



HAL
open science

Quand et Comment Toucher un Humain ? Un Modèle de Décision pour un Agent Touchant

Fabien Boucaud, Indira Thouvenin, Catherine I Pelachaud

► **To cite this version:**

Fabien Boucaud, Indira Thouvenin, Catherine I Pelachaud. Quand et Comment Toucher un Humain ? Un Modèle de Décision pour un Agent Touchant. Workshop sur les Affects, Compagnons artificiels et Interactions (WACAI 2020), CNRS, Université Toulouse Jean Jaurès, Université de Bordeaux, Jun 2020, Saint Pierre d'Oléron, France. hal-02933486

HAL Id: hal-02933486

<https://inria.hal.science/hal-02933486>

Submitted on 8 Sep 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Quand et Comment Toucher un Humain ?

Un Modèle de Décision pour un Agent Touchant

Fabien Boucaud[†]

Sorbonne Université, Université de
Technologie de Compiègne
CNRS UMR 7253 Heudiasyc
57 Av. Landshut, 60200
Compiègne, France
fabien@boucaud.net

Indira Thouvenin

Sorbonne Université, Université
de Technologie de Compiègne
CNRS UMR 7253 Heudiasyc
57 Av. Landshut, 60200
Compiègne, France
indira.thouvenin@utc.fr

Catherine Pelachaud

Sorbonne Université
CNRS UMR 7222 ISIR
UPMC Campus Jussieu, 75005
Paris, France
catherine.pelachaud@upmc.fr

ABSTRACT

Le toucher est un sens dont on connaît de mieux en mieux l'importance dans le développement social et le bien-être général de l'être humain, de son enfance à son âge le plus avancé. C'est un sens dont on sait qu'il participe activement à la facilitation de l'établissement des relations sociales et il constitue un canal extrêmement puissant en terme de communication émotionnelle. Les agents conversationnels animés (ACA) se voient quant à eux munis de plus en plus de capacités sociales et émotionnelles. Ils peuvent créer de l'entente (*rapport*) avec des humains en exprimant leurs pensées et leurs émotions à travers des modalités aussi bien verbales que non-verbales. Le toucher semble donc constituer une modalité tout à fait pertinente pour les ACA, qui en sont encore très peu munis.

Munir des ACA de capacités de toucher social pose un certain nombre de défis qui commencent seulement à être étudiés. La forme d'un ACA peut être extrêmement différente d'une implémentation à une autre, et la façon dont on peut les munir d'un sens du toucher en est grandement influencée. Comment élaborer une boucle interactive complète dans laquelle un agent conversationnel animé peut à la fois être touché par un humain, interpréter ce toucher, choisir comment toucher l'humain en retour, et réaliser ce toucher sur l'humain ? En particulier, comment modéliser l'interaction sociale basée sur le toucher ? Quels touchers sont acceptables dans quels situations ? Comment toucher un humain de manière crédible et cohérente ?

Les travaux présentés ici proposent un modèle de décision adapté à un agent conversationnel animé *touchant*, capable d'utiliser le toucher social et affectif pour interagir avec un être humain dans une boucle interactive complète, dans un contexte d'environnement immersif.

MOTS-CLEF

Modèle computationnel d'émotion, Environnement immersif, Agent conversationnel animé, Toucher social

1 Introduction

Si le toucher est resté pendant longtemps l'un des sens les moins étudiés par la science, en particulier dans le contexte social, pour des raisons culturelles aussi bien philosophiques que religieuses, c'est un sujet d'étude qui attire de plus en plus d'attention. L'importance du toucher dans le développement social de l'individu a été montrée, aussi bien dans son enfance qu'au cours du reste des étapes de sa vie [14]. C'est qu'il semble y avoir un lien fort entre émotion et toucher, qui favorise la création d'une relation entre deux personnes [10,40]. Et en effet, là où les travaux de Hertenstein et ses collègues [18,19] ont montré la façon dont le toucher pouvait permettre, à lui seul, d'exprimer des émotions reconnaissables, même par un inconnu, les meilleurs résultats semblent toujours avoir lieu lorsque les deux personnes se connaissent déjà et ont une relation intime ou d'amitié proche [8,30,40]. L'utilisation de petits compagnons robotiques tactiles a même montré leurs bienfaits pour le bien-être et la santé des personnes isolées socialement [33]. Le toucher social, c'est-à-dire utilisé dans un contexte social d'interaction interpersonnelle, semble donc être un puissant canal de communication des émotions, et un facilitateur du développement des relations sociales.

Dans un contexte où les recherches dans le champ de l'informatique émotionnelle (*affective computing*) sur les agents conversationnels animés (ACA) ont montré à la fois la capacité de ces agents à créer du lien avec un humain [17], et l'intérêt que cela peut avoir, par exemple pour permettre la communication dans des situations de stigmatisation [26,27], le toucher social pourrait ainsi s'avérer une modalité d'interaction extrêmement pertinente. Dans quelle mesure la modalité tactile permettrait-elle à un agent conversationnel animé de plus facilement développer un lien social et exprimer des émotions avec un humain ?

Introduire le toucher social comme nouvelle modalité d'interaction entre ACA et humain pose cependant de nombreux défis loin d'être aisés à résoudre. Les ACA peuvent prendre de multiples formes selon leurs supports (robot, VR, écran d'ordinateur ou de smartphone, etc.) et chaque support doit

répondre de manière spécifique aux questions générales qui seraient : comment faire en sorte de pouvoir toucher et être touché par un ACA ? Quand et comment toucher un humain de manière adaptée à la situation et acceptable ?

Pour s'approcher au mieux de l'expérience du toucher social entre humains dans un contexte d'environnement immersif, nous avons choisi d'utiliser un système de type CAVE™ [9]. L'avantage de ce système de réalité mixte est qu'il permet à la fois de plonger l'humain dans un environnement immersif à échelle réelle, et de le laisser garder la vision de son propre corps. Cela lui permettra, dans un contexte de toucher social, de se sentir correctement incarné et d'être libre d'utiliser les mouvements avec lesquels il est familier.



