de *couverture comportementale* et d'*intelligibilité des influences*.

La notion de couverture désigne en première acception l'extension du domaine des phénomènes psychologiques concernés par une implémentation informatique donnée. La plupart des travaux récents portent sur des études de cas particuliers (par exemple, elles traitent un ou deux traits de personnalité seulement). En psychologie, la question de la couverture est centrale et plusieurs larges taxonomies de traits de personnalité ont été proposées, comme le modèle 16PF de Catell [5] ou le modèle FFM [11]. Cependant, du point de vue computationnel, ces modèles ont deux inconvénients : ils sont trop génériques et pas assez descriptifs des comportements effectifs des personnes qui portent les traits. En conséquence, nous étendrons la notion de couverture à celle de couverture comportementale (intuitivement, comment se traduit en pratique le fait d'être une personne ambitieuse, conciliante, pessimiste *etc.*).

La notion d'intelligibilité désigne la transparence du mécanisme d'implémentation des phénomènes psychologiques. L'essentiel des travaux portant sur les influences des traits psychologiques ont recourt à des implémentations procédurales ou à des heuristiques "enfouies dans

le code" qui ne facilitent pas la discussion critique, en particulier avec les psychologues ; or celle-ci est indispensable car la nature de l'influence de la psychologie sur le raisonnement est encore mal connue. Il faut donc proposer une architecture flexible et observable, qui permette de tester et de suivre facilement des hypothèses sur les relations traits/raisonnement.

Dans cet article, nous proposons une architecture dédiée à l'implémentation informatique et à l'expérimentation des influences des traits de personnalité sur le raisonnement rationnel d'agents artificiels, en ayant pour objectifs de faciliter la couverture comportementale ainsi que l'intelligibilité des influences telles que définies ci-dessus. Tout d'abord, la Section 2 décrit la procédure suivie pour garantir la couverture comportementale, en enrichissant la taxonomie de traits et facettes FFM/NEOPI-R de schémas décrivant les comportements effectifs des agents. Dans la Section 3, nous traitons le problème de l'intelligibilité en montrant comment traits de personnalité et opérateurs d'influence peuvent être liés de manière déclarative, dans le cadre d'une architecture appelée *moteur de personnalité* préalablement introduite. Nous traitons ensuite dans la Section 4 la question de l'instanciation de l'architecture définie, en usant ici du cadre des agents BDI avec un modèle d'action proche d'INDIGOLOG, et en montrant comment il est alors possible d'éliciter pour ce modèle des classes d'opérateurs d'influence. La possibilité d'une implémentation concrète est finalement démontrée dans la Section 5 par le biais d'une étude de cas portant sur le trait FFM *Conscientiousness*, en définissant la liste d'opérateurs d'influence associés aux schémas comportementaux de ce trait et la matrice d'activation liant ces schémas et opérateurs. Enfin, la Section 1.2 situe notre approche par rapport à des travaux connexes récents, avant de conclure sur les perspectives ainsi ouvertes.

1.2 Travaux connexes

Agents cognitifs. Depuis plusieurs années, les architectures de type BDI on servi de base à la construction d'agent cognitifs, en particulier sous l'influence des chercheurs en systèmes multi-agents (SMA). En effet, ceux-ci ont exploré depuis le début la notion « d'agent cognitif » au moyen de théories destinées à modéliser les *capacités* de raisonnement des agents en s'appuyant sur des architectures BDI : les termes *belief*, *desire* et *intention* ne sont pas réellement pris ici au sens psychologique ; ils correspondent de fait à des notions qui restent propres à l'action rationnelle (faits, buts, plan en cours).

Récemment, les chercheurs en SMA se sont intéressés à l'intégration de notions plus proches de la psychologie dans des architectures cognitives d'agents, par exemple, en ajoutant une couche à des plateformes SMA déjà existantes, comme dans le cas de CoJACK [18][9] pour la plateforme JACK qui prend en compte des paramètres émulant certaines caractéristiques de la *physiologie* (et non la psychologie) humaine comme : la durée d'apprentissage ; les limites mémorielles (*e.g.* « perdre une croyance » si l'activation est basse ou encore « oublier l'action suivante » dans une procédure) ; la remémoration vague de croyances (*fuzzy retrieval*) ; la limitation du champ d'attention ; ou encore l'utilisation d'opérateurs, appelés modérateurs, qui altèrent le raisonnement cognitif lui-même. Là encore, les aspects psychologiques ne sont pas pris en compte ou marginalement.

Agents conversationnels. Les agents conversationnels définis par Maes [16] dès 1994, sont souvent utilisés dans las applications d'assistance aux usagers [28], en particulier dans l'Internet [25] et mettent l'accent sur les aspects dialogiques [1]. Des études ont donc été menées pour ajouter des émotions aux architectures d'agents BDI, par exemple pour prendre en compte des notions comme : la peur, l'anxiété ou encore la confiance en soi. Ceci a été effectué par l'ajout de paramètres supplémentaires comme les désirs fondamentaux, les capacités et les ressources [21]. Plus récemment, les travaux de Pélachaud *et al.*, fondés sur l'architecture d'agent conversationnel GRETA [20], font intervenir des modèles de personnalité pour l'expression des émotions (faciales, gestuelles *etc.*). Poursuivant cela dans l'architecture FATIMA, Paiva *et al.* [8] affirment qu'il faut construire des agents qui ont des apparences physiques bien distinctes mais aussi des comportements psychologiques bien différenciés.
Ces études s'intéressent aux émotions (influence des traits sur le déclenchement et l'expression posturale) plus qu'au raisonnement et l'action rationnels, et les traits y sont traités cas par cas, ce qui donne des architectures très ciblées. Enfin, la relation trait→influence étant procédurale ("règles en dur") la flexibilité et l'intelligibilité sont faibles.

