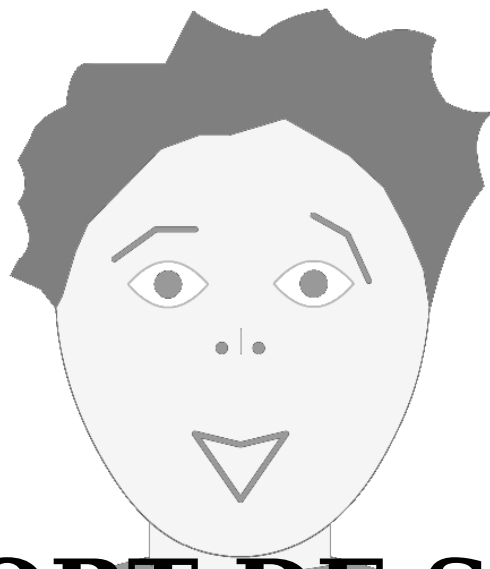




VALORIA



RAPPORT DE STAGE

01

Expression et dynamique des émotions
Application sur un avatar virtuel

Céline Jost
Master Recherche Informatique
Systèmes intelligents auto-organisés
Application à l'animation par ordinateur, la robotique et au Web Mining

Encadrant : Dominique Duhaut

Février - Juin 2009

Résumé

L'expression, la reconnaissance et la manipulation d'émotions sont les clefs d'une bonne interaction émotionnelle entre un robot et un patient dans le cadre de la thérapie assistée par robot. En effet, le robot devient un compagnon affectif qui se doit d'être émotionnellement intelligent, car pour apporter du réconfort, il faut pouvoir analyser et comprendre l'état émotionnel du patient. La finalité étant de pouvoir réagir à cet état émotionnel pour apporter ce dont le patient a besoin.

Le projet EmotiRob traite cette thématique en fabriquant un robot-peluche capable d'apporter du réconfort aux personnes fragilisées. Le robot analyse l'état émotionnel de l'enfant à travers son discours et réagit en fonction de son propre état. On s'intéresse ici à l'expression des émotions car si l'enfant n'est pas capable de reconnaître l'émotion du robot alors il ne peut pas être influencé par l'interaction émotionnelle.

L'objectif de ce travail est donc de créer un avatar virtuel coordonné avec le robot en ce qui concerne son état émotionnel. Cet avatar est la conscience du robot et il peut être transporté avec soit, bien plus facilement que le robot. La cible de l'avatar est en effet des dispositifs mobiles. En partant du travail effectué sur l'expression statique des émotions, on va dans un premier temps le reporter sur l'avatar et dans un deuxième temps compléter l'étude par une expression dynamique. L'avatar doit avoir l'air « vivant ». Pour finir le dernier objectif de ce travail est d'apporter un caractère et une personnalité à l'avatar.

Le travail effectué sur l'avatar sera reporté sur le robot afin d'augmenter son impact auprès des personnes fragilisées.

Mots clés : émotions, interaction intelligence, avatar virtuel, expression statique, expression dynamique, simulation.

Abstract

We know that a good emotional interaction between a robot and a patient is defined by capabilities to express emotion, capabilities to identify and to control emotion. This emotional interaction is used in the RAT - Robot Assisted Therapy - which allows to give comfort to weak people. Indeed, nowadays robots become real companions for people, they have a real impact on people, that is why they have to be emotionally intelligent.

The only thing to understand is the importance of this emotional intelligence, because in order to give comfort, in order to help people, in order to control others' emotions, robot has to be able to analyze someone and to understand his emotional state. The purpose of this interaction is to bring patients exactly what they need.

The EmotiRob project deals with RAT. The purpose of this project is to create a n automated teddy bear being able to give comfort to weak people. This robot analyzes people emotional state from his talk and reacts taking account its own emotional state.

This work takes interest in emotion expression because it is important that people recognize the robot emotion. Otherwise it implies a bad emotional interaction and people could not be influenced.

The purpose of the current work is to create a virtual avatar which is able to express emotion. This avatar has to be coordinate with the robot in order they both are in the same emotional state at the same instant. This avatar represents the conscience of the teddy bear. The advantage is that we can easily transport the avatar, actually its target is mobile.

This work is based on the work about statical facial expression of the robot. First, we report that on the avatar and secondly we complete it with dynamical facial expression. We want the avatar seems to be alive. The third part of this work is to add a temperament and a personality to the avatar.

The work on dynamical expression will be report on robot in order to increase its impact on weak people.

Key words : emotions, intelligent interaction, virtual avatar, statical expression, dynamical expression, simulation.

Remerciements

Je tiens à remercier mon encadrant, **Dominique Duhaut**, qui m'a aidé tout au long de ce stage, qui a eu la patience de me comprendre, la pédagogie de me faire avancer et qui m'a offert des conditions de travail agréables.

Je souhaite remercier aussi **Brigitte Le Pévédic** qui m'a également encadré sur ce stage.

Je tiens également à remercier **Sébastien Saint-Aimé** qui m'a intégré au sein de ce projet, qui a toujours pris le temps de répondre à mes requêtes et de travailler avec moi.

Je remercie également **Thibaut Le Naour**, qui partage le bureau dans lequel je travaille, pour avoir eu la patience de m'écouter et pour n'avoir pas bronché de jouer mon cobaye quelques fois.

Je souhaite aussi remercier **toute l'équipe de travail** de Vannes et de Lorient, qui m'a accueilli chaleureusement et qui m'a intégré directement.

J'adresse mes sincères remerciements à **tous les enseignants** qui m'ont permis d'arriver jusque là.

Je tiens particulièrement à rendre hommage à **mes parents**, qui ont toujours été là pour moi et m'ont toujours soutenu.

Et pour finir, je tiens à remercier **mon entourage** qui m'a supporté et épaulé au mieux.

Table des matières

1. Introduction.....	6
1.1. Problème scientifique général.....	6
1.2. Problématique d'Emotirob.....	6
1.3. Plan du rapport.....	7
2. État de l'art	8
2.1. Introduction.....	8
2.2. Les émotions.....	9
2.2.1. Que sont les émotions ?.....	9
2.2.2. Intérêt des émotions dans les interactions homme-machine.....	12
2.2.3. Expression des émotions.....	13
2.2.4. Mesure des émotions sur autrui.....	14
2.3. Quelques implémentations.....	15
2.3.1. Simulation et détection d'émotions.....	15
2.3.2. Quelques robot émotifs.....	18
2.4. Conclusion.....	19
3. Expression - Mise en œuvre d'Art-e	21
3.1. Sourcils + bouche (échange avec le robot).....	21
3.1.1. Les sourcils.....	21
3.1.2. La bouche.....	22
3.2. Les yeux	23
3.3. La tête et le tronc.....	25
3.3.1. La tête.....	25
3.3.2. Le tronc.....	25
3.4. Ensemble d'expressions faciales d'Art-e.....	26
3.4.1. La joie.....	27
3.4.2. La colère.....	27
3.4.3. La tristesse.....	28
3.4.4. La surprise.....	28
3.4.5. Le dégoût.....	29
3.4.6. La peur.....	30
3.4.7. Expression neutre	30
4. Dynamique des émotions.....	31
4.1. Outils de base à la dynamique.....	31
4.1.1. La respiration.....	32
4.1.2. Le clignement des yeux.....	33
4.1.3. Le mouvement du regard.....	34
4.1.4. Les rictus.....	36
4.2. Passage d'une émotion à une autre.....	36
4.2.1. Types d'affichages des émotions.....	36
4.2.2. Réaction émotionnelle des composants.....	39
4.2.3. Étapes d'une réaction émotionnelle d'Art-e.....	41
4.3. Commandes de contrôle.....	41
5. Particularisation d'Art-e.....	41
5.1. Caractère.....	42
5.2. Personnalité.....	42
6. Évaluation et premiers tests	44
6.1. Évaluation.....	44
6.2. Premier test.....	45
6.3. Deuxième test.....	45
7. Conclusion.....	46
7.1. Travaux effectués.....	46
7.2. Travaux restants et futurs.....	47
8. Références.....	48

Index des illustrations

Illustration 1: Synoptique du projet Emotirob.....	8
Illustration 2: Scherer - Organisation de l'émotion.....	10
Illustration 3: Modèle d'émotion de D. Duhaut.....	12
Illustration 4: Le robot Paro de T. Shibata.....	12
Illustration 5: Modèle de Plutchik.....	14
Illustration 6: Expressions faciale de iCat de Philips.....	18
Illustration 7: Kismet.....	18
Illustration 8: Paramètres caractérisant le sourcil.....	21
Illustration 9: Paramètres caractérisant la bouche.....	22
Illustration 10: Paramètres caractérisant les yeux.....	23
Illustration 11: Mécanisme permettant de fermer l'œil.....	24
Illustration 12: Paramètres caractérisant la tête.....	25
Illustration 13: Les 2 plans de coupe de l'être humain.....	26
Illustration 14: Expressions faciales représentant la joie.....	27
Illustration 15: Expressions faciales représentant la colère.....	28
Illustration 16: Expressions faciales représentant la tristesse.....	28
Illustration 17: Expressions faciales représentant la surprise.....	29
Illustration 18: Expressions faciales représentant le dégoût.....	29
Illustration 19: Expressions faciales représentant la peur.....	30
Illustration 20: Expressions faciales neutres.....	30
Illustration 21: Diagramme d'état d'Arte.....	31
Illustration 22: Fonction de simulation de respiration.....	32
Illustration 23: Exemple de respiration pour $a=1$ avec un déphasage de $-PI/2$	33
Illustration 24: Diagramme d'état du clignement d'œil.....	34
Illustration 25: Surface de la cornée graphique.....	35
Illustration 26: Exemple de trajectoire possible avec une courbe de Bézier de degré 3.....	35
Illustration 27: Simulation de l'hésitation.....	37
Illustration 28: Les 8 "préférences" du MBTI.....	43
Illustration 29: Trois instances d'Arte en waiting mode.....	44
Illustration 30: Trois instance d'Arte en listening mode (+emotion mode).....	45

Index des tables

Texte 1: Calcul de la durée d'une phase du clignement de l'œil.....	33
Texte 2: Interpolation exponentielle de l'image n°step sur nbImage.....	37
Texte 3: Interpolation logarithmique de l'image n°step sur nbImage.....	37
Texte 4: Interpolation break linear de l'image n°step sur nbImage.....	38
Texte 5: Interpolation break exponentielle de l'image n°step sur nbImage.....	38
Texte 6: Interpolation break logarithmique de l'image n°step sur nbImage.....	39
Texte 7: Changement d'oméga et de a au cours du temps.....	41

1. Introduction

1.1. Problème scientifique général

Les robots ont souvent été utilisés pour construire des produits, pour déplacer des objets ou pour transformer spécifiquement un objet. Mais actuellement, on commence à assimiler le robot à un compagnon, un partenaire. Plus précisément, le domaine du RAT (« Robot Assisted Therapy ») a prouvé que les robots compagnons pouvaient donner un certain confort moral et psychologique aux personnes fragiles.

Mais pour que le robot compagnon réussisse à apporter ce confort, il est nécessaire qu'il y ait une bonne interaction entre lui et la personne fragilisée. Le plus difficile est de savoir ce qu'est une bonne interaction. Rosaling W. Picard évoque la notion d'intelligence émotionnelle : « *l'intelligence émotionnelle consiste à la capacité de reconnaître, d'exprimer et d'avoir des émotions, associée à la capacité de réguler ses émotions, de les exploiter pour un but constructif et de manipuler habilement les émotions des autres.* »

Suivant cette définition un être intelligent émotionnellement connaît les émotions, les utilise, les reconnaît, les manipule. Un être intelligent est capable de jouer avec les émotions de son interlocuteur.

Le problème résulte dans le fait qu'on ne sait pas exactement ce que sont les émotions. Et même si on est capable de les exprimer, on est moins capable de les reconnaître. Dans le contexte de la RAT, nous sommes face à deux difficultés majeures : la reconnaissance de l'état émotif de l'interlocuteur (la personne fragilisée) et l'expression d'une émotion par le robot. C'est ce deuxième point que nous allons étudier dans ce stage.

1.2. Problématique d'Emotirob

Le projet Emotirob a pour objectif la conception d'un robot-peluche capable d'apporter du réconfort à des enfants fragilisés. Le projet se compose de trois parties principales et interdépendantes à savoir

- ◆ la reconnaissance et la compréhension de la parole de l'enfant,
- ◆ l'interaction émotionnelle entre l'enfant et le robot,
- ◆ l'interaction cognitive entre l'enfant et le robot.

Le sujet de mon stage porte sur le deuxième point. On va s'intéresser au réconfort psychologique que l'on peut apporter à un enfant fragilisé. Or le réconfort psychologique est lié à la qualité du lien affectif que l'enfant peut avoir avec le robot. Comme ce lien dépend de l'émotion, de l'humeur et des sentiments, on doit s'intéresser à ces trois notions. On commencera par ce travail à étudier l'expression et la dynamique des émotions.

Pour étudier la dynamique des émotions on peut procéder de deux façons :

- ◆ utilisation de la Motion Capture : la motion capture consiste à récupérer les mouvements d'une personne par capteur. Cette solution n'est pas retenue car ce travail est déjà effectué au sein du laboratoire VALORIA. De plus, la motion capture nécessite forcément qu'on reporte les données sur un avatar virtuel dont la morphologie est la même que la personne qui a enregistré les mouvements et c'est trop contraignant.
- ◆ Construction directe consistant à faire soi même l'avatar virtuel, à fabriquer l'expressivité et la dynamique en utilisant des outils mathématiques. Cette solution est retenue, elle nous permettra d'augmenter les marges de manœuvre.

D'autres raisons motivent la construction de notre propre avatar :

- ◆ Nous voulons que l'avatar soit la conscience du robot, un compagnon virtuel pour l'enfant. Le robot et l'avatar doivent pouvoir communiquer et se coordonner émotivement.
- ◆ De ce fait, l'avatar doit pouvoir suivre l'enfant partout. Nous pensons donc à le déployer sur un appareil mobile.
- ◆ Sur l'appareil mobile que nous utilisons pour le stage, il n'existe aucun avatar ouvert sur la programmation. Cela explique le besoin de le fabriquer complètement.

1.3. Plan du rapport

Dans un premier temps, ce rapport présente l'état de l'art de « la robotique et l'intelligence émotionnelle », c'est une bibliographie qui présente les émotions de façons générales. L'intérêt de cette généralité a été de vérifier de quelle manière la recherche avait évolué afin de déterminer comment se place le projet Emotirob par rapport à la recherche.

Dans un deuxième temps nous présenterons Art-e, l'avatar virtuel réalisés au sein de ce stage et capable d'exprimer des émotions et capable de coopérer avec Emi, le robot compagnon. L'avatar possède au niveau de sa bouche et de ses sourcils les mêmes degrés de liberté que le robot, afin de faciliter leur coordination. Il possède des degrés de liberté supplémentaires qui lui permettent d'exprimer des émotions à sa manière. Et il se voit pourvu d'yeux et d'un tronc qui permettent de jouer sur d'avantage de paramètres pour exprimer des émotions.

Nous verrons que Art-e est capable de cligner des yeux, de respirer, de regarder les mouches voler et d'avoir des rictus. C'est une proposition de dynamique des émotions qui sera décrite dans la suite du rapport.

De plus, un des objectifs de ce travail est de pouvoir différencier plusieurs instance d'Art-e, pour cette raison il est doté d'un caractère et d'une personnalité, qui permet à chaque instance de ressentir et d'exprimer les émotions de façon différente.

Dans la dernière partie du rapport, on parlera de l'évaluation de l'avatar et des résultats actuels. Puis la conclusion rappellera le travail effectué et le travail restant à faire.

2. État de l'art

2.1. Introduction

Dans notre société, les équipements informatiques et robotiques occupent une place prépondérante. Ils participent particulièrement à simplifier la vie des êtres humains. Tout a commencé avec les premières machines automatisées dans les entreprises et maintenant presque tout est automatisé, les tableaux de bords des avions ou de tous les véhicules motorisés, la domotique etc. Mais le besoin humain est plus complexe. Nous avons pu nous en rendre compte à travers deux films qui ont montré l'intérêt de l'être humain pour une interaction riche avec les équipements robotiques. Dans le film *I, Robot* de *Alex Proyas*, les robots vivent parmi les humains, ils descendent les poubelles, promènent les animaux domestiques et sont les compagnons idéaux des personnes âgées. Ils sont capables de dialoguer avec les humains et de prendre des décisions. Malheureusement, le film ouvre une perspective tragique sur le fait que les robots puissent se rebeller contre les humains et devenir dangereux. Cela note la tendance commune des êtres humains à être attiré par les robots, mais craintif. Dans le film *A.I. Artificial Intelligence*, de *Steven Spielberg*, les robots ont aussi une importance capitale dans la vie quotidienne, ils assurent la plupart des tâches domestiques. Ce film est différent du premier en ce sens où il parle d'un garçon de onze ans, David, qui est en fait un robot créé par un professeur. *C'est en réalité le premier androïde sensible capable de développer un vaste répertoire d'émotions et de souvenirs.*

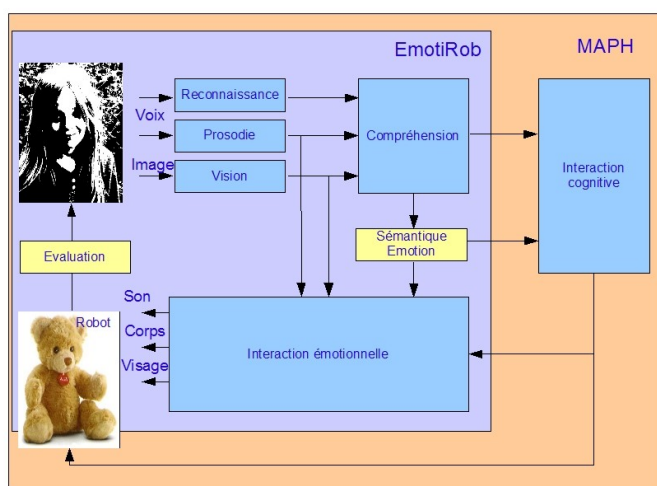


Illustration 1: Synoptique du projet EmotiRob

affective.

Mon travail consiste à reprendre le stage de Thi Hai Ha Dang, sur la définition et la modélisation de l'émotion et d'étudier les méthodes d'analyse de l'état émotionnel d'un individu. Je ne vais pas faire une recherche approfondie sur la théorie des émotions puisque cela a déjà été fait dans [1], [2] et [3]. Mais puisque les modules traitant l'émotion dépendent directement des autres modules (prosodie, vision, compréhension, interaction cognitive) et que le traitement associé influence directement le comportement du robot, j'ai décidé de faire l'état de l'art de toutes les techniques existantes pour créer une interaction émotionnelle avec l'utilisateur ce qui revient à créer une intelligence émotionnelle du robot. Rosalind W. Picard disait en 2001 que *l'intelligence émotionnelle consiste à la capacité de reconnaître, d'exprimer et d'avoir des émotions, associée à la capacité de réguler ses émotions, de les exploiter pour un but constructif et de manipuler habilement les émotions des autres.*

Par cette étude bibliographique, je cherche donc à répondre aux questions suivantes :

Qu'est-ce que l'intelligence émotionnelle (incluant qu'est-ce qu'une émotion ?) ? Quels sont les intérêts des émotions dans une interaction homme-machine ? A quoi reconnaît-on une

émotion sur autrui ? Quels sont les méthodes d'émotions existant actuellement ? Existe-t-il des robots doté d'intelligence émotionnelle ? Si oui, à quel niveau en sont-ils aujourd'hui ?

La suite du document sera composée de la façon suivante :

La première partie de cette bibliographie expliquera ce que sont les émotions, c'est un sujet difficile car plus de cent définitions existent et aucun consensus n'a encore été trouvé. Ensuite nous étudierons l'intérêt des émotions dans les interactions homme-machine afin de déterminer de l'importance de cette relation et des conséquences qu'elle peut avoir sur les humains. Puis nous verrons que les émotions peuvent être représentées par des expressions faciales, mais aussi par de la couleur et par d'autres moyens. Et comme nous voulons créer une interaction bidirectionnelle nous clôturerons cette première partie par l'étude des mesures de l'émotion sur autrui, car si nous ne sommes pas capables d'évaluer l'émotion de notre interlocuteur, il n'y aura aucun moyen d'agir sur ses émotions.

Dans la deuxième partie, nous nous intéresserons aux émotions appliquées à la robotique, nous commencerons donc par étudier les travaux existant sur la simulation et la détection des émotions, puis nous présenterons quelques robots sachant exprimer leurs émotions.