Figure 1: **Interaction entre un humain et un agent dans une salle immersive, vue de l'extérieur**

Pour qu'un agent puisse interagir de manière dynamique et autonome avec un humain dans un environnement immersif, il a besoin de trois capacités [15] : pouvoir percevoir son environnement et ce qu'il s'y passe ; pouvoir prendre en compte ce qu'il perçoit pour ensuite décider par lui-même du comportement adapté à adopter en réponse ; pouvoir réaliser le comportement choisi (une interface haptique, une animation). Avec ces trois capacités, représentées dans la figure 2, il devient possible d'avoir une interaction dynamique au long cours. En effet, la boucle interactive est dite complète : l'humain touche l'agent, l'agent perçoit ce toucher, choisi comment

répondre et effectue son toucher à son tour, et l'humain peut de nouveau toucher ou non pour continuer le cycle d'interaction.

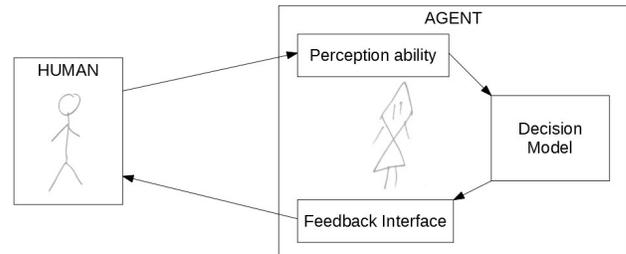


Figure 2: **Boucle interactive complète entre un humain et un agent**

A travers des travaux préliminaires, nous avons déjà étudié les possibilités pour toucher et être touché par un agent conversationnel animé virtuel [5], à travers d'un côté un module de perception du toucher pour l'agent (constitué de formes géométriques virtuelles détectant les collisions entre le corps de l'humain et le corps virtuel de l'agent), et de l'autre une interface haptique (sous forme de manchon muni de moteurs vibrants) pour que l'agent puisse envoyer des signaux haptiques (donc des touchers) à l'humain. Ces travaux ont montré que si les sensations haptiques manquaient de réalisme, elles étaient tout de même assez bien reconnues en termes de types de toucher (un coup, une caresse, etc.) En revanche, il y avait beaucoup à redire sur le comportement de l'agent, qui n'était absolument pas crédible, en particulier lorsqu'il recevait le toucher. Nous proposons donc de nous pencher ici sur la question du modèle de décision de l'agent, qui pourra lui permettre de gagner en autonomie et en crédibilité au cours de l'interaction.

2 Etat de l'art

2.1 L'interprétation du toucher chez l'humain

Le sens du toucher possède deux caractéristiques fondamentales qui rendent son interprétation compliquée, d'un point de vue social et émotionnel. En effet, les principes d'équipotentialité et d'équifinalité s'appliquent [20] : un type de toucher spécifique peut exprimer plusieurs émotions différentes, et deux types de toucher différents peuvent exprimer la même émotion. Par exemple, un coup pourrait à la fois avoir lieu dans un contexte de familiarité amicale, ou d'encouragement sportif, et dans un contexte de colère envenimée ; de même, une caresse et une étreinte peuvent être indifféremment utilisées pour exprimer de l'amour. Cela signifie qu'une simple reconnaissance du type de toucher ne suffit pas à interpréter correctement le sens social et émotionnel d'un toucher, et que, comme c'est déjà avec la plupart des comportements non-verbaux, d'autres éléments doivent être pris en compte pour désambiguïser le sens d'un toucher.

Un premier élément de désambiguïser le toucher est bien sûr le reste des signaux communicatifs utilisés. Le toucher prend

place dans des interactions sociales riches de toutes les modalités utilisées par les humains pour s'exprimer, de la parole à tous les types de communication non-verbale. Le toucher est une dimension de ces modalités et peut donc permettre d'aider à l'interprétation globale d'un comportement, de même que le comportement global, et l'état émotionnel, aident à mieux comprendre le sens d'un toucher. Toutefois, si les ACA peuvent utiliser de plus en plus de ces modalités d'expression (nous utilisons la plateforme GRETA d'animation d'ACA [37]), l'utilisation de l'environnement immersif de type CAVE™ rend plus difficile d'un point de vue technique la perception de ces éléments chez l'humain. En effet, l'humain y est libre de se déplacer à sa guise et porte des lunettes de 3D stéréoscopique, ce qui rend l'utilisation de reconnaissance des expressions faciales compliquée à mettre en place, et l'utilisation d'interfaces tangibles est limitée par le fait que tout objet physique amené dans la salle immersive peut briser la projection de la 3D.

Comme mentionné précédemment, viennent ensuite le contexte et en particulier la relation entre les deux personnes en train d'interagir. Un toucher d'un coéquipier n'a pas le même sens que le même toucher réalisé par un adversaire dans un match [21], et au sein d'un couple le toucher prend de nombreuses nuances parfois très fines. Pouvoir prendre en compte le contexte d'interaction est donc essentiel pour pouvoir toucher l'humain de manière à la fois compréhensible, et surtout acceptable. En effet, étant un sens particulièrement intime, tout toucher n'est pas toujours bienvenu. Des études ont montré le niveau général de tolérance au toucher en fonction de la zone du corps et de la relation avec la personne qui nous touche [39], dont il ressort que les femmes ont globalement plus de facilité à toucher et être touché aussi bien par un homme qu'une femme, là où les hommes acceptent peu le toucher d'un autre homme. Cela nous motive à dans un premier temps utiliser un agent de genre féminin pour améliorer l'acceptabilité de ses touchers. Toutefois la question de l'influence du genre sur l'interprétation du toucher est complexe et évolue potentiellement au cours du temps, du fait des évolutions culturelles, il faut donc y faire attention [11,13,23].