Perception des traits. De manière symétrique à la problématique de l'expression des traits, dis-

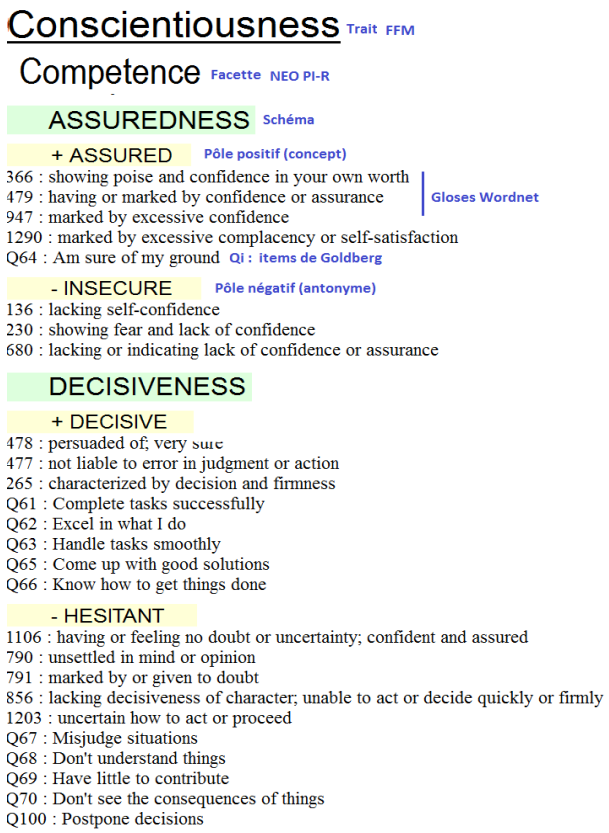

FIGURE 1 – Extrait Web de FFM/NEOPI-R/BS

cutée ci-dessus, il faut soulever le problème de la perception effective des traits implémentés par les usagers. Étant donné un agent x muni d'une personnalité $P(x)$, exprimée dans une instance de moteur de personnalité $\mathtt{MP}_i$, trois questions se posent : a) des sujets humains considèrent-ils les séquences comportementales comme des sources (*cues*) pour déceler la personnalité des autres ? Et si oui, b) sont-ils capables de distinguer deux agents rationnellement équivalents (but et situation identiques) mais ayant des personnalités différentes ? Et si oui, c) sont-ils capables de découvrir les traits attribués à ces agents ? Pour répondre, des expérimentations dédiées restent à mener. Il est cependant déjà possible de donner une réponse positive pour a), puisque cela correspond au processus « d'attribution causale » étudié de longue date en psychologie [15], et confirmé par des études récentes [22][4], portant cependant sur les émotions et non sur les traits.

2 Schémas psychologiques

Plusieurs familles de théories psychologiques décrivant la personnalité d'un individu ont été proposées depuis longtemps : la psychanalyse freudienne, les types et les traits, l'approche humaniste de Maslow et Rogers, la théorie socio-cognitive de Bandura, *etc.* Parmi celles-ci, les descriptions à base de traits ont été largement utilisées en informatique affective (*affective computing* de R. Picard au MIT) ont montré qu'elles étaient bien adaptées pour les agents cognitifs de J. Gratch à USC [13]. Nous suivrons donc cette approche de la personnalité, en utilisant la taxonomie FFM [12] qui prédomine dans les études computationnelles (*cf.* [14]).

Comme toutes les autres taxonomies de traits visant à décrire l'ensemble de la personnalité humaine, FFM présente cependant un inconvénient majeur d'un point de vue computationnel : sa généricité, puisqu'elle ne définit que cinq grandes classes de traits, souvent appelées O.C.E.A.N. (pour Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism). C'est la raison pour laquelle d'autres chercheurs en psychologie [6, 30, 31] ont fait des expériences afin d'introduire un second niveau de raffinement à FFM à l'aide de *facettes bipolaires* (c'est-à-dire le concept et son antonyme) associées à chaque trait. Nous avons opté pour celle définie par Costa et McCrae, FFM/NEOPI-R, en raison du nombre supérieur de facettes qu'elle propose (6 facettes par traits, soit 30 au total).

Malheureusement FFM/NEOPI-R reste encore très générique. Ceci pose un problème au moment de l'implémentation informatique car on ne dispose pas de correspondance psychologiquement testée entre les facettes et les comportements effectifs des individus en action : en effet, chaque facette FFM/NEOPI-R est définie par une simple glose [1]).

C'est la raison pour laquelle lors de travaux précédents nous avons été amenés à enrichir la taxonomie FFM/NEOPI-R, introduisant ainsi la ressource FFM/NEOPI-R/BS [2][2], où chaque facette est décomposée en concepts appelés *schémas comportementaux* (ou *schémas* en abrégé). Notons que FFM/NEOPI-R/BS n'est pas un nouvelle taxonomie mais seulement un *enrichissement* de FFM/NEOPI-R dont elle hérite des analyses de fond (analyses factorielles issues soit de test de personnalité, soit d'analyses lexicales). La description complète de cette approche et la documentation de FFM/NEOPI-R/BS sont disponibles en tant que ressource libre d'accès sur le

1. Phrase courte, décrivant généralement un élément de sémantique lexicale, comme pour les définitions des sens d'un mot dans les dictionnaires ou pour les définitions des *synsets* de la base lexicale WordNet.
2. BS (Behavioral Schemes) pour schémas comportementaux.

Web [3]. Chaque schéma y est défini par un couple d'ensembles de gloses : l'ensemble décrivant le concept du schéma et l'ensemble décrivant son antonyme. Ces ensembles ont été constitués par l'analyse de deux types de ressources croisées : tout d'abord nous avons eu recourt à la base lexicale Wordnet pour récolter et catégoriser les gloses définissant les synsets associés à 1 055 adjectifs de personnalité ; ensuite les catégories obtenues ont été alignées avec 300 gloses (appelées *q-items*) tirées de questionnaires utilisés en psychologie par Goldberg [4]. Chaque catégorie et son antonyme définissent alors un schéma comportemental bipolaire, souvent dénoté par le concept associé au pôle positif. La taxonomie FFM/NEOPI-R/BS (dont la construction est détaillée dans [2]) est constituée de 69 schémas bipolaires, contenant en tout 766 gloses distinctes. La figure 1 donne un extrait pour deux des schémas contenus dans le facette Conscientiousness/Competence de la ressource.

L'outil FFM/NEOPI-R/BS permet non seulement de disposer de deux fois plus de concepts pour décrire une personnalité (69 *vs.* 30), mais surtout propose pour chacun d'entre eux plusieurs gloses, ce qui se révèle utile de deux manières complémentaires : 1) côté psychologique : lors de la définition d'une personnalité, l'expert peut s'appuyer sur ces gloses pour annoter l'individu x avec le schéma +ASSURED ou -HESITANT par exemple ; 2) côté computationnel, les gloses associées aux schémas permettent d'éliciter les opérateurs d'influence ainsi que la matrice d'activation schémas/opérateurs (*cf.* l'étude de cas décrite dans la Section 5). Ainsi côté psychologique, il est permis de ne pas être d'accord avec une annotation particulière pour x, ou encore, côté computationnel, avec une association particulière entre un schéma et un opérateur, mais dans les deux cas, le processus est rendu plus précis et plus clair (par exemple que dans les études décrites à la Section 1.2) par l'architecture que nous proposons dans la section suivante.

