



HAL
open science

Modélisation des processus émotionnel dans la prise de décision

Karim Mahboub

► **To cite this version:**

Karim Mahboub. Modélisation des processus émotionnel dans la prise de décision. Intelligence artificielle [cs.AI]. Université du Havre, 2011. Français. NNT : . tel-00696675

HAL Id: tel-00696675

<https://theses.hal.science/tel-00696675>

Submitted on 13 May 2012

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

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

UNIVERSITÉ DU HAVRE

PRES Normandie université

École doctorale SPMII

Laboratoire d'informatique, de traitement de
l'information et des systèmes

THÈSE

présentée en vue d'obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DU HAVRE

Discipline : Informatique

Karim MAHBOUB

MODÉLISATION DES PROCESSUS ÉMOTIONNELS DANS LA PRISE DE DÉCISION

Thèse soutenue le 22 novembre 2011 devant le jury composé de :

Rapporteurs :	Cyril FONLUPT	LISIC, Université du Littoral
	Abderrafiâa KOUKAM	SET, UTBM
	Gérard H.E. DUCHAMP	LIPN, Université de Paris 13
Examineurs :	Aladdin AYESH	Université de Montfort, Royaume-Uni
	Sophie BAUDIC	INSERM, Hôpital Ambroise PARÉ
Directeur :	Cyrille BERTELLE	LITIS, Université du Havre
Co-directrices :	Evelyne CLÉMENT	Psy.NCA, Université de Rouen
	Véronique JAY	LITIS, Université du Havre

REMERCIEMENTS

Une thèse est comme une pièce de théâtre : entre les lieux, les personnages évoluant en ces lieux, un scénario finalement assez imprévisible, et un personnage principal qui doit mener l'intrigue jusqu'à son terme, le temps passe et voit ainsi défiler les acteurs qui arrivent, partent, interagissent, le tout au service du déroulement de l'histoire. Je voudrais rendre hommage aux personnes qui ont fait partie de mon monde de ces quelques dernières années. Voici donc cette page de remerciements, qui fera certainement appel à plus de ressources émotionnelles que tout le reste de ce document.

Tout d'abord, je souhaiterais remercier mon directeur de thèse, M. Cyrille BERTELLE dont le véritable calme olympien, associé à sa formidable capacité d'analyse font de lui un directeur de thèse parfaitement adapté à ce genre de thématique. Merci de m'avoir supporté pendant toutes ces années. Non pas que je sois particulièrement ingérable, mais je pense que mon travail a dû déteindre sur moi (surtout vers la fin!) ... un peu comme si l'émotion était finalement une sorte de maladie scientifiquement transmissible! Je ne puis oublier de remercier Véronique JAY pour sa présence absolument indispensable, et aussi pour m'avoir rappelé au moins à 14 reprises que *fondamental* ne prend qu'un seul *m*! Merci infiniment à Evelyne CLÉMENT d'avoir participé à cette aventure et de m'avoir enseigné les bases psychologiques nécessaires à la compréhension de ces *phénomènes* émotionnels (^_^)!

Merci aux membres de mon jury de thèse grâce auxquels ces années de travail auront pu se concrétiser par une soutenance d'une grande intensité émotionnelle. Merci à mes rapporteurs que sont Cyril FONLUPT, Abderrafiâa KOUKAM et Gérard H.E. DUCHAMP d'avoir accepté de partager mon travail et ainsi pu émettre des critiques d'une grande pertinence au sujet d'un thème de recherche relativement difficile et atypique. Merci infiniment à Sophie BAUDIC pour toutes ces conversations « émotionnées ». Merci par ailleurs d'avoir joué le rôle de présidente de jury au pied levé. Je tiens à remercier tout particulièrement Aladdin AYESH, tout d'abord pour avoir fait le trajet d'Angleterre jusqu'au Havre, mais aussi pour la qualité de nos collaborations, et surtout pour son éternelle bonne humeur! *Thank you so much!*

Merci à ma famille, à mes parents qui ont toujours cru en moi (je ne sais pas pourquoi d'ailleurs!), à mes sœurs et à mon frère dont la seule présence suffit à me remonter le moral (c'est magique!). Merci à Ichem pour ses conseils qui

ne manquent pas de « piquant ». Merci à Philippe et Marie-Claire pour leurs encouragements.

Merci à mes amis, proches et moins proches, qui m'ont accompagné dans cette aventure. Merci tout d'abord à Haifa, Djamila et Guilhelm, mes trois compagnons de bureau, pour leur chaleur humaine (en toute amitié bien sûr !). Un grand merci à ceux qui ont clairement contribué au bon environnement hautement scientifique régnant au labo : merci à Butcher, Kamikaze, Martine, Virus, Obama, Djamilator, MasterKekette, Ciao16Coups, Razi3l, Mobutu et Astra pour ces parties de *Urban Terror*, véritable sas de décompression en période de crise ! Merci à toute l'équipe RI2C pour l'excellente atmosphère qui règne au LITIS. Merci à Damien, Stéfan, Antoine, Laurent, Jean-Luc, Nathalie, Claire, Aziz ... Merci aussi à Omar et Rawan pour l'ambiance des premières années.

Merci à tous ceux que j'ai oubliés (qu'ils m'en excusent !). Merci au lecteur qui d'aventure découvrirait ce manuscrit au détour d'un chemin ...

Merci enfin à ma petite fleur que j'aime de tout mon cœur !

TABLE DES MATIÈRES

TABLE DES FIGURES	11
INTRODUCTION	13
I QU'EST-CE QUE L'ÉMOTION ?	19
1 DÉFINITIONS DE L'ÉMOTION	23
1.1 ORIGINE DU TERME	24
1.2 DEUX GRANDES CLASSES DE MODÈLES	27
1.2.1 Modèles hiérarchiques	28
1.2.2 Modèles componentiels	28
1.3 LES DIFFÉRENTES PERSPECTIVES D'ÉTUDE	29
1.3.1 Perspective darwinienne	29
1.3.2 Perspective jamesienne	29
1.3.3 Perspective cognitiviste	30
1.3.4 Perspective socio-constructiviste	30
2 HISTORIQUE DES THÉORIES DE L'ÉMOTION	31
2.1 JAMES-LANGE <i>versus</i> CANNON-BARD	32
2.2 ZAJONC <i>versus</i> LAZARUS-MANDLER	36
2.3 SCHÉMAS DE BASE	39
2.3.1 La théorie de James-Lange	39
2.3.2 La théorie de Cannon-Bard	40
2.3.3 La théorie de Schachter-Singer	40
2.3.4 La théorie de Lazarus	40
2.3.5 Conclusion	41
3 VERS UNE VISION CONTEMPORAINE DES ÉMOTIONS	43
3.1 LES ÉMOTIONS DE BASE	44
3.2 EKMAN ET LES MODÈLES ÉVOLUTIONNISTES	47
3.3 LES THÉORIES COGNITIVISTES	48
3.3.1 La théorie de Schachter et Singer (1962)	49
3.3.2 Le modèle OCC	49
3.3.3 La « roue des émotions » de Robert Plutchik	50

3.3.4	Conclusion	52
3.4	LES THÉORIES DE L' <i>appraisal</i>	53
3.4.1	La théorie de Magda Arnold (1945)	54
3.4.2	La théorie de Frijda	55
3.4.3	Le modèle d'évaluation cognitive de Roseman	55
3.4.4	La théorie de Lazarus	56
3.4.5	La théorie de Scherer	57
3.4.6	Conclusion	58
3.5	LES THÉORIES SOCIO-CONSTRUCTIVISTES	58
4	L'ÉMOTION AU SEIN DES PROCESSUS CÉRÉBRAUX	61
4.1	PHINEAS GAGE ET LE CORTEX PRÉFRONTAL	62
4.2	LE CONCEPT DE CIRCUIT NEURONAL	65
4.2.1	Le circuit de la peur	66
4.2.2	Le circuit du plaisir	67
4.3	CONCLUSION	68
5	ÉMOTIONS ARTIFICIELLES ET MODÈLES INFORMATIQUES	73
5.1	BILAN SUR LA REPRÉSENTATION DE L'ÉMOTION NATURELLE	75
5.2	NÉCESSITÉ DE LA MODÉLISATION DE L'ÉMOTION	76
5.3	MODÈLES INFORMATIQUES DE L'ÉMOTION	78
5.3.1	Le modèle EMA	79
5.3.2	Le <i>CogAff</i> de Aaron Sloman	80
5.3.3	La machine pensante de Alain Cardon	81
5.3.4	Le modèle <i>Cathexis</i> de Velásquez	85
5.4	CONCLUSION SUR L'ÉMOTION ARTIFICIELLE	86
5.4.1	Réflexions sur l'émotion	87
5.4.2	Définition de l'émotion artificielle	87
II LE PROBLÈME DES <i>Cascades</i>		89
6	MODÉLISATION	95
6.1	RÈGLES DU JEU	95
6.2	REPRÉSENTATION DE LA GRILLE DES <i>Cascades</i>	98
6.3	GESTION DES STRATÉGIES	99
6.3.1	Graphe de stratégies	100
6.3.2	<i>Pattern-matching</i> des stratégies	101
6.4	PROCESSUS DE RÉOLUTION	103
6.4.1	Algorithme fourni	104
6.4.2	Sélection des stratégies	105
6.4.3	Algorithme principal	106
7	RÉSULTATS	111

7.1	EXEMPLE DE RÉOLUTION	113
7.1.1	Étape 1	113
7.1.2	Étape 2	116
7.1.3	Étape finale	117
7.2	PLANS D'EXPÉRIENCE	119
7.2.1	Sensibilité aux paramètres	119
7.2.2	Impact de l'émotion	122
8	CONCLUSION ET PERSPECTIVES	127
8.1	CONCLUSION	127
8.2	PERSPECTIVES	128
III	ÉMOTION ET LOGIQUE FLOUE	131
9	REPRÉSENTATION DE L'ENVIRONNEMENT	135
9.1	DESCRIPTION DU MILIEU	135
9.2	ROBOT	137
9.3	IMPLANTATION DU ROBOT	138
10	IMPLANTATION ET RÉSULTATS	145
10.1	MODÈLE ÉMOTIONNEL	145
10.2	DESCRIPTION DE L'IMPLANTATION	147
10.3	EXEMPLE D'EXÉCUTION	151
10.3.1	Étape $t = 0$	151
10.3.2	Étape $t = 15$	151
10.3.3	De l'étape $t = 21$ à l'étape $t = 27$	154
10.3.4	De l'étape $t = 60$ à l'étape $t = 70$	154
10.3.5	De l'étape $t = 170$ à l'étape $t = 190$	154
10.3.6	Étape $t = 260$	155
IV	LE MODÈLE GAEA	157
11	UNE APPROCHE « ÉMERGENTISTE »	161
11.1	L'INTELLIGENCE COMPUTATIONNELLE	162
11.2	LES ALGORITHMES ÉVOLUTIONNISTES	163
11.3	LA PROGRAMMATION GÉNÉTIQUE	165
11.4	LA PROGRAMMATION GÉNÉTIQUE LINÉAIRE	167
12	MODÉLISATION	169
12.1	ENVIRONNEMENT	171
12.1.1	Le robot	171
12.1.2	Les proies	173

12.1.3	Les prédateurs	173
12.2	STRUCTURE D'UN PROGRAMME	173
12.2.1	Les registres mémoire	175
12.2.2	Les instructions du programme	176
12.3	MOTEUR GÉNÉTIQUE	178
12.3.1	Algorithme principal	179
12.3.2	Initialisation	180
12.3.3	Sélection	180
12.3.4	Reproduction	182
12.3.5	Remplacement	185
12.4	AJOUT D'HEURISTIQUES	186
13	IMPLANTATION ET RÉSULTATS	191
13.1	INTERFACE GRAPHIQUE	192
13.1.1	Création d'une simulation	192
13.1.2	Chargement d'une simulation	194
13.1.3	Manipulation d'une simulation	194
13.1.4	Environnement 3D	196
13.2	PREMIERS RÉSULTATS	199
13.3	RÔLE DES HEURISTIQUES	203
14	CONCLUSION ET PERSPECTIVES	205
14.1	CONCLUSION	205
14.2	PERSPECTIVES	206
	CONCLUSION	209
	BIBLIOGRAPHIE	222
	ANNEXES	225
A	LE MODÈLE OCC	225
B	LE MODÈLE DE MILLENSON	227
C	<i>Psychologically Grounded Emotion Model for a NPC Fuzzy Controller</i>	229

TABLE DES FIGURES

3.1	Les <i>Basic Emotions</i>	45
3.2	Le problème de la verbalisation émotionnelle selon SCHERER	47
3.3	Les six <i>basic emotions</i> (émotions basiques) de EKMAN : colère, peur, dégoût, surprise, joie et tristesse	48
3.4	Attributs émotionnels selon le modèle OCC (Version originale en annexe A)	50
3.5	La « roue des émotions » de PLUTCHIK	51
3.6	Principales tendances à l'action ainsi que leurs fonctions et émotions respectives	55
3.7	Modèle de Roseman	56
4.1	Phineas GAGE et sa fameuse barre de fer	63
4.2	Le cerveau de Phineas GAGE	63
4.3	Schémas du cortex préfrontal	69
4.4	Les deux circuits de PAPEZ (Source : http://lecerveau.mcgill.ca)	70
4.5	Localisation des organes associés au circuit de la peur (Source : site internet <i>Anatomography</i> maintenu par le <i>Life Science Databases</i> URL : http://lifesciencedb.jp)	71
4.6	Zones cérébrales associées au système de récompense (Source : http://lecerveau.mcgill.ca)	72
5.1	Principaux modèles informatiques de l'émotion (GRATCH, [54])	75
5.2	Le modèle H-CogAff de Aaron SLOMAN	81
5.3	Le corps de la machine pensante et ses deux systèmes générateurs d'émotions et de pensée ([18])	82
5.4	Structure générale d'un agent aspectuel	83
5.5	Dimensions de l'espace morphologique	84
5.6	Organisation des agents	85
5.7	Les deux robots YUPPY et KISMET	86
5.8	Automate comportemental de la <i>Gambling Task</i>	92
5.9	Modèle d' <i>appraisal-coping</i> adapté de LAZARUS ([69])	93
5.10	Modèle d' <i>appraisal-coping</i> appliqué au problème des <i>Cascades</i>	93
6.1	Grille initiale, première étape de résolution et solution de la grille du problème des <i>Cascades</i>	97

6.2	Élève résolvant le problème des <i>Cascades</i> ([12])	97
6.3	Une case et son voisinage	98
6.4	Exemple de stratégie. Pour trouver le contenu de la case rouge, il faut additionner les valeurs des deux cases noires du dessus	99
6.5	Exemples de stratégies triangulaires	102
6.6	Exemples de stratégies linéaires	102
6.7	Exemple d'application d'une stratégie d'inférence	103
7.1	Fenêtre principale de l'application	112
7.2	Étape initiale	114
7.3	Graphe avant et après la pénalisation et grille de l'étape 1	115
7.4	Évolution du graphe de l'étape 2 avant et après la pénalisation	116
7.5	Exemples de deux résolutions de grilles à l'étape 2	117
7.6	État du graphe et de la grille à l'étape finale	118
7.7	Nombre de grilles résolues (sur 50 grilles) en fonction de Z_n , pour différentes valeurs de Z_e et Q	120
7.8	Nombre moyen d'étapes de résolution en fonction de Z_n , pour différentes valeurs de Z_e et Q	121
7.9	Nombre de grilles résolues (sur 50 grilles) en fonction de Z_n , pour différentes valeurs de Z_e et Q (version cognitive)	124
7.10	Nombre moyen d'étapes de résolution en fonction de Z_n , pour différentes valeurs de Z_e et Q (version cognitive)	125
8.1	Environnement proie-prédateur modélisé	134
9.1	Voisinage d'un individu	136
9.2	Capteurs du robot	138
9.3	Carte de l'environnement	139
9.4	Vue à la troisième personne	140
9.5	Diagramme de classe de l'environnement	141
9.6	Diagramme de classe des capteurs du robot	142
10.1	Modèle tridimensionnel de l'intensité émotionnelle de MILLENSON (Version originale en annexe B)	146
10.2	Boucle principale d'exécution	148
10.3	Exemple de situation locale d'un agent et sa représentation	150
10.4	Évolution des vecteurs \mathbf{J} et \mathbf{B}_e , ainsi que de l'angle de rotation $\delta\Theta$, en fonction du temps	152
10.5	Étape initiale ($t = 0$)	153
10.6	Étape $t = 15$	153
10.7	Étape $t = 21$	153
10.8	Étape $t = 27$	153
10.9	Étape $t = 65$	153

10.10	Étape $t = 184$	153
10.11	Étape $t = 280$	155
11.1	Les sous-catégories de l'intelligence computationnelle	163
11.2	Expression sous forme d'arbre syntaxique	166
12.1	Capteurs de l'individu	172
12.2	Organisation générale du robot	174
12.3	Fonctionnement du saut conditionnel. Le branchement peut s'effectuer avant la condition-objectif (à gauche) ou après (à droite)	178
12.4	Sélection des individus par tournoi	182
12.5	Croisement entre deux individus-programmes	183
12.6	Les mutations	184
12.7	Organisation générale du robot (avec heuristiques)	189
13.1	Fenêtre de création d'une nouvelle simulation – Paramètres du moteur génétique	192
13.2	Fenêtre de création d'une nouvelle simulation – Paramètres des programmes génétiques	193
13.3	Fenêtre principale de GAEA — onglet <i>Population</i>	195
13.4	Fenêtre principale de GAEA — onglet <i>Hall of Fame</i>	196
13.5	Fenêtre principale de GAEA — onglet <i>Individual</i>	197
13.6	Fenêtre de chargement d'un environnement 3D	198
13.7	Environnement de simulation proies-prédateurs en 3D	199
13.8	Évolution de la <i>fitness</i> maximale, moyenne et minimale au sein de la population en fonction de la génération	201
13.9	Évolution de la longueur de programme maximale, moyenne et minimale au sein de la population en fonction de la génération	202
13.10	Évolution de la <i>fitness</i> maximale, minimale et moyenne au sein de la population en fonction de la génération (avec heuristiques)	203
13.11	Évolution de la longueur de programme maximale, minimale et moyenne au sein de la population en fonction de la génération (avec heuristiques)	204
A.1	Attributs émotionnels selon le modèle OCC	225
B.1	Modèle tridimensionnel de l'intensité émotionnelle de MILLENSON	227

INTRODUCTION

Les choses ne sont pas toujours ce qu'elles paraissent. Beaucoup se laissent tromper par les apparences. Rares sont ceux capables de voir derrière le masque.

— Jean de LA FONTAINE

Émotion . . . Par sa simple sonorité, ce mot d'allure onirique évoque pourtant un immense éventail d'impressions, de ressentis, de modes d'expression bien réels, faisant place tantôt aux rires, tantôt aux larmes. Par ailleurs, parler ou écrire l'*émotion*, c'est aussi accéder à un monde où l'humain tente d'interpréter, de concrétiser par les mots ce qu'il ressent, afin de communiquer sur ses impressions. Cette dualité nous ouvre les portes d'un univers profondément humain, subjectif par nature, mais dont l'immensité nous permet de toucher du doigt l'extraordinaire complexité de l'esprit, dont nous allons tenter de percer un des grands mystères.

L'émotion est omniprésente ! La publicité utilise sans cesse ses codes pour vanter les mérites de ses produits, la presse en tire régulièrement parti pour appâter le chaland, le cinéma transporte le spectateur à l'aide d'images représentant invariablement l'amour, la peur ou la stupéfaction. L'être humain tente pourtant depuis longtemps de camoufler son émotion, souvent considérée comme synonyme de faiblesse ou de manque de contrôle de soi par le cartésianisme ambiant. La société inconsciemment friande d'émotion, à l'instar du sportif adepte de sensations fortes, admet néanmoins sa présence au travers de la promotion de l'art, la recherche du bien-être, la lutte contre le stress, ou encore ce qui nous intéresse plus particulièrement ici, l'attention suscitée de nos jours par l'étude scientifique de l'émotion. Ce regain d'intérêt concerne aussi bien les sciences humaines, parmi lesquelles la psychologie ou la sociologie, que la médecine avec la neurologie ou la neurophysiologie.

L'informatique n'est pas en reste, car elle tente depuis plusieurs décennies de concevoir une intelligence artificielle dotée de capacités d'apprentissage, d'analyse et de décision de plus en plus sophistiquées, intégrant de plus en plus profondément une dimension émotionnelle. Et si les processus émotionnels étaient

la pièce manquante dans l'élaboration d'une entité artificielle intelligente et performante? Les récentes avancées en neurologie semblent en tout cas aller en ce sens ([22]). La mise en place de processus émotionnels nous permettrait de prendre des décisions rapides et en adéquation avec la situation lorsque celle-ci réclame une réaction urgente. En revanche, si le contexte ne requiert aucune précipitation, les décisions ont tendance à s'appuyer sur des processus plus rationnels impliquant raisonnement et déduction. De l'éventualité de l'émotion à sa nécessité, il n'y a qu'un pas.

*The question is not whether intelligent machines can have any emotions, but whether machines can be intelligent without emotions.*¹

— Marvin MINSKY

Le but de ce travail de thèse est d'élaborer plusieurs modèles de l'émotion, basés sur les connaissances généralement admises concernant les processus émotionnels, afin de les mettre en place au sein de programmes informatiques. Ces programmes se heurteront à un obstacle particulièrement délicat à surmonter et dont nous tenterons de fournir un éclairage modeste : il s'agit de la question des liens existant entre processus émotionnels et cognitifs. Si au xx^e siècle, cognition rimait avec raison, au xxi^e siècle, cognition rimera plutôt avec émotion, car si d'aucuns considèrent l'émotion et la cognition comme parfaitement distinctes — à la manière du behaviorisme des années 60 —, il est d'usage aujourd'hui de penser les deux activités indissociables l'une de l'autre.

Il ne s'agit certainement pas ici de créer un modèle totalement fonctionnel directement implanté sur un formidable robot doué des plus puissantes facultés humaines. Le phantasme du robot humanoïde, entretenu par l'industrie du cinéma à gros budget, nous fait souvent oublier la nature même de ce qu'est une Machine. Le terme est issu du grec dorien *mekhane* (μηχανή) qui signifie « dispositif ingénieux, moyen ». Il s'agit donc à l'origine d'un procédé mécanique visant à effectuer une ou plusieurs tâches spécifiques. Par ailleurs, le *robot*, terme inventé au début du xx^e siècle et dont l'étymologie d'origine slave est associée notamment à l'esclave (*robotnik* en tchèque) ou au travail pénible (*robot*), désigne plus précisément un dispositif *mécatronique* — c'est-à-dire qui allie la mécanique, l'électronique et l'informatique — qui lui permet d'effectuer des tâches généralement plus risquées ou difficiles. Les robots sont en somme des machines particulièrement évoluées et équipées d'un côté d'une partie mécanique et électronique, et d'un autre côté d'un programme informatique adapté aussi bien à leur structure qu'à leur fonction. Rappelons au passage qu'un programme est une suite d'instructions-machine créées par l'Homme et dont les opérations per-

1. « La question n'est pas de savoir si les machines intelligentes peuvent avoir une émotion, mais de savoir si les machines peuvent être intelligentes sans émotion. » ([92], NDR)

mettent de manipuler dans un certain but une machine, en l'occurrence ici un robot, dans les limites de sa capacité physique et computationnelle.

Cette limitation qui fait le malheur de nombreux informaticiens — bien que pourtant régulièrement repoussée — doit nous sensibiliser sur la nécessité de concevoir l'émotion de manière novatrice, car l'étonnante complexité du substrat neuro-anatomique cérébral, caractéristique de tout être humain biologiquement constitué, ne saurait être comparée à l'inflexibilité d'un ensemble de composants électroniques inertes dont l'activité n'est que le fruit des prévisions d'un programmeur averti. Et la différence est réelle : comment transposer l'émotion dans le domaine du calculable ? Comment concilier la logique d'une approche anthropomorphique, démarche naturelle lorsqu'il s'agit de phénomènes concernant l'Humain, avec le fonctionnement *a priori* déterministe d'une machine pré-programmée ? Dans ce cas, peut-on encore parler d'émotion ? Une distinction doit alors être établie entre l'émotion dite *naturelle* qui anime les êtres vivants, et ce que l'on appelle désormais l'*émotion artificielle*.

Cette dernière est aujourd'hui en plein essor et mobilise la recherche en informatique et en robotique principalement en raison de l'échec des approches cognitives traditionnelles. En effet, les débuts de l'intelligence artificielle ont vu apparaître le *cognitivism*, approche utilisant des mécanismes d'inférence logique à base de règles formelles sur des faits et axiomes. Puis s'est révélée la nécessité d'inclure aux modèles développés l'idée de perception et d'action, permettant l'arrivée des *agents* et autres robots dotés de capteurs et d'effecteurs. Enfin, le comportement social et la notion d'incertitude sont venus compléter la liste des fonctionnalités réputées indispensables à la quête d'une véritable intelligence. Toutefois, le raisonnement pur — et en particulier la logique — est un processus bien trop gourmand en ressources cérébrales pour être suffisamment rapide dans l'immense majorité de nos prises de décision quotidiennes. Comme nous l'avons mentionné précédemment, les découvertes récentes en neurologie ont permis de démontrer l'importance de l'émotion dans la prise de décision chez l'être humain. Par conséquent, l'intégration des processus émotionnels dans de nombreux modèles et théories développés de nos jours suit logiquement cette voie et témoigne de l'intérêt grandissant suscité par la contribution éventuelle de l'émotion.

Cette thèse s'inscrit dans la continuité d'un précédent travail de master ([78] et [79]) qui a vu naître un modèle de l'émotion reprenant la *Gambling Task* ([11]), un jeu d'argent expérimental impliquant de profondes réactions émotionnelles. L'objectif consistait alors en l'élaboration d'une simulation de jeu tenant compte de l'évolution de l'activité émotionnelle de l'individu en situation de prise de décision. Les résultats se révélant encourageants, deux nouveaux modèles ont fait leur apparition et sont décrits ici dans le cadre de ce travail de thèse en informatique.

Le premier, intitulé *problème des Cascades*, vise à simuler la résolution d'un

exercice de mathématique par un enfant d'une dizaine d'année. L'intérêt de cette approche est double. D'une part, les manifestations émotionnelles émanant d'enfants sont souvent moins complexes à expliciter que celles d'adultes, tout en restant riches et instructives. En effet, un enfant a l'avantage de pouvoir mettre en place des stratégies de résolution développées et intuitives, et possède les capacités de recul suffisantes pour pouvoir expliquer son comportement *a posteriori*. D'autre part, le problème à modéliser, à l'instar de la *Gambling Task*, est une activité solitaire, dont la résolution ne subit pas d'autres perturbations que celles générées par le problème lui-même. En ce sens, le travail de modélisation s'en trouve simplifié. Il en résulte un programme informatique mêlant cognition et émotion par l'utilisation d'un graphe décisionnel à marquage émotionnel, et dont la dynamique reposant sur un algorithme fourni décrit le fonctionnement d'un individu prenant sa décision parmi différentes stratégies.

Le second modèle réalisé emprunte une toute autre perspective, et prend le parti qu'une dynamique émotionnelle peut être modélisée par l'émergence d'un programme adaptatif. En effet, la création d'un programme informatique par sélection génétique peut être assimilable à la construction du câblage cérébral responsable notamment des différentes routes de l'émotion que l'on observe chez l'être humain. Par conséquent, l'idée est de simuler le déplacement d'un robot prénommé GAEA, équipé de capteurs et d'effecteurs, et plongé dans un environnement hostile de type proies-prédateurs. Ce robot sera en outre doté d'un programme interne régissant son comportement. L'utilisation d'un algorithme de programmation génétique linéaire sera responsable de la sélection progressive des individus-programmes les mieux adaptés à la survie.

Pour les deux modèles présentés ci-dessus, l'objectif majeur sera de déterminer si la composante émotionnelle modélisée peut se révéler utile et efficace dans le cadre de la prise de décision. La temporalité jouant un rôle critique dans de nombreux cas, en particulier pour GAEA et son environnement proies-prédateurs, il va s'agir de démontrer la contribution éventuelle que peut révéler une approche émotionnelle, plus rapide et plus automatique qu'une approche cognitive, dans la réalisation de modèles informatiques d'aide à la décision.

Ce manuscrit de thèse se compose de quatre grandes parties, dont la première est une forme d'état de l'art pluridisciplinaire autour de l'émotion, les trois autres parties décrivant les trois contributions réalisées durant cette thèse.

Dans une première partie, on présentera les différentes approches de l'émotion en s'attardant sur les différentes catégories conceptuelles de l'émotion, son histoire, ainsi que les modèles en psychologie, neurologie et informatique dont elle a fait l'objet au cours de ces dernières décennies. La seconde partie traitera du *problème des Cascades*, décrit précédemment. Il sera suivi en troisième partie d'un modèle développé en collaboration avec William BLEWITT de l'université de Montfort au Royaume-Uni, qui utilise la logique floue appliquée au modèle émotionnel de MILLENSON, le tout appliqué au système proies-prédateurs de

GAEA, afin de créer un robot doté d'un comportement émotionnel efficace. Enfin, l'ultime quatrième partie sera consacrée au modèle GAEA, ainsi qu'à une description des perspectives que l'on peut proposer dans la continuité de cette approche.

Quoi qu'il en soit, beaucoup d'émotions en perspective !

I

QU'EST-CE QUE L'ÉMOTION ?

La définition de l'émotion est souvent sujette à de nombreuses polémiques, et il n'est pas exagéré de considérer qu'il y a autant de définitions de l'émotion que de scientifiques travaillant sur le sujet. Dès 1981, Paul et Anne KLEINGINNA ([61]) en relevaient plus de 140 définitions. D'une manière générale, les processus émotionnels sont par nature très complexes, et le fait de déterminer une délimitation théorique à la portée même de ce que l'on appelle une émotion est quasiment impensable tant les connaissances physiologiques et comportementales sont incomplètes à ce jour. Ainsi, là où un neurologue par exemple, sera attaché à des notions de facteurs somatiques ou d'activation cérébrale, un sociologue, par ailleurs, aura une vision bien plus globale, et déterminera des valeurs liées à des paramètres sociaux donnés pour étudier par exemple un mouvement de panique de masse ou encore l'anxiété face à une situation donnée. Il n'en reste pas moins que les différentes approches conceptuelles apportent des points de vue complémentaires, qui permettent de mieux cibler la difficulté du problème.

*Everyone knows what an emotion is, until asked to give a definition.*²

— B. FEHR & J.A. RUSSELL ([42])

Dans l'histoire des sciences en général, le passage de l'observation d'un phénomène à sa modélisation s'est la plupart du temps soldé par une sorte de « refonte » complète des connaissances du phénomène, et ce afin de pouvoir mieux l'appréhender dans le domaine du calculable. Ainsi, lorsqu'un géographe décide de se représenter une dynamique de population quelconque dans un environnement urbain par exemple, il est fréquent de constater qu'il va utiliser des outils mathématiques ou informatiques pour modéliser les entités de son système ainsi que les différents liens entre ces entités. Il va donc devoir « adapter » sa vision du phénomène aux outils de modélisation qui lui sont disponibles.

Dans le cas de mécanismes psychologiques, la tâche s'avère beaucoup plus ardue. En effet, le manque d'informations à propos de l'activité à modéliser, en l'occurrence l'émotion, est un premier rempart difficile à surmonter. Quand bien même les avancées majeures dans le domaine des neurosciences nous permettent aujourd'hui d'identifier globalement les zones cérébrales sous-tendant l'activité émotionnelle, les correspondances comportementales entre émotion et décision sont pour ainsi dire méconnues. Si l'on ajoute à cela l'extraordinaire complexité de la dynamique liée à cette mécanique, caractérisée principalement par sa rapidité et son instabilité, on comprend alors qu'il est délicat de le représenter en tenant compte de tous ces éléments. Fort heureusement, les nombreuses recherches scientifiques ont permis à l'Homme de visualiser l'émotion de différentes manières, dans tous les domaines, et par des points de vue variés.

Cette première partie a une fonction double : tout d'abord présenter une ap-

2. « Chacun sait ce qu'est une émotion, jusqu'à ce qu'on lui demande de la définir. » (NDR)

proche chronologique qui décrit les avancées scientifiques concernant l'émotion depuis un peu plus d'un siècle, tout en adoptant une attitude pluridisciplinaire, laquelle s'avère indispensable si l'on souhaite cibler l'objet d'étude avec une précision suffisante.

Dans un premier temps, il s'agira de définir ce que l'on entend par *émotion* à travers les différentes classifications que l'on peut en faire et s'il est possible de généraliser les processus émotionnels en étudiant de quelles façons ils peuvent être abordés. Deuxièmement, et dans une perspective plus chronologique, nous allons aborder les origines de l'étude de l'émotion, qui débute vers la fin du XIX^e siècle jusqu'à nos jours. Dans un troisième temps seront traités quelques modèles parmi les plus représentatifs de ces cinquante dernières années, et dont l'héritage reste encore très présent aujourd'hui. Un quatrième chapitre sera consacré à certains aspects neurologiques, notamment le concept de circuit neural, ainsi qu'une zone cérébrale clef : le cortex préfrontal. Enfin, le dernier chapitre sera l'occasion d'extraire une définition de l'émotion que je qualifierai d'acceptable, afin de traiter plus en profondeur le passage de l'émotion naturelle à la question de l'émotion artificielle et de son utilité au travers de quelques grands modèles informatiques.

DÉFINITIONS DE L'ÉMOTION



SOMMAIRE

1.1	ORIGINE DU TERME	24
1.2	DEUX GRANDES CLASSES DE MODÈLES	27
1.2.1	Modèles hiérarchiques	28
1.2.2	Modèles componentiels	28
1.3	LES DIFFÉRENTES PERSPECTIVES D'ÉTUDE	29
1.3.1	Perspective darwinienne	29
1.3.2	Perspective jamesienne	29
1.3.3	Perspective cognitive	30
1.3.4	Perspective socio-constructiviste	30

Feelings are not supposed to be logical. Dangerous is the man who has rationalized his emotions.^a

— David BORENSTEIN

a. « Les sentiments ne sont pas supposés être logiques. Dangereux serait l'homme qui rationaliserait ses émotions. » (NDR)

Classer les divers points de vue d'un phénomène est l'un des meilleurs moyens de l'aborder dans une perspective logique, et l'on pourrait définir l'émotion en dressant la liste de tout ce qui est aujourd'hui considéré comme manifestation émotionnelle et, par déduction, ce qui ne l'est pas. Néanmoins, quand bien même les activités de recherche autour de l'émotion sont telles que l'on peut assez facilement en identifier ses différents aspects, il n'est pas question ici de faire un descriptif exhaustif de toutes les formes émotionnelles que l'on rencontre habituellement. Effectivement, même si l'usage voudrait que l'on cite

toutes les émotions possibles et imaginables avant de les trier par catégories selon des critères physiologiques tels que des manifestations gestuelles ou verbales, ou encore tout simplement les caractériser de façon manichéenne, c'est-à-dire émotions *positives* et *néglatives*, il n'en reste pas moins que l'intériorité du sentiment émotionnel, que la langue anglaise appelle *feeling* en référence au ressenti propre et non pas aux signes extérieurs qui en sont issus, nous oblige à observer les émotions d'une façon plus fondamentaliste.

Pour ce faire, j'ai donc considéré que les techniques de classification générales utilisées notamment en sociologie des émotions afin d'étudier l'émotion uniquement d'un point de vue externe ne sont pas *a priori* utilisables si l'on souhaite comprendre les processus émotionnels dans leur intériorité. Je vais par conséquent tenter une approche plurielle qui consiste à observer l'émotion sous des angles différents.

Dans un premier temps, je vais présenter deux exemples de classifications typiques des émotions : le premier classe les modèles de l'émotion selon leur forme et leur mode d'évolution, et le second, dans une optique plus structurale, les trie selon leur organisation. Dans un second temps seront étudiées de manière plus chronologique les différentes perspectives d'étude de l'émotion. Mais auparavant, il nous faut expliciter l'origine du mot *émotion* ainsi que sa réelle portée sémantique.

1.1 ORIGINE DU TERME

Étymologiquement, le terme *émotion* ([emosjõ]) vient du mot *émouvoir*, lui-même issu du mot latin *emovere*, c'est-à-dire *ex-* (« vers l'extérieur ») associé à *movere* (« mouvoir, agiter »). Le mot *emovere* signifie donc « ôter d'un lieu, déplacer », mais aussi « remuer, ébranler » ([47]), ce qui dénote une forte expression visible de l'extérieur. D'ailleurs, le terme *émotion*, apparu au xvii^e siècle, désignait un sentiment fort et profond, bouleversant en somme. Au xix^e siècle, son sens s'affaiblit et désigne simplement un sentiment quelconque. Le dictionnaire Trésor de la Langue Française ([2]) nous donne entre autres les définitions suivantes :

1. *Veilli*. Mouvement assez vif.
 - « *L'émotion de l'air* ».
 - *Au fig.* « *Émotion collective, populaire* ».
2. Conduite réactive, réflexe, involontaire vécue simultanément au niveau du corps d'une manière plus ou moins violente et affectivement sur le mode du plaisir ou de la douleur.
 - « *Éprouver, ressentir une émotion* ».

Commençons par la deuxième définition. Celle-ci se rapporte à notre expérience quotidienne de l'émotion, à savoir l'idée d'activité automatique, involontaire, et qui peut s'apparenter au plaisir ou à la douleur. La notion fondamentale décrite ici est la dualité *esprit – corps*, présente depuis l'Antiquité, mais qui reste à ce jour difficile à discerner. L'émotion serait alors un mécanisme hybride. D'un côté, le ressenti au niveau du corps, qui peut se manifester par un ensemble d'expressions corporelles typiques (par exemple le sourire, les sourcils qui froncent, les mains qui tremblent ...) ou par une activité viscérale caractéristique (un nœud dans l'estomac, le rythme cardiaque qui s'accélère, la transpiration ...). D'un autre côté, le ressenti de l'esprit qui s'apparente à l'affect, c'est-à-dire les mécanismes psychologiques issus des stimuli environnementaux et influençant le comportement. Il faut alors parler de *sentiment* émotionnel.

Bien que la première définition semble succincte, on remarque pourtant l'utilisation du concept de mouvement. En effet, si l'on revient à l'étymologie du terme, le principe de mouvement est central dans l'émotion, si bien qu'il est directement à l'origine de l'idée de bouleversement psychologique associé. Aujourd'hui, l'usage répété du mot en a considérablement réduit la portée, et le terme *émotion* s'est désolidarisé *a priori* de la notion de mouvement. En réalité, cette idée n'a jamais quitté l'esprit des intellectuels scientifiques, et il faut revenir à cette célèbre dualité *esprit – corps* énoncée précédemment pour la retrouver :

Je considère que les émotions sont, en essence, constituées par des changements survenant dans l'état du corps, induits dans ses nombreux organes par les terminaisons nerveuses issues d'un système neural spécifique, lorsque celui-ci répond aux pensées évoquées par un phénomène ou un événement donnés. Un grand nombre des changements qui s'opèrent dans l'état du corps — comme la variation de la couleur de la peau, de la posture corporelle et de l'expression faciale — sont, en fait, perceptibles par un observateur extérieur [...]. Certains autres changements dans l'état du corps ne sont perceptibles que par l'individu chez lequel ils se produisent.

— Antonio DAMASIO ([23], p.194)

L'étendue de la réalité sémantique qui entoure le concept d'émotion ne nous permet pas à l'heure actuelle d'en avoir une définition claire, précise, et surtout universelle, et ce pour deux raisons principales. D'une part, il est très difficile, voire impossible, de délimiter précisément l'activité émotionnelle et de la distinguer de ses pairs dans la construction du processus cérébral dans son ensemble. En effet, si l'on place l'émotion dans un cadre évolutionniste, on peut alors distinguer quatre grandes classes de processus cérébraux : les motivations, les émotions, les cognitions et, de façon plus marginale, la conscience (MAYER, SALOVEY & CARUSO, 2000 [88]).

- Les **motivations** sont des processus de base, en réponse à des états internes

de type physiologique tels que la faim, la soif, ou encore les désirs sexuels. Ils ont un cycle temporel plus ou moins déterminé, qui correspond à une logique de fonctionnement conditionnel simple. Ainsi, si un individu a soif, il lui suffit de boire effectivement pour épancher sa soif ;

- Les **émotions** sont considérées comme des processus plus complexes qui répondent à des perceptions internes — réelles ou imaginaires — de changements environnementaux dans le but de fournir une réponse adéquate. Elles se situent donc dans un cycle temporel beaucoup moins rigide et peuvent conditionner plusieurs réponses comportementales différentes pour un problème donné. Ainsi, la peur par exemple, laisse le choix à deux alternatives distinctes — l'affrontement ou la fuite — ce qui par conséquent confère à l'émotion un premier niveau de flexibilité, ce qui est primordial dans le processus de décision ;
- Les **cognitions** sont des productions cérébrales très évoluées qui visent à apprendre de l'environnement, et à en résoudre les situations qui peuvent s'avérer problématiques. Elles se composent de l'apprentissage, de la mémoire et de la résolution de problèmes par l'intermédiaire de processus tels que la planification ou encore le langage. Les cognitions permettent de satisfaire les motivations et d'optimiser l'état émotionnel en conservant, autant que faire se peut, une émotion positive ;
- La **conscience** se traduit par deux processus : la conscience de soi et la conscience du monde. La conscience de soi est la perception interne de l'individu, de sa personnalité et de ses actes. La conscience du monde, au contraire, passe par la perception du monde, des choses — vivantes ou inanimées — de l'environnement, et d'autrui dans la société.

Alors que l'identification de ces concepts paraît relativement aisée, qu'en est-il de la frontière qui relie chacune de ces notions ? Concrètement, si l'on sait assez facilement identifier l'émotion de peur dans une situation de danger quelconque, que peut-on dire par exemple de la peur d'un serpent ? À quel moment passe-t-on de la conscientisation de l'événement concret (la vue du serpent) à celui du ressenti du danger (l'émotion de peur) ? Ce problème se pose aussi bien du point de vue qualitatif que quantitatif, et nécessite une grande précision dans la définition des différentes dimensions du comportement émotionnel.

D'un autre côté, il existe une autre entité qui entretient des rapports très particuliers avec les émotions et qui va nous nécessiter une attention toute particulière (voir partie 2.2 sur le débat entre ZAJONC et LAZARUS-MANDLER) : il s'agit des *cognitions* ou *processus cognitifs*. La détermination des liens entre émotion et cognition fait l'objet de vives polémiques dans le domaine des sciences cognitives depuis environ une cinquantaine d'années. D'une manière générale, et dans une acception unanimement admise, les cognitions représentent les processus cérébraux responsables de la gestion des connaissances en général. Cela

peut aller de la simple mémoire d'événements particuliers aux planifications les plus élaborées, leur point commun étant la mise en lien d'informations apprises et stockées sous forme de données réutilisables en temps voulu par l'individu. La difficulté inhérente à l'existence des processus cognitifs est que ceux-ci forment une sorte de ciment conceptuel entre les stimuli issus de l'environnement et les décisions prises par l'individu en rapport avec ses connaissances, ce qui complique le travail d'isolement des deux processus, d'un côté l'activité émotionnelle, et de l'autre l'activité cognitive.

Cognition

Ensemble des processus cérébraux liés à la connaissance (mémoire, langage, etc.), mais aussi à son acquisition et à sa mobilisation (stockage, récupération, association, etc.).

Aujourd'hui, l'on admet volontiers qu'il existe incontestablement et dans chacun d'entre nous un être cognitif responsable de notre apprentissage, accompagné d'un être émotionnel qui ajoute une dimension intuitive à nos perceptions. Ces deux entités sont restées longtemps indépendantes pour la majeure partie de la communauté scientifique. En effet, depuis l'Antiquité, l'émotion était considérée comme un parasite à la raison, empêchant tout discernement rationnel par des sentiments tout aussi douloureux qu'inutiles. En somme, on n'accordait qu'une place minimale à l'analyse émotionnelle, alors qu'en pratique, émotion et cognition restent deux concepts indissociables et en interaction étroite ; nous le savons aujourd'hui. Depuis les années soixante, certains auteurs ont donc décidé de modéliser plus précisément les processus émotionnels, en se basant sur des concepts issus notamment de la psychologie cognitive.

1.2 DEUX GRANDES CLASSES DE MODÈLES

Des récentes approches qui ont été étudiées dans le domaine de la modélisation de l'émotion, deux principaux types de modèles sont à distinguer : les modèles *hiérarchiques* et *componentiels* (voir BAUDIC & DUCHAMP, 2006 [9]). Ces deux catégories diffèrent en ce sens qu'elles dénotent la façon dont on peut concevoir les liens qui existent entre types d'émotions. Ainsi, soit elles sont organisées hiérarchiquement, soit elles se placent dans une relation de complémentarité.

1.2.1 Modèles hiérarchiques

D'après la définition du modèle hiérarchique, les émotions tiennent leur origine des stades de développement cérébral antérieurs dont la dynamique est centrée sur l'adaptation à l'environnement. Elles sont donc organisées hiérarchiquement, entre des dimensions émotionnelles primordiales ou primaires à la base du système (voir DENTON, 2006 [29]), et de nombreuses émotions dites secondaires construites à partir des premières. Ainsi, des émotions telles que la joie ou la peur conditionnent l'apparition d'émotions plus sophistiquées qui elles-mêmes permettent l'élaboration de processus tels que la jalousie ou la fierté.

Les processus situés dans les plus bas niveaux ont alors un rôle adaptatif important et sont directement liés aux fonctions sensorielles et viscérales telles que la soif ou la douleur. Néanmoins, ces fonctions peuvent être considérées différemment des émotions et être classées séparément comme appartenant à la famille des *motivations* ([88]). Elles peuvent aussi se retrouver englobées dans les processus émotionnels comme faisant partie intégrante du mécanisme émotionnel.

La classification de GRIFFITHS ([55], [20]) est l'une des nombreuses approches hiérarchiques que l'on peut trouver en psychologie. GRIFFITHS décrit trois formes distinctes d'émotions qu'il appelle *kinds of emotions* :

- les *affect program responses* ou réponses du programme d'affect sont les émotions au sens de Charles DARWIN. Elles sont rapides, automatiques, adaptées et surtout universellement reconnues dans toutes les cultures. Ainsi, la peur, le dégoût, ou encore la surprise, sont considérées comme appartenant à cette catégorie ;
- les *higher cognitive emotions* ou émotions de plus haut niveau cognitif tiennent compte des croyances et désirs de l'individu. Elles sont par conséquent plus complexes et variées que les précédentes et sont intimement liées aux processus cognitifs, ce qui leur confère une adaptabilité plus importante à différents problèmes tels que la coordination sociale ;
- les émotions de type *socially sustained pretense* que l'on peut traduire par « prétexte soutenu socialement » et que GRIFFITHS appelle aussi *disclaimed action*, c'est-à-dire « action réfutée ». Cette catégorie émotionnelle concerne toutes les pressions sociales exercées sur l'individu telles que l'éthique ou encore les interdits sociaux. Elle vise à tenir compte du vécu, non seulement du point de vue de l'histoire personnelle de l'individu, mais aussi de la culture qui lui a été transmise à travers les générations.

1.2.2 Modèles componentiels

Dans les modèles componentiels, les émotions ont différentes facettes qualitatives (SCHERER, 2001 [121]). Ce que l'on appelle la « triade de réponses émo-

tionnelles » se compose de trois composantes principales pour la production de l'émotion : l'expérience subjective, les réponses physiologiques périphériques et l'expression motrice. Certains théoriciens y ajoutent les composants cognitifs et motivationnels. L'approche componentielle vise à déterminer le rôle relatif de chacun de ces composants. Les émotions sont alors le produit de l'ensemble de ce processus, de la perception cognitive à la réponse effective. LAZARUS (1982, [68]) et SCHERER (1984, [115]) sont généralement associés à cette approche, ainsi que la plupart des modèles de l'évaluation cognitive que nous verrons un peu plus tard (voir section 3.4).

1.3 LES DIFFÉRENTES PERSPECTIVES D'ÉTUDE

Les travaux de recherche sur la modélisation de l'émotion sont relativement récents et ont débuté réellement dans les années cinquante. Néanmoins, et ce depuis plus d'un siècle, quatre perspectives d'études ont été empruntées pour l'analyse du fonctionnement émotionnel :

- La perspective darwinienne
- La perspective jamesienne
- La perspective cognitiviste
- La perspective socio-constructiviste

Ces différents mouvements de recherche sont en fait intimement liés aux définitions, études et autres explications que l'on attribue à l'émotion. Chacune de ces perspectives met en avant ses propres hypothèses et théories sur la nature, les causes et les conséquences de l'émotion.

1.3.1 Perspective darwinienne

En premier lieu, la perspective darwinienne considère l'émotion comme un processus, un phénomène évolutif, universel, optimisant la survie. Elle est issue des travaux de Charles DARWIN qui en 1872 écrit *L'expression des émotions chez l'homme et l'animal* ([27], traduction française [28]). Dans sa vision évolutionniste, DARWIN affirme qu'il existe un nombre fini d'émotions, universellement reconnues dans toutes les cultures, et qui permettent la survie de l'espèce grâce à leur pouvoir d'adaptation, c'est-à-dire leur capacité à répondre de façon appropriée aux exigences environnementales. EKMAN, IZARD ou encore TOMPKINS sont les principaux ambassadeurs de ce mouvement.

1.3.2 Perspective jamesienne

La perspective jamesienne est introduite par William JAMES en 1884 ([59]). D'après lui, les émotions sont issues de modifications somatiques liées à la per-

ception du contexte. Ce sont ensuite ces changements physiologiques qui induisent la perception d'une émotion particulière. Ainsi, l'expression « *j'ai peur donc je tremble* » devient « *je tremble donc j'ai peur* ». Cette conception, révolutionnaire pour l'époque bien que fortement remise en question par Walter CANNON en 1920, est récemment revenue sur le devant de la scène grâce au neurologue Antonio DAMASIO dans les années 90. Je reviendrai plus en détail sur les travaux de JAMES un peu plus tard (section 2.1).

1.3.3 Perspective cognitiviste

La lignée des psychologues cognitivistes tels que LAZARUS, FRIJDA, SCHERER, ROSEMAN, ou encore ORTONY, CLORE et COLLINS est représentative de la psychologie cognitive. Ici, processus cognitifs et émotionnels sont intimement liés. Les émotions sont alors conçues comme produits d'un processus d'évaluation cognitive. Cette perspective, apparue dans les années soixante avec des auteurs comme LAZARUS va littéralement envahir les travaux de la deuxième partie du xx^e siècle et introduire définitivement la notion de cognition dans le débat des liens entre émotions et cognitions.

1.3.4 Perspective socio-constructiviste

Enfin, et dans une autre mesure, la perspective socio-constructiviste est une mise en forme sociale du changement physiologique et de son interprétation. Les émotions sont des productions culturelles qui tiennent leur sens et leur cohérence de règles sociales apprises. Le principal représentant de ce mouvement est James R. AVERILL ([4]) dans les années 80.

HISTORIQUE DES THÉORIES DE L'ÉMOTION

2

SOMMAIRE

2.1	JAMES-LANGE <i>versus</i> CANNON-BARD	32
2.2	ZAJONC <i>versus</i> LAZARUS-MANDLER	36
2.3	SCHÉMAS DE BASE	39
2.3.1	La théorie de James-Lange	39
2.3.2	La théorie de Cannon-Bard	40
2.3.3	La théorie de Schachter-Singer	40
2.3.4	La théorie de Lazarus	40
2.3.5	Conclusion	41

*Words are, of course, the most powerful drug used
by mankind.^a*

— Rudyard KIPLING

a. « Les mots, c'est évident, sont la plus puissante
drogue utilisée par l'humanité. »

En 1872, le naturaliste britannique Charles DARWIN (1809~1882) publie *L'expression des émotions chez l'homme et l'animal* ([27], [28]) qui marque le début de l'étude scientifique de l'émotion. Dans ce livre, DARWIN affirme que les expressions physiologiques du visage et du corps sont primordiales chez l'animal au même titre que pour l'homme. Il montre alors que la fonction première des émotions est l'adaptation à l'environnement. Les émotions sont ainsi considérées comme innées ou héréditaires ; elles sont donc universelles. Plus tard, dans la seconde moitié du xx^e siècle, de nombreux anthropologues et autres scientifiques du courant social viendront relativiser cette théorie, notamment dans l'importance qu'accorde DARWIN à l'inné par rapport à l'acquis. En effet, les dimensions sociale et culturelle s'avèreront bien plus conséquentes, tant du point

de vue de leur expression que du point de vue de leur contrôle. Quoi qu'il en soit, l'œuvre de DARWIN marque un véritable tournant dans la science de l'émotion, et aura une profonde influence dans la conception des théories de la fin du XIX^e siècle, et dans tout le XX^e siècle.