2.2. Les émotions

2.2.1. Que sont les émotions ?

D'un point de vue cérébral, les émotions sont gérées par l'hippocampe, qui enregistre le contexte des événements, et par l'amygdale, qui retient la connotation émotionnelle qui leur est attachée [8]. On apprend que lorsque quelqu'un se retrouve dans la même contexte d'une situation qui a provoqué une forte réaction émotionnelle, alors le cerveau se souvient de cette réaction et aura tendance à la refaire surgir. On apprend également que les stimulus passent par deux circuits, un court et un long. Le circuit court permet de réagir vite au stimulus avant même qu'il n'ait été analysé, tandis que le circuit long permet de vérifier la situation et soit de renforcer la réaction, soit de la neutraliser en cas d'erreur de jugement. Par exemple, si on marche sur la route et qu'on aperçoit une forme sur le sol (en pensant à un serpent), la première réaction sera la peur, puis l'analyse de la situation (on se rend compte que c'est un morceau de bois) neutralisera cette peur. Les émotions apparaissent comme des modes de fonctionnement du cerveau permettant de réagir rapidement face à une décision à prendre. Nous pouvons donc remarquer qu'un processus d'émotion a un **module de réaction**, un **module de réflexion** et un **module d'apprentissage et de mémorisation**.

D'un point de vue psychologique, il existe beaucoup de définition, beaucoup d'axes de théories. Cependant, il est possible de trouver certaines corrélations. [8] Concernant les théories des émotions, le point de départ est toujours un stimulus ou une perturbation, quelque chose qui va faire changer la situation. De ces changements s'effectue un traitement qui aboutit à une réaction émotionnelle. La plupart des différences entre les théories portent sur ces fameux traitements. Cependant, on retrouve souvent la notion d'évaluation. L'organisme évalue la situation à la réception du stimulus et réagit en fonction de ses connaissances. Ces dernières sont basées sur de la subjectivité ou alors parfois on parle de mémoire émotionnelle (au même sens que le point de vue cérébral l'expliquait). De plus, une émotion a une nature et une intensité. Cette explication est semblable à la définition des différentes théories hormis quelques unes qui sont totalement différentes.

Selon Scherer [4], l'émotion est confondue avec sentiment, passion et état affectif, elle a longtemps été vue comme une désorganisation et une désadaptation alors que selon lui les émotions sont, au contraire, décisives pour l'adaptation de l'individu. Il apporte un œil nouveau en apportant toute une dimension autour de l'émotion. Cette dernière est vue comme un processus de composants qui englobe tous les sous-systèmes organiques et qui est principalement dirigé par des changements dans l'évaluation de l'information reçue (voir l'illustration 2).

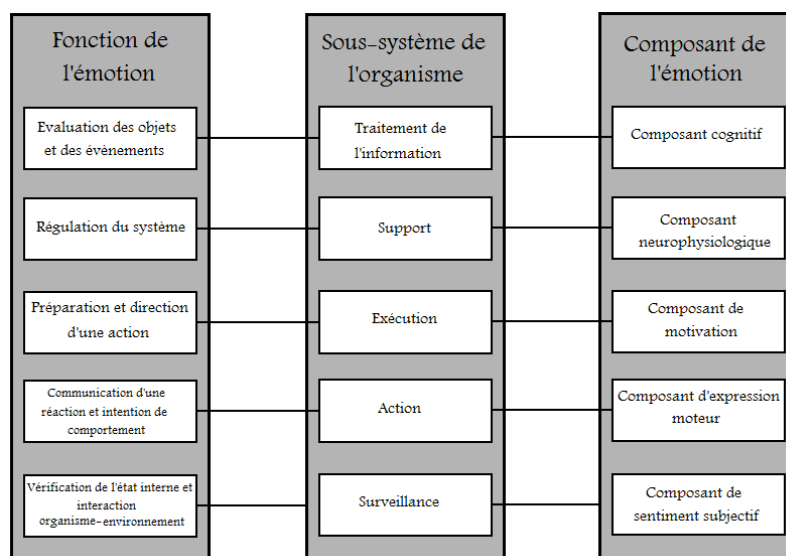


Illustration 2: Scherer - Organisation de l'émotion

L'émotion serait le résultat du travail en équipe des cinq sous-systèmes de l'organisme qui normalement fonctionnent de façon relativement indépendante. Ce travail synchrone serait lancé à la suite d'un stimulus interne ou externe. Le composant « composant de sentiment subjectif » serait le plus important. Une grande part du traitement émotionnel se fait de façon consciente, en particulier l'évaluation cognitive qui donne une signification à l'état affectif de l'individu. Le processus émotionnel est relativement lent, ce qui permet un traitement de l'information sophistiqué, mais il est toujours préparé à répondre à des situations d'urgence. *L'idée générale est que l'évaluation subjective d'une situation, d'un événement ou d'un stimulus détermine la nature (la qualité et la quantité) de la réaction émotionnelle. La force du système d'évaluation humain réside dans le fait que des niveaux différents d'évaluation sont disponibles et que le traitement peut être adapté à la nature du stimulus ou des besoins interprétatifs de l'organisme.*

Scherer explique qu'il ne faut pas confondre l'émotion avec d'autres phénomènes affectifs comme les sentiments, les humeurs ou les attitudes. Nous voyons sur l'illustration 2 que le sentiment est géré par le « composant de sentiment subjectif » et fait donc partie de l'émotion. En réalité, Scherer distingue quatre types de phénomènes affectifs qui diffèrent de l'émotion :

- ◆ Les préférences : aimer ou ne pas aimer un stimulus, le préférer à un autre.
- ◆ Les attitudes : c'est le fait d'avoir des croyances et des dispositions envers des objets ou des personnes spécifiques. Les attitudes provoquent des émotions.
- ◆ Les dispositions affectives : elles décrivent la disposition d'une personne à ressentir certaines humeurs plus souvent que d'autres.
- ◆ Les positions interpersonnelles : c'est un style affectif spontanément adopté lors de l'interaction avec d'autres personnes.

Beaucoup de recherches se font actuellement depuis le modèle de Scherer. A ces notions, il faut rajouter que l'on distingue les émotions ressenties des émotions exprimées [7] car un individu peut jouer avec ses émotions. Il peut les supprimer, les intensifier, les atténuer, les masquer ou les remplacer. De plus, les individus ressentent généralement plusieurs émotions en même temps. On distingue également les émotions primaires des émotions secondaires [8]. Les émotions primaires sont des réactions innées comme la peur, ce sont des émotions générées selon une composante non cognitive, tandis que les émotions secondaires sont le résultat d'évaluations cognitives. Chaque théoriciens a donné sa vision des émotions primaires, ce qui n'aide pas à arriver à un consensus.

Par exemple, nous trouvons une comparaison dans [7] :

- ◆ Izard 1971 : Colère, dégoût, joie, peur, surprise, tristesse, mépris, honte, intérêt, culpabilité, timidité.

- ◆ Plutchik 1980 : Colère, dégoût, joie, peur, surprise, tristesse, acceptation anticipation.
- ◆ Tomkins 1980 : Colère, dégoût, joie, peur, surprise, mépris, honte, intérêt, anxiété.
- ◆ Panksepp 1989 : Colère dégoût, tristesse, intérêt.
- ◆ Ekman 1992 : Colère, dégoût, joie, peur, surprise, tristesse, mépris, honte, intérêt, culpabilité, embarras, respect, excitation

Beaucoup de chercheurs considèrent qu'il y a six émotions de bases (primaires ou fondamentales selon les textes) : colère, dégoût, joie, peur, tristesse et surprise.

Un étude de Ekman propose une sorte de définition ou norme identifiant une émotion primaire :

- ◆ elle possède un signal universel distinct,
- ◆ elle est présente chez d'autres primates que l'humain,
- ◆ elle a une configuration propre de réactions physiologiques,
- ◆ elle est associée à des évènements déclencheurs universels distincts,
- ◆ elle a des réponses émotionnelles ou des composantes convergentes,
- ◆ elle est rapidement déclenchée,
- ◆ elle est évaluée automatiquement et
- ◆ elle apparaît spontanément.

En plus de Scherer, il est expliqué dans [0] que les théories d'Ortony et Lazarus sont très importantes.

La théorie d'Ortony considère les émotions comme des réactions de valence sur les évènements, les agents et les objets, qui, eux-mêmes, sont évalués en fonction des buts, des standards et des attitudes d'un individu. Cette théorie a l'avantage d'être proche de l'aspect informatif mais elle ne tient pas compte de l'intensité des émotions.

La théorie de Lazarus apporte la notion d'évaluation cognitive (*appraisal*) et d'adaptation (*coping – la faculté de faire face*). Lazarus définit l'évaluation cognitive comme un processus adaptatif qui permet de conserver ou de modifier la relation entre l'agent (ses croyances, ses buts) et le monde (ses contraintes, ses modifications), cela permet de maintenir l'équilibre. Il définit également l'évaluation primaire, qui évalue la pertinence et la congruence d'un événement, et l'évaluation secondaire, qui évalue ce qui doit être fait en réponse à un événement. La notion de coping a deux rôles : l'adaptation *problem-focused* qui essaie de résoudre le problème ou le nier pour minimiser son effet et l'adaptation *emotion focused* qui essaie de résoudre l'effet que le problème a sur la personne. L'état mental de l'individu est basé sur ses buts et ses croyances. Ses buts représentent ses intentions, les choses qu'il veut réaliser et les croyances représentent sa vision du monde, de façon subjective, ce qu'il pense que les choses représentent.

D. Duhaut [1] propose de définir l'émotion comme *un processus qui caractérise l'ensemble des réactions physiologiques et psychologiques d'un être humain à un événement pris à un instant donné*.

Un événement peut être résumé comme un changement externe ou interne (l'étude complète se trouve dans [1]).

En tant que processus, une émotion suit un algorithme :

- ◆ Perception de l'événement,
- ◆ Réaction à l'événement,
- ◆ Expression de l'événement.

Cet algorithme est répété sur tous les événements que l'individu identifie.

Son modèle d'émotion s'appuie sur les théories citées ci-dessus, celles de Scherer, Ortony et Lazarus [0].

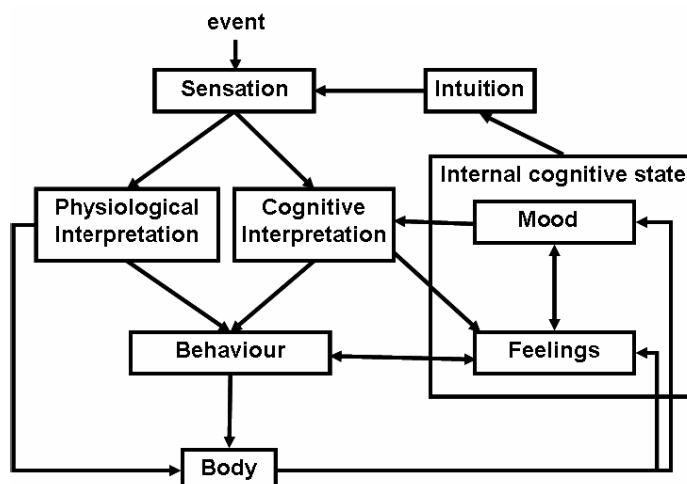


Illustration 3: Modèle d'émotion de D. Duhaut

L'entrée du modèle se fait par la sensation (déclenchée par un événement). Le modèle prend en compte à la fois l'interprétation physiologique et cognitive pour désigner le comportement adopté. Mais il propose également un module d'intuition, d'humeur et de sentiments qui rend le modèle complet et très intéressant.

2.2.2. Intérêt des émotions dans les interactions homme-machine

Le but de ce chapitre est d'analyser les recherches actuelles concernant l'intérêt des émotions dans les interactions homme-machine. Cette étude déterminera les cas de recherches qui ont permis de comparer les interactions entre le robot et les humains.



Illustration 4: Le robot Paro de T. Shibata

Une expérience a été effectuée avec le robot nommé Paro de T. Shibata développé par l'institut national japonais des sciences et techniques industrielles avancées [2]. Le laboratoire Valoria a testé ce robot sur des enfants handicapés, et des adolescents ayant des troubles du comportement. Paro s'est révélé très utile pour faciliter l'interaction entre les enfants et les thérapeutes et a également apporté de l'affection aux enfants. Il a également permis de se rendre compte qu'un robot ne doit pas être trop lourd et doit pouvoir offrir une interaction dans les deux sens : humain →

robot et robot → humain. Cette hypothèse est largement confirmée par [9] qui annonce qu'il est important de savoir reconnaître l'émotion d'une personne car cela a des applications conséquentes en médecine, dans les distractions et dans l'interaction homme-machine.

Dans son travail sur les Agents Conversationnels Animés (humanoïdes virtuels), dans [5] on apprend que les expressions émotionnelles créent l'illusion de vie et augmente le réalisme de l'agent. De plus, on dénote l'importance de l'empathie qui est la capacité de se mettre mentalement à la place de quelqu'un d'autre afin de pouvoir ressentir les mêmes émotions. Cela permet d'améliorer [5] *la perception de l'agent, d'induire des émotions positives et d'augmenter les performances et l'engagement de l'utilisateur dans la réalisation d'une tâche durant l'interaction*. Par contre, il semblerait que l'expression d'émotions non empathiques n'aurait peu ou pas d'impact sur l'interaction. Mais l'interaction homme-machine est différente si elle s'effectue entre un agent virtuel et un humain ou entre un robot et un humain. C'est ce que prouve [12] à travers ses expériences. Des sujets ont été présentés à un robot virtuel (la vidéo du robot en temps réel) et à un vrai robot. Certains aspects de l'interaction ont été observés comme la rencontre entre le robot et le sujet, leur coopération, la confiance et l'espace personnel. Le test s'effectuait de la façon suivante : d'abord, le sujet rencontre le robot, celui-ci pointe des livres du doigt et une destination pour que le sujet les déplace, d'abord sur une étagère éloignée, puis à la poubelle, puis sur une étagère juste derrière le robot. Les résultats ont montré que les sujets ont été plus à l'aise pour effectuer une action non habituelle avec le robot physique

(mettre les livres à la poubelle). De plus, l'absence de présence physique est un frein à la bonne interaction homme-machine car le temps de réaction est plus long et les consignes moins bien comprises. La présence physique du robot semble avoir développé une meilleure confiance en lui et un respect semble s'être instauré ; en effet lorsque les sujets ont du amener les livres derrière le robot, les seuls à l'avoir contourné sont ceux qui étaient en présence du robot physique. La conclusion de cette étude est que *la présence du robot est crucial dans l'interaction homme-machine, notamment pour développer la confiance et le respect de l'humain envers la machine*. De plus, [17] ajoute que l'échange d'information émotionnelle joue un rôle important dans l'interaction sociale humaine. Cette idée est développée par [19] dans son expérience avec les enfants autistes. On s'est rendu compte que seule une véritable interaction homme-machine peut permettre aux enfants de se développer socialement. Dans le cas où l'enfant devait se contenter de regarder le robot jouer, les enfants ont eu tendance à se refermer sur eux mêmes et à exprimer des émotions négatives. Le comportement du robot a complètement influencé le comportement de l'enfant autiste.

Dans [20], on ouvre une perspective supplémentaire en argumentant en faveur de la personnalité des robots. Dans [0], on explique que la personnalité a une incidence directe sur la façon dont nous agissons tous les jours, c'est ce qui nous pousse à choisir notre façon d'agir. Il semblerait que la personnalité ainsi que la gestuelle, le regard, le discours, le son et la proximité d'un robot sont des facteurs déterminant pour son acceptabilité des humains [20]. De plus, la personnalité serait une méthode utile pour faciliter l'interaction homme-machine parce que les humains pourraient prévoir et comprendre le comportement du robot et pourraient s'intéresser à des robots en fonction de leurs préférences.

Pour exprimer une émotion, il faut déjà que l'expression soit perçue par l'utilisateur donc on se propose d'explorer les moyens d'expression des émotions.

2.2.3. Expression des émotions

Ce paragraphe a pour objectif de faire la liste (non exhaustive) des moyens existantes pour représenter les émotions de façon à ce qu'elles soient bien perçues par un individu.

La première façon de percevoir des émotions est dans le discours des gens. Il existe un corpus de mots pour exprimer les émotions dont une liste de 256 mots peut être lue dans [1]. A titre d'exemple, nous trouvons toute une diversité de mots comme « décourager », « malaise », « trac »... Mais selon [10], la communication humaine se trouve à 80% dans l'expression faciale et les mouvements du corps. L'expression des émotions se fait donc beaucoup de manière non verbale.

Dans [9], on explique que Scherer propose de représenter les émotions suivant deux axes : le degré d'excitation (*Arousal*) et le degré de valence (*Valence*). L'excitation est une donnée partant du calme et de l'apaisement et allant jusqu'à l'activité et l'excitation, tandis que la valence est un axe dont la valeur négative correspond à ce qui est déplaisant et la valeur positive correspond à ce qui est plaisant. De [1] nous apprenons que Breazeal ajoute une troisième dimension qui est l'attitude (*stance*) qui est une donnée partant de l'attrait et allant jusqu'à la fuite.

Scherer dans [4] dit que l'expression faciale ne permet pas de définir des émotions précises, pareil pour la musique qui ne permet de définir exactement l'émotion suggérée, cette idée est rejoint par [10] dont l'étude sur la musique a révélé que la colère et la joie avait donné les mêmes résultats. Cependant, dans le cadre de mon stage, l'utilisation de capteur est impossible donc l'étude de l'expression faciale est nécessaire. Les émotions sont révélées à travers différentes régions du visage, chacune reflétant généralement une émotion [7]. L'expression faciale résultant du ressenti de deux émotions est différente de celle résultant d'une émotion ressentie et d'une émotion feinte. Une étude des expressions faciales de six émotions primaires est donné par [23] :

- ◆ Joie : émotion positive caractérisée par le coin des yeux élevés, les lèvres tirées, apparition de rides partant du bas du nez jusqu'aux côtés des lèvres et allant vers l'extérieur des coins des yeux.
- ◆ Surprise : émotion positive caractérisée par la bouche ouverte, les sourcils sont courbés et élevés, les yeux sont grands ouverts et l'apparition de rides horizontales traversant le front.

- ◆ Dégout : émotion négative caractérisée par le nez froncé, les pommettes élevées, les lèvres supérieures élevées, les sourcil baissé et des lignes formées sous les paupières inférieurs.
- ◆ Peur : émotion négative caractérisée par les sourcils qui se tirent et s'élèvent, les lèvres légèrement tirées en arrière et la bouche ouverte, les paupières supérieures mises en avant et les rides au milieu du front.
- ◆ Colère : émotion négative caractérisée par les sourcils qui se baissent et se rapprochent, les lèvres restent rapprochées, les yeux ressortent, les narines s'écartent et les paupières supérieures et inférieures se plissent.
- ◆ Tristesse : émotion négative caractérisée par le coin des lèvres qui tombent, la lèvre inférieure peut trembler, les côtés extérieurs des sourcils sont tirés vers le bas et la peau en dessous des paupières devient triangulaire et s'élève.



Illustration 5: Modèle de Plutchik

La couleur serait très expressive en ce qui concerne les émotions. Il est montré que les couleurs à courtes ondes radio sont plus plaisantes que les couleurs à longue onde radio [10], ce qui signifie que l'on se sent mieux auprès des couleurs chaudes. Pourtant il ressort également qu'un état d'anxiété est associé au rouge et jaune, qui sont plus excitants que le vert et le bleu. Plutchik a poussé l'étude plus loin en comparant les émotions à une palette de couleurs [8]. Il fait correspondre ses huit émotions primaires aux couleurs primaires, et les émotions complexes à des couleurs secondaires. Par exemple, le mépris résulte de la colère et du dégoût. Il fait également varier ces émotions en intensité.

La couleur a été utilisée dans le stage de l'année dernière pour représenter les émotions (à travers une interface graphique). La plupart des personnes interrogées ont réussi à interpréter les émotions simulées par ces couleurs [2].

Jocelyn Scheirer et Rosalind W. Picard proposent dans [10] un nouveau système pour exprimer les émotions en utilisant des objets affectifs notamment elles étudient les niveaux de tension et d'excitation d'une personne à travers la préhension d'un objet. Si l'utilisateur serre fort l'objet on aura tendance à penser qu'il n'est pas relaxé. L'expérience a été menée sur un téléphone, le *touch phone*, qui affiche une icône qui change de couleur en fonction du niveau de pression qu'effectue la main sur l'appareil. Les résultats montrent qu'un dégradé de couleur et le jeu sur les contrastes permettent d'exprimer des paramètres de changements émotionnels. Les objets affectifs ont le pouvoir de rendre visible certaines émotions et de changer la façon dont les personnes communiquent entre eux.

2.2.4. Mesure des émotions sur autrui

Afin de pouvoir obtenir une interaction homme-machine efficace, il est nécessaire de connaître l'état émotif de l'interlocuteur afin de réagir. C'est pourquoi la mesure des émotions est importante. Elle se fait par différents moyens. Nous excluons les questionnaires adressés aux sujets, car il est facile de mentir et de ne pas exprimer exactement ce que l'on pense [11]. [20] nous apprend qu'une étude a montré que si les paroles et la gestuelle sont contradictoire, alors l'observateur donne l'importance à la gestuelle dans 64% des cas (Preuve de Cassel et al.).