En effet, la culture est un élément important qui joue également sur l'interprétation et l'acceptabilité du toucher. Les différentes sociétés ont des rapports au toucher différents [13,35], qu'il convient de prendre au mieux en compte pour une interaction basée sur le toucher social. Toutefois, de la même manière que pour le genre, ceci est à nuancer par la tolérance au toucher personnelle de chaque individu, qui diffère aussi indépendamment de sa culture d'origine et de son genre. Ainsi, le concept de "*touch avoidance*" a été forgé pour représenter l'aversion au toucher des individus [1,2,29]. C'est un autre élément qu'il convient de prendre en compte pour une interaction respectueuse de l'individu à même de créer du lien entre l'agent et l'humain.

Deux questions se posent alors : comment prendre en compte tous ces éléments d'interprétation, et notamment le côté émotionnel de l'interaction ? Et comment représenter et mesurer

le niveau de relation entre l'humain et l'agent, en temps réel, au fil de l'interaction ?

2.2 Les modèles computationnels d'émotion

Comme nous l'avons vu, toucher et émotion semblent très liés, le toucher affectif étant d'ailleurs un champ d'étude à part entière dans la recherche scientifique. Pouvoir prendre en compte l'état émotionnel de celui ou celle avec qui on interagit nous permet de mieux interpréter ses actions et avoir un agent capable d'avoir un état émotionnel le rendrait, dans ce contexte, bien plus crédible aux yeux d'un humain.

A ce titre, il nous semble pertinent de nous intéresser à un modèle qui puisse prendre en compte ce versant émotionnel de l'interaction humain-agent. Les émotions sont toutefois un concept complexe, encore aujourd'hui très discuté, et dont la définition précise et la manière dont elles interagissent avec la cognition restent incertaines [22]. Les théories psychologiques des émotions sont nombreuses, mais s'organisent d'une manière générale en trois grands courants [28] : les théories qui représentent les émotions de façon discrète, les théories dimensionnelles de l'émotion et les *appraisal theories*, qu'on peut traduire comme théories d'évaluation cognitive.

Les principales différences entre ces paradigmes théoriques concernent la façon dont les émotions sont représentées, et donc générées. Dans le cas des théories discrètes, les émotions sont considérées « en bloc », avec des labels comme la colère, la joie, etc. Ces émotions peuvent être plus ou moins complexes, mais sont alors pensées comme s'organisant en combinaisons d'émotions primaires, qui auraient des expressions spécifiques, innées et universellement identifiables [38]. Le reproche généralement fait à ces théories est qu'elles ont tendance à gommer la complexité des émotions et leur déclenchement du point de vue cognitif, et l'universalité de ces expressions primaires des émotions est souvent remise en question. Aujourd'hui encore, il n'est pas clair que ce qu'on appelle quotidiennement la « colère » soit bien un phénomène spécifique, inné, non-composite. Cela reste une notion floue [22], dont beaucoup de sous-composants (l'humeur, le « sentiment », la sensation physiologique, etc.) semblent participer. De leur côté, les théories dimensionnelles de l'émotion considèrent que les émotions peuvent être représentées dans un espace à plusieurs dimensions en fonction de ses propriétés. La représentation la plus connue est sans doute celle du PAD (*pleasure-arousal-dominance*) qui représente les émotions selon le plaisir (aussi appelé valence), l'excitation (*arousal*) et la dominance ressentis dans un environnement ou pour qualifier un événement [31]. Ce modèle fonctionne assez bien avec un certain ensemble d'émotions, mais toutes ne sont pas forcément représentables de manière convaincantes avec ces trois dimensions. D'ailleurs, le choix de ces dimensions est somme toute relativement arbitraire et ne tient pas forcément beaucoup plus compte de l'influence de la culture ou des objectifs particuliers de l'individu et la façon dont

cela entre en ligne de compte dans la génération des émotions. Ce paradigme est globalement aisé à implémenter informatiquement, étant donné la présence de valeurs quantifiables qui peuplent l'espace PAD, et il propose une plus fine représentation de la complexité de chaque émotion que les théories discrètes. C'est ce qui fait que ces modèles de représentation dimensionnelle de l'émotion est très utilisé pour la reconnaissance, là où les théories discrètes alimentent les travaux de simulation (visuelle) de comportements émotionnels. Toutefois, ce sont véritablement les *appraisal theories* qui forment à l'heure actuelle le paradigme principal des théories des émotions. Dans ces théories, l'émotion est générée en fonction de l'évaluation d'un événement extérieur par rapport à la subjectivité de l'individu, donc, intrinsèquement, ses buts, sa culture et sa personnalité. Le nombre des paramètres d'évaluation et leur nature changent d'une théorie à l'autre, mais le principe de l'*appraisal* reste le même et un certain nombre de paramètres sont communs ou similaires (la valence est notamment très souvent présente). Avec les théories d'*appraisal*, on voit bien le lien qu'il peut y avoir entre cognition et émotions, le processus d'évaluation pouvant être déjà considéré comme un processus cognitif.

De même que les théories psychologiques des émotions sont nombreuses, les implémentations de ces théories sous forme de modèles computationnels d'émotions sont également nombreuses, peut-être même plus dans la mesure où plusieurs modèles computationnels peuvent implémenter une même théorie, quoique pas forcément dans son intégralité. Ces modèles computationnels d'émotion semblent tout à fait adaptés à un modèle de décision pour des agents capables d'utiliser le toucher social, dans la mesure où ils permettent de prendre en compte l'émotion aussi bien de l'humain (modèle d'empathie [36]) que de l'agent, en générant ses émotions en fonction des événements. Si les modèles computationnels d'émotion sont déjà nombreux à l'heure actuelle [25,32], la majorité suit les théories d'*appraisal* qui sont les plus intéressantes du point de vue de la génération computationnelle des émotions et de l'interaction, grâce à l'évaluation cognitive, tout en fournissant une base de variables quantifiables facilement manipulables par la machine. L'avantage de ces modèles déjà élaborés est qu'ils ont déjà fourni des résultats convaincants et validés, et il convient d'essayer au maximum de poursuivre ces travaux existants et les étendre aux nouveaux cas d'étude, plutôt que de repartir chaque fois de zéro avec un nouveau modèle [34].