Dès lors, les théories du courant biologique des émotions, desquelles seront issues celles de TOMKINS (1962, [127]), de IZARD (1972, [57]), ou encore d'EKMAN (1982, [34]), se verront accompagnées, une douzaine d'années plus tard après DARWIN, d'un nouveau courant psychologique des émotions dès la fin du XIX^e siècle. Les nouvelles théories de ce nouveau courant n'ont qu'un rapport indirect avec celles de DARWIN, en ce sens qu'elles considèrent bien plus les aspects subjectifs qu'expressifs dans l'étude des émotions. En effet, depuis l'Antiquité grecque, les émotions faisaient principalement l'objet d'études philosophiques, mais à partir de la fin du XIX^e siècle, l'analyse émotionnelle se fera essentiellement de façon empirique, par l'expérimentation scientifique ou bien par l'observation de sujets pathologiques ou normaux.

Parmi les nombreuses théories développées au cours du XX^e siècle, beaucoup d'entre elles s'opposaient féroce­ment les unes aux autres. Ces divergences ont généré un certain nombre de grands débats d'idées dont les deux principaux sont les conflits JAMES-LANGE *versus* CANNON-BARD, et ZAJONC *versus* LAZARUS-MANDLER. Le premier débat a permis de délaiss­er les anciennes visions autour de l'émotion, et de permettre aux nouvelles conceptions, plus modernes, de voir le jour, tandis que le second débat a relancé la problématique essentielle des liens entre émotion et cognition, longtemps exclue de l'étude et encore d'actualité au début du XXI^e siècle.

2.1 LE DÉBAT JAMES-LANGE *versus* CANNON-BARD : LES PREMIÈRES GRANDES DÉFINITIONS DE L'ÉMOTION

Dès les années 1880 voit le jour la première théorie scientifique de la psychologie des émotions (DARWIN étudiait plus spécifiquement les formes expressives des émotions) grâce au psychologue et philosophe américain William JAMES (1842~1910). En 1884, dans son article *Qu'est-ce qu'une émotion?* ([59]), JAMES considère l'émotion comme un changement physiologique faisant suite à la perception corporelle d'un phénomène en général. En effet, les changements périphériques suivent directement la perception du stimulus, et c'est la perception de ces changements qui constitue l'émotion. Pour prendre une image caractéristique de JAMES, la vision d'un ours (stimulus) va nous conduire à fuir (changements physiologiques), et ce sont les perceptions de ces perturbations somatiques, liées à la course, qui génèrent en nous le sentiment de peur (émotion). Quant au problème, toujours d'actualité, de savoir si l'émotion naît des centres moteurs et sensoriels, ou bien d'une potentielle zone spécifique du cer-

veau, JAMES opte pour la première solution, stipulant que la génération des processus émotionnels est décentralisée et spécifiquement locale. Un an plus tard, Carl LANGE (1834~1900), physicien et psychologue danois publie séparément en 1885 un article intitulé *Les émotions : une étude psycho-physiologique* ([67]) partageant globalement les mêmes conclusions que JAMES. Le seul point de désaccord se situe au niveau de la centralisation de la production émotionnelle. En effet, LANGE émet l'hypothèse de l'existence d'un centre spécifique du cerveau appelé le *centre vasomoteur*, qui commanderait le système vasculaire. Il décrit alors l'émotion comme la perception de changements environnementaux qui viendrait exciter le centre vasomoteur, qui déclencherait une modification vasculaire particulière, à partir de laquelle d'éventuels changements physiologiques secondaires peuvent avoir lieu. Ces deux théories sont dites périphéralistes car elles considèrent que la source des émotions se trouve être en périphérie, c'est-à-dire au niveau de la perception de la modification corporelle captée par les récepteurs. On peut enfin noter que les deux psychologues s'accordent à dire que des réponses physiologiques différentes induisent des émotions différentes, et que par conséquent la notion de différenciation des émotions est parfaitement admise. Nous verrons plus tard que ce concept est souvent soumis à controverses.

Contrairement aux travaux de JAMES et LANGE, les théories de CANNON et BARD sont considérées comme centralistes dans la mesure où l'origine de l'émotion est le système nerveux central. Les modifications physiologiques ne sont alors que la conséquence de l'émotion. Afin de contester les travaux de JAMES-LANGE, Walter CANNON (1871~1945), psychologue américain, organise en 1915 un débat différé avec JAMES (décédé en 1910) au cours duquel il démontre expérimentalement que de fortes émotions telles que la peur ou la fureur naissent dans le cortex cérébral et entraînent de profonds changements corporels. Ces changements typiques sont acquis génétiquement et constituent ainsi des modifications d'utilité biologique. CANNON montre aussi que ces changements corporels sont produits par une grande quantité d'adrénaline sécrétée et sont similaires à une injection volontaire d'adrénaline. Cette substance, par son action sur l'organisme permet de mieux préparer l'individu aux comportements liés à la colère ou à la peur. En effet, l'adrénaline permet l'augmentation du taux de sucre dans le sang et facilite sa circulation dans les différents organes inhérents à l'effort physique, à savoir le cœur, les poumons, le système nerveux central et les membres. Par conséquent, d'après CANNON, les effets induits par l'adrénaline sont adaptés car conformes aux réactions de fuite ou de combat, lesquelles exigent une activité physique considérable en cas de danger immédiat. Il considère donc, à l'instar de la théorie évolutionniste de DARWIN, que l'émotion est un optimisant la survie, chez l'animal comme chez l'homme. Plus tard, en 1927, il publie un article, "*The JAMES-LANGE theory of emotions: A critical examination and an alternative theory*"¹ ([17]) à l'encontre de la rétroaction viscérale. Pour lui, l'induction de

1. « *La théorie des émotions de JAMES-LANGE : critique et théorie alternative* »

changements périphériques ne suffit pas, à elle seule, à produire de l'émotion. Il propose alors le *thalamus* comme centre des réactions émotionnelles : celui-ci serait responsable de la régulation des décharges physiologiques et du ressenti affectif. De cette façon, il déclencherait alors des réponses physiologiques différenciées. La théorie de CANNON diffère donc grandement de l'interprétation émotionnelle de JAMES-LANGE, notamment sur le rôle du feedback viscéral, qui pour CANNON, n'intervient pas dans la création d'émotions, contrairement au thalamus, qui par son excitation, induirait la production émotionnelle. De plus, contrairement à la théorie de JAMES-LANGE, CANNON considère que les modifications corporelles sont étrangement analogues et indépendantes de l'émotion ressentie. Ainsi, la grande joie, la grande peur, le grand chagrin, produisent les mêmes changements corporels, car leur objectif est le même : préparer le corps à l'exécution d'un effort physique important (fuite en cas de peur, ou attaque en cas de fureur). Cependant, si les réponses physiologiques associées aux émotions sont identiques, comment faire pour ressentir l'*émotion différenciée* que l'on connaît ? Il faudra attendre 1964 et les travaux de Stanley SCHACHTER ([113]) pour trouver une réponse à cette question, par l'introduction du nouveau concept d'*étiquetage cognitif* dépendant du contexte, qui permet l'émotion différenciée malgré l'apparente ressemblance des modifications physiologiques, et qui va permettre à la théorie de JAMES-LANGE de se développer en parallèle des conceptions de CANNON-BARD.

Entre-temps, l'acharnement de CANNON et de ses adeptes aura eu raison des conclusions de JAMES et de LANGE. En effet, si la théorie de JAMES-LANGE s'avère exacte, comment expliquer qu'un individu dont les liens entre le cortex cérébral et les viscères sont rompus continue à ressentir des émotions ? Ce raisonnement a été étudié par plusieurs auteurs. En 1924, le neurologue anglais Alexander WILSON (1877~1937) décrit des expérimentations qui s'opposent aux conclusions de JAMES-LANGE ([134]). Il montre par exemple des cas de patients dont les manifestations physiologiques de rires ou de pleurs ne reflètent pas du tout leurs réels ressentis émotionnels internes. Ainsi, ces patients souffraient de ne pouvoir exprimer des émotions qu'ils ressentent subjectivement. Ils pouvaient par exemple exprimer leur peine par de grands éclats de rire, ou leur gaieté par des pleurs. WILSON expose aussi des cas de paralysies émotionnelles du visage, c'est-à-dire de patients souffrant de la perte de manifestations émotionnelles extérieures au niveau des muscles du visage mais qui n'ont aucun problème à éprouver leurs émotions de l'intérieur. D'après WILSON, ils étaient particulièrement affectés de ne pouvoir extérioriser physiquement leurs ressentis émotionnels authentiques. Ces divers cas de paralysie du visage ont amené WILSON à réfuter les aspects périphéralistes de JAMES et LANGE. Par ailleurs, Charles DANA (1852~1935), neurologue américain, raconte, en 1921, le cas d'une femme d'une quarantaine d'années s'étant brisé le cou au niveau des troisième et quatrième vertèbres cervicales, et qui par conséquent souffrait d'une totale paralysie des

muscles du squelette, du tronc et des quatre membres, et d'une complète insensibilité de tout le corps à partir du cou. Elle décédera un an plus tard. Durant ce laps de temps, DANA observa attentivement les manifestations émotionnelles de la patiente avant d'en arriver à la conclusion qu'aucune modification sensible au niveau de son comportement n'était à signaler, bien qu'il ne lui était pas possible d'utiliser ses membres et notamment la musculature de son cou. Toute réaction potentiellement émotionnelle inhérente au corps étant parfaitement impossible, DANA en conclut la réfutation de la théorie de JAMES-LANGE au profit de celle de CANNON.

Plus tard en 1924, Gregorio MARAÑÓN, médecin espagnol, appuiera la théorie de JAMES-LANGE au travers de ses recherches ([86]), et apportera sa pierre au débat, au même titre que d'autres auteurs tels que HOHMANN qui, par l'étude de patients paraplégiques, démontre que ces derniers éprouvent moins d'émotions que des patients valides, et rejoindra ainsi le point de vue de MARAÑÓN. Il faudra tout de même attendre l'arrivée de SCHACHTER et SINGER en 1962 ([114]) pour que les théories de JAMES-LANGE prennent un nouvel envol. En effet, ils tentent de répondre, entre 1957 et 1964, à CANNON et ses partisans en soumettant l'idée de JAMES qui consiste à dire que l'éveil physiologique est nécessaire à la production de l'émotion. Toutefois, ils ne sont pas en accord avec JAMES sur la différenciation des modifications physiologiques et vont suivre les théories de CANNON pour arriver à la conclusion que les changements somatiques ne sont en aucun cas différenciés. On en revient alors à l'idée d'*étiquetage cognitif* selon lequel la différenciation s'opère par des facteurs cognitifs et par le rôle du contexte. En conclusion, il faut pour SCHACHTER et SINGER deux conditions à la production d'émotions : les modifications somatiques, et l'étiquetage cognitif dépendant du contexte. C'est la première fois dans les courants scientifiques autour de l'émotion que les cognitions sont associées à l'émotion dans l'élaboration de la réponse comportementale. Les années soixante vont alors servir de terreau à l'apparition d'un nouveau débat, vingt années plus tard, non pas sur la définition même de l'émotion mais sur les liens entre émotions et cognitions. D'un côté sera défendue la primauté de l'émotion sur la cognition, sous couvert d'antériorité phylogénétique (issue de l'histoire évolutive de l'espèce) et ontogénétique (issue du développement propre de l'individu), et de l'autre, la primauté de la cognition sur l'émotion avec l'idée que l'interaction entre l'être humain et l'environnement passe nécessairement par le système cognitif, lequel met en place des processus perceptifs et cognitifs, rapides et automatiques.

2.2 ZAJONC *versus* LAZARUS-MANDLER : ÉMOTIONS ET COGNITIONS

Robert ZAJONC (1923~2008), psychologue social à l'université Stanford, et Richard LAZARUS (1922~2002), psychologue cognitiviste de l'université de Californie à Berkeley, ont marqué le renouveau de la réflexion autour de l'émotion à partir des années 80 en y adjoignant un concept clef : la *cognition*. Pour ZAJONC, les émotions sont prépondérantes ([136]), tandis que LAZARUS considère que la cognition est nécessaire à l'émotion ([68]). L'hypothèse de base de ZAJONC suppose que le premier comportement d'un individu envers un objet ou une autre personne est une réaction émotionnelle d'approche ou d'évitement, et qu'en ce qui concerne le ou les processus cognitifs qui s'ensuivent, plus tardifs, soit ils en résultent, soit ils en sont complètement indépendants. Par conséquent, les émotions n'ont aucunement besoin de cognition pour exister. Par ailleurs, ZAJONC utilisera cet argument pour affirmer que la prise de conscience des modifications physiologiques n'est pas une nécessité à l'apparition d'une émotion. À l'issue de nombreuses recherches expérimentales, il finira par défendre l'idée selon laquelle le système cognitif et le système émotionnel sont en partie indépendants, mais que s'il existe un lien entre les deux, c'est l'émotion qui prévaut. Ainsi, le système cognitif est relégué à la seconde place en raison de son apparition plus tardive, aussi bien phylogénétiquement qu'ontogénétiquement, ce qui viendra renforcer le mouvement darwiniste de la fin du xx^e siècle.

Pour appuyer l'idée d'antériorité phylogénétique et ontogénétique de l'émotion, ZAJONC ajoute que celle-ci s'inscrit dans une logique de survie de l'espèce. En effet, la sélection naturelle ayant permis aux processus émotionnels de subsister au cours de l'évolution des espèces, l'émotion est donc adaptative, car dans le cas contraire, d'autres schémas d'action auraient été sélectionnés. D'après ZAJONC, si un individu se trouve dans une position de danger immédiat, il ne va ni s'attarder à contempler la situation ou le stimulus, ni évaluer la dangerosité réelle de la situation. À l'inverse, une action rapide de fuite ou de combat est biologiquement souhaitable pour la survie de l'individu : l'investissement cognitif est donc minimal. Dans une autre mesure, le caractère adaptatif des émotions conforte l'idée selon laquelle les émotions ne sont pas contrôlables par la cognition, celle-ci étant plus longue à se mettre en œuvre.

L'autre argument majeur de ZAJONC s'appuie sur des données expérimentales basées sur le jugement de préférence pour évaluer les réactions émotionnelles, et le jugement de reconnaissance pour évaluer les réactions cognitives. Dans le cadre d'une expérience réalisée avec l'aide de ses collaborateurs ([66]), il présente aux sujets dix polygones irréguliers, cinq fois chacun, avec une durée de présentation de une milliseconde. Dès lors, les sujets affirmaient avoir vu quelque chose, mais n'étaient pas certains de ce qu'ils voyaient. Dans un second temps, on présente aux sujets des couples de polygones composés chacun d'un

stimulus-cible, c'est-à-dire un stimulus déjà présenté, et d'un distracteur, à savoir, un stimulus non présenté auparavant. Les sujets devaient alors exprimer des jugements de préférence (« *Quel stimulus préférez-vous ?* ») ou des jugements de reconnaissance (« *Quel stimulus avez-vous déjà vu ?* »), dans une situation de choix forcé parmi les deux figures constitutives de la paire qui leur était présentée. Les résultats montrent que les stimuli-cibles ne sont choisis plus fréquemment qu'ils ne l'auraient été au hasard que dans la phase de jugement de préférence ! En d'autres termes, l'émotion joue un rôle discriminatoire sur la cible, et ce sans reconnaissance préalable. ZAJONC prouve ainsi que le jugement de préférence se produit alors même que le sujet n'a pas consciemment accès au jugement de reconnaissance.

En parallèle, les recherches menées par SEAMON, BRODY et KAUFF viennent conforter la théorie de ZAJONC ([122]) en montrant que les jugements de préférence des cibles surviennent en réponse à des présentations de stimuli plus courtes que celles que réclament les jugements de reconnaissance. Pour des durées de présentations très courtes, ni les jugements de préférence, ni les jugements de reconnaissance ne se distinguent de la réponse au hasard, ce qui suppose qu'un temps minimal est nécessaire à chacun de ces deux types de jugement. Pour des durées de présentations un peu plus longues, les jugements de préférence s'expriment avec une prévalence supérieure au hasard, là où la reconnaissance ne change pas. Pour des durées de présentations encore plus élevées, la reconnaissance des cibles apparaît enfin, et plus la durée s'allonge, plus la courbe de reconnaissance se rapproche de la courbe de préférence, la croise, puis la dépasse lorsque la durée dépasse 200 millisecondes. Les auteurs utilisent alors ces résultats d'expérience pour démontrer expérimentalement la primauté de l'émotion sur la cognition.

Pour Richard LAZARUS, c'est le contraire : la cognition prime sur l'émotion. En 1982, il met en lumière un malentendu avec ZAJONC sur la définition même du concept de cognition. Celui-ci indique que les cognitions sont des processus délibérés, rationnels et conscients, là où LAZARUS, MANDLER et d'autres indiquent qu'elles peuvent être rapides, non contrôlables et inconscientes. À cette époque, beaucoup de scientifiques considèrent le débat inutile, dans la mesure où les deux opposants ne peuvent se mettre d'accord sur une définition unique de la cognition, et l'affrontement aurait pu en finir là. En réalité, l'existence de processus cognitifs non contrôlables et inconscients tarde à être officiellement démontrée. Pourtant, c'est cette cognition qui permet d'analyser rapidement une situation et d'entraîner une réaction émotionnelle avant la conscientisation effective par l'individu.

Georges MANDLER (1924~) entre alors en jeu en critiquant les théories de ZAJONC par des expérimentations analogues. Il commence par remettre en question le statut de jugements de préférence exprimés à la suite de brèves présentations d'un stimulus, que KUNST-WILSON et ZAJONC décrivent dans leurs travaux

([66]), en démontrant qu'il n'existe pas de spécificité du traitement affectif relative à l'activation des représentations de formes sans signification. Après avoir reproduit à l'identique les résultats obtenus par ZAJONC, MANDLER a non seulement redémontré la préférence des sujets pour les formes ayant déjà été présentées précédemment sans pour autant les reconnaître, mais aussi des jugements de familiarité pour ces formes comparables aux jugements de préférence. En d'autres termes, cette nouvelle interprétation des mêmes résultats part du principe que les jugements affectifs ne sont pas nécessairement indépendants des jugements cognitifs. En effet, les théories de la reconnaissance ([3], [58], [84]) distinguent deux mécanismes que sont le *processus d'activation* et le *processus d'élaboration*, dans le traitement de l'information. Cette dichotomie permet de discerner les effets de la familiarité au sein des processus émotionnels. MANDLER démontre alors que les formes non familières produisent les résultats initiaux. Ceci est dû à l'absence de contexte qui empêche toute élaboration et dont la seule conséquence est la plus grande accessibilité de la représentation activée, phénomène que JACOBY a appelé *dextérité perceptive* (*perceptual fluency*), et que l'on peut définir par la tendance à juger plus favorablement les stimuli traités le plus rapidement.

MANDLER pense ainsi que tout jugement rapide sur n'importe quelle dimension pertinente visera à choisir la plus distincte des deux formes, à savoir celle ayant déjà été présentée précédemment. Le problème est que la présentation antérieure d'un objet peut générer à la fois de la préférence et de la familiarité en l'absence de reconnaissance. En outre, il démontre que les résultats obtenus par KUNST-WILSON et ZAJONC sont parfaitement indépendants du type de stimulus testé et sont uniquement liés à l'activation de la représentation de formes quelles qu'elles soient. Son idée est la suivante : seule une présentation antérieure peut déclencher les représentations du stimulus et permettre alors à cette activation de s'étendre à tout jugement pertinent lié aux stimuli. Il a par ailleurs montré que les jugements de brillance ou de noirceur sont clairement favorisés par des présentations antérieures. Les résultats observés ([85]) révèlent que les jugements de brillance, et de la même manière les jugements de noirceur, sont effectués en l'absence de reconnaissance du stimulus, avec la même probabilité que les jugements de préférence. MANDLER en déduit donc que la pertinence des décisions est peut être liée à l'activation de la représentation du stimulus. Par contre, à l'inverse des expérimentations de KUNST-WILSON et ZAJONC, MANDLER n'émet aucune conclusion quant aux discriminations affectives spécifiques, c'est-à-dire qu'il y aurait primauté des processus émotionnels avant tout traitement cognitif.

Le débat qui a animé entre autres ZAJONC, LAZARUS et MANDLER reste, encore aujourd'hui, toujours actif. Par sa virulence, il démontre non seulement la difficulté à observer les émotions et cognitions en général, mais aussi le manque d'harmonie dans la définition des différents concepts utilisés. Pour autant, mal-

gré ces limitations, de nombreux modèles de l'émotion ont pu naître durant ces dernières décennies.

2.3 SCHÉMAS DE BASE

Au regard de la profusion des modèles de l'émotion présents dans la littérature, il me paraît important, avant toute opération de modélisation, de récapituler certaines théories en les visualisant d'un point de vue de leur déroulement temporel. En effet, si l'on effectue une liste des processus liés de près ou de loin à l'activité émotionnelle, seuls quelques concepts-clefs sont généralement concernés :

- le **stimulus** initial à l'origine du déclenchement de l'émotion dans son ensemble ;
- l'**éveil** (*arousal*) qui représente ici les modifications physiologiques ;
- l'**interprétation**, autrement dit la traduction inconsciente des mécanismes physiologiques ;
- l'**évaluation** (*appraisal*), à savoir l'évaluation cognitive de la situation ;
- le **ressenti** émotionnel, souvent confondu avec l'émotion au sens large, mais qui ici indique le sentiment (*feeling*), l'expérience émotionnelle.

2.3.1 La théorie de James-Lange

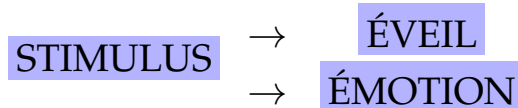
Apparue dès le XIX^e siècle, la théorie de William JAMES et Carl LANGE place l'éveil physiologique au début du processus d'évaluation, juste avant l'interprétation. Seulement ensuite intervient le ressenti émotionnel associé à cette interprétation. Si l'éveil ne se déclenche pas ou si la pensée n'intervient pas à l'issue de cet éveil, aucune expérience émotionnelle n'est possible.

STIMULUS → **ÉVEIL** → **INTERPRÉTATION** → **ÉMOTION**

Exemple de situation : « Vous errez dans une ruelle sombre tard dans la nuit. Soudain, vous entendez des bruits de pas derrière vous. Vous commencez à trembler, vos pulsations cardiaques s'accroissent et votre respiration devient plus profonde. Vous remarquez inconsciemment ces bouleversements physiologiques, en les interprétant comme une préparation du corps à affronter une situation qui suscite la peur. Enfin, vous ressentez la peur. »

2.3.2 La théorie de Cannon-Bard

En réponse à la théorie de JAMES-LANGE, Walter CANNON et Philip BARD proposent un schéma selon lequel l'éveil physiologique et le ressenti émotionnel se réalisent simultanément, sans tenir compte des aspects cognitifs et comportementaux dont on suppose qu'ils arrivent *a posteriori*.



Exemple : « Vous errez dans une ruelle sombre tard dans la nuit. Soudain, vous entendez des bruits de pas derrière vous. Vous commencez à trembler, vos pulsations cardiaques s'accroissent et votre respiration devient plus profonde. En même temps que vous ressentez ces manifestations physiologiques, vous êtes envahi par la peur. »

2.3.3 La théorie de Schachter-Singer

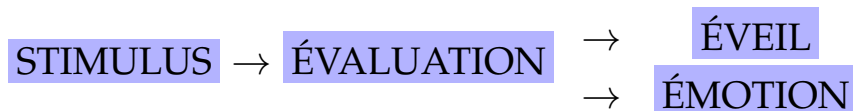
SCHACHTER et SINGER introduisent pour la première fois la notion d'étiquetage cognitif qui permet la différenciation de l'émotion ressentie. Par conséquent, cette dernière arrive logiquement après la rationalisation de l'éveil physiologique.



« Vous errez dans une ruelle sombre tard dans la nuit. Soudain, vous entendez des bruits de pas derrière vous. Vous commencez à trembler, vos pulsations cardiaques s'accroissent et votre respiration devient plus profonde. Tout en ressentant ces manifestations physiologiques, vous réalisez qu'elles sont dues au fait que vous êtes seul dans une rue sombre. Cette situation vous paraît dangereuse et par conséquent, vous êtes envahi par la peur. »

2.3.4 La théorie de Lazarus

Enfin, Richard LAZARUS propose une théorie qui considère en priorité les processus cognitifs. Ainsi, la pensée arrive juste après le stimulus et déclenche en parallèle l'éveil physiologique et l'émotion.



« Vous errez dans une ruelle sombre tard dans la nuit. Soudain, vous entendez des bruits de pas derrière vous et vous vous dites qu'il peut s'agir d'un individu mal intentionné. Par conséquent, vous commencez à trembler, vos pulsations cardiaques s'accroissent et votre respiration devient plus profonde, alors que vous êtes envahi par la peur. »

2.3.5 Conclusion

Ces quelques schémas récapitulatifs révèlent bien la difficulté à se représenter ne serait-ce que la temporalité de l'émotion. De plus, une confusion existe entre d'un côté le ressenti émotionnel (ou le *feeling*) purement interne et subjectif, et d'un autre côté les processus dits émotionnels qui visent davantage à identifier les différentes dynamiques produisant les réponses que l'on qualifie généralement d'émotionnelles, c'est-à-dire principalement les réponses comportementales (expressions faciales, gestes, paroles . . .) et les aspects décisionnels.

Pourtant, ces représentations suffisent souvent à concevoir des modèles informatiques de base. En effet, l'articulation logique des différentes composantes, tant au niveau structurel qu'au niveau temporel, nous permet de comprendre le parcours de l'information. Malheureusement, si le ressenti émotionnel s'avère important pour la conception de modèles d'interaction homme-machine, le processus émotionnel, lui, n'est pas clairement indiqué et force est de constater que bien peu de scientifiques délimitent formellement les bornes de l'émotion en tant qu'activité interne dynamique. Dans cette thèse, on considérera l'émotion comme une entité présente finalement dans toutes les composantes citées précédemment. Ainsi, il s'agit du fruit de l'activité globale du système, dont la structure et la temporalité jouent un rôle primordial d'une part dans le fonctionnement et la sensibilité du système, et d'autre part dans l'évolution et la rétroaction du système sur lui-même.

VERS UNE VISION CONTEMPORAINE DES ÉMOTIONS

3

SOMMAIRE

3.1	LES ÉMOTIONS DE BASE	44
3.2	EKMAN ET LES MODÈLES ÉVOLUTIONNISTES	47
3.3	LES THÉORIES COGNITIVISTES	48
3.3.1	La théorie de Schachter et Singer (1962)	49
3.3.2	Le modèle OCC	49
3.3.3	La « roue des émotions » de Robert Plutchik	50
3.3.4	Conclusion	52
3.4	LES THÉORIES DE L' <i>appraisal</i>	53
3.4.1	La théorie de Magda Arnold (1945)	54
3.4.2	La théorie de Frijda	55
3.4.3	Le modèle d'évaluation cognitive de Roseman	55
3.4.4	La théorie de Lazarus	56
3.4.5	La théorie de Scherer	57
3.4.6	Conclusion	58
3.5	LES THÉORIES SOCIO-CONSTRUCTIVISTES	58

*Maybe if we tell the truth about the past, we can tell
the truth about the present.^a*

— Ken LOACH

a. « Si nous osons dire la vérité sur le passé, peut-être
oserons-nous dire la vérité sur le présent. »

Malgré les différences conceptuelles qui existent entre les théories de JAMES, LANGE et CANNON, il semble tout de même important de noter que tous prennent en compte les stimuli environnementaux comme étant directement responsables

de la production émotionnelle, et que toute émotion est accompagnée de manifestations physiologiques, dont certaines sont visibles extérieurement. Néanmoins, les aspects communicationnels ne sont en aucun cas abordés, et l'on pourrait se demander si l'émotion ne serait pas étroitement liée aux processus cognitifs permettant à l'individu de déterminer son environnement selon ses propres critères, et ainsi générer des processus émotionnels qui lui sont propres et personnels, conformes à son vécu événementiel et à ses motivations. Dans cette optique sont apparus les modèles cognitifs aux environs des années soixante, qui ont pour la première fois donné la part grande aux aspects cognitifs, parallèlement aux productions émotionnelles. Ces deux grands débats auxquels la communauté scientifique assistera tout au long du xx^e siècle formeront le squelette de l'évolution des théories scientifiques de l'émotion. Par ailleurs, la seconde moitié de ce siècle verra apparaître de nombreux travaux sur l'émotion, et ce malgré les disparités entre domaines scientifiques. Dès lors, psychologues, sociologues, neurologues, philosophes, vont parfois aller jusqu'à transcender leurs cultures respectives afin de concevoir les théories qui marqueront l'empreinte scientifique de l'émotion jusqu'au début du xxi^e siècle. Voici leur histoire.

3.1 LES ÉMOTIONS DE BASE

La plupart des modèles de l'émotion présents dans la littérature n'entrent pas nécessairement dans une catégorie bien précise. En effet, comme nous l'avons vu précédemment, que ce soit la classe de modèle utilisée (hiérarchique ou componentielle) ou bien la perspective empruntée (darwinienne, jamesienne, cognitivienne ou socio-constructiviste), elles n'en restent pas moins des pistes de recherche ayant certaines caractéristiques communes. Parmi ces caractéristiques, beaucoup de chercheurs considèrent l'émotion comme un processus multi-dimensionnel, recouvrant un panel d'émotions différentes appelées *émotions de base* ou *émotions primaires*. Ces émotions de base issues des travaux de DARWIN permettent, par leur combinaison, l'expression émotionnelle dans son ensemble.

Comme nous pouvons le constater au sein du tableau des *basic emotions* tiré des travaux d'ORTONY et TURNER (voir figure 3.1, [99]), chaque auteur possède sa propre liste de dimensions affectives dont le but est d'exprimer de la manière la plus complète l'éventail des possibles émotionnels. On pourrait par ailleurs y ajouter une dernière ligne dictée par la médecine traditionnelle chinoise qui affirme que le psychique est indissociable du corps et se manifeste par cinq émotions, chacune associée à un des éléments des *cinq phases* (ou en chinois, les *wǔxíng*, 五行) :

- la **colère**, associée au *bois* (*mù*, 木);
- la **joie**, associée au *feu* (*huǒ*, 火);

	Basic Emotions	Basis for Inclusion
Plutchik	Acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise	Relation to adaptive biological processes
Arnold	Anger, aversion, courage, dejection, desire, despair, fear, hate, hope, love, sadness	Relation to action tendencies
Ekman, Friesen and Ellsworth	Anger, disgust, fear, joy, sadness, surprise	Universal facial expressions
Frijda	Desire, happiness, interest, surprise, wonder, sorrow	Forms of action readiness
Gray	Rage and terror, anxiety, joy	Hardwired
Izard	Anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, surprise	Hardwired
James	Fear, grief, love, rage	Bodily involvement
McDougall	Anger, disgust, elation, fear, subjection, tender-emotion, wonder	Relation to instincts
Mowrer	Pain, pleasure	Unlearned emotional states
Oatley and Johnson-Laird	Anger, disgust, anxiety, happiness, sadness	Do not require propositional content
Panskepp	Expectancy, fear, rage, panic	Hardwired
Tomkins	Anger, interest, contempt, disgust, distress, fear, joy, shame, surprise	Density of neural firing
Watson	Fear, love, rage	Hardwired
Weiner and Graham	Happiness, sadness	Attribution independent

FIG. 3.1 – LES *Basic Emotions*

- la **pensée**, que l'on peut aussi traduire par réflexion, introspection ou souci, associée à la *terre* (*tǔ*, 土);
- la **tristesse** ou l'**anxiété**, associée au *métal* (*jīn*, 金);
- la **peur**, associée à l'*eau* (*shuǐ*, 水).

Ces *basic emotions* illustrent trois phénomènes que je qualifierais de symptomatiques dans la mesure où ils conditionnent de manière profonde la façon dont l'émotion est analysée de nos jours.

Premièrement, les processus émotionnels, le plus souvent, ne peuvent être décrits que de manière discrète au sens mathématique du terme, en définissant par une liste finie d'entités les étapes émotionnelles par lesquelles un individu peut naviguer. Les émotions de base posent ainsi un ensemble fini de pôles

sémantiques distincts qui ont pour principal avantage d'offrir une description simple et concise de la nature d'un événement émotionnel quel qu'il soit. Néanmoins, cette énumération d'axes émotionnels prédéterminés présente l'inconvénient majeur de rendre plus difficile la définition des espaces inter-émotionnels, et ce malgré l'utilisation d'agrégats émotionnels dans une perspective hiérarchique ou componentielle.

En second lieu, l'utilisation d'un vocabulaire spécifique permet à chacun de comprendre l'état émotionnel d'un individu et facilite grandement les relations sociales ainsi que l'empathie. Cependant, on peut émettre certaines critiques à l'égard de la simple utilisation de termes conventionnels car ils présentent une dimension subjective réelle qui peut fortement perturber l'aspect générique de l'émotion. Ainsi, par exemple, si un individu s'exclame « *J'ai peur !* » lors d'une rencontre avec un anaconda en pleine forêt amazonienne, l'émotion ressentie n'aura rien de comparable avec un jeune homme qui « *a peur de n'avoir assez d'argent pour payer le loyer* ». Et pourtant, le terme utilisé est le même dans les deux situations ! Cette subtilité linguistique qui fait le charme de la langue française, entre autres langages, permet de décrire l'émotion de manière simple ... mais redoutablement imprécise !

Enfin, la troisième difficulté est liée à la pertinence de la verbalisation émotionnelle. En effet, nous savons que l'émotion est interne, mais jusqu'à quel point peut-on déterminer sa part consciente ? SCHERER répond à cette question à l'aide d'un diagramme de VENN (figure 3.2) représentant trois zones qui décrivent schématiquement la répartition de l'activité émotionnelle ([120]). La zone (A) représente la part des processus totalement inconscients. SCHERER affirme que la plupart des processus entrent dans cette catégorie. Vient ensuite la région (B) qui identifie la part des représentations conscientes qui ne peuvent être verbalisées. Enfin, la zone (C) détermine la part des processus potentiellement verbalisables.

De son côté, Rosalind PICARD, auteur de *l'affective computing*, expose le problème de la façon suivante : « *Think about it: you have better access to your innermost feelings than anyone, but you still do not always know how to "recognize" or label what you are feeling.* ¹ » (Picard, 2003, [102])

Malgré la grande faiblesse dont est victime l'émotion en termes de *conscientisation* et de verbalisation, et donc de représentation, l'utilisation d'expressions dites *de base* comme formalisation générique est pourtant très répandue en psychologie, en neurologie, et par conséquent en informatique. Nous verrons néanmoins que certains modèles parviennent à outrepasser ces limitations.

1. « Si on y pense : vous avez mieux que quiconque accès à vos sentiments les plus profonds, mais vous ne savez pourtant pas toujours comment reconnaître ou étiqueter ce que vous ressentez. » (NDR)

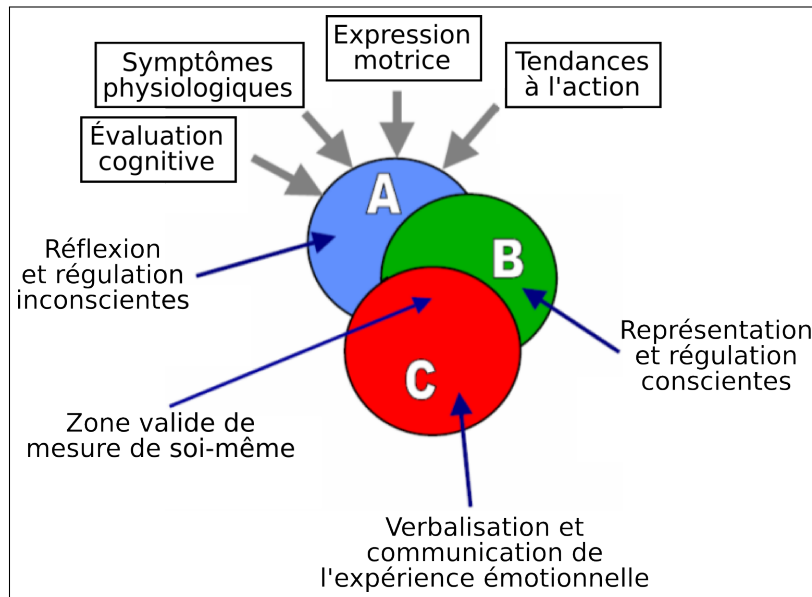


FIG. 3.2 – LE PROBLÈME DE LA VERBALISATION ÉMOTIONNELLE SELON SCHERER

3.2 EKMAN ET LES MODÈLES ÉVOLUTIONNISTES

Les modèles dits évolutionnistes sont directement issus de la perspective darwinienne décrite précédemment (partie 1.3.1). L'émotion est alors vue comme un panel d'états primaires et secondaires dont le but est avant tout la survie. Bien que Charles DARWIN ait entamé le débat il y a plus de 125 ans, des auteurs contemporains s'inspirent encore aujourd'hui de ses travaux. Parmi ceux-ci, EKMAN est sans doute le plus charismatique.

Paul EKMAN (1934~), psychologue américain, a marqué le xx^e siècle par ses travaux sur les émotions et leurs liens avec les expressions faciales humaines ([36]). Dès les années soixante-dix, il décrit six classes d'émotions basiques : la colère, la peur, le dégoût, la surprise, la joie et la tristesse (voir figure 3.3). Chacune de ces émotions est associée à un ensemble de caractéristiques expressives universellement reconnues. Ainsi, un homme en colère froncera les sourcils alors qu'un enfant joyeux montrera son plus beau sourire. EKMAN et FRIESEN ([37]) ont non seulement exposé les différentes expressions faciales de ces émotions basiques, mais ont aussi déterminé l'évolution de ces expressions en fonction de l'intensité de l'émotion. Le résultat de leurs travaux sera compilé au sein du système FACS (*Facial Action Coding System*²) dans lequel les différents mouvements du visage sont décomposés en 46 *action units*³ (ou AU) dont chacune corres-

2. Système de codage de l'activité faciale (NDR)

3. Unités d'action

pond à la contraction ou la détente d'un ou plusieurs muscles du visage ([38]). Par exemple, l'émotion de peur correspond à la liste d'*action units* suivante :

- AU₁, lever les sourcils intérieurs ;
- AU₂, lever les sourcils extérieurs ;
- AU₂₀, tension des lèvres ;
- AU₂₆, abaissement de la mâchoire inférieure.



FIG. 3.3 – LES SIX *basic emotions* (ÉMOTIONS BASIQUES) DE EKMAN : COLÈRE, PEUR, DÉGOÛT, SURPRISE, JOIE ET TRISTESSE

En somme, dans le cadre de ses travaux, EKMAN se place dans une perspective darwinienne en indiquant que la caractéristique universelle des expressions faciales est liée à l'instinct de communication qui vise à informer les autres individus non seulement de son état émotionnel propre, mais aussi d'informations utiles sur la situation générale. Ainsi, un groupe d'individus sera informé par l'expression faciale apeurée d'un de ses membres de la proximité d'un danger par exemple.

3.3 LES THÉORIES COGNITIVISTES

Alors que les théories de la perspective jamesienne considèrent que des stimuli causés par une activation physiologique au niveau des viscères, du centre vasomoteur ou du thalamus déclenchent la création d'émotion, les théories dites

évolutionnistes affirment que les émotions sont le fruit de l'évolution et sont pré-programmées génétiquement. Ces deux approches ne font aucunement mention des aspects cognitifs éventuellement présents dans le déroulement du processus émotionnel. À l'inverse, pour les théories cognitivistes, la cognition joue un rôle capital dans le déclenchement de l'émotion. Celle-ci n'est alors plus uniquement considérée comme un processus automatique ou irrationnel, elle est aussi le fruit d'un processus cognitif directement issu de l'évaluation de la situation par l'individu.

3.3.1 La théorie de Schachter et Singer (1962)

SCHACHTER et SINGER ont succédé à Magda ARNOLD dans le fait de considérer l'émotion comme résultant de l'activité physiologique et des processus cognitifs en parallèle. D'après eux, l'état émotionnel nécessite la présence de deux composantes que sont l'activation physiologique d'un côté, et une cognition appropriée à cet état d'activation de l'autre ([114]). Ensuite, le contexte est interprété selon l'expérience de l'individu, lequel peut ou pas comprendre et définir ses sentiments. Dans le cas où le sujet possède une cognition explicative immédiate de son état, il déterminera clairement son état émotionnel. Dans le cas contraire, il cherchera à interpréter son état en se comparant à autrui. SCHACHTER introduit la notion de « besoin évaluatif » lié à un besoin de comparaison sociale. Il affirme ainsi qu'un individu se doit d'avoir une explication cohérente immédiate de son état émotionnel. Sinon, il utilise les réactions émotionnelles d'autrui comme cible de comparaison. L'émotion serait donc directement fonction du contexte social.

3.3.2 Le modèle OCC : l'émotion selon l'origine de son déclenchement

Le modèle OCC est un des modèles de l'émotion les plus utilisés à ce jour, notamment dans le cadre des systèmes computationnels. Issu de ses trois concepteurs psycho-cognitivistes que sont Andrew ORTONY, Gerald L. CLORE et Allan COLLINS ([98]), il spécifie les états mentaux de 22 émotions-types, classées en onze couples antagonistes, tels que « joie / détresse », « fierté / honte », ou encore « admiration / reproche ». Les émotions de même catégorie sont déclenchées dans des situations analogues, correspondant à trois grandes causes de déclenchement (voir figure 3.4) : les conséquences d'événements, l'action d'agents et la perception d'un objet.

- les émotions étant la **conséquence d'événements** extérieurs, et affectant potentiellement le but de l'individu, telles que la joie, l'espoir ou la peur ;
- les émotions issues de l'**action d'agents**, c'est-à-dire de sa propre action ou de l'action d'un autre individu, et affectant un principe, une règle ou un standard, telles que la honte, la fierté ou le reproche ;

- les émotions liées à la **perception d'un objet** particulier (animé ou inanimé, concret ou abstrait), telles que l'amour ou la haine.

Le modèle OCC définit ainsi trois classes d'émotions, chacune regroupant des émotions-types déclenchées par le même aspect du monde.

		+	-
Conséquences d'événements	Pour les autres	Heureux pour	Ressentiment
		Jubilation	Pitié
	Pour soi-même	Espoir	Peur
		Joie	Détresse
Action des agents	Soi-même	Fierté	Honte
		Satisfaction	Remords
		Gratitude	Colère
	Autre agent	Admiration	Reproche
		Satisfaction	Remords
		Gratitude	Colère
Aspects des objets		Amour	Haine

FIG. 3.4 – ATTRIBUTS ÉMOTIONNELS SELON LE MODÈLE OCC
(VERSION ORIGINALE EN ANNEXE A)

Par ailleurs, un critère d'évaluation, appelé *variable centrale*, est associé à chaque classe et permet de déterminer le type et l'intensité de l'émotion déclenchée suivant les buts, principes et préférences de l'individu. D'autres critères d'évaluation, appelés *variables d'intensité* influençant l'intensité des émotions, sont définis. Plus tard, ORTONY décidera de simplifier le modèle en regroupant certains types d'émotions, portant à cinq le nombre d'émotions positives et à six le nombre d'émotions négatives ([97]).

3.3.3 La « roue des émotions » de Robert Plutchik

Le modèle émotionnel de Robert PLUTCHIK (1927~2006), professeur émérite à l'*Albert Einstein College of Medicine* de New York, a pour but non seulement de définir un certain nombre d'émotions dites basiques, mais aussi des associations émotionnelles pour illustrer le concept d'émotions composées ([103], [104]). PLUTCHIK considère ainsi huit émotions primaires : peur, surprise, tristesse, dégoût, colère, anticipation, joie et acceptation, avec lesquelles il compose d'autres formes d'émotions dites *mixtes*. Par exemple, la joie et l'acceptation se combinent en amour, qui peut alors être plus ou moins intense selon les situations (voir figure 3.5).

Dans cette « roue des émotions », les processus émotionnels s'organisent selon leur opposition polaire, autour de deux dimensions principales :

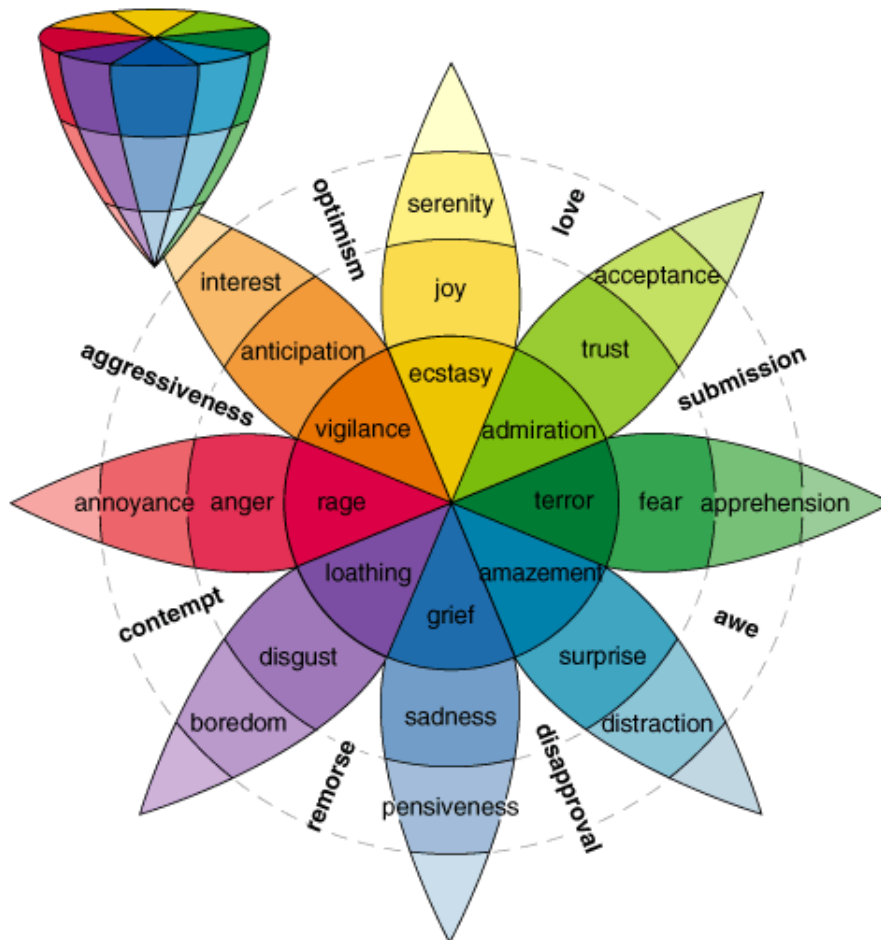


FIG. 3.5 – LA « ROUE DES ÉMOTIONS » DE PLUTCHIK

- l'angle d'orientation par rapport au centre de la roue détermine le type même de l'émotion identifiée, sachant que les différents types d'émotions sont classés par similarité tout autour du centre, et que deux émotions antagonistes auront des positions opposées sur la roue ;
- la distance par rapport au centre de la roue, ou la dimension verticale du cône si l'on considère le volume (en haut à gauche sur la figure) permet de définir l'intensité de cette émotion.

En réalité, PLUTCHIK a créé ce modèle à la suite de sa théorie psycho-évolutionniste des émotions basiques qu'il met au point en 1980 ([103]). Cette théorie admet 10 postulats :

1. *The concept of emotion is applicable to all evolutionary levels and applies to animals as well as to humans.*
2. *Emotions have an evolutionary history and have evolved various forms of expression in different species.*
3. *Emotions served an adaptive role in helping organisms deal with key survival issues posed by the environment.*
4. *Despite different forms of expression of emotions in different species, there are certain common elements, or prototype patterns, that can be identified.*
5. *There is a small number of basic, primary, or prototype emotions.*
6. *All other emotions are mixed or derivative states; that is, they occur as combinations, mixtures, or compounds of the primary emotions.*
7. *Primary emotions are hypothetical constructs or idealized states whose properties and characteristics can only be inferred from various kinds of evidence.*
8. *Primary emotions can be conceptualized in terms of pairs of polar opposites.*
9. *All emotions vary in their degree of similarity to one another.*
10. *Each emotion can exist in varying degrees of intensity or levels of arousal.*

Traduction (NDR) :

1. Le concept d'émotion s'applique à tous les stades de l'évolution, aussi bien aux animaux qu'aux êtres humains.
2. Les émotions possèdent leur histoire évolutive et se sont développées de diverses manières au sein des différentes espèces.
3. Les émotions ont joué un rôle adaptatif dans l'aide aux organismes pour lutter contre les difficultés clés imposées par leur environnement.
4. Malgré les différentes formes d'expression des émotions au sein des différentes espèces, un certain nombre d'éléments communs ou de schémas standard peuvent être identifiés.
5. Il existe un petit nombre d'émotions basiques, primaires ou standard.
6. Toutes les autres émotions sont mixtes ou dérivées ; c'est-à-dire qu'elles sont des combinaisons, des mélanges ou des composés des émotions primaires.
7. Les émotions primaires sont des construits hypothétiques ou des états idéalisés dont les propriétés et caractéristiques peuvent être déduites à partir de diverses sources.
8. Les émotions primaires peuvent être conceptualisées sous forme de paires antagonistes.
9. Toutes les émotions varient dans leur degré de similitude entre elles.
10. Chaque émotion peut s'exprimer à des degrés d'intensité ainsi que des niveaux d'activation variables.

3.3.4 Conclusion

Malgré l'utilisation d'un mode de description simple et clair, souvent à base d'émotions primaires, les théories purement cognitivistes n'indiquent cependant par réellement comment sont stockés les éléments cognitifs, ni la nature des liens

entre processus émotionnels et cognitifs pour parvenir à l'aboutissement d'une décision.

Afin de remédier à cette faiblesse, les modèles dits de l'*appraisal* ont introduit la notion d'évaluation cognitive.

3.4 LES THÉORIES DE L'*appraisal*

Appraisal signifie en anglais « évaluation » dans le sens « *appréciation d'un événement particulier* », et c'est bien d'évaluation cognitive dont il s'agit ici. Aujourd'hui, la plupart des théories de l'émotion s'appuient sur des critères d'évaluation cognitive pour déterminer l'émotion. En effet, un ours en cage au sein d'un zoo ne fera qu'attirer la curiosité, tandis qu'un promeneur rencontrant ce même ours en pleine forêt déclenchera pour l'individu une réaction de peur. Cet exemple récapitule parfaitement la vision de William JAMES mais introduit en parallèle l'idée d'une perception contextualisée qui induit par conséquent une signification différente.

Les chercheurs associés à cette perspective ont déterminé qu'il existait cinq composantes qui déterminent la succession de changements corporels et psychiques caractérisant une émotion :

- **L'évaluation cognitive**, mentionnée précédemment, qui permet à l'individu d'évaluer à quel point un événement particulier révèle une pertinence affective ;
- **L'expression motrice** qui se caractérise par des modifications physiques adaptées, par exemple l'expression faciale ;
- **La réaction du système nerveux périphérique** préparant l'individu à une réaction urgente, par exemple l'augmentation de la fréquence cardiaque pour pouvoir courir et échapper à un agresseur, ou au contraire l'affronter ;
- **La tendance à l'action** qui consiste à vouloir précipiter ou éviter un événement et qui permet d'influer sur les interactions environnementales ;
- **Le sentiment subjectif** permettant la conscientisation et la verbalisation du ressenti émotionnel.

Dans cette approche, l'émotion est vue comme un processus déclenché par une évaluation subjective d'un événement. Le type et l'intensité de cette émotion sont alors déterminés par la perception de cette événement ainsi que par l'évaluation d'un ensemble de variables appelées *variables d'évaluation*. L'attention est particulièrement portée sur la détermination de ces variables dont les valeurs dépendent de l'état mental de l'individu (but, croyances), de son profil psychologique (personnalité, préférences), et de facteurs situationnels et culturels ([69], [118]). C'est ainsi qu'une même situation peut déclencher deux émotions différentes chez deux individus distincts.