Mesures physiques

Les émotions provoquent beaucoup de perturbations au niveau de l'organisme. Les mesures physiques proposent d'évaluer ces perturbations pour connaître l'état émotif d'une personne. Différents capteurs permettent cette évaluation (de [8] et [9]) :

- ◆ Activité électromyographique : ce capteur mesure la tension des muscles. Utilisé sur le corps il permet de noter le niveau de stress d'un individu, sinon utilisé sur le visage, il mesure la tension faciale. Les contractions des muscles faciaux signalent les expressions faciales avant même qu'elles ne soient visibles à l'œil nu et même si

- le sujet tente de feinter ses émotions.
- ◆ Activité électrodermale : ce capteur mesure les variations électriques de la peau afin d'évaluer la conductance de la peau. Celle-ci devient un meilleur conducteur d'électricité quand le corps a reçu des stimulus internes ou externes.
 - ◆ Fréquence respiratoire : on mesure l'expansion et la contraction de la cavité abdominale autour du diaphragme (*Hall effect respiration*). Des excitations émotionnelles engendrent des respirations plus profondes. Attention au stress qui à la fois augmente la respiration sauf pour les stress ponctuels qui provoquent un arrêt momentané de la respiration.
 - ◆ Pression sanguine : on évalue la pression dans les artères, lorsque celle-ci augmente on dénote une excitation émotionnelle.
 - ◆ Électrocardiogramme : mesure le rythme cardiaque. La valence émotionnelle pourrait être prédite par la fréquence cardiaque. Une valence négative sera associée à un rythme cardiaque inférieur à celui d'une valeur positive.

A partir des quatre premières mesures, Rosalind W. Picard [9] propose des formules statistiques permettant de capter l'émotion d'une personne.

Mesures par observation

Les principales observations permettant d'extrapoler les émotions d'un sujet sont l'écoute de la voix et l'étude des expressions faciales. N'ayant trouvé aucune recherches évoquant l'étude de la voix, ce point sera consacrée aux émotions perçues sur le visage.

Ekman et Friesen ont créé en 1978 la méthode de FACS (*Facial Action Coding System*) qui codifie de façon standard tous les mouvements possibles sur un visage afin de décrire les actions faciales (*Par exemple : baisser les paupières, bouger les lèvres*). Cette méthode est reprise par beaucoup de chercheurs pour observer les expressions faciales.

Mais les émotions ne passent pas qu'à travers la voix et le visage. Toute la gestuelle, la posture est à prendre en compte, ainsi que le mouvement en lui même. Est-il fluide, hésitant, saccadé ? Toutes les caractéristiques de la gestuelle du corps peuvent dénoter une émotion. L'étude de l'observation est complétée dans le chapitre II.1 sur la simulation et la détection d'émotions.

2.3. Quelques implémentations

2.3.1. Simulation et détection d'émotions.

Dans ce paragraphe, nous allons voir les résultats de recherches sur la simulation d'émotions par les robot et sur la détection des émotions des humains.

L'article [7] se concentre sur la modélisation de la combinaison des émotions qui concerne la superposition ou séquence de plusieurs émotions fondamentales (colère, dégoût, peur, joie, tristesse et surprise). Chaque combinaison conduit à une expression faciale différente. Le modèle est fondé sur la logique floue. Les émotions sont révélées à travers différentes régions du visage. Les émotions négatives sont principalement perçues à travers la région supérieure du visage alors que les émotions positives sont perçues à travers la partie inférieure. Le visage est découpé en deux zones, par une ligne passant par les yeux, on distingue donc la région supérieure (qui exprimera les émotions négatives) et la région inférieure (qui exprimera les émotions positives). Ainsi lors de la superposition de deux émotions, on applique une expression faciale dans la région supérieure et une autre dans la région inférieure, sachant qu'il y a des règles de priorités. Par exemple, la colère a une prédominance supérieure à la peur qui elle-même a une prédominance supérieur à la joie. Ainsi, si l'on veut représenter la colère et la peur, on représentera la colère dans la région supérieure et la peur dans la région inférieure. Cela revient à dire qu'entre deux émotions, celle qui est prédominante sera exprimée dans la région supérieure du visage. Pour ce faire, on utilise les règles de logique floue, notamment en utilisant les opérateurs « > » et « ≤□> » pour l'ordonnancement des émotions.

D'autres techniques ont consisté à étudier des vidéos, d'annoter toutes les remarques concernant [6]:

- ◆ les émotions, grâce à un vecteur d'émotion. Par exemple, 56% de désespoir, 33% de colère et 11% de tristesse. Le vecteur est utilisé pour associer une combinaison de

comportement des yeux et des sourcils.

- ◆ Les mouvements des yeux : direction du regard (haut, bas, gauche, droite), la durée du regard.
- ◆ Les mouvements des sourcils (froncés ou relevés avec la durée) et l'intensité.

Après avoir annoté les caractéristiques des émotions, le but de l'expérience est de faire rejouer ces émotions par un agent virtuel.

En robotique, les composantes statiques et dynamiques sont importantes. Des positions particulières appartiennent à la composante statique. La façon de passer d'une position à une autre appartient à la composante dynamique et est déterminante dans le réalisme du mouvement. Ainsi pour améliorer l'expressivité des, quatre annotations sont rajoutées : la fluidité, la force, la vitesse et l'expansion spatiale du mouvement. Ces expériences ont permis de rejouer les émotions très bien mais cette version brute est bien trop expérimentale pour pouvoir l'utiliser dans une interaction bilatérale. Cependant, cela reste une ouverture possible.

Mais pour obtenir une réelle interaction bilatérale, la meilleure façon est l'observation de l'humain participant à l'interaction. Ainsi, [14] propose une solution assez complète. Cette solution consiste en un système de vision active qui inclut la détection de visage, du *tracking*, de la reconnaissance et de l'analyse d'expression faciale. Le système cherche des visages humains sur la vidéo, cible les visages et les suit et analyse les paramètres d'émotions. Le système est composé de trois modules :

- ◆ Détection de visage et repérage : La caméra est l'œil du robot, elle commence par détecter les visages dans son champ de vision et ensuite elle suit le mouvement de ce visage en temps réel (*tracking*).
- ◆ Reconnaissance du visage : le système est capable de s'adapter à différentes conditions lumineuses et à différentes orientations de visage. L'utilisation d'une caméra PTZ permet de zoomer sur un visage et de reconnaître le visage ensuite, par son identité, grâce à une base de données qui stocke toutes les identités de visages connus.
- ◆ Reconnaissance d'expression faciale selon six émotions primaires : joie, tristesse, peur, dégoût, surprise et colère. Chacune de ces expressions est associée à une seule expression faciale. La reconnaissance d'expression faciale est divisée en deux catégories principales : l'approche orientée cible, qui permet de déterminer l'expression faciale en fonction de l'image du visage, et l'approche orientée geste, qui utilise l'information temporelle d'une séquence d'images d'un mouvement d'expression faciale.

Le système utilise un modèle en trois dimensions pour analyser les muscles faciaux et les propriétés élastiques de la peau et un perceptron multi-couche pour l'analyse de l'expression faciale.

Le temps moyen de reconnaissance est de une seconde, pour un taux de bonne reconnaissance de 94%.

Pour être capable d'exprimer des émotions empathiques [5], on doit être capable de simuler le processus de déclenchement des émotions de l'utilisateur. Pour ce faire, il faut à la fois connaître l'état mental de l'utilisateur (buts et croyances) et les circonstances dans lesquelles les émotions se sont déclenchées. La méthode consiste à identifier les caractéristiques des situations dans lesquelles un utilisateur exprimer des émotions durant son interaction avec un agent conversationnel. A partir de ces connaissances, l'agent peut déterminer dans quelles situations l'utilisateur peut ressentir des émotions et donc dans quelles situations l'agent devrait exprimer des émotions empathiques. Selon l'article, si un individu pense qu'un événement affecte un de ces buts, alors cet événement va générer une émotion négative, si, au contraire, le but est facilité, alors l'émotion générée est positive. Il y a quatre cause possibles amenant à l'échec d'un but :

- ◆ les conflits de croyance : lorsqu'un interlocuteur pense que l'autre a un but qu'il n'a pas.
- ◆ Les conflits de but : lorsqu'un interlocuteur a un but que l'autre pense déjà réalisé.
- ◆ L'information de faisabilité : un interlocuteur a informé l'autre que le but est réalisable.
- ◆ L'information d'infaisabilité : un interlocuteur a informé l'autre que le but n'est pas réalisable.

L'expérience a été faite sur 47 interactions de communication et les déductions suivantes ont

été trouvées :

Les émotions négatives sont majoritairement déclenchées suite à l'échec d'un but (dans 45 cas sur 47), mais également :

- ◆ suite à l'échec d'un but causé par un conflit de but,
- ◆ suite à l'échec d'un but causé par un conflit de croyance,
- ◆ suite à l'échec d'un but alors qu'il y avait eu une information sur sa faisabilité,
- ◆ suite à un conflit de croyance et
- ◆ suite à un conflit de but.

Lorsqu'une émotion négative a déjà été déclenchée, le nombre d'échec de but pour en déclencher une autre diminue (c'est la notion de coping vue dans la définition) ; en effet une première émotion est déclenchée par 3 ou 4 échecs d'un même but tandis qu'une deuxième émotion est déclenchée en moyenne par 2 échecs.

En conclusion, on retient qu'il y a deux types de déclenchement d'une émotion négative : l'échec ou la menace du but. Ainsi, on note l'importance qu'une interaction homme-machine soit bilatérale et que le robot soit réellement capable de comprendre l'humain afin de pouvoir réagir à ses propos et l'aider à réaliser son but ou du moins ne pas l'empêcher de l'atteindre.

L'article [20] apporte un axe de recherche supplémentaire en introduisant une personnalité du robot. La personnalité est basée sur le *Myers-Briggs Type Indicator (MBTI)* et les émotions sont contrôlées par des paramètres de taille, de vitesse et de fréquence de la gestuelle du robot. Insérer une personnalité dans le robot permet une meilleure acceptabilité auprès des humains. Le MBTI regroupe quatre paramètres de personnalité :

- ◆ Extraversion (E) et Introversiion (I) : attitudes pour rassembler l'énergie
- ◆ Sensation (S) et Intuition (N) : fonction pour collecter l'information
- ◆ Pensées (T) et Sentiments (F) : fonction pour prendre des décisions
- ◆ Jugement (J) et Perception (P) : style de vie pour s'adapter aux circonstances.

L'expérience a montré que la gestuelle d'un robot influence la perception de l'humain concernant la personnalité du robot. Les paramètres Extraversion – Introversiion et Pensées – Sentiments sont observables à travers la gestuelle du robot. Par exemple, l'extraversion est associée à une gestuelle large, des mouvements rapides et une fréquence de mouvement élevée tandis que l'introversiion est associée à une gestuelle de petite taille, des mouvements lents et une basse fréquence de mouvements. Au final, les robots avec la personnalité d'extraversion ont été considérés comme plus agréables et plus compétents comparés à ceux qui avaient la personnalité d'introversiion.

Pendant cette recherche a une limitation concernant les interactions bilatérales mais elle offre un point de vue intéressant concernant l'insertion d'une personnalité dans un robot.

Pour finir, des chercheurs ont essayé de créer un robot génétique avec une personnalité, c'est un concept assez novateur [21]. Le robot contient plusieurs chromosomes dont les gènes contribuent à la définition de sa personnalité. Pour éviter les problèmes, liés à la complexité du système, qui empêchent la fiabilité, la variabilité et la cohérence du système, l'article propose un algorithme évolutionniste. Le robot est un agent virtuel qui vit dans un mobile. L'interaction est basée sur des stimuli envoyés par les touches du clavier. L'architecture interne est composée de quatre modules primaires :

- ◆ Module de perception : ce module peut reconnaître et évaluer l'environnement et par conséquent envoyer l'information au module d'état interne,
- ◆ Module d'état interne : concerne l'unité de motivation, l'unité d'homéostasie et l'unité d'émotion.
- ◆ Module de sélection de comportement : est utilisé pour choisir un comportement correct, basé sur l'état interne du robot.
- ◆ Module moteur : permet d'exécuter un comportement sélectionné dans

l'environnement 3D.

Ce qui est intéressant c'est que le robot a 14 chromosomes répartis de la manière suivante :

- ◆ 6 pour la motivation : curiosité, intimité, monotonie, fuite, gourmandise et désir de contrôle.
- ◆ 3 pour l'homéostasie : fatigue, somnolence et faim.
- ◆ 5 pour l'émotion : joie, tristesse, colère, peur et neutre.

Le problème c'est que le robot répond à des stimuli pré-fabriqués et que ça ne correspond pas à une vraie interaction bilatérale mais ça donne une perspective d'ouverture vers une intelligence émotionnelle basée en partie sur les algorithmes génétiques.

2.3.2. Quelques robot émotifs

Dans ce chapitre, trois robots vont être décrit. Ce chapitre n'a pas pour objectif de faire une étude complète sur les robots et leur modèle d'émotion car ce travail a déjà été réalisé par Thi Hai Ha Dang dans son rapport de stage [2]. Les robots présentés sont ceux qui ont participé à certaines expériences décrites dans le chapitre précédent.

iCat

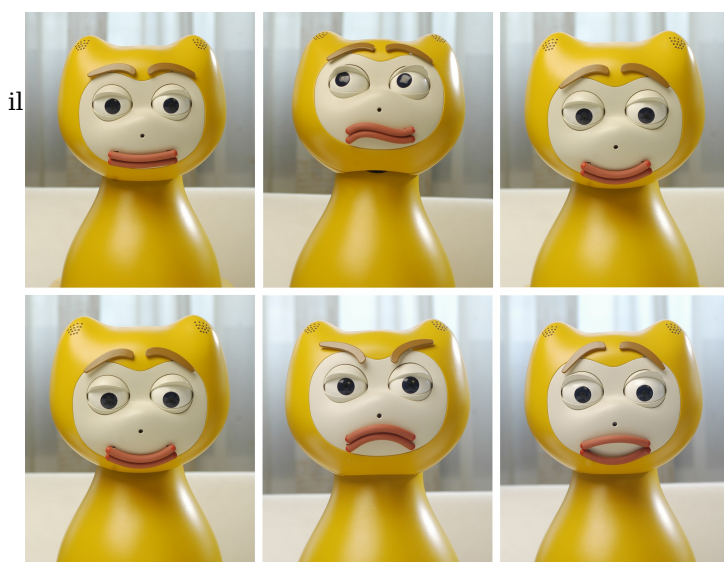


Illustration 6: Expressions faciale de iCat de Philips

iCat est un robot développé par Philips et a été utilisé dans [17], est capable d'interagir avec un humain, en lui expliquant des règles de jeu. Il est capable de comprendre ce que dit l'humain (du moins il comprend son nom). Lorsqu'on joue avec iCat, celui-ci décrit le coup qu'il doit faire et le joueur exécute ce coup. iCat est capable de se rendre compte si le joueur a joué un autre coup que celui qui était voulu. On voit sur l'image les différentes expressions faciales du robot. On voit que ses expressions ressemblent un peu à celles des émoticônes utilisées dans le chat de discussion et non au visage humain, comme par exemple le « s » que forme la

bouche sur la deuxième image du haut. Cependant les expressions sont tout de même reconnaissables.

Kismet

Kismet est le robot de Cynthia Breazeal du MIT. Il peut engager une interaction expressive avec les gens. Il possède quinze degrés de liberté pour bouger son visage afin de générer des expressions faciales. Il peut même apprendre de ses erreurs afin de ne plus les reproduire.

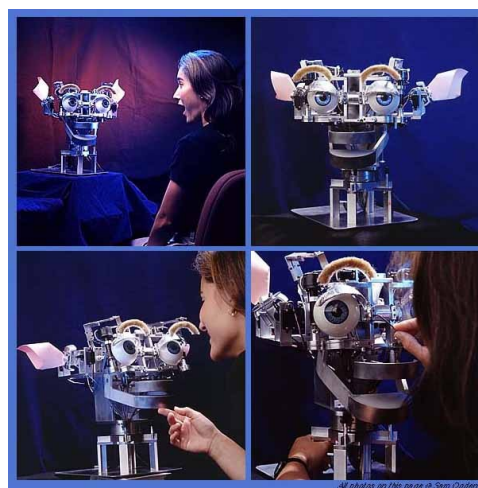
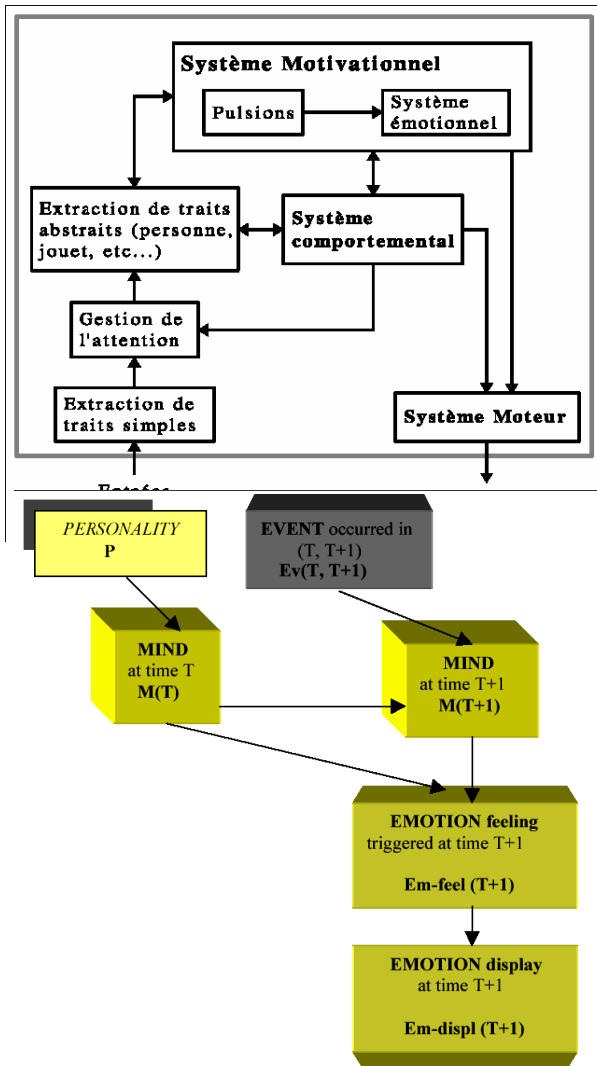


Illustration 7: Kismet



Greta

Greta est un agent conversationnel qui a été utilisé pour l'article [6]. Lorsque la réussite d'un but est menacé ou que l'atteinte au but a échoué, alors les croyances de l'agent change et cela provoque des émotions. Ces déclenchements d'émotions sont modélisés par un réseau dynamique de croyance. L'intensité de l'émotion déclenchée dépend à la fois de l'incertitude des croyances et de l'importance du but.



2.4. Conclusion

Cette étude montre qu'il n'y a, aujourd'hui, aucun consensus, ni aucune norme, ni aucun modèle permettant d'implémenter une interaction émotionnelle entre un robot et un humain. Créer de l'intelligence émotionnelle est quelque chose de difficile et aucun travail n'a réellement abouti à ce jour sur le sujet. Le modèle d'émotion sur lequel je travaillerai pendant mon stage [0] et [2] n'a pas besoin de se remettre en cause bien qu'il soit déjà ancien (il a deux ans) puisqu'il est au moins aussi bon que les autres modèles existants [2].

Pour répondre aux questions posées au début de cette étude, l'intelligence émotionnelle est une notion relativement bien définie puisqu'elle décrit une interaction homme → robot et robot → homme concernant les émotions, aussi bien la faculté de connaître ses émotions, de les contrôler, mais aussi la faculté de comprendre les émotions des autres et de pouvoir agir dessus. Le problème ne réside donc pas dans la compréhension de l'intelligence émotionnelle mais dans la compréhension de l'émotion elle-même. En effet, elle est l'objet de plus de cent définitions qui peuvent parfois être en désaccord par exemple en ce qui concerne les changements physiologiques liés aux émotions. Certains diront que les changements physiologiques engendrent des émotions et d'autres diront que les émotions engendrent des changements physiologiques. Peut-être faut-il tout simplement penser que les deux théories sont bonnes et que l'on doit considérer des changements en amont et en aval de la période émotionnelle. L'important serait de

prendre en compte la façon dont les émotions fonctionnent à l'intérieur du cerveau et d'essayer de reproduire ces actions au plus proche de la réalité. Il faut certainement prendre en compte la mémoire émotionnelle qui est présente à l'intérieur de chaque humain en fonction de son vécu, il faudrait aussi considérer une réaction émotionnelle instinctive et une autre basée sur la réflexion pour illustrer au mieux cette notion d'émotions qui protègent du danger. La vision de Scherer est à suivre, en ce sens qu'il considère l'émotion comme quelque chose de complexe et liée à l'humeur, au sentiment, à la perception du monde, à la subjectivité. Ces différentes notions font la différence entre les réactions émotionnelles des uns et des autres.