Notre seconde interrogation reste toutefois entière : comment, à partir d'un modèle computationnel d'émotions, représenter et mesurer le niveau de la relation entre un humain et un agent, en temps réel ?

2.3 La notion de *rapport*

Il peut sembler présomptueux de penser qu'un agent peut réellement établir une relation avec un humain dans le cadre d'une interaction de courte durée. La relation étant un concept prenant a priori son sens plutôt sur le long-terme, on peut se demander

quelle autre dimension on peut utiliser pour mesurer le niveau du lien entre l'humain et l'agent.

Un concept fréquemment utilisé dans le domaine des ACA est celui de *rapport*, qu'on peut traduire par "entente" en français. Le *rapport* se constitue même dans des interactions de courte durée et désigne une sensation "d'attention mutuelle, de positivité et de coordination" [41]. Une riche littérature existe quant aux comportements non-verbaux qui interviennent dans la création du rapport, non seulement dans le domaine des interactions sociales entre humains [24,41], mais aussi dans celui des interactions entre humains et agents [6,7,16]. Ces travaux montrent comment le *rapport* peut être construit grâce à du mimétisme dans les mouvements du corps et l'expression d'attention à travers à la fois les gestes et la posture. Mais le *rapport* implique aussi des stratégies verbales, dont le *small talk* est un élément important [3,4]. Les modèles de création de rapport pour ACA ont exploré tous ces éléments, et il apparaît que trois mesures supplémentaires interviennent dans le niveau de rapport : l'appréciation, la solidarité (le niveau auquel les deux individus partagent des opinions similaires) et la familiarité (niveau d'intimité des informations qu'un individu se sent à l'aise de divulguer à un autre individu).

On pourrait se demander pourquoi ne pas juste s'intéresser à l'appréciation, la solidarité et la familiarité, mais le *rapport* offre une mesure plus complète, quantifiable dans le temps, qui rend compte de l'évolution de ces trois sous-mesures essentielles au fil d'une interaction. D'autres composantes du rapport, comme l'attention et la positivité (du domaine de l'affectif) sont également très intéressantes pour prendre la mesure de l'évolution du ressenti mutuel dans l'interaction. Il nous semble donc que le *rapport* est une notion tout à fait pertinente pour évaluer le niveau de la relation entre un humain et un agent, et ainsi permettre à un agent d'évaluer ce niveau en temps réel et se baser dessus pour déterminer si un type de toucher serait acceptable et approprié.

3 Modèle de Décision pour un Agent Touchant

A partir de notre étude de la littérature, nous nous sommes tournés vers la constitution d'un modèle de décision basé sur un modèle computationnel d'émotion, étendu pour prendre en compte le niveau de *rapport* entre l'humain et l'agent. Pour les entrées du modèle, un module de perception faisant office de sens du toucher pour l'agent a été développé dans des travaux préliminaires [5]. Le modèle prendra donc en entrée les différents inputs liés à ce sens du toucher de l'agent ainsi que les actions de l'humain faisant avancer la situation d'interaction (dialogue, gestes, actions sur l'environnement, etc.)

La figure 3 détaille le fonctionnement du modèle que nous avons élaboré. Il étend une architecture de modèle computationnel d'émotion basé sur la théorie d'*appraisal*. Les inputs permettent ainsi à la fois de générer et faire évoluer l'état émotionnel de l'agent selon les règles d'*appraisal* et en particulier les buts et la

personnalité de l'agent ; et à la fois de mettre à jour les différentes connaissances de l'agent sur la situation, le niveau de *rapport*, l'état émotionnel de l'humain, etc. A partir à la fois de sa base de connaissance et de son état émotionnel, des règles de décision vont ensuite permettre à l'agent de choisir le comportement à adopter, et notamment en termes de toucher.