Les modèles décrits ici illustrent cette théorie, en montrant pourquoi seuls les événements pertinents pour les objectifs, les valeurs ou les besoins d'un individu sont en mesure de déclencher une émotion particulière et adaptée à la situation.

3.4.1 La théorie de Magda Arnold (1945)

Magda ARNOLD est sans doute la pionnière des théories d'évaluation cognitive. À l'inverse de JAMES ou CANNON, elle affirme que le cerveau n'est pas uniquement doué de processus réflexes mais qu'il participe activement au décodage des informations émotionnelles en général. Ainsi, ARNOLD parle de références à des événements passés en tant qu'expériences émotionnelles, ou encore de prévisions des conséquences possibles de la situation ([1]).

Le modèle de ARNOLD se présente sous la forme d'une succession d'étapes. Tout d'abord, directement issue du stimulus déclencheur, on a une évaluation préliminaire, en fonction des expériences similaires passées et de la prévision événementielle des actions possibles. Pour illustrer ses propos, ARNOLD prend l'exemple de la peur d'un pilote de bombardier qui voit apparaître dans le ciel des avions de combat. D'après son expérience passée, le pilote identifie ces avions comme des ennemis. Cette perception entraîne une évaluation cognitive des conséquences possibles, en l'occurrence une bataille entre lui et eux. Ensuite survient ce que ARNOLD appelle une tendance à l'action dirigée « vers » ou « contre » la situation. Ainsi, le pilote a une attitude émotionnelle qui le pousse à combattre ou bien à fuir ses ennemis. En conséquence, une attitude émotionnelle spécifique de l'émotion en jeu va initier des impulsions nerveuses au niveau du cortex via les centres nerveux (thalamus, hypothalamus), lesquelles vont ensuite déclencher des patterns physiologiques responsables de changements périphériques précis. Pour en revenir à notre pilote de bombardier, les changements périphériques liés à la peur du combat se font sentir sous la forme d'une accélération du rythme cardiaque et de la respiration, à la suite d'une activation corticale, et finalement, des signes extérieurs tels que l'expression faciale de la peur, ou encore une tension musculaire importante apparaissent.

Cependant, le processus émotionnel ne s'arrête pas là. En effet, les modifications physiologiques sont reportées vers le cortex à travers les afférences sensorielles, puis perçues, puis analysées de nouveau. C'est cette réévaluation cognitive qui produira le label émotionnel final. Dans le cas de notre pilote, les changements périphériques caractéristiques de la peur seront donc perçus par le cortex sous forme de rétroaction physiologique, puis réévalués par le pilote par rapport à ses expériences passées, pour ensuite traduire son état par le sentiment de « peur ».

3.4.2 La théorie de Frijda : l'émotion comme tendance à l'action

Nico FRIJDA (1927~), psychologue hollandais à l'université d'Amsterdam, a passé une grande partie de sa vie à étudier l'émotion. Après s'être concentré sur les expressions faciales dans les années cinquante, à une époque dans laquelle l'approche behavioriste est en pleine expansion, il définit le concept de *tendance à l'action*, théorie décrite dans son livre intitulé *The Emotions* ([46]). Pour lui, les émotions représentent des tendances à la réalisation des intérêts, un intérêt étant la disposition d'un individu à préférer certains états de l'environnement ou du sujet lui-même. Par ailleurs, une des composantes préalables à la création du processus émotionnel est la modification de la préparation à l'action dans le cadre de la réalisation d'un intérêt, par exemple, la propension à affronter ou à fuir. Comme FRIJDA considère qu'une émotion et une tendance à l'action se confondent, chaque émotion est alors représentée par une tendance à l'action spécifique (voir figure 3.6).

Tendance d'action	Fonction	Émotion
Évitement	Protection	Peur
« être avec »	Accès à la consommation	Plaisir, confiance
Attention	Orientation vers les stimulations	Intérêt
Rejet	Protection	Dégoût
Inattention	Sélection	Indifférence
Attaque-menace	Récupération du contrôle	Colère
Interruption	Réorientation	Choc, surprise

FIG. 3.6 – PRINCIPALES TENDANCES À L'ACTION AINSI QUE LEURS FONCTIONS ET ÉMOTIONS RESPECTIVES

D'après FRIJDA, les émotions surviennent lorsqu'un intérêt est menacé. Associées chacune à une tendance à l'action, elle permettent de remédier à cette menace. L'exemple le plus caractéristique est la fuite ou l'évitement d'un individu face à un animal dangereux, lequel va déclencher une émotion de peur.

3.4.3 Le modèle d'évaluation cognitive de Roseman

Le modèle de Ira ROSEMAN (1953~), psychologue à l'université Rutgers (New Jersey), définit 5 critères d'évaluation qui, suivant leurs valeurs, permettent de représenter les états mentaux de 13 émotions distinctes ([109], [108]). Par exemple, l'espoir correspond à l'état mental où un événement a été évalué comme incertain mais cohérent avec les buts de l'agent. Les événements sont divisés en deux catégories :

- Les *motive-consistent events* ou événements cohérents avec l'un des buts de l'individu ;

- Les *motive-inconsistent events* ou événements qui menacent l'un des buts de l'individu.

La simplicité du modèle permet de le traduire facilement en un ensemble de règles déterminant quels états mentaux déclenchent quelles émotions chez un agent (voir figure 3.7).

		Positive Emotions Motive-Consistent		Negative Emotions Motive-Inconsistent		
		Appetitive	Aversive	Appetitive	Aversive	
Circumstance-caused	Unexpected	Surprise				
	Uncertain	Hope		Fear		Low Control Potential
	Certain	Joy	Relief	Sadness	Distress	High Control Potential
	Uncertain	Hope		Frustration	Disgust	
Certain	Joy	Relief				
Other-caused	Uncertain	Liking		Dislike		Low Control Potential
	Certain					
	Uncertain			Anger	Contempt	High Control Potential
	Certain					
Self-caused	Uncertain	Pride		Regret		Low Control Potential
	Certain					
	Uncertain			Guilt	Shame	High Control Potential
	Certain					

FIG. 3.7 – MODÈLE DE ROSEMAN

3.4.4 La théorie de Lazarus : théorie relationnelle, motivationnelle et cognitive des émotions

L'idée générale de Richard LAZARUS (1922~2002) est de considérer que l'émotion ne peut pas être uniquement décrite selon les points de vue individuel ou environnemental en tant que deux domaines distincts ([69]). Ainsi, il insiste sur l'aspect relationnel de sa théorie au travers de l'influence mutuelle d'un sujet et de son environnement ([71]).

Néanmoins, seules les interactions susceptibles d'affecter le bien-être de l'individu seraient porteuses d'un message émotionnel. On a alors deux types d'émotions qui se profilent :

- Les *émotions positives* issues de l'interaction sujet-environnement qui favorisent la réalisation des intérêts individuels ou interpersonnels ;
- Les *émotions négatives* issues d'événements qui retardent ou empêchent l'atteinte des objectifs.

L'aspect motivationnel apparaît dans le désir de réalisation de ses buts en adaptation avec son environnement. Ainsi, LAZARUS détermine deux formes distinctes de motivations :

- Une *variable motivationnelle* sous-jacente à une hiérarchie de buts à atteindre ;
- Une *variable transactionnelle* qui définit les dispositions de l'individu à honorer ses buts compte tenu des différentes contraintes, demandes et ressources présentes au sein de l'environnement.

Enfin, l'aspect cognitif est issu de l'utilisation des connaissances de l'individu, et de son analyse des changements environnementaux. LAZARUS, KANNER et FOLKMAN ([70]) distinguent trois types d'évaluation cognitive :

- Les *évaluations primaires* correspondent à l'analyse du degré d'implication pour le bien-être de l'individu. Ainsi, soit la situation n'est pas pertinente (elle n'est pas liée à un intérêt personnel), soit elle est positive (bénéfique ou désirable), soit elle est négative (douloureuse, stressante ...);
- Les *évaluations secondaires* concernent l'analyse des ressources disponibles pour faire face à la situation. LAZARUS et LAUNIER ([71]) affirment qu'elles sont directement liées aux stratégies mises en place par le sujet ;
- Les *réévaluations* se rapportent aux modifications des évaluations primaires et secondaires dans une perspective d'adaptation aux changements de la situation. Ainsi, soit l'individu s'intéresse à ce qui a changé dans sa relation environnementale, soit il s'adapte au stress induit.

Il faut noter cependant que LAZARUS considère que l'évaluation cognitive n'est pas la seule cause de la production émotionnelle. En effet, il faut aussi distinguer un état subjectif particulier (lié à la personnalité), les tendances à l'action (préparation de l'organisme à répondre comportementalement à la situation, cf. FRIJDA, SCHERER), et une configuration de réponses physiologiques.

Les émotions seraient donc non seulement fonction de l'évaluation cognitive interne et étroitement liées aux conditions environnementales, mais aussi fortement influencées par des adaptations subjectives, lesquelles sont couplées à des normes et règles sociales apprises.

3.4.5 La théorie de Scherer : L'émotion définie comme un processus complexe

Klaus SCHERER (1943~), psychologue suisse, affirme que le processus émotionnel est composé de cinq sous-systèmes fonctionnels : les systèmes perceptivo-cognitif, neurophysiologique, motivationnel, moteur et de contrôle. L'émotion est alors définie comme un changement général synchronisé de tous ces sous-systèmes, en réaction à l'évaluation d'un événement particulier ([119]). SCHERER étudie tout particulièrement le système perceptivo-cognitif qui s'occupe de l'évaluation de l'environnement. D'après lui, un individu qui évalue systématiquement son environnement déclenche l'évaluation séquentielle d'un

certain nombre de variables appelées les *Stimulus Evaluation Checks* (SECs). Il existe alors cinq SECs principaux :

- l'évaluation de la nouveauté : « *Qu'est-ce qui a changé chez moi ou dans mon environnement ?* » ;
- l'évaluation du plaisir intrinsèque : « *La situation actuelle est-elle plaisante ou déplaisante ?* » ;
- l'évaluation de la pertinence : « *La situation actuelle est-elle significative ? Affecte-t-elle mes buts ?* » ;
- l'évaluation de la capacité à faire face : « *Ai-je un contrôle sur la situation ?* » ;
- l'évaluation de la compatibilité des normes : « *La situation est-elle compatible avec les conventions sociales ainsi que mes principes personnels ?* »

Une émotion particulière est déclenchée en fonction du résultat de l'activation de ces SECs, les conditions de déclenchement étant liées à des combinaisons spécifiques de SECs. Par ailleurs, SCHERER a développé en 1993 un système expert qui assure la conversion entre résultats d'évaluation des SECs et émotion déclenchée ([116]).

3.4.6 Conclusion

Les modèles de l'*appraisal* offrent globalement une plus grande précision dans la description des relations entre émotion et cognition grâce à la mise en place d'une forme d'interprétation cognitive qui conditionne le déclenchement émotionnel. Là encore, ce procédé ne répond pas concrètement au problème de la prise de décision mais introduit toutefois une coloration émotionnelle aux processus cognitifs, tout en proposant une notion de contextualisation de l'émotion.

Mais afin d'élargir encore le champs d'investigation des phénomènes influençant l'émotion, certains scientifiques sont allés jusqu'à accroître l'étendue de cette contextualisation en y introduisant une dimension sociale.

3.5 LES THÉORIES SOCIO-CONSTRUCTIVISTES

Tous les modèles psychologiques présentés jusqu'ici dépeignent une vision intra-individuelle de l'émotion, c'est-à-dire une étude des processus émotionnels, des liens entre les cognitions et émotions, entre le raisonnement et le ressenti, ou encore de la description des émotions de base, en tant que piliers structurels de la compréhension des états émotionnels. Bien que certains modèles tiennent compte, bien entendu, de l'univers social comme source d'activation potentielle, on ne peut que constater l'évidence de la tournure intrapsychique

que défendent les scientifiques de l'émotion. Cependant, quelques modestes auteurs se sont penchés sur la socialité de l'émotion et considèrent que celle-ci pourrait se révéler être le produit de constructions sociales.

C'est en tout cas le point de vue de James AVERILL, de l'université du Massachusetts, qui considère que l'émotion est principalement dépendante du contexte social dans lequel elle intervient. On ne parle plus ici de caractéristiques innées ou universelles des émotions, mais simplement de la dimension sociale déterminée par les normes, l'éthique et les autres règles sociales localement admises. L'opposition aux concepts d'émotions primaires ou de base est par conséquent très forte.

AVERILL affirme que les processus émotionnels sont en fait des *rôles sociaux transitoires*, c'est-à-dire un ensemble de réponses socialement pré-programmées et alignées sur des normes sociales ou des attentes spécifiques partagées ([4]). Cependant, pour que la définition de l'émotion soit complète, il faut y adjoindre l'évaluation des stimuli émotionnels et l'expression de passivité. Les *évaluations émotionnelles* correspondent à l'appréciation personnelle sur ce qui est désirable et ce qui ne l'est pas. Quant à l'*expression de passivité*, elle est liée à la nature même de l'émotion qui est alors considérée non pas comme une action volontaire mais comme une passion subie.

Si l'émotion n'est plus effectivement assimilée qu'à un ensemble de réponses sociales constituées de normes et d'attentes partagées au sein d'un groupe, et dont la signification est simplement déterminée par l'étiquetage émotionnel que l'individu applique à son comportement, la perspective linguistique entre alors nécessairement en jeu dans la diversité de l'expression émotionnelle. En effet, Anna WIERZBICKA (1938~), linguiste à l'université nationale australienne, émet l'hypothèse, par exemple, que les concepts *happiness, sadness, fear, anger* et *disgust*⁴ sont des créations culturelles de la langue anglaise qui ne peuvent représenter un état émotionnel universel. De la même façon, les travaux de James RUSSELL, professeur de psychologie à l'université de Californie à Los Angeles, tendent à démontrer que le concept d'émotion ainsi que sa catégorisation usuelle en émotions de base ne sont pas clairement universels ([110]).

Cependant, de l'aveu même de AVERILL, une des limites principales de l'approche socio-constructiviste fondée sur l'analyse de rôle sociaux se révèle être l'absence de lien direct avec les problèmes qui se manifestent généralement dans les théories traditionnelles de l'émotion. Toutefois, bien que ces théories aux perspectives quelque peu globalisantes fassent preuve d'une certaine imprécision quant aux mécanismes émotionnels étudiés, elles ont tout de même le mérite de prendre position au sommet de l'échelle des niveaux d'analyse de l'émotion, composée de bas en haut des approches biologiques, physiologiques, cognitives et sociales. Cela pose aussi la question de la véritable nature de l'émotion : n'est-elle le produit que du fonctionnement corporel interne soumis aux

4. joie, tristesse, peur, colère et dégoût

3. VERS UNE VISION CONTEMPORAINE DES ÉMOTIONS

pressions environnementales, ou bien puise-t-elle la majeure partie de sa force dans la dynamique des relations interpersonnelles ?

L'ÉMOTION AU SEIN DES PROCESSUS CÉRÉBRAUX

4

SOMMAIRE

4.1	PHINEAS GAGE ET LE CORTEX PRÉFRONTAL	62
4.2	LE CONCEPT DE CIRCUIT NEURONAL	65
4.2.1	Le circuit de la peur	66
4.2.2	Le circuit du plaisir	67
4.3	CONCLUSION	68

*The distance between insanity and genius is
measured only by success.^a*

— *Tomorrow Never Dies*, Bruce FEIRSTEIN

a. « La distance entre la folie et le génie ne se mesure
que par le succès. » (NDR)

Bien que complexe à étudier, l'émotion est une activité relativement bien localisée, bien qu'elle soit relativement diffuse au sein du système nerveux central. Pour autant, le système nerveux périphérique participe à son élaboration, mais jusqu'à quel point ? On en revient alors à l'éternel débat entre William JAMES et Carl LANGE d'un côté, et Walter CANNON et Philip BARD de l'autre, débat qui a déchaîné les passions au début du xx^e siècle et qui revient sur le devant de la scène depuis quelques années (LEVENSON *et al*, [75] et NIEDENTHAL, [93]). Dans tous les cas, il est clair que les structures neuronales sont responsables de la création de l'activité cérébrale dans son ensemble, et tout particulièrement de l'émotion. Par conséquent, la compréhension du fonctionnement du substrat neuro-anatomique se révèle être indispensable au processus de modélisation de l'activité émotionnelle.

Depuis l'Antiquité, l'étude du cerveau s'est cristallisée autour de la dualité entre corps et esprit, entre cognition et émotion. Cette dernière souffrira d'un

héritage culturel malheureux qui confinerait l'étude des sentiments et autres processus dits émotionnels dans des domaines de la psychologie et des maladies mentales, au profit de la cognition et des thématiques liées au raisonnement et à la connaissance qui, elles, seront rangées du côté des sciences neurologiques. L'étude de l'émotion est alors complètement négligée par la neurobiologie principalement pour cette raison, ainsi que pour un autre motif plus subtil directement lié à la définition même de l'émotion, c'est-à-dire la difficulté qu'éprouve le chercheur à distinguer le processus émotionnel du ressenti émotionnel.

Pour des scientifiques tels que Joseph LEDOUX, émotion et sentiment sont une seule et même entité : « *An emotion is a subjective experience, a passionate invasion of consciousness, a feeling*¹ » ([73]). À l'inverse, Antonio DAMASIO sépare les deux processus en invoquant la possibilité de mesurer certaines émotions à l'aide des expressions faciales ou corporelles ainsi que les manifestations internes comme la production d'hormones ou les battements cardiaques, contrairement aux sentiments qui sont alors perçus comme des ressentis subjectifs complètement personnels. En somme, les émotions sont des comportements tandis que les sentiments sont des pensées ([25]).

Malgré les désaccords sémantiques et le flou conceptuel ambiant, la neurologie tente d'aborder le problème de manière progressive. Ainsi, comme le dit Françoise LOTSTRA : « *En neuroanatomie, la distinction est clairement établie entre expression et conscience émotionnelle sous le contrôle de structures différentes. Les biologistes de l'émotion s'accordent à reconnaître que les comportements émotionnels ainsi que leurs réseaux neuronaux, garants de la survie de l'espèce, ont été sélectionnés au cours de l'évolution* » ([76]). C'est donc dans l'exploration de ces réseaux, et plus particulièrement la découverte des circuits neuronaux spécifiques à chaque émotion que vont se démarquer les neurosciences.

4.1 PHINEAS GAGE ET LE CORTEX PRÉFRONTAL

Bien que l'étude du cerveau ait fasciné de nombreux scientifiques depuis l'Antiquité, on peut estimer l'apparition de la neuroanatomie des émotions aux alentours de 1850. À cette époque, un certain Phineas GAGE fait son apparition dans l'*American Journal of the Medical Sciences*. Il s'agit d'un cas clinique dans lequel on observe une corrélation entre une lésion cérébrale localisée et un profond changement de personnalité. Cette histoire sera exhumée presque un siècle et demi plus tard par Antonio DAMASIO, dans son ouvrage *L'erreur de Descartes* en 1994 ([22], [23]).

Phineas GAGE (1823~1860), 25 ans, dirige une équipe sur un chantier de construction de voies ferrées dans le Vermont. Un jour de 1848, il est victime

1. « Une émotion est une expérience subjective, un envahissement passionné de conscience, un sentiment » (NDR)



FIG. 4.1 – PHINEAS GAGE ET SA FAMEUSE BARRE DE FER

d'un accident alors qu'il fait sauter une mine pour creuser la roche. « *Distract, il commence à bourrer la poudre avec sa barre de fer, alors que son aide n'a pas encore versé le sable. Presque instantanément, cela met le feu à la charge explosive et la mine lui saute à la figure* » ([23], p. 22). La barre de fer lui traverse alors littéralement le crâne, de la joue gauche jusqu'au-dessus de la tête (voir figure 4.2). Étonnamment, Phineas GAGE survécut au drame mais sa personnalité se vit totalement transformée : d'un homme sociable, fin et habile en affaires, il devint d'humeur changeante, capricieux, asocial et grossier. Dès lors, il changea continuellement de travail, passant d'éleveur de chevaux à conducteur de diligences. Il devint même l'attraction du cirque Barnum à New York un an après son accident. Il meurt d'une crise d'épilepsie en 1860, soit 12 ans après son accident.

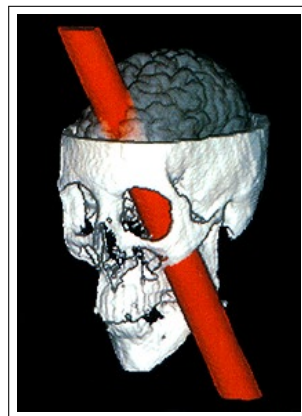


FIG. 4.2 – LE CERVEAU DE PHINEAS GAGE

Son corps est exhumé en 1867 par son médecin John HARLOW mais ce dernier

ne peut pas à l'époque en tirer d'informations substantielles. Il faudra attendre Antonio et Hanna DAMASIO en 1994 pour voir le dossier GAGE ressortir des archives, avec notamment des analyses plus poussées sur la trajectoire exacte de la barre à partir du crâne et de la barre à mine du patient conservés au *Warren Anatomical Museum* de l'université de Harvard.

Cette perte soudaine de tout respect à l'égard des conventions sociales et des règles morales antérieurement apprises liée à une lésion cérébrale sévère, et ceci sans perte apparente des capacités intellectuelles fondamentales ni du langage peuvent indiquer qu'il existe une région du cerveau responsable se rapportant « *de façon spécifique à des caractéristiques propres à l'homme, parmi lesquelles : la capacité d'anticiper l'avenir et de former des plans d'action en fonction d'un environnement social complexe ; le sentiment de responsabilité vis-à-vis de soi-même et des autres ; et la possibilité d'organiser sa survie, en fonction de sa libre volonté* » ([23]).

C'est en tout cas ce que propose DAMASIO après l'étude du cas GAGE parallèlement aux autres observations de patients analogues rapportées ([26], [10]). Le cortex préfrontal est une zone située à l'avant du cerveau (voir figure 4.3) représentant environ un tiers du volume du cortex cérébral. Il est composé notamment du *cortex orbito-frontal* (numéros 10, 11 et 12 sur les schémas) situé juste au-dessus des orbites. « *Sa lésion entraîne une série de modifications comportementales parmi lesquelles on relève l'apparition de traits caractéristiques de la personnalité antisociale : impulsivité, irresponsabilité, absence de conscience sociale, manque d'empathie, investissement excessif dans la recherche des plaisirs. Des expressions émotionnelles déconnectées du contexte social, des émotions perturbées telles que euphorie, irritabilité, exubérance, sensibilité excessive, sont également observées (Zald & Kim, 1996, [137])* » ([76]).

En plus du cortex orbito-frontal, le cortex préfrontal est composé de deux autres zones impliquées dans le comportement. La première, le *cortex cingulaire antérieur* (numéroté 24), affecte directement la motivation. Par conséquent, son altération entraîne une apathie profonde et une diminution importante des capacités cognitives et affectives. On peut en outre observer une connexion directe entre cette région du cortex et le noyau *accumbens*, zone cérébrale qui reçoit la dopamine, souvent citée comme la molécule naturelle du plaisir. Cette particularité entraîne inexorablement un lien entre motivation et plaisir. En effet, le désir de mettre en place certains comportements vitaux est associé au plaisir qu'ils produisent par l'activation d'un système hédonique analogue à un système de récompense chez l'animal, et que l'on identifie comme la *voie dopaminergique mésocorticolimbique*, qui part du tronc cérébral sur lequel repose les deux hémisphères cérébraux jusqu'au cortex frontal et à l'*accumbens*.

La seconde zone impliquée dans le comportement est le *cortex préfrontal latéral* (numéros 9 et 46), responsable de nombreuses capacités cognitives telles que la planification, la mémoire de travail, la capacité d'abstraction, etc. La schizophrénie est souvent liée à une lésion dans cette région du cerveau.

En outre, la communauté scientifique a pu identifier certains circuits neuro-naux impliqués dans les processus émotionnels. En effet, les progrès de l'imagerie médicale et les avancées récentes dans la compréhension des mécanismes sous-tendant l'activité cérébrale ont permis de visualiser certaines *routes* de l'émotion, dont deux ont fait l'objet de nombreuses études : les circuits de la *peur* et du *plaisir*.

4.2 LE CONCEPT DE CIRCUIT NEURONAL

Afin de comprendre les différentes étapes de la découverte des différentes routes de l'émotion, il faut remonter aux années trente. En effet, dès 1937 James PAPEZ (1883~1958) imagine un circuit en anneau au niveau de la partie médiane du cerveau, qui relierait plusieurs structures du système limbique parmi lesquelles le cortex cérébral temporal et cingulaire, le thalamus (figure 4.5(b)), l'hypothalamus (figure 4.5(c)) et certaines de leurs interconnexions ([100]). Cette hypothèse est dans un premier temps confirmée par les expériences de KLÜVER et BUCY ([62], [63]) qui montrent qu'une ablation bilatérale des lobes temporaux chez les singes macaques déclenche une *cécité psychique*, c'est-à-dire une absence de réaction de peur et une perte des émotions en général. Ce syndrome, appelé syndrome de KLÜVER et BUCY se caractérise en outre par une perte des interactions sociales et une profonde altération comportementale.

Après la seconde guerre mondiale, une théorie nouvelle intitulée *théorie du cerveau triunique* proposée par Paul MACLEAN en 1949 viendra bouleverser les connaissances neuroscientifiques ([77]). Au cours de l'évolution, trois cerveaux seraient apparus successivement : le *cerveau reptilien* ou cerveau primitif, le *cerveau paléo-mammalien* ou limbique apparu avec les premiers mammifères, et enfin le *cerveau neo-mammalien* ou neocortex propre aux *Australopithèques* africains. Ce modèle de cerveau à trois couches est aujourd'hui dépassé mais a néanmoins permis de découvrir l'influence du *cortex*, centre de la raison, sur l'hypothalamus, organe des instincts et des pulsions. On découvrira plus tard l'autre rôle de l'hypothalamus, petite glande de la taille d'une noisette dépendante du thalamus, qui garantit l'*homéostasie*, c'est-à-dire l'équilibre du corps en termes de sécrétion d'hormones et de supervision du fonctionnement de la plupart des organes internes. L'homéostasie est en outre inscrite dans le système nerveux autonome, ce qui signifie qu'elle échappe complètement à la volonté dictée par le cortex cérébral. Il faut donc noter un lien finalement assez ténu entre d'un côté les processus relevant de l'homéostasie, incontrôlables car instinctifs, et d'un autre côté l'activité émotionnelle qui n'est que partiellement influençable par la cognition.

Bien que la logique voudrait que les processus émotionnels s'expriment dans une région cérébrale bien précise, il n'en est rien. À l'instar de l'activité visuelle ou motrice ou encore langagière, il n'y a pas en réalité de centres émotionnels.

Joseph LEDOUX et Jeff MULLER préfèrent parler d'*unités cérébrales fonctionnelles* distinctes pour chaque émotion et résultant de l'évolution des espèces ([74]). Ces unités sont ensuite reliées entre elles pour former un système plus général en fonction de la contribution qu'elles apportent à ce dernier ; elles sont donc indispensables. Les chercheurs se sont penchés sur deux émotions simples, la peur et le plaisir, qui ont l'avantage d'être observables sur un grand nombre d'espèces animales.

4.2.1 Le circuit de la peur

Joseph LEDOUX a beaucoup apporté à la formalisation du circuit lié à l'émotion de peur. Il a ainsi pu démontrer le rôle très important que joue l'*amygdale*, petit noyau en forme d'amande situé dans la région antérieure de chaque lobe temporal (voir figure 4.5(a)) et connecté avec différentes zones du cerveau inhérentes à la peur. La spécificité de ce circuit réside dans sa dualité, car en effet, il existe deux circuits, l'un assez long et parcourant successivement le thalamus, le cortex cérébral et enfin l'amygdale, et l'autre, plus court, passant directement du thalamus à l'amygdale (voir figure 4.4). Dans un cas comme dans l'autre, c'est l'amygdale qui est responsable de l'expression corporelle (transpiration, accélération du pouls, paralysie du corps) lors d'un événement provoquant la peur.

LEDoux illustre ses propos ([72]) à l'aide de l'exemple d'un promeneur dans un bois qui aperçoit vaguement un bâton qui pourrait s'avérer être un serpent. L'image incertaine de ce bâton est alors perçue par le thalamus qui déclenche l'amygdale qui met en place les réactions corporelles liées à la peur. En parallèle, le thalamus envoie l'information au cortex visuel afin que celui-ci décrypte l'image de manière plus précise. Si le cortex visuel confirme l'image d'un bâton, le thalamus désactive l'amygdale, ce qui a pour effet d'atténuer les manifestations biologiques de la peur. Si au contraire le bâton s'avère être un authentique serpent, le cortex visuel renforce l'action de l'amygdale, laquelle maintiendra les manifestations physiologiques de la peur.

Il faut bien voir que l'émotion est un processus particulièrement rapide facilitant la survie de l'espèce. Dans l'exemple précédent, le promeneur a des chances de survie bien meilleures s'il assimile un vulgaire bâton à un serpent venimeux, plutôt que de réagir à la morsure d'une vipère avec quelques millisecondes de retard. Ainsi, le circuit court passant directement du thalamus à l'amygdale est privilégié dans les cas où la réaction se doit d'être rapide, par exemple la fuite ou encore la défense. D'ailleurs, il ne faut que 12 millisecondes seulement chez le rat pour qu'une stimulation sonore parvienne jusqu'à l'amygdale par la voie courte. En ce qui concerne le circuit long thalamus-cortex-amygdale, ce temps est multiplié par deux. En effet, bien que le passage par le cortex assure une analyse plus précise de l'information perçue, le chemin à parcourir étant allongé

en conséquence, la réponse arrive plus tardivement ce qui peut se révéler dommageable en cas de danger immédiat.

On comprend alors aisément les répercussions d'une éventuelle lésion de l'amygdale dont les victimes, d'après DAMASIO, « *sont incapables de se protéger contre les risques sociaux simples et moins simples.* » ([24]), car leur émotion est affectée au point de ne plus être capables de reconnaître la peur. Il est par conséquent bien plus difficile pour eux de détecter le danger et leur adaptabilité en société est alors radicalement remise en question.

4.2.2 Le circuit du plaisir

On associe souvent le circuit du plaisir à l'équivalent du système de récompense chez l'animal. En effet, c'est par le système hédonique que la volonté d'effectuer certains comportements vitaux est associée à un plaisir particulier dans le but de mettre en place ces comportements plus efficacement. En réalité, ce circuit correspond à une voie dite *dopaminergique*, c'est-à-dire qu'elle suit le parcours de la *dopamine*, neurotransmetteur produit entre autres par l'*aire tegmentale ventrale* (notée ATV sur la figure 4.6) et qui a pour cible une zone cérébrale particulière : le *nucleus accumbens septi* (qui signifie en latin « noyau appuyé contre le septum ») plus connu sous le nom *noyau accumbens* (fig. 4.6).

Tout événement plaisant tel que l'éloignement d'un prédateur ou la découverte de nourriture, active la voie dopaminergique qui déclenche la libération de dopamine au sein du noyau accumbens. La dopamine entre également en jeu dans l'ingestion des drogues (nicotine, alcool, cannabis, entre autres), lesquelles activent l'accumbens de la même façon. On constate par ailleurs que les drogues, au même titre que tous les phénomènes de dépendances et d'addiction, jouent un rôle de renforcement du système hédonique au point d'altérer celui-ci de manière à entraîner de fortes modifications comportementales. En effet, chez un sujet sain, le système hédonique fournit la motivation nécessaire à la réalisation d'actions optimisant la survie, tels que la reproduction ou l'alimentation, mais surtout il renforce ces agissements par l'activation d'une sensation de plaisir. La consommation de drogues détourne ce système et génère un comportement altéré, souvent au service de l'accoutumance et de la servilité.

Ainsi, nous pouvons voir que la neurologie actuelle a su déterminer le rôle de certaines zones du cerveau dans l'expression de l'émotion, ainsi qu'un ensemble de circuits neuronaux permettant la communication entre ces zones dans le but de faire naître les comportements adaptés. Tout ceci implique non seulement un lien direct entre l'émotion et certaines régions cérébrales spécifiques, mais aussi l'importance des interactions entre d'un côté le cortex préfrontal, et d'un autre côté l'amygdale et les structures limbiques de manière générale.

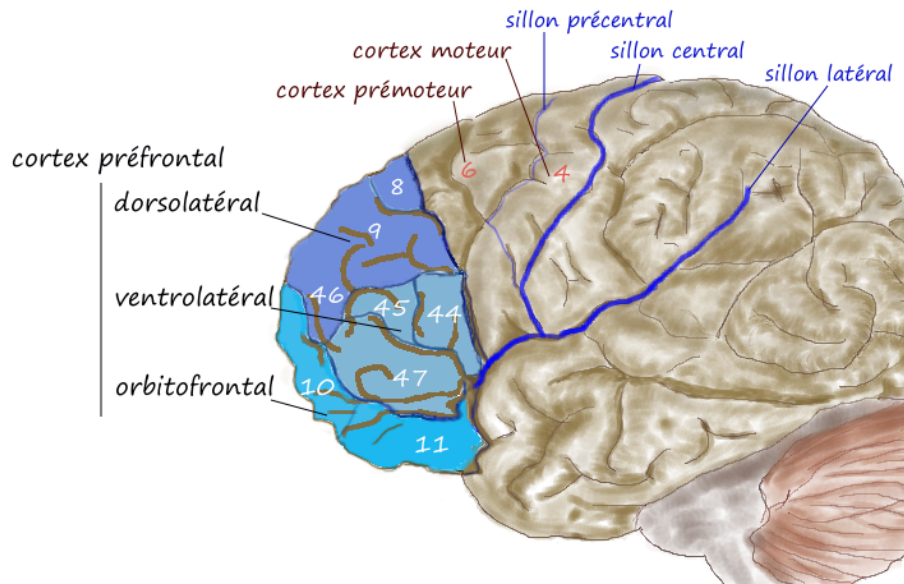
En outre, l'influence du cortex sur ces régions sous-corticales — et en particulier l'amygdale — n'a jamais cessé de croître tout au long de l'évolution

humaine ([76]). Cependant, les circuits partant du cortex vers l'amygdale sont bien moins nombreux que ceux qui partent de l'amygdale vers le cortex, ce qui explique probablement l'impact encore énorme de l'émotion sur la pensée et le raisonnement.

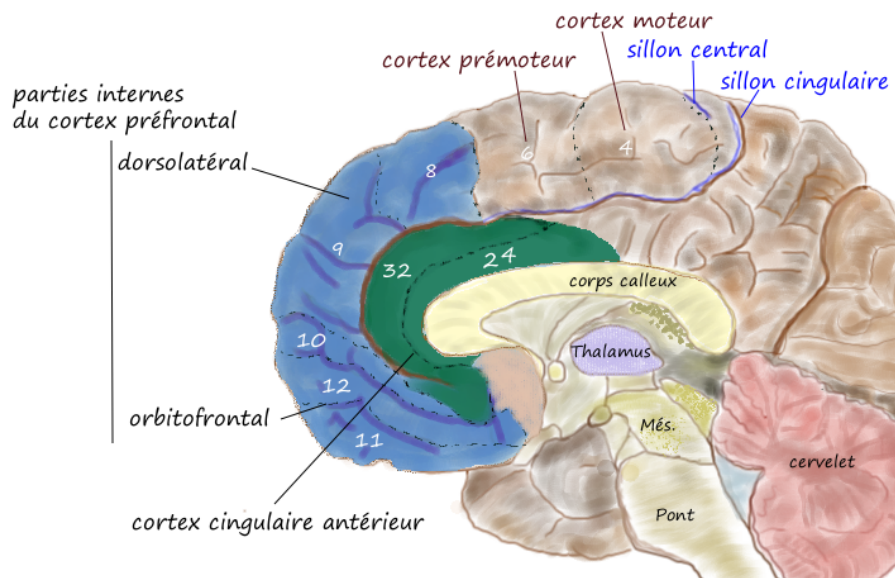
4.3 CONCLUSION

Ce chapitre permet de mettre en lumière les différents niveaux d'organisation dans lesquels s'inscrivent les processus émotionnels, à l'instar de nombreux autres systèmes complexes. En effet, l'émotion s'avère être d'une part génétiquement programmée, comme le suggère la présence de circuits neuronaux entre les différentes zones cérébrales participant à l'activité émotionnelle. D'autre part, ce câblage de bas niveau est accompagné d'une émergence émotionnelle comportementale, sous-tendue par l'expérience et les interactions environnementales. C'est l'action cumulée du substrat neuro-anatomique associée à cette dynamique émergente qui à court terme déterminera les décisions futures de l'individu, et à long terme, sa personnalité, comme l'illustre l'histoire de Phineas GAGE dont l'accident malheureux a considérablement perturbé ses capacités décisionnelles ainsi que ses traits de caractère.

L'informaticien doit donc penser la modélisation émotionnelle en deux temps. Tout d'abord déterminer une structure de base sur laquelle reposeront les processus émotionnels et éventuellement cognitifs. Ensuite, décrire la dynamique du système qui devra tenir compte du support afin de générer une activité auto-adaptative.



(a) Vue latérale



(b) Vue médiale

FIG. 4.3 – SCHÉMAS DU CORTEX PRÉFRONTAL

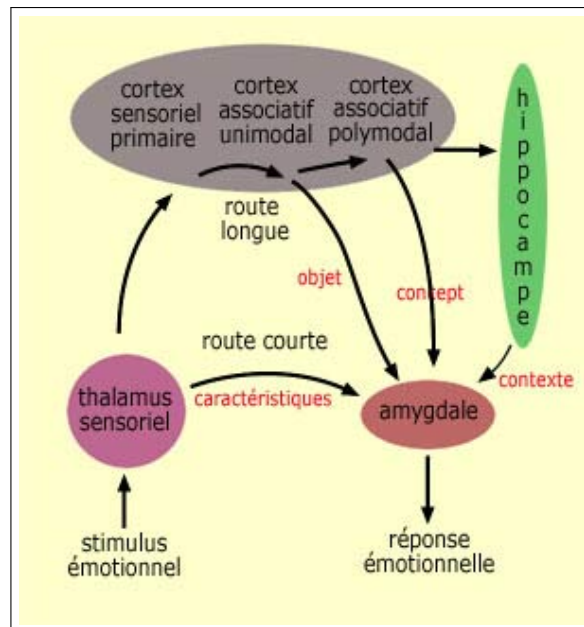
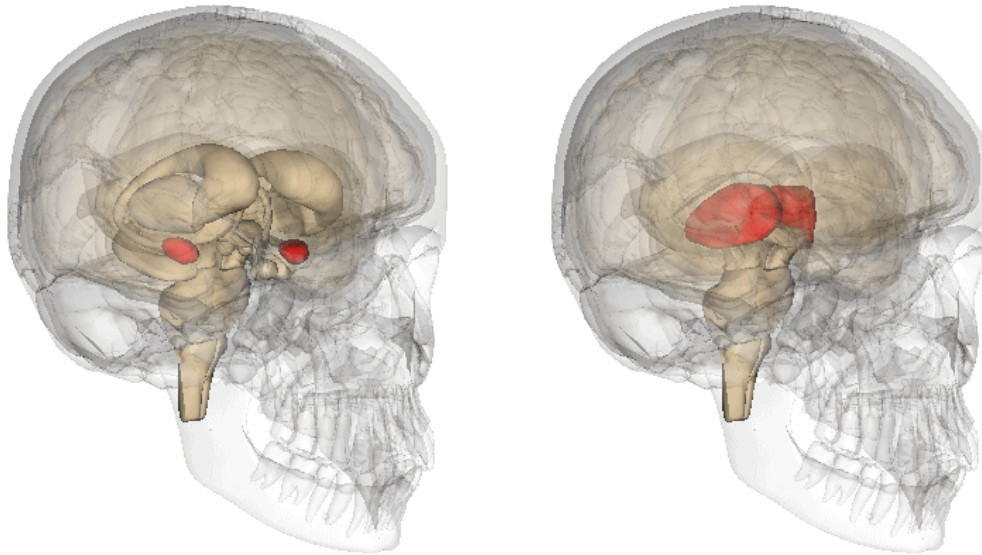
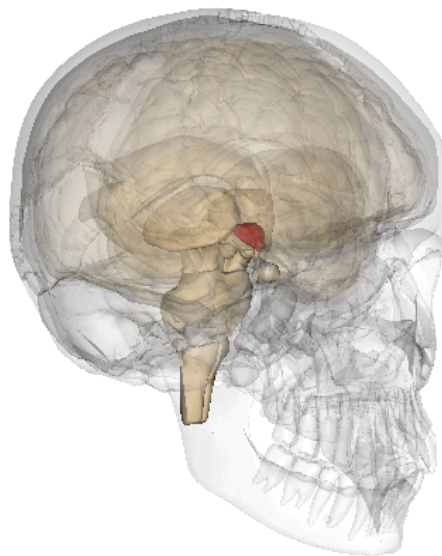


FIG. 4.4 – LES DEUX CIRCUITS DE PAPEZ
(Source : <http://lecerveau.mcgill.ca>)



(a) Amygdale

(b) Thalamus



(c) Hypothalamus

FIG. 4.5 – LOCALISATION DES ORGANES ASSOCIÉS AU CIRCUIT DE LA PEUR
(Source : site internet *Anatomography* maintenu par le *Life Science Databases*
URL : <http://lifesciencedb.jp>)

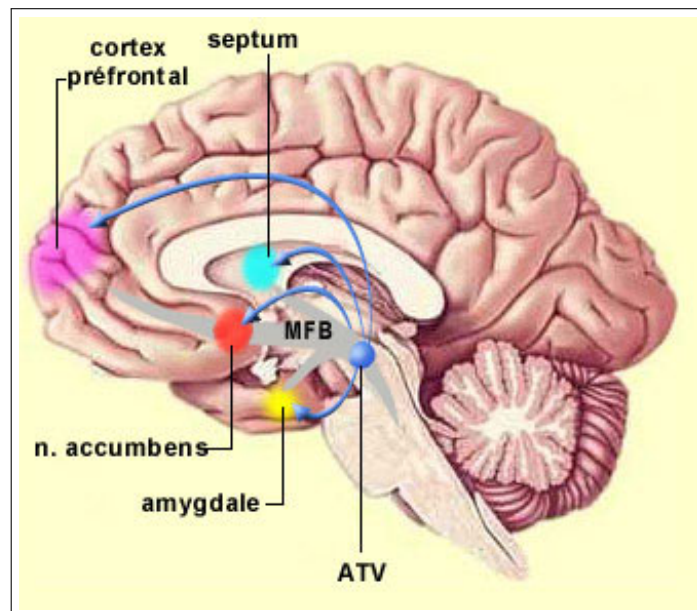


FIG. 4.6 – ZONES CÉRÉBRALES ASSOCIÉES AU SYSTÈME DE RÉCOMPENSE
(Source : <http://lecerveau.mcgill.ca>)

ÉMOTIONS ARTIFICIELLES ET MODÈLES INFORMATIQUES

5

SOMMAIRE

5.1	BILAN SUR LA REPRÉSENTATION DE L'ÉMOTION NATURELLE	75
5.2	NÉCESSITÉ DE LA MODÉLISATION DE L'ÉMOTION	76
5.3	MODÈLES INFORMATIQUES DE L'ÉMOTION	78
5.3.1	Le modèle EMA	79
5.3.2	Le <i>CogAff</i> de Aaron Sloman	80
5.3.3	La machine pensante de Alain Cardon	81
	5.3.3.1 Système générateur d'émotions	82
	5.3.3.2 Système générateur de pensées	84
5.3.4	Le modèle <i>Cathexis</i> de Velásquez	85
5.4	CONCLUSION SUR L'ÉMOTION ARTIFICIELLE	86
5.4.1	Réflexions sur l'émotion	87
5.4.2	Définition de l'émotion artificielle	87

*Patience et longueur de temps
font plus que force ni que rage.*

— *Le loup et le rat*, Jean de LA FONTAINE

Malgré l'engouement scientifique autour de l'émotion tout au long du xx^e siècle, peu de programmes informatiques modélisant l'émotion de façon artificielle ont été développés, comparativement au grand nombre de modèles décrits en sciences humaines. Une des causes de l'incompatibilité apparente entre émotion et informatique est sans doute historique. Depuis les débuts de l'intelligence artificielle de la seconde partie du xx^e siècle, l'émotion n'a guère réussi à trouver sa place. Le mouvement profondément cognitiviste apparu à cette époque se voulait traditionnellement centré autour du pur raisonnement

rationnel, notamment avec l'apparition du *General Problem Solver (GPS)* en 1957 par Herbert SIMON, Cliff SHAW, et Allen NEWELL ([41]). Cependant, ce modèle ne s'applique qu'à la résolution d'un nombre très limité de problèmes, d'où la nécessité de rechercher de nouveaux formalismes, parmi lesquels les *systèmes experts*. Ces derniers furent utilisés principalement pour l'aide à la décision dans des domaines très spécifiques tels que le diagnostic médical — avec par exemple des modèles comme *Mycin* (1975, [123]) — au moyen de bases de connaissances et de règles dictées par des experts scientifiques du domaine concerné (GIARRATANO et RILEY, 2004, [48]).

À cette époque, les recherches étaient globalement axées sur les aspects cognitifs — apprentissage et stockage de l'information entre autres — par l'utilisation de symboles et de systèmes d'inférences manipulant ces symboles. L'émotion était alors considérée comme un « parasite », une faiblesse humaine, face à la cognition qui, elle, détenait la connaissance et la raison. Il faudra attendre les années 90, avec notamment les découvertes en neuropsychologie d'Antonio DAMASIO ([23]), pour que l'informatique s'intéresse de plus près aux émotions, grâce aux avancées majeures des techniques d'observation en neurologie.

Comme l'indique le schéma récapitulatif de la figure 5.1, la plupart des modèles s'inscrit dans une logique cognitive. Ainsi, ils étudient tout particulièrement les liens entre émotion et cognition dans le cadre de la prise de décision. Toutefois, d'autres types d'approches se développent dans le but d'observer le phénomène sous d'autres angles : d'une part une perspective communicationnelle qui vise à observer les manifestations extérieures de l'émotion pour en déterminer des règles de communication, et d'autre part, une idée plus *émergentiste* de l'émotion par l'étude des mécanismes biologiques impliquant les processus émotionnels à des niveaux supérieurs.

Toujours est-il que tous ces modèles voient se développer une véritable volonté de formaliser l'émotion et de l'incorporer au sein des différents modèles développés, au même titre que les processus cognitifs. Un langage spécifique pour décrire les émotions indépendamment du modèle utilisé a même été conçu pour l'occasion : *EmotionML* (pour *Emotion Markup Language*, [5]).

Ce chapitre a pour vocation de présenter les différents types d'approches modélisant l'émotion artificielle en informatique. Dans un premier temps sera traitée la question de la nécessité de modéliser l'émotion, au travers de quelques pistes de réflexion. Ensuite, nous verrons quelques modèles de l'émotion, dont la plupart sont issus des travaux de la perspective cognitive et des théories de l'*appraisal*, mais pas uniquement. En troisième lieu, nous terminerons ce tour d'horizon par une conclusion générale sur l'émotion artificielle. Mais tout d'abord, pour fixer les idées, faisons un petit bilan sur l'émotion naturelle !

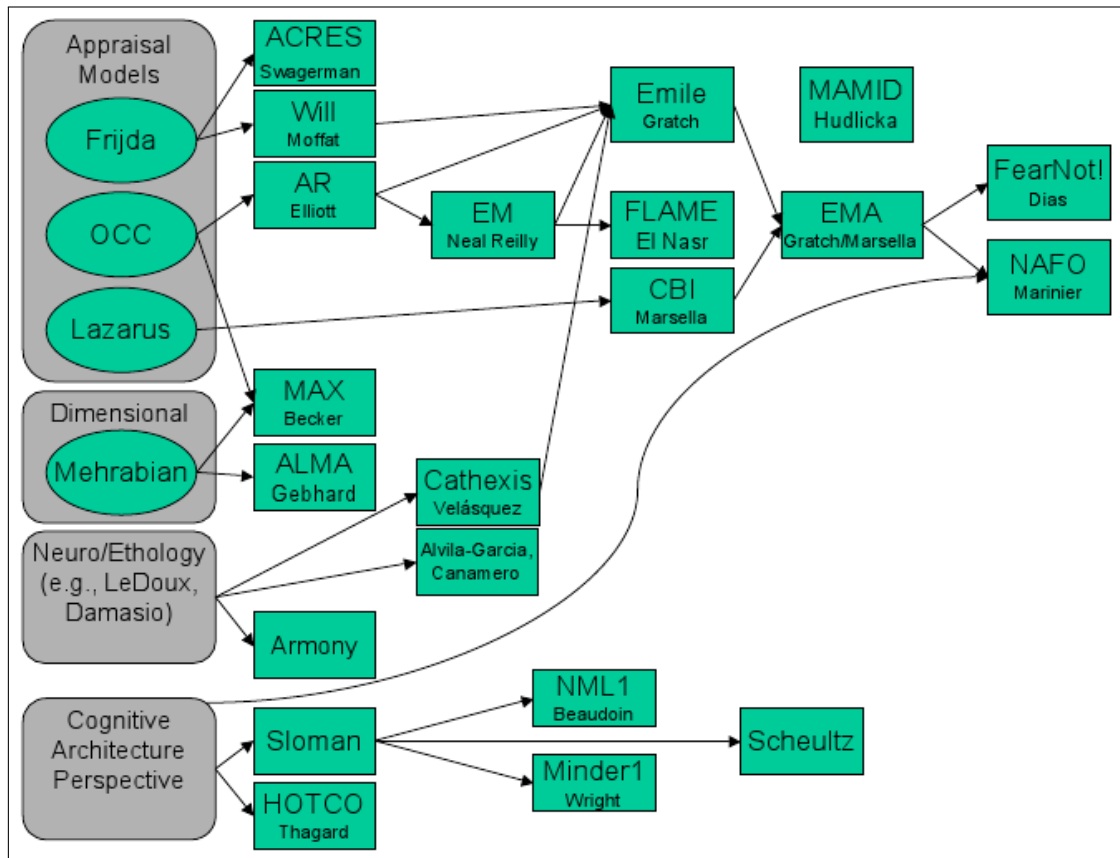


FIG. 5.1 – PRINCIPAUX MODÈLES INFORMATIQUES DE L'ÉMOTION
(GRATCH, [54])

5.1 BILAN SUR LA REPRÉSENTATION DE L'ÉMOTION NATURELLE

De mon point de vue d'informaticien, j'ai décidé d'insister sur un certain nombre de notions importantes afin d'établir ma propre définition de l'émotion. Tout d'abord, si l'origine de l'émotion est bien le système nerveux, il ne s'agit visiblement pas d'une émergence directement issue de l'héritage génétique, car si tel était le cas, une grande partie des « routes de l'émotion » parcourant le cerveau auraient clairement été identifiées. En réalité, quand bien même certains circuits tels que celui de la peur sont aujourd'hui connus, l'écrasante majorité des processus émotionnels complexes restent diffus au sein du cortex cérébral, car leur évolution est intimement liée à l'expérience de chacun.

Émotion

Ensemble de processus automatiques, parfois réflexes, permettant à un individu de manifester des comportements auto-adaptatifs, conjointement à son activité cognitive. L'émotion peut se déclencher suite à des perceptions externes provenant de changements environnementaux mais aussi en réponse à des perturbations internes de l'organisme. La structure du fonctionnement émotionnel est partiellement transmise génétiquement sous la forme de « routes de l'émotion » dédiées, mais son fonctionnement est essentiellement le fruit de l'apprentissage issu de l'expérience de l'individu.

Cette définition permet de résumer les principaux processus sous-jacents qui entrent en jeu lors de l'expérience émotionnelle. Dans le cadre de ce travail de thèse, elle nous servira de base conceptuelle à la création de modèles en informatique. Cependant, et avant d'entamer un quelconque travail de modélisation, il nous reste à répondre à une question primordiale : pourquoi cherche-t-on à modéliser l'émotion ?

5.2 NÉCESSITÉ DE LA MODÉLISATION DE L'ÉMOTION

Comme nous l'avons vu précédemment, l'émotion possède deux formes différentes :

- une forme « externe », comportementale, basée sur des signaux émotionnels circulant entre deux ou plusieurs individus communicants ;
- une forme « interne », plus fonctionnelle, selon laquelle, d'après les récentes études, les mécanismes émotionnels seraient indispensables dans le processus de prise de décision.