3 Moteurs de personnalités

3.1 Structure d'un moteur de personnalité

Composantes. Un moteur de personnalité $\mathtt{MP}$ est défini comme une structure à 4 composantes $\mathtt{MP} = \langle \mathtt{O}, \mathtt{W}, \Omega_{\mathtt{W}}, \mathtt{M} \rangle$ où :
– $\mathtt{O}$ est *ontologie de personnalité*, qui doit permettre la description précise de personnalités. Dans la suite, nous utiliserons à fin d'exemple l'ensemble Σ de schémas bipolaires de FFM/NEOPI-R/BS (*cf.* Section 2), on aura donc $|\Sigma| = 69$. L'ensemble des positions positives (*resp.* négatives) est noté $+\Sigma$ (*resp.* $-\Sigma$) et leur union est $\pm\Sigma$ tel que $\pm\Sigma = +\Sigma \cup -\Sigma$ et $|\pm\Sigma| = 138$;
– $\mathtt{W} = \langle \mathtt{Wk}, \mathtt{Wm}, \mathtt{Wc}, \mathtt{Wr} \rangle$ est un *modèle de monde d'agents* qui inclut : leur structure de connaissances rationnelles $\mathtt{Wk}$ (*knowledge base*) ; leur structure mentale $\mathtt{Wm}$ (*mind*) ; leurs protocole de communication externe $\mathtt{Wc}$ (*e.g.* type FIPA *etc.*) ; leur processus de décision rationnelle $\mathtt{Wr}$. Par exemple, $\mathtt{Wr}$ peut être un modèle fondé sur l'approche BDI ou un modèle encore plus spécifique comme celui défini à la Section 4.1 ;
– $\Omega_{\mathtt{W}}$ est un ensemble d'*opérateurs d'influence* sur des règles de $\mathtt{Wr} \cup \mathtt{Wc} = \mathtt{Wrc}$;
– $\mathtt{M}$ est une *matrice d'activation* qui établit une relation sur $\pm\Sigma \times \Omega_{\mathtt{W}}$.
Les composantes $\mathtt{O}$ et $\mathtt{W}$ sont considérées comme des ressources *données* tandis que les composantes $\Omega_{\mathtt{W}}$ et $\mathtt{M}$ doivent être élicitées à partir de $\mathtt{O}$ et $\mathtt{W}$ en fonction du cas à implémenter, selon le processus décrit à la Section 5.2.

Opérateurs. Étant donné un modèle de raisonnement rationnel $\mathtt{W}$, les opérateurs d'influence $\omega \in \Omega_{\mathtt{W}}$ sont des méta-règles qui contrôlent ou altèrent la partie non structurelle de $\mathtt{W}$, *i.e.* $\mathtt{Wrc}$. Formellement, toute règle portant sur $\mathtt{Wrc}$ est un opérateur d'influence, cependant seules seront considérées comme pertinentes les règles qui auront une interprétation possible en termes des schémas de $\mathtt{O}$. En conséquence, l'élicitation de tels opérateurs est un processus de $\mathtt{Wrc} \times \mathtt{O} \longmapsto \Omega_{\mathtt{W}}$ plutôt que $\mathtt{Wrc} \longmapsto \Omega_{\mathtt{W}}$.
La plupart des opérateurs sont activés en "tout ou rien" (la règle est appliquée ou pas), mais certains peuvent prendre des paramètres d'activation. On considérera deux cas fréquents :
– Une *intensité* ou force d'application est fournie ;
– Certains opérateurs peuvent fonctionner en mode inverse ou en mode antonyme et prennent alors une *direction*, notée par le préfixe +/- (+ est souvent omis).

Matrice d'activation. Étant donné un ensemble de schémas $\sigma \in \pm\Sigma$ et un ensemble d'opérateurs $\omega \in \Omega_{\mathtt{W}}$, le(s) concepteur(s) d'un moteur de personnalité particulier doit expliciter comment $\pm\sigma_i$ sont associés à ω_i, c'est-à-dire quels schémas activent quels opérateurs. Ces relations, qui peuvent varier d'un concepteur à l'autre, sont établies au moyen d'une matrice $\mathtt{M}$

3. http ://perso.limsi.fr/jps/research/rnb/toolkit/taxo-glosses/taxo.htm
4. http ://ipip.ori.org/newNEOKey.htm

TABLE 1 – Valeurs des niveaux d'activation $\lambda_{i,j}$

Valeur	Opérateur actif	Direction	Intensité
2	oui	+	forte
1	oui	+	modérée
0	non	N/A	N/A
-1	oui	-	modérée
-2	oui	-	forte

dont les éléments sont appelés des *niveaux d'activation* $\lambda_{i,j}$, telle que $\mathtt{M} = \pm\Sigma \times \Omega_{\mathtt{W}}$ (*cf.* Table 1).

3.2 Instanciation d'un moteur

Après avoir défini un moteur de personnalité particulier $\mathtt{MP}_0$, on dispose d'une structure symbolique qui peut alors être instanciée dans diverses situations effectives $\mathtt{MP}_0(\mathtt{T}_i, P(x_i))$ qui sont définies par deux paramètres indépendants :
– $\mathtt{T}_i$ est un *topique applicatif* ;
– $P(x_i)$ est un *profil de personnalité*.

Topique applicatif. On utilisera le terme 'topique' $\mathtt{T}_i$ pour désigner une instance de $\mathtt{W}$ qui a pour rôle de fournir : a) un ensemble d'actions disponibles $\alpha_i \in \mathtt{A}(\mathtt{T}_i)$ (où $\mathtt{A}$ est une API) ; b) des informations dépendantes de l'application définissant les modalités d'exécution des actions. Par exemple, posons l'opérateur $\omega_{+safe} \doteq Sort(\{\alpha_i\}, \prec_{danger})$ qui trie un ensemble d'actions de la plus sûre à la plus dangereuse. Il a besoin que $\mathtt{T}_i$ lui fournisse une fonction de mesure $\mu_{danger} : \mathtt{A}(\mathtt{T}_i) \longmapsto [0, 1]$.