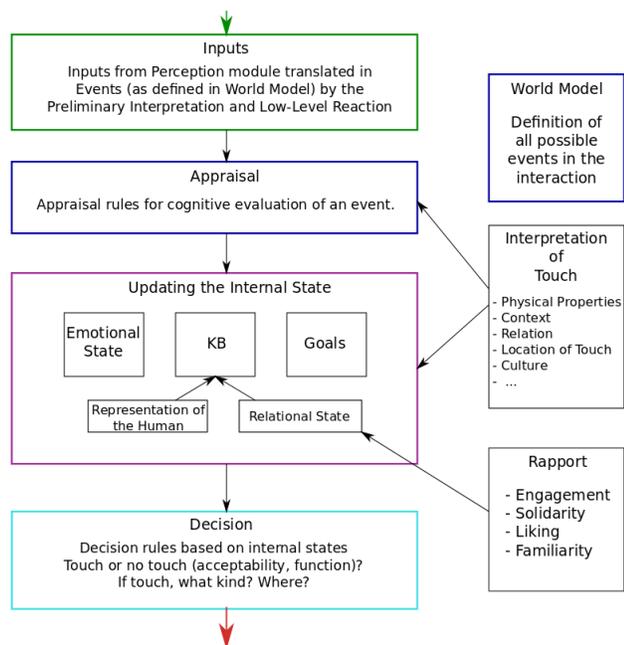


Figure 3: **Architecture générale du modèle de décision**

Plus particulièrement, la figure 4 détaille les dimensions sur lesquelles nous souhaitons baser le processus décisionnel lié au toucher social. Comme vu précédemment, nous modélisons l'évaluation du *rapport* entre l'humain et l'agent. A cela nous ajoutons la notion d'engagement, pour mesurer le niveau d'implication de l'humain dans l'interaction. L'engagement se définit généralement par le niveau d'attention et l'implication émotionnelle d'un individu au cours de l'interaction. Etant lié à l'attention, l'engagement influence donc la mesure du *rapport*, de même que le *rapport* influence la mesure de l'engagement, sur sa composante émotionnelle. Sur le schéma apparaissent également les notions de *static* et *dynamic touch avoidance*, qui font référence à la *touch avoidance* mentionnée dans l'état de l'art, soit la tolérance au toucher de l'humain. La conjonction de la culture et des préférences personnelles de chaque individu forment son niveau de tolérance au toucher : est-il à l'aise avec le fait d'être touché ? Par qui ? Nous considérons que cette tolérance peut se différencier en une composante statique, la baseline a priori qu'on peut évaluer à travers des questionnaires spécifiques avant le début de l'interaction, et une composante dynamique, liée à l'évolution de la relation entre l'individu et son interlocuteur. Le niveau de tolérance au toucher de base n'est pas modifiée par cette évolution, mais peut être contrebalancée par la mesure dynamique, qui est alors évaluée à travers l'engagement et le

rapport. L'agent se voit donc premièrement muni dans ses connaissances d'une estimation *a priori* de la tolérance de base au toucher de l'humain, qui lui aura été fournie en amont de l'interaction (*static touch avoidance*) grâce à l'utilisation d'un questionnaire préliminaire dédié à l'estimation de cette mesure. C'est ensuite au cours de l'interaction qu'il évaluera une tolérance dynamique en fonction de l'évolution de ses mesures du *rapport* et de l'engagement entre lui et l'humain. Ces deux éléments seront pris en compte pour la phase de décision.

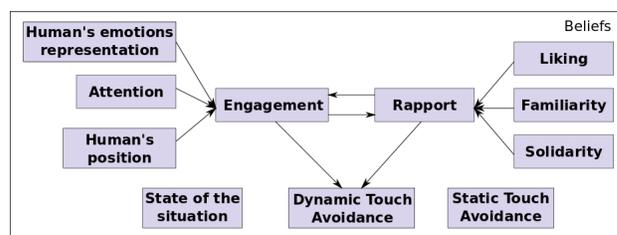


Figure 4: **Détail des dimensions prises en compte dans les connaissances de l'agent**

Ces différentes mesures et leur estimation représentent en elles-mêmes des défis scientifiques qui ne sont pas encore résolus dans l'ensemble, même en dehors du contexte du toucher, défis qui ne sont pas l'objet de notre travail. Une partie essentielle du travail à venir sera donc, à partir de modèles validés comme celui du *rapport* de Zhao et. al. [42] sur lequel nous nous basons, de déterminer dans quelle mesure le toucher peut influencer ces différentes dimensions de l'interaction, et inversement dans quelle mesure chacune de ces dimensions peut être prise en compte pour moduler les règles de décision du toucher (équations numériques pour estimer *rapport* et engagement dans notre contexte). A titre d'exemple, on ne sait pas encore de quelle manière la notion de solidarité ou celle de familiarité peuvent être corrélées avec des événements de toucher de l'humain vers l'agent ou inversement. Ce sera l'objet de nos travaux à venir que de raffiner le modèle au niveau des interactions entre les différentes dimensions prises en compte. Une première tentative de définition des influences entre la base de connaissance de l'agent et ses décisions est présentée en figure 5.

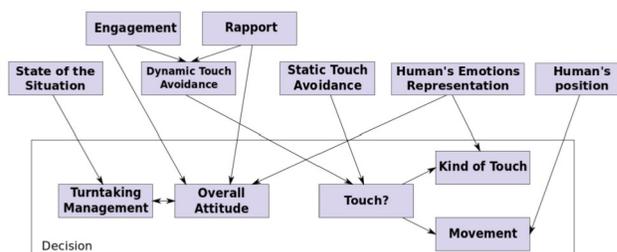


Figure 5: **Influence des différents éléments de la base de connaissances de l'agent sur sa décision**

4 Prochaines étapes

4.1 Implémentation

Le modèle ainsi envisagé repose sur une base architecturale de modèle computationnel d'émotion, qui est ensuite étendue pour prendre en compte les spécificités du toucher social. La prochaine étape est de passer à l'implémentation de ce nouveau modèle. Sur la base de cette architecture, il reste de nombreux choix possibles quant à la nature précise de la décision (réseaux bayesiens, modèle par apprentissage, etc.) Comme nous souhaitons contribuer à l'avancement des travaux sur les modèles computationnels d'émotion, une revue de l'état de l'art des modèles existants nous a conduit à considérer l'utilisation de FAtiMA [12], qui est un outil à la fois open-source, modulaire donc facilement modifiable selon nos besoins, et déjà validé et utilisé dans de nombreux travaux de recherche.

FAtiMA est un modèle computationnel d'émotions fondé sur l'évaluation cognitive (*appraisal*) et qui utilise des algorithmes à base de règles d'inférences sur une base de connaissances pour fonctionner. Ce genre de modèle est naturellement restreint sur le versant de la flexibilité et la capacité d'adaptation face à des situations non-prévues dans les règles. Nous pourrions donc envisager de développer un modèle qui serait plutôt basé sur une décision via *machine learning*, plus souple et robuste en cas de situations hors du commun. Toutefois, ce genre d'algorithmes sont moins intelligibles et il n'est pas toujours évident de déterminer sur quels critères une décision a été prise. Dans ce contexte, nous pensons qu'il est pour le moment plus intéressant de rester sur un modèle à base de règles, pour lequel nous pourrions déterminer des règles de toucher à partir de la littérature du toucher social. Cela constituera ainsi une contribution d'ordre également théorique pour la recherche sur le toucher social. Des travaux ultérieurs pourront se concentrer sur la constitution d'un modèle alternatif mettant l'accent sur l'apprentissage.