Concernant la question sur l'intérêt des émotions dans une interaction homme-machine, il est largement prouvé des biens fondés d'une telle interaction. Autant dans un but thérapeutique, que dans un but d'acceptabilité du robot. Le robot apporte également une présence bénéfique, mais pour cela il faut que l'humain ressente l'intérêt du robot à son égard.

Pour finir sur la question de l'existant, il y a beaucoup de propositions, beaucoup de robots qui expriment des émotions, ou alors repèrent les émotions des humains. Mais là encore, aucun robot n'est vraiment aboutit et ne peut avoir une réelle interaction intelligente et émotionnelle avec un interlocuteur humain.

3. Expression - Mise en œuvre d'Art-e

Art-e est le nom de l'avatar qui exprime des émotions dans le projet Emotirob. Art-e est un assemblage de **composants** qui participent à l'expression de ses émotions.

On utilise le terme composant pour désigner :

- ◆ un élément de l'avatar qui participe à la réaction émotionnelle,
- ◆ un élément indépendant capable de fonctionner seul ou avec d'autres composants,
- ◆ un élément de structure de l'avatar,
- ◆ un élément qui peut contenir et gérer d'autres composants (composite).

Les composants actuels, qui sont les sourcils, la bouche, les yeux, la tête et le tronc, sont décrits dans les chapitres suivants.

Un des objectifs de l'avatar est de coordonner les émotions avec Emi (le robot). Il est donc demandé qu'il possède au moins les mêmes capacités d'expression, à savoir les sourcils et la bouche, ce qui représente 6 degrés de liberté. L'avatar possède en plus des yeux, une tête et un tronc qui participent aux expériences émotives.

3.1. Sourcils + bouche (échange avec le robot)

La première étape obligatoire dans la construction d'Art-e a été de le doter de sourcils et d'une bouche ayant les mêmes degrés de liberté que ceux du robot afin que les deux acteurs puissent échanger des informations sur la position de ces composants.

Les sourcils du robot possèdent 1 degré de liberté chacun et sa bouche en possède 4. On retrouve ces mêmes degrés de liberté dans les sourcils et la bouche d'Art-e. De cette manière Emi et Art-e échangent des informations sur ces points là lors de leur coordination.

3.1.1. Les sourcils

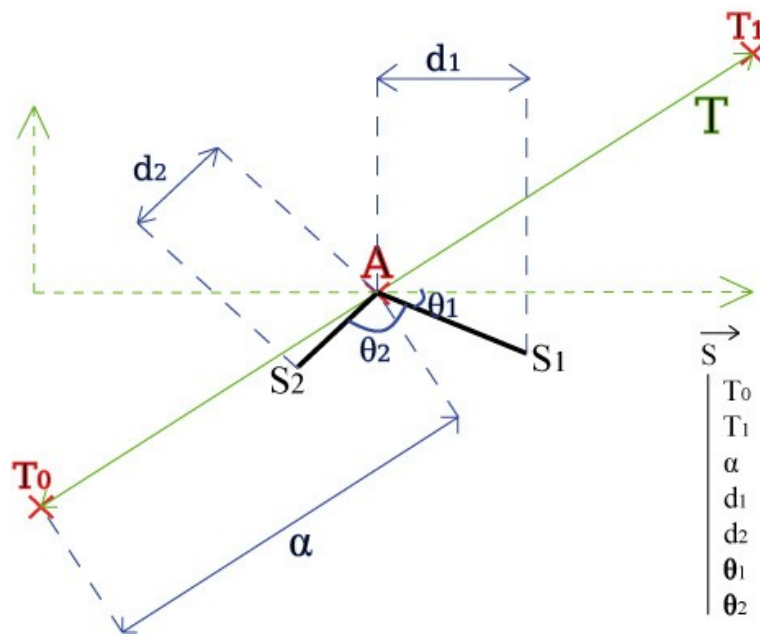


Illustration 8: Paramètres caractérisant le sourcil

Un sourcil est composé d'un point d'articulation A qui relie deux segments S_1 et S_2 .

Un sourcil est caractérisé par 7 valeurs :

- ◆ T_0 et T_1 sont les deux points formant le segment $T = [T_0T_1]$ qui représente la

trajectoire suivi par le point A.

- ◆ α représente le pourcentage de déplacement de A le long de la trajectoire T.
- ◆ d_1 et d_2 sont les longueurs de chacun des deux segments qui composent le sourcil.
- ◆ Θ_1 est l'angle formé par l'axe des abscisses et le segment S_1
- ◆ Θ_2 est l'angle formé par le segment S_1 et S_2

La position P_t de A à l'instant t se calcule de la façon suivante (sur le schéma $t=\alpha$) :

$$x_{P_t} = (1-t)T_0 + tT_1 \quad \text{et}$$

$$y_{P_t} = (1-t)T_0 + tT_1$$

Les coordonnées du point extrémité de S_1 se calcule à l'aide de d_1 et Θ_1 :

$$x_1 = x_{P_t} + d_1 \times \cos(\theta_1) \quad \text{et}$$

$$y_1 = y_{P_t} + d_1 \times \sin(\theta_1)$$

Les coordonnées du point extrémité de S_2 se calcule à l'aide de d_2 et $(\Theta_1 + \Theta_2)$:

$$x_2 = x_{P_t} + d_2 \times \cos(\theta_1 + \theta_2) \quad \text{et}$$

$$y_2 = y_{P_t} + d_2 \times \sin(\theta_1 + \theta_2)$$

Pour informer le robot de la position du sourcil, on lui transmettra la valeur α . Il pourra alors déplacer son sourcil en fonction de sa position minimum et maximum sur le même principe : $p = (1-t)min + tmax$ (fonction d'interpolation linéaire).

3.1.2. La bouche

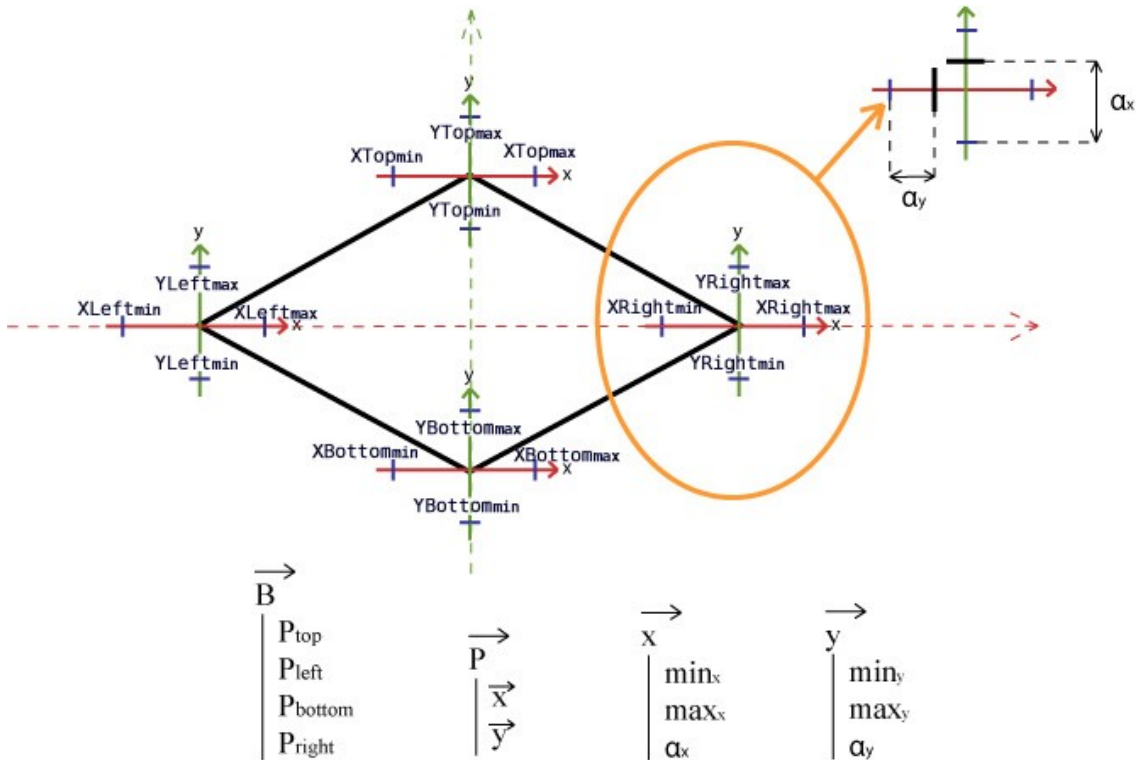


Illustration 9: Paramètres caractérisant la bouche

La bouche possède 4 points de contrôle qui sont caractérisés par 6 paramètres chacun.

- ◆ Y_{min} et Y_{max} : valeurs indiquant les bornes de déplacement du point de contrôle

sur l'axe des ordonnées..

- ◆ Xmin et Xmax : valeurs indiquant les bornes de déplacement du point de contrôle sur l'axe des abscisses.
- ◆ α_x (respectivement α_y) représente le pourcentage de déplacement du point le long de l'axe des abscisses (respectivement des ordonnées).

Les coordonnées d'un point de contrôle P pour un α donné se calcule de la façon suivante :

$$x_p = (1 - \alpha_x)x_{min} + \alpha_x x_{max}$$

$$y_p = (1 - \alpha_y)y_{min} + \alpha_y y_{max}$$

4 valeurs au total sont transmises au robot :

- ◆ La valeur en x de α concernant les points de gauche et droite
- ◆ La valeur en y de α concernant les points du haut et du bas

En effet, ce dernier a juste besoin de connaître la position relative du point par rapport aux minimum et maximum, soit α , car l'avatar et le robot ne possède pas les mêmes types de valeurs, mais savent tous deux interpréter α .

3.2. Les yeux

Une émotion exprimée avec les sourcils et la bouche est la base de notre système, mais elle n'est pas assez précise et ne suffit pas à obtenir une expression pertinente (reconnue par une majorité de personnes).

En effet, l'état de l'art a montré la difficulté de reconnaître une émotion basée sur les sourcils et la bouche.

L'ajout des yeux a donc pour but d'augmenter la capacité d'expression de l'avatar et de faciliter la reconnaissance de l'émotion qu'il exprime.

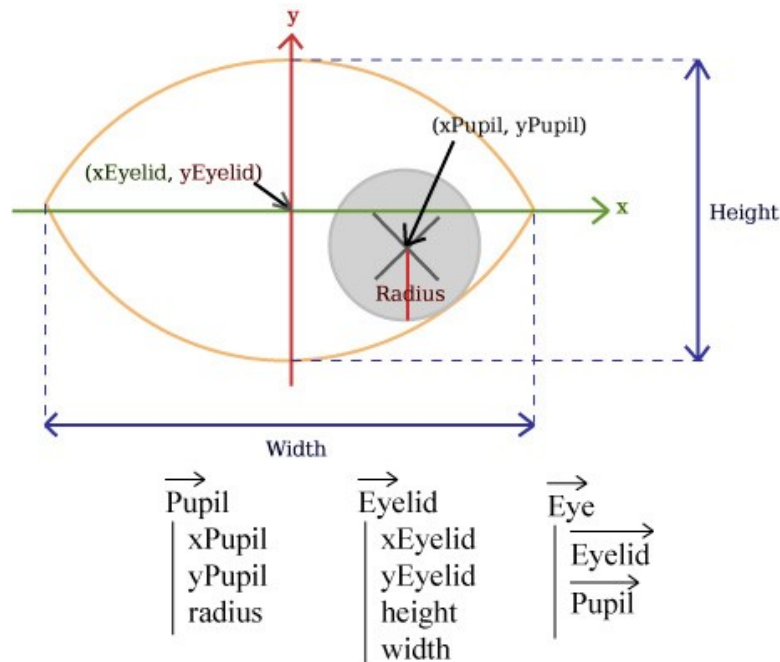


Illustration 10: Paramètres caractérisant les yeux

Un œil possède 10 paramètres :

- ◆ Les coordonnées en x et en y de la pupille de l'œil, **xPupil** et **yPupil** en fonction de l'axe (x,y) de l'œil.
- ◆ Le rayon **radius** de la pupille.
- ◆ Les coordonnées en x et en y de l'œil qui sert à tracer les paupières : **xEyelid** et **yEyelid**.
- ◆ La hauteur maximale du contour de l'œil : **height**.
- ◆ La largeur maximale du contour de l'œil : **width**.
- ◆ **small_eye**, **normal_eye** et **big_eye** : définit la hauteur et la largeur d'un petit, moyen ou grand œil.

Les deux arcs de cercle qui constitue le contour de l'œil sont tracés par une courbe de Bézier de degré 2 dont :

- ◆ Le point de départ est : $x = xEyelid - (width/2)$ et $y = yEyelid$,
- ◆ Le point d'arrivée est : $x = xEyelid + (width/2)$ et $y = yEyelid$,
- ◆ Le point de contrôle est
 - ◆ Pour le contour du haut : $x = xEyelid$ et $y = yEyelid + (height/2)$
 - ◆ Pour le contour du bas : $x = xEyelid$ et $y = yEyelid - (height/2)$

En faisant varier le centre de la pupille au cours du temps, le regard change de direction.

En faisant varier **height** et **width** au cours du temps, la taille de l'œil change, ça peut permettre de simuler l'étonnement par exemple.

Pour fermer la paupière, on utilise un dernier paramètre :

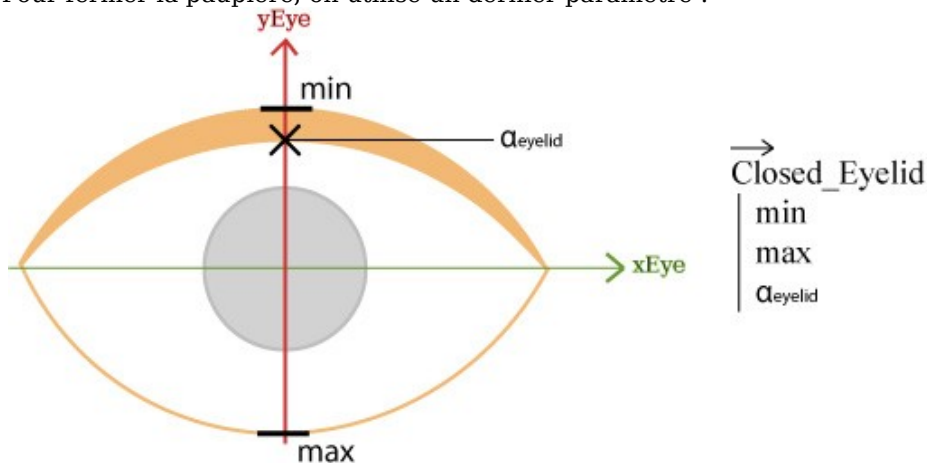


Illustration 11: Mécanisme permettant de fermer l'œil

- ◆ La valeur **min** est le point de contrôle du haut
- ◆ La valeur **max** est le point de contrôle du bas.
- ◆ α_{Eyelid} est le pourcentage de déplacement en y sur le segment [min max], dont la direction est inverse à l'axe des ordonnées.

Lorsque $\alpha_{Eyelid} = \mathbf{min}$ alors l'œil est ouvert, et lorsque $\alpha_{Eyelid} = \mathbf{max}$ alors l'œil est fermé.

En faisant varier α_{Eyelid} au cours du temps, on peut simuler le clignement de l'œil.

On dispose alors d'un mécanisme assez complet pour pouvoir rendre l'œil réaliste et émotif.

3.3. La tête et le tronc

3.3.1. La tête

La tête est un composant composite. Cela signifie qu'elle est un composant et qu'elle contient aussi des composants.

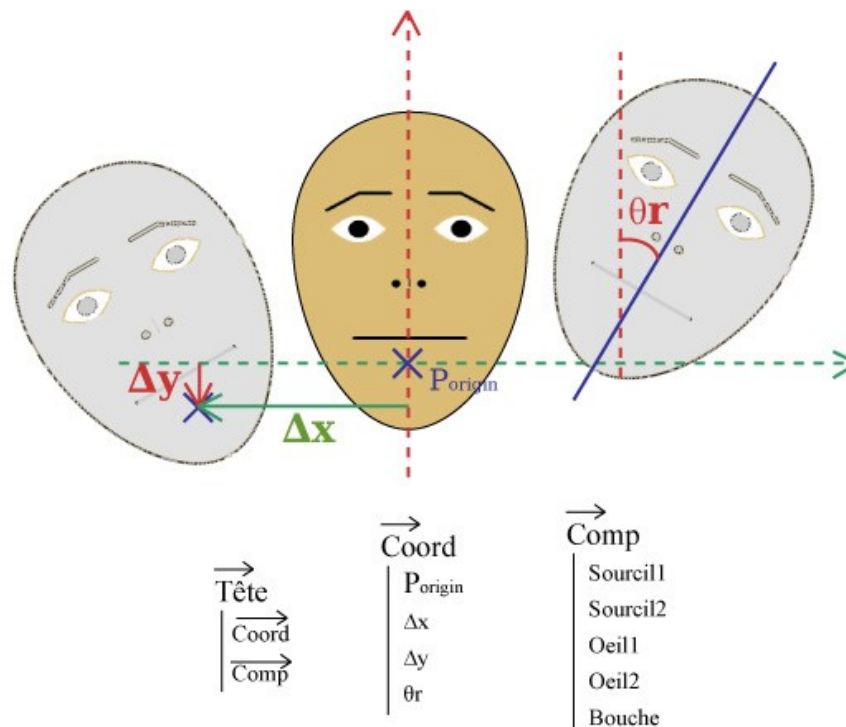


Illustration 12: Paramètres caractérisant la tête

La tête est caractérisée par 4 valeurs + un ensemble de composants :

- ◆ **P_{origin}** est le point d'origine de la tête.
- ◆ **Δx** représente la translation en x du point d'origine.
- ◆ **Δy** représente la translation en y du point d'origine.
- ◆ **Θ_r** est l'angle de rotation de la tête.
- ◆ L'ensemble de composants contenus = {sourcil, sourcil, œil, œil, bouche}

Les composants contenus héritent des valeurs **Δx**, **Δy** et **Θ_r** mais la réciproque est fausse. Cela signifie que :

- ◆ le modification d'une coordonnée de la tête entraîne la modification de la même coordonnée de tous les composants contenus,
- ◆ par contre une modification d'une coordonnée d'un composant contenu n'entraîne aucune modification dans le composant conteneur.

Art-e penche la tête lorsque celle ci effectue une rotation.

3.3.2. Le tronc

Le tronc de l'avatar est caractérisé par les mêmes valeurs de coordonnées que la tête, à savoir un point d'origine, une valeur de translation en x, en y et un angle de rotation.

Il est ajouté pour pouvoir simuler la respiration.

Pour choisir la technique employée pour simuler la respiration, il faut quelques pré-requis d'anatomie humaine. Cette prise de connaissance montrera qu'on ne peut pas reproduire le vrai mécanisme de la respiration sur Art-e et qu'il est nécessaire de simuler la respiration d'une autre manière.

Le centre de gravité du corps humain se situe à l'intersection de trois plans : le plan sagittal, la plan frontal et le plan transversal.

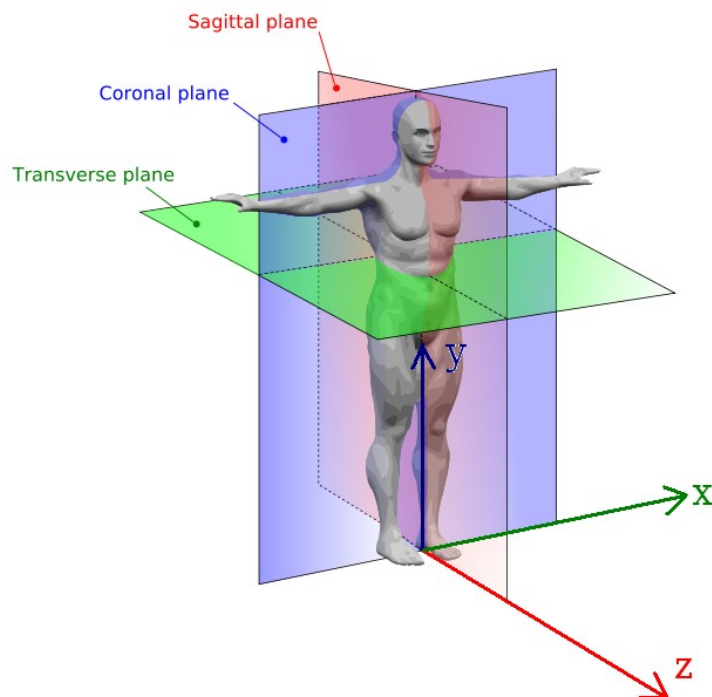


Illustration 13: Les 2 plans de coupe de l'être humain

Dans le monde 3D, chacun de ces plans correspond à un des axes x , y , et z . La respiration a deux acteurs principaux, qui sont la cage thoracique et l'abdomen et qui agissent sur l'axe des y **ET** surtout sur l'axe des z [32].