Ces deux visions antagonistes, bien que reflétant le même processus, n'ont pas la même portée, ce qui crée parfois quelques confusions scientifiques en termes de compréhension des processus émotionnels selon le point de vue adopté.

D'une part, certains auteurs réfutent l'intérêt de la modélisation des émotions, à l'instar de John MCCARTHY : « *Robots should not be equipped with human-like emotions* » ([89])¹ dans la mesure où cela induirait une complexité supplémentaire à la compréhension du comportement du robot dans sa relation avec les êtres humains. Ce souci d'apparition d'une complexité additionnelle est purement comportemental et ne se préoccupe que du lien extérieur humain-robot

1. « Les robots ne devraient pas être équipés d'émotions humaines » (NDR)

dans la perspective d'une utilité potentielle de l'émotion. D'autre part, dans cette même perspective, d'aucuns considèrent que l'émotion est indispensable à une communication plus naturelle entre humains et robots, à l'image de l'*Affective Computing* ([101]), car elle introduit un mécanisme d'empathie chez le robot, qui lui permet d'interpréter les manifestations émotionnelles de son interlocuteur humain, afin de mieux adapter sa réponse et de produire un comportement plus approprié. On comprend bien ici que d'un point de vue uniquement communicationnel (externe), l'implantation d'une mécanique émotionnelle chez un robot, bien que complexe, s'avère nécessaire.

Contrairement aux affirmations de MCCARTHY, Marvin MINSKY adopte un point de vue plus fonctionnel (interne) en affirmant qu'on ne peut créer une intelligence comparable à celle des êtres humains sans y intégrer l'émotion ([92]) car celle-ci serait indispensable aux mécanismes internes de prises de décisions. Cette idée s'oppose à la vision traditionnelle qui prétend que l'émotion est un parasite à la décision sous-tendue par la raison. En d'autres termes, le robot parfait se devrait de prendre les décisions les plus pertinentes possibles sans faire preuve d'aucun mécanisme émotionnel, ce que la psychologie appelle « cognition froide ». Mais alors qui a raison ? DAMASIO a répondu partiellement à la question en démontrant expérimentalement que chez l'être humain, les processus émotionnels sont indispensables à la prise de décision ([22], [23]).

Finalement, cette dualité *externe/interne* explique doublement, si ce n'est la nécessité, l'intérêt que la science peut avoir à s'intéresser de près aux mécanismes de l'émotion, ne serait-ce que pour apporter à certains domaines de recherche un nouvel éclairage. Par exemple, une meilleure prise en charge de patients dépressifs ou une possible intégration des résultats des modèles de l'émotion en neurologie afin d'expliquer les liens potentiels entre zones topographiques du cerveau, tels pourraient être des arguments en faveur de l'étude scientifique de l'émotion.

Malgré l'apparente évidence du rôle crucial de l'émotion comme partie intégrante de l'intelligence humaine (et animale), il nous faut nous poser la question une bonne fois pour toutes : est-il intéressant de représenter l'émotion dans nos modèles informatiques d'aide à la décision ? Ma réponse est *oui* ! Au-delà de la simple curiosité scientifique, il me semble que l'émotion doit être étudiée de près et implantée sur des robots, ne serait-ce que pour pouvoir découvrir si ces robots ont un meilleur comportement que leurs équivalents « purement cognitifs » et s'ils sont ou non mieux adaptés à leur environnement. De plus, les résultats produits seraient en mesure de confirmer ou d'infirmer des hypothèses ou théories formulées dans des domaines tels que la psychologie, la philosophie ou encore l'intelligence artificielle.

5.3 MODÈLES INFORMATIQUES DE L'ÉMOTION

La recherche scientifique actuelle a démontré jusqu'à présent le rôle crucial de l'émotion dans l'adaptation d'un individu à son environnement en termes de processus cognitifs et de régulation des interactions sociales ([118]). Dans ce contexte, *l'intelligence émotionnelle* ([112]) représente justement cette capacité, que possèdent de nombreuses espèces animales, de comprendre naturellement les mécanismes émotionnels, leurs influences, leurs conséquences, mais aussi d'être capables de contrôler son état émotionnel afin de mieux s'intégrer socialement et atteindre plus facilement ses objectifs ([49]).

Dans cette optique, les recherches en informatique se sont penchées sur la question de l'émotion dans le but d'augmenter les performances computationnelles avec, en ligne de mire, la résolution de problèmes complexes. Ainsi est née *l'informatique affective* (ou *affective computing*) dont l'intérêt est porté sur la modélisation de l'émotion ainsi que la conception de systèmes permettant de reconnaître, d'exprimer et de synthétiser des émotions ([101]).

L'informatique affective distingue deux perspectives d'étude. La première s'inscrit dans une logique de reconnaissance des émotions d'un utilisateur par l'utilisation et l'analyse de données physiologiques et somatiques à l'aide de capteurs. Ces données peuvent ainsi mesurer par exemple les signaux visuels (expressions faciales, comportement, ...), les indices auditifs (hauteur et intonation de la voix), la réponse électrodermale, les pulsations cardiaques, ...

La deuxième perspective d'étude étudie plus spécifiquement la modélisation de l'émotion au sein d'un agent virtuel ou robotisé, et son impact sur le comportement de cet agent. Les recherches se portent alors sur les origines du déclenchement de l'émotion, ou bien les influences entre émotion et cognition afin d'augmenter l'efficacité comportementale, ou encore l'expression émotionnelle qui vise à transmettre les caractères de l'émotion, d'un programme à son utilisateur.

Cette deuxième perspective est celle qui nous concerne dans le cadre de ce travail de thèse. Cependant, quand bien même elle nous permet, par définition, de mieux appréhender les influences comportementales et décisionnelles, de nombreux modèles utilisent cette approche dans le but de construire un agent virtuel plus réaliste, plus crédible, pour faciliter la communication homme-machine. Il est clair que la vraisemblance d'un agent apporte des avantages non négligeables, notamment dans les cas où les interactions entre machine et utilisateur sont primordiales. Pour autant, ce paramètre doit rester secondaire face à la nécessité de concevoir une émotion interne la plus authentique et la plus riche possible.

Les quelques modèles informatiques présentés dans cette section sont assez représentatifs des différents points de vue computationnels adoptés ces quinze dernières années. Le modèle EMA de GRATCH et MARSELLA s'inspire en grande

partie des travaux en psychologie autour de l'*appraisal*. Le *CogAff* de SLOMAN, quant à lui, est issu des théories cognitivistes qui tiennent compte des différentes composantes cognitivo-émotionnelles entrant en jeu dans le processus de décision. La machine pensante de Alain CARDON présente elle aussi un modèle constitué de plusieurs niveaux d'organisation des émotions, des cognitions, et plus généralement des processus de pensée. Enfin, le modèle *Cathexis* de VELÁSQUEZ se place davantage dans la lignée de la perspective darwinienne en proposant un système décrivant les émotions à un niveau plus profond, proche des stimuli environnementaux. Voici une description de ces modèles.

5.3.1 Le modèle EMA

Proposé par Jonathan GRATCH et Stacy MARSELLA ([53]), le modèle EMA (pour *EMotion and Adaptation*) est un modèle cognitif générique qui plonge un agent virtuel de type BDI² dans un environnement complexe qui met en avant la construction émotionnelle ainsi que l'interaction sociale. Le système comporte deux processus principaux : un processus d'évaluation (*appraisal*) et un processus d'adaptation (*coping*). Le processus d'évaluation permet de déterminer les émotions déclenchées par un événement, alors que le processus d'adaptation décide quel est le comportement à adopter pour faire face aux émotions de forte intensité.

À partir de l'idée centrale de *fenêtres* et de *variables d'appraisal*, grâce auxquelles les valeurs émotionnelles des événements ainsi que des processus internes et externes sont capturées, le modèle génère des états émotionnels qui seront ensuite assemblés pour produire une humeur globale (*mood*). Pour ce faire, le programme utilise les capacités de raisonnement offertes par sa structure BDI, laquelle repose sur des concepts tels que la probabilité (*likelihood*) ou la désirabilité (*desirability*). Le composant d'évaluation est basé sur les travaux de Jonathan GRATCH et son modèle *Émile* ([51], [52]), ce dernier étant adapté de l'*Affective Reasoner* de Clark ELLIOTT ([39], [40]), qui lui-même est inspiré du modèle OCC ([98]). Grâce à *Émile*, GRATCH propose un modèle de raisonnement émotionnel dans lequel le concept d'*évaluation basée sur des plans* (*plan-based appraisal*) est central. Il s'agit en fait d'un mode d'évaluation dans lequel différents plans sont chargés en mémoire et organisés de façon générique. Par ce biais, la relation entre les événements extérieurs et l'état interne de l'agent est par nature indépendante du domaine concerné.

Par ailleurs, le modèle EMA permet de mettre en place diverses stratégies d'adaptation. Ainsi, face à une émotion négative, l'individu peut opter pour une stratégie de réinterprétation positive au cours de laquelle il tente de se concentrer sur les conséquences positives de l'événement ayant déclenché l'émotion négative. Il peut aussi opter pour une stratégie dite d'acceptation où l'événement

2. qui signifie en anglais *Belief-Desire-Intention* (croyance-désir-engagement)

néгатif est considéré comme inévitable. L'individu choisit alors d'abandonner son intention première qui initialement a déclenché l'émotion négative.

L'agent virtuel peut aussi adopter plusieurs stratégies d'adaptation simultanées, chacune d'entre elles permettant de modifier la situation actuelle en constante réévaluation, ce qui déclenche d'autres émotions successives, et ainsi de suite, dans une approche auto-adaptative. Par ailleurs, le modèle EMA a été intégré au sein d'un ensemble de soldats artificiels dans le cadre d'un environnement virtuel d'entraînement militaire ([53]).

5.3.2 Le *CogAff* de Aaron Sloman

Aaron SLOMAN, philosophe à l'université de Birmingham, défend l'idée qu'un système informationnel n'a pas besoin de corps pour être émotionnel. Il a conçu un système cognitif appelé *The Cognition and Affect Project* ou *CogAff* ([125]). Les émotions y sont considérées comme structures de contrôle qui viennent interrompre une tâche en cours, afin de concentrer les ressources attentionnelles sur une tâche plus urgente. On déclenche alors un signal d'alarme qui permet au système de passer au-dessus de tout processus provenant de n'importe quel module du modèle.

La figure 5.2 décrit le *H-CogAff* ou *Human CogAff*, présenté comme un cas particulier du *CogAff* appliqué à l'homme et dans lequel on distingue les trois couches nécessaires à l'élaboration du modèle :

- Les composants réactifs (*reactive components*) propres à toute espèce réactive telle que l'insecte. Cette couche s'occupe de la perception des différents événements de l'environnement et y répond automatiquement. Par conséquent, il s'agit du siège des émotions primaires. C'est ici que l'on retrouve le système d'alarme mentionné précédemment ;
- Les mécanismes volontaires (*deliberative mechanisms*) pour la considération d'événements non existants. Ils représentent ce que SLOMAN appelle les *what-if mechanisms*. Peu d'êtres vivants sont doués de ces facultés. À ce niveau sont représentés les processus rationnels pour la constitution de plans, la prise de décision et l'allocation de ressources. C'est le siège des émotions secondaires ;
- Les capacités dites de *meta-management* pour l'examen, l'évaluation et l'influence des informations internes. Ce sont des processus réflexifs impliquant l'observation et le contrôle de soi que seule une poignée d'espèces possède. SLOMAN y définit les émotions tertiaires, sortes de méta-émotions, telles que la culpabilité qui va orienter l'attention vers soi-même ([124]).

SLOMAN décrit ainsi un modèle de type incrémental, en couches, dans lequel l'homme tient la plus haute place, au sommet de la pyramide de la complexité émotionnelle.

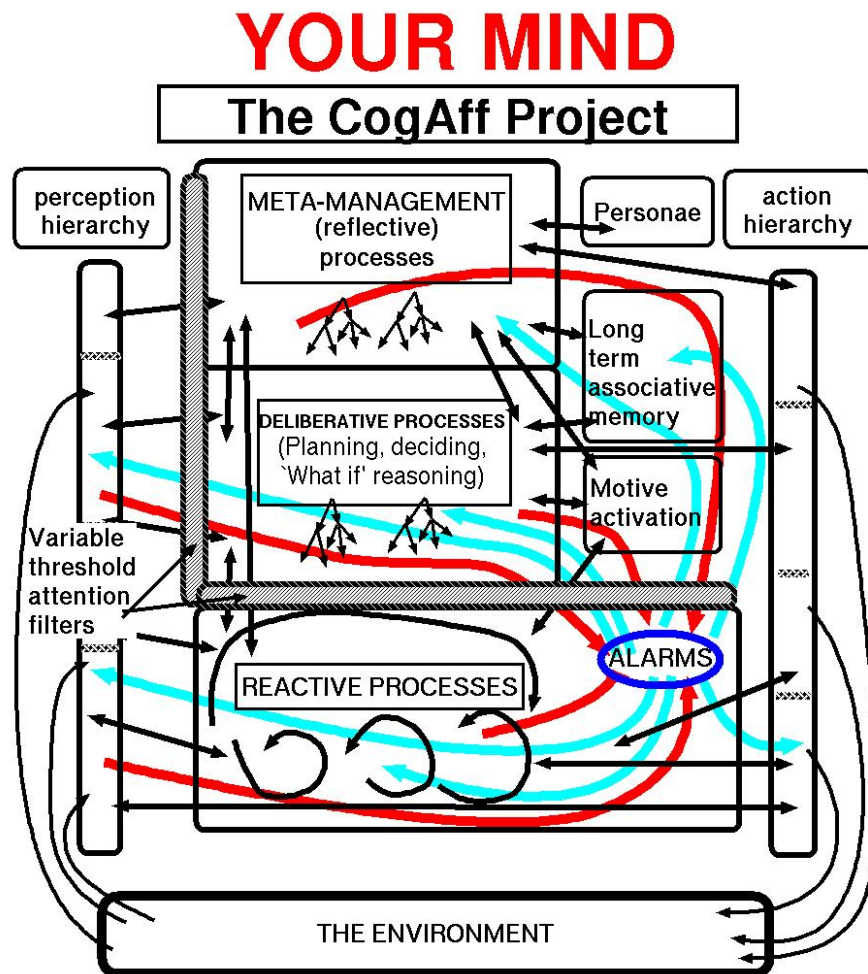


FIG. 5.2 – LE MODÈLE H-COGAFF DE AARON SLOMAN

5.3.3 La machine pensante de Alain Cardon

Pour Alain CARDON, ancien professeur d'informatique à l'université du Havre et ancien membre permanent du laboratoire d'informatique de Paris 6 (LIP 6), le cerveau humain est incapable de produire des connaissances strictement objectives, car le système méso-limbique est toujours activé et lié au néocortex. Par conséquent, les processus de pensée sont systématiquement emprunts d'émotions. De plus, la connaissance est liée à une intention de produire du sens sur quelque chose à un moment donné : elle est fortement contextualisée. Ainsi, l'idée est de produire un système qui génère de la connaissance intentionnelle et située, donc dotée d'émotions.

Dans le cadre de la conception d'une machine pensante, Alain CARDON décrit une activité artificielle cérébrale dans son exécution, son fonctionnement, dans la perspective de placer le phénomène dans le domaine du calculable, à

différentes échelles d'organisation ([18], [19]). Pour ce faire, il utilise un système multi-agents composé de deux modules principaux que sont un *système générateur d'émotions* et un *système générateur de pensées*, liés à l'environnement par un ensemble d'*éléments d'interface*, à savoir des capteurs et effecteurs (voir figure 5.3).

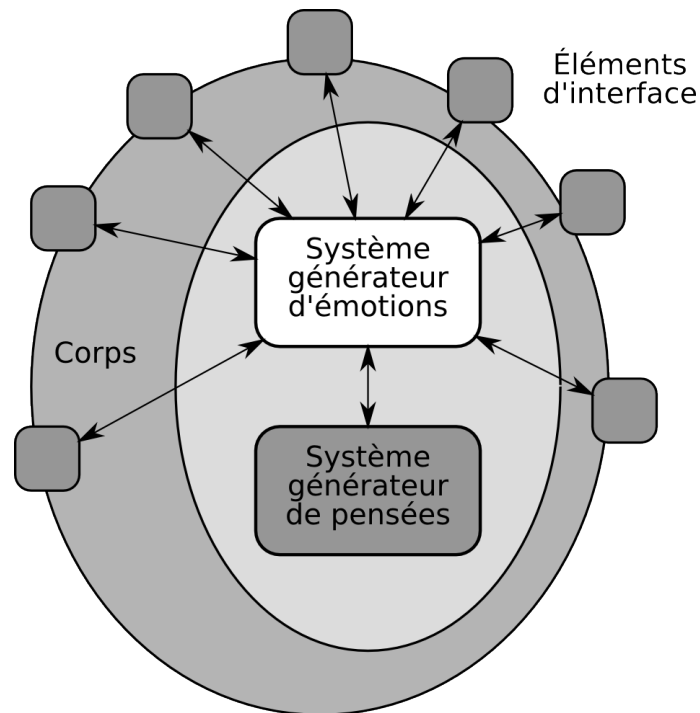


FIG. 5.3 – LE CORPS DE LA MACHINE PENSANTE ET SES DEUX SYSTÈMES GÉNÉRATEURS D'ÉMOTIONS ET DE PENSÉE ([18])

5.3.3.1 Système générateur d'émotions

Alain CARDON définit que le corps et la conscience sont indissociables de façon absolue, c'est-à-dire qu'on ne peut construire une conscience sans qu'elle ne soit intimement couplée à un substrat matériel qui interagit avec elle et son environnement. Ainsi, l'on se doit de posséder tout d'abord, un robot composé de beaucoup de capteurs et d'effecteurs, nous permettant de créer un lien avec l'extérieur. Le système générateur d'émotions s'appuie sur la corporéité même du robot ainsi que sur son organisation interne, afin de générer un ensemble de *cartes corporelles* composées d'*agents logiciels*. L'élément de base de la représentation est l'*agent aspectuel*.

Un agent aspectuel est une entité logicielle proactive, communicante, évolutive, qui va éventuellement prendre des informations provenant de certains capteurs sensoriels et en envoyer à des effecteurs. Cet agent aura la tâche de

représenter un caractère symbolique et d'agir selon ses intérêts propres (voir figure 5.4).

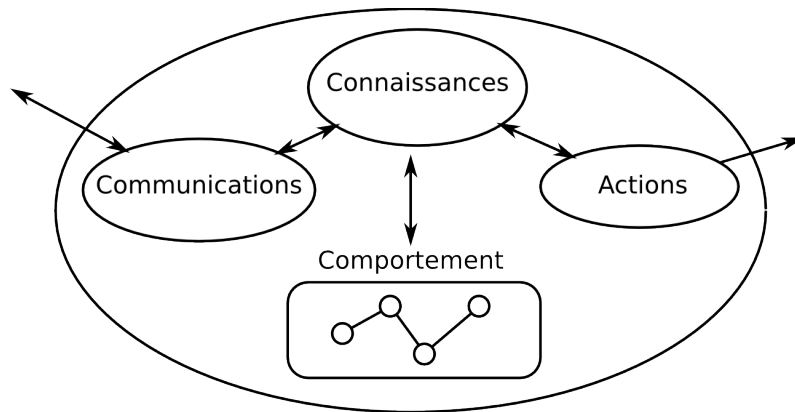


FIG. 5.4 – STRUCTURE GÉNÉRALE D'UN AGENT ASPECTUEL

Les agents aspectuels ont la particularité, comme beaucoup d'autres types d'agents logiciels, de former entre eux des agrégats d'entités fortement communicantes. Néanmoins, de simples organisations d'agents ne suffisent pas à représenter les caractères d'une émotion artificielle, et pour augmenter l'étendue de la représentation aspectuelle, une structure d'auto-observation a été greffée sur ces agents, elle-même composée d'agents dits *agents morphologiques* qui vont permettre au système de réifier non seulement les caractères propres des agents, mais aussi leurs actions organisationnelles. Un agent morphologique aura alors pour fonction d'analyser le comportement des agents aspectuels, afin d'en dégager la géométrie selon un certain nombre de critères (voir figure 5.5).

L'idée sous-jacente de cette structure d'agents à plusieurs niveaux est de réaliser des cartes émotionnelles composées des agents aspectuels, des agents morphologiques, ainsi que d'un couplage fort entre ces deux groupes d'agents appelé *processus miroir*. Une carte émotionnelle a donc l'avantage de s'auto-contrôler grâce à l'intervention des agents de morphologie, de s'adapter, et surtout de se synchroniser avec d'autres cartes. En effet, elle a la particularité d'être spécialisée et c'est l'union de milliers de cartes du même type, mais pour d'autres fonctionnalités, qui permettra d'avoir une vision générale du mécanisme émotionnel, la synchronisation des cartes émotionnelles entre elles se faisant à l'aide d'agents de synchronisation. Sur la figure générale de l'organisation des agents (figure 5.6), les *agents d'interface* sont les agents directement en liaison avec les capteurs et effecteurs, et les *agents d'évocation* s'occupent de générer du sens lié à l'auto-apprentissage et à la mémorisation.

Aspect externe de l'agent	Suprématie
	Indépendance
	Persistance
Aspect interne de l'agent	Facilité
	Vélocité
État organisationnel de l'agent	Intensité du flux interne d'activation
	Complexification
	Fréquence communicationnelle
	Écart organisationnel
Ouverture du système	Transport d'information

FIG. 5.5 – DIMENSIONS DE L'ESPACE MORPHOLOGIQUE

5.3.3.2 Système générateur de pensées

En parallèle, le système générateur de pensées est une architecture logicielle analogue à celle du système générateur d'émotions, mais plus complexe. Comme le disait NIETZSCHE : « *Une pensée ne me vient pas quand je veux mais quand elle veut* ». Le modèle est donc conçu comme ayant la volonté de penser et se voit doté d'une mémoire incrémentielle et sélective. Il possède en outre une structure proactive, auto-observatrice et évolutive au même titre que la composante émotionnelle avec laquelle il est couplé. Son but est de transcrire de façon sémantique et langagière les informations venues du système émotionnel, ainsi que ses connaissances acquises

Le système général est donc finalement composé d'un ensemble de composantes dynamiques synchronisées par des agents particuliers. Ainsi, l'architecture du système entier est en constante évolution, en parallèle avec le corps. C'est en quelque sorte une co-construction avec le robot lui-même.

En conclusion, l'idée de Alain CARDON est de concevoir un modèle que lui-même qualifie de « système multi-agents massif » dont le principe consiste à créer un nombre considérable d'agents, ayant leur rôle et leurs objectifs propres, et de les organiser à différentes échelles. Le système trouvera ses motivations conscientes dans l'origine même de ses tendances comportementales initiées par les agents.

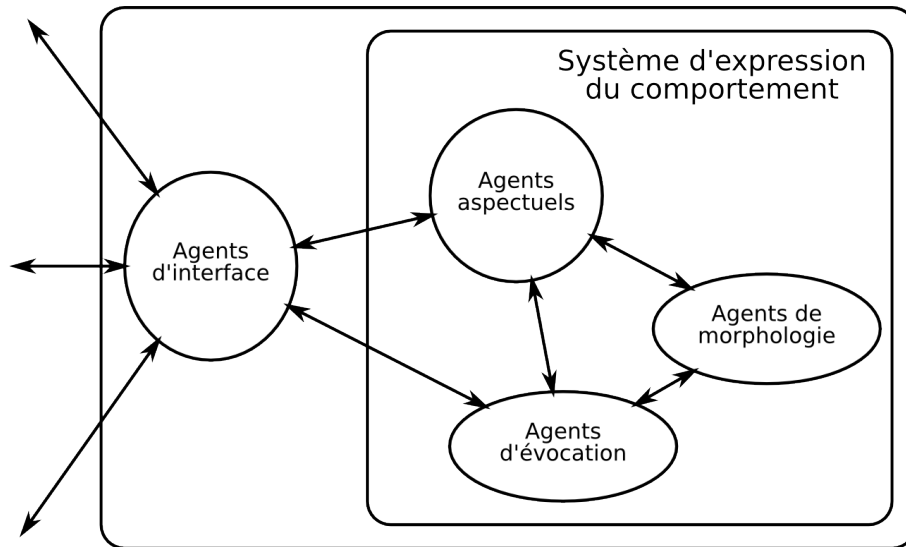


FIG. 5.6 – ORGANISATION DES AGENTS

5.3.4 Le modèle *Cathexis* de Velásquez

Dans son approche, Velásquez ([129], [130]) présente un modèle computationnel d'émotions basiques et complexes dont la dynamique met l'accent en particulier sur la dépendance au temps ainsi que l'influence des émotions sur le comportement et la motivation. L'architecture de *Cathexis* distingue les *émotions* de l'*humeur* et du *tempérament* autour de l'utilisation de quatre sous-systèmes : le sous-système comportemental, le sous-système de la génération d'émotions, le sous-système moteur et le sous-système de stimuli internes. Ces sous-systèmes sont exposés à des stimuli externes provenant de l'environnement et le système réagira par des actions sur cet environnement. *Cathexis* met une emphase particulière sur l'intensité de l'émotion générée.

Les émotions sont modélisées sous forme d'un réseau de *proto-spécialistes*. Chacun de ces spécialistes représente une famille d'émotions spécifiques, c'est-à-dire une catégorie parmi les émotions primaires ou basiques — à savoir la colère, la peur, la tristesse, la joie, le dégoût et la surprise —, et observe les stimuli externes et internes à travers des capteurs qui influent sur l'intensité de ses émotions. Lorsque cette intensité dépasse un certain seuil, la sortie du spécialiste est activée vers le système de comportement mais les autres spécialistes au sein du réseau peuvent aussi l'inhiber ou l'activer.

VELÁSQUEZ distingue les émotions des humeurs en termes de niveaux d'activation, c'est-à-dire que les humeurs sont considérés comme des affects possédant des niveaux d'activation inférieurs à ceux des émotions. En ce qui concerne les différents tempéraments, ils sont modélisés par des seuils différents d'activation des émotions. Plus tard, VELÁSQUEZ a ajouté à *Cathexis* l'apprentissage

d'une conduite émotionnelle suivant les événements soumis à l'individu afin de mémoriser les causes d'une émotion ([131]).

L'architecture *Cathexis* a fait l'objet de deux implantations robotiques (voir figure 5.7) : une sur un agent nommé YUPPY ([132]) et l'autre au sein d'un robot social plus expressif nommé KISMET ([16]).

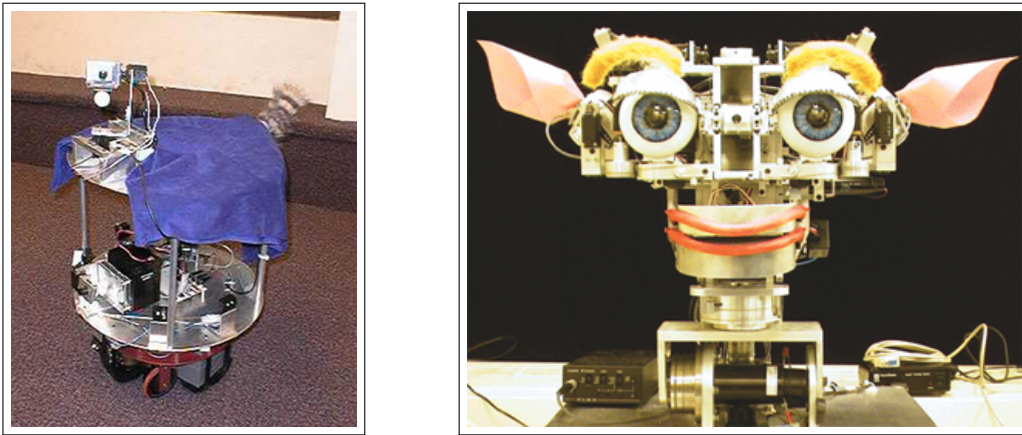


FIG. 5.7 – LES DEUX ROBOTS YUPPY ET KISMET

5.4 CONCLUSION SUR L'ÉMOTION ARTIFICIELLE

À l'instar des théories développées en psychologie, les modèles présentés ci-dessus ont tous pour objectif de décrire un fonctionnement émotionnel qui tend à se rapprocher de l'émotion naturelle. La différence se situe dans la réalisation effective de programmes simulant ce fonctionnement sur des machines, et introduisant dès lors une rupture conceptuelle, nous faisant passer d'émotion naturelle à émotion artificielle. L'exercice qui consiste à modéliser l'émotion naturelle informatiquement pose un ensemble de problèmes étrangers à ce que l'on observe chez l'être humain. En effet, même si nous avons précédemment considéré l'émotion humaine comme une émergence de l'activité cérébrale, inhérente à la structure même des réseaux neuronaux, il est bien entendu inconcevable de recréer les conditions de cette activité au sein d'un ordinateur, tant la complexité à représenter serait colossale, aussi bien qu'incontrôlable. La corporéité dans laquelle s'inscrit l'émotion étant tout à fait indissociable de l'émotion elle-même, il me paraît aussi bien difficile qu'inutile de tenter l'aventure de la modélisation d'un réseau neuronal authentique, en espérant simuler un jour un ersatz d'émotion analogue à ce qui est observé chez *Homo sapiens* que nous sommes.

5.4.1 Réflexions sur l'émotion

D'un point de vue phénoménologique, il me paraît difficile d'étudier un objet que l'on ne peut mesurer qu'indirectement. En effet, les processus émotionnels ne sont globalement observables que par l'intermédiaire des expressions faciales qu'ils déclenchent, ainsi que par leurs facultés à guider nos décisions et nos actes, ces deux types d'observations pouvant plus ou moins être évaluées d'un point de vue physiologique et psychologique.

Cependant, la communauté scientifique considère que la création d'un modèle de l'émotion passe nécessairement par l'analyse complète de ces manifestations. Bien entendu, loin de moi l'idée de considérer que ces manifestations n'existent pas ou que leur étude est inutile, bien au contraire. Seul le développement des connaissances sur l'Homme et sa dynamique nous permettront d'avancer dans la compréhension de l'intelligence humaine. Néanmoins, je considère que l'on cherche à modéliser un phénomène ne possédant pas de réalité explicitement observable, tant il est diffus au sein de la mécanique complexe cérébrale. En somme, il n'y a pas d'émotion, il n'y a que des manifestations de l'émotion. En effet, à l'instar de nombreux processus émergents, les émotions proviennent en grande partie de l'activité globale du système, qui par sa dynamique se structure en différents niveaux d'organisation.

Il faut alors chercher non pas à modéliser des activités émotionnelles ni même cognitives, mais plutôt essayer, à partir des informations que l'on possède sur les facultés intellectuelles en général, de se représenter ce qu'est un processus artificiel intelligent. Tout ceci suppose de faire abstraction de tous les processus ontogénétiques ou phylogénétiques humains, lesquels ne peuvent s'appliquer à la création d'une entité informatique constituée uniquement de mémoire et d'instructions. Il s'agit là de la principale différence entre le cerveau et le circuit électrique. Ce dernier, aussi élaboré que possible, possède systématiquement une logique calculable à partir de laquelle, souvent, il a été conçu. Le cerveau quant à lui, ne possède pas de logique propre. Il est polyvalent et peut s'adapter à une situation donnée, mais ne peut en aucun cas être vu comme un processus déterministe dont les fonctionnalités sont localisées au sein d'une zone cérébrale spécifiquement dédiée. Par conséquent, il ne peut être étudié comme un système global ayant un objectif, une finalité.

5.4.2 Définition de l'émotion artificielle

D'un point de vue d'informaticien, l'émotion peut prendre invariablement la forme de *sur-couche* structurelle, de *sous-couche*, ou encore de module indépendant. Lorsque l'on parle de sur-couche émotionnelle, il s'agit d'un ensemble de processus qui vont venir accompagner ou perturber (positivement ou négativement) une structure cognitive déjà en place, visant à résoudre le problème posé. Une sous-couche consiste par ailleurs en une base émotionnelle au-dessus de la-

quelle des processus cognitifs viendraient se former et se verraient littéralement conditionnés par les circonstances émotionnelles courantes. Enfin, l'émotion vue en tant qu'organisation modulaire indépendante aurait alors une existence parallèle à un module cognitif, l'un et l'autre étant exécutés successivement.

Ces trois approches feront l'objet de trois modèles réalisés dans le cadre de cette thèse : le premier modèle, intitulé problème des Cascades, décrit l'émotion comme une sur-couche sous forme de marquage sur un graphe cognitif ; le second modèle utilise un module émotionnel à base de logique floue, en parallèle avec un module cognitif interfacé avec un environnement proie-prédateur représenté ; enfin, le troisième modèle, nommé GAEA, part du principe que l'émotion est une sous-couche émergente au sein de laquelle s'organisent les processus cognitifs. Avant de présenter ces trois travaux de modélisation, je tiens tout d'abord à formuler clairement ce que je considère comme faisant partie des processus émotionnels artificiels.

Émotion artificielle

Ensemble de processus préprogrammés ou non s'exécutant au sein d'une machine, favorisant la prise de décision, et permettant au système de s'adapter à l'environnement. Cette émotion artificielle est le fruit des entrées-sorties du programme ainsi que de son activité interne propre, et fait souvent l'objet d'une collaboration avec une structure cognitive, à l'aide de laquelle le système fait face aux problématiques introduites par son environnement. En outre, elle s'inscrit dans une logique de programmation plus ou moins explicite, mais toujours dans le domaine du calculable.

Les trois parties exposées ci-après sont l'illustration de cette définition, et présentent les trois modèles de l'émotion développés durant cette thèse.

II

LE PROBLÈME DES *Cascades* : UN MODÈLE À BASE DE GRAPHE DÉCISIONNEL À MARQUAGE ÉMOTIONNEL

L'objectif de ce travail de thèse est de proposer plusieurs approches autour de la représentation de l'émotion par la création de modèles qui tendent à représenter les processus émotionnels ainsi que leur dynamique. Un précédent modèle a déjà fait l'objet d'une étude lors de mes travaux de master à l'université du Havre ([78] et [79]). Il s'agissait alors de reproduire le déroulement d'une session d'un jeu appelé la *Gambling Task*³. Créée par BECHARA et DAMASIO, cette expérience est basée sur la réaction émotionnelle d'un individu face à la perte ou au gain d'une somme d'argent. Dès lors, le but étant pour lui de maximiser ses gains et de minimiser ses pertes, il va rapidement se prendre au jeu et développer un ensemble de réactions émotionnelles visant à faire face à la situation ([11]).

La *Gambling Task* se présente comme un jeu utilisant quatre paquets de cartes et une cagnotte de départ. À chaque tour, le joueur doit choisir de tirer une carte dans l'un des paquets. En fonction du contenu de la carte, l'expérimentateur lui annonce alors s'il a perdu ou gagné de l'argent, l'objectif étant d'amasser un maximum d'argent à l'issue de la partie. Le jeu en apparence aléatoire est en réalité préparé de manière à différencier le contenu des paquets. En somme, on propose au joueur deux paquets à risque élevé et deux paquets à moindre risque, les premiers ayant plutôt tendance à apporter davantage d'argent qu'ils n'en font perdre (et *vice versa* pour les seconds paquets). On distingue alors des profils de joueurs dont le comportement est conditionné par leur faculté à « oser tenter le diable » pour potentiellement gagner plus, mais aussi perdre beaucoup, voire énormément pour certaines cartes présentes dans les paquets les plus risqués !

La modélisation adoptée pour simuler une partie de *Gambling Task* faisait usage d'un automate comportemental représentant les quatre paquets de cartes, assorti d'un ensemble de probabilités de se mouvoir d'un paquet à l'autre ou bien de continuer à piocher des cartes au sein du même paquet (voir figure 5.8). Ces probabilités sont dynamiques et dépendent de l'état émotionnel de l'individu, représenté par un modèle OCC, ainsi que de l'histoire de chaque paquet, indiqué par un système de score.

Le modèle proposé dans cette partie va plus loin dans la démarche de représentation de l'émotion en y introduisant la notion d'*appraisal* (évaluation cognitive) dans la façon d'appréhender la résolution d'un problème donné. En effet, les théories de l'*appraisal*, issues de la perspective cognitive décrite précédemment, ont affirmé leur capacité à apporter une couleur cognitive aux processus émotionnels, et ont prouvé à maintes reprises qu'émotion et cognition sont les deux faces d'une seule et même pièce : elles forment toutes deux la pierre angulaire de la prise de décision.

Le travail présenté ici se place dans le cadre de la résolution de problème au sein d'un exercice de mathématiques intitulé *Problème des Cascades*. Le modèle

3. *Gambling Task* signifie « tâche de jeu », le terme *jeu* étant ici utilisé pour désigner le jeu d'argent. Par ailleurs, le terme *gamble* peut être traduit par « parier » ou « jouer de l'argent »

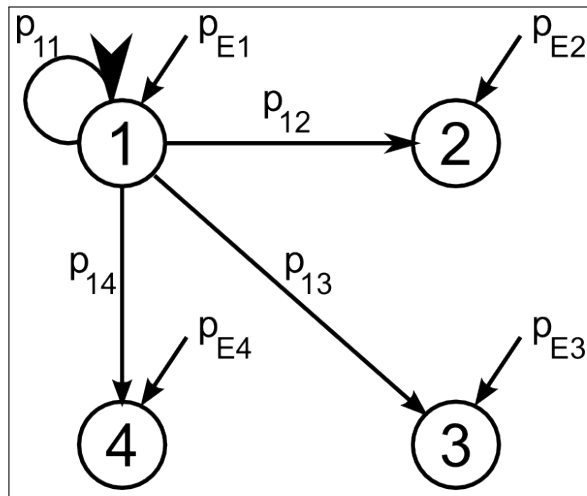


FIG. 5.8 – AUTOMATE COMPORTEMENTAL DE LA *Gambling Task*

qui en est issu est une forme d'interprétation d'une théorie nommée *appraisal-coping*, chère à Richard LAZARUS ([69]), qui définit deux concepts majeurs (voir figure 5.9) : les processus de l'*appraisal* consistent en l'évaluation cognitive de la situation par rapport à ses conséquences pour l'organisme (bien-être) et décrivent la façon dont un individu évalue son environnement, et les processus du *coping* sont des stratégies d'ajustement qui orientent les réponses en fonction des conséquences évaluées. Le *coping*, en somme, indique comment l'individu va faire face à la situation, d'une part au travers de stratégies dites *problème-centrées* qui concernent la modification de l'environnement, et d'autre part au travers de dispositions dites *émotion-centrées* qui vont venir modifier certains aspects cognitifs et émotionnels de l'individu.

Appliqué au problème des Cascades, ce modèle considère la grille à compléter en tant qu'environnement de travail, ainsi que le graphe de stratégies comme responsable de l'activité cognitivo-émotionnelle sous tendant les processus de planification et de décision (voir figure 5.10).

Le premier chapitre de cette partie va s'intéresser aux règles inhérentes au problème des *Cascades* et présenter le contexte dans lequel il s'inscrit. Ensuite sera présenté le modèle réalisé, visant à reproduire l'activité de résolution du problème. Enfin, seront traités les résultats et éventuelles perspectives.

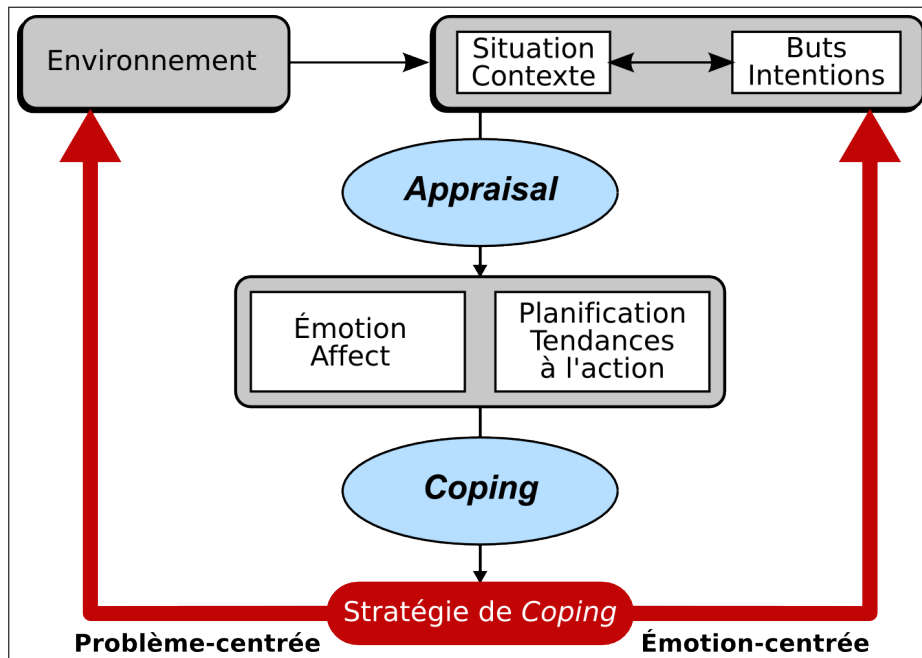


FIG. 5.9 – MODÈLE D'*appraisal-coping* ADAPTÉ DE LAZARUS ([69])

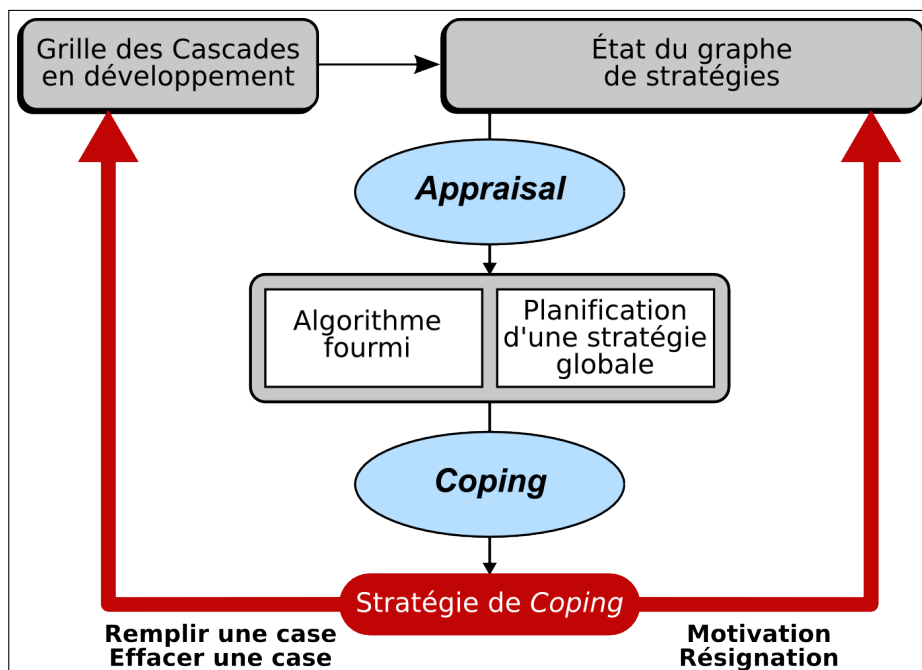


FIG. 5.10 – MODÈLE D'*appraisal-coping* APPLIQUÉ AU PROBLÈME DES CASCADES

MODÉLISATION

6

SOMMAIRE

6.1	RÈGLES DU JEU	95
6.2	REPRÉSENTATION DE LA GRILLE DES <i>Cascades</i>	98
6.3	GESTION DES STRATÉGIES	99
6.3.1	Graphe de stratégies	100
6.3.2	<i>Pattern-matching</i> des stratégies	101
6.4	PROCESSUS DE RÉOLUTION	103
6.4.1	Algorithme fourni	104
6.4.2	Sélection des stratégies	105
6.4.3	Algorithme principal	106

Deep in the human unconscious is a pervasive need for a logical universe that makes sense. But the real universe is always one step beyond logic.^a

— Frank HERBERT

a. « Dans l'inconscient profond de l'homme existe le besoin irrépressible d'un univers logique qui aurait un sens. Mais l'univers réel a toujours une longueur d'avance sur la logique. » (NDR)

6.1 RÈGLES DU JEU

Le problème des Cascades est un test d'arithmétique proposé à des élèves de CM2 qui vise à comprendre le comportement émotionnel d'enfants de 10 ans face à un exercice donné dont la complexité élevée induit pour beaucoup d'entre

eux un comportement émotionnel significatif. Cet exercice utilise une grille de nombres dans laquelle les enfants devront respecter une règle mathématique simple. Cette expérience, menée par le laboratoire Psy.NCA de l'université de Rouen ([12]) nous a appris un certain nombre de choses sur le comportement des élèves :

- ils tendent à développer, tout au long de l'exercice, des stratégies globales et locales, même s'il s'avère qu'elles présentent parfois certaines incohérences. Ils sont néanmoins, la plupart du temps, parfaitement capables d'expliquer leurs choix *a posteriori* ;
- ces enfants présentent des caractéristiques émotionnelles manifestement différentes, en fonction notamment de leur confiance en eux lors de la résolution. Alors que certains opèrent par essais-erreurs, d'autres finissent par abandonner délibérément, tant ils estiment que leurs résultats sont faux.

Le modèle proposé ici tente de représenter le plus fidèlement possible les différentes étapes de résolution de ce problème en tenant compte non seulement des stratégies employées, mais aussi de l'état cognitivo-émotionnel de l'enfant, celui-ci étant amené assez régulièrement à changer de stratégie globale durant l'exercice. Le programme issu de ce modèle doit ainsi, en accord avec son environnement, c'est-à-dire le problème en cours de résolution, et son état interne, *i.e.* cognitif et émotionnel, adapter son comportement à la bonne résolution de l'exercice.

On propose à chaque enfant la grille initiale des *Cascades* (voir figures 6.1 et 6.2), ainsi que l'énoncé suivant :

*Chaque case contient la somme des nombres situés au-dessus d'elle.
Trouve les nombres qui manquent dans la grille ci-dessous.*

Une des premières remarques que l'on peut faire à l'issue des expériences est le fait que les enfants utilisent absolument toutes sortes de stratégies pour résoudre le problème donné. En réalité, c'est la non-compréhension de l'énoncé (tout du moins en début d'expérience) qui force l'élève à construire sa propre interprétation des règles de l'exercice. Ce processus d'approche par essais-erreurs révèle ainsi des stratégies insoupçonnées, qui vont mener certains enfants à résoudre la grille entièrement et de façon cohérente avec leurs propres règles malgré l'inexactitude de la solution finalement proposée.

Finalement, l'objectif de ce travail est de stocker les valeurs de chacune des cases de la grille des *Cascades*, ainsi que les différentes stratégies mises en œuvre dans le processus de remplissage de ces cases. Nous obtenons alors deux niveaux de description distincts : d'un côté une représentation graphique de la grille des *Cascades* en elle-même, et d'un autre côté un graphe de stratégies utilisé pour organiser l'ensemble des *patterns* possibles.

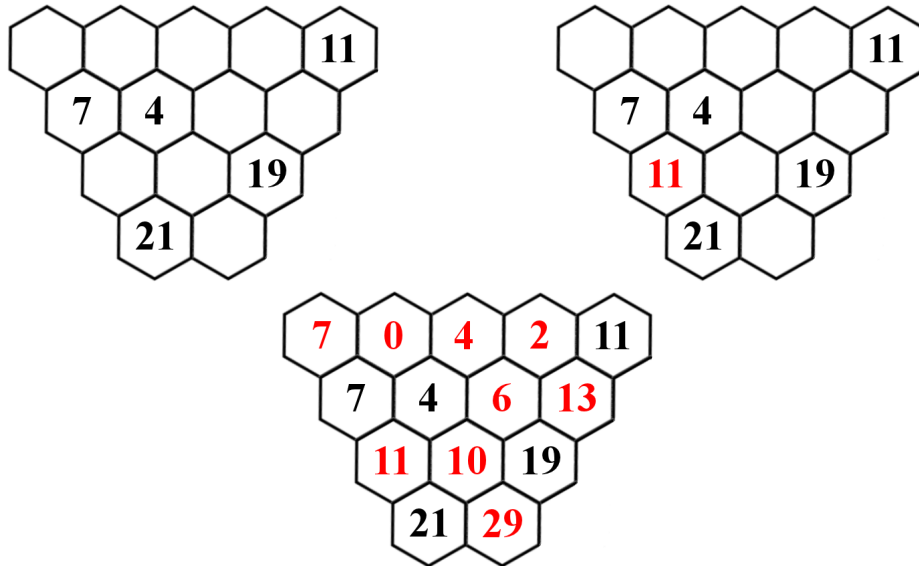


FIG. 6.1 – GRILLE INITIALE, PREMIÈRE ÉTAPE DE RÉOLUTION ET SOLUTION DE LA GRILLE DU PROBLÈME DES *Cascades*

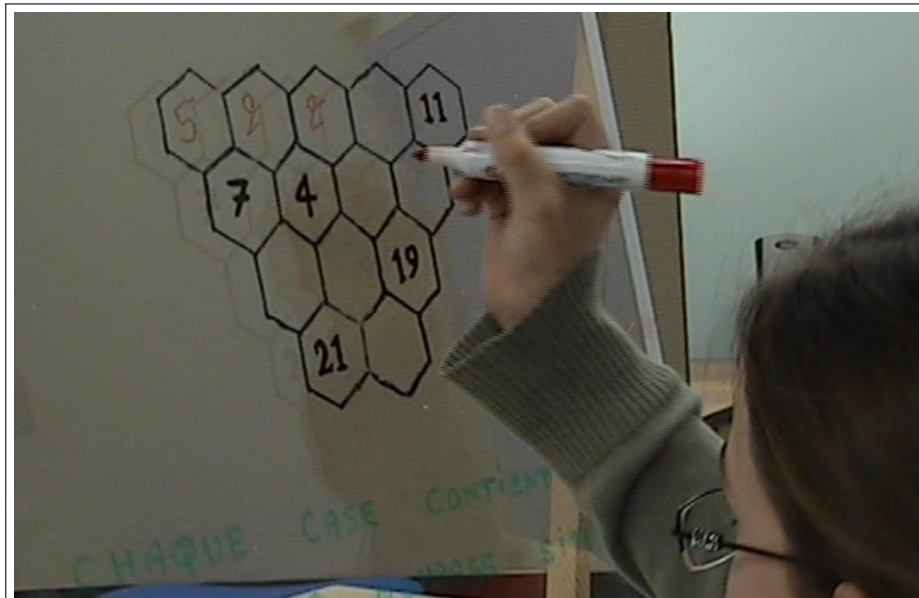


FIG. 6.2 – ÉLÈVE RÉSOUVANT LE PROBLÈME DES *Cascades* ([12])

6.2 REPRÉSENTATION DE LA GRILLE DES *Cascades*

Techniquement parlant, la grille des *Cascades* est un graphe, c'est-à-dire une structure logique composée de nœuds et de liens, tel que chaque nœud $n \in N$ représente un item conceptuel de l'environnement, en d'autres termes une case à remplir dans la grille, et chaque lien $l \in L$ est une fonction entre deux de ces nœuds. Ce graphe étant supposé être une description physique de la grille des *Cascades*, il nous faut considérer la représentation des cases à remplir, mais aussi des relations spatiales entre ces cases (« à gauche, à droite, en-dessous... »). Entrent alors en jeu les liens l entre les nœuds qui vont décrire cette relation de voisinage entre deux cases (voir figure 6.3).

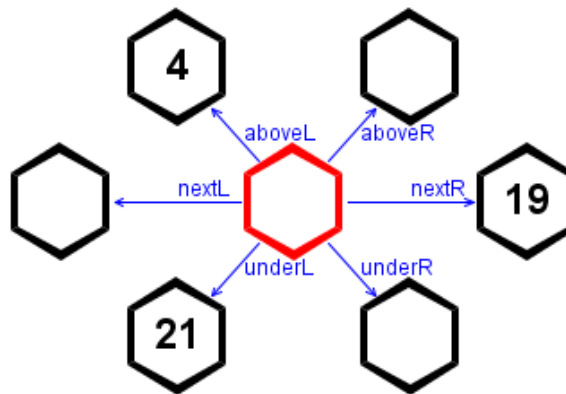


FIG. 6.3 – UNE CASE ET SON VOISINAGE

Une distinction est faite, par exemple, entre *aboveL* (*above left* qui signifie « au-dessus à gauche ») et *aboveR* (*above right* qui signifie « au-dessus à droite ») car certains opérateurs tels que la soustraction ne sont pas commutatifs et distinguent par conséquent leur opérande de droite de celle de gauche. Il est donc indispensable de clarifier les liens séparant les différentes cases afin d'éviter toute confusion entre les nombres à manipuler (ce que certains enfants ne font pas toujours par ailleurs!).