Profil de personnalité. Étant donné un individu x_i sa personnalité, dénotée $P(x_i)$, fera office de profil de personnalité de l'agent dans la situation instanciée. Intuitivement, les profils de personnalité sont souvent définis par des ensembles de mots ; des adjectifs ou des adverbes qui décrivent le comportement d'une personne agissant seule ou en société (*e.g.* on dit de $Pierre$ qu'il est *attentif*, *travailleur* et *passif*). La recherche sur les taxonomies de traits a permis d'établir des profils techniquement plus précis. En particulier, l'usage de la taxonomie FFM/NEOPI-R/BS permet de produire des profils bien fondés, selon la définition suivante :

Étant donné un individu x, sa personnalité $P(x)$ est définie par un ensemble de $|\Sigma|$ fonctions $\mathrm{p}(\sigma_i) : \Sigma \longrightarrow \{+, \asymp, -\}$ où :

- $\asymp$ signifie que vis-à-vis du schéma σ_i, le comportement de x n'est pas significativement déviant d'un comportement moyen ;
- $+$ signifie que le comportement de x est significativement déviant d'un comportement moyen, selon le pôle positif ;
- $-$ respectivement selon le pôle négatif.

Notation Si on considère les 69 schémas de Σ, les personnes ont tendance à exhiber un comportement moyen dans la grande majorité des cas. En conséquence $P(x)$ est un vecteur creux dont les éléments majoritaires ont pour valeur $\asymp$, aussi $P(x)$ est préférablement donné comme l'ensemble des schémas différents de $\asymp$. Par exemple, pour $Pierre$ on aura $P(Pierre)$ = {+ATTENTIVE, +HARDWORKER, -PASSIVE}, ignorant ainsi les 66 autres schémas pour lesquels le comportement de $Pierre$ n'est pas particulièrement notable.

4 Support des influences

4.1 Modèle de raisonnement

Choix du modèle. Les architectures d'agents rationnels furent initialement fondées sur des systèmes de règles comme SOAR de Laird puis ACT-R d'Anderson qui apporte des aspects déclaratifs à SOAR. Plus récemment, la théorie BDI de Bratman, Rao et Georgeff met en avant la primauté des désirs et des intentions dans l'action rationnelle. Les plateformes BDI sont nombreuses et utilisent des formalismes logiques (KGP, MINERVA, AGENTSPEAK, CAN ; avec des règles : 2APL ; ou temporelle : MetateM, la famille Golog) ou encore Java (JASON, JADEX, JADE, JACK) – *cf.* [17] pour une revue.

En principe, chacun de ces modèles offre, selon des modalités qui leur sont propres, des opportunités d'expression des influences issues des traits psychologiques. Dans cette étude, nous appliquerons ce principe dans le cadre BDI, en se focalisant sur le modèle INDIGOLOG [10] qui apporte à CAN les aspects situationnels dans un cadre déclaratif (en Prolog). Le modèle INDIGOLOG étant complexe, on se limite à ses principaux opérateurs dans une version simplifiée (*cf.* Table 2), suffisante toutefois pour illustrer notre propos.

Cycle BDI de l'agent. Le processus de raisonnement rationnel d'un agent de type BDI est typiquement un cycle composé de cinq étapes :

1. *Délibération* Étant donné un agent muni d'un

ensemble de désirs $\gamma_i \in \mathbb{G}$, éventuellement contradictoires, lors de la délibération l'agent sélectionne un sous ensemble d'éléments γ_i^* dès lors appelés *buts*, tels que 1) l'agent a l'intention de les faire advenir ; 2) ils ne sont pas mutuellement contradictoires.

2. *Planification* Étant donné le but en cours d'accomplissement γ^* dans γ_i^*, l'agent génère un plan rationnel $\pi^* \in \mathbb{P}$ qui a γ^* dans ses postconditions : $\gamma \in post(\pi^*)$. Un plan est une expression symbolique bien formée dans le langage de plans $\mathcal{L}_\pi$, défini dans le paragraphe suivant. Il est composé de sous-plans $\pi_i \in \mathbb{P}$ et d'actions terminales $\alpha_i \in \mathbb{A}$.

3. *Décision* Bien que le planificateur propose une expression unique π^* afin d'accomplir le but courant γ^*, cette expression contient souvent des alternatives (sous-plans ou actions équivalents vis-à-vis du but) ainsi que des options (sous-plans ou actions indifférents vis-à-vis but). En conséquence les plans sont sous-déterminés et l'agent doit choisir entre les alternatives et décider s'il exécute ou non les options. En tout état de cause, le résultat de l'étape de décision est l'élection de l'action terminale suivante à exécuter α^* dans le plan π^*.

4. *Exécution* L'agent exécute l'action élue α^*. Dans la plupart des architectures BDI cette étape est implicite (c'est-à-dire vue comme une boîte noire). Un point original de notre architecture est que nous la prenons en compte ; ceci est possible grâce aux données fournies par $\mathtt{T}_i$. Cela permet d'offrir un support aux opérateurs d'influence qui portent sur la manière dont les actions sont exécutées (notions d'intensité et de gradualité).

5. *Évaluation* L'agent reçoit du topique applicatif des informations sur les conséquences effectives de l'exécution de α^*. Cela permet à l'agent d'évaluer ses performances et de les comparer à ses attentes pour en tirer des conséquences 1) en termes rationnels pour le prochain tour du cycle BDI et 2) en termes psychologiques pour la mise à jour de son état mental $\mathtt{Wm}$.

Langage de plan. Le langage de plan $\mathcal{L}_\pi$ permet de définir des expressions composées de deux sortes d'entités : sous-plans et actions terminales. Les actions sont alors définies en termes des fonctions dépendantes de $\mathtt{T}_i$. Dans la suite, s'il n'y a pas ambiguïté, nous dirons plans pour sous-plans, actions pour actions terminales, et fonctions pour fonctions dépendantes de $\mathtt{T}_i$.