Notre travail a ainsi à la fois vocation à proposer un système original permettant d'interagir par le toucher au sein d'une boucle interactive complète avec un ACA, et une base de règles de décision quant à l'utilisation du toucher social, qui pourront servir pour des modèles de décision ultérieurs ainsi que pour comparaison dans des études plus générales sur le toucher social.

4.2 Scénario d'interaction

Une question essentielle à ce genre d'étude d'interactions est : dans quelle situation d'interaction le modèle sera-t-il mis à l'épreuve ? Etant donné que les modèles de décision à base de règles sont limités dans leur flexibilité vis-à-vis des situations non-définies à l'avance par les règles de décision, et étant également donnée la complexité de définir un corpus de règles exhaustif pour la variété des interactions basées sur le toucher social possible, il convient de définir une situation d'interaction précise et bien limitée sur laquelle définir nos règles. Il s'agit donc de déterminer précisément le scénario d'interaction, qui nous permettra de mettre le système à l'épreuve.

Comme nous cherchons essentiellement à mesurer l'impact du modèle et du toucher sur la création de relation sociale entre un humain et un agent, ainsi que la communication des émotions, la situation de mise en oeuvre du modèle doit permettre une interaction émotionnelle intéressante. Dans cette optique, nous envisageons un scénario de narration, dans lequel l'agent conversationnel animé raconte un souvenir émotionnel à l'humain, tandis que l'humain l'aide à poursuivre son récit en interagissant avec lui, et notamment à l'aide du toucher.

Un important défi avec un tel type de scénario, par essence très social et peu orienté sur une tâche à accomplir, est de déterminer comment le toucher va pouvoir naturellement prendre place. Cela requiert à la fois un travail précis de scénarisation du déroulé de l'interaction, une utilisation à bon escient de tout l'environnement et des capacités de l'agent, et une prise en compte du scénario dans la mise au point des règles de décision. Il reste ensuite à déterminer quelles sont les mesures objectives qui peuvent valider les performances du modèle.

4.3 Evaluation du modèle

La littérature du toucher social et de ses technologies n'a pas encore produit de *guidelines* claires ou d'étude précise concernant des questions pourtant essentielles de l'ordre de : quand tel type de toucher est-il utilisé ? Dans quel objectif ? Pour exprimer quelle émotion ? Pour combien de temps ? De même sur le versant de l'acceptabilité, il n'est pas encore clair dans quelle mesure un toucher venant d'un agent virtuel est acceptable, ni dans quelle mesure des humains sont susceptibles de trouver de l'intérêt à communiquer socialement par le toucher avec un ACA virtuel.

C'est donc sur ce premier point que l'on peut formuler un critère d'évaluation de notre modèle, et en particulier de notre base de règles. Quel est le niveau de crédibilité du comportement tactile de l'agent ? Les touchers qu'il entreprend sont-ils considérés comme acceptables par l'humain ?

Sur le versant émotionnel et relationnel, le *rapport* étant au cœur de notre modèle, le plus pertinent serait de mesurer la perception du *rapport* par l'humain et de la confronter à l'évaluation en temps réel faite par le modèle de l'agent, afin de confirmer une éventuelle coïncidence des deux mesures. De même, l'intérêt du toucher dans l'interaction pourra être mesuré en interrogeant le participant humain sur son ressenti au cours de l'interaction, en particulier sur les dimensions de présence sociale, co-présence, et appréciation de l'expérience et de l'agent. Ces éléments pourront être mesurés à travers l'utilisation de questionnaires validés sur ces concepts déjà largement discutés dans la littérature des ACA et de la réalité virtuelle, ainsi que par des relevés objectifs de direction du regard et déplacements de l'humain dans l'environnement virtuel, rythme cardiaque, etc.

5 Conclusion

Mettre en place une boucle interactive complète pour des interactions humain-agent incluant le toucher social requiert à la fois des capacités de perception, de décision et de réalisation. Nous avons ici présenté un modèle de décision basé sur une architecture de modèle computationnel d'émotions qui puisse prendre en compte le toucher social et en particulier son versant émotionnel, pour permettre à un agent d'interpréter des touchers, d'évaluer le niveau de *rappport* entre lui et l'humain, et de choisir, en conséquence, le toucher à réaliser ou pas sur l'humain pour poursuivre l'interaction.

Le modèle présenté reste à implémenter et évaluer, mais nous pensons qu'il permettra de faire avancer les recherches sur le toucher social simulé en contribuant d'une part à l'élaboration d'un corpus de règles de toucher affectif, et d'autre part à la mise en place technique d'une interaction dans les deux sens (l'humain touche l'agent, l'agent touche l'humain). Cette implémentation constituera donc un premier jalon dans la mise en place d'une boucle interactive complète incluant les interactions de toucher, et des développements incrémentaux permettront de peu à peu améliorer la crédibilité et la performance de ce modèle, ou d'ouvrir de nouvelles perspectives de développement (*machine learning*, autre contexte d'interaction, robotique, etc.)

REMERCIEMENTS

Ce travail a pu être réalisé grâce au financement de l'ANR dans le contexte du projet Social Touch (ANR-17-CE33-0006). Nous remercions également le FEDER et la région Hauts-de-France pour le financement de la salle immersive TRANSLIFE du laboratoire UMR CNRS Heudiasyc.