Δ étant défini comme l'ensemble des *fonctions spatiales* possibles entre deux cases, nous obtenons alors :

$$\Delta = \left\{ \begin{array}{l} aboveL \\ aboveR \\ nextL \\ nextR \\ underL \\ underR \end{array} \right\}$$

Ainsi, la grille de représentation des *Cascades* M (pour *map* en anglais) est un graphe $M = (N, L)$, N étant l'ensemble des nœuds et L l'ensemble des liens

entre ces nœuds, tels que :

$$M = (N, L)$$

$$L = \{(n, n', \delta) / n, n' \in N, \delta \in \Delta\}$$

Pour pouvoir effectivement modifier le contenu des cases de la grille, il nous faut définir une fonction $f : N \rightarrow \mathbb{Z}$ sur les nœuds, dont le but sera de calculer la valeur numérique au sein de chaque nœud/case dans la grille. Finalement, la solution au problème des *Cascades* consiste à trouver la bonne fonction pour chaque nœud n appartenant à N .

6.3 GESTION DES STRATÉGIES

Afin de modéliser les aspects logiques de la résolution, on utilise un ensemble de stratégies prédéterminées. En effet, après qu'un enfant a décidé quelle case il va modifier, il doit aussi construire le raisonnement nécessaire au calcul effectif du contenu de la case (« *Quelles cases vais-je additionner ou soustraire ?* »). Les stratégies entrent alors en jeu en tant que représentations locales au service de l'état courant de la résolution. Pour ce faire, on applique un algorithme de *pattern-matching* dont le but sera de tenter d'appliquer une stratégie à un certain endroit de la grille. Si la forme de la stratégie sélectionnée recouvre parfaitement la zone spécifique choisie, le programme l'applique. Sinon, il en sélectionne une autre. La figure 6.4 est un exemple de *pattern* de stratégie.

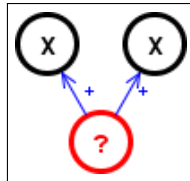


FIG. 6.4 – EXEMPLE DE STRATÉGIE. POUR TROUVER LE CONTENU DE LA CASE ROUGE, IL FAUT ADDITIONNER LES VALEURS DES DEUX CASES NOIRES DU DESSUS

Formellement, l'ensemble des stratégies S s'appliquant localement au sein de la grille se définit ainsi :

$$S = \{(N', L', op) / N' \subset N, L' \subset L, op : L' \rightarrow O\}$$

où $N' \subset N$ est l'ensemble des nœuds de la grille sélectionnés par la stratégie courante $s \in S$ dans le cadre du *pattern-matching*, $L' \subset L$ est l'ensemble des liens sélectionnés par cette même stratégie courante s , et op est une fonction qui permet de relier chaque lien à un opérateur spécifique dans O , ce dernier étant l'ensemble des opérateurs mathématiques pouvant s'appliquer entre deux nœuds.

En réalité, N' regroupe deux ensembles de nœuds particuliers : N'_g correspondant aux nœuds sélectionnés mais déjà remplis dans la grille, et N'_c correspondant aux nœuds à calculer.

$$N' = N'_g \cup N'_c$$

Chaque lien $l \in L'$ dans une stratégie possède un unique point d'origine appelé la *racine* de la stratégie. Cette racine identifie, la plupart du temps, le nœud que l'on cherche à remplir à travers la fonction f . Dans ce cas, la valeur de ce nœud-racine est la conséquence de l'application de la stratégie, des nœuds environnants jusqu'à lui, tel que l'on peut voir dans la figure 6.4. Cependant, il arrive parfois que le nœud racine soit le point de départ du calcul de la stratégie (voir figure 6.7). Dans ce cas, c'est à l'aide de ce nœud que l'on peut remplir les autres cases de la stratégie.

6.3.1 Graphe de stratégies

Maintenant que nous savons ce qu'est une stratégie, il nous faut définir une représentation dynamique utilisée dans le cadre du processus de résolution. Pour chaque stratégie s , on associe une variable k (de l'anglais *knowledge value*) dans l'intervalle $[0; 1]$ représentant sa valeur de pondération cognitive. C représente ainsi l'ensemble des stratégies associées à une valeur k dans $[0; 1]$ telles que :

$$C \subset S \times [0; 1]$$

En parallèle, E décrit l'ensemble des liens à connotation émotionnelle entre les stratégies, tels que chaque $e \in E$ est défini comme suit :

$$e = (c, c', \eta)$$

$$\text{avec : } \begin{aligned} c, c' &\in C \\ \eta &\in [0; 1] \end{aligned}$$

où η correspond à une valeur émotionnelle sur chaque lien qui relie deux stratégies. Finalement, les deux ensembles C (stratégies associées à des valeurs cognitives k) et E (ensemble de liens unissant ces stratégies et associés à des valeurs émotionnelles η) nous permettent de définir ce que l'on appellera dorénavant le *graphe de stratégies* $G = (C, E)$ dont le rôle est de structurer les relations entre stratégies et grâce auquel il sera possible de résoudre la grille des *Cascades*.

En résumé, ce modèle permet de définir deux structures mathématiques importantes :

- une grille M représentant graphiquement le problème des *Cascades* ;
- un graphe de stratégies G utilisé pour organiser les stratégies et leurs relations les unes par rapport aux autres.

Notre propos est de postuler que l'émotion joue un rôle de catalyseur sur les liens entre stratégies qui seront alors plus facilement activables dans un processus de décision. Le graphe de stratégies représente par conséquent le support d'une exploration guidée par l'émotion.

6.3.2 *Pattern-matching* des stratégies

Comme nous l'avons vu précédemment, le rôle de la grille est de se représenter un schéma visuel des *Cascades*. Quant aux stratégies, chacune d'entre elles se doit d'être capable de remplir effectivement une case particulière en utilisant un *pattern* logique spécifique.

Un *pattern* de stratégie est un diagramme basique construit à partir de 3, 4 ou 5 cases, représentant les items requis pour produire un calcul local (addition ou soustraction). Par conséquent, ce diagramme se compose, la plupart du temps, d'une case vide et de plusieurs cases considérées comme déjà remplies ou données au début de l'exercice, le but étant de trouver quel nombre doit être inscrit dans la case vide.

Grâce aux résultats d'expériences effectuées par le laboratoire Psy.NCA, nous avons pu identifier un très grand nombre de *patterns* logiques utilisés par les élèves durant l'exercice. En effet, la plupart des enfants ne comprenant pas l'énoncé de l'épreuve, ils tentent d'utiliser beaucoup de *patterns* alternatifs dont certains sont très loin du résultat attendu.

Ainsi, deux principaux types de *patterns* ont été extraits de l'analyse des résultats : des *patterns triangulaires* et des *patterns linéaires* (voir figures 6.5 et 6.6). Les *patterns* triangulaires sont des structures à trois cases dans quasiment toutes les configurations possibles. Ils sont eux-mêmes subdivisés en deux catégories que sont les formes en *diamant* (en anglais *diamond*) et les formes en *pyramide* (*pyramid*). Le nom d'une stratégie triangulaire est ainsi composé de l'un de ces deux termes (*diamond* ou *pyramid*) suivi généralement de la position de la case à remplir (en rouge sur les figures) selon le code suivant :

- U pour *under* (dessous) ;
- L pour *left* (gauche) ;
- R pour *right* (droite) ;
- A pour *above* (dessus) ;
- I pour *infer* (inférence).

Les *patterns* linéaires, quant à eux, sont composés de 3, 4 et jusqu'à 5 cases alignées dans toutes les directions possibles. Elles portent le nom *line* (ligne) précédé de **u** (pour *up*, haut) pour les lignes partant du bas vers le haut — de gauche à droite — et **d** (pour *down*, bas) pour les lignes allant du haut vers le bas — toujours de gauche à droite. En outre, les lignes horizontales sont précédées

de **h** (pour *horizontal*). Leur nom se compose enfin du nombre de cases qu'elles utilisent, en l'occurrence 3, 4 ou 5, et indique la position de la case à trouver en utilisant **L** ou **R** selon la même logique que celle décrite ci-dessus.

Le graphe de stratégies comprend ainsi au total 22 stratégies de forme et de composition variées. Cependant, il s'agit là uniquement d'une base de travail. En effet, il existe de nombreux autres *patterns* que l'on pourrait potentiellement intégrer dans ce modèle, preuve s'il en est de l'extraordinaire inventivité de nos charmantes têtes blondes !

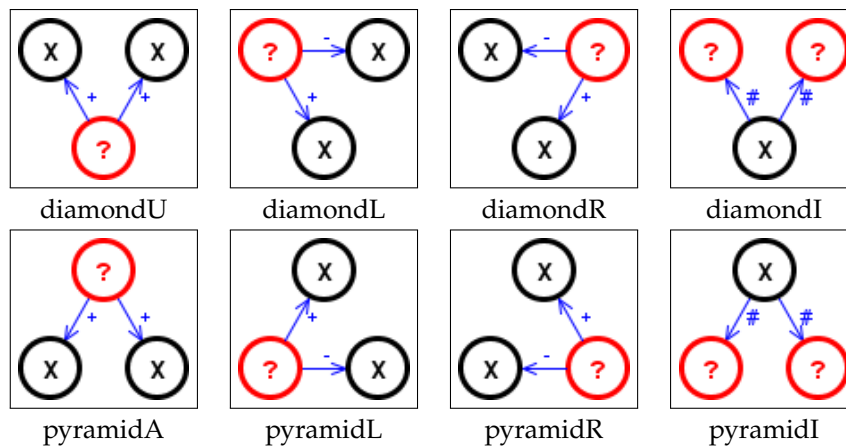


FIG. 6.5 – EXEMPLES DE STRATÉGIES TRIANGULAIRES

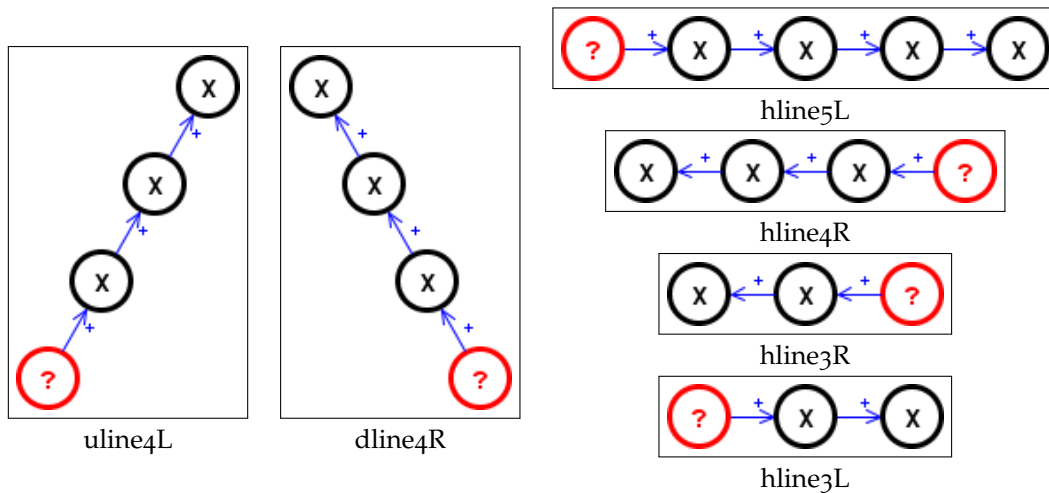


FIG. 6.6 – EXEMPLES DE STRATÉGIES LINÉAIRES

L'interprétation d'un *pattern* est la façon de connecter les cases ensemble, dans le but de savoir quelle case doit être additionnée ou soustraite pour trouver la réponse. Là encore, il existe un ensemble d'opérateurs utilisables entre deux cases :

l'addition (+), la soustraction (−) ou l'inférence (#). En effet, il arrive parfois que les élèves génèrent les deux opérandes à partir d'un opérateur en « devinant » quels nombres doivent être inscrits dans les cases. Par exemple, si deux cases vides se situent au-dessus d'une troisième case marquée d'un 7, un enfant peut penser : « $7 = 3 + 4$ donc pourquoi ne pas mettre les valeurs 4 et 3 dans les deux cases du dessus ? » (voir figures 6.5 et 6.7).

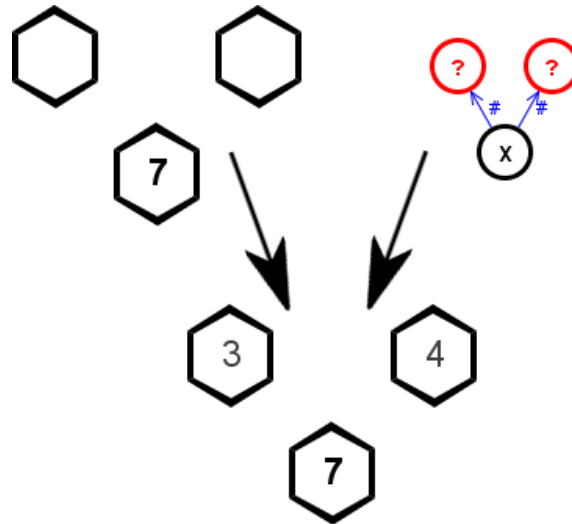


FIG. 6.7 – EXEMPLE D'APPLICATION D'UNE STRATÉGIE D'INFÉRENCE

Cet opérateur est utilisé par les élèves lorsqu'ils ne trouvent pas de moyen déterministe pour arriver à la solution. La jeune fille sur la photographie issue de l'expérience (voir figure 6.2) utilise cette méthode avec la règle $7 = 5 + 2$. Cela lui permet ensuite de déduire le 2 juste à droite avec la règle $2 = 4 - 2$ en utilisant la stratégie *diamondR* (voir figure 6.5).

6.4 PROCESSUS DE RÉOLUTION

Le mécanisme de résolution de la grille consiste à décider quelle stratégie utiliser pour quelle case. Dans toute activité cérébrale impliquant une composante cognitive et émotionnelle, le marquage émotionnel se doit d'être un processus dynamique évoluant non seulement en fonction du temps, mais aussi de la situation. En d'autres termes, l'expression de l'émotion d'un individu doit être à l'image de l'environnement en constante évolution. Cette dynamique adaptant facilement son comportement à la rétroaction environnementale peut être représentée grâce au principe de *stigmergie*, introduit par le biologiste français Pierre-Paul GRASSÉ en 1959 ([50]).

Stigmergie

La stigmergie est une forme de communication indirecte souvent observée dans un environnement auto-organisé. Les individus communiquent entre eux par l'intermédiaire de l'environnement, ce qui crée alors une activité émergente intelligente ou intelligence distribuée.

Le modèle des Cascades met en place ce principe de stigmergie à travers un modèle célèbre d'intelligence en essaim : un *algorithme fourmi* ([31], [14]).

6.4.1 Algorithme fourmi

Les fourmis, par leur dynamique, sont capables de discrétiser l'espace au sein d'un graphe en déposant des phéromones sur les chemins menant aux endroits les plus attractifs. Appliqué à notre graphe de stratégies, ce procédé a l'avantage de faire émerger les meilleures stratégies à l'aide de deux techniques : le *renforcement* des arêtes en cas de découverte d'une bonne stratégie ou bien la *pénalisation* des arêtes en cas d'échec d'une stratégie. Ainsi, la rétroaction positive se joue au niveau des fourmis qui viennent déposer leurs phéromones au niveau des arêtes, mais aussi du renforcement des valuations des sommets lors de la phase de résolution. Quant à la rétroaction négative, elle se manifeste par l'évaporation des phéromones sur les arêtes, ainsi que la dévaluation des sommets à l'étape de résolution. Grâce à ces rétroactions positives et négatives sur le graphe de stratégies, le processus de résolution de la grille des *Cascades* se développe de manière équilibrée, jusqu'à la découverte d'une solution cohérente.

Voici maintenant le fonctionnement de l'algorithme fourmi en tant que tel. Soit \mathcal{A} l'ensemble des fourmis réparties dans le graphe. Chaque fourmi $a \in \mathcal{A}$ démarre sa course sur un nœud i au hasard pendant la phase d'initialisation, et a une probabilité $p_{i,j}$ de se déplacer vers un autre nœud j à chaque étape de l'algorithme, tel que :

$$p_{i,j} = \frac{k_j}{\sum_{j \neq i} k_j} \quad (6.1)$$

où k_j est la valeur de connaissance associée à la stratégie sur le nœud j . Ainsi, si l'on considère une arête (i, j) , une fourmi est potentiellement attirée par la valeur k sur la stratégie située de l'autre côté de cette arête.

Lorsqu'une fourmi a se déplace le long d'une arête (i, j) , elle y dépose une certaine quantité $q_{i,j}^a$ de phéromones qui viennent compléter la quantité $\eta_{i,j}$ de phéromones déjà présente sur l'arête. La mise à jour de la quantité $\eta_{i,j}$ de phéromones sur l'arête (i, j) , de l'étape $t - 1$ à l'étape t , se fait par la formule suivante :

$$\eta_{i,j}^{(t)} = (1 - \rho)\eta_{i,j}^{(t-1)} + Q_{i,j} \quad (6.2)$$

$$\text{avec : } Q_{i,j} = \sum_{a \in \mathcal{A}} q_{i,j}^a$$

où ρ est le coefficient d'évaporation des phéromones et $Q_{i,j}$ est la somme des phéromones q^a déposées par les fourmis à chaque étape. L'algorithme fourni permet ainsi de marquer le graphe pour y faire émerger des chemins significatifs qui seront prioritairement empruntés afin d'élaborer une solution à partir d'une sélection de stratégies.

Contrairement aux algorithmes fourmis traditionnels, ce marquage de phéromones ne rétroagit pas directement sur la phase d'exploration des fourmis décrites par la formule 6.1. En effet, un tel renforcement dans cette formule n'a pas de sens puisque, à cette étape, il n'y a pas d'évaluation de solutions potentielles. D'ailleurs, le calcul de $p_{i,j}$ ne tient pas compte des valeurs $\eta_{i,j}$ contenues sur les arêtes du graphe. L'analyse du choix des stratégies potentielles aura lieu dans une phase ultérieure et c'est seulement à cette étape que le dépôt de phéromones aura un impact sur le processus global de recherche de solutions.

6.4.2 Sélection des stratégies

Lorsqu'un certain nombre de pas de temps est atteint par l'algorithme fourni, le graphe est figé et l'on observe les valeurs η sur chaque arc. Ainsi débute le processus de sélection dont l'intérêt est d'extraire les stratégies qui seront ensuite utilisées pour la résolution.

Les deux sommets situés de part et d'autre de l'arc dont la valeur η est la plus élevée, c'est-à-dire les sommets correspondant à l'arête qui a été la plus visitée, sont sélectionnés et placés dans une liste L . On obtient ainsi une liste de stratégies ajoutées successivement, au fur et à mesure du bon déroulement du processus de sélection. Cette liste contient alors tous les sommets extraits du graphe de stratégies selon le mécanisme de sélection des arêtes dont la valeur η est la plus élevée, en sachant qu'un sommet ajouté doit impérativement être connecté aux sommets déjà sélectionnés dans la liste. Dès lors, on obtient une liste de stratégies L dont l'ordre a une importance capitale car si deux stratégies sont appliquées sur la grille des *Cascades* dans un sens ou dans un autre, l'état final de la grille peut être complètement différent. Autrement dit, les stratégies sélectionnées au sein du graphe constituent une liste de priorité dont l'ordre peut déterminer, ou pas, l'état final de résolution de la grille des *Cascades*. On constate alors deux situations distinctes :

- si l'ordre dans lequel un ensemble de stratégies est appliqué au cours de la résolution a un impact sur l'état final de la grille, cela signifie que les

patrons (patterns) respectifs de deux (ou plus) de ces stratégies sont suffisamment proches pour pouvoir être utilisés dans les mêmes situations de résolution. Elles ont de grandes chances de modifier la grille avec des valeurs différentes, et par là même, biaiser l'avenir de la résolution. Par conséquent, elles sont déclarées *conflictuelles* ;

- si l'ordre dans lequel les stratégies sont employées n'a aucune influence sur l'état final de la grille, il s'agit alors d'un cas dans lequel les stratégies n'entrent pas en conflit car leurs patrons s'utilisent dans des situations de résolution complètement différentes. Elles sont donc déclarées *compatibles*.

Afin de détecter les éventuels conflits au sein de la liste L de stratégies sélectionnées, on effectue la résolution dans l'ordre indiqué par ladite liste, puis l'on recommence cette même résolution en changeant l'ordre des stratégies, jusqu'à tester ainsi toutes les possibilités d'ordre. Si, à l'issue de toutes ces résolutions, les états finaux de la grille ne sont pas rigoureusement identiques, on en déduira que la dernière stratégie ajoutée dans cette liste n'est pas compatible avec la liste L , c'est-à-dire qu'elle entre en conflit avec au moins une des stratégies présentes dans cette liste. Il faudra alors pénaliser cette dernière stratégie par rapport aux autres stratégies de L , et considérer globalement l'ensemble des stratégies de la liste comme un échec de solution potentielle.

Finalement, à partir de la liste L de stratégies sélectionnées, le programme teste toutes les configurations possibles, à savoir tous les ordres possibles, afin de résoudre la grille, jusqu'à ce que :

- tous les ordres de résolutions mènent au même résultat final et la grille est complète. Cette situation idéale implique l'arrêt de l'algorithme ;
- tous les ordres de résolutions mènent au même résultat final mais la grille n'est toujours pas complète. Il nous faut alors extraire une nouvelle stratégie à partir du graphe, afin de continuer la résolution ;
- les ordres de résolution mènent à différents états finaux, ce qui signifie que l'ordre avec lequel les stratégies sont appliquées a un impact sur le résultat final. Dans ce cas, la dernière stratégie est déclarée incompatible avec les autres. Elle est par conséquent pénalisée par la suppression des arêtes existantes entre elle et les autres stratégies de la liste. En outre, toutes les stratégies précédemment sélectionnées de la liste L seront pénalisées en décrémentant leur valeur de connaissance k . Enfin, la liste L est vidée et la boucle principale du programme peut reprendre à zéro par l'exécution de l'algorithme fourni au sein du graphe de stratégies nouvellement modifié.

6.4.3 Algorithme principal

Comme décrit précédemment, l'algorithme principal (voir algorithme 1) utilise une liste de stratégies L dont l'ordre peut avoir un impact sur la résolution

Algorithme 1 Résolution des *Cascades*

```

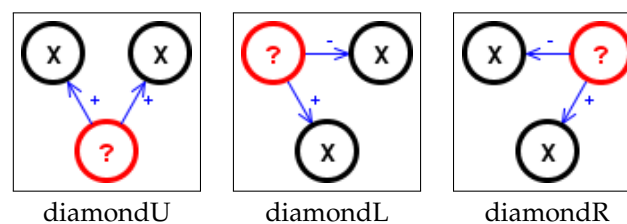
1:  $s \leftarrow \text{null}$  // stratégie extraite courante
2:  $L \leftarrow []$  // liste des stratégies sélectionnées
3:  $\text{continuer} \leftarrow \text{vrai}$ 
4: Tant que (la grille des Cascades n'est pas remplie) Faire
5:   exécuter l'algorithme fourni pendant  $n$  étapes
6:    $s \leftarrow \text{extraireStrategie}(L)$ 
7:    $L \leftarrow [s]$  // on ajoute  $s$  à la liste  $L$ 
8:    $\text{continuer} \leftarrow \text{vrai}$ 
9:   Tant que ( $\text{continuer} = \text{vrai}$ ) Faire
10:     $s \leftarrow \text{extraireStrategie}(L)$ 
11:     $L \leftarrow L + [s]$ 
12:    Pour (tous les ordres possibles de  $L$ ) Faire
13:      résoudre la grille en utilisant  $L$ 
14:      stocker les états finaux de la grille
15:    Fin Pour
16:    Si (les états finaux sont tous identiques) Alors
17:      Si (la grille des Cascades est remplie) Alors
18:        STOP! La résolution est un succès!
19:      Fin Si
20:      Si (la liste  $L$  est pleine) Alors
21:        pénaliser toutes les stratégies dans  $L$ 
22:        renforcer les stratégies connectées à  $L$ 
23:         $\text{continuer} \leftarrow \text{faux}$ 
24:      Fin Si
25:      Sinon
26:        pénaliser toutes les stratégies dans  $L$ 
27:        renforcer les stratégies connectées à  $L$ 
28:        pénaliser les arêtes entre la dernière stratégie dans  $L$  et les autres stratégies
        de  $L$ 
29:         $\text{continuer} \leftarrow \text{faux}$ 
30:      Fin Si
31:    Fin Tant que
32:  Fin Tant que

```

des Cascades. Si les états finaux obtenus par les différents ordres de la liste L ne sont pas tous parfaitement identiques, alors la pénalisation des sommets et des arêtes se met en place ainsi que le renforcement des stratégies connectées à L . En outre, l'ajout de stratégie dans L s'interrompt. Si à l'inverse les états finaux issus de la liste L sont identiques, trois cas se présentent : si la grille est complète, l'algorithme s'arrête sur un succès ; si la liste L est pleine seule, la pénalisation des stratégies a lieu et l'ajout de nouvelles stratégies s'interrompt ; enfin, si L n'est pas pleine, l'algorithme poursuit son extraction de nouvelles stratégies.

Par ailleurs, cette extraction permet la construction de la liste L à l'aide d'une fonction d'extraction nommée *extraireStrategie(L)* décrite dans l'algorithme 2. Cette fonction distingue la situation dans laquelle la liste de stratégies passée en paramètre est vide de celle où elle contient au moins une stratégie. Ceci est nécessaire dans la mesure où la méthode utilisée est différente. Dans le premier cas, si L est vide, il va s'agir de trouver l'arête *maxA* la plus forte dans l'ensemble du graphe et de retourner aléatoirement un des deux sommets qu'elle relie. Dans le deuxième cas, si L contient déjà une ou plusieurs stratégies, l'algorithme choisit parmi les arêtes a connectées aux sommets de L une des arêtes les plus fortes et teste si celle-ci ne relie pas deux stratégies s et s_2 déjà présentes dans L . Si s_2 n'appartient pas à L , il est stocké dans *maxS* jusqu'à être éventuellement remplacé par un autre sommet plus avantageux. En fin d'algorithme, *maxS* est retourné et viendra se joindre aux autres stratégies de la liste L .

Dans l'algorithme principal utilisé (algorithme 1), l'unique condition d'arrêt est le remplissage de la grille des *Cascades*, peu importe le résultat final. En effet, quand bien même l'état final de la grille est faux par rapport aux instructions de l'exercice, il sera cohérent avec les différentes stratégies mises en œuvre et par conséquent, parfaitement pertinent. Si l'on devait trouver la bonne solution à la grille des *Cascades* utilisée jusqu'à présent et obtenir le résultat final attendu (voir la 3^{ème} grille de la figure 6.1), le programme devrait utiliser les trois stratégies suivantes :



Pour autant, l'objectif de ce modèle est de simuler le comportement d'un enfant face à un exercice de mathématiques relativement complexe pour lui, et l'intérêt ne réside pas du tout dans une quelconque optimisation du problème spécifique des *Cascades*. L'interprétation de l'énoncé de l'exercice joue alors un rôle primordial dans la manière dont seront mises en avant les stratégies utilisées. Dans notre cas, l'état initial du graphe de stratégies représente la situation

Algorithme 2 extraireStrategie (L : liste de stratégies)

```

1:  $a \leftarrow null$  // arête temporaire
2:  $maxA \leftarrow null$  // arête maximale
3:  $s \leftarrow null$  // stratégie temporaire
4:  $s2 \leftarrow null$  // stratégie temporaire
5:  $maxS \leftarrow null$  // sommet connecté à l'arête maximale
6: Si ( $L$  est vide) Alors
7:   Pour (chaque arête  $a$  du graphe) Faire
8:     trouver l'arête  $a$  la plus forte et la mettre dans  $maxA$ 
9:   Fin Pour
10:  Retourne un des deux sommets de  $maxA$  aléatoirement
11: Sinon
12:  Pour (chaque stratégie  $s$  de  $L$ ) Faire
13:    Pour (chaque arête  $a$  connectée à  $s$ ) Faire
14:      Si ( $a$  est plus forte que  $maxA$ ) Alors
15:         $s2 \leftarrow$  second sommet de  $a$  (différent de  $s$ )
16:        Si ( $s2$  n'est pas dans  $L$ ) Alors
17:           $maxA \leftarrow a$ 
18:           $maxS \leftarrow s2$ 
19:        Fin Si
20:      Fin Si
21:    Fin Pour
22:  Fin Pour
23:  Retourne  $maxS$ 
24: Fin Si

```

dans laquelle se trouve l'élève après avoir lu la consigne. Cet état peut être généré aléatoirement ou programmé manuellement à l'aide de fichiers de configuration. L'évolution du graphe de stratégies est alors assimilable à la recherche combinatoire d'une solution parmi l'ensemble des stratégies disponibles.

La recherche guidée par l'émotion va favoriser l'apparition de chemins « émotionnels » d'exploration pour se substituer à une exploration combinatoire exhaustive qui ne relève pas du processus cognitif d'un humain.

RÉSULTATS

7

SOMMAIRE

7.1	EXEMPLE DE RÉOLUTION	113
7.1.1	Étape 1	113
7.1.2	Étape 2	116
7.1.3	Étape finale	117
7.2	PLANS D'EXPÉRIENCE	119
7.2.1	Sensibilité aux paramètres	119
7.2.1.1	Nombre de grilles de Cascades résolues	119
7.2.1.2	Nombre moyen d'étapes de résolution	121
7.2.2	Impact de l'émotion	122
7.2.2.1	Nombre de grilles de Cascades résolues	123
7.2.2.2	Nombre moyen d'étapes de résolution	125

L'Homme est une prison où l'âme reste libre.

— Victor Hugo

L'application résultant du modèle des *Cascades* a été développée en Java, à l'aide de la librairie *GraphStream* afin d'assurer la gestion du graphe de stratégies ([33], [32]).

La fenêtre principale (voir figure 7.1) se compose de la grille des *Cascades* en haut à droite, ainsi que du graphe de stratégies en haut à gauche. Au sein de la grille, les nombres donnés en début d'exercice sont en noir là où les nombres ajoutés par le programme sont affichés en gris. En ce qui concerne le graphe de stratégies, plus les nœuds sont rouges et plus leur valeur cognitive k est élevée. De la même manière, plus les arêtes sont vertes et plus leur valeur émotionnelle η est grande. La zone en bas à gauche de la fenêtre est réservée aux patterns

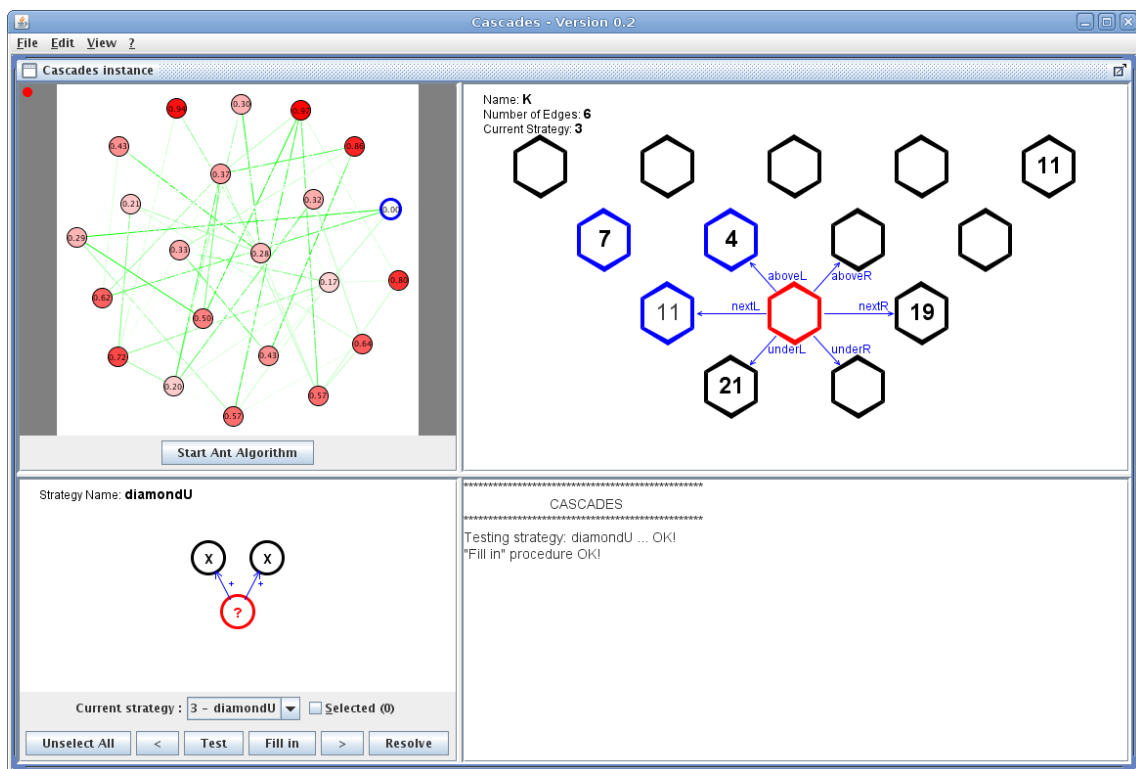


FIG. 7.1 – FENÊTRE PRINCIPALE DE L'APPLICATION

de stratégies. Enfin, en bas à droite se trouve un panneau dans lequel s'inscrit progressivement l'état de l'application en temps réel.

Je vais dans un premier temps effectuer la description d'un exemple typique de résolution par l'exposition des éléments majeurs de l'algorithme, parmi lesquels les listes de stratégies ainsi que les graphes et grilles des Cascades. Dans un second temps, il va s'agir de généraliser l'approche par des plans d'expérience dont l'intérêt réside dans l'étude de la sensibilité aux paramètres de l'algorithme, notamment les coefficients de pénalisation utilisés.

7.1 EXEMPLE DE RÉOLUTION

Cet exemple de simulation débute par l'utilisation d'un graphe de stratégies aléatoire (voir figure 7.2), ce qui signifie que le graphe de stratégies est initialisé avec des valeurs aléatoires sur les nœuds et sur les arêtes. Nous aurions cependant très bien pu charger en mémoire un graphe à l'aide d'un fichier en y inscrivant au préalable toutes les valeurs sur les sommets et arêtes.

7.1.1 Étape 1

Pendant son exécution, l'algorithme tente de trouver une liste de stratégies potentielles afin de remplir la grille des Cascades. À l'issue de la première étape, le programme propose la liste [*uline3L*, *hline4L*, *uline4R*, *dline4R*, *diamondR*] et affiche les résultats tels que représentés sur la figure 7.3. On y distingue deux graphes représentant l'état du graphe de stratégies avant et après la pénalisation des stratégies sélectionnées dans le cadre de la résolution. Pour comprendre le déroulement de l'algorithme, il faut bien saisir le fait que l'étape de sélection débute par l'identification de l'arête la plus forte. Il s'agit dans cet exemple de celle reliant les stratégies *uline3L* et *hline4L*, ce qui explique leur apparition en début de liste de sélection. Ensuite est ajoutée la stratégie *uline4R* reliée à *hline4L* par l'arête la plus forte **parmi les arêtes connectées à l'une des deux stratégies de la liste**. De la même façon on été ajoutées *dline4R* et *diamondR* jusqu'à atteindre la taille maximale de la liste, c'est-à-dire 5 stratégies — valeur qui est bien sûr paramétrable.

On constate par ailleurs qu'aucune case de la grille des Cascades n'est remplie. En effet, aucune des stratégies sélectionnées ne permet de calculer le contenu ne serait-ce que d'une seule case de la grille, le *pattern-matching* d'aucune d'entre elles ne pouvant s'appliquer ici. Il faut donc pénaliser les sommets du graphe représentant ces stratégies en leur appliquant un certain coefficient Z_n — dans cet exemple de résolution, on ne garde que 25% de la valeur initiale des sommets sélectionnés. Cela se traduit sur le graphe par une diminution de la couleur rouge des sommets concernés. Le programme peut alors entamer une nouvelle étape de résolution par une nouvelle exécution de l'algorithme fourni.

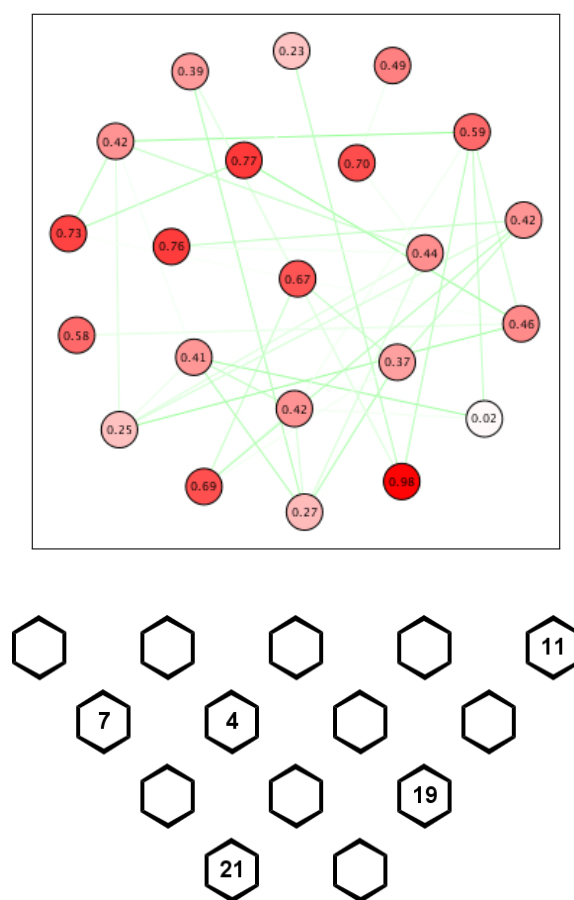


FIG. 7.2 – ÉTAPE INITIALE

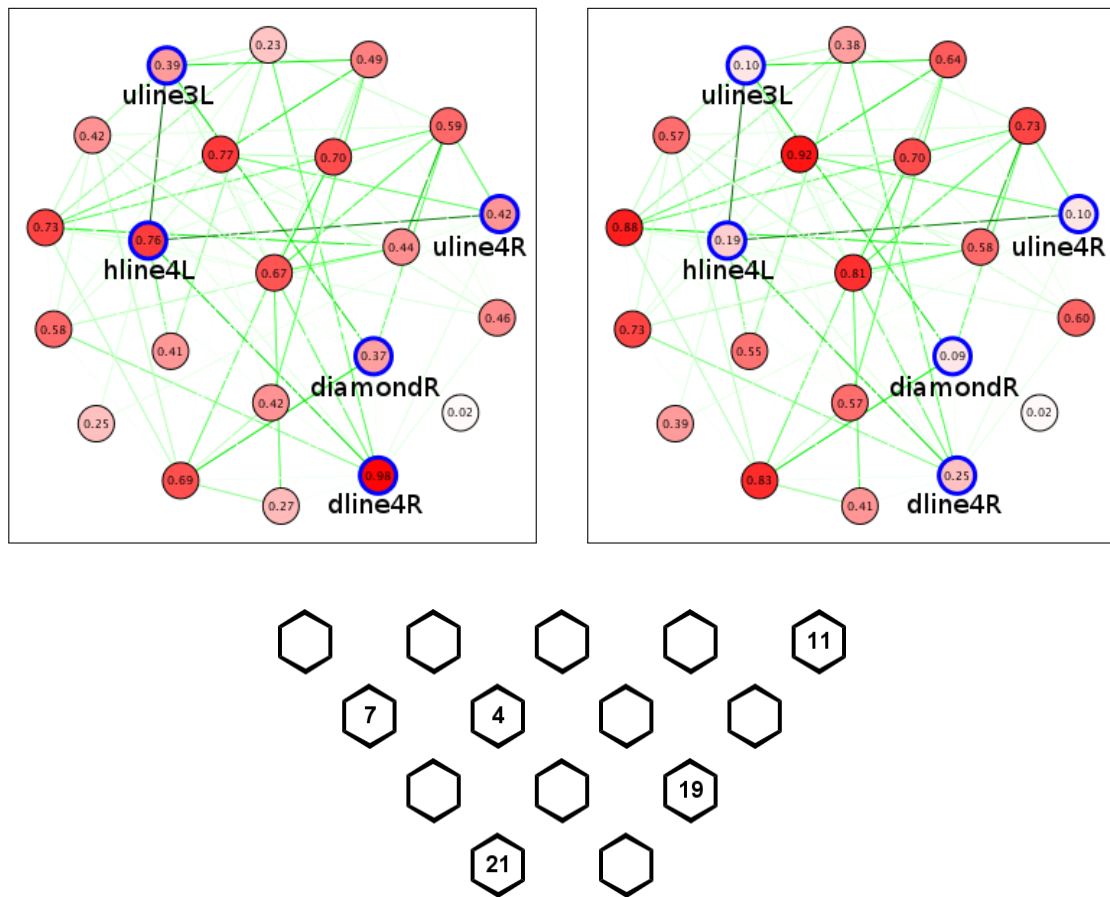


FIG. 7.3 – GRAPHE AVANT ET APRÈS LA PÉNALISATION ET GRILLE DE L'ÉTAPE 1

7.1.2 Étape 2

La seconde étape de résolution s'achève par la sélection des stratégies suivantes : $[dline4R, hline5L, pyramidA, uline3L, diamondI]$ et affiche le graphe de la figure 7.4. L'évolution du graphe montre bien la pénalisation des stratégies. En outre, les grilles calculées par l'algorithme sont différentes, c'est-à-dire que l'ordre d'application des stratégies sélectionnées a un impact sur l'état de la grille de Cascades finale.

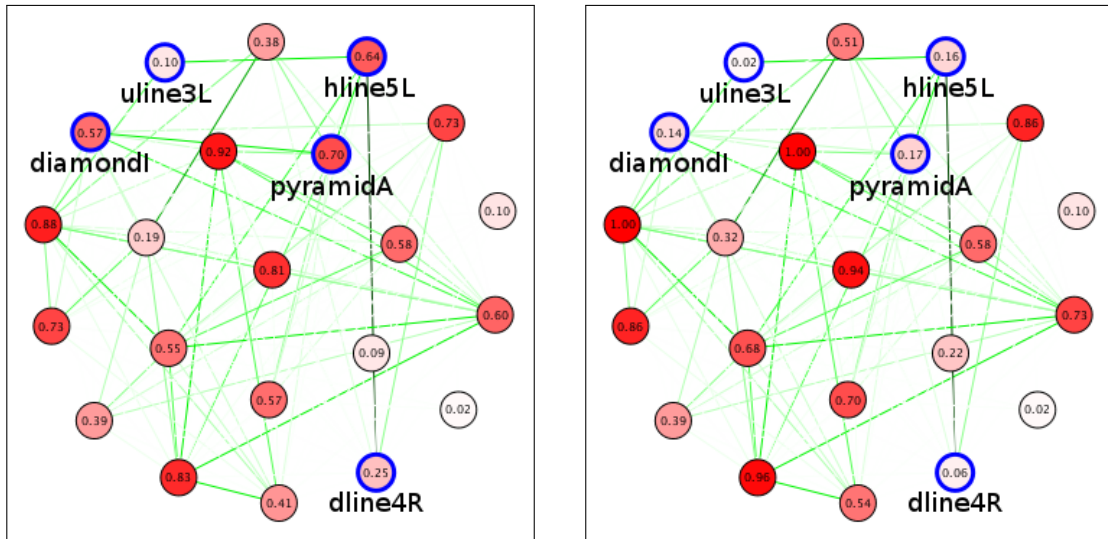
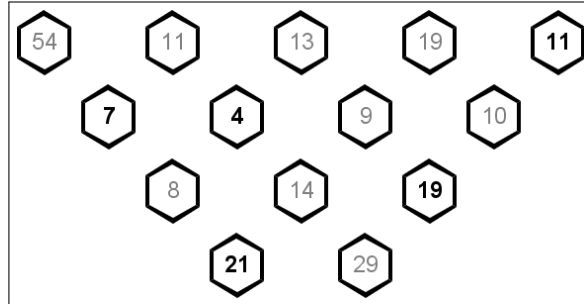


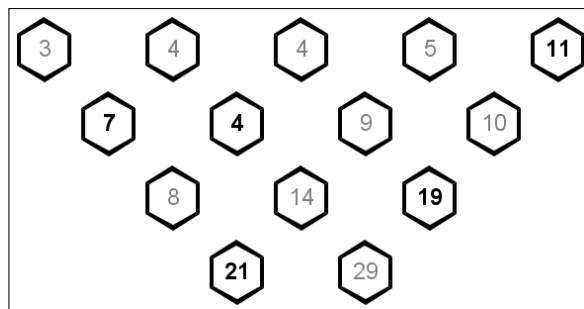
FIG. 7.4 – ÉVOLUTION DU GRAPHE DE L'ÉTAPE 2 AVANT ET APRÈS LA PÉNALISATION

En effet, bien que l'ensemble des stratégies sélectionnées suffise à remplir totalement la grille, cela ne suffit pas car la solution trouvée se doit d'être déterministe. À titre d'exemple, étudions le cas dans lequel la liste de stratégies est appliquée dans cet ordre : $[dline4R, hline5L, pyramidA, uline3L, diamondI]$, puis dans ce nouvel ordre : $[dline4R, hline5L, uline3L, diamondI, pyramidA]$. Les deux grilles générées sont alors très différentes (voir figure 7.5).

À l'issue de l'algorithme s'effectue la pénalisation des stratégies sélectionnées, comme lors de l'étape précédente. Cependant, et contrairement à l'étape précédente, l'ordre des stratégies ayant un impact sur le résultat final, certaines arêtes seront elles aussi modifiées. Et plus précisément les arêtes reliant la dernière stratégie ajoutée, à savoir *diamondI*, aux autres stratégies sélectionnées, car *diamondI* est alors considérée comme incompatible avec ses consœurs et doit affaiblir ses liens avec ces dernières. Ainsi, les arêtes concernées seront pénalisées d'un facteur Z_e qui dans cet exemple vaut 20%.



Liste : [*dline4R, hline5L, pyramidA, uline3L, diamondI*]



Liste : [*dline4R, hline5L, uline3L, diamondI, pyramidA*]

FIG. 7.5 – EXEMPLES DE DEUX RÉOLUTIONS DE GRILLES À L'ÉTAPE 2

7.1.3 Étape finale

Au bout de 13 étapes de résolution, l'algorithme propose finalement la liste suivante : [*hline5R, uline4R, uline4L, diamondI*], liste qui permet au programme non seulement de remplir complètement la grille, mais de manière unique, indépendamment de l'ordre des stratégies (voir figure 7.6).

La grille finale obtenue est donc parfaitement cohérente avec la liste de stratégies utilisée, malgré le fait qu'elle soit fautive compte tenu de l'exercice demandé. Néanmoins, il ne faut pas perdre de vue la finalité du modèle qui, à partir d'une certaine interprétation de l'exercice, fût-elle incorrecte, permet à un élève de trouver une solution au problème des Cascades.

La résolution décrite ici se conclut au bout de 13 étapes, ce qui est plutôt rapide en moyenne (nous le verrons juste après). Il arrive pourtant que l'algorithme tourne indéfiniment sans parvenir à extraire du graphe une liste de stratégies convenable. Ce cas est relativement rare mais apparaît sous certaines conditions, notamment lorsque les paramètres utilisés pour la résolution sont mal choisis et ne permettent pas au graphe de faire émerger une solution appropriée.

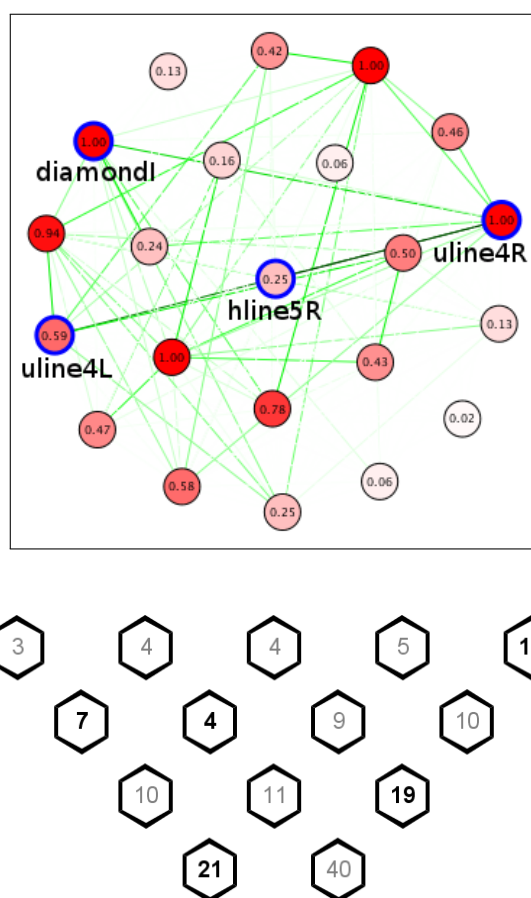


FIG. 7.6 – ÉTAT DU GRAPHE ET DE LA GRILLE À L'ÉTAPE FINALE

7.2 PLANS D'EXPÉRIENCE

La sensibilité aux paramètres est un problème récurrent dans les modèles à base d'intelligence collective. En effet, bien qu'une légère fluctuation de valeur n'occasionne pas nécessairement de profonds changements qualitatifs, elle peut altérer grandement la dynamique du système. À titre d'exemple, si j'avais modifié ci-dessus le paramètre Z_n — relatif à la pénalisation des sommets du graphe — le faisant passer de 25% à 20%, le programme aurait pu se conclure en 6 étapes seulement au lieu de 13, sans pour autant proposer une grille finale différente.

D'un autre côté, il reste une question majeure à laquelle il manque une réponse : quel est le rôle joué par l'émotion dans ce modèle ? En effet, bien que l'on ait incarné l'émotion par une valuation des arêtes du graphe de stratégies, quelle serait la conséquence de la suppression pure et simple de ce dispositif ? Nous répondrons à ces questions à l'issue d'une comparaison entre deux modèles de résolution du problème des Cascades : l'un, traditionnel, tenant compte des aspects émotionnels aussi bien que cognitifs au sein du graphe de stratégies, et l'autre, plus statique, prenant uniquement en considération le côté cognitif.

7.2.1 Sensibilité aux paramètres

De manière générale, les paramètres d'un modèle n'ont pas tous la même importance car leur perturbation ne provoque pas le même impact sur l'évolution du modèle. Il en est de même pour le problème des Cascades implanté ici. Les graphiques qui vont suivre montrent les résultats obtenus à partir d'une série d'expériences effectuées à l'aide de paramètres différents.

7.2.1.1 Nombre de grilles de Cascades résolues

L'objectif de la première expérience consiste à mesurer l'efficacité du modèle en fonction des paramètres. Pour ce faire, le programme compte le nombre de grilles de Cascades que le modèle est capable de résoudre pour chaque valeur de paramètre. Un nombre fixé de graphes de stratégies — en l'occurrence ici 50 — générés aléatoirement en début de simulation vont être testés. Si le nombre d'étapes de résolution dépasse un certain seuil — 150 étapes dans cet exemple —, la résolution est considérée comme ayant échoué. Si à l'inverse celle-ci s'achève par le remplissage en règle de la grille de Cascades, il s'agira alors d'un succès qui sera comptabilisé comme tel.

La figure 7.7 représente un ensemble de tentatives de résolution du problème des Cascades réalisées en changeant successivement les paramètres Z_n (pénalisation des nœuds), Z_e (pénalisation des arêtes) et Q (quantité d'ajout de phéromones sur les arêtes par les fourmis). Pour chaque graphique, on mesure

en ordonnée le nombre de résolutions fructueuses parmi 50 tests ; les valeurs en abscisse représentent l'évolution du paramètre Z_n dans l'intervalle $[0;1]$, tandis que les différentes courbes de couleur désignent chacune une valeur de Z_e , toujours entre 0 et 1. Enfin, chaque graphique illustre une valeur différente de Q allant de 0.2 en haut à gauche, à 0.8 en bas à droite.