TABLE 2 – Opérateurs procéduraux des plans

Nom	Op.	Description générale ($x_i = a_i \vert \pi_i$)
seq	;	$x_1; x_2$: Done(x_1) = précond pour exécuter x_2
alt	$\vert$	$x_1 \vert x_2$: un seul x est tiré au sort et exécuté
par	$\Vert$	$(x_1 \Vert x_2) \equiv (x_1; x_2) \vert (x_2; x_1)$ les x sont exhaustivement tirés au sort et exécutés
case	$\mapsto$	$\mathrm{guard}_1 \mapsto x_1, \mathrm{guard}_2 \mapsto x_2$ guard_i sont des préconds explicites pour que x_i soit exécutable. Si plusieurs gardes sont vraies, une est tirée au sort et son x est exécuté

Les plans sont des expressions de la forme :
$\pi = \langle p_i, \gamma_i, pre, post, attr, corps \rangle$ où
– p_i est un identifiant unique de plan. Ci-dessous les plans sont dénotés π ou p_i.
– pre est l'ensemble des préconditions rationnelles du plan (noté $pre(p_i)$), *i.e.* les faits qui *doivent* être vrais dans le monde pour que π soit applicable.
– $post$ est l'ensemble des postconditions rationnelles du plan (noté $post(p_i)$), *i.e.* les faits qui seront *nécessairement* ou *possiblement* vrais dans le monde après l'exécution de π.
– $\gamma_i \subseteq post(p_i)$ est l'ensemble des buts *associés au plan*. Il faut noter que les agents n'ont pas forcément l'intention de rendre vraies toutes ces postconditions mais seulement celles qui sont en intersection avec γ_i (*cf.* le scénario du bombardier stratégique de Bratman [3]).
– $attr$ est l'ensemble des attributs définissant le plan au regard des influences psychologiques.
– $corps$ est une expression définissant comment le plan est implémenté en termes de sous-plans et d'actions terminales, à l'aide d'opérateurs de composition définis dans la Table 2.

Les actions sont des expressions de la forme :
$\alpha = \langle a_i, \gamma_i, pre, post, attr, appel \rangle$ où $appel$ est un appel de fonction $f_i \in \mathbb{F}$, exécutable dans $\mathtt{T}_i$ et où les autres éléments sont définis comme pour les plans ci-dessus.

Les fonctions sont fournies par $\mathtt{T}_i$ via une API $\mathbb{F}$ de la forme : $\varphi = \langle f_i, args, params \rangle$ où
– f_i est un identifiant unique.
– $args$ est l'ensemble des descriptions des arguments *rationnels* de la fonction f_i. Ils sont strictement opératoires.
– $params$ est l'ensemble des descriptions des arguments *modaux* de la fonction f_i. Ils correspondent aux propriétés non fonctionnelles (*e.g.* vitesse d'exécution, consommation de ressources, *etc.*). On posera $args \cap params = \emptyset$.

Les postconditions $post(a_i)$ de l'action a_i peuvent être décomposées en deux catégories :
– les postconditions nécessaires (notées $post_N$) sont forcément vraies après l'exécution de a_i,

TABLE 3 – Classes d'opérateurs d'influence pour le cycle BDI

Classe (**C**ode)	Opérande	Etape BDI	Description générale
Goal	γ	Délib (1)	Contrôle la gestion des buts : élicitation à partir des désirs, ordonnancement dans la pile des buts.
Preference	$\alpha\|\pi$	Déci (3)	Comparaison et tri en ordre total ou partiel d'ensembles de $\alpha\|\pi$ considérés comme *rationnellement* équivalents. Utilisation : implémentation des attitudes mentales a propos d'entités préférées/détestées comme décrit Section 3.2–topique applicatif.
Filter	$\alpha\|\pi\|\gamma$	Délib (1), Déci (3)	Décision de conserver ou de rejeter un élément $\gamma\|\alpha\|\pi$ dans la suite du cycle BDI. Alors que les préférences fonctionnent par comparaison d'éléments, les filtres implémentent les attitudes mentales intrinsèques de l'agent vis-à-vis d'un élément.
Scheduling	π	Exéc (4)	Contrôle la phase de scheduling des plans, selon deux modes : les *politiques* d'ordonnancement portent sur les arguments des opérateurs (*e.g.* pour $S_p\|S_o\|S_d$ dans $CorpsDecl$) ; les *altérateurs* modifient la structure (*e.g.* remplacer un seq par un par est associé au schéma -DISORGANIZED de la Figure 2 via l'opérateur S_{order}).
Modal	$\varphi\|\alpha\|\pi$	Exéc (4)	Contrôle l'exécution des actions selon deux modes, activables ou non en fonction des données fournies par le topique $\mathtt{T}_i$: la *gradualité* porte sur le degré d'accomplissement final d'une l'action ; l'*intensité* (*cf.* Section 3.1) sur son déroulement.
Option	α	Exéc (4)	Ajout d'actions, non nécessaires au succès du plan, avant ou après l'exécution de l'action courante. Elles modifient les pre/post conditions mais n'altèrent pas les conditions menant aux buts (*e.g.* appliquer l'option close-door après open-door).
e**X**pectation	$\alpha\|\pi\|\gamma$	Eval (5)	Définition des attitudes mentales de l'agent au sujet des résultats *possibles* des actions exécutées selon trois mode principaux : *espérance*, *crainte* et *neutre*. Note : ils ont une attitude neutre vis-à-vis des résultats *nécessaires* de leurs actions.
Appraisal	$\alpha\|\pi\|\gamma$	Eval (5)	Définition de l'impact sur le modèle mental de l'agent de la différence entre les résultats espérés/craints et les résultats constatés après l'exécution d'une action.

e.g. étant donnée a_0 associée à la fonction lancer-un-dé, alors $face = i$; $i \in [1..6]$ appartient à $post_N(a_0) \subseteq post(a_0)$;
– les postconditions possibles ($post_P$) ne sont que potentiellement vraies *e.g.* $face = 6 \in post_P(a_0) \subseteq post(a_0)$.

Le corps d'un plan a pour syntaxe :

$$\begin{array}{ll} \text{Corps} & \rightarrow \text{CorpsProc} \mid \text{CorpsDecl} \mid a_i \\ \text{CorpsProc} & \rightarrow \text{ProcOp}[\text{Corps}|p_i|a_i, \ldots] \\ \text{CorpsDecl} & \rightarrow \langle S_p, S_o, S_d \rangle \\ S_p|S_o|S_d & \rightarrow \{p_i|a_i, \ldots\} \end{array}$$

où :
– CorpsProc est composé d'opérateurs procéduraux (ProcOp) tels que définis dans la Table 2.
– CorpsDecl est une expression déclarative composée de trois ensembles d'éléments $p_i|a_i$: S_p (éléments préférés), S_o (élements optionnels) et S_d (élements à exécuter par défaut [5]). Dans chacun des ensembles, les éléments sont considérés par le planificateur comme rationnellement équivalents vis-à-vis du but poursuivi. Par contre la distinction entre 'préférés', 'optionnels' et 'défauts' offre une opportunité aux opérateurs d'influence de s'exercer au moment de la sélection de l'action à exécuter.