Art-e est contraint car il ne possède que les axes x et y .

Le chapitre sur la dynamique des émotions proposera une alternative pour simuler la respiration dans un espace à deux dimensions de façon crédible.

3.4. Ensemble d'expressions faciales d'Art-e

L'ensemble d'expression faciale se base sur l'ensemble utilisé et validé par Sébastien Saint-Aimé [26] qui concerne les sourcils et la bouche. Les yeux, la tête et le tronc sont des spécificités d'Art-e. On retrouve les six émotions de base : joie, tristesse, dégoût, colère, peur et surprise, ainsi que l'expression neutre.

L'ensemble des expressions faciales d'une même émotion sont stockés dans un tableau à deux dimensions : `tableau[codeEmotion][indexFacialeExpression]`. Chaque émotion ne possède pas le même nombre d'émotion.

Une expression est choisie aléatoirement parmi l'ensemble stocké dans le tableau. Les figures se ressemblent à quelques variations près, cela permet d'accroître la crédibilité d'Art-e qui n'exprime pas toujours de la même manière. Il faut aussi pouvoir différencier une émotions très forte d'une émotion retenue, modérée...

3.4.1. La joie

Le code de la joie est 0. Les expressions faciales stockées dans le tableau sont visibles ci-dessous. En le comparant avec les autres tables, on peut remarquer que $\text{tableau}[\text{« joie »}][6] = \text{tableau}[\text{« tristesse »}][5]$. Une expression faciale peut être associée à plusieurs émotions. C'est une des raisons pour laquelle on ne peut pas se contenter des sourcils et de la bouche pour exprimer des émotions.

La joie est principalement représentée par une bouche ouverte et souriante, et des sourcils ouverts sur l'intérieur.

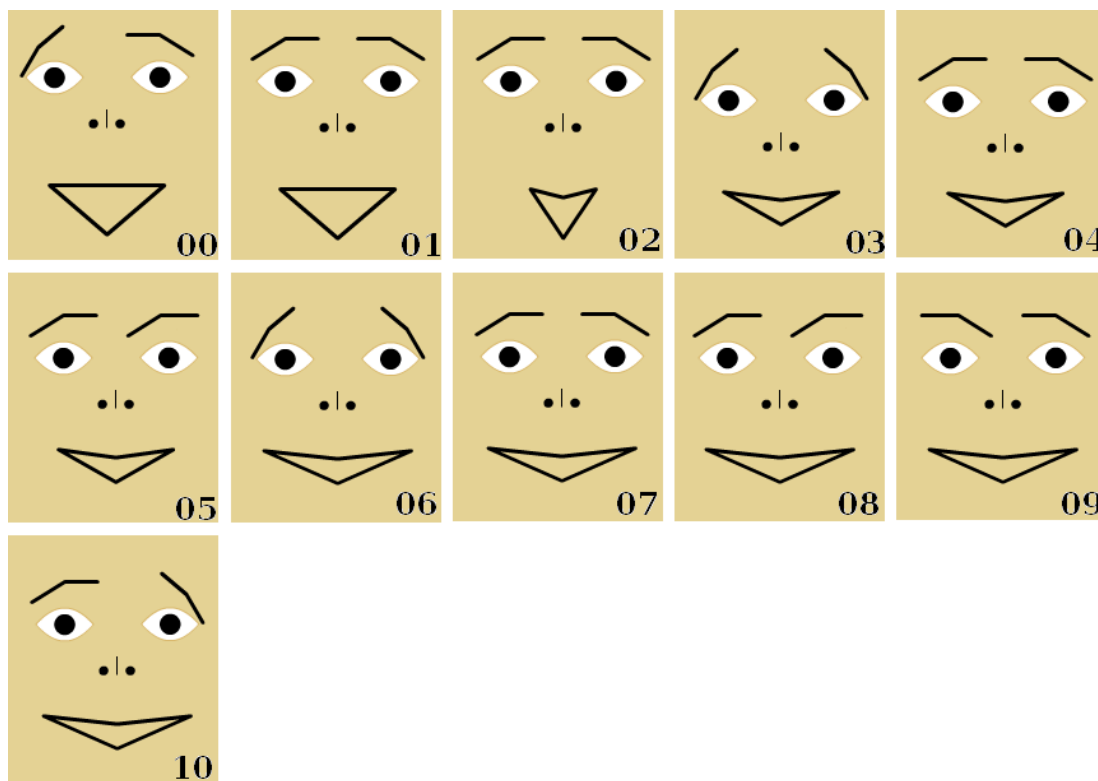


Illustration 14: Expressions faciales représentant la joie

3.4.2. La colère

Le code de la colère est 1.

La colère est principalement représentée par une bouche ouverte ou fermée, mais qui tombe sur les côtés, à l'inverse d'un sourire. Les sourcils sont ouverts sur l'extérieur à l'inverse de la joie. De plus, la colère est également exprimé avec un plissement d'œil et une légère inclinaison de la tête.

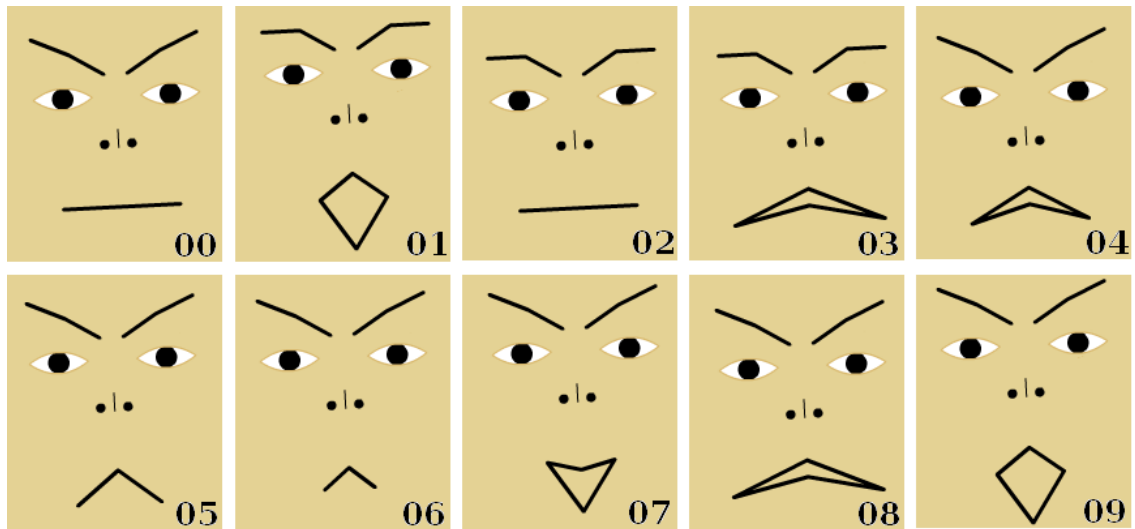


Illustration 15: Expressions faciales représentant la colère

3.4.3. La tristesse

Le code de la tristesse est 2.

La tristesse est principalement représentée par une bouche ouverte ou fermée, mais qui tombe sur les côtés, comme la colère. Les sourcils sont ouverts sur l'intérieur comme la joie. De plus, la tristesse est également exprimé avec un plissement d'œil sans inclinaison de la tête, comme le fait la colère.

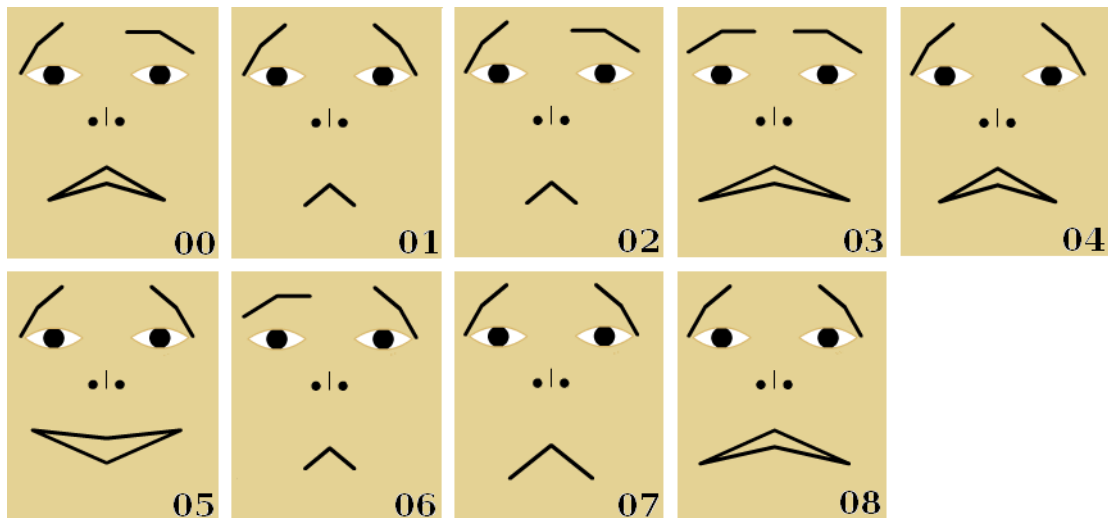


Illustration 16: Expressions faciales représentant la tristesse

3.4.4. La surprise

Le code de la surprise est 3.

La tristesse est principalement représentée par une bouche grande ouverte ou souriante. Les sourcils sont ouverts sur l'intérieur comme la joie. De plus, la surprise est également exprimé par des yeux grands ouverts et une légère inclinaison de la tête.

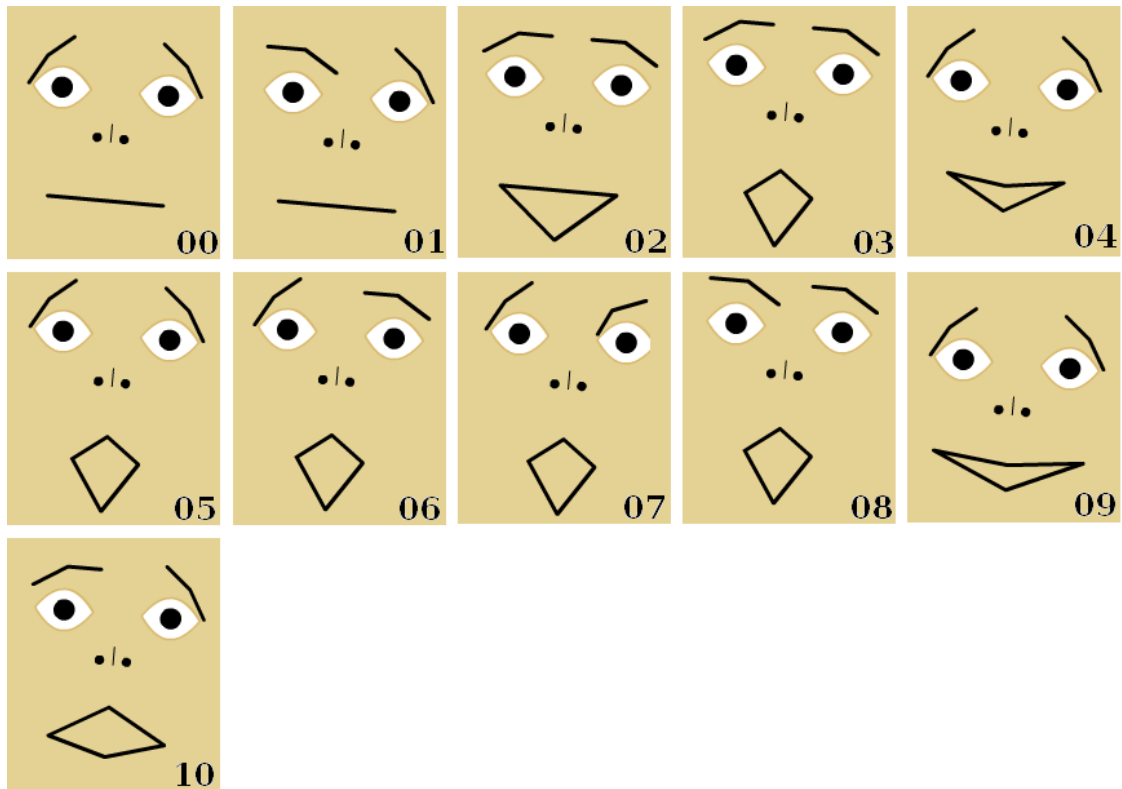


Illustration 17: Expressions faciales représentant la surprise

3.4.5. Le dégoût

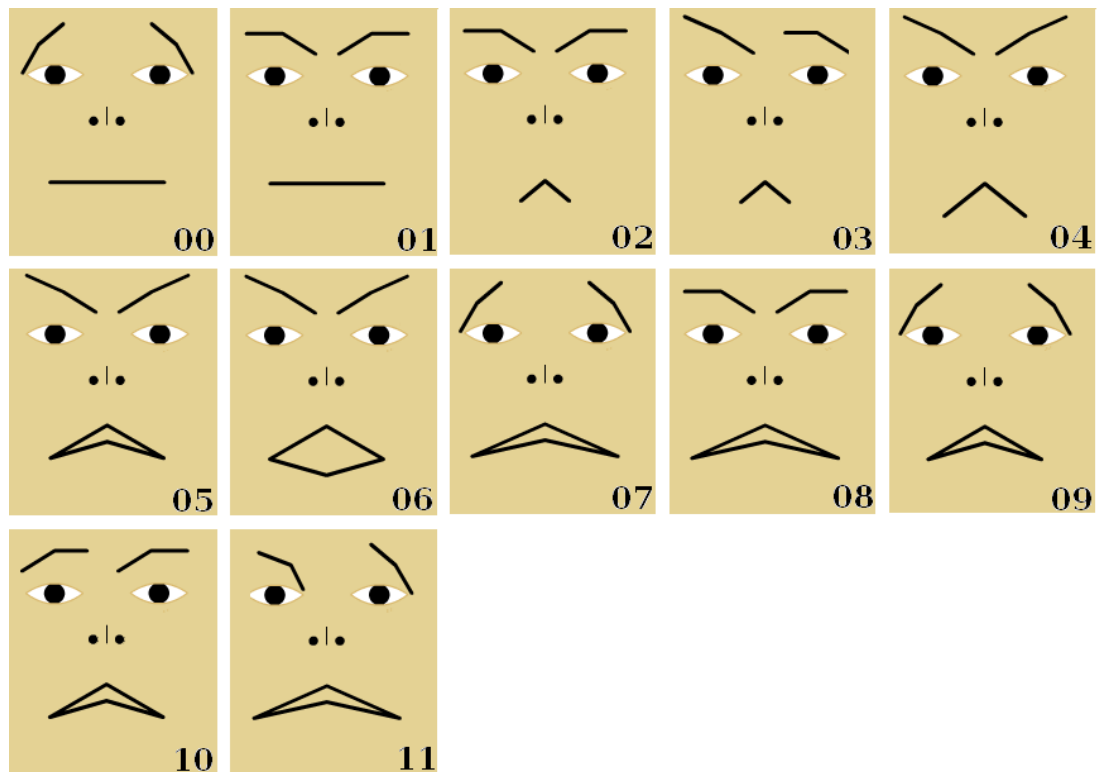


Illustration 18: Expressions faciales représentant le dégoût

Le code du dégoût est 4.

Le dégoût est principalement représentée par une bouche tombante sur le côté comme la tristesse. Les sourcils sont ouverts sur l'intérieur et sur l'extérieur. De plus, les yeux sont plissés.

3.4.6. La peur

Le code de la peur est 5.

L'expression faciale de la peur ressemble beaucoup à la tristesse mais avec des yeux grands ouverts.

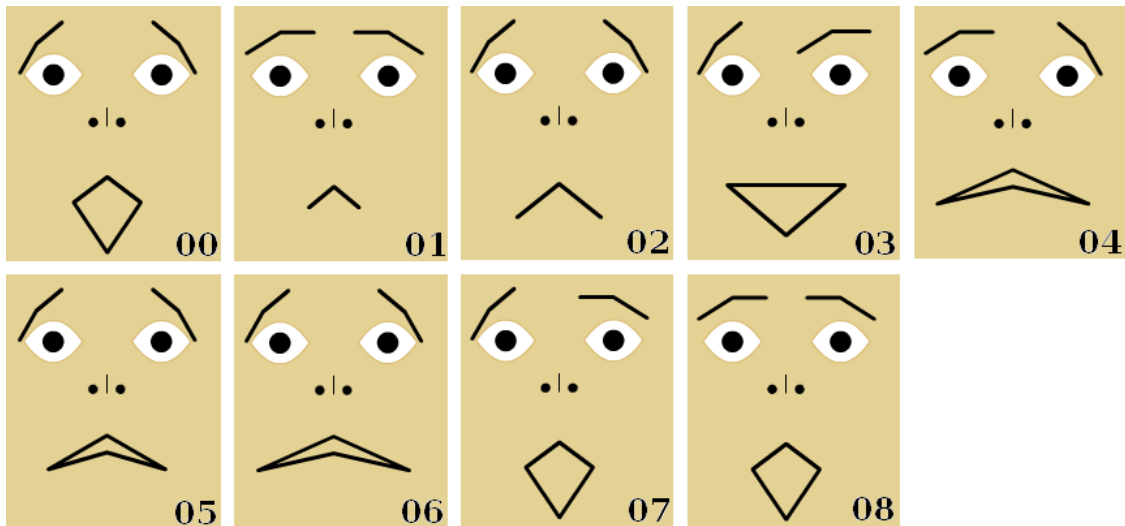


Illustration 19: Expressions faciales représentant la peur

3.4.7. Expression neutre

Le code de l'expression neutre est 6. Il y a peu d'expressions neutres, avec peu de variation. Les sourcils sont ouverts sur l'intérieur et là bouche est droite, légèrement ouverte ou souriante.

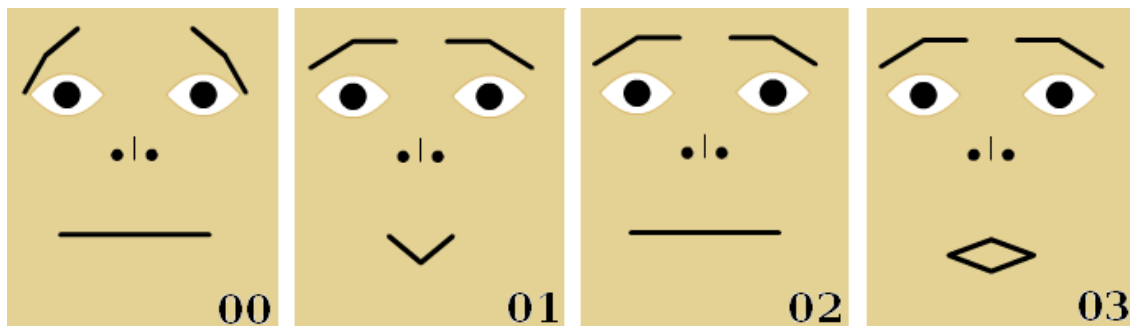


Illustration 20: Expressions faciales neutres

4. Dynamique des émotions

Dans un souci de crédibilité et de réalisme [27], on ne veut pas se contenter d'un avatar qui affiche statiquement des figures émotives. Il faut mettre au point un avatar capable de « vivre » à l'écran et capable d'exprimer des émotions de différentes manières d'où un besoin de dynamique.

Art-e possède trois modes comportementaux.

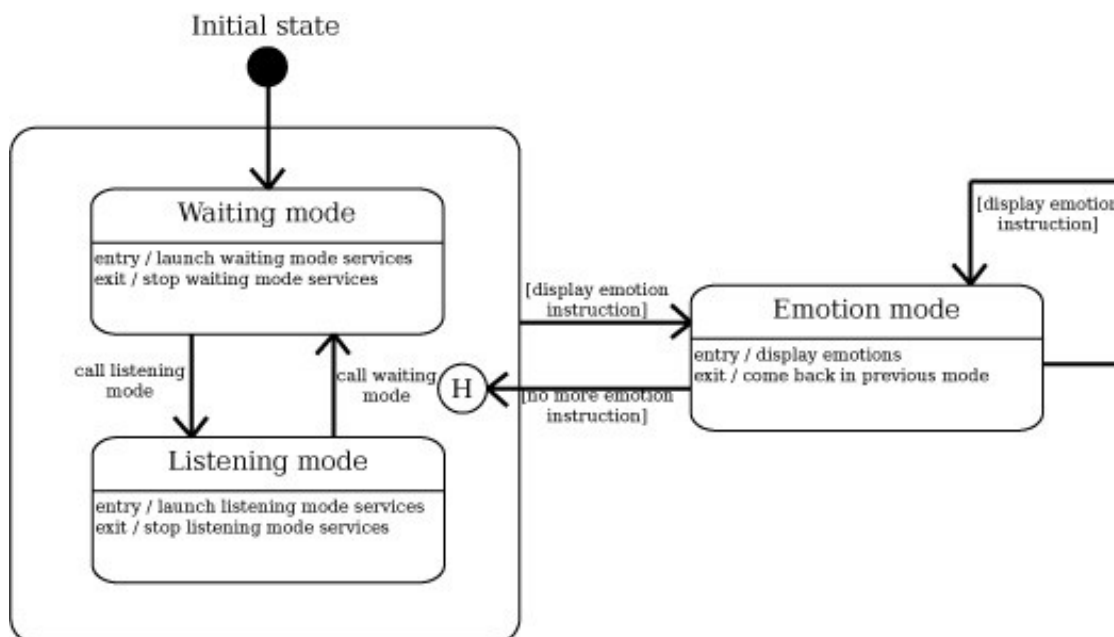


Illustration 21: Diagramme d'état d'Art-e

Le **waiting mode** représente le comportement de l'avatar lorsqu'il n'exprime rien de particulier. Il est libre de ses motivations et n'est pas concentré sur un interlocuteur. C'est le mode par défaut, à l'état initial.