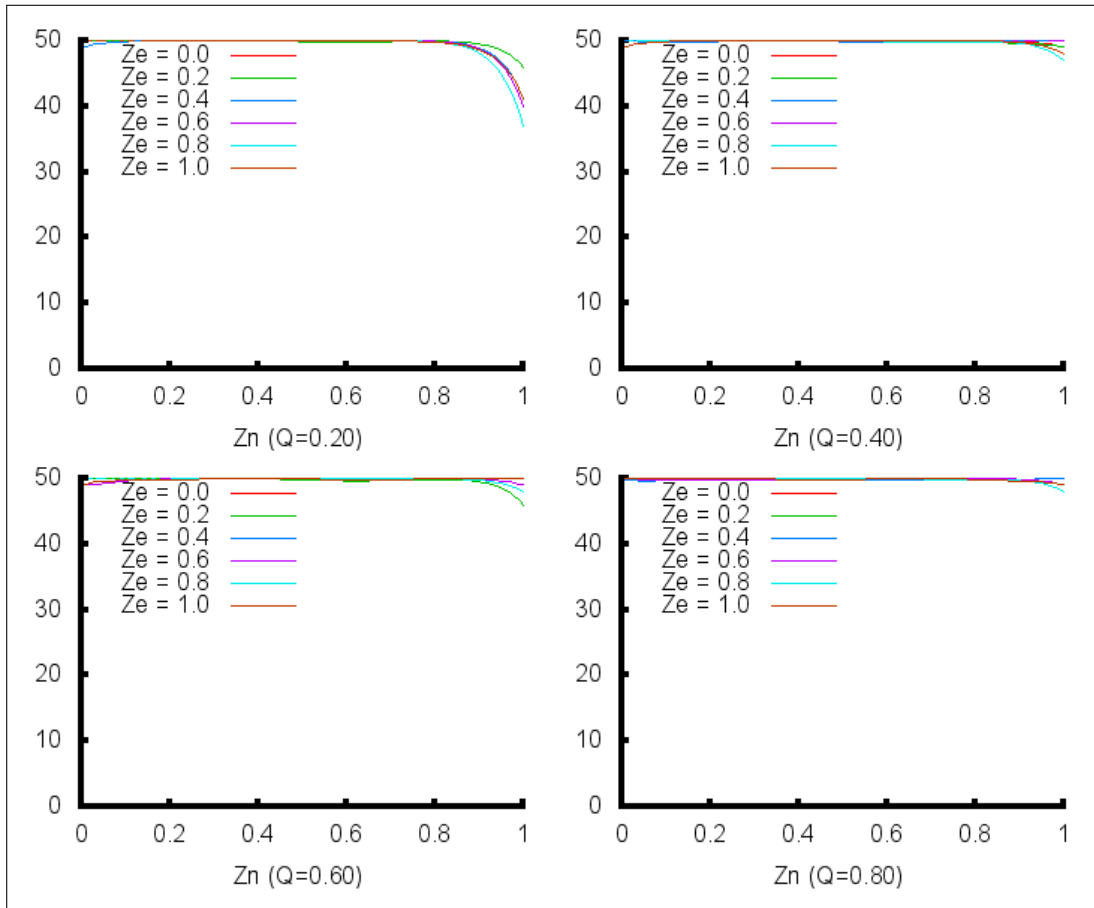


FIG. 7.7 – NOMBRE DE GRILLES RÉSOLUES (SUR 50 GRILLES) EN FONCTION DE Z_n , POUR DIFFÉRENTES VALEURS DE Z_e ET Q

Les résultats obtenus nous permettent de faire plusieurs remarques. Tout d'abord, il faut noter que l'algorithme utilisé pour la résolution des Cascades est particulièrement efficace dans la mesure où toutes les valeurs de paramètres permettent de résoudre pratiquement 100% des grilles testées. En outre, l'impact des différents paramètres est quasiment nul à moins de considérer les extremums, notamment pour $Z_n = 1$ (c'est-à-dire pour une pénalisation inexistante des stratégies). Cela signifie que l'algorithme utilisé n'est en général pas influencé qualitativement par les variations de paramètres et témoigne de la robustesse du modèle.

On constate néanmoins de meilleurs résultats aux extremums lorsque Q est élevée. Ceci est cohérent dans la mesure où le dépôt de phéromones fait alors indirectement office de balance à l'absence (ou à l'excès) de pénalisation sur les sommets.

7.2.1.2 Nombre moyen d'étapes de résolution

Afin de mesurer plus précisément la qualité des solutions obtenues, le programme a également pris en compte le nombre d'étapes de résolution nécessaires au remplissage des différentes grilles de Cascades. L'expérience réalisée est donc rigoureusement identique à la précédente. Pour différentes valeurs de paramètres, différents graphes de stratégies sont testés et analysés et on mesure la moyenne des nombres d'étapes, en sachant que l'on ne considère que les grilles terminées avec succès (c'est-à-dire pratiquement toutes pour cette expérience !) dans le calcul de la moyenne (voir figure 7.8).

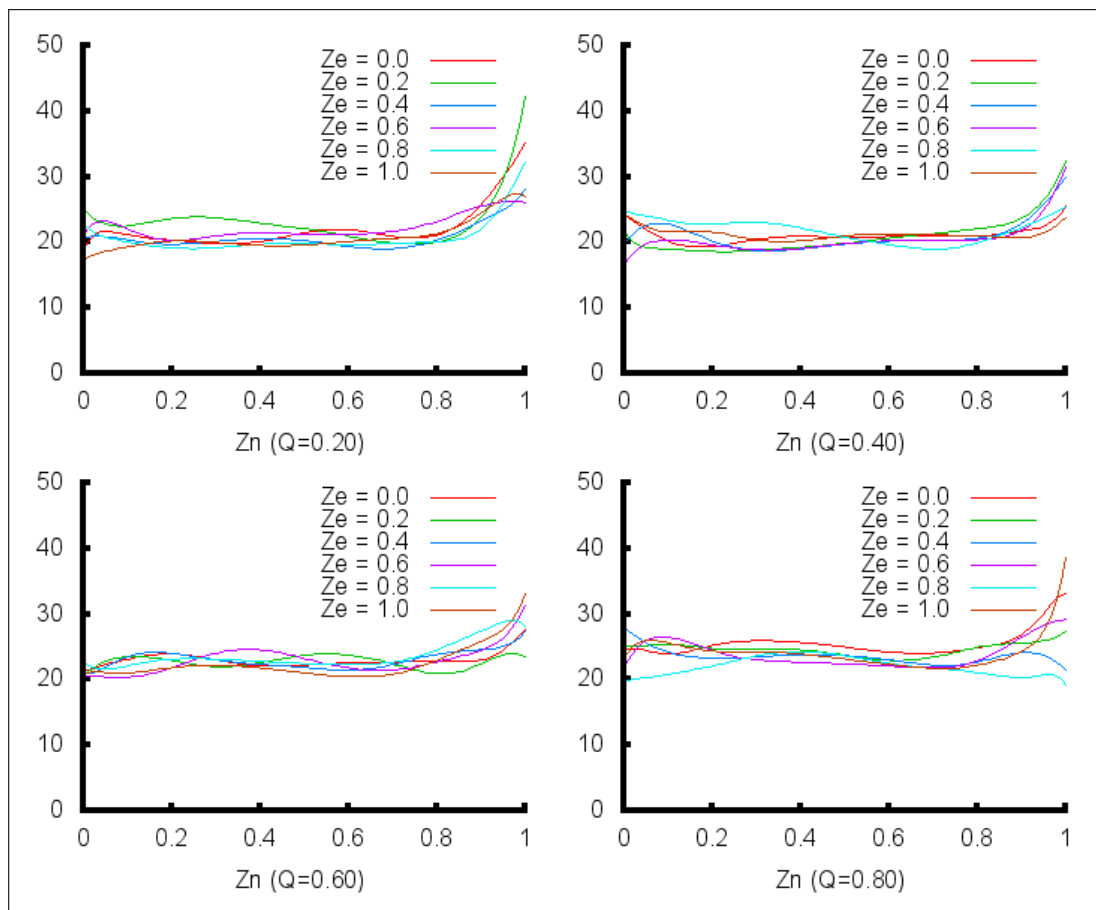


FIG. 7.8 – NOMBRE MOYEN D'ÉTAPES DE RÉOLUTION EN FONCTION DE Z_n , POUR DIFFÉRENTES VALEURS DE Z_e ET Q

Là encore, on constate une certaine stabilité des résultats malgré les variations de paramètres. Z_e et Z_n n'ont toujours que peu d'impact sur la simulation, excepté pour des valeurs extrêmes de Z_n . Toutefois, une augmentation de Q provoque un léger accroissement du nombre d'étapes de résolution, ce qui, malgré le paradoxe apparent compte tenu des résultats précédents — nous avons démontré qu'une forte valeur de Q réduisait les erreurs introduites par des valeurs de paramètres extrêmes —, se révèle une nouvelle fois parfaitement cohérent : l'affaiblissement du dépôt de phéromones a un effet retardateur sur la dynamique du graphe de stratégies qui à son tour provoque un ralentissement dans l'extraction des stratégies utiles à la résolution. Ainsi, pour $Q = 0.2$, environ 20 étapes sont nécessaires contre 23 étapes pour $Q = 0.8$. L'augmentation du nombre d'étapes de résolution est cependant relativement faible et finalement assez peu significative.

7.2.2 Impact de l'émotion

Nous avons vu précédemment que ce modèle assimilait l'émotion à l'action des phéromones sur le graphe, ce qui permet d'obtenir une dynamique aussi bien qualitative — la structure du graphe dépend de la présence ou non des arêtes entre les sommets — que quantitative — les valeurs sur les arêtes viennent renforcer l'importance de certains chemins du graphe. Nous avons également défini les aspects cognitifs comme étant présents sur les sommets du graphe, car d'une part leur contenu permet de produire la connaissance nécessaire à la résolution de la grille des Cascades, mais d'autre part leur pondération au sein du graphe permet d'introduire la dynamique essentielle à la sélection des solutions les plus pertinentes.

Parallèlement à cette première simulation, une deuxième expérience a été effectuée, dans laquelle les aspects émotionnels sont supprimés, au profit du contenu cognitif. Pour ce faire, la sélection des stratégies ne se fait désormais plus en fonction des valuations sur les arêtes comme cela était fait auparavant, mais directement par le classement des sommets en fonction de leur valeur cognitive, comme le décrit l'algorithme 3.

Toutefois, l'émotion entre tout de même en jeu lors du déroulement de l'algorithme puisque l'on utilise les valuations sur les arêtes afin de déterminer le voisinage des stratégies sélectionnées. Ce voisinage permet ainsi, en cas de pénalisation des sommets, de redistribuer les pertes aux sommets les plus proches, c'est-à-dire ceux dont les arêtes intermédiaires sont évaluées au-delà d'un certain seuil prédéfini. Bien qu'importants, les enjeux liés au rôle joué ici par l'émotion restent très limités face aux impacts cognitifs développés dans cette seconde expérience.

Algorithme 3 extraireStrategie2 (L : liste de stratégies)

```

1:  $s \leftarrow null$  // stratégie temporaire
2:  $maxS \leftarrow null$  // sommet maximal
3: Si ( $L$  est vide) Alors
4:   Pour (chaque sommet  $s$  du graphe) Faire
5:     trouver le sommet  $s$  le plus fort et le mettre dans  $maxS$ 
6:   Fin Pour
7: Sinon
8:   Pour (chaque sommet  $s$  du graphe) Faire
9:     Si ( $s$  n'est pas dans  $L$ ) Alors
10:      Si ( $s$  est plus fort que  $maxS$ ) Alors
11:         $maxS \leftarrow s$ 
12:      Fin Si
13:    Fin Si
14:  Fin Pour
15: Fin Si
16: Retourne  $maxS$ 

```

7.2.2.1 Nombre de grilles de Cascades résolues

À l'instar de la simulation précédente, l'efficacité de l'algorithme sera mesurée au travers du nombre de graphes de stratégies ayant fait l'objet d'un succès, et plus précisément à l'aide du nombre moyen d'étapes de résolution nécessaires. Les résultats sont profondément différents, comme le révèle la figure 7.9.

Contrairement aux graphiques précédents, Z_n joue un rôle majeur dans l'évolution du nombre de grilles résolues, et pour cause : la pénalisation des sommets affecte directement la sélection des stratégies puisqu'elle modifie la valeur des sommets qui elle-même conditionne cette sélection. Ainsi, pour une valeur de Z_n dans l'intervalle $[0; 0.75]$ seules 40% à 60% des grilles sont résolues, avec un minimum local pour la valeur $Z_n = 0.4$. Au-delà de 0.75, la fonction atteint son maximum aux alentours de $Z_n = 0.85$ avec un peu plus de 80% de grilles résolues. Cette valeur de Z_n correspond à une pénalisation « idéale » des sommets qui permet à la sélection de stratégies de s'effectuer correctement. Puis à partir de $Z_n = 0.85$, la fonction entame une chute vertigineuse jusqu'à atteindre 0% en fin d'intervalle ($Z_n = 1$). Ce dernier résultat nous conforte dans l'idée qu'une absence totale de pénalisation sans compensation de la part de l'émotionnalité introduite par les arêtes du graphe rend celui-ci parfaitement inerte, et donc inefficace.

Par ailleurs, le paramètre Q ne joue cette fois-ci absolument aucun rôle dans la mesure où le dépôt de phéromone n'a qu'une fonction structurelle au sein du graphe, et n'est d'aucun secours dans la sélection des stratégies.

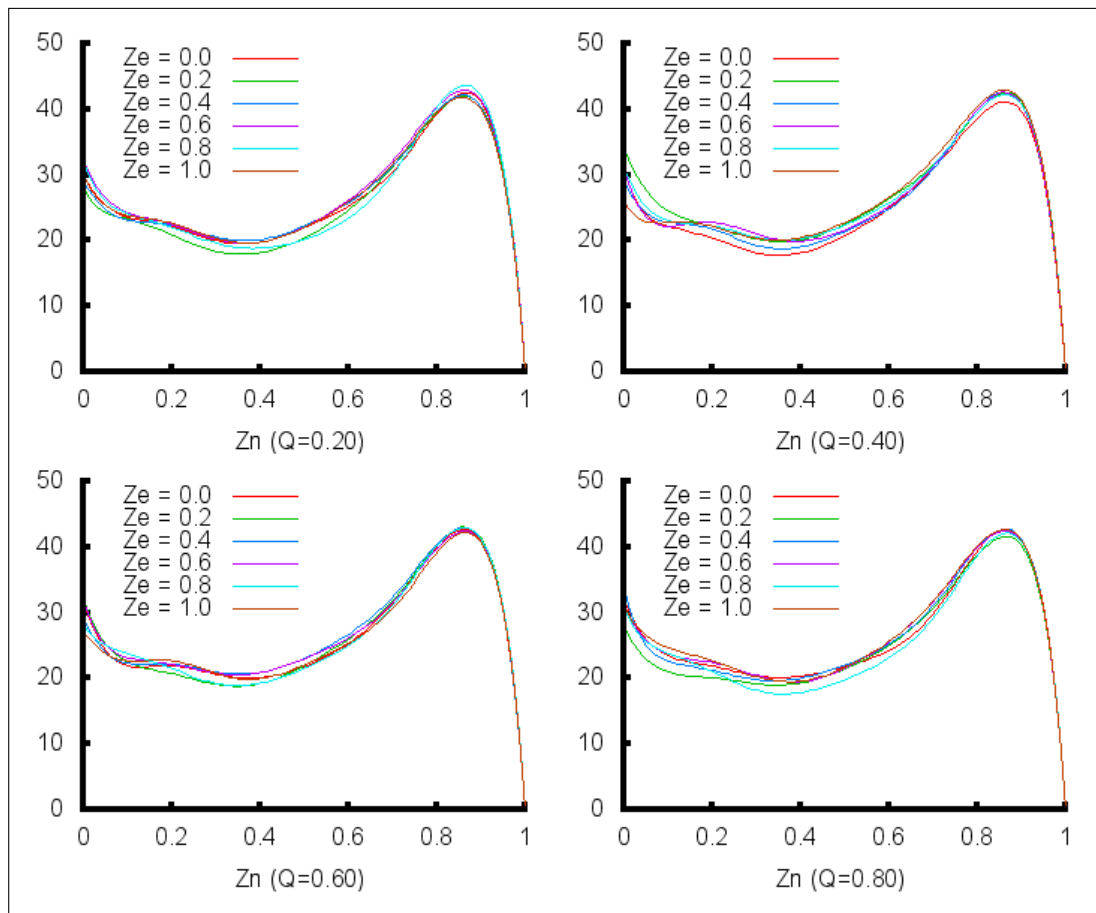


FIG. 7.9 – NOMBRE DE GRILLES RÉSOUES (SUR 50 GRILLES) EN FONCTION DE Z_n , POUR DIFFÉRENTES VALEURS DE Z_e ET Q (VERSION COGNITIVE)

7.2.2.2 Nombre moyen d'étapes de résolution

De la même façon qu'auparavant, le nombre moyen d'étapes de résolution a été étudié pour déterminer avec plus de précision le comportement de ce modèle comparativement au précédent (voir figure 7.10). Là encore, on ne considère que les grilles terminées avec succès, ce qui a une réelle incidence ici.

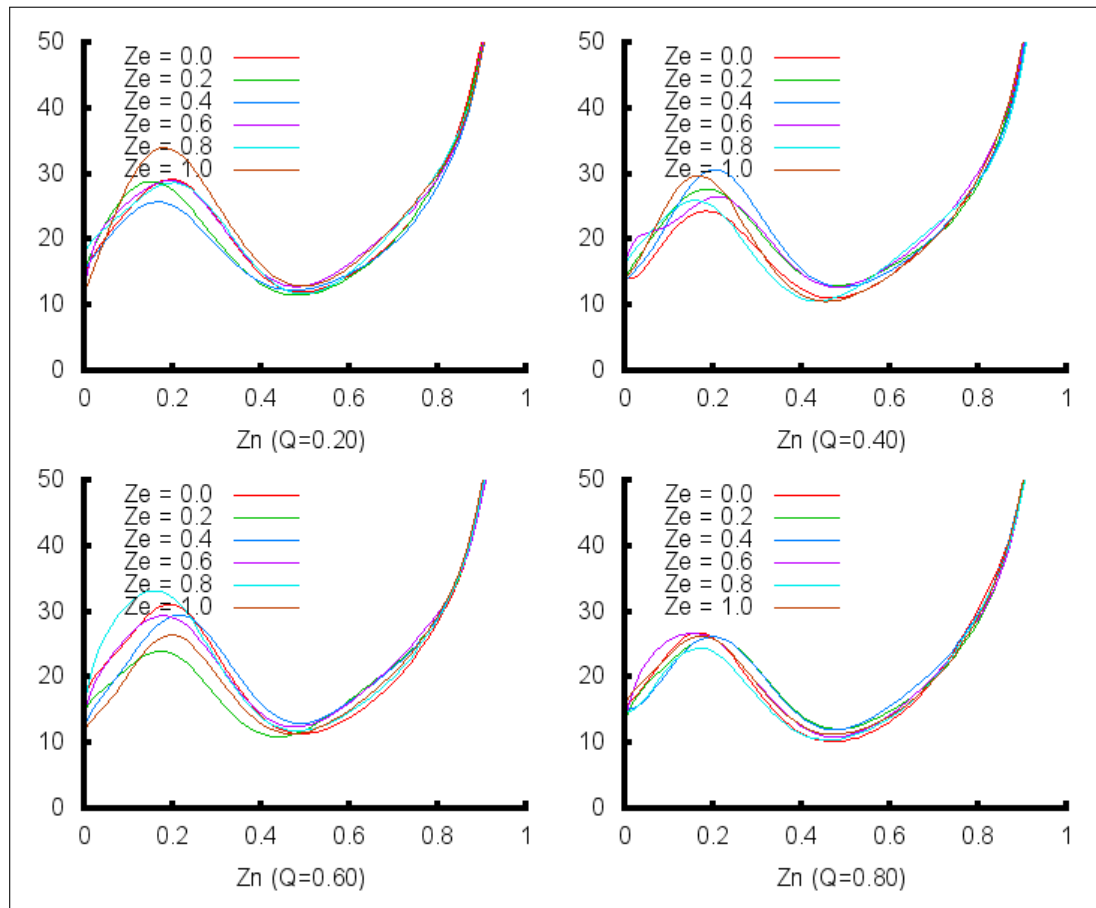


FIG. 7.10 – NOMBRE MOYEN D'ÉTAPES DE RÉOLUTION EN FONCTION DE Z_n , POUR DIFFÉRENTES VALEURS DE Z_e ET Q (VERSION COGNITIVE)

Sur la majeure partie des courbes, pour Z_n appartenant à l'intervalle $[0;0.8]$, le nombre moyen d'étapes oscille entre 12 et 28 environ, le minimum — 12 étapes donc — étant atteint soit dans le cas d'une pénalisation totale des sommets sélectionnés ($Z_n = 0$) soit pour une pénalisation moyenne ($Z_n = 0.5$). Malgré ce score somme toute meilleur que celui du modèle précédent, qui était de 20 étapes en moyenne, le fait de recouper les résultats obtenus avec les graphiques précédents révèle une relative faiblesse de ce modèle cognitif par rapport à son *alter ego* émotionnel.

Certes, 12 étapes représentent une excellente performance compte tenu de

la combinatoire qu'impose le problème des Cascades, mais le nombre de grilles résolues sous ces conditions n'est que de 60% dans le premier cas (pour $Z_n = 0$) et de guère plus de 40% dans le second cas ($Z_n = 0.5$). La quasi-totalité des grilles résolues assurées par le modèle émotionnel confirme l'adaptabilité apportée par la dynamique des arêtes du graphe. Quant à la position $Z_n = 0.85$ qui permet au programme de résoudre plus de 80% des grilles, à savoir le maximum dans cette expérience, il ne faut pas perdre de vue que le nombre moyen d'étapes associé est de 35 environ, ce qui représente une augmentation de 75%. Ce modèle parvient donc à obtenir un pourcentage de grilles résolues relativement élevé, mais au détriment de l'efficacité de résolution.

Jusqu'à présent, la modification de Z_e (pénalisation des arêtes) n'avait eu aucune répercussion visible sur les résultats obtenus. Dans ce dernier exemple, on constate une assez grande variation de niveau aux alentours de $Z_n = 0.2$, faisant passer le nombre moyen d'étapes de 23 à 35 en fonction des graphiques. Cette variation assez importante témoigne du bruit introduit par la pénalisation des arêtes, signe qu'une perturbation de la diffusion des valeurs cognitives, seul rôle joué par l'émotion au cours de l'algorithme, peut grandement altérer l'issue de la résolution. Elle n'a néanmoins aucun impact sur le nombre de grilles résolues.

CONCLUSION ET PERSPECTIVES

8

SOMMAIRE

8.1 CONCLUSION	127
8.2 PERSPECTIVES	128

*We ought to hate very rarely, as it is too fatiguing;
remain indifferent to a great deal, forgive often and
never forget.^a*

— Sarah BERNHARDT

a. « Il faut haïr très peu, car c'est trop fatigant ; il faut ignorer beaucoup, pardonner souvent, et ne jamais oublier. » (NDR)

8.1 CONCLUSION

Comme nous venons de le voir, la modification des paramètres utilisés au sein d'un programme peut influencer profondément la dynamique du graphe, surtout dans le cadre d'un modèle purement cognitif, au point que l'on peut parler littéralement de « profil » d'individu. De la même manière, à paramètres égaux, l'état initial du graphe de stratégies peut grandement déterminer l'issue de la résolution. En effet, bien que la variation des paramètres étudiés ne provoque pas toujours des bouleversements significatifs, il ne faut pas perdre de vue l'aspect probabiliste du système. L'exemple de résolution décrit à la section 7.1 prouve qu'un graphe de stratégie peut très bien se conclure en 13 étapes sous certaines conditions et plus tard en 6 étapes sous des conditions très légèrement différentes.

C'est là qu'entre en jeu l'émotion ! Malgré la grande variété des paramètres testés, les divers graphes de stratégies expérimentés — générés aléatoirement de surcroît — se sont tous soldés par un succès, c'est-à-dire qu'ils ont tous trouvé une solution en moins de 150 étapes de résolution — à quelques paramètres extrêmes près — et en seulement une vingtaine d'étapes en moyenne. Ce constat n'est que l'illustration de la grande capacité d'adaptation dont est capable de faire preuve ce modèle émotionnel.

À l'inverse, la version cognitive de ce modèle, plus rigide et moins dynamique se révèle être plus décevante. Certes, elle permet d'obtenir plus de 80% de réussite, mais seulement sous contraintes et sans proposer d'efficacité de résolution réellement plus intéressante.

En conclusion, nous pouvons dire que ce modèle met en avant un mécanisme d'émergence de solution au sein d'un problème modélisé à l'aide d'un graphe. Le travail ainsi produit permet de fournir un outil de représentation susceptible d'être utilisé dans le cadre d'éventuelles expérimentations dont la réalisation dépasse le cadre de cette thèse. Néanmoins, le libre paramétrage de la majeure partie des variables du programme permet finalement de rapidement se rapprocher de résultats d'expérience, et ainsi favoriser la création de profils d'individus.

8.2 PERSPECTIVES

Comme souvent dans le développement de modèles qui par définition sont réductionnistes, plusieurs fonctionnalités potentielles peuvent être imaginées, permettant de se rapprocher davantage de la réalité.

Par exemple, il peut s'agir d'améliorations au niveau de la gestion des incohérences lors de la résolution. En effet, un élève aura tendance à faire assez souvent des erreurs de calcul, ou bien à changer de stratégie au cours de la résolution, sans effacer les cases déjà remplies, ou encore à ajouter un nombre aléatoirement dans une case afin de débloquer une situation difficile et générer ainsi des nombres de manière non déterministe, comme illustré par la figure 6.7, vue précédemment, qui traite du mécanisme d'inférence mis en place potentiellement par les élèves.

De même, pourquoi ne pas considérer que l'algorithme de résolution d'une grille de Cascades est terminé dès que l'on a trouvé une solution, sans tenir compte de l'ordre dans lequel les stratégies sélectionnées sont appliquées ? Cela serait plus conforme aux réactions générales des élèves mais poserait un ensemble de problèmes additionnels difficiles à modéliser. Par exemple, il faudrait de toutes manières décider de l'ordre de la liste de stratégies mais dans ce cas, il faudrait introduire une notion de préférence qui permettrait de définir les différentes priorités dans les stratégies testées, ce qui implique une nouvelle dimension cognitive complexe et fortement liée au profil de l'individu. D'où l'intérêt de

ne se limiter qu'à une liste « simple » sans notion de priorité, mais qui implique alors une sélection de stratégies non conflictuelles. Un autre exemple d'adaptation nécessaire serait la détermination précise de la localisation des cases à remplir pour le *pattern-matching* des stratégies. En clair, si on introduit une préférence pour le choix dans l'ordre des stratégies à appliquer, pourquoi ne pas faire de même pour la position de la case à remplir ? Ceci n'est valable bien sûr que s'il existe plusieurs choix de cases à partir d'une seule stratégie.

III

ÉMOTION ET LOGIQUE FLOUE

Le travail présenté dans cette partie correspond au développement d'un environnement de simulation de type proies-prédateurs utilisé au sein de deux modèles. Le premier correspond à un travail de collaboration avec William BLEWITT de l'université De Monfort (Leicester, Royaume-Uni) intitulé "*Psychologically Grounded Emotion Model for a NPC Fuzzy Controller*"¹ (voir annexe C). Le second, appelé GAEA, fait l'objet de la quatrième partie de ce document et s'applique à la modélisation d'une entité génétiquement adaptée (voir partie IV).

Le modèle émotionnel de BLEWITT est à base de logique floue et s'inspire du modèle de MILLENSON ([91]) afin de définir un état émotionnel, tandis que le second modèle propose une vision émergentiste de l'émotion à base de programmation génétique.

Dans le cadre de ces deux travaux, on modélise un environnement centré sur un individu plongé dans un environnement hostile, le but pour l'individu étant d'optimiser sa survie. Il s'agit d'un système proie-prédateur dans lequel l'individu central se devra de se nourrir de proies afin d'augmenter son capital d'énergie vitale simulé ici par une jauge de vie, tout en évitant ses prédateurs qui n'hésiteront pas à lui ôter la vie. La figure 8.1 représente cet environnement.

En somme, l'objectif visé est de simuler le comportement d'un robot équipé de capteurs et d'effecteurs, c'est-à-dire de capacités de récupération d'informations de son environnement ainsi que de capacités d'actions, tout en essayant de s'adapter au maximum aux aléas situationnels invoqués par l'instabilité du milieu. Le terme de *robot* prend ici tout son sens dans la mesure où il s'agit de développer des aspects décisionnels chez une entité artificielle en environnement difficile et qui se doit d'évoluer selon un objectif très précis — sa propre survie en l'occurrence.

Dans un premier temps, je présenterai l'environnement en termes de représentation objet et d'affichage, avant de présenter brièvement le modèle émotionnel sous-jacent qui non seulement détermine l'émotion du robot, mais aussi son comportement. La dernière partie sera, quant à elle, réservée à l'implantation ainsi qu'aux résultats obtenus.

1. « Modèle psychologique de l'émotion pour le contrôle flou d'un personnage non joueur »

REPRÉSENTATION DE L'ENVIRONNEMENT

9

SOMMAIRE

9.1	DESCRIPTION DU MILIEU	135
9.2	ROBOT	137
9.3	IMPLANTATION DU ROBOT	138

Aucun mortel ne peut garder un secret. Si les lèvres restent silencieuses, ce sont les doigts qui parlent. La trahison suinte par tous les pores de sa peau.

— Sigmund FREUD

L'environnement de la simulation, tel que décrit précédemment, possède un certain nombre d'entités en interaction permanente, à savoir des proies, des prédateurs, et bien sûr, l'entité robotisée que l'on cherche à modéliser. Toutes ces entités correspondent à ce que l'on va appeler des *créatures* de l'environnement, c'est-à-dire qu'elles sont non seulement des objets ayant une position dans l'espace, mais elles sont également dotées de capacités supplémentaires de perception et de mobilité.

9.1 DESCRIPTION DU MILIEU

Dans un premier temps, comme leur nom l'indique, les proies sont situées tout en bas de la chaîne alimentaire et sont par ailleurs les plus nombreuses. Leur principal rôle est d'alimenter le robot en lui fournissant l'énergie nécessaire à sa survie. Les proies se déplacent systématiquement en groupe selon un algorithme d'intelligence en essaim de *boids* décrit par Craig REYNOLDS ([107]).

Un **boïd** est une entité se déplaçant en groupe par attraction et répulsion de ses congénères, de ses prédateurs et d'éventuels éléments du décor, en fonction de son voisinage (voir figure 9.1). Il possède un champ de vision limité (en jaune sur la figure) qui lui permet de détecter la présence de congénères ou d'ennemis dans son voisinage visuel. Par ailleurs, il dispose d'une région tout autour de lui décrivant un cercle (en orange sur la figure) qui fait office de zone de détection auditive d'individus amis ou ennemis.

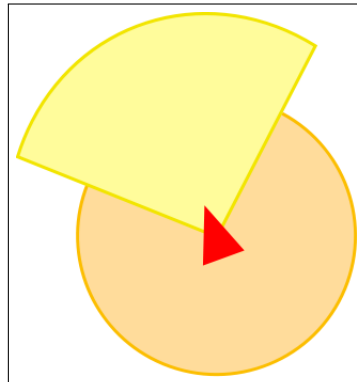
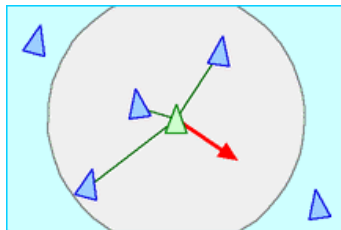
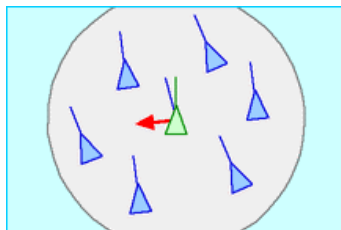


FIG. 9.1 – VOISINAGE D'UN INDIVIDU

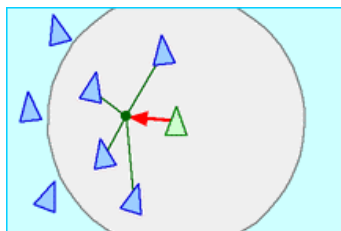
Les forces guidant le mouvement des *boïds* leur permettent de conserver une forme en essaim malgré l'éventuelle présence de prédateurs. Elles sont au nombre de trois.



Séparation : force qui évite la trop grande proximité des *boïds* entre eux



Alignement : force qui permet à chaque *boïd* d'aligner son vecteur déplacement sur celui de ses voisins



Cohésion : force qui permet à chaque *boïd* de se diriger vers la position moyenne de ses congénères voisins

Ainsi, les proies se déplacent en groupe de façon auto-organisée en n'ayant comme prédateur réel que l'entité émotionnelle principale que l'on cherche à modéliser. En effet, les individus que l'on appelle prédateurs ne sont que les prédateurs du robot principal mais ne menacent en aucun cas les proies à la base du système. Ce modèle à trois couches d'individus nous permet d'avoir une représentation simple de relation « mangeur / mangé » centré sur le robot, qui entretient donc un rapport de force suffisamment équilibré pour forcer parfois le robot à affronter sa peur des prédateurs pour pouvoir se nourrir.

En ce qui concerne les prédateurs, ils sont équipés d'un champ de vision et d'audition comparable à celui des proies. Leur caractéristique principale est d'être attirés par le robot dès que celui-ci pénètre dans leur champ de détection visuelle ou auditive. Sinon, leur comportement par défaut consiste à errer linéairement dans l'environnement sans but précis.

Toutes les créatures du système, quelles qu'elles soient, sont bornées par les limites de l'environnement au-delà desquelles elles ne peuvent se déplacer.

9.2 ROBOT

Le robot principal dont on cherche à modéliser le comportement est équipé de capteurs et d'effecteurs lui permettant de capturer les informations de l'environnement et de rétroagir sur lui. Les capteurs sont de deux types, auditifs et visuels, et sont localisés tout autour du robot selon un angle paramétrable. La figure 9.2 présente les capteurs visuels (en bleu) localisés à l'avant du robot, ainsi que les capteurs auditifs (en marron) situés à 360 degrés autour du robot. L'intérêt d'avoir plusieurs capteurs ainsi répartis permet de discrétiser l'espace de détection et ainsi, de savoir avec une plus grande précision d'où provient le stimulus d'entrée. Par ailleurs, on observe que les capteurs se colorent en jaune si une proie a été détectée, et en noir si la détection concerne un prédateur.

Ainsi, sur la figure 9.2, on constate que le robot voit trois proies sur sa gauche et un prédateur sur sa droite. En outre, il entend une quatrième proie derrière lui à sa gauche. Cette disposition se vérifie sur la carte de l'environnement dessinée en parallèle par le programme (voir figure 9.3). Par ailleurs, la vue à la troisième personne (vue principale) permet de se placer juste derrière le robot (voir figure 9.4). Ainsi, il nous est possible d'avoir une vue globale de la carte, de la même façon qu'une vue subjective du point de vue du robot. Bien que la détection des individus n'introduise aucune notion de distance, celle-ci est tout de même renforcée en cas de double détection visuelle et auditive, qui indiquerait alors que la proie ou le prédateur est très proche du robot.

En ce qui concerne les effecteurs, nous pourrions les définir comme des dispositifs permettant l'action du robot sur son environnement. Ici, il s'agit principalement du déplacement, pour lequel deux types d'action sont envisageables : la modification de la vitesse de déplacement, ainsi que la rotation latérale du

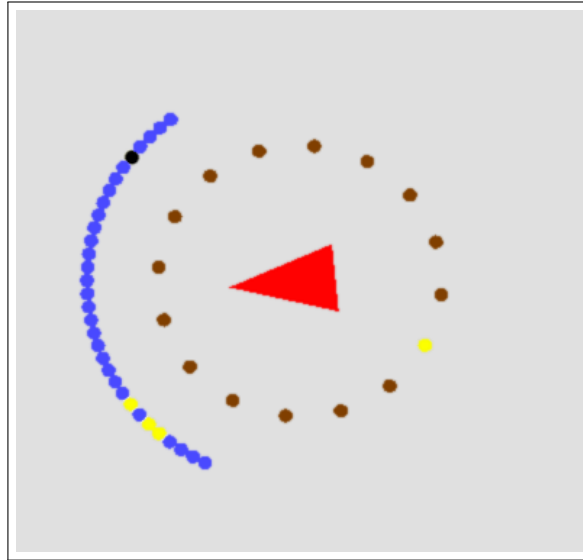


FIG. 9.2 – CAPTEURS DU ROBOT

robot afin de lui permettre de changer sa direction. Bien entendu, et ce dans le but de se rapprocher de conditions réelles de déplacement, les valeurs possibles d'action ont été limitées pour simuler les contraintes physiques liées au matériel. Par conséquent, le robot possède une vitesse de déplacement vers l'avant ou vers l'arrière ainsi qu'une vitesse angulaire maximale. Une autre action possible de la part du robot consiste à se nourrir afin d'augmenter son capital énergétique nécessaire à sa survie.

Toutes ces capacités de détection et d'action, associées à un contrôle des contraintes strict à base de seuils permettent au robot d'avoir une interface simple et efficace aux aléas environnementaux, qu'ils soient de nature critique (échapper à un prédateur) ou informationnelle (décider de poursuivre un groupe de proies).

De manière générale, qu'il s'agisse des capteurs ou des effecteurs, le système de perception est parfaitement générique, et permet d'introduire de nouveaux dispositifs de détection ou d'action sans modifier le programme. Ainsi, il nous aurait été parfaitement possible d'ajouter des capteurs tactiles ou de température, de la même façon que des effecteurs permettant par exemple le déplacement en hauteur ou pourquoi pas un armement spécifique pour se défendre contre ses adversaires. Tout est possible !

9.3 IMPLANTATION DU ROBOT

Le programme à proprement parler a été conçu en langage Java, en utilisant une bibliothèque 3D nommée JOGL pour Java OpenGL ([111]). L'écran principal

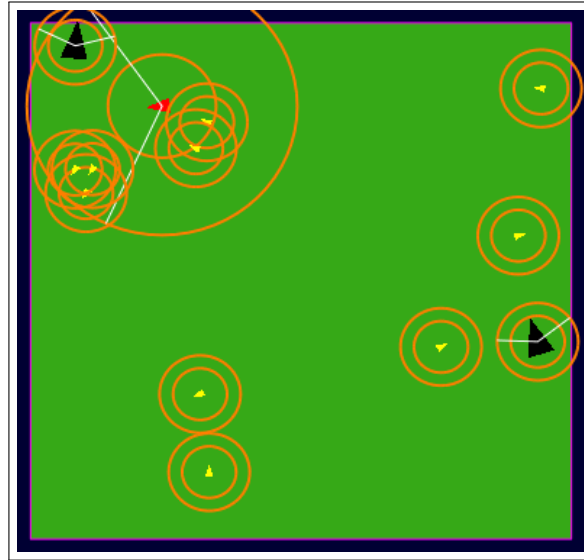


FIG. 9.3 – CARTE DE L'ENVIRONNEMENT

affiche la vue à la troisième personne (figure 9.4), accompagnée des deux vues précédemment mentionnées, à savoir les capteurs du robot (figure 9.2) et la carte de l'environnement (figure 9.3).

Le programme s'organise autour d'une classe *World* qui possède tous les éléments du monde, c'est-à-dire les *Creature(s)* (proies, prédateurs et robot). Chaque créature, héritant à juste titre de la classe *Creature*, possède entre autres une vitesse de déplacement, ainsi que des valeurs de rayon et de distance pour la détection visuelle et auditive. Quant au robot, identifié par la classe *Individual*, il est équipé de nombreux capteurs de détection (*VisualSensor* et *AudioSensor*, héritant de la classe *Sensor*), mais aussi de trois effecteurs (*MoveEffector* pour le déplacement, *TurnEffector* pour la rotation et *EatEffector* pour le nourrissage, tous trois héritant de la classe *Effector*).

Toutes ces informations sont résumées dans un diagramme de classe en deux parties, l'une relative aux informations autour de la classe *World*, figure 9.5, et l'autre spécifiquement sur les capteurs, figure 9.6.

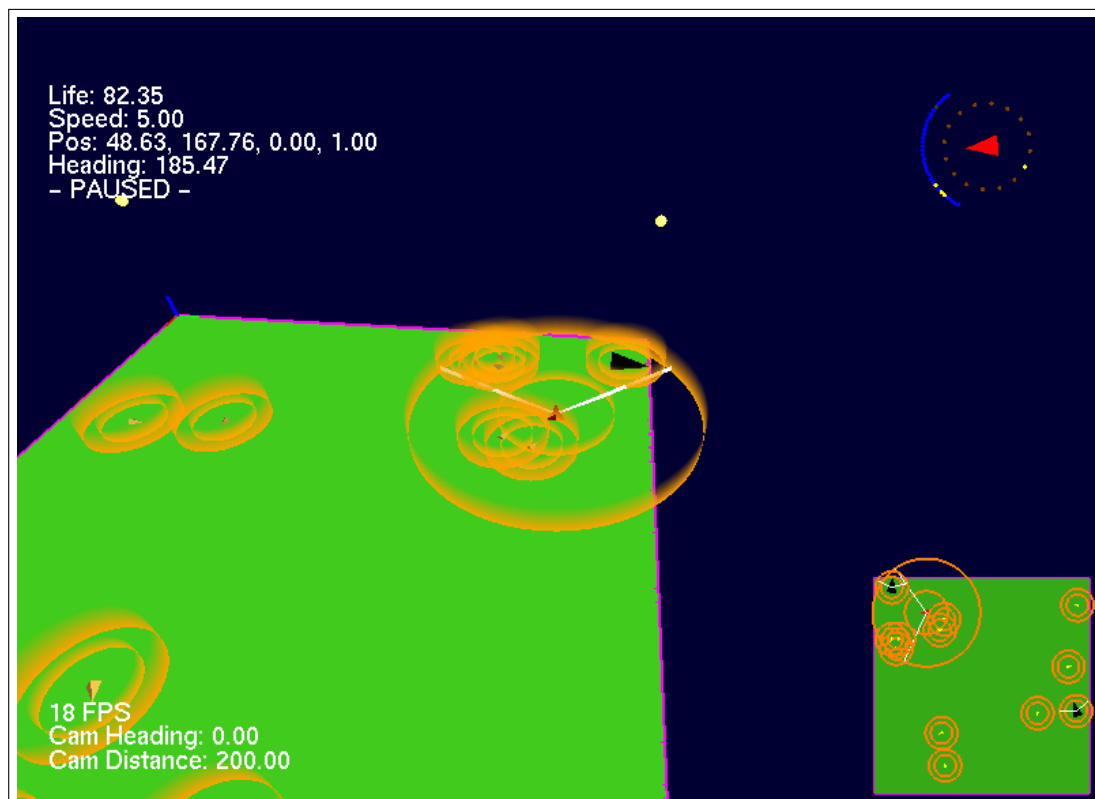


FIG. 9.4 – VUE À LA TROISIÈME PERSONNE

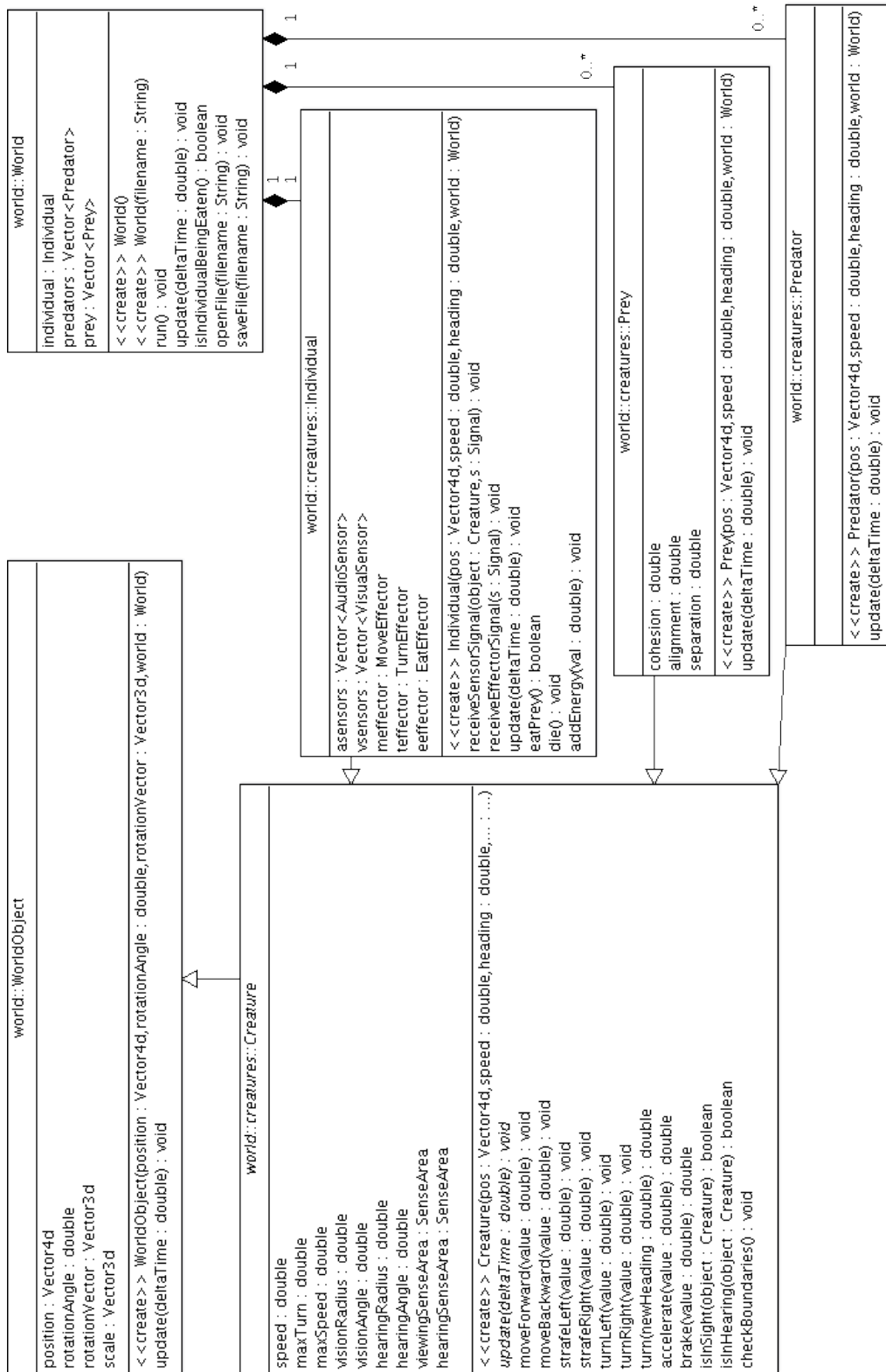


FIG. 9.5 – DIAGRAMME DE CLASSE DE L'ENVIRONNEMENT

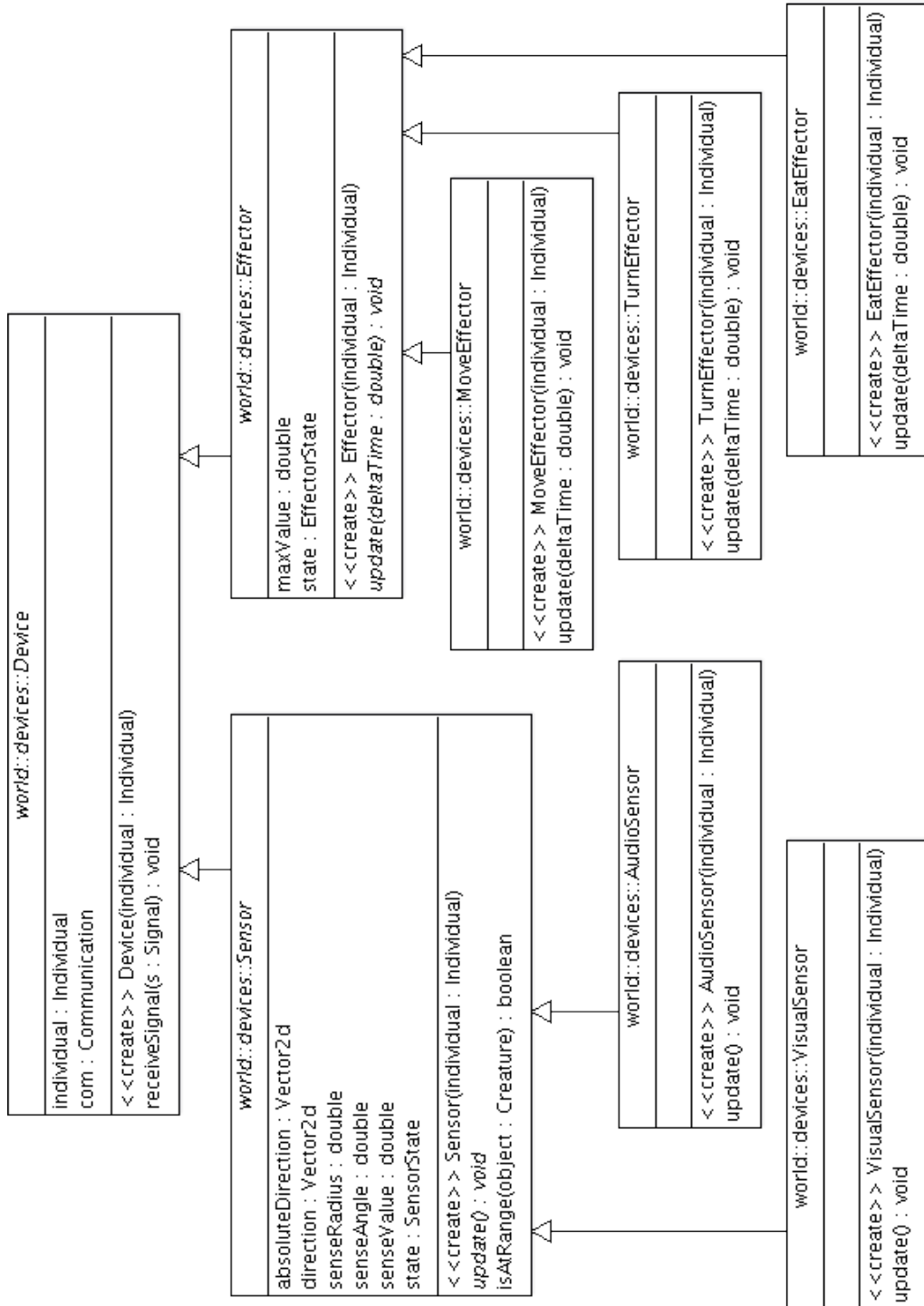


FIG. 9.6 – DIAGRAMME DE CLASSE DES CAPTEURS DU ROBOT

Dans le premier diagramme, il faut noter que tous les objets appartenant au monde héritent de la classe-mère *WorldObject*, littéralement « objet du monde », et ceci afin de garantir qu'ils sont tous localisables en termes de position et de rotation au sein de l'environnement. Bien que ce diagramme ne mentionne que les créatures, il existe néanmoins d'autres objets, peut-être plus secondaires, qui viennent peupler l'environnement, tels que les lumières, ou encore la caméra. J'ai volontairement choisi de ne pas les faire apparaître à l'intérieur du schéma.

En ce qui concerne les capteurs et effecteurs, ils sont tous deux définis par la super-classe *Device* (qui peut être traduit par « dispositif »). En effet, qu'il s'agisse d'un capteur ou d'un effecteur, chaque dispositif possède un module de communication, lequel lui permet de transmettre et de recevoir des signaux de la part des autres entités de l'environnement ou du robot lui-même. Ainsi, si une proie passe à proximité de l'individu, un signal sera envoyé à un capteur visuel, lequel préviendra l'individu en lui transmettant l'information sur l'intitulé (une proie) et sur le type (visuel) de la détection. De la même façon, si le robot décide de se déplacer, il enverra un signal à son effecteur de mouvement (de la classe *MoveEffector*), celui-ci se chargeant de faire se mouvoir le robot selon des règles strictes respectant certaines contraintes.