5. Ce sont des éléments $p_i|a_i$ tels que $pre(p_i|a_i) \neq \emptyset$. En effet pour être rationnellement corrects, les plans doivent à chaque étape posséder au moins un $p_i|a_i$ exécutable ; S_d permet d'éviter les blocages.

4.2 Classes d'opérateurs

Les opérateurs d'influence peuvent être appliqués à quatre types d'opérandes : des buts (γ), des actions (α), des plans (π) ou des fonctions (φ). Étant donné un modèle particulier de raisonnement, on peut distinguer plusieurs grandes classes d'opérateurs. Pour le modèle de raisonnement défini à la Section 4.1, il est possible d'exhiber huit classes dont les descriptions générales sont listées dans la Table 3.

5 Etude de cas

5.1 Moteur de personnalité : $\mathtt{MP}_\mathbf{C}$

Afin de donner une première validation de l'approche proposée, nous détaillons dans cette section une étude de cas pour un moteur de personnalité $\mathtt{MP}_\mathbf{C} = \langle \mathtt{O}, \mathtt{W}, \Omega_\mathtt{W}, \mathtt{M}_\mathbf{C} \rangle$ composé des éléments suivants :
– O est FFM/NEOPI-R/BS, restreinte pour cette étude au trait FFM Conscientiousness, choisi comme exemple car il est représentatif des phénomènes que l'on rencontre dans les quatre autres traits de FFM.
– W n'est pas entièrement détaillé dans cet article (pour Wk et Wm voir [26] et pour Wc [27]).

TABLE 4 – 30 opérateurs pour Conscientiousness

$Classe_{nom}$	Glose (description générale) du pôle +
$G_{forget}+$	oublie γ (*e.g.* efface de la pile des buts)
G_{quick}	délibère rapidement (*e.g.* prend toujours le premier élément d'une liste d'alternatives)
$G_{deduct}+$	contrôle la 'profondeur' de déduction
$G_{prudent}$	prend en compte les conséquences de γ
$G_{nowaste}$	déteste les γ qui consomment des ressources
$G_{knowledge}+$*	possède une grande base de faits
$G_{reason}+$	* possède une grande base d'heuristiques
P_{easy}	préfère les $\gamma\|\pi$ faciles
P_{safe}	préfère les $\gamma\|\pi$ non dangereux
P_{clean}	préfère les $\gamma\|\pi$ propres/nets
$P_{ambitious}$	préfère les $\gamma\|\alpha\|\pi$ ambitieux
F_{lawful}	rejette les $\gamma\|\alpha\|\pi$ non légaux
F_{safe}	rejette les $\gamma\|\alpha\|\pi$ dangereux
S_{order}	exécute $\alpha\|\pi$ dans l'ordre donné
$S_{predictable}$	exécute $\alpha\|\pi$ toujours dans le même ordre
M_{focus}	exécute $\alpha\|\pi$ avec concentration mentale
$M_{precision}$	exécute $\alpha\|\pi$ avec précision
$M_{tenacity}$	persévère si difficultés/échecs lors de $\alpha\|\pi$
M_{quick}	exécute $\alpha\|\pi$ rapidement
M_{dopar}	contrôle la gradualité de $\alpha\|\pi$ (*cf.* Table 3-M)
$M_{nowaste}$	ne gaspille pas de ressources lors de $\alpha\|\pi$
M_{clean}	exécute $\alpha\|\pi$ de manière 'propre/nette'
$O_{extra}+$	tendance à ajouter des actions optionnelles
O_{play}	ajoute des actions 'agréables' à $\alpha\|\pi$
$O_{cleanup}$	ajoute des actions de 'nettoyage' après $\alpha\|\pi$
$X_{success}$	attend normalement le succès de $\gamma\|\alpha\|\pi$
X_{hope}	espère fortement que $poss(\alpha\|\pi)$ advienne
$X_{choiceConf}$	sûr de ses choix lors l'élaboration de $\gamma\|\alpha\|\pi$
$A_{responsible}$	se sent responsable des résultats de $\gamma\|\alpha\|\pi$
$A_{success}$	a une sensibilité forte au succès

Les opérateurs ont aussi un pôle négatif (antonyme) *e.g.* $-O_{cleanup}$ = "qui ajoute des actions qui salissent ou mettent en désordre après $\alpha|\pi$ "
\+ Opérateur n'ayant pas d'antonyme
* Opérateur appartenant aussi au raisonnement rationnel

Le processus de décision rationnelle Wr est celui décrit Section 4.1.
– $\Omega_{\mathtt{W}}$ est la liste des opérateurs d'influence associés au trait FFM Conscientiousness. La liste des traits élicités est décrite dans la Table 4.
– $\mathtt{M}_{\mathbf{C}}$ est la sous-matrice d'activation reliant les schémas du trait Conscientiousness aux opérateurs associés tels que décrits dans la Table 4.

5.2 Processus de construction

Opérateurs. Le processus de construction des opérateurs d'influence est mené en confrontant les ensembles de gloses associés aux schémas de O avec les classes d'influences engendrées par Wr. Par exemple, si on considère les gloses contenues dans le schéma ASSUREDNESS (pôles + et -) de la Table 1 et qu'on les confronte à la Table 3, on peut éliciter des opérateurs comme par exemple : P_{safe} (classe **P**reference), $M_{tenacity}$ (classe **M**odal), X_{hope} (classe e**X**pectation), *etc.* Ces opérateurs ont un préfixe de direction (+/-) *e.g.* $-P_{safe}$ veut dire « qui recherche les actions dangereuses »). Pour l'implémentation effective d'un opérateur comme P_{safe} par la classe **P**reference voir la Section 3.2–topique applicatif.