Le **listening mode** représente le comportement de l'avatar lors d'un échange, par exemple s'il regarde une vidéo, écoute une histoire ou un interlocuteur. Dans ce mode, l'avatar est concentré et ses mouvements sont moins libres. Lorsqu'on entre dans ce mode, à l'appel d'une fonction, les services du waiting mode sont automatiquement stoppés et ceux du listening mode sont lancés.

L'**emotion mode** permet à Art-e de passer d'une émotion à une autre. Ce mode est lancé lorsqu'une instruction est envoyée à Art-e pour qu'il affiche une émotion. Les instructions peuvent se cumuler. Une fois qu'elles ont toutes été interprétées, Art-e sort de ce mode et retourne dans le mode précédent (waiting ou listening mode).

La suite de ce chapitre décrit les services mis au point pour mettre en œuvre ces trois modes.

4.1. Outils de base à la dynamique

Art-e possède quatre outils pour mettre en place la dynamique :

- ◆ **La respiration** : Le chapitre 3, page 14, a montré que les émotions ont une influence non négligeable sur le système respiratoire, c'est donc parce que l'on rajoute la respiration à l'avatar.
- ◆ **Le clignement des yeux** : Le clignement des yeux est important concernant le réalisme d'un agent [29], car il peut déterminer l'état de fatigue, de stress, de concentration d'un individu.

- ◆ **Le mouvement du regard** : On met en place un outil de *watching grass grow* (regarder les mouches voler) qui permet d'augmenter la crédibilité de l'avatar.
- ◆ **Les rictus** : On permet aussi à l'avatar de pouvoir avoir quelques rictus car le visage n'est jamais statique. Il est même possible de représenter des tics en provoquant les rictus à intervalle régulier.

4.1.1. La respiration

Le chapitre 3.3.2 a montré qu'on ne peut pas reproduire le vrai mécanisme de la respiration sur Art-e. On propose d'utiliser la fonction sinus qui permet d'obtenir un mouvement fluide de va et vient ou de « soufflet » [25]. La fonction sinus est utile car elle permet de contrôler la fréquence et l'amplitude de la courbe.

La fonction utilisée est la suivante : $f(t) = a \sin(\omega t + \varphi)$

- ◆ t représente le temps, c'est une valeur entre 0 et 1. Ce paramètre a été normalisé car, dans ce cas, on obtient exactement n périodes pour $\omega = n \cdot \pi$.
- ◆ a représente l'amplitude de la fonction.
- ◆ ω représente la fréquence. On la divise par π , afin d'obtenir une valeur qui indique le nombre de période pour une unité de temps.
- ◆ φ Représente le déphasage de la fonction.

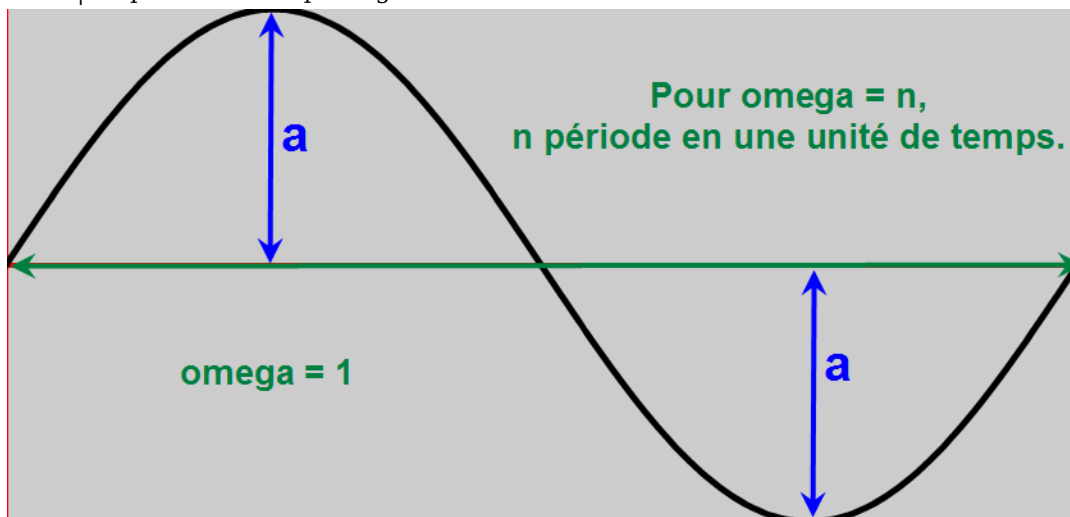


Illustration 22: Fonction de simulation de respiration

L'intérêt de ce mécanisme est de pouvoir varier facilement en fonction du temps en jouant sur la fréquence (facteur oméga) et l'amplitude (facteur a) de la sinusoïde (la variation sera expliquée dans « Passage d'une émotion à une autre »).

Mais la respiration est asymétrique. Certes, on peut assimiler la respiration à un mouvement cyclique mais la phase expiratoire est deux fois plus longue que la phase insufflatoire [25] – sachant que ces durées peuvent être légèrement modifiées.

C'est pour cela que le mécanisme d'Art-e respecte ces proportions en changeant la valeur de a en fonction qu'il soit en phase d'inspiration ou d'expiration :

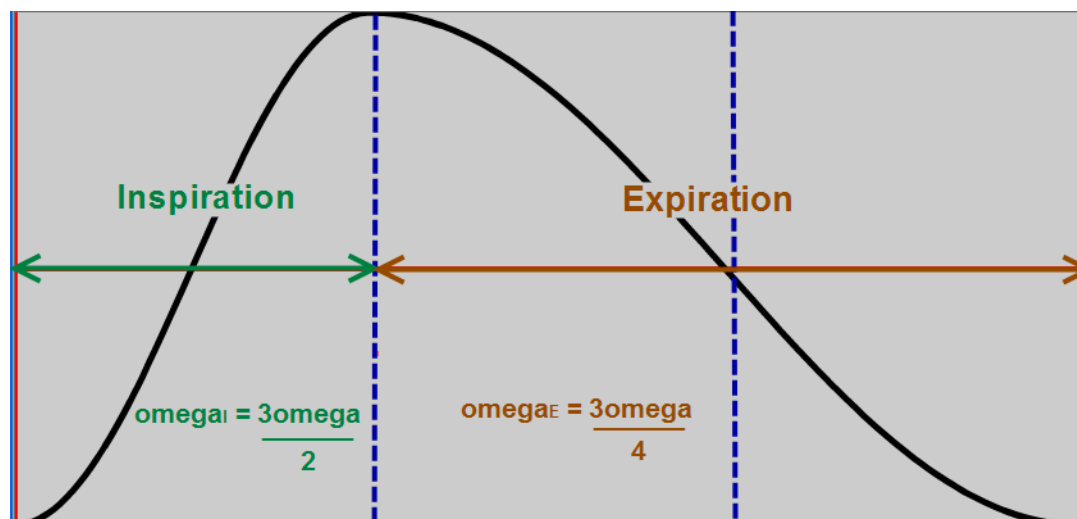


Illustration 23: Exemple de respiration pour $a=1$ avec un déphasage de $-PI/2$

Pour la simulation on a choisi une unité de temps égale à 5 secondes, ce temps représente la moyenne (inspiration + expiration) calculé sur des individus respirant calmement. Une unité de temps correspond au temps que mettra la sinusoïde pour parcourir n période(s) pour un oméga fixé à $n2\pi$.

Remarque : la respiration est le seul service qui tourne pendant les trois modes d'Art-e (waiting, listening and emotion mode).

4.1.2. Le clignement des yeux

Le clignement des yeux est un phénomène important pour l'être humain et peut caractériser en partie l'état de l'individu. S. Garten explique que le nombre de clignement augmente avec la fatigue entre autres [28] et que chaque personne cligne différemment les paupières. Dans son ouvrage, il explique la durée des différentes phases de clignement de la paupière.

- ◆ Phase 1 : la descente de la paupière est très brusque et rapide (durée entre 0.075 et 0.091 secondes).
- ◆ Phase 2 : l'œil reste fermé pour une durée assez variable, les durées les plus courtes sont d'environ 0.15 secondes.
- ◆ Phase 3 : le relèvement de la paupière est plus lent que sa descente et dure environ 0.17 secondes.
- ◆ Phase 4 : la durée entre deux clignements n'est pas clairement indiqué, mais il explique que la fréquence des clignements est très différentes et variables.

De plus il est important de noter que lorsque un individu est fortement concentré il ne cligne moins des yeux.

Art-e s'inspire de ce mécanisme et de cette notion de temps variable. Il y a huit paramètres qui permettent d'implémenter le clignement des paupières.

- ◆ **MIN_DURATION_CLOSING** et **MAX_DURATION_CLOSING** : temps de descente de la paupière.
- ◆ **MIN_DURATION_BETWEEN** et **MAX_DURATION_BETWEEN** : temps de pause entre la descente et la remontée de la paupière.
- ◆ **MIN_DURATION_OPENING** et **MAX_DURATION_OPENING** : temps de remontée de la paupière.
- ◆ **MIN_DURATION_AFTER** et **MAX_DURATION_AFTER** : temps d'attente avant la prochaine descente de paupière.

A partir de ces paramètres, on calcul quatre valeurs qui représentent les durées de chacune des phases décrites précédemment.

```
Duration := min + random*(max - min)
//random is a number between 0 and 1
```

Texte 1: Calcul de la durée d'une phase du clignement de l'œil

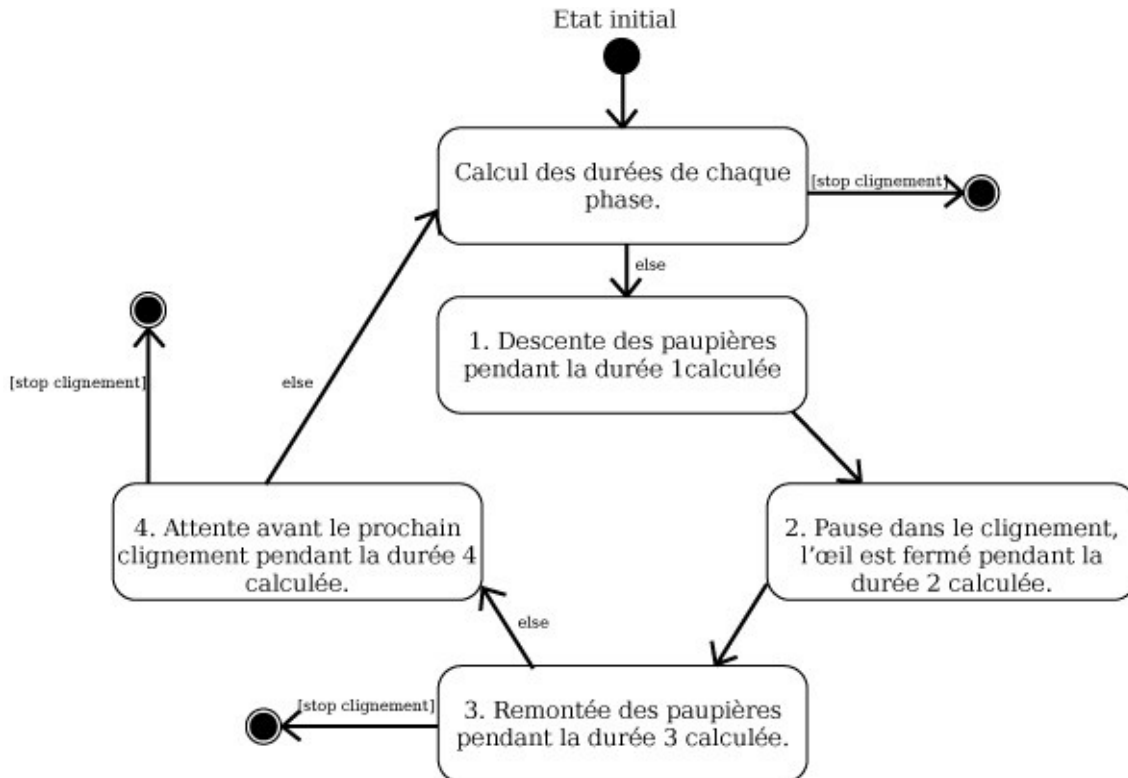


Illustration 24: Diagramme d'état du clignement d'œil

- ◆ On obtient pour les quatre phases une **durée** comprise entre le minimum et le maximum, qui est la durée d'exécution de cette phase.
- ◆ **Le clignement est lancé** ce qui correspond à l'exécution séquentielle des quatre phases.
- ◆ Lorsque le clignement est terminé, on **calcule quatre nouvelles durées**,
- ◆ puis **on recommence** le clignement.

Il n'y a que deux conditions pour pouvoir stopper le clignement :

- ◆ soit on stoppe explicitement le service
- ◆ soit l'*emotion mode* est lancé.

Le clignement des yeux reprend automatiquement lorsque l'*emotion mode* est terminé.

Dans les deux cas, on ne peut pas stopper le clignement de l'œil entre le début de la phase 1 et la fin de la phase 3. Si un ordre est envoyé pendant cette période là, il ne sera exécuté qu'à la fin de la période 3.

Remarque : le clignement des yeux est lancé pendant le waiting et le listening mode.

4.1.3. Le mouvement du regard

Art-e n'a pas le regard fixe, [29] c'est un des aspects qui rend un avatar réaliste. Un individu qui ne fait rien à tendance à regarder les mouches voler cela se caractérise par un mouvement pseudo aléatoire du regard. Le mouvement du regard possède trois modes :

- ◆ **BEZIER** : Mouvement pseudo aléatoire dont le point d'arrivée est totalement

hasardeux.

- ◆ **NEUTRAL** : Mouvement pseudo aléatoire dont le point d'arrivée est le centre de l'œil.
- ◆ **BREAK** : Pause dans le mouvement, permettant de fixer le regard.

Ainsi le mouvement du regard est une suite infinie dont chaque membre est un des trois modes précédents choisis aléatoirement.

Il y a neuf paramètres pour le mouvement du regard :

- ◆ On définit la durée **minimum** et **maximum** de ces trois modes
- ◆ On définit également la **fréquence d'utilisation** de chacun des modes.

Le mouvement pseudo aléatoire utilise une courbe de Bézier de degré 3 :

$$f(t) = P_0(1-t)^3 + 3P_1t(1-t)^2 + 3P_2t^2(1-t) + P_3t^3$$

La courbe commence à P_0 et se termine en P_3 en passant près de P_1 et P_2 , cela permet d'obtenir un grand choix de mouvements possibles. L'application prend la position actuelle du centre de la pupille comme point de départ et choisi trois points aléatoires qui appartiennent à l'espace des points possibles définis par la surface de la cornée graphique.

On retire de l'espace des points possibles une bande de surface qui suit la bordure et dont la largeur est égale au rayon de la pupille *π. De cette manière on évite que la pupille ne déborde du contour de l'œil.

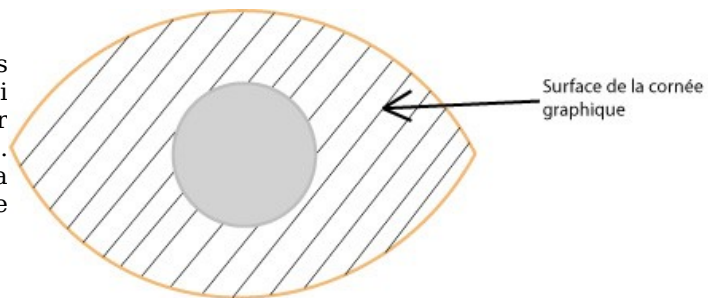


Illustration 25: Surface de la cornée graphique

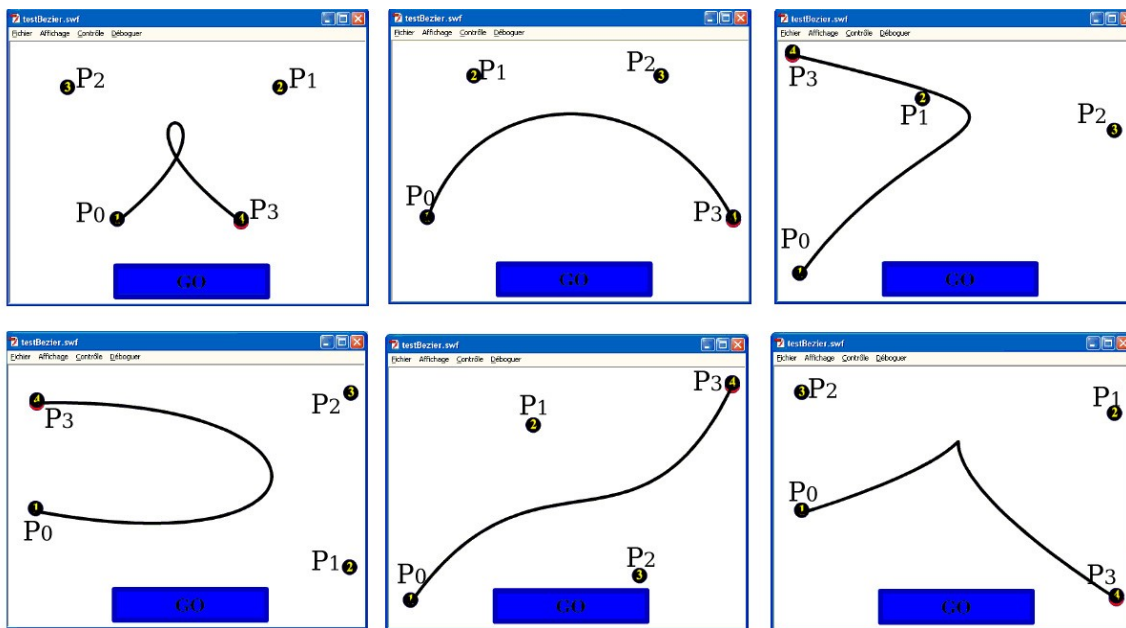


Illustration 26: Exemple de trajectoire possible avec une courbe de Bézier de degré 3

Cette illustration montre qu'en 6 placements différents de points, on obtient des trajectoires déjà complètement différentes. Cette fonction est également utilisée pour le retour au neutre sauf que seulement deux points sont choisis aléatoirement puisque P_3 est égale au point origine.

On transmet au robot les points utilisées par la courbe et celui-ci va utiliser la direction empruntée par la courbe pour en faire des mouvements de la tête. Ça donnera également l'impression que le robot regarde ailleurs.

Remarque : le watching grass grow n'est lancé que pendant le waiting mode.

4.1.4. Les rictus

L'outil **rictus** fonctionne sur le même principe que le mouvement du regard. Il existe trois actions possibles :

- ◆ **RICTUS** : représente des petites variations de positions des sourcils et de la bouche, qui s'assimile à des rictus.
- ◆ **NEUTRAL** : retour à une expression faciale neutre.
- ◆ **BREAK** : Aucune variation d'émotion, pause dans le mouvement

Chacun de ces actions est également gérées par trois paramètres :

- ◆ une durée **minimum** et **maximum** pour chaque action,
- ◆ un **pourcentage d'utilisation** de chacune des actions.

L'outil **rictus** concerne donc les sourcils et la bouche. Elle fonctionne selon le principe suivant :

1. Choix d'un point aléatoire du composant parmi les points autorisés à être bougés.
2. Choix d'une direction en x ou en y, la direction est égale
 - ◆ à -1 si on veut effectuer une variation dans le sens inverse de l'axe,
 - ◆ 1 si la variation se fait dans le sens de l'axe et
 - ◆ 0 si aucune variation n'est demandée.
3. Choix d'une valeur de variation en x et en y choisie aléatoirement parmi un minimum et un maximum fixé, ça peut être un nombre de pixel ou un pourcentage d'avancement entre 0 et 1.
4. On calcule la variation en x et en y : variation = direction * valeur
5. Puis on teste indépendamment pour x et pour y
 - ◆ si le point actuel + la variation dépasse la borne maximum ou minimum
 - ◆ alors on soustrait la variation au point actuel pour qu'il se trouve dans la direction opposé.
 - ◆ Dans le cas où dans une direction ou une autre le point futur dépasse les bornes, alors la variation ne se fait pas.