REFERENCES

- [1] Peter A. Andersen and Kenneth Leibowitz. 1978. The development and nature of the construct touch avoidance. *J Nonverbal Behav* 3, 2 (December 1978), 89–106. DOI:<https://doi.org/10.1007/BF01135607>
- [2] T. W. Bickmore, R. Fernando, L. Ring, and D. Schulman. 2010. Empathic Touch by Relational Agents. *IEEE Transactions on Affective Computing* 1, 1 (January 2010), 60–71. DOI:<https://doi.org/10.1109/T-AFFC.2010.4>
- [3] Timothy Bickmore and Justine Cassell. 2001. Relational Agents: A Model and Implementation of Building User Trust. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '01)*, ACM, New York, NY, USA, 396–403. DOI:<https://doi.org/10.1145/365024.365304>
- [4] Timothy W. Bickmore and Justine Cassell. 1999. Small Talk and Conversational Storytelling In Embodied Conversational Interface Agents.
- [5] Fabien Boucaud, Quentin Tafiani, Catherine Pelachaud, and Indira Thouvenin. 2019. Social Touch in Human-agent Interactions in an Immersive Virtual Environment: In *Proceedings of the 14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, SCITEPRESS - Science and Technology Publications*, Prague, Czech Republic, 129–136. DOI:<https://doi.org/10.5220/0007397001290136>
- [6] Justine Cassell, Alastair J. Gill, and Paul A. Tepper. 2007. Coordination in conversation and rapport. In *Proceedings of the Workshop on Embodied Language Processing - EmbodiedNLP '07*, Association for Computational

- Linguistics, Prague, Czech Republic, 41–50. DOI:<https://doi.org/10.3115/1610065.1610071>
- [7] Justine Cassell and Kristinn R. Thorisson. 1999. The power of a nod and a glance: Envelope vs. emotional feedback in animated conversational agents. *Applied Artificial Intelligence* 13, 4–5 (May 1999), 519–538. DOI:<https://doi.org/10.1080/088395199117360>
- [8] Jonas Chatel-Goldman, Marco Congedo, Christian Jutten, and Jean-Luc Schwartz. 2014. Touch increases autonomic coupling between romantic partners. *Front. Behav. Neurosci.* 8, (2014). DOI:<https://doi.org/10.3389/fnbeh.2014.00095>
- [9] Carolina Cruz-Neira, Daniel J. Sandin, and Thomas A. DeFanti. 1993. Surround-screen Projection-based Virtual Reality: The Design and Implementation of the CAVE. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '93)*, ACM, New York, NY, USA, 135–142. DOI:<https://doi.org/10.1145/166117.166134>
- [10] Anik Debrot, Dominik Schoebi, Meinrad Perrez, and Andrea B. Horn. 2013. Touch as an Interpersonal Emotion Regulation Process in Couples' Daily Lives: The Mediating Role of Psychological Intimacy. *Pers Soc Psychol Bull* 39, 10 (October 2013), 1373–1385. DOI:<https://doi.org/10.1177/0146167213497592>
- [11] Valerian J. Derlega, Robin J. Lewis, Scott Harrison, Barbara A. Winstead, and Robert Costanza. 1989. Gender differences in the initiation and attribution of tactile intimacy. *J Nonverbal Behav* 13, 2 (June 1989), 83–96. DOI:<https://doi.org/10.1007/BF00990792>
- [12] João Dias, Samuel Mascarenhas, and Ana Paiva. 2014. FAtiMA Modular: Towards an Agent Architecture with a Generic Appraisal Framework. In *Emotion Modeling: Towards Pragmatic Computational Models of Affective Processes*, Tibor Bosse, Joost Broekens, João Dias and Janneke van der Zwaan (eds.). Springer International Publishing, Cham, 44–56. DOI:https://doi.org/10.1007/978-3-319-12973-0_3
- [13] Rosemarie Dibiase and Jaime Gunnoe. 2004. Gender and Culture Differences in Touching Behavior. *The Journal of Social Psychology* 144, 1 (February 2004), 49–62. DOI:<https://doi.org/10.3200/SOCP.144.1.49-62>
- [14] Tiffany Field. 2003. *Touch*. MIT, Cambridge, Mass.; London.
- [15] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. A survey of socially interactive robots. *Robotics and Autonomous Systems* 42, 3 (March 2003), 143–166. DOI:[https://doi.org/10.1016/S0921-8890\(02\)00372-X](https://doi.org/10.1016/S0921-8890(02)00372-X)
- [16] Jonathan Gratch, Sin-Hwa Kang, and Ning Wang. 2013. Using Social Agents to Explore Theories of Rapport and Emotional Resonance. In *Social Emotions in Nature and Artifact*, Jonathan Gratch and Stacy Marsella (eds.). Oxford University Press, 181–197. DOI:<https://doi.org/10.1093/acprof:oso/9780195387643.003.0012>
- [17] Jonathan Gratch, Ning Wang, Anna Okhmatovskaia, Francois Lamothe, Mathieu Morales, R. J. van der Werf, and Louis-Philippe Morency. 2007. Can Virtual Humans Be More Engaging Than Real Ones? In *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments (Lecture Notes in Computer Science)*, Springer Berlin Heidelberg, 286–297.
- [18] Matthew J. Hertenstein, Rachel Holmes, Margaret McCullough, and Dacher Keltner. 2009. The communication of emotion via touch. *Emotion* 9, 4 (2009), 566–573. DOI:<https://doi.org/10.1037/a0016108>
- [19] Matthew J. Hertenstein, Dacher Keltner, Betsy App, Brittany A. Bulleit, and Ariane R. Jaskolka. 2006. Touch communicates distinct emotions. *Emotion* 6, 3 (2006), 528–533. DOI:<https://doi.org/10.1037/1528-3542.6.3.528>