Facettes ⇒		Competence		Orderliness		Self-Discipline			Achievement -striving	
Schémas ⇒ Opérateurs ⇓	Agent générique (défaut) ⊕ ⊖	ASSURED ⊕ INSECURE ⊖	DECISIVE ⊕ HESITANT ⊖	METHODICAL ⊕ DISORGANIZED ⊖	TIDY ⊕ MESSY ⊖	CONCERNED ⊕ UNCONCERNED ⊖	ATTENTIVE ⊕ INATTENTIVE ⊖	MINDFUL ⊕ FORGETFUL ⊖	AMBITIOUS ⊕ UNAMBITIOUS ⊖	HARDWORKER ⊕ LAZY ⊖
G_{forget}	1							0 2		0 2
G_{quick}	1		2 -2							1 -2
G_{deduct}	1									
$G_{prudent}$	1									
$G_{nowaste}$	1									
$G_{knowledge}$	1									
G_{reason}	1									
P_{easy}	1									0 -2
P_{safe}	1	0 2								
P_{clean}	1				2 -1					
$P_{ambitious}$	1								2 0	
F_{lawful}	1									
F_{safe}	1						2 0			
S_{order}	1			2 -1						
$S_{predictable}$	1			2 -1						
M_{focus}	1	2 -1	1 0	2 -1		2 0	2 0		2 0	2 -2
$M_{precision}$	1		1 0			2 0	2 0		2 0	2 -2
$M_{tenacity}$	1	2 -2	2 -1			2 0			2 -2	2 -1
M_{quick}	1		1 -2							1 -2
M_{dopar}	1			2 -2					0 2	2 -2
$M_{nowaste}$	1			2 0						
M_{clean}	1				2 -1					1 -2
O_{extra}	1			0 2						2 0
O_{play}	1									0 0
$O_{cleanup}$	1				2 0					2 0
$X_{success}$	1	2 -2							2 0	
X_{hope}	1	2 -2							2 0	
$X_{choiceConf}$	1	2 -2	2 -2							
$A_{responsible}$	1					2 0				2 -1
$A_{success}$	1								2 -2	1 -1

FIGURE 2 – Extrait de la sous matrice d'activation $\mathtt{M}_{\mathbf{C}}$ (facettes 1 – 4).

Matrice d'activation. La matrice d'activation est alors construite en associant les schémas de O aux opérateurs au moyen de poids $\lambda_{i,j}$ définissant l'intensité d'activation de l'opérateur ainsi que sa direction (+/-). Pour le moteur $\mathtt{MP}_{\mathbf{C}}$, un exemple de sous-matrice d'activation $\mathtt{M}_{\mathbf{C}}$ est donné (sous forme d'un extrait limité aux quatre premières facettes du trait Conscientiousness) dans la Figure 2.

Pour ne pas surcharger la lecture, les valeurs $\lambda_{i,j}$ correspondant à un comportement considéré comme moyen/normal sont factorisées en tête dans la colonne « agent générique (défaut) » et représentés par des cellules vides.

Définition d'une personnalité cohérente. La personnalité $P(x)$ d'un individu x, telle que définie Section 3.2, utilise la matrice M pour activer les

opérateurs qui à leur tour influencent le cycle de délibération de l'agent. La cohérence des traits dans $P(x)$ est une question appartenant au domaine de la psychologie : par exemple, les traits +AMBITIOUS et -INSECURE co-occurrent rarement, tandis que -MESSY et -LAZY co-occurrent fréquemment. Néanmoins, en théorie, pour un x donné, les psychologues peuvent choisir des traits qui activent un opérateur w_i de manière conflictuelle. Par exemple, si x est à la fois +AMBITIOUS et -LAZY, alors selon $\mathtt{M_C}$, il active les opérateurs M_{focus} et $M_{precision}$ avec les poids +2 et -2. Dans d'autres cas, l'écart est moindre (*e.g.*-1 et -2) et peut être résolu automatiquement (*e.g.* le plus fort l'emporte). La question de la cohérence, essentielle en psychologie, trouve dans notre modèle une représentation formelle et un moyen d'expérimenter et d'observer les effets des choix.

5.3 Discussion

– *Pertinence des opérateurs* La méthode de construction des opérateurs garantit que les opérateurs élicités sont pertinents : émanant des schémas, ils sont activés au moins une fois de manière non triviale (*i.e.* $\forall \omega_i \exists j$ tel que $\lambda_{i,j} \neq$ agent-générique (i)).
– *Complétude des opérateurs* La méthode de construction des opérateurs ne garantit pas que tous les opérateurs possibles soient bien élicités ; en pure théorie ce n'est pas possible et c'est bien là que FFM/NEOPI-R/BS joue un rôle crucial concernant la question de la couverture : fondée sur des ressources lexicales larges et reconnues, FFM/NEOPI-R/BS couvre, pour ce que la littérature en connait, les comportements associés aux traits et permet donc de limiter le risque de silence.
– *Validation des poids* Les poids $\lambda_{i,j} \in \mathtt{M_C}$ sont proposés par des annotateurs et à ce titre, ils sont susceptibles d'être discutés entre experts informaticiens et psychologues. Notre approche a le mérite de faire ressortir ce problème fondamental, bien souvent 'enfouis dans le code' des approches procédurales citées Section 1.2. Dans notre cas, le recours à une méthode déclarative via une matrice de poids plutôt que des règles procédurales, améliore l'intelligibilité et le suivi de l'association traits/comportements. De plus, elle facilite grandement le dialogue avec les psychologues devant *in fine* valider les choix.
– *Évaluation du modèle* Dans cet article, nous proposons une approche pour traiter le phénomène d'influence des traits sur les actions tel que issu de la littérature. Il ne s'agit donc pas d'évaluer directement un modèle particulier (composé d'un modèle rationnel, des opérateurs est des poids spécifiques associés) par une expérimentation. Notre objectif ici est double : montrer la faisabilité (il existe des points d'influence dans le raisonnement rationnel) et se placer au niveau de la couverture (potentiellement la matrice d'activation couvre tout OCEAN).

6 Conclusion

La notion de moteur de personnalité est fondée sur une approche qui lie de manière déclarative et flexible (via les poids de la matrice d'activation) les traits d'une taxonomie à un ensemble d'opérateurs d'influence exhibés par un modèle de raisonnement donné. Cette approche doit permettre d'observer et de discuter l'impact des choix effectués, en accord avec les psychologues.

Nous avons appliqué notre approche d'un côté au modèle BDI (INDIGOLOG) et de l'autre à la taxonomie FFM/NEOPI-R/BS qui enrichit FFM de schémas comportementaux assurant la précision de la couverture du domaine psychologique. Nous envisageons d'appliquer l'approche à d'autres modèles BDI et à d'autres parties de FFM (*e.g.* les traits interpersonnels, amorcés dans [27]). De plus, des expérimentations portant sur les questions b) et c) liées à la perception des traits par les usagers restent à mener.