SOMMAIRE

10.1	MODÈLE ÉMOTIONNEL	145
10.2	DESCRIPTION DE L'IMPLANTATION	147
10.3	EXEMPLE D'EXÉCUTION	151
10.3.1	Étape t = 0	151
10.3.2	Étape t = 15	151
10.3.3	De l'étape t = 21 à l'étape t = 27	154
10.3.4	De l'étape t = 60 à l'étape t = 70	154
10.3.5	De l'étape t = 170 à l'étape t = 190	154
10.3.6	Étape t = 260	155

There is no refuge from memory and remorse in this world. The spirits of our foolish deeds haunt us, with or without repentance.^a

— Gilbert PARKER

a. « On ne peut échapper ni aux souvenirs ni aux remords. Leur ombre continue de nous hanter, ce sont nos repentirs. »

10.1 MODÈLE ÉMOTIONNEL

Le modèle émotionnel a été développé dans le cadre des travaux de thèse de William BLEWITT. Basé sur le modèle de MILLENSON, l'émotion est représentée spatialement suivant trois axes thématiques : la *colère*, la *peur* et le *plaisir* ([91]).

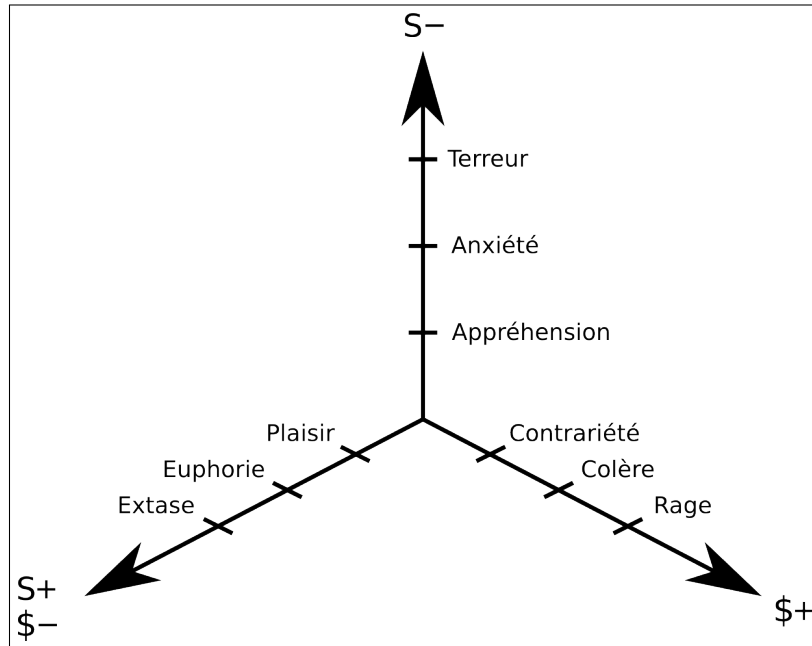


FIG. 10.1 – MODÈLE TRIDIMENSIONNEL DE L'INTENSITÉ ÉMOTIONNELLE DE MILLENSON
(VERSION ORIGINALE EN ANNEXE B)

Chaque axe donne une graduation de l'intensité de l'émotion de base qu'il représente. Un état émotionnel est alors une position dans cet espace tridimensionnel (voir figure 10.1).

Le premier axe de représentation désigne la **colère**, que MILLENSON décrit comme l'expérience issue de la *suppression de stimuli positifs* (notée \$+), et divisée en trois émotions d'intensités graduelles que sont la *contrariété*, la *colère* et la *rage*. De la même façon, le second axe symbolisant la **peur** et représentant l'*ajout de stimuli négatifs* (noté S-), se découpe en trois émotions : l'*appréhension*, l'*anxiété* et la *terreur*. Enfin, le troisième axe désigne le **plaisir** et exprime à la fois l'*ajout de stimuli positifs* et la *suppression de stimuli négatifs* (notés S+ et \$-), à l'aide là encore de trois émotions d'importance croissante : le *plaisir*, l'*euphorie* et l'*extase*.

Le module MATLAB développé par BLEWITT ([87]) utilise la *logique floue* (*fuzzy logic* en anglais), introduite par Lotfi ZADEH en 1965 ([135]). À l'inverse de la logique booléenne qui ne propose que deux valeurs possibles à une affirmation, à savoir *vrai* ou *faux*, la logique floue permet de définir un ensemble de valeurs de vérité et offre ainsi un degré de précision plus important tout en restant relativement vague. Cette méthode est donc particulièrement adaptée à la modélisation de l'émotion.

Par une série de *fuzzifications* et *défuzzifications*, c'est-à-dire de décompositions de valeurs numériques exactes en valeurs floues, et *vice versa*, un état émotionnel de départ en trois dimensions est converti en un vecteur à neuf dimensions

représentant les neuf émotions décrites par MILLENSON :

$$\left\{ \begin{array}{l} x_1 : \text{contrariété} \\ x_2 : \text{colère} \\ x_3 : \text{rage} \\ y_1 : \text{appréhension} \\ y_2 : \text{anxiété} \\ y_3 : \text{terreur} \\ z_1 : \text{plaisir} \\ z_2 : \text{euphorie} \\ z_3 : \text{extase} \end{array} \right.$$

L'idée principale derrière cette décomposition réside dans le déclenchement progressif des trois émotions d'une composante à l'aide d'un système de seuils. Ainsi, une petite contrariété pourra se transformer en colère, puis en rage selon l'évolution de la situation, l'ensemble des seuils jouant le rôle de profil émotionnel. Une description complète du modèle se trouve en annexe C.

10.2 DESCRIPTION DE L'IMPLANTATION

Le travail effectué ici consiste à créer en langage Java l'environnement proie-prédateur décrit précédemment, ainsi qu'à utiliser le modèle de MILLENSON développé en MATLAB dans une boucle itérative qui permet la communication des deux processus, la sortie de l'un étant redirigée vers l'entrée de l'autre (voir figure 10.2). Cette boucle met en œuvre un certain nombre de vecteurs représentant les différentes étapes de construction des états du système.

Il faut noter la différence fondamentale entre E_M et E_J , qui permet de bien comprendre le fonctionnement du programme. E_M est l'état émotionnel du robot, tandis que E_J représente l'impact émotionnel d'un stimulus donné J . Ainsi, E_J permet de construire E_M mais ces deux vecteurs ne sont pas équivalents :

$$\mathbf{E}_M(t+1) = \frac{(u\mathbf{E}_M(t) + v\mathbf{E}_J)}{u+v}$$

où u et v sont des paramètres de pondération.

Afin de déterminer le comportement de l'agent, le programme Java prend en entrée le vecteur \mathbf{E}_M produit par le module MATLAB, c'est-à-dire les neuf variables de sortie x_{1-3} , y_{1-3} et z_{1-3} .

Dans le but de relier ces neuf émotions distinctes au comportement de l'agent, chaque triplet d'émotions est associé à une facette comportementale. Par exemple, les composantes x_{1-3} , c'est-à-dire le triplet $\{\text{contrariété}, \text{colère}, \text{rage}\}$,

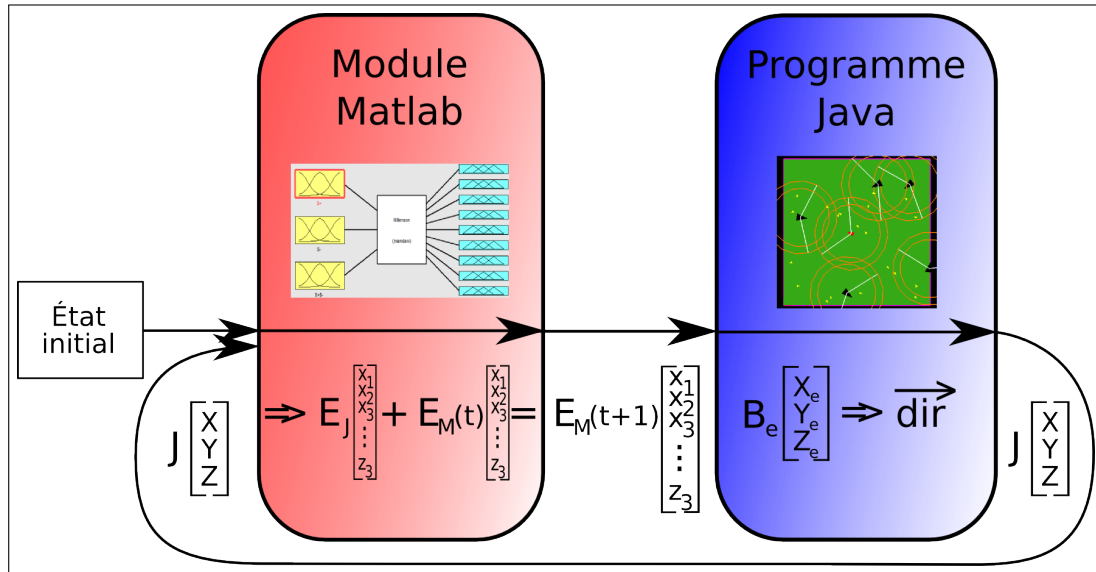


FIG. 10.2 – BOUCLE PRINCIPALE D'EXÉCUTION

représentent la suppression de stimuli positifs, d'après le modèle de MILLENSON. Par conséquent, le modèle associe ces trois émotions au désir accru de l'agent de capturer des proies pour se nourrir. Comme ces émotions distinctes représentent différentes intensités liées à la suppression de stimuli positifs, le calcul utilisé pondère chaque émotion différemment. Toutefois, afin de conserver la logique de MILLENSON selon laquelle les composantes émotionnelles sont organisées dans l'ordre croissant de leur magnitude, les valeurs de pondération se doivent de respecter la même règle. Ainsi, les trois valeurs w_{x_1} , w_{x_2} et w_{x_3} , coefficients de pondération associés respectivement aux valeurs x_1 , x_2 et x_3 s'organisent telles que $w_{x_1} \leq w_{x_2} \leq w_{x_3}$.

Finalement, ces trois composantes émotionnelles sont regroupées en une seule, représentant une valeur de pondération générale X_e appliquée à l'attraction des proies pour l'agent. On a alors :

$$X_e = \frac{w_{x_1}x_1 + w_{x_2}x_2 + w_{x_3}x_3}{w_{x_1} + w_{x_2} + w_{x_3}}$$

De la même manière, les six autres émotions sont représentées afin de générer les variables similaires Y_e et Z_e :

$$Y_e = \frac{\sum_{n=1}^3 w_{y_n} y_n}{\sum_{n=1}^3 w_{y_n}} \quad Z_e = \frac{\sum_{n=1}^3 w_{z_n} z_n}{\sum_{n=1}^3 w_{z_n}}$$

Les trois variables X_e , Y_e et Z_e permettent ainsi de former le vecteur de pondération \mathbf{B}_e , lequel servira à nuancer le comportement de l'agent :

$$\mathbf{B}_e = \begin{bmatrix} X_e \\ Y_e \\ Z_e \end{bmatrix}$$

Alors que X_e est appliquée à la suppression de stimuli positifs, à savoir l'attractivité des proies, la composante Y_e est quant à elle liée à l'application de stimuli négatifs, c'est-à-dire à la peur des prédateurs qui pousse l'agent à fuir le danger. Enfin, la variable Z_e représentative du triplet $\{\text{plaisir, euphorie, extase}\}$ agit comme une force retardatrice sur le vecteur directeur principal. Ceci s'explique par le fait que plus la situation courante est agréable, moins l'agent sera susceptible de vouloir la changer.

Le vecteur direction \vec{dir} est par ailleurs calculé à partir des positions des proies (\vec{Pos}_f) et des prédateurs (\vec{Pos}_p), par rapport à la position de notre agent (\vec{Pos}_{agent}), F représentant l'ensemble des proies (pour l'anglais *food*) et P l'ensemble des prédateurs. L'ensemble des proies forme le vecteur attraction \vec{att} et les prédateurs créent un effet répulsif \vec{rep} sur la direction finale :

$$\vec{dir} = X_e \vec{att} + Y_e \vec{rep} + Z_e (\vec{att} + \vec{rep})$$

avec :

$$\vec{att} = \sum_{f \in F} (\vec{Pos}_f - \vec{Pos}_{agent})$$

$$\vec{rep} = \sum_{p \in P} (\vec{Pos}_{agent} - \vec{Pos}_p)$$

Ce vecteur direction nous permet ainsi de calculer le nouveau vecteur déplacement nécessaire au robot pour mettre à jour sa position dans l'environnement.

L'étape suivante consiste à calculer le vecteur stimulus \mathbf{J} qui doit être envoyé au module MATLAB. Comme nous l'avons annoncé précédemment, l'émotion est directement issue de la situation spatiale de l'agent, c'est-à-dire des positions des proies et prédateurs locaux. Si nous considérons une situation dans laquelle l'environnement local est composé de deux proies f_1 et f_2 , et de deux prédateurs p_1 et p_2 , il nous est possible de calculer chaque composante du vecteur stimulus, X représentant la suppression des stimuli positifs ($\sum\{\$+\}$), Y l'addition de stimuli négatifs ($\sum\{S-\}$), et Z aussi bien l'addition de stimuli positifs que la suppression de stimuli négatifs ($\sum\{S+, \$-\}$), comme le montre la figure 10.3.

Dans cet exemple, le robot possède deux individus dans son champ de vision, un de chaque type, ainsi que deux autres entités dans son rayon auditif. Pour chacun de ces individus, on calcule la valeur $\delta dist$ qui correspond à la différence

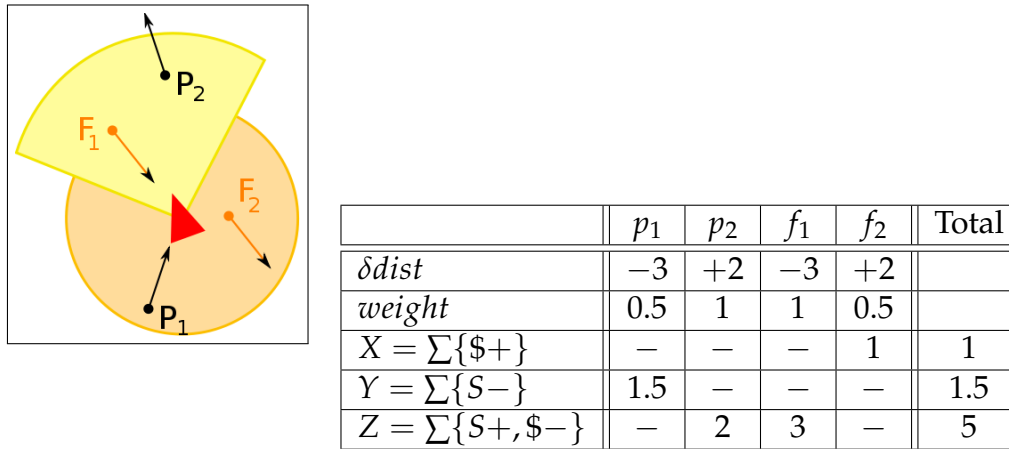


FIG. 10.3 – EXEMPLE DE SITUATION LOCALE D'UN AGENT ET SA REPRÉSENTATION

entre la distance courante et la distance précédente. Par conséquent, si cette valeur est négative, cela signifie que l'individu correspondant se rapproche du robot, et *vice versa*.

Ensuite, il nous faut considérer l'importance relative de chaque individu sur le comportement du robot. Pour ce faire, une valeur de pondération est fixée pour chaque individu détecté par le robot, en fonction de la nature même de la détection. Dans cet exemple, la valeur 1 est donnée pour les individus situés dans l'aire visuelle du robot, alors que 0.5 est attribuée aux entités situées dans la zone auditive uniquement. Cette distinction spatiale liée à la détection permet de considérer un prédateur, par exemple, comme deux fois plus significatif si celui-ci est visible que s'il est simplement audible. Bien entendu, les valeurs 1 et 0.5 sont arbitraires et dépendent directement de l'implantation et de l'importance relative que l'on souhaite porter à un stimulus sensitif particulier. Aussi, les valeurs de pondération présentées au sein du tableau de la figure 10.3 sont simplement indiquées dans un but illustratif.

Finalement, nous calculons les trois composantes X , Y et Z constitué de Z_F (pour les proies) et Z_P (pour les prédateurs) qui vont nous permettre de former le vecteur stimulus J :

$$\mathbf{J} = \begin{bmatrix} X \\ Y \\ Z = Z_F + Z_P \end{bmatrix}$$

avec :

$$X = \sum_{f \in F} |\delta dist_f \times weight_f| \quad \text{si } \delta dist_f \geq 0$$

$$Y = \sum_{p \in P} |\delta dist_p \times weight_p| \quad \text{si } \delta dist_p < 0$$

$$Z_F = \sum_{f \in F} |\delta dist_f \times weight_f| \quad \text{si } \delta dist_f < 0$$

$$Z_P = \sum_{p \in P} |\delta dist_p \times weight_p| \quad \text{si } \delta dist_p \geq 0$$

Ce vecteur stimulus sera ensuite réintégré en entrée du module MATLAB afin d'entamer une nouvelle itération.

10.3 EXEMPLE D'EXÉCUTION

Cette partie décrit un exemple d'exécution commenté, afin de comprendre comment se comporte le robot. Les courbes de la figure 10.4 décrivent l'évolution du vecteur stimulus \mathbf{J} et du vecteur de pondération \mathbf{B}_e , ainsi que de l'angle de rotation $\delta\Theta$, en fonction du temps.

10.3.1 Étape $t = 0$

L'état initial (figure 10.5) présente l'environnement de départ, constitué de 20 proies (triangles jaunes) et de 5 prédateurs (triangles noirs). Le robot (triangle rouge) possède 5 proies dans son champ de perception mais celles-ci se trouvent toutes derrière lui. Elles sont donc dans son champ auditif mais pas visuel, ce qui est suffisant pour être détectées. L'émotionnalité est à zéro pour toutes les composantes (X, Y, Z), c'est-à-dire que le vecteur \mathbf{J} transmis au module MATLAB pour la première itération vaut $(0, 0, 0)$, signe d'un état émotionnel complètement neutre.

10.3.2 Étape $t = 15$

On observe peu après (jusqu'à t_{15} environ) que le robot tente de se retourner le plus rapidement possible, afin de capturer les proies qui lui échappent (figure 10.6). Ce comportement engendre deux conséquences. La première est le bouleversement de la valeur $\delta\Theta$ qui se retrouve à son maximum dans le négatif. En effet, la vitesse angulaire maximale est de 10° par étape, c'est-à-dire 10° pour une rotation vers la gauche et -10° pour une rotation vers la droite. Le deuxième élément est la légère augmentation de la composante X (la colère) due à l'éloignement des proies. Cette même composante décroît aussitôt après que le robot s'est retourné. En parallèle, la composante Z (le plaisir) commence sa montée lorsque le robot découvre les proies dans son champ visuel et se rapproche d'elles progressivement.

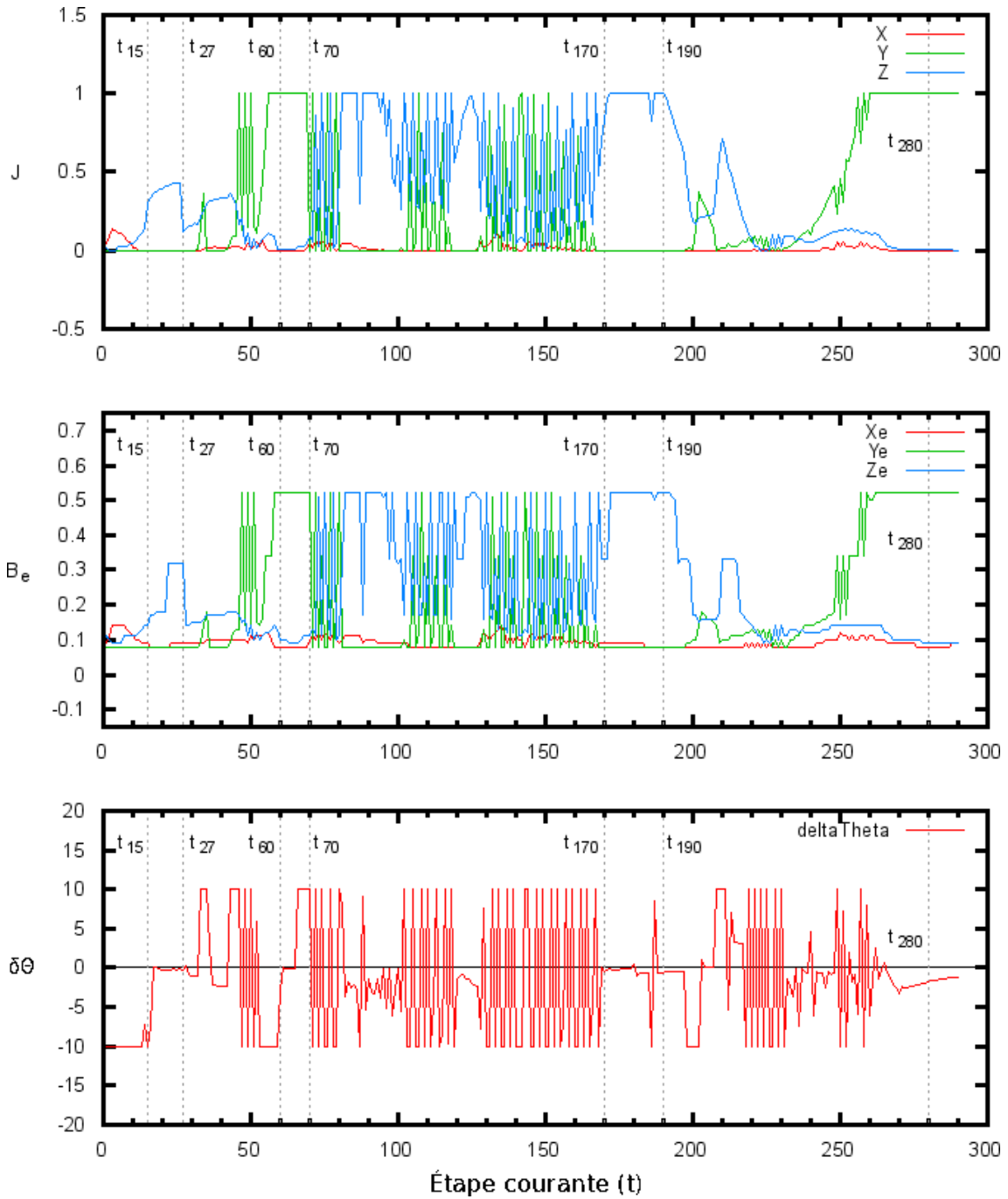


FIG. 10.4 – ÉVOLUTION DES VECTEURS \mathbf{J} ET \mathbf{B}_e , AINSI QUE DE L'ANGLE DE ROTATION $\delta\Theta$, EN FONCTION DU TEMPS

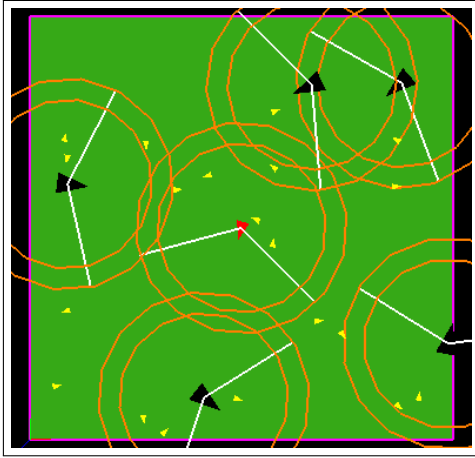


FIG. 10.5 - ÉTAPE INITIALE ($t = 0$)

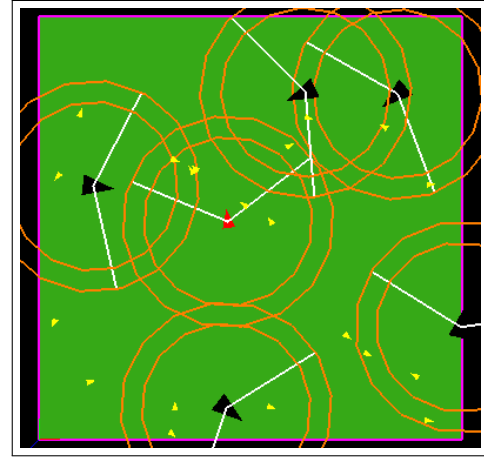


FIG. 10.6 - ÉTAPE $t = 15$

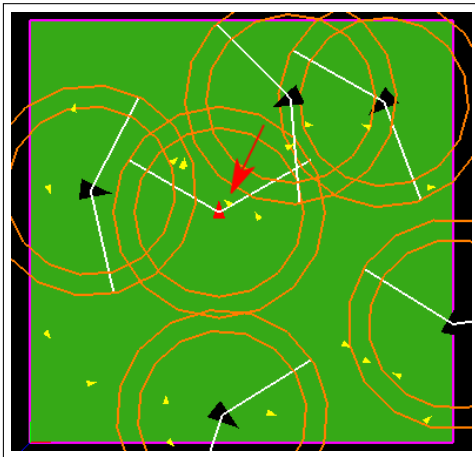


FIG. 10.7 - ÉTAPE $t = 21$

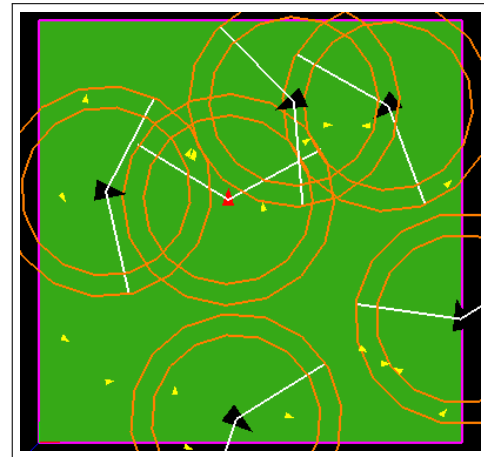


FIG. 10.8 - ÉTAPE $t = 27$

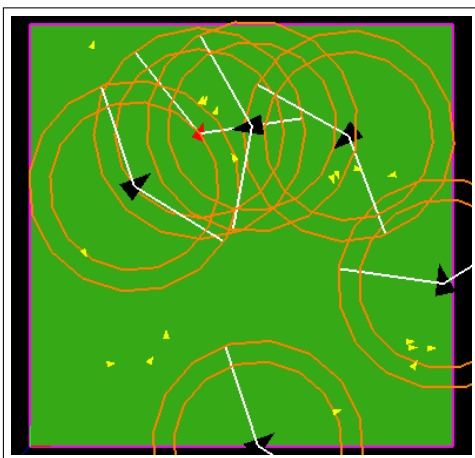


FIG. 10.9 - ÉTAPE $t = 65$

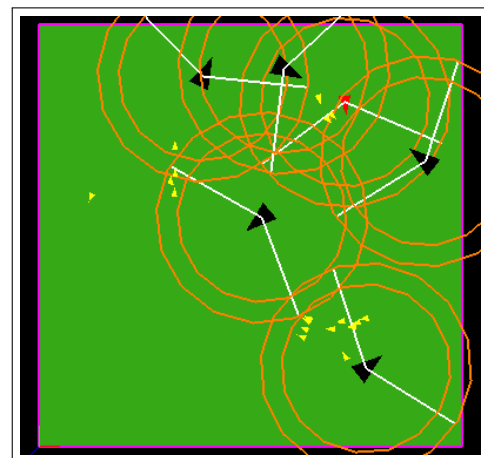


FIG. 10.10 - ÉTAPE $t = 184$

10.3.3 De l'étape $t = 21$ à l'étape $t = 27$

Entre la 21^{ème} et la 27^{ème} étape, le robot se nourrit de sa première proie (indiquée par la flèche rouge sur la figure 10.7), ce qui a pour conséquence une baisse instantanée du plaisir (Z) liée à la disparition de ladite proie. En effet, le plaisir n'est influencé que par la présence de proies mais en aucun cas par la satisfaction de les consommer. Ce comportement qui pourrait sembler être à l'opposé de ce que l'on attend d'un robot anthropomorphique est en réalité parfaitement cohérent dans la mesure où le but est uniquement la survie par la recherche de nourriture. Le résultat aurait été tout autre si l'objectif du programme avait été centré sur la simulation de l'état émotionnel d'un être vivant en particulier qui, dans le cas de la consommation de proies, possède un réel intérêt biologique à ressentir du plaisir à se nourrir.

10.3.4 De l'étape $t = 60$ à l'étape $t = 70$

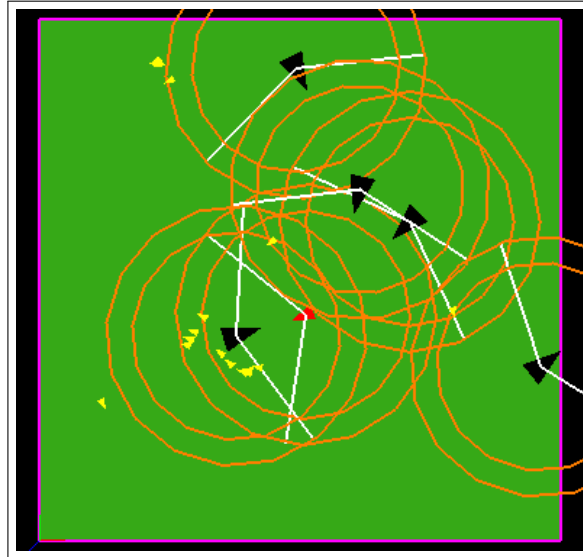
Entre les étapes 60 et 70, on assiste à une période pendant laquelle la peur est le sentiment dominant (voir figure 10.9). En effet, le robot se trouve dans une situation de poursuite d'un certain nombre de proies regroupées en un essaim en fuite. Par ailleurs, un prédateur se rapproche dangereusement par la droite et tente logiquement de dévorer notre robot. L'état émotionnel consécutif à cette configuration met naturellement la peur, représentée par la composante Y , en position haute. Le robot préférera finalement contourner la zone de danger en empruntant un chemin plus à gauche, tout en essayant de revenir sur l'essaim de proie ensuite.

Par ailleurs, on constate que le graphe des valeurs angulaires $\delta\theta$ se présente souvent sous la forme d'oscillations rapides, de -10 à 10 , dénotant un déplacement en zigzag très rapide. Ce genre de période de balancement est en réalité dû à un passage entre les zones d'action de deux prédateurs, lesquels repoussent à chaque étape le déplacement du robot, ce dernier essayant de se frayer un passage entre les deux entités menaçantes. Il arrive aussi que l'oscillation soit due à un choix difficile entre deux proies situées à égale distance du robot.

10.3.5 De l'étape $t = 170$ à l'étape $t = 190$

À l'inverse, la période allant de $t = 170$ à $t = 190$ se caractérise par une domination du sentiment de plaisir, lequel se traduit par une valeur maximale de la composante Z . Cette valeur est liée, là-encore, à la proximité d'un groupe de proies qui ne cesse de se rapprocher (voir figure 10.10). Comme le souligne l'angle de rotation $\delta\theta$, le robot se déplace de façon pratiquement droite, presque en parallèle du déplacement des proies qu'il poursuit. Ce comportement est lié à la présence du prédateur juste derrière lui sur sa droite, qui risquerait de se rapprocher si le robot se décidait à aller plus franchement sur ses proies.

10.3.6 Étape $t = 260$

FIG. 10.11 – ÉTAPE $t = 280$

Aux alentours de l'étape $t = 260$ et jusqu'à la fin de la simulation, la terreur ressentie par le robot est le sentiment dominant. En effet, on assiste à une scène finale durant laquelle se met en place un face-à-face dangereux avec un prédateur (voir figure 10.11). Bien entendu, cette mise en danger n'est pas anodine : le robot pourchassait un grand nombre de proies quand soudain est apparu un ennemi juste en face de lui. Dans cette configuration, proies et prédateur se sont croisés, laissant la place libre à ce dernier pour dévorer notre pauvre robot. Moralité : la véritable terreur ressentie par le robot n'a pas suffi à contrecarrer la force d'attraction exercée par les proies, en grand nombre, situées derrière le prédateur.

IV

LE MODÈLE GAEA : UNE VISION *émergentiste* DE L'ÉMOTION

Historiquement, dans l'étude de l'émotion, il est d'usage de considérer les manifestations corporelles, telles que les expressions faciales, la transpiration ou encore les battements cardiaques, en tant que point de départ de la mesure expérimentale. Cependant, si ces manifestations physiologiques sont potentiellement conséquentes à l'état émotionnel (« *J'ai peur, et par conséquent, mon rythme cardiaque s'accélère* »), il est fallacieux de penser que l'inverse est vrai (« *J'ai le cœur qui bat vite, donc cela signifie nécessairement que j'ai peur* »), car plusieurs états émotionnels différents peuvent mener aux mêmes manifestations physiologiques. Bien entendu, certains signes ne trompent pas, et de nombreuses études analysant les expressions faciales, par exemple celles de Paul EKMAN ([37], [38]) tentent de déterminer un état émotionnel par la description et la classification de mouvements musculaires faciaux, avec le postulat que ces mouvements sont universellement reconnus et assimilés par tous. Pour autant, on ne peut avancer avec certitude l'état émotionnel d'un individu en relevant uniquement ces informations, étant donnés les différents profils émotionnels présents au sein d'une population. Et quand bien même cela eût été possible, il est impossible d'en déterminer l'éventuelle décision qui s'ensuivra.

Ce problème, central en psychologie, s'explique par l'internalité de l'émotion. En effet, comme nous l'avons vu précédemment, l'émotion est un état cérébral interne global ayant également des conséquences sur les aspects cognitifs et de façon plus générale, sur le comportement. Cette vision appelée *top-down* et qui consiste à appréhender le phénomène en partant du « haut » (manifestations physiologiques visibles de l'extérieur) vers le « bas » (état émotionnel interne difficile à mesurer), n'est donc pas suffisante car limitée uniquement à ce que l'on peut observer.

Par conséquent, l'idée d'une perspective plus « émergentiste », c'est-à-dire dont le but est avant tout de favoriser l'émergence, et qui ne part plus de ce qui se voit mais de ce qui est à l'origine, s'oppose aux modèles traditionnels de type *top-down* : ce sont les approches *bottom-up*, pour lesquelles l'émotion provient de l'activité du système qui produit des manifestations comportementales et apprend à être plus efficace lors de la résolution d'un problème posé. Afin de concevoir un tel système, le modèle GAEA utilise un algorithme évolutionniste qui permet à l'émotion, tout comme la cognition, de faire l'objet d'une construction cérébrale génétiquement programmée. Cette émotion qui s'auto-construit est alors le fruit du processus d'adaptation du système, qui renforce l'efficacité de la résolution.

L'objectif de cette partie est, dans un premier temps, de décrire cette vision du bas vers le haut, que l'on va qualifier d'« émergentiste », et qui dénote la complexité des processus émotionnels. Pour ce faire, un algorithme de programmation génétique linéaire sera utilisé. Ensuite, on se propose de concevoir un modèle qui adhère à ce principe d'émergence, tout en proposant une situation expérimentale dans laquelle la décision doit se faire rapidement et de manière

adéquate. Ce modèle utilise une version améliorée du système proie-prédateur de la partie précédente (partie III). Le principe d'urgence introduit par le système ainsi que la pression subie par l'individu représenté seront la clef de voûte de cette étude. Les résultats obtenus viendront confirmer l'efficacité du modèle.

UNE APPROCHE « ÉMERGENTISTE »

11

SOMMAIRE

11.1 L'INTELLIGENCE COMPUTATIONNELLE	162
11.2 LES ALGORITHMES ÉVOLUTIONNISTES	163
11.3 LA PROGRAMMATION GÉNÉTIQUE	165
11.4 LA PROGRAMMATION GÉNÉTIQUE LINÉAIRE	167

*Life obeys no logic, why do we want to deduce its
meaning with logic?^a*

— Gao XINGJIAN

a. « La vie n'obéit à aucune logique, pourquoi voulons-nous déduire sa signification par la logique ? » (NDR)

La science moderne n'a pas coutume de dire que l'émotion n'est qu'un processus émergent et n'est en aucun cas précisément représenté dans le système nerveux. En effet, afin de pouvoir étudier un phénomène précis, il est d'usage non seulement d'en faire mention *a priori*, mais surtout de le mettre en évidence en l'identifiant, en le symbolisant au sein du système. À l'inverse, l'idée ici est de pouvoir laisser le modèle s'exprimer et faire de lui-même émerger ce qui représente pour lui l'émotion, sans le guider vers une ligne de conduite particulière, ni utiliser d'heuristique trop directive qui le pousserait à une convergence forcée, ou pire, qui le ferait sombrer irrémédiablement vers un minimum local. Dans l'approche proposée ici, on utilise un outil bien connu de l'univers de l'optimisation combinatoire : les *algorithmes évolutionnistes*, et plus particulièrement la *programmation génétique linéaire*. Celle-ci va nous permettre de générer une population d'*individus-programmes* de plus en plus adaptés au problème de départ. À chaque génération de programmes, le taux d'adaptation moyen progresse par pression-sélection des individus en fonction de leur efficacité.

Ainsi, de la même manière que le câblage cérébral d'un individu conditionne sa capacité d'adaptation à une situation donnée, un programme génétiquement adapté, par sa structure, permet de résoudre des problèmes, même complexes. L'émotion, quant à elle, est le fruit de cette auto-organisation, une sorte d'état interne en constante activité, et dont la dynamique permet la résolution ou dans une autre mesure l'aide à la décision.

GAEA est une application qui vise à modéliser une entité fondamentalement réactive, à savoir un robot virtuel dont le comportement repose sur sa capacité à répondre à des besoins environnementaux sans que l'informaticien n'ait explicitement programmé en son sein des comportements prévus. L'acronyme GAEA signifie par ailleurs *Genetically Adaptive Emotional Agent*, ce qui signifie *agent émotionnel génétiquement adapté*, et dénote en l'occurrence l'importance de l'aspect auto-adaptatif du modèle. En effet, il est question ici d'utiliser plus particulièrement un algorithme de programmation génétique linéaire afin de construire le comportement du robot.

Ce chapitre vise à comprendre le parcours épistémologique entourant le modèle GAEA, qui débute par l'*intelligence computationnelle*, catégorie générale réunissant l'ensemble des approches bio-inspirées, en passant par les *algorithmes évolutionnistes*, subdivision qui regroupe tous les modèles mettant en place des processus évolutifs, et allant jusqu'aux algorithmes de *programmation génétique linéaire*, encore très peu usités de nos jours, mais qui possèdent des propriétés remarquables. L'objectif final sera de faire émerger un comportement émotionnel à une entité programmatique baignée dans un environnement complexe en constante mutation, avec le postulat qu'à partir de stimuli simples, un robot doté d'un programme évolutif peut progressivement manifester des signes comportementaux analogues à des réactions émotionnelles.

11.1 L'INTELLIGENCE COMPUTATIONNELLE

L'intelligence computationnelle, issue de l'intelligence artificielle, est une famille de procédés, le plus souvent bio-inspirés, et combinant plusieurs domaines tels que l'adaptation, l'apprentissage, ou encore l'évolution, afin de créer des programmes dont le comportement peut être considéré comme intelligent.

L'IEEE (*Institute of Electrical and Electronics Engineers*) a officiellement découpé l'intelligence computationnelle en quatre thèmes distincts que sont :

- **Les réseaux de neurones** introduits par McCulloch et Pitts en 1943 ([90]), sont des modèles inspirés du fonctionnement des neurones biologiques qui tirent parti de leur connectivité et de leur capacité d'apprentissage pour fournir, d'une part, un outil de classification rapide et puissant, et d'autre part, un mécanisme perceptif particulièrement générique et adaptatif.

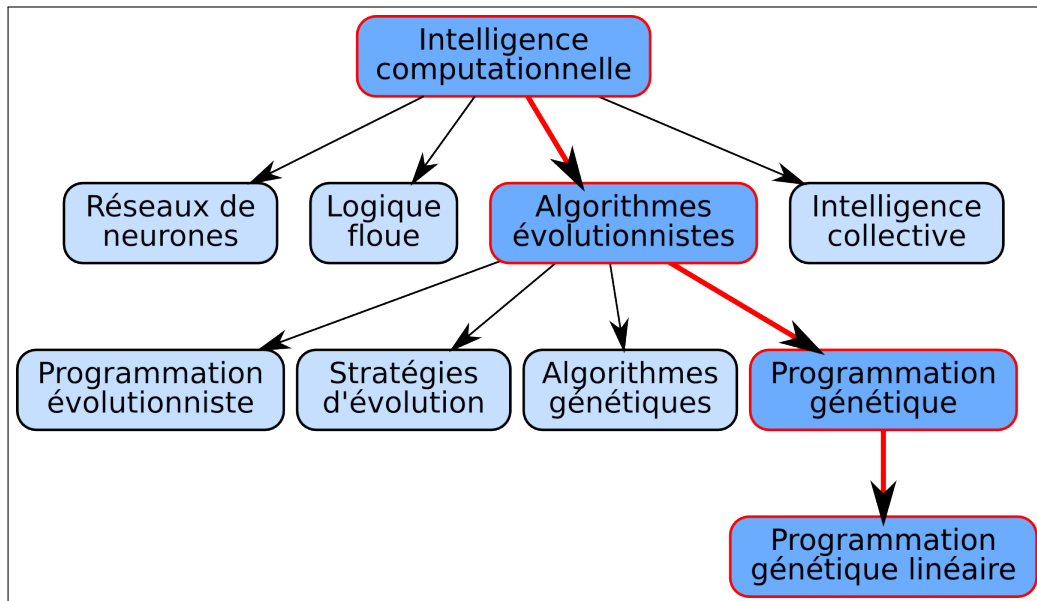


FIG. 11.1 – LES SOUS-CATÉGORIES DE L'INTELLIGENCE COMPUTATIONNELLE

- **Les systèmes à contrôle flou** utilisent la logique floue, formalisée par Lotfi ZADEH en 1965 ([135]). Il s'agit d'une théorie formelle et mathématique qui étend la notion de logique booléenne en ce sens où une condition peut être dans un état différent de *vrai* ou *faux*.
- **L'intelligence collective** est une méta-heuristique qui considère que la capacité cognitive d'une communauté résulte de l'ensemble des interactions de ses membres. L'optimisation par essaim de particules ou les algorithmes de colonies de fourmis en sont les exemples parmi les plus célèbres.
- **Les algorithmes évolutionnistes** sont la lignée dans laquelle s'inscrit le modèle GAEA (voir ci-après).

Les disciplines relatives à l'intelligence computationnelle peuvent être représentées au sein d'un schéma hiérarchique (voir figure 11.1).

11.2 LES ALGORITHMES ÉVOLUTIONNISTES

Les *algorithmes évolutionnistes* (ou *évolutionnaires*) sont une famille d'algorithmes utilisée pour résoudre des problèmes d'optimisation, et dont le fonctionnement s'inspire directement de la théorie de l'évolution. En ce sens, ils manipulent des populations de solutions potentielles qu'ils font évoluer, dans l'optique de trouver les meilleurs résultats.

D'après la théorie de l'évolution des êtres vivants, chère à DARWIN ([27], [28]), le cycle de la vie tend à produire des individus de plus en plus adaptés à leur

environnement par croisements successifs et par la conservation des caractéristiques génétiques les plus efficaces, compte tenu des exigences naturelles. Par conséquent, ils possèdent un certain nombre de propriétés essentielles, parmi lesquelles :

- les individus possèdent un génotype qui permet de définir leurs caractéristiques ;
- les individus présents dans une population donnée sont tous différents ; ils ont par conséquent une capacité d'adaptation plus ou moins grande par rapport à leur environnement ;
- les individus se doivent de transmettre tout ou partie de leur génotype à leur descendance, afin de garantir une certaine pérennité dans la conservation des caractères les plus efficaces ;
- les individus les plus adaptés doivent pouvoir se reproduire globalement plus souvent et plus efficacement, ceci afin de permettre à leurs caractéristiques de se répandre davantage dans la population.

Il faut bien distinguer le génotype des individus, transmis de générations en générations, et le phénotype, qui est l'efficacité réelle de l'individu, et dont la mesure est communément appelée *fitness*. En outre, les algorithmes évolutionnistes définissent un certain nombre d'opérateurs permettant, à chaque étape de l'algorithme, de manipuler les différents individus au sein de la population courante, en fonction de leur *fitness*. Il en existe de deux sortes : les *opérateurs de sélection*, associés aux étapes de sélection et de remplacement, ainsi que les *opérateurs de variation*, liés aux croisements et aux mutations.

On peut alors généraliser le fonctionnement d'un algorithme évolutionniste de la façon suivante :

1. **initialisation** d'une population d'individus ;
2. **sélection** d'une partie des individus de la population ;
3. création de nouveaux individus en appliquant, parfois selon une certaine probabilité, un certain nombre d'opérateurs parmi lesquels :
 - ★ **copie** : on duplique un individu, sans aucune modification ;
 - ★ **croisement** ou **recombinaison** : on échange une partie du génotype entre certains individus ;
 - ★ **mutation** : on remplace/modifie le génotype de la descendance ;
 - ★ **insertion** : on intègre un certain nombre d'individus générés aléatoirement au sein de la population.
4. **évaluation** du degré d'adaptation de chaque individu ;
5. **remplacement** de la population courante par la nouvelle population ;
6. si un certain **critère d'arrêt** n'est pas atteint, on retourne à l'étape 2 ;

Néanmoins, ce schéma d'allure assez généraliste se révèle en réalité bien incomplet. En effet, la programmation d'un algorithme évolutionniste s'accompagne systématiquement de spécificités inhérentes au protocole utilisé, et comporte un ensemble de paramètres et de contraintes qui font de chaque algorithme un modèle de convergence précis et complexe. Ainsi, on ne parlera pas ici de la manière dont ces algorithmes gèrent par exemple l'information génétique — souvent à l'aide d'une structure chromosomique — ni du processus utilisé pour calculer le degré d'adaptation (*fitness*) d'un individu de la population, qui dépend bien souvent du problème posé.

D'une manière générale, on considère que les algorithmes évolutionnistes ont fait leur apparition dans les années soixante, décrivant ainsi trois grandes familles d'algorithmes :

- **Les stratégies d'évolution** proposées par Ingo RECHENBERG en 1965 ([105], [106]) sont véritablement les premières méta-heuristiques de type évolutionniste appliquées à des problèmes d'optimisation. À l'origine, l'algorithme utilise un ensemble de vecteurs de variables réelles qu'il manipule à l'aide d'opérateurs de sélection et de mutation.
- **La programmation évolutionniste** est née l'année suivante, en 1966, de l'imagination de FOGEL, OWENS et WALSH ([43]). Il s'agit d'une méthode d'optimisation stochastique pour la conception d'automates à états finis.
- **Les algorithmes génétiques** ont été proposés pour la première fois en 1975 par John HOLLAND ([56]) dans le cadre de l'optimisation combinatoire. Leur but est d'obtenir une solution approchée à un problème, en un temps raisonnable, par « bonds » sélectifs au sein d'une population d'individus. Par ailleurs, l'opérateur d'enjambement (*crossover*) a été spécialement introduit, en complément des opérateurs de sélection et de mutation.

Plus tard, la **programmation génétique** sera mise à l'honneur. Il s'agit en réalité d'une évolution de la programmation évolutionniste, dont John KOZA notamment proposera un formalisme au début des années 90 ([64], [65]).

11.3 LA PROGRAMMATION GÉNÉTIQUE

La programmation génétique, appelée *genetic programming* en anglais (et parfois raccourcie en *GP*), est une sous-catégorie des algorithmes évolutionnistes. Issue, à l'origine, de la programmation évolutionniste, elle est en quelque sorte une méthode d'apprentissage automatique utilisant les concepts fondamentaux des algorithmes génétiques appliqués à des programmes. En effet, là où la programmation évolutionniste vise à optimiser une succession d'actions à effectuer — d'où l'utilisation d'automates notamment — la programmation génétique va directement manipuler les instructions d'un programme, c'est-à-dire des opérateurs et opérandes, et selon le modèle choisi, des fonctions spécifiques.

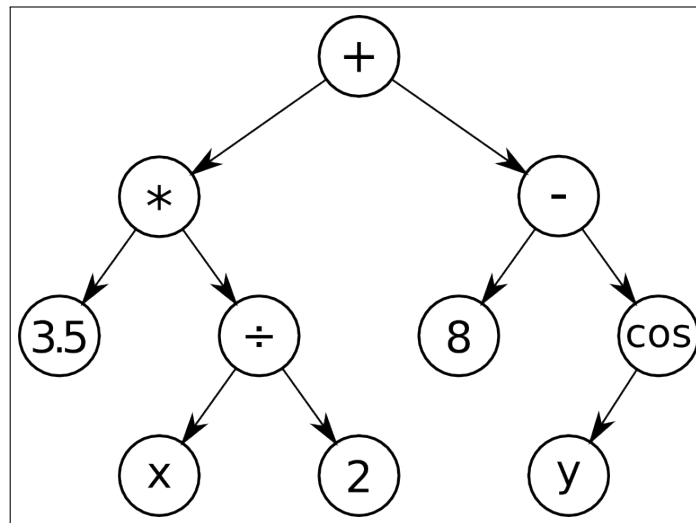


FIG. 11.2 – EXPRESSION SOUS FORME D'ARBRE SYNTAXIQUE

Traditionnellement, la programmation génétique est en mesure de résoudre les mêmes types de problèmes que d'autres techniques d'apprentissage automatique (*machine learning*) telles que les réseaux de neurones. Elle est en particulier adaptée aux problèmes d'exploration de données (*data mining*), c'est-à-dire la recherche de régularités au sein d'une grande quantité de données. En outre, la programmation génétique est applicable à d'autres domaines tels que les problèmes de contrôle, le traitement du signal, ou encore le traitement d'image.

Ces dernières décennies, différents types de programmations génétiques ont émergé, toutes basées sur le principe d'évolution de programmes. De cette manière, on peut dénombrer plusieurs catégories de modèles que sont les arbres syntaxiques aussi connue sous le nom *tree-based GP* ou *TGP* (KOZA [64] [65]), la représentation linéaire ou *linear GP* ou *LGP* (BANZHAF [7] [6], NORDIN [94], BRAMEIER [15]), sous forme de graphe ou *graph-based GP* (TELLER [126]), ou mêlant forme linéaire et arbres, connue sous le nom de *linear-tree GP* (KANTSCHIK [60]).

Toutefois, la forme la plus répandue et la plus emblématique est sans nul doute la représentation sous forme d'arbres, qui correspond à l'expression d'un langage de programmation fonctionnelle. Les fonctions sont alors situées dans les nœuds internes de l'arbre, tandis que les feuilles contiennent les constantes et variables d'entrée. Par exemple, la figure 11.2 utilise les variables d'entrée x et y pour représenter l'expression suivante :

$$(3.5 \times \frac{x}{2}) + (8 - \cos(y))$$

Les applications associées à la programmation génétique sont très variées, et vont de la reconnaissance d'images, à la planification de déplacements de robots,

en passant par la génération d'arbres de décision, la classification de protéines, ou encore la synthèse de circuits électroniques.

11.4 LA PROGRAMMATION GÉNÉTIQUE LINÉAIRE

La programmation génétique linéaire (LGP) allie la puissance et la flexibilité de la programmation génétique sous forme d'arbre avec la représentation itérative de tout code machine traditionnel, permettant notamment l'utilisation de sauts conditionnels et d'instructions manipulant la mémoire. En ce sens, la LGP se rapproche de l'ordinateur physique par le respect du principe de séquenceur inhérent à l'architecture de VON NEUMANN. En outre, la manipulation de code source en tant que matériel génétique n'est pas sans rappeler l'ADN des cellules vivantes.

Effectivement, tout comme l'ADN (acide désoxyribonucléique) permet la production des protéines nécessaires au fonctionnement de la cellule, le code source induit à terme le comportement du programme, notamment au travers de la manipulation des valeurs des registres mémoire. De plus, l'ADN possède des *introns*, c'est-à-dire des portions de gène non codantes dont le rôle est encore mal connu mais qui s'apparenteraient à un ensemble de matériels génétiques redondants permettant d'introduire une certaine combinatoire, en particulier lors de l'épissage (étape dans l'évolution de l'ARN). De la même façon, un programme généré par croisements de codes successifs possède lui aussi des introns, à savoir des groupes d'instructions dont l'intérêt peut paraître *a priori* nul, mais dont la fonction est double : ils servent de véritable réservoir de symboles pour les générations de programmes futures, et ils constituent des barrières efficaces qui réduisent les effets parfois destructeurs des croisements et mutations. Il existe deux types d'introns :

- **Les introns syntaxiques** qui sont des portions de code jamais évaluées, car victimes de branchements conditionnels systématiques ;
- **Les introns sémantiques** qui sont en revanche normalement évalués, mais qui n'apportent rien à l'exécution du programme. Par exemple, $x + (3 - 3)$.