Concernant les sourcils, le seul point actuellement autorisé à bouger est t (voir illustration 1). Concernant la bouche, on autorise les quatre points de contrôle à bouger en x et en y.

Ces variations sont transmis au robot qui va pouvoir les adapter à ses sourcils et à sa bouche.

Remarque : Art-e n'a des rictus que pendant le waiting mode.

4.2. Passage d'une émotion à une autre

Lorsqu'Art-e doit exprimer une émotion, il passe de son expression faciale actuelle à une expression faciale caractérisant l'émotion à exprimer. Le passage d'une émotion à une autre peut se faire de différentes façons afin de mettre en évidence différents comportements. La suite de ce paragraphe décrira ces types d'affichages ainsi que la réaction émotionnelle des composants qui est différentes pour chacun d'eux. Ce paragraphe se terminera par une description des étapes de la réaction émotionnelle globale d'Art-e.

4.2.1. Types d'affichages des émotions

Lorsqu'Art-e doit exprimer une émotion, il doit passer de l'expression faciale actuelle à une expression faciale caractérisant l'émotion à exprimer. Pour passer de l'une à l'autre, l'application utilise une interpolation linéaire sur chaque point de chaque figure :

$$P_t = P_{origin}(1-t) + P_{destination}t$$

- ◆ P_t est le point interpolé.
- ◆ P_{origin} est le point de départ qui sera interpolé.
- ◆ $P_{destination}$ est le point d'arrivée de l'interpolation.
- ◆ t est le pourcentage d'interpolation effectué pour passer de P_{origin} à $P_{destination}$.

Chaque interpolation suit cette fonction, mais on applique une autre fonction sur t si bien que l'on peut avoir des interpolations plus complexes qu'une interpolation linéaire.

Lorsqu'on lance une interpolation, on indique :

- ◆ **nbImage** : la durée de l'interpolation en nombre d'images par secondes,
- ◆ l'expression faciale de départ et d'arrivée,
- ◆ le type d'interpolation,
- ◆ l'intensité de l'émotion.

Les paramètres pour calculer les points interpolés à chaque image sont :

- ◆ **step** : le nombre d'images passées depuis le début de l'interpolation,
- ◆ **P1** : le point de départ,
- ◆ **P2** : le point d'arrivée.

Les types d'interpolations sont **linear**, **exp**, **log**, **break_linear**, **break_exp** et **break_log**.

L'interpolation **linear** renvoie la valeur $step/nbImage$. C'est une interpolation linéaire classique.

L'interpolation **exp** agit sur l'intervalle $[-3 ; 2]$ car c'est l'intervalle de l'exponentielle qui évolue le plus puis calcule le nouveau t à chaque image, puis appelle l'interpolation linéaire avec ce t .

```
chgt_repere := -3;
pas_de_val := 5/nbImage;
val_max := nbImage/pas_de_val;
t := exp(step*pas_de_val+chgt_repere) / exp(val_max+chgt_repere);
```

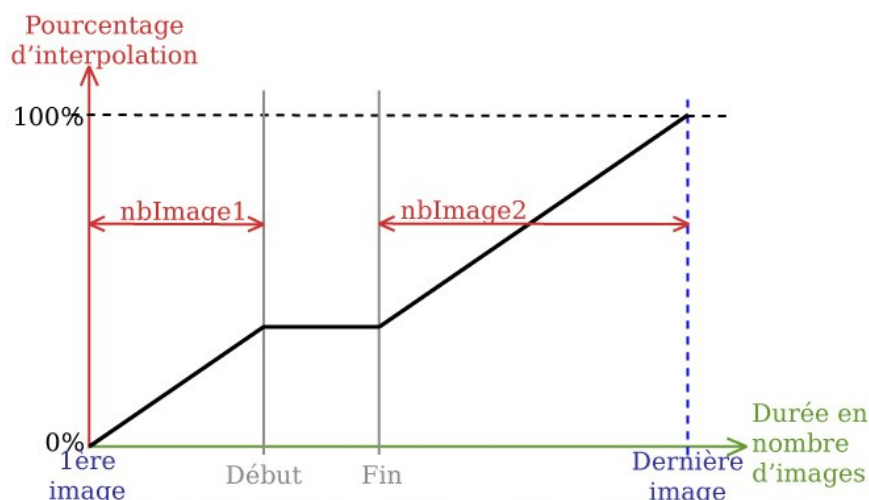
Texte 2: Interpolation exponentielle de l'image n°step sur nbImage

L'interpolation **log** calcule le nouveau t à chaque image, puis appelle l'interpolation linéaire avec ce t .

```
t := log(step) / log(nbImage);
```

Texte 3: Interpolation logarithmique de l'image n°step sur nbImage

L'interpolation **break_linear** simule une hésitation sur la base de l'interpolation **linear**.



Durée réelle de l'interpolation : $nbImage1 + nbImage2$

Illustration 27: Simulation de l'hésitation

Le principe est le suivant :

- ◆ Choix aléatoire du début de la pause
- ◆ Choix aléatoire de la fin de la pause
- ◆ Calcul d'une nouvelle durée d'interpolation qui exclue la pause puisque pendant ce temps aucune progression de l'interpolation n'est effectuée.

L'interpolation lancée sera la même qu'une interpolation linéaire sauf que pendant la pause, l'image sera figée.

```

A l'initialisation :
breakPointDeparture := nombre au hasard entre 0 et nbImage - 10;
breakPointArrival := nombre au hasard > breakPointDeparture et nbImage -5;
newNbImage := nbImage - breakPointArrival + breakPointDeparture;

A chaque image n°step:
si step < breakPointDeparture
  alors t := step/newNbImage;
sinon
  si step >= breakPointDeparture && step < breakPointArrival
    alors t = breakPointDeparture /newNbImage;
  sinon
    si step >= breakPointArrival
      alors newStep := step - breakPointArrival + breakPointDeparture;
         t := newStep / newNbImage;
    fin si
  fin si
fin si

```

Texte 4: Interpolation break linear de l'image n°step sur nbImage

L'interpolation **break exponentielle** simule une hésitation sur la base de l'interpolation **exponentielle**. Le principe est le même que pour l'interpolation `break_linear`. L'algorithme est le même sauf pour le calcul de `t`. Les différences sont notées en gras.

```

Même initialisation que pour break_linear

A chaque image n°step:

chgt_repere := -3;
pas_de_val := 5/nbImage;
val_max := nbImage/pas_de_val;
val_i := i*pas_de_val;
t := exp(step*pas_de_val+chgt_repere) / exp(val_max+chgt_repere);

si step < breakPointDeparture
  alors t := exp(val_i+chgt_repere)/exp(val_max+chgt_repere);
sinon
  si step >= breakPointDeparture && step < breakPointArrival
    alors val_i = breakPointDeparture*pas_de_val;
       t = exp(val_i+chgt_repere) /exp(val_max+chgt_repere);
  sinon
    si step >= breakPointArrival
      alors newStep := step - breakPointArrival + breakPointDeparture;
         val_i = newStep*pas_de_val;
         t := exp(val_i+chgt_repere) / exp(val_max+chgt_repere);
    fin si
  fin si
fin si

```

Texte 5: Interpolation break exponentielle de l'image n°step sur nbImage

L'interpolation **break_logarithmique** simule une hésitation sur la base de l'interpolation **logarithmique**. Le principe est le même que pour l'interpolation `break_linear`. L'algorithme est le même sauf pour le calcul de t . Les différences sont notées en gras.

```
Même initialisation que pour break_linear

A chaque image n°step:
si step < breakPointDeparture
  alors t := log(step)/log(newNbImage) ;
sinon
  si step >= breakPointDeparture && step < breakPointArrival
    alors t = log(breakPointDeparture) /log(newNbImage) ;
  sinon
    si step >= breakPointArrival
      alors newStep := step - breakPointArrival + breakPointDeparture;
         t := log(newStep) / log(newNbImage) ;
    fin si
  fin si
fin si
```

Texte 6: Interpolation break logarithmique de l'image n°step sur nbImage

Ces six types d'interpolation permettent de donner des comportements différents à Art-e, on peut ainsi imaginer un avatar qui réagisse très vite aux émotions, ou au contraire qui met du temps à les laisser paraître etc.

4.2.2. Réaction émotionnelle des composants

Chaque composant (sourcil, bouche, œil, tête et tronc) est capable d'exprimer à une émotion grâce à un ou plusieurs comportements pré-enregistrés (dans un fichier xml).

- ◆ **Composants sourcil et bouche** : lors d'une expérience émotive, Art-e choisit aléatoirement parmi toutes les figures possibles d'une même émotion (expressions faciales du chapitre 4.5).
- ◆ **Composant œil** : à chaque émotion est associée une ou plusieurs tailles d'œil différentes. Art-e choisit aléatoirement parmi les tailles proposées.
- ◆ **Composant tête** : à chaque émotion est associée un ou plusieurs triplet composés comme suit :
 - ◆ une valeur de translation en x,
 - ◆ une valeur de translation en y,
 - ◆ une valeur de rotation.
- ◆ **Composant tronc** : à chaque émotion on choisit un duo de valeurs composés comme suit :
 - ◆ une nouvelle valeur d'oméga ou la même.
 - ◆ Une nouvelle valeur de a ou la même.

Pour changer oméga ou a au cours du temps, il est nécessaire de résoudre le système suivant :

Soit f_1 la courbe actuelle et f_2 la courbe avec les nouvelles valeurs.
 Soit a_1 et ω_1 les valeurs actuelles et a_2 et ω_2 les nouvelles valeurs.
 On sait que $f_1(t_1) = a_1 \sin(\omega_1 t_1 + \varphi_1)$ et $f_2(t_2) = a_2 \sin(\omega_2 t_2 + \varphi_2)$.
 On sait aussi que $f_1(t_1) = f_2(t_2)$ et $t_1 = t_2 = t$.
 Ce qui signifie que $f_1(t) = a_2 \sin(\omega_2 t + \varphi_2)$ et $\frac{f_1(t)}{a_2} = \sin(\omega_2 t + \varphi_2)$

Comme $\sin(u) = \sin(\pi - u)$

$$\text{alors } \arcsin\left(\frac{f_1(t)}{a_2}\right) = \omega_2 t + \varphi_2$$

$$\pi - \omega_2 t - \varphi_2$$

$$\text{On en déduit } \varphi_2 = \arcsin\left(\frac{f_1(t)}{a_2}\right) - \omega_2 t$$

$$\pi - \omega_2 t - \arcsin\left(\frac{f_1(t)}{a_2}\right)$$

On calcule maintenant les coordonnées de deux sommets afin de tester les cas.

$$\min Courbe_2 = a_2 \sin\left(\frac{3\pi}{2}\right) \quad \text{et} \quad \max Courbe_2 = a_2 \sin\left(\frac{\pi}{2}\right)$$

Si $a_2 < a_1$ et si ($f_1(t) < \min Courbe_2$ ou $f_1(t) > \max Courbe_2$)

alors pas de solution car le point de jonction avec la courbe f2 se trouve au dessus du maximum de f2.

on attend que le temps passe jusqu'à ce que la solution existe.

Sinon

$$\omega = \omega_2$$

$$a = a_2$$

si $a_2 < a_1$ et ($f_1(t) \geq \min Courbe_2$ et $f_1(t) \leq \max Courbe_2$)

alors

si $\sin'(\omega t + \varphi) > 0$ dérivée positive, courbe croissante

$$\text{alors } \varphi = \arcsin\left(\frac{f_1(t)}{a}\right) - \omega t$$

$$\text{sinon } \varphi = \pi - \arcsin\left(\frac{f_1(t)}{a}\right) - \omega t$$

fin si

sinon si $a_2 = a_1$ ne rien faire...

sinon si $a_2 > a_1$

si $\sin'(\omega t + \varphi) > 0$ dérivée positive, courbe croissante

$$\text{alors } \varphi = \arcsin\left(\frac{f_1(t)}{a}\right) - \omega t$$

$$\text{sinon } \varphi = \pi - \arcsin\left(\frac{f_1(t)}{a}\right) - \omega t$$

fin si

fin si

fin si

Texte 7: Changement d'oméga et de a au cours du temps

Cet algorithme permet donc de modifier l'amplitude et la fréquence de la respiration sans

qu'il n'y ait de sursaut dans la courbe.

4.2.3. Étapes d'une réaction émotionnelle d'Art-e

Lorsque l'avatar exprime une émotion, il y a plusieurs étapes qui entrent en jeu :

- ◆ Étape 1 : **L'expression de l'émotion** :
 - ◆ Art-e reçoit l'ordre d'exprimer une émotion pendant une durée nbImage.
 - ◆ Art-e choisit un type d'affichage parmi ceux qu'il peut choisir (cf 4.2.1)
 - ◆ Chaque composant choisit une réaction émotionnelle parmi les réactions possibles.
 - ◆ Art-e exprime l'émotion pendant la durée donnée en paramètres.
- ◆ Étape 2 : **Pause dans l'expression** : L'expression est figée pendant un temps calculé préalablement en utilisant :
 - ◆ *min_duration_fixed_emotion* : temps minimum pendant lequel Art-e doit figer l'émotion après une réaction émotionnelle.
 - ◆ *max_duration_fixed_emotion* : temps maximum pendant lequel Art-e doit figer l'émotion après une réaction émotionnelle.
- ◆ Étape 3 : **Retour au neutre** : L'expression retourne au neutre. C'est une interpolation dont le type est également choisi parmi les types possibles. Le temps de retour au neutre est choisi parmi :
 - ◆ *min_duration_return_neutral* : temps minimum pendant lequel Art-e retourne à une expression neutre.
 - ◆ *max_duration_return_neutral* : temps maximum pendant lequel Art-e retourne à une expression neutre.

4.3. Commandes de contrôle

Art-e est un avatar qui attend des instructions à exécuter. Le format des instructions qu'il attend est un quadruplet de quadruplet :

- ◆ Sourcil[émotion, durée, type d'interpolation, intensité]
- ◆ Bouche[émotion, durée, type d'interpolation, intensité]
- ◆ Œil[émotion, durée, type d'interpolation, intensité]
- ◆ Tête[émotion, durée, type d'interpolation, intensité]

Cela signifie qu'on envoie quatre informations à chacun des composants émotionnels de l'avatar.

- ◆ L'émotion est à choisir parmi la joie, la tristesse, le dégoût, la colère, la peur, la surprise et l'émotion neutre.
- ◆ La durée concerne le temps que va mettre la figure à passer de l'émotion actuelle à la nouvelle.
- ◆ Le type d'interpolation indique la façon dont le changement d'émotion va s'effectuer et est à choisir parmi : linear, exp, log, break_linear, break_exp et break_log.
- ◆ L'intensité est un facteur d'expression, qui est une valeur comprise entre 0 et 1, on peut dire que l'intensité est un pourcentage d'expression. Par exemple, si l'intensité est égale à 1 alors Art-e exprimera les émotions pleinement, tandis que si l'intensité est à 0.2, cela donnera l'impression d'une grande réserve, puisqu'il n'exprimera que des petites émotions, beaucoup moins visibles.

5. Particularisation d'Art-e

L'objectif initial du travail est de créer plusieurs instances d'Art-e qui évoluent différemment. Cette évolution est caractérisée par

- ◆ un caractère,
- ◆ une personnalité.

La suite de ce paragraphe définira ce que l'on entend par caractère et personnalité et

expliquera les paramètres sur lesquels jouer pour différencier plusieurs instances d'Art-e.

5.1. Caractère

On considère que le caractère représente *la façon d'exprimer ses émotions*. Une personne peut laisser passer un temps très court entre la sensation et l'expression d'une émotion, alors qu'une autre peut exprimer plus lentement l'émotion.

Les six types d'interpolation vu précédemment jouent un rôle dans l'expression de ce caractère. En effet, ces interpolations sont six manières différentes d'afficher les émotions.

- ◆ Affichage linéaire : Art-e exprime ces émotions, c'est la réaction normale, sans plus.
- ◆ Affichage exponentiel : Art-e doit exprimer une émotion. Ce n'est d'abord pas réellement visible sur son visage (courbe faiblement croissante). Puis l'émotion apparaît de plus en plus rapidement. Il est espéré que cela représente une personne qui ne réagit pas tout de suite à l'évènement.
- ◆ Affichage logarithmique : lorsqu'Art-e doit exprimer une émotion, il commence très très rapidement, puis ralenti petit à petit. Cela peut donner l'impression d'une personne qui réagit très très rapidement, puis se ravise.
- ◆ Chacun des ses trois types d'affichage est associé à une fonction qui effectue le même calcul mais en plaçant quelque part une pause. Cela permet de simuler l'hésitation dans les trois cas.

En définissant le pourcentage d'apparition de ces six types d'affichages pour chaque instance d'Art-e, on l'oblige à choisir certains types. Il pourra ainsi exprimer les émotions selon ce pourcentage d'apparition et cela permet de voir émerger des caractères.

Les instructions à envoyés à Art-e sont simplifiées. Elles deviennent un quadruplet de triplet où l'on ne précise plus le type d'interpolation puisqu'Art-e la choisit automatiquement.

En plus du type d'interpolation, pour faire émerger un caractère, on prévoit une différenciation concernant :

- ◆ Tailles de l'œil : en changeant la taille du **small_eye**, **normal_eye** et **big_eye**, on modifie l'expression des émotions. Si **big_eye** est très proche de **normal_eye**, alors l'expression de la surprise sera moins marquée.
- ◆ Clignement de l'œil : en changeant les durées minimums et maximums pour chacune des quatre étapes, on crée un avatar plus ou moins stressé, plus ou moins fatigué...
- ◆ La respiration : on peut changer la durée de la respiration pour un avatar plus ou moins calme. On peut également changer la valeur de oméga et de a en fonction des émotions.

5.2. Personnalité

On considère que la personnalité représente *la façon de ressentir ses émotions*. Elle n'a pas encore été totalement étudiée pour Art-e.

Cette partie doit reprendre le travail de Thi Hai Ha Dang [30], excepté que l'on a Art-e à la place des cercles de couleurs. Les paramètres d'urgence, de danger et d'affection devront être complétés pour pouvoir générer les six émotions de base et il faut également créer l'interpréteur de ces paramètres – qui effectue des calculs pour trouver l'émotion associée.

L'interpréteur est indépendant de Art-e, c'est une sur couche qui prend en entrée des données sur l'environnement et sort les informations dont a besoin Art-e pour exprimer une émotion.

La personnalité se base sur le MBTI. On rappelle que le MBTI permet de générer 16 personnalités différentes, basées sur le choix d'une préférence dans chacune des 4 dimensions (**energy**, **information**, **decisions** et **lifestyle**.)

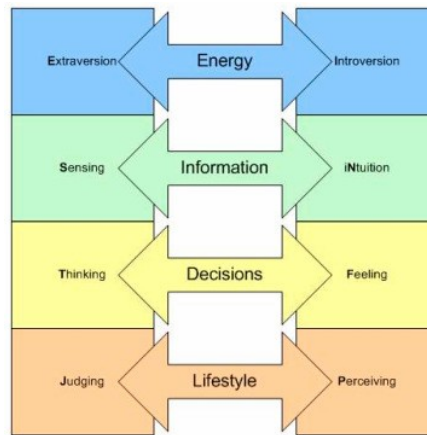


Illustration 28: Les 8 "préférences" du MBTI

Art-e peut exprimer ses émotions avec extraversion ou introversion. Pour chaque émotion + l'expression neutre on associe :

- ◆ un **minIntensity** qui est une valeur entre 0 et 1 et qui représente l'intensité minimum de l'émotion et
- ◆ un **maxIntensity** qui représente l'intensité maximum de l'émotion.

En bornant ainsi l'intensité de la réaction émotionnelle d'Art-e, on donne l'impression d'un Art-e plus ou moins introverti.

6. Évaluation et premiers tests

6.1. Évaluation

Pour évaluer l'avatar, on propose de placer 3 instances d'Art-e ayant un caractère différent à l'écran et de les faire regarder une vidéo. On évalue leur caractère, ils réagissent donc aux mêmes émotions mais de façons différentes.