- [20] Matthew J. Hertenstein, Julie M. Verkamp, Alyssa M. Kerestes, and Rachel M. Holmes. 2006. The communicative functions of touch in humans, nonhuman primates, and rats: a review and synthesis of the empirical research. *Genetic, social, and general psychology monographs* 132, 1 (2006), 5–94. DOI:https://doi.org/10.3200/MONO.132.1.5-94
- [21] Gijs Huisman, Jan Kolkmeier, and Dirk Heylen. 2014. With Us or Against Us: Simulated Social Touch by Virtual Agents in a Cooperative or Competitive Setting. In *Intelligent Virtual Agents* (Lecture Notes in Computer Science), Springer International Publishing, 204–213.
- [22] Carroll E. Izard. 2010. The Many Meanings/Aspects of Emotion: Definitions, Functions, Activation, and Regulation. *Emotion Review* 2, 4 (October 2010), 363–370. DOI:https://doi.org/10.1177/1754073910374661
- [23] Stanley E. Jones. 1986. Sex differences in touch communication. *Western Journal of Speech Communication* 50, 3 (December 1986), 227–241. DOI:https://doi.org/10.1080/10570318609374230
- [24] Marianne LaFrance. 1979. Nonverbal Synchrony and Rapport: Analysis by the Cross-Lag Panel Technique. *Social Psychology Quarterly* 42, 1 (March 1979), 66. DOI:https://doi.org/10.2307/3033875
- [25] Jerry Lin, Marc Spraragen, and Michael Zyda. 2012. Computational Models of Emotion and Cognition. In *Advances in Cognitive Systems*, 59–76.
- [26] Gale M. Lucas, Jonathan Gratch, Aisha King, and Louis-Philippe Morency. 2014. It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior* 37, (August 2014), 94–100. DOI:https://doi.org/10.1016/j.chb.2014.04.043
- [27] Gale M. Lucas, Albert Rizzo, Jonathan Gratch, Stefan Scherer, Giota Stratou, Jill Boberg, and Louis-Philippe Morency. 2017. Reporting Mental Health Symptoms: Breaking Down Barriers to Care with Virtual Human Interviewers. *Front. Robot. AI* 4, (2017). DOI:https://doi.org/10.3389/frobt.2017.00051
- [28] Stacy Marsella, Jonathan Gratch, and P Petta. 2010. Computational models of emotion. *A Blueprint for Affective Computing-A Sourcebook and Manual* (January 2010), 21–46.
- [29] Matthew M. Martin and Carolyn M. Anderson. 1993. Psychological and biological differences in touch avoidance. *Communication Research Reports* 10, 2 (December 1993), 141–147. DOI:https://doi.org/10.1080/08824099309359927
- [30] Sarah McIntyre, Athanasia Mounou, Rebecca Boehme, Peder Mortvedt Isager, Frances Lau, Ali Israr, Ellen A. Lumpkin, Freddy Abnoui, and Haakan Olausson. 2019. Affective touch communication in close adult relationships. DOI:https://doi.org/10.17605/OSF.IO/7XRWC
- [31] Albert Mehrabian and James A. Russell. 1974. *An approach to environmental psychology*. MIT Press, Cambridge, Mass.
- [32] Desmond C. Ong, Jamil Zaki, and Noah D. Goodman. 2019. Computational Models of Emotion Inference in Theory of Mind: A Review and Roadmap. *Topics in Cognitive Science* 11, 2 (2019), 338–357. DOI:https://doi.org/10.1111/tops.12371
- [33] Sandra Petersen, Susan Houston, Huanying Qin, Corey Tague, and Jill Studley. 2016. The Utilization of Robotic Pets in Dementia Care. *JAD* 55, 2 (November 2016), 569–574. DOI:https://doi.org/10.3233/JAD-160703
- [34] Rainer Reisenzein, Eva Hudlicka, Mehdi Dastani, Jonathan Gratch, Koen Hindriks, Emiliano Lorini, and John-Jules Meyer. 2013. Computational Modeling of Emotion: Towards Improving the Inter- and Intradisciplinary Exchange. *IEEE Transactions on Affective Computing* 4, n° 3 (July 2013), 246–266. DOI:https://doi.org/10.1109/T-AFFC.2013.14
- [35] Martin S. Remland, Tricia S. Jones, and Heidi Brinkman. 1995. Interpersonal Distance, Body Orientation, and Touch: Effects of Culture, Gender, and Age. *The Journal of Social Psychology* 135, 3 (June 1995), 281–297. DOI:https://doi.org/10.1080/00224545.1995.9713958
- [36] Sérgio Hortas Rodrigues, Samuel Mascarenhas, João Dias, and Ana Paiva. 2015. A Process Model of Empathy For Virtual Agents. *Interact Comput* 27, 4 (July 2015), 371–391. DOI:https://doi.org/10.1093/iwc/iwu001
- [37] Fiorella de Rosis, Catherine Pelachaud, Isabella Poggi, Valeria Carofiglio, and Berardina De Carolis. 2003. From Greta's mind to her face: modelling the dynamics of affective states in a conversational embodied agent. *International Journal of Human-Computer Studies* 59, 1 (July 2003), 81–118. DOI:https://doi.org/10.1016/S1071-5819(03)00020-X
- [38] Klaus R. Scherer. 2000. Psychological models of emotion. In *The neuropsychology of emotion*. Oxford University Press, New York, NY, US, 137–162.
- [39] Juulia T. Suvilehto, Enrico Glerean, Robin I. M. Dunbar, Riitta Hari, and Lauri Nummenmaa. 2015. Topography of social touching depends on emotional bonds between humans. *Proc Natl Acad Sci USA* 112, 45 (November 2015), 13811–13816. DOI:https://doi.org/10.1073/pnas.1519231112
- [40] Erin H. Thompson and James A. Hampton. 2011. The effect of relationship status on communicating emotions through touch. *Cognition and Emotion* 25, 2 (February 2011), 295–306. DOI:https://doi.org/10.1080/02699931.2010.492957
- [41] Linda Tickle-Degnen and Robert Rosenthal. 1990. The Nature of Rapport and Its Nonverbal Correlates. *Psychological Inquiry* 1, 4 (October 1990), 285–293. DOI:https://doi.org/10.1207/s15327965pli0104_1
- [42] Zhao R., Papangelis A., Cassell J. 2014. Towards a Dyadic Computational Model of Rapport Management for Human-Virtual Agent Interaction. In: Bickmore T., Marsella S., Sidner C. (eds) *Intelligent Virtual Agents*. IVA 2014. Lecture Notes in Computer Science, vol 8637. Springer, Cham