Références

[1] F. Bouchet and J. P. Sansonnet. Caractérisation de requêtes d'Assistance à partir de corpus. In Jérôme Lang, Yves Lespérance, David Sadek, and Nicolas Maudet, editors, *Actes des Quatrièmes Journées Francophones Modèles Formels de l'Interaction (MFI' 07)*, volume 8 of *Annales du LAMSADE*, pages 269–276, Paris, France, 2007. Université Paris Dauphine.

[2] F. Bouchet and J. P. Sansonnet. Classification of wordnet personality adjectives in the NEO PI-R taxonomy. In *Fourth Workshop on Animated Conversational Agents, WACA 2010*, pages 83–90, Lille, France, 2010.

[3] Michael E Bratman. *Intentions, Plans, and Practical Reason*. Harvard University Press, Cambridge, MA, 1987.

[4] S. Campano, E. de Sevin, and N. Sabouret. The "resource" approach to emotions. In *Poster at AAMAS 2012*, Valencia, 2102.

[5] R. B. Cattell, H. W. Eber, and M. M. Tatsuoka. *Handbook for the sixteen personality factor questionnaire (16 PF)*. Champain, Illinois, 1970.

[6] P. T. Costa and R. R. McCrae. *The NEO PI-R professional manual*. Odessa, FL : Psychological Assessment Resources, 1992.

[7] Antonio R. Damasio. Descartes error : Emotion, reason and the human brain. New York : G.P., 1994. Putnam's Sons.

[8] Tiago Doce, Joao Dias, Rui Prada, and Ana Paiva. Creating individual agents through personality traits. In *Intelligent Virtual Agents (IVA 2010)*, volume 6356 of *LNAI*, pages 257–264, Philadelphia, PA, 2010. Springer-Verlag.

[9] Rick Evertsz, Franck E. Ritter, Paolo Busetta, and Matteo Pedrotti. Realistic behaviour variation in a BDI-based cognitive architecture. In *Proc. of SimTecT'08*, Melbourne, Australia, 2008.

[10] Giuseppe Giacomo, Yves Lespŕance, Hector J. Levesque, and Sebastian Sardina. IndiGolog : A high-level programming language for embedded reasoning agents. In Amal El Fallah Seghrouchni, Jaergen Dix, Mehdi Dastani, and Rafael H. Bordini, editors, *Multi-Agent Programming : Languages, Platforms and Applications*, pages 31–72. Springer, 2009.

[11] L. R. Goldberg. An alternative description of personality : The big-five factor structure. *Journal of Personality and Social Psychology*, 59 :1216–1229, 1990.

[12] Lewis R. Goldberg. Language and individual differences : The search for universal in personality lexicons. *Review of personality and social psychology*, 2 :141–165, 1981.

[13] Jonathan Gratch and Stacy Marsella. A domain-independent framework for modeling emotion. *Journal of Cognitive Systems Research*, 5(4) :269–306, 2004.

[14] Oliver P. John, Richard W. Robins, and Lawrence A. Pervin, editors. *Handbook of Personality : Theory and Research*. The Guilford Press, 3rd edition, 2008.

[15] Harold H. Kelley. The processes of causal attribution. *American Psychologist*, 28(2) :107–128, 1973.

[16] Pattie Maes. Agents that reduce work and information overload. *Commun. ACM*, 37(7) :30–40, 1994.

[17] V. Mascardi, D. Demergasso, and D. Ancona. Languages for programming BDI-style agents : an overview. In F. Corradini, F. De Paoli, E. Merelli, and A. Omicini, editors, *Proceedings of WOA 2005*, pages 9–15, Camerino, MC, Italy, 2005. Pitagora Editrice Bologna.

[18] Emma Norling and Franck E. Ritter. Towards supporting psychologically plausible variability in Agent-Based human modelling. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multi-Agent Systems*, 2004.

[19] Andrew Ortony, Gerald L. Clore, and Allan Collins. *The Cognitive Structure of Emotions*. Cambridge, UK, cambridge university press edition, 1988.

[20] Catherine Pelachaud. Some considerations about embodied agents. In *Proc. of Achieving Human-Like Behavior in Interactive Animated Agents, The 4th International Conference on Autonomous Agents*, Barcelona, Spain, 2000.

[21] David Pereira, Eugenio Oliveira, and Nelma Moreira. Formal modelling of emotions in BDI agents. In F. Sadri and K. Satoh, editors, *Proceedings of CLIMA-VIII*, volume 5056 of *LNAI*, pages 62–81, Porto, Portugal, 2008. Springer-Verlag.

[22] R. Pfeifer. The"Fungus Eater" Approach to Emotion : A View from Artificial Intelligence. *Cognitive Studies*, 1 :42–57, 1994.

[23] D. Rousseau. Personality in computer characters. In *AAAI Technical Report WS-96-03*, pages 38–43, 1996.

[24] D. Rousseau and B. Hayes-Roth. Personality in synthetic characters. In *Technical report, KSL 96-21*. Knowledge Systems Laboratory, Stanford University, 1996.

[25] N. Sabouret and J. P. Sansonnet. Un langage de description de composants actifs pour le Web sémantique. *Information - Interaction - Intelligence Review, Revue I3*, 3(1) :9–36, 2003.

[26] J. P. Sansonnet and F. Bouchet. Joint handling of rational and behavioral reactions in assistant conversational agents. In Helder Coehlo, Rudi Studer, and Michael Wooldridge, editors, *Proc. of the 19th European Conference on Artificial Intelligence (ECAI 2010)*, pages 1049–1050, Lisbon, Portugal, 2010. IOS Press.

[27] J. P. Sansonnet and F. Bouchet. Managing personality influences in dialogical agents. In *Submitted at ECAI 2012*, Montpellier, 2102.

[28] J. P. Sansonnet, D. Leray, and J. C. Martin. Architecture of a framework for generic assisting conversational agents. In *International Conference on Intelligent Virtual Agents IVA'06*, volume 4133 of *LNAI*, pages 145–156, CA : Marina del Rey, 2006. Springer.

[29] G. Saucier and F. Ostendorf. Hierarchical subcomponents of the big five personality factors : A cross-language replication. *Journal of Personality and Social Psychology*, 76(4) :613–627, 1999.

[30] Christopher J. Soto and Oliver P. John. Ten facet scales for the big five inventory : Convergence with NEO PI-R facets, self-peer agreement, and discriminant validity. *Journal of Research in Personality*, 43(1) :84–90, 2009.