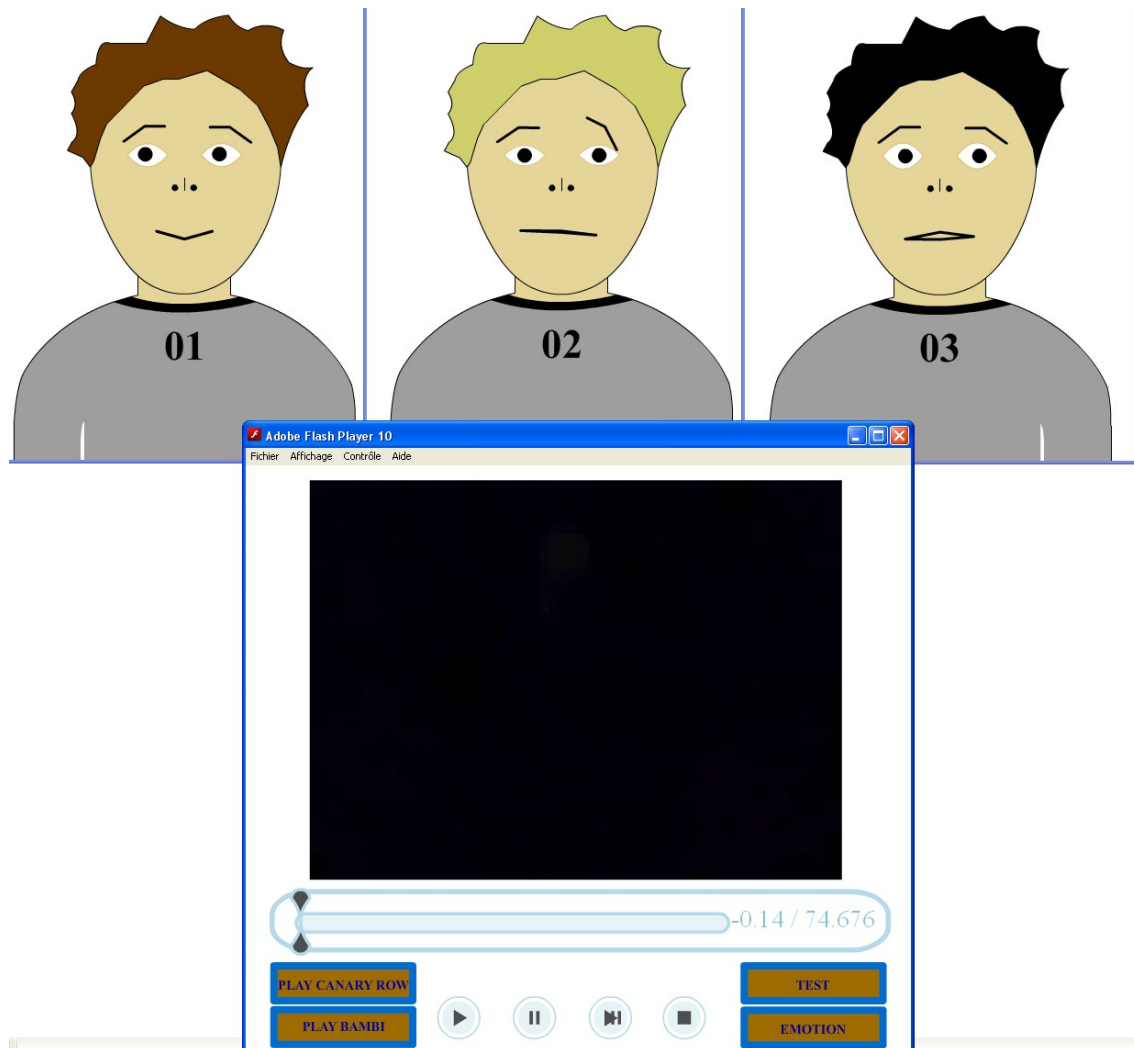


Illustration 29: Trois instances d'Art-e en **waiting mode**

Afin de pouvoir valider la dynamique des émotions, il est nécessaire d'évaluer l'avatar et de poser les questions suivantes, dans un premier temps :

- ◆ Voyez-vous une différence entre les trois Art-e ?
- ◆ Quel mot représente le mieux Art-e 01 ?
- ◆ Quel mot représente le mieux Art-e 02 ?
- ◆ Quel mot représente le mieux Art-e 03 ?

Dans une deuxième phase, on souhaite évaluer la réponse émotionnelle des trois instances. Leur réaction est-elle crédible ? Quelle réaction ont-ils eu ?

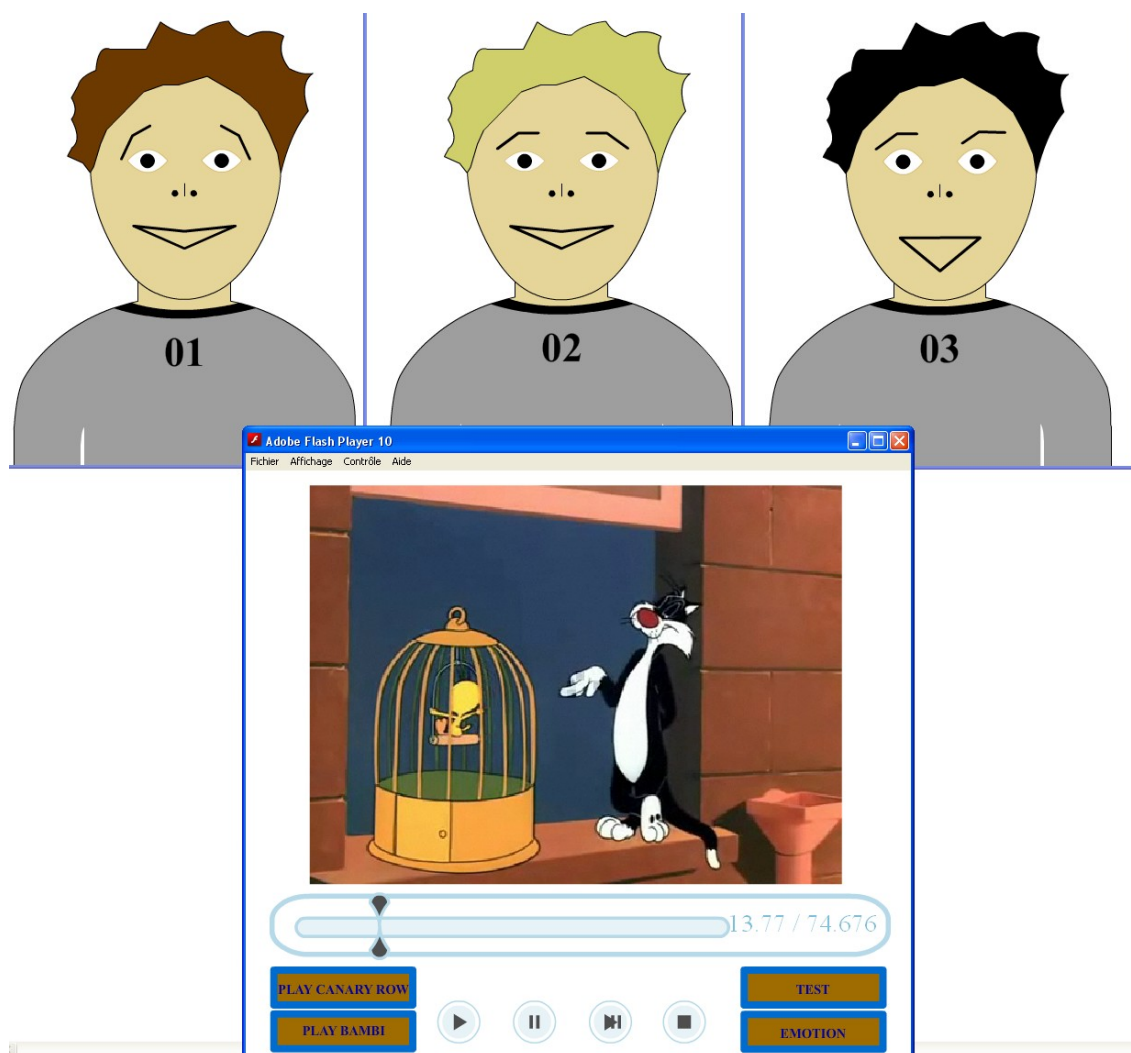


Illustration 30: Trois instance d'Art-e en **listening mode (+emotion mode)**

6.2. Premier test

Les premiers tests effectués sur une dizaine de personnes ne sont pas concluants. Les différences que l'on souhaite faire percevoir ne sont pas assez prononcées. Les personnes interrogées ont vu des différences :

- ◆ « Les sourcils n'ont pas la même position » ou « La bouche n'est pas pareil », ce qui est normal puisqu'ils expriment la même émotion, mais en choisissant une expression faciale au hasard.
- ◆ « Ils ne regardent pas aux mêmes endroits », c'est normal aussi puisque les yeux effectuent des mouvements pseudo aléatoires.
- ◆ Des bugs d'affichage ont aussi empêché des personnes de voir des résultats.

6.3. Deuxième test

En ce moment, nous sommes en train de revoir les paramètres car une petite différence peut avoir de gros effet.

Le deuxième test est en cours.

7. Conclusion

Le travail consistait à créer un avatar capable d'exprimer des émotions. Les émotions se basent sur le travail préalablement validé sur la bouche et les sourcils. On rajoute des yeux et un tronc pour tenter de différencier les émotions. En effet, il résulte du travail précédent que :

- ◆ Il y a beaucoup d'expressions faciales associées à une émotion, et chaque émotion ne possède pas forcément le même nombre d'expression. L'expression neutre par exemple ne possède que 4 figures différentes, contre une dizaine pour les autres.
- ◆ Dans deux groupes différents, on peut trouver la même expression faciale, ce qui empêche l'interprétation de l'émotion. D'où le besoin de rajouter d'autres composants émotionnels.

Le travail d'expression devait être complété par un travail sur la dynamique des émotions. Le but étant de rendre l'avatar « vivant ». Le travail effectué doit être utilisé par Emi (le robot compagnon) pour lui donner un comportement.

La deuxième partie du travail consistait à développer un avatar capable de se comporter de façon unique. Cela signifie que si l'on dispose trois avatars sur l'écran, ils doivent évoluer différemment.

En travaillant sur la personnalité, nous avons découvert au cours du stage qu'un caractère différent permettait d'avoir des avatars qui se comportaient différemment.

7.1. Travaux effectués

Dans un premier temps, un avatar statique a été créé. C'est une sorte de boîte à outil qui exprime des émotions quand on le lui commande. Il est également capable d'afficher le scénario d'émotions commandé par l'interface d'Emi.

L'avatar est paramétrable et tous les paramètres sont stockés dans un fichier xml afin de proposer une solution simple et internationale pour les modifier.

Dans un second temps, on a ajouté le type d'affichage des émotions. Cela signifie que lorsqu'on ordonne à Art-e d'exprimer une émotion, une interpolation est effectuée entre l'expression faciale courante et l'expression faciale représentant l'émotion. Trois principaux types d'affichage émergent du travail effectué :

- ◆ L'affichage linéaire qui représente une interpolation linéaire classique.
- ◆ L'affichage exponentielle qui pourrait représenter un Art-e assez nerveux, qui réagit après un petit temps de réaction, mais très rapidement.
- ◆ L'affichage logarithmique qui pourrait représenter un Art-e assez mou, il réagit tout de suite, mais de plus en plus lentement.

Pour chacun de ces trois types d'affichages, Art-e peut simuler l'hésitation.

Dans un troisième temps, Art-e s'est vu doté de 4 outils afin de le rendre « vivant ».

- ◆ La respiration qui se base sur une sinusoïde, mais qui est légèrement modifiée car l'inspiration dure deux fois moins longtemps que l'expiration. A chaque fois qu'Art-e exprime une émotion, cela a un impact sur la respiration (données stockés dans un fichier xml).
- ◆ Le clignement des yeux qui respectent les quatre phases : descente de la paupière, pause, remontée de la paupière pause. Chaque instance d'Art-e cligne des yeux à sa façon selon des paramètres stockés dans un fichier xml. Lorsqu'il exprime une émotion, Art-e ne cligne pas des yeux.
- ◆ Mouvement aléatoire du regard, Art-e peut regarder les mouches voler. Ce mouvement suit la trajectoire d'une courbe de Bézier de degré 3. Le mouvement du regard est une suite aléatoire et répétitive de trois actions : le regard retourne au regard neutre (Art-e regarde droit devant), le regard est fixé quelque part, le regard suit un mouvement pseudo aléatoire.
- ◆ Les rictus se basent sur des petites variations effectuées au niveau des sourcils et de la bouche. Cette action est aussi une suite aléatoire de trois actions : retour à une expression faciale neutre, pause dans l'affichage de l'expression et petit rictus.

Dans un dernier temps, on donne la possibilité à Art-e de choisir comment il

affichera son émotion et avec quel intensité. On stocke dans un fichier xml, les valeurs relatives.

Le fait de pouvoir faire afficher les émotions de trois façons différentes sur trois Art-e différents permet de simuler un caractère différent puisqu'ils expriment les émotions différemment.

7.2. Travaux restants et futurs

Avant toutes choses, le travail présent doit être validé en rendant la différence entre les instances d'Art-e plus pertinente.

Un des objectifs de la suite de ce travail est de classer les expressions faciales associées à une émotion afin de :

- ◆ leur donner un ordre d'importance. La sélection de l'expression faciale serait donc un choix aléatoire mais pondéré, favorisant la sélection d'expression classées comme plus importante.
- ◆

Il reste à mettre en place la personnalité. Pour se faire, l'idée est de partir du travail de Thi Hai Ha Dang [2] et de calculer l'émotion à exprimer en fonction de certains paramètres. Elle a choisi d'utiliser le danger, l'urgence et l'affection et d'en générer la peur, la joie et la tristesse qu'elle exprime sous forme de cercles de couleur. En poursuivant dans son optique, on souhaite rajouter les trois autres émotions primaires (le dégoût, la surprise et la colère). En rajoutant un paramètre d'imprévu, on peut atteindre ces émotions qui ont en commun le fait d'apparaître lors d'un événement non souhaité.

Afin de représenter le modèle émotionnel du projet, il est prévu de rajouter l'état de l'avatar (humeur et sentiment) qui a une influence sur la perception cognitive de l'évènement, ce qui permet à deux instances différentes d'interpréter différemment l'évènement et de ressentir différemment l'émotion

Dans un avenir moins proche, il est souhaitable de modifier l'avatar actuel, afin de lui donner d'avantage de degré de liberté et plus de réalisme.

Le monde 2D de l'avatar est une grosse contrainte car certains mouvements importants ne peuvent pas être représentés, par exemple baisser la tête ou simuler le travail de la cage thoracique dans le mécanisme de la respiration, d'où ce besoin de passer en 3D.

Pour améliorer son expressivité, on peut lui rajouter des bras, puisque des études ont montré que les émotions étaient plus facilement reconnus lorsque le corps participait à l'expérience émotionnelle, plutôt que le visage seulement.

8. Références

- [site] Site du projet Emotirob, <http://www-valoria.univ-ubs.fr/emotirob/index.html>
- [0] T.-H.-H. D., S. Letellier-Zarshenas, D. Duhaut. *Comparison of recent architectures of emotions*. In 10th International Conference on Control, Automation, Robotics and Vision – ICARCV 2008. Hanoi, Vietnam, December 2008
- [1] Dominique Duhaut, *Évaluation Handbook du projet ANR Emotirob*, laboratoire VALORIA Janvier 2008
- [2] Thi Hai Ha Dang, *Etude sur la définition et la modélisation d'émotion*, mémoire de stage de fin d'études, laboratoire VALORIA, Août 2008
- [3] Sébastien Saint-Aimé, *Génération dynamique d'émotions*, rapport de stage, laboratoire VALORIA, Septembre 2006
- [4] Scherer, K.R. & Sangsue, J, *Le système mental en tant que composant de l'émotion*, édition G. Kirouac, 2004.
- [5] Magalie Ochs, Catherine Pelachaud et David Sadek, *Les conditions de déclenchement des émotions d'un agent conversationnel empathique*, Workshop Francophone sur les Agents Conversationnels Animés (WACA), Toulouse, 2006.
- [6] Martin, Abrilian, Devillers, Lamolle, Mancini et Pelachaud, *Du corpus vidéo à l'agent expressif*, Utilisation des différents niveaux de représentation multimodale et émotionnelle, revue d'intelligence artificielle ISSN 0992-499X, 2006.
- [7] Ochs, Niewiadomski, Pelachaud et Sadek, *Expressions intelligentes des émotions*, Revue d'Intelligence Artificielle (RIA), Editions spéciale "Interaction Emotionnelle", 20(4-5) : 607-620, sept. 2006.
- [8] Audrey Rivière et Benoît Godet, *L'affective computing ; rôle adaptatif des émotions dans l'interaction homme-machine*, Travail d'Étude et de Recherche, 2002-2003
- [9] Rosaling W. Picard, Elias Vyzas et Jennifer Healey, *Toward machine emotional intelligence : analysis of affective physiological state*, 2001
- [10] Jocelyn Scheirer et Rosalind W. Picard, *Affective objects*, Janvier 1999
- [11] De Klaus R. Scherer, *What are emotions ? And how can they be measured ?*, 2005
- [12] Wilma A. Bainbridge, Justin Hart, Elizabeth S. Kim et Brian Scassellati, *The effect of presence on human-robot interaction*, 17th IEEE International Symposium on Robot and Human Interactive Communication, Août 2008
- [13] Kohei Ogawa et Tetsuo Ono , *ITACO : Constructing an emotional relationship between human and robot*, 17th IEEE International Symposium on Robot and Human Interactive Communication, Août 2008.
- [14] Shuzhi Sam Ge, Hooman Aghaebrahimi Samani, Yin Hao Janus Ong et Chang Chieh Hang, *Active affective facial analysis for human-robot interaction*, 17th IEEE International Symposium on Robot and Human Interactive Communication, Août 2008
- [15] Sébastien Saint-Aimé, Brigitte Le Pévédic et Dominique Duhaut, *Emotirob : an emotional interaction model* 17th IEEE International Symposium on Robot and Human Interactive Communication, Août 2008
- [16] Minoru Hashimoto, Hiromi Kondo et Yukisama Tamatsu, *Effect of emotional expression to gaze guidance using a face robot*, 17th IEEE International Symposium on Robot and Human Interactive Communication, Août 2008,
- [17] Iolanda Leite, André Pereira, Carlos Martinho et Ana Paiva , *Are emotional robot more fun to play with ?* 17th IEEE International Symposium on Robot and Human Interactive Communication, Août 2008,
- [18] Hsuan-Kuan Huang, Hung-Hsiu Yu, Yao-Jiunn Chen, Yaw-Nan Lee, *Development of an interactive robot with emotional expressions and face detection*, 17th IEEE

- International Symposium on Robot and Human Interactive Communication, Août 2008,
- [19] David Feil-Seifer et Maja J Matarié, *B³IA : A control architecture for autonomous robot-assisted behavior intervention for children with autism spectrum disorders*, 17th IEEE International Symposium on Robot and Human Interactive Communication, Août 2008
- [20] Heeyoung Kim, Sonya S. Kwak et Myungsuk Kim, *Personality design of sociable robots by control of gesture design factors*, 17th IEEE International Symposium on Robot and Human Interactive Communication, Août 2008,
- [21] Kang-hee Lee, Hyun-Sik Shim, Woo-Sup han, Kwang-Choon Kim et Jong-Hwan Kim, *Evolutionary algorithm for a genetic robot's personality*, 17th IEEE International Symposium on Robot and Human Interactive Communication, Août 2008,
- [22] Anja Austermann et Seiji Yamada, « *Good robot* », « *bad robot* » - *Analyzing users' feedback in a human-robot teaching task*, 17th IEEE International Symposium on Robot and Human Interactive Communication, Août 2008,
- [23] TD sur l'intelligence artificielle <http://www-desir.lip6.fr/~beyniera/downloads/IADM1/iad-td6.pdf>, Laboratoire informatique de Paris, mars 2008.
- [24] Site traitant en parti de l'émotion : <http://library.thinkquest.org/25500/index2.htm>, projet étudiants, 1999.
- [25] Laurent Mougey, Contribution à l'étude de l'écoulement de l'air dans les fosses nasales humaines, Thèse soutenue à l'INSA de Lyon, France, Décembre 2000.
- [26] S. Saint-Aimé, B. Le-Pévédic, D. Duhaut. Générer des émotions avec 6 degrés de liberté. In 1ère Conférence Internationale sur l'accessibilité et les systèmes de suppléance aux personnes en situation de handicap, ASSISTH 2007. Toulouse, France, Novembre 2007
- [27] Arjan Egges, Real-time Animation of Interactive Virtual Humans, Thèse soutenue à la faculté des Sciences de Genève, 2006.
- [28] Siegfried Garten, Zur Kenntniss des zeitlichen Ablaufes der Lidschläge (Étude de la durée des clignements des paupières), Pflügers Archiv European Journal of Physiology, Volume 71, Numbers 9-10, Pages 477-491, mai 1898
- [29] Scott A.King, Animating Speech in Games, Motion in Games, First International Workshop, MIG 2008, Edition Springer, page 239, Pays-Bas, Juin 2008.
- [30] Thi-Hai-Ha Dang, Sabine Letellier-Zarshenas, Dominique Duhaut. Grace – Generic Robotic Architecture To Create Emotions. In 11th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines – CLAWAR 2008. Coimbra, Portugal, Septembre 2008
- [31] Diane Perchet, Modélisation in-silico des voies aériennes : reconstruction morphologique et simulation fonctionnelle. Thèse soutenue à l'Université René Descartes – Paris V, France, Novembre 2005.
- [32] Emmanuel Promayon, Modélisation et Simulation de la Respiration, Thèse soutenue à l'université de Grenoble I, France, Novembre 1997.