L'intérêt d'utiliser ici les principes de la programmation génétique réside dans la constitution d'une activité émergente en assemblant les briques de base d'un programme informatique, en l'occurrence un ensemble d'*instructions-machine*. Le fait d'utiliser spécifiquement une forme linéaire permet de respecter la logique impérative du code, lequel exécute ses instructions une à une, sans hiérarchie particulière. Par conséquent, là où la TGP (programmation génétique sous forme d'arbre) sera plus adaptée à une formulation mathématique simple à représenter à l'aide par un arbre syntaxique, la LGP vise davantage une structuration plus séquentielle dans laquelle toute ligne de code possède le même statut. Par ailleurs, j'ai volontairement représenté les programmes du modèle de

manière non procédurale — c'est-à-dire qu'il n'existe pas de notion de bloc délimitant certaines zones de code — ne nécessitant donc pas l'usage d'une structure en arbre syntaxique.

Après avoir décrit l'approche émergentiste utilisée ainsi que les mécanismes généraux inhérents à la programmation génétique linéaire, découvrons à présent le fonctionnement du modèle GAEA.

SOMMAIRE

12.1	ENVIRONNEMENT	171
12.1.1	Le robot	171
12.1.2	Les proies	173
12.1.3	Les prédateurs	173
12.2	STRUCTURE D'UN PROGRAMME	173
12.2.1	Les registres mémoire	175
12.2.2	Les instructions du programme	176
12.2.2.1	Les affectations	176
12.2.2.2	Les conditionnelles	177
12.3	MOTEUR GÉNÉTIQUE	178
12.3.1	Algorithme principal	179
12.3.2	Initialisation	180
12.3.3	Sélection	180
12.3.3.1	Sélection initiale	180
12.3.3.2	Évaluation des programmes	181
12.3.3.3	Tournoi	181
12.3.4	Reproduction	182
12.3.4.1	Les croisements	182
12.3.4.2	Les mutations	183
12.3.5	Remplacement	185
12.4	AJOUT D'HEURISTIQUES	186

*There is no fire like passion, there is no shark like
hatred, there is no snare like folly, there is no torrent
like greed.^a*

— BUDDHA

a. « Il n'est pas de feu comme la passion, il n'est pas de requin comme la haine, il n'est pas de piège comme la folie, il n'est pas de torrent comme l'avidité. » (NDR)

D'ordinaire, les approches de type *bottom-up* jouent sur une activité verticale, du bas vers le haut, particulièrement forte afin de permettre la remontée d'informations qui vont ensuite se structurer dynamiquement dans les niveaux supérieurs. Dans ce projet, GAEA utilise un moteur génétique qui manipule un ensemble de programmes considérés comme des individus au sein d'une population en constante évolution. Chaque *individu-programme* se verra alors implanté au sein d'un robot baigné dans un système *proie-prédateur* en guise d'environnement simulé. Ce programme va alors utiliser les informations transmises par ses capteurs, s'exécuter en effectuant une à une les instructions qui le composent, et enfin envoyer ses résultats à des effecteurs responsables de la mobilité du robot.

En résumé, l'application GAEA, dont l'objectif est de déterminer la suite d'instructions la plus adaptée, va créer un grand nombre de programmes internes, tous implantés dans des robots différents et baignant dans des environnements successifs aléatoires, qui seront ensuite manipulés par des opérateurs génétiques afin d'être sélectionnés et croisés entre eux. Ce mécanisme va ainsi permettre à la population de programmes d'évoluer progressivement.

Pour bien comprendre le modèle dans son ensemble, il faut néanmoins scinder son fonctionnement en trois parties distinctes, et que l'on va traiter successivement. En premier lieu, il s'agit de définir l'*environnement*, c'est-à-dire le type de problème que chaque individu-robot de la population devra résoudre de la manière la plus efficace. Dans un second temps sera décrite la structure d'un individu et notamment le programme interne qui le compose. Ensuite, ce que l'on appelle le *moteur génétique*, à savoir l'algorithme général qui encadre la manipulation des différents programmes, sera explicité dans les détails. Enfin, nous aborderons l'ajout optionnel d'heuristiques permettant de faciliter l'évolution de la population.

12.1 ENVIRONNEMENT

L'environnement représenté est analogue à celui utilisé dans la partie précédente (chapitre 9), à savoir une zone finie dans laquelle vont évoluer un certain nombre d'entités appelées *créatures*. Parmi ces créatures, l'individu central est un robot doté de capteurs et effecteurs, et dans lequel un individu-programme sélectionné au sein de la population décrite précédemment sera implanté. En tout, il existe trois sortes de créatures représentées :

- le **robot** lui-même dont le comportement est entièrement déterminé par l'exécution de son programme génétique interne et dont la survie est conditionnée par sa capacité à se nourrir et à éviter ses prédateurs ;
- des **proies** organisées en essaims dont l'objectif est de fuir son prédateur, à savoir le robot. Leur présence est indispensable car elle permet à notre robot de se nourrir ;
- des **prédateurs** se déplaçant de façon aléatoire, mais qui se dirigeront vers notre robot, si celui-ci passe dans leur champ audio-visuel.

Là encore, l'intérêt de cette plateforme réside dans la contrainte imposée au robot afin de limiter au maximum ses possibilités et ainsi évaluer son taux d'adaptabilité de manière plus efficace. En effet, plus l'environnement soumis au robot sera dangereux et complexe, plus sa capacité de survie sera mise à l'épreuve. Il faut noter que les prédateurs et les proies n'interagissent pas directement entre eux.

En outre, des fonctionnalités additionnelles ont été implantées, notamment des capteurs ainsi que des attributs supplémentaires. L'application 3D a elle aussi été grandement améliorée.

12.1.1 Le robot

Le robot est l'entité principale de la simulation, ainsi que la créature centrale du système proie-prédateur représenté ici. En effet, il s'agit de l'individu dont l'objectif sera d'optimiser le comportement à l'aide d'un programme interne s'adaptant aux aléas environnementaux.

En plus des caractéristiques proposées précédemment que sont les capteurs visuels et auditifs (voir chapitre 9), le robot est doté d'un troisième type de dispositif : des capteurs tactiles. Ces derniers ont été ajoutés afin de permettre au robot de savoir s'il est en contact avec une proie ou l'un des quatre murs de la simulation, et ainsi obtenir des informations plus précises sur sa position au sein de l'environnement. Voici finalement la liste des capteurs disponibles :

- Les **capteurs visuels** permettent au robot de visualiser ses proies et prédateurs, ainsi que les murs d'enceinte de l'environnement, et décrivent un

champ de vision, dont la précision est proportionnelle au nombre de ces capteurs. Ils ont l'avantage de pouvoir détecter un obstacle à distance ;

- Les **capteurs auditifs** possèdent un rayon de détection plus réduit, mais décrivent néanmoins un champ auditif à 360 degrés autour du robot. Ils perçoivent les bruits produits par le déplacement des proies et prédateurs. Là encore, le nombre de capteurs permet de déterminer de façon plus précise la localisation des obstacles ;
- Les **capteurs tactiles** ont un rayon de détection réduit à une distance de contact. Leur intérêt est double : ils permettent l'alimentation en proies du robot, lequel ne peut se nourrir que si une proie est en contact direct, et donnent aussi des informations concernant la position exacte des murs de l'environnement, notamment lorsque ceux-ci se situent en-dehors du champ de vision du robot.

Sur le schéma de représentation (figure 12.1), les capteurs ont une couleur par défaut. Ils sont énumérés ici par ordre décroissant de leur distance par rapport au robot : **bleu** pour les capteurs visuels, **vert** pour les capteurs auditifs, et **marron** pour les capteurs tactiles. Si l'un d'eux détecte une présence, il se colore automatiquement de la couleur de l'obstacle : *jaune* pour une proie, *noir* pour un prédateur, et *rose* pour un mur. Leur nombre est librement paramétrable et détermine la précision de la sensibilité des capteurs. En effet, plus leur nombre est important et plus le robot sera capable de déterminer précisément d'où vient l'information détectée. D'un autre côté, la masse d'information à traiter sera bien plus importante et nécessitera plus de ressources et de temps de calcul.

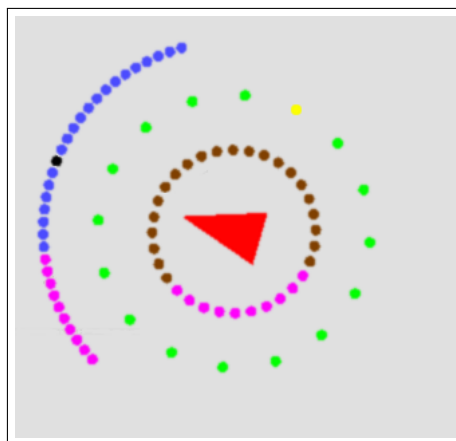


FIG. 12.1 – CAPTEURS DE L'INDIVIDU

Les informations détectées par les capteurs seront ensuite transmises au programme interne implanté au sein du robot sous la forme de *registres mémoire* d'entrée. Une fois ces registres d'entrée initialisés, le programme interne s'exécute. À l'issue de cette opération entrent alors en jeu les effecteurs, qui ont la

responsabilité de mettre en mouvement le robot en fonction des valeurs de sortie calculées par le programme. Au total, il existe trois effecteurs :

- **Un effecteur d'accélération** permet au robot de se déplacer linéairement, d'avant en arrière. La vitesse de déplacement est bornée. Une accélération négative fait ici office de marche arrière ;
- **Un effecteur de rotation** permet à l'individu de tourner sur lui-même afin de changer sa direction. Le robot peut alors pivoter vers la gauche ou la droite avec une vitesse de rotation maximale ;
- **Un effecteur d'alimentation** donne la possibilité au robot de se nourrir lorsqu'une proie est en contact avec lui.

Mais ce n'est pas tout ! En plus de ses capteurs et effecteurs, chaque robot possède un programme interne caractéristique, lui permettant de guider son comportement, et dont je développerai le fonctionnement dans la section suivante, consacrée au moteur génétique.

12.1.2 Les proies

Les proies représentent le centre d'intérêt du robot. Elles possèdent un champ de vision et d'audition propre et se déplacent en essaim, de la même manière que dans le modèle précédent (voir chapitre 9). Leur objectif est double : échapper au robot tout en restant relativement groupées. Il faut ajouter que l'instinct grégaire des proies ne doit pas être trop important afin qu'elles puissent se répartir dans tout l'environnement et augmenter leur visibilité. Enfin, elles doivent être en nombre suffisant afin de susciter l'intérêt du robot et permettre à celui-ci de vivre durablement.

12.1.3 Les prédateurs

Les prédateurs sont des créatures peu nombreuses mais particulièrement redoutables dans la mesure où leur champ de vision est gigantesque. Cela leur permet de menacer notre robot plus souvent et plus facilement, ce dernier ayant des capacités de détection plus limitées. Pour compenser cette supériorité, leur vitesse de déplacement maximale a été revue à la baisse. Enfin, leur comportement par défaut consiste simplement à se mouvoir en ligne droite et à changer de direction au contact d'un des quatre murs de l'environnement.

12.2 STRUCTURE D'UN PROGRAMME

Le schéma de la figure 12.2 décrit le fonctionnement général d'un robot de type GAEA. Il est constitué d'attributs, ainsi que d'un programme interne.

L'unique attribut représenté ici est l'énergie, c'est-à-dire une valeur qui détermine la santé du robot — les autres attributs désignent par exemple la vitesse de déplacement ou l'angle de vue, mais ne représentent pas des valeurs pertinentes liées à l'évolution interne du robot. L'énergie diminue en fonction du temps, jusqu'à atteindre zéro, signe que l'individu n'est plus en mesure de se déplacer. Il sera donc considéré comme mort. Par contre, la consommation de proies permet d'augmenter le capital d'énergie disponible et ainsi de rallonger son espérance de vie.

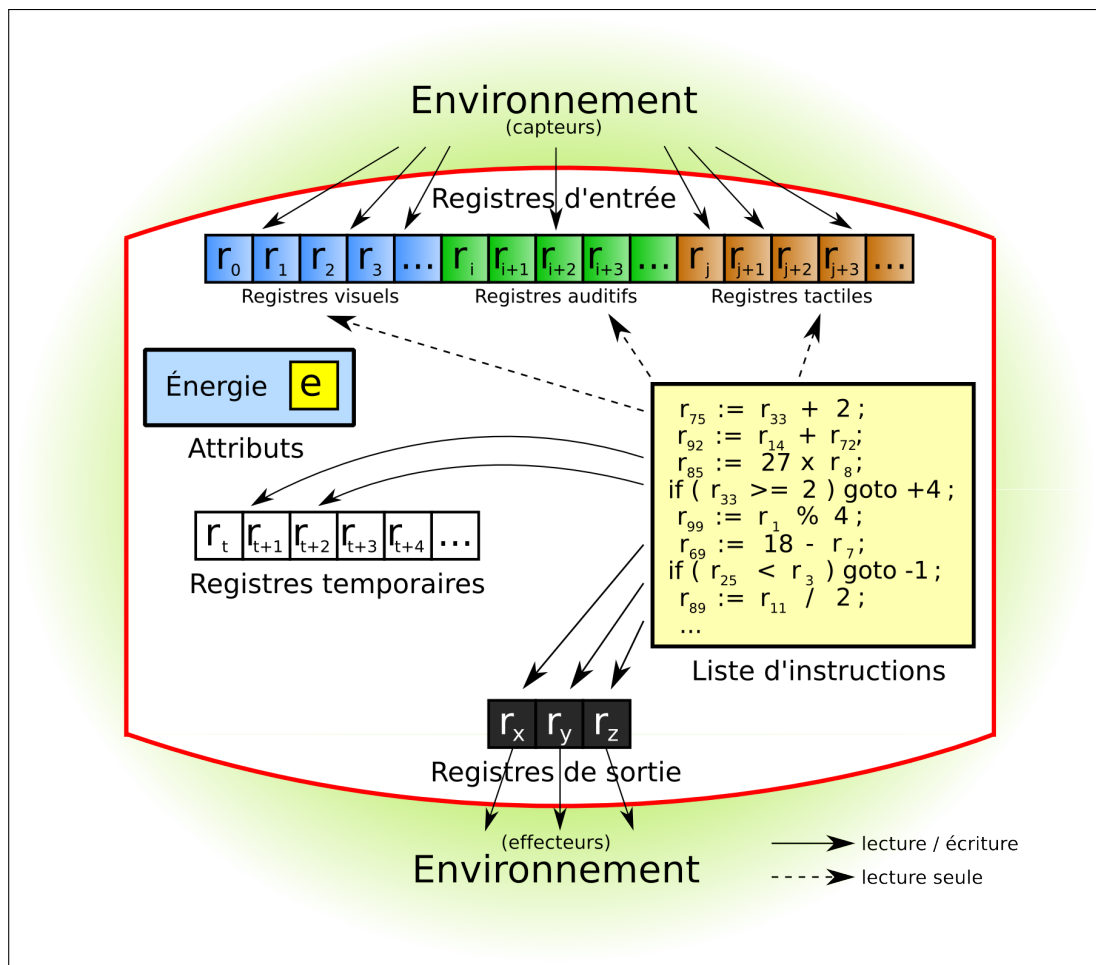


FIG. 12.2 – ORGANISATION GÉNÉRALE DU ROBOT

En outre, un robot se compose d'un programme interne, composé d'une part de registres mémoire dont les entrées et sorties permettent de s'interfacer avec l'environnement via les capteurs et effecteurs du robot, et d'autre part, d'une liste d'instructions machine permettant de manipuler la valeur des registres. L'objectif de cette section est de décrire plus précisément les composantes de tout programme contenu au sein d'un individu.

12.2.1 Les registres mémoire

Comme nous l'avons vu ci-dessus, l'individu modélisé possède des capteurs visuels et auditifs lui permettant de récupérer les informations de l'environnement. Ces informations seront ensuite transmises à son programme interne sous la forme de registres mémoire. Les registres sont par définition le moyen de stockage et de modification de l'information numérique. Ils représentent un degré de liberté non contrôlé du système et servent de base à l'exécution des processus d'auto-organisation. En somme, ces registres constituent les variables locales du programme génétique et sont par conséquent librement modifiables par celui-ci. Afin de garantir leur facilité d'accès, les registres possèdent tous le même nom r , suivi de leur numéro d'identification unique. Le modèle GAEA utilise trois types de registres mémoire (voir figure 12.2) :

- **Les registres d'entrée** dont la valeur dépend de l'état des capteurs visuels, auditifs et tactiles du robot, transmis à chaque étape de la simulation. Ils sont accessibles en lecture seule par le programme interne pour le calcul des valeurs des autres registres. Les registres d'entrée sont eux-mêmes subdivisés en catégories correspondant aux capteurs du robot. Par conséquent, à chaque capteur visuel sera attribué un registre visuel en début de mémoire ; il en est de même pour les capteurs auditifs et tactiles. Ainsi, il existe autant de registres d'entrée que de capteurs ;
- **Les registres temporaires** sont utilisés de façon provisoire par le programme interne, afin de calculer les valeurs de sortie. Contrairement aux registres d'entrée, ils sont accessibles en lecture mais aussi en écriture. De même que pour les registres d'entrée, le nombre de registres temporaires est libre et configurable, bien que relativement limité afin d'éviter une éventuelle explosion combinatoire ;
- **Les registres de sortie** recueillent les valeurs calculées par le programme interne et sont assimilables à valeurs de retour pour celui-ci. Par conséquent, tout comme les registres temporaires, ils sont accessibles et modifiables. Par ailleurs, les trois registres de sortie correspondent aux trois effecteurs du robot, lesquels sont responsables de ses trois actions possibles (déplacements d'avant en arrière, rotation sur lui-même et consommation de nourriture).

Afin de faciliter la discrétisation des valeurs, GAEA limite le type des registres à des entiers relatifs, situés entre une valeur minimale et une valeur maximale déterminées par l'utilisateur en début de simulation. Cela signifie, entre autres, qu'un registre dont la valeur est amenée à franchir une des deux bornes imposées par le modèle se verra automatiquement attribuer la valeur de la borne.

De la même façon, si une instruction du programme exécute une division par zéro — et ceci est valable aussi pour un modulo — aucune erreur ne sera

relevée, mais la valeur retournée sera constante, fixée par l'utilisateur, typiquement 0 ou une des deux valeurs minimale ou maximale autorisées. Les situations dans lesquelles une division par zéro peut se produire sont extrêmement difficiles à prédire car elles sont le plus souvent le fruit d'un mauvais concours de circonstances concernant la valeur du dénominateur. Par conséquent, elles sont inévitables et ne peuvent être considérées comme des erreurs de programmation.

À la fin de l'exécution d'un programme, l'état des trois registres de sortie est finalement envoyé aux trois effecteurs, lesquels seront alors responsables du comportement du robot au sein de l'environnement.

12.2.2 Les instructions du programme

Traditionnellement, un programme informatique est composé d'une suite d'instructions distinctes. Pour des raisons de simplicité et de rapidité de convergence, GAEA n'utilise que deux types d'instructions : l'*affectation* et la *conditionnelle*.

Ces lignes ordonnées de code constituant le programme interne du robot sont capables d'accéder aux valeurs contenues au sein des registres mémoire et permettent éventuellement de modifier celles-ci. Les affectations en particulier sont les seules instructions prenant en paramètre un *registre résultat* modifiable, et dont elles peuvent successivement altérer la valeur, à chaque unité de temps correspondant à l'exécution d'une instruction.

Généralement, un programme est doté d'un *compteur ordinal* qui détermine quelle instruction est exécutée, ce qui permet à tout moment de connaître l'état d'avancement du programme, ainsi que de permettre les branchements, conditionnels ou non, qui viennent rompre le flot d'exécution afin d'accéder à d'autres parties du programme. Ici, le compteur ordinal débute naturellement à la première instruction, l'exécution du programme s'effectuant séquentiellement, à chaque unité de temps, à moins qu'une conditionnelle dont le test s'avère positif ne vienne déplacer le compteur ordinal.

On peut noter que GAEA n'utilise pas de *sauts non-conditionnels*, car ceux-ci supposent l'usage d'étiquettes, lesquelles constitueraient une troisième catégorie d'instructions, ainsi qu'une source potentielle d'erreurs au sein des programmes.

12.2.2.1 Les affectations

L'affectation consiste à mettre à jour la valeur d'un registre en lui appliquant le résultat d'un calcul particulier. Ce calcul doit être de la forme :

$$r_i := o_1 f^a o_2$$

r_i représente le registre résultat, c'est-à-dire la variable que l'on souhaite modifier. Il ne peut s'agir que d'un registre temporaire ou de sortie, car les registres

d'entrée sont uniquement accessibles en lecture. En outre, f^a est un opérateur arithmétique d'arité 2 parmi les fonctions mathématiques traditionnelles (+, −, ×, ÷, ainsi que le modulo) et o_1 et o_2 sont ses opérandes, c'est-à-dire une constante de type numérique dans l'intervalle autorisé ou bien un autre registre — cette fois parmi tous les types de registres car tous sont au moins accessibles en lecture.

Afin d'optimiser l'efficacité de l'exécution d'une affectation, le modèle interdit l'utilisation de deux constantes pour o_1 et o_2 . Ainsi, une instruction du type $r_{12} := 8 + 17$ est irrecevable.

Exemple d'affectation :

$$r_{85} := 27 \times r_8$$

12.2.2.2 Les conditionnelles

La structure conditionnelle permet l'utilisation de ce que la programmation impérative appelle le *saut conditionnel* ou *branchement conditionnel* — aussi appelé *goto* pour « aller vers » en anglais. Il s'agit simplement du déplacement de la séquence d'exécution à un autre endroit du programme, lui-même identifié par un saut conditionnel. Dans le cas de GAFA, il s'effectue de la façon suivante :

$$\text{if } (o_1 f^b o_2) \text{ goto } +n$$

Les deux opérandes o_1 et o_2 peuvent être là encore des registres ou des constantes, mais néanmoins, l'opérateur f^b qui les sépare est un opérateur booléen d'arité 2. Si la condition est vérifiée, le saut conditionnel (*goto*) s'opère et la séquence d'exécution est déplacée à la n -ième structure conditionnelle suivante ou précédente, selon le signe du saut (voir figure 12.3). Si en revanche la condition n'est pas vérifiée, l'exécution du programme se poursuit normalement après l'instruction conditionnelle. Il faut toutefois noter un cas particulier : si la condition est vraie et que n vaut zéro, le saut conditionnel force le programme à s'arrêter.

L'alternative selon laquelle le saut s'effectue juste avant ou juste après la condition-objectif est paramétrable. De même, il est possible d'interdire les sauts négatifs, ce qui a pour conséquence d'éliminer les boucles. En effet, les sauts dirigés vers une instruction conditionnelle précédente occasionnent souvent l'exécution répétée d'une même portion de code. Ce comportement peut engendrer la création de boucles infinies, notamment si le test effectué concerne un ou des registres non modifiés au sein de la boucle ; cette situation n'ayant aucune chance de provoquer une inversion du résultat du test. Afin de garantir l'achèvement du programme, l'algorithme détermine une valeur représentant le nombre maximal d'instructions que l'on peut effectuer. Une fois cette limite dépassée, l'exécution du programme prend fin automatiquement.

Exemples de structures conditionnelles :

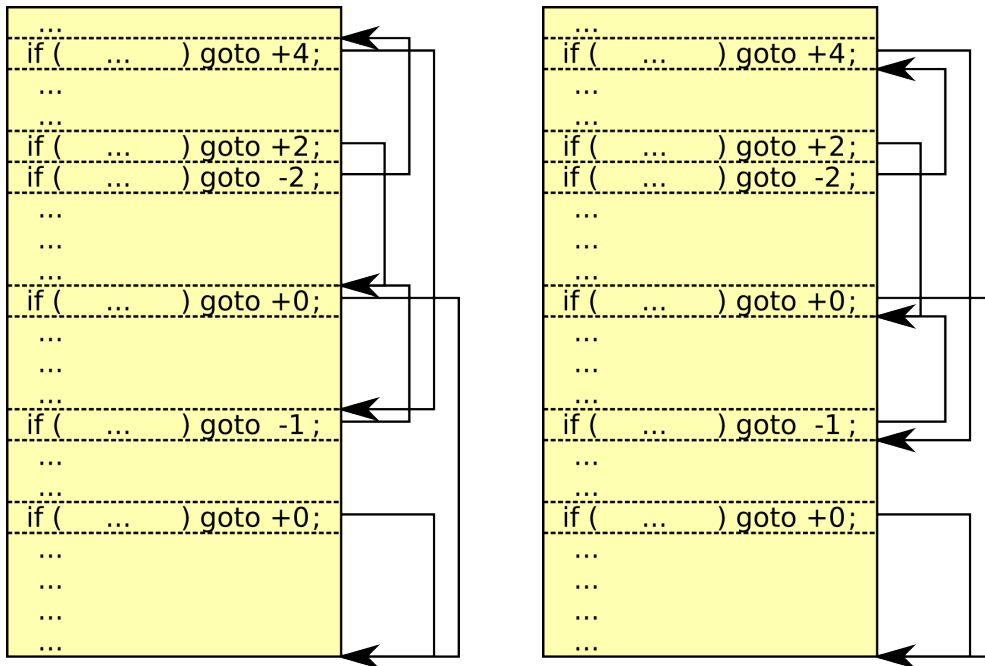


FIG. 12.3 – FONCTIONNEMENT DU SAUT CONDITIONNEL. LE BRANCHEMENT PEUT S’EFFECTUER AVANT LA CONDITION-OBJECTIF (À GAUCHE) OU APRÈS (À DROITE)

```
if (r25 ≥ 2) goto +3
if (r3 < r2) goto -2
```

Ainsi, les deux seules manières d’achever l’exécution d’un programme sont soit d’atteindre le nombre maximal d’instructions autorisées, soit d’atteindre l’ultime ligne de code.

Dans la suite de ce document, tous les résultats présentés utiliseront les sauts négatifs, afin de produire des programmes plus riches par l’apparition de boucles. En outre, les branchements se feront systématiquement après les conditions-objectifs afin d’éviter, notamment dans le cas d’une boucle, d’avoir à réaliser deux tests successifs.

12.3 MOTEUR GÉNÉTIQUE

Le moteur génétique est la partie centrale de GAEA. Il permet de manipuler la population d’individus-programmes au cours des différentes générations, gère l’évaluation de ces programmes et applique sur certains d’entre eux un ensemble d’opérations génétiques.

Voici une description du fonctionnement du moteur génétique de GAEA, découpée en cinq grandes parties. Tout d’abord, nous verrons comment se déroule

l'algorithme principal. Dans un second temps, on parlera de *l'initialisation* des programmes de la population. Ensuite, seront développées plus précisément les étapes de *sélection* utilisées. Dans une quatrième partie, on décrira les opérateurs génétiques relatifs à la *reproduction*, que sont les croisements et les mutations. Enfin, la phase finale du *remplacement* sera abordée, ce qui permet d'achever le renouvellement de la population.

12.3.1 Algorithme principal

Le déroulement de l'algorithme du moteur génétique est à l'image de nombreux algorithmes évolutionnistes. En effet, il se compose des grandes étapes définies précédemment que sont : l'initialisation de la population, la sélection d'une partie des individus, la création de nouveaux individus, et enfin le remplacement d'anciens individus par des nouveaux.

Algorithme 4 Algorithme principal du moteur génétique

```

1: population[] : tableau de programmes // Population de programmes
2: nbMondes : entier // Nombre de mondes créés à chaque génération
3: prgmSelection[] : tableau de programmes // Programmes sélectionnés
4: prgmGagnants[] : tableau de programmes // Programmes ayant gagné le tournoi
5: prgmNouveaux[] : tableau de programmes // Programmes nouvellement créés
6:
7: population ← initialisation()
8: Répéter
9:   genererMondes(nbMondes)
10:  // Sélection
11:  prgmSelection ← selection(population)
12:  evaluation(prgmSelection)
13:  prgmGagnants ← tournoi(prgmSelection)
14:  // Reproduction
15:  prgmNouveaux ← croisement(prgmGagnants)
16:  mutation(prgmNouveaux)
17:  // Remplacement
18:  evaluation(prgmNouveaux)
19:  remplacement(population, prgmNouveaux)
20: Jusqu'à atteindre un certain nombre de générations

```

Ce fonctionnement général est décrit dans l'algorithme 4 qui résume la manière dont sont articulées les différentes phases du moteur génétique. Il se compose tout d'abord d'une étape d'initialisation de la population d'individus représentés à l'aide d'un tableau. Ensuite, la boucle principale fait évoluer cette population à l'aide de fonctions manipulant des tableaux de programmes, les-

quels sont transmis de fonction en fonction jusqu'au stade de remplacement final.

Il est à noter que chaque étape de cet algorithme fait accessoirement l'objet d'éventuelles opérations de stockage des résultats à l'aide de fichiers journalisés (fichiers *log*) afin de conserver un historique des actions effectuées, ainsi que d'une sauvegarde systématique des programmes et mondes générés, avec un objectif de réutilisation ou de test *a posteriori*.

12.3.2 Initialisation

Comme mentionné précédemment, la population est un tableau composé initialement de programmes aléatoires. Pour ce faire, chaque case du tableau appellera un constructeur de programme aléatoire, sorte d'usine de production générant un ensemble d'instructions accompagné d'une liste de registres mémoire liés à ces instructions. Il suffit ensuite d'attribuer à chacun de ces individus un identifiant unique afin de suivre plus facilement son évolution au sein de sa population.

En outre, au début de chaque tour de boucle principale apparaît une étape dite de *génération des mondes* consistant simplement à créer l'ensemble des mondes virtuels — c'est-à-dire les positions des proies et prédateurs au sein de l'environnement — nécessaires aux étapes futures d'évaluation des programmes.

L'utilisation de mondes aléatoires différents pour chaque boucle de simulation n'est pas anodine. En effet, l'objectif du modèle GAEA étant de créer des programmes auto-adaptatifs, il faut à tout prix éviter d'utiliser systématiquement les mêmes positions initiales pour les proies et prédateurs car cela reviendrait à générer un profil de programmes adaptés uniquement à certaines configurations et pas à d'autres. D'autre part, les mondes générés doivent être identiques au sein d'une même génération afin de garantir une certaine équité en termes d'évaluation, notamment lors de l'étape du tournoi.

12.3.3 Sélection

La sélection des individus consiste à déterminer quels seront les parents des futurs enfants qui rejoindront la population. Cette sélection se fait en trois étapes : la *sélection initiale*, l'*évaluation* des programmes et l'étape du *tournoi*.

12.3.3.1 Sélection initiale

La sélection initiale consiste simplement à choisir aléatoirement dans la population un certain nombre de programmes. Ce nombre est d'ailleurs paramétrable. Ce choix arbitraire permet de désigner un ensemble d'individus représentatifs de la population, afin de les faire participer à un tournoi.

12.3.3.2 Évaluation des programmes

À l'issue de la sélection initiale a lieu une étape primordiale : l'évaluation des programmes. Cela consiste à confronter tous les individus sélectionnés à l'ensemble des mondes générés au commencement de chaque boucle principale. À l'issue de chaque simulation est attribué un *score* correspondant à leur durée de vie, ce qui implique que plus l'individu survivra longtemps au sein de ce monde, plus son score sera élevé.

En effet, lors de chaque simulation, la position des créatures de l'environnement est mise à jour selon les règles de déplacement propres à chacun : les prédateurs avancent linéairement et traquent le robot, les proies se déplacent selon leur algorithme de *boids* (REYNOLDS, [107]), et le robot est conditionné par le résultat de ses registres de sortie qui guident ses effecteurs. La simulation prend fin lorsque le robot est capturé par un prédateur ou si son énergie atteint zéro.

Mais ce n'est pas tout ! L'ensemble des scores relatifs aux différents mondes testés fait l'objet du calcul de ce que l'on appelle la *fitness* du programme. En effet, afin d'atténuer les éventuels effets positifs (ou négatifs) que peuvent avoir certains mondes favorables (ou défavorables) à l'égard de certains individus, la *fitness* est le résultat de la moyenne des scores d'un individu, ôtés du meilleur et du moins bon score. Il s'agit donc d'une sorte de « méta-score » qui détermine le taux d'adaptabilité général d'un programme. Ainsi, plus un programme possèdera une *fitness* élevée, plus il aura de chances d'intégrer ou de réintégrer la génération suivante.

Il faut noter au passage que l'étape d'évaluation est la phase la plus lente de l'algorithme, car elle réclame l'exécution de chaque programme sélectionné pour chaque monde. Toutefois, l'indépendance de chaque simulation effectuée permet une parallélisation aisée.

12.3.3.3 Tournoi

L'étape du tournoi arrive en troisième position dans la phase de sélection. Lors d'un tournoi, les programmes ayant fait l'objet de la sélection initiale participent à une succession de duels dans lesquels une *roulette biaisée* détermine quel est le vainqueur du duel. Il s'agit d'une roulette classique dans laquelle chaque individu est représenté par une portion proportionnelle à la valeur de sa *fitness*. On effectue ensuite un tirage au sort homogène traditionnel sur cette roulette. Si un programme est tiré au sort, il gagne un duel et se voit autorisé à participer à l'étape suivante du tournoi. À l'inverse, s'il perd son duel, il est définitivement éliminé du tournoi (voir figure 12.4).

Lorsqu'un tournoi s'achève, il doit rester au moins deux vainqueurs, lesquels serviront de parents aux enfants à venir. En effet, dès lors que la sélection des parents est effectuée, les programmes vainqueurs du tournoi sont croisés deux à deux pour produire leur descendance lors de l'étape dite de la *reproduction*.

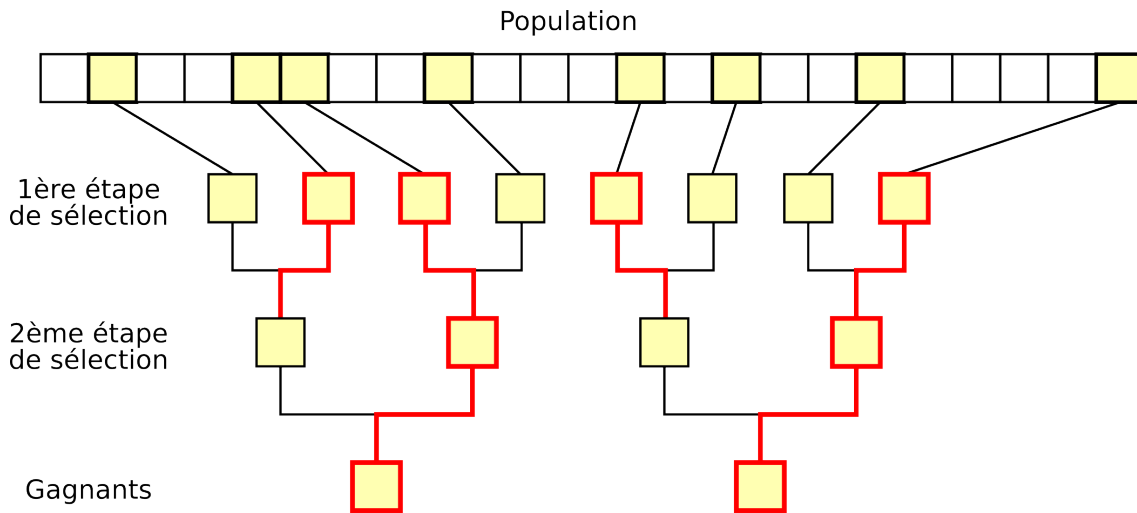


FIG. 12.4 – SÉLECTION DES INDIVIDUS PAR TOURNOI

12.3.4 Reproduction

Dans les algorithmes évolutionnistes interviennent systématiquement des *opérateurs génétiques* dont le rôle est de manipuler, de génération en génération, les individus peuplant la population. Le modèle GAEA quant à lui utilise en premier lieu un opérateur de *croisement*, afin de procéder à la recombinaison des parents vainqueurs du tournoi et de produire leurs enfants, et dans un second temps un opérateur de *mutation* en vue d'altérer le génome des enfants à naître.

12.3.4.1 Les croisements

Un croisement (enjambement, *crossover* ou encore recombinaison) est une manipulation consistant à produire des individus dont le patrimoine génétique sera en partie commun à celui de leurs deux parents. Comme indiqué sur la figure 12.5, le procédé réside dans le choix de deux portions de code appelées *segments* — une dans chacun des deux parents — que l'on va ensuite intervertir afin de produire les descendants.

Tout commence par la sélection d'un point particulier que l'on appelle *chiasma* (noté p_1 ou p_2 sur le schéma) en sachant que la distance (notée d_c) entre les deux *chiasmata* relativement à la première ligne de code (Position 0) doit être inférieure à une valeur maximale donnée, choisie par l'utilisateur. De la même façon, la différence de longueur (d_l) entre les deux segments (l_1 et l_2) ne doit pas dépasser un maximum, fixé là aussi par l'utilisateur.

À l'issue d'un croisement de code, des erreurs de programmation peuvent se produire. En effet, il peut arriver que les valeurs des sauts conditionnels pointent vers des conditions-cibles désormais inexistantes. Ainsi, un saut d'une valeur +4 ne peut être cohérent que s'il existe effectivement au moins 4 conditions entre lui

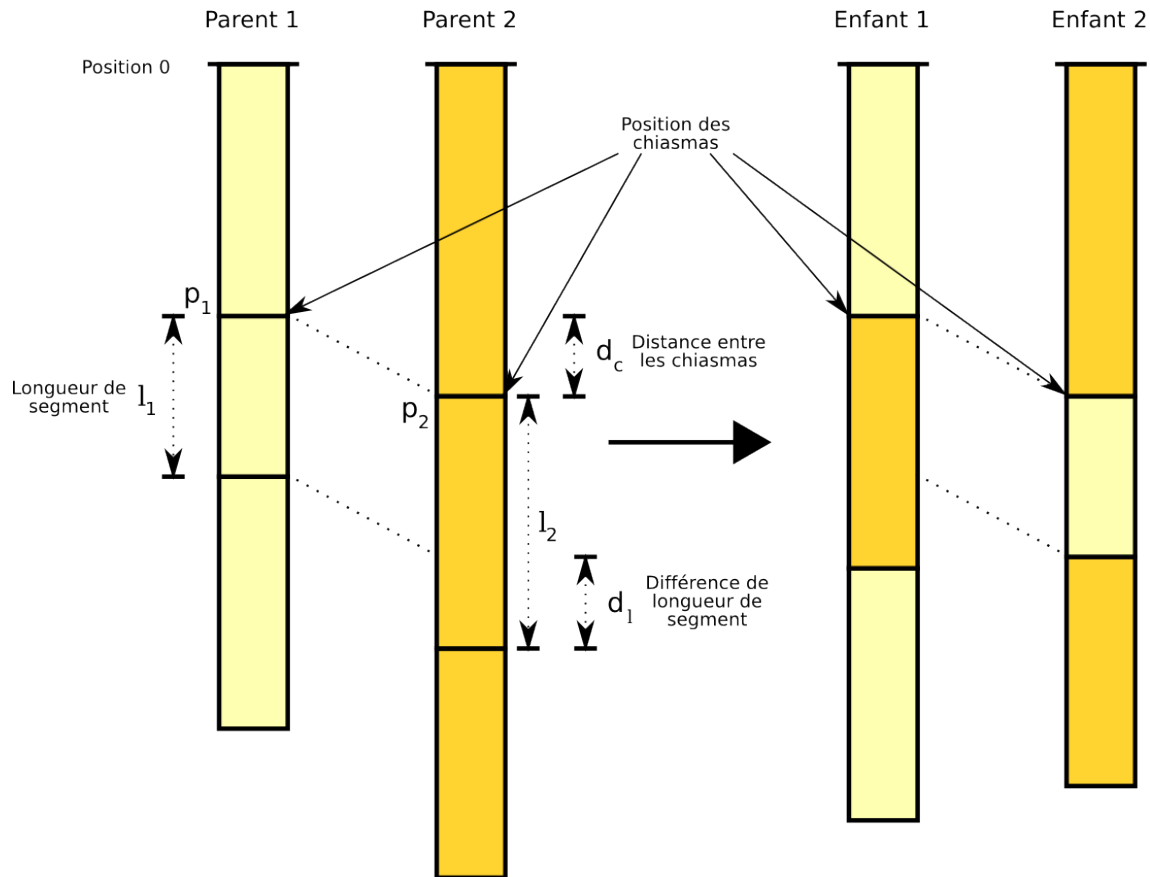


FIG. 12.5 – CROISEMENT ENTRE DEUX INDIVIDUS-PROGRAMMES

et la dernière ligne du programme. Or, il peut arriver que le croisement ait lieu dans cette zone et fasse disparaître une ou plusieurs conditions, auquel cas la valeur de saut devient illogique car trop élevée. Afin de remédier à ce problème, il suffit alors d'identifier les sauts erronés et les modifier par une valeur aléatoire mais cohérente.

Cependant, dans un processus de reproduction, effectuer un simple croisement ne suffit pas ! Un algorithme génétique quel qu'il soit doit en effet engendrer suffisamment de nouveauté dans le matériel génétique produit pour pouvoir explorer d'autres possibles et agrandir ainsi l'espace de recherche. Les mutations entrent alors en jeu.

12.3.4.2 Les mutations

Une mutation permet de modifier volontairement le patrimoine génétique d'un individu afin d'éviter la convergence prématurée de l'algorithme et ainsi garantir une certaine diversité dans l'exploration des génotypes possibles. GA EA

utilise deux types de mutations : les *macro-mutations* et les *micro-mutations* (voir figure 12.6).

Une macro-mutation modifie toute une zone de code et la remplace par un ensemble d'instructions générées aléatoirement. La taille de la zone de code supprimée (l_s sur le schéma 12.6(a)) ainsi que le nombre d'instructions insérées (l_i) sont maximisés afin d'éviter que l'altération des programmes ne soit trop importante.

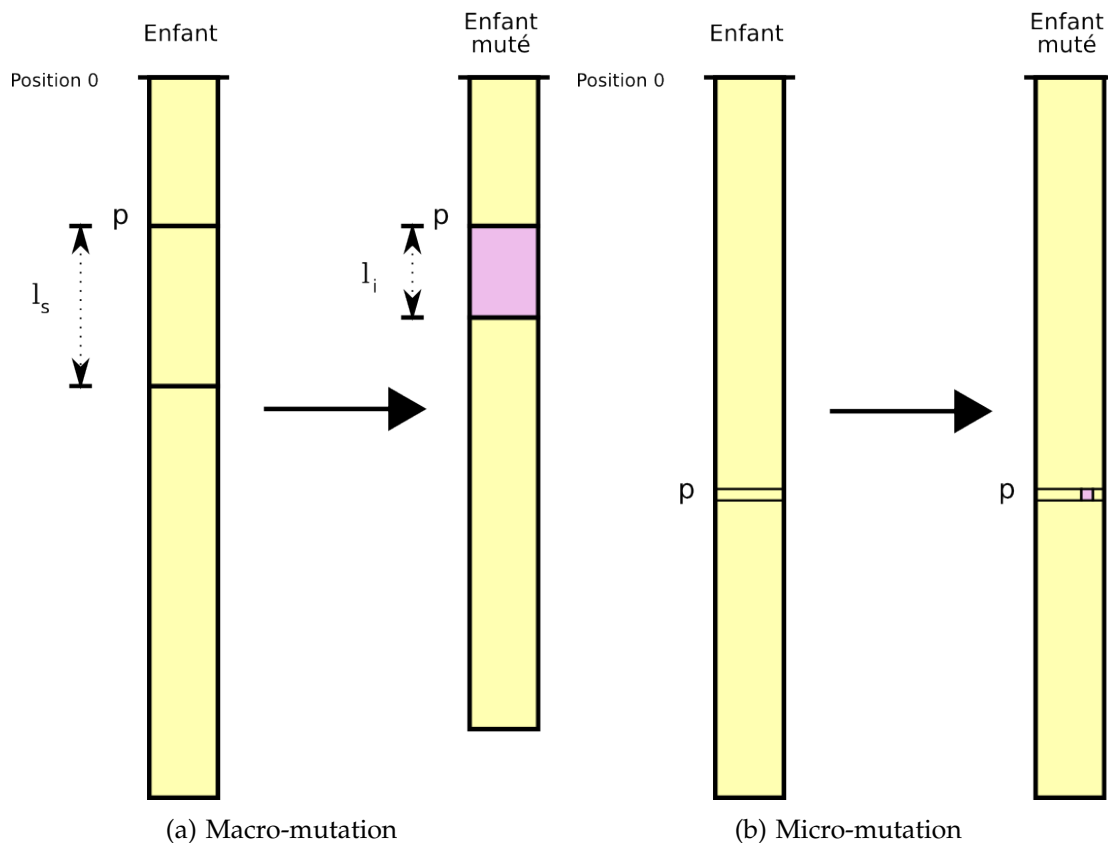


FIG. 12.6 – LES MUTATIONS

Une micro-mutation, comme son nom l'indique, ne transforme qu'une petite portion d'un programme (voir figure 12.6(b)). En l'occurrence, il s'agit ici de ne modifier qu'un opérande ou un opérateur au sein d'une instruction sans changer le type de cette dernière. Par souci de cohérence, un opérateur par exemple sera nécessairement remplacé par un autre opérateur du même type (arithmétique ou booléen), et ce afin de conserver une syntaxe correcte. Néanmoins, un registre pourra se voir converti en valeur constante et inversement.

À l'instar des croisements, les mutations peuvent produire des erreurs de sauts conditionnels, lesquelles seront corrigées de la même façon.

12.3.5 Remplacement

Finalement, lorsque les nouveaux programmes issus des processus de croisement et de mutation sont nés, ils entrent dans une phase dite de remplacement qui détermine s'ils sont autorisés ou non à prendre place au sein de la population.

Avant toute chose, les programmes nouvellement créés sont évalués, exactement de la même manière que les individus précédents l'ont été lors de l'étape de sélection, et en utilisant les mêmes mondes générés en début de boucle principale. Lorsque cette évaluation se termine, l'algorithme de remplacement est exécuté (voir algorithme 5).

Algorithme 5 Remplacement de la population

```

1: population[] : tableau de programmes // Population de programmes
2: prgmGagnants[] : tableau de programmes // Programmes ayant gagné le tournoi
3: prgmPerdants[] : tableau de programmes // Programmes ayant perdu le tournoi
4: prgmNouveaux[] : tableau de programmes // Programmes nouvellement créés
5:
6: // On trie les nouveaux programmes par ordre de fitness décroissant
7: trierInverse(prgmNouveaux)
8: // On trie les gagnants par ordre de fitness croissant
9: trier(prgmGagnants)
10: // On laisse les perdants dans l'ordre dans lequel ils ont été éliminés
11:
12: p ← 0 // indice du tableau des perdants
13: n ← 0 // indice du tableau des nouveaux
14: Tant que ((p < taille(prgmPerdants)) et (n < taille(prgmNouveaux))) Faire
15:   Si (prgmPerdants[p].fitness() ≤ prgmNouveaux[n].fitness()) Alors
16:     remplacer prgmPerdants[p] par prgmNouveaux[n] dans population
17:     n ← n + 1
18:   Fin Si
19:   p ← p + 1
20: Fin Tant que
21:
22: g ← 0 // indice du tableau des gagnants
23: Tant que ((g < taille(prgmGagnants)) et (n < taille(prgmNouveaux))) Faire
24:   Si (prgmGagnants[g].fitness() ≤ prgmNouveaux[n].fitness()) Alors
25:     remplacer prgmGagnants[g] par prgmNouveaux[n] dans population
26:     n ← n + 1
27:   Fin Si
28:   g ← g + 1
29: Fin Tant que

```

L'objectif de cet algorithme est d'introduire progressivement les nouveaux

nés dans la population en remplacement des anciens individus, par ordre de mérite. Pour ce faire, on confronte chaque nouveau programme aux perdants du tournoi, puis aux gagnants de ce même tournoi. Au préalable, les nouveaux programmes sont stockés dans un tableau, par ordre décroissant de leur *fitness*, de la plus grande à la plus petite, afin de placer en priorité les meilleurs nouveaux individus. En parallèle, les gagnants du tournoi sont eux aussi stockés dans un tableau, mais par ordre croissant de *fitness*, dans le but de n'éliminer les gagnants qu'en commençant par le moins bon. Quant aux perdants du tournoi, ils sont stockés dans l'ordre dans lequel ils ont été éliminés.

L'algorithme débute tout d'abord par les perdants du tournoi. Si un nouveau programme possède une *fitness* meilleure que le premier perdant du tournoi, le nouveau prendra la place du perdant. Dans le cas contraire, on passe au perdant suivant, et ainsi de suite. En fin de boucle, s'il reste des nouveaux programmes à placer, on passe ensuite aux gagnants du tournoi, jusqu'à ce que tous les nouveaux individus aient été introduits dans la population, ou que l'on ait passé en revue l'intégralité du tableau des gagnants.

À la fin d'une génération, plusieurs scénarii potentiels peuvent se dérouler pour chaque individu :

- un programme peut gagner le tournoi, être utilisé en tant que parent et rester au sein de la population ;
- un programme peut gagner le tournoi, être utilisé en tant que parent et se voir éradiqué de la population lors de la phase de remplacement ;
- un programme peut perdre le tournoi et rester malgré tout au sein de la population si les nouveaux individus possèdent globalement une faible *fitness* ;
- un programme peut perdre le tournoi et se voir éradiqué de la population lors de la phase de remplacement ;
- un programme peut ne pas avoir fait l'objet d'une sélection et par conséquent rester au sein de la population.

12.4 AJOUT D'HEURISTIQUES

Notre modèle doit faire émerger l'émotion sans qu'elle ne soit initialement construite. Cependant, afin de faire converger le modèle plus rapidement, il nous est possible de décrire indirectement l'état du système à l'aide d'un ensemble d'heuristiques sous la forme d'attributs particuliers liés au robot. En somme, on introduit un certain nombre d'indices détectables par le robot, non pas pour l'avertir d'un danger quelconque par exemple — ce qui reviendrait à influencer physiquement le comportement —, mais à simplement fournir les informations susceptibles de se révéler pertinentes pour faciliter l'émergence émotionnelle.

Le modèle précédent voit par conséquent ses registres d'entrée composés non seulement des capteurs habituels, mais aussi de zones mémoire initialisées à l'aide de valeurs particulières qui déterminent l'état interne du robot : l'énergie, la douleur et le plaisir.

- **L'énergie** est une composante essentielle qui joue ici le rôle de jauge de santé et exerce une pression sur le robot afin que celui-ci se nourrisse. En effet, le niveau d'énergie décroît en fonction du temps, et s'il tombe à zéro, la simulation s'arrête aussitôt, l'individu étant alors considéré comme mort. À l'inverse, si celui-ci consomme une proie, il se verra attribuer une certaine quantité d'énergie lui permettant ainsi de vivre plus longtemps ;
- **La douleur**, comme son nom l'indique, mesure la souffrance du robot au cours de sa vie. Elle est liée à l'apparition de deux phénomènes que sont la peur ressentie par la menace d'un prédateur à proximité, et la difficulté engendrée par le manque d'énergie vitale pour se déplacer. Le calcul de la douleur se fait donc en identifiant la position du prédateur le plus proche — si prédateur détecté il y a —, laquelle va déterminer l'importance de la peur ressentie. D'un autre côté, si l'énergie disponible tombe en-dessous d'un certain seuil, la douleur sera incrémentée d'une valeur inversement proportionnelle au taux d'énergie. Ce procédé permet d'informer le robot de son état interne, tout en l'aidant à prendre conscience des prédateurs qui l'entourent ;
- **Le plaisir** quant à lui symbolise la satisfaction de poursuivre une ou des proies. Il est proportionnel à la distance qui sépare le robot de la proie la plus proche — uniquement si le robot a détecté une proie.

Toutes ces mesures ont tendance à redéfinir le concept d'adaptabilité, la situation de départ consistant simplement à survivre le plus longtemps possible, et la seconde version avec heuristiques optimisant le comportement en proposant l'ajout de valeurs subjectives en guise de capteurs. Il s'agit en quelque sorte d'enrichir la réalité observée par le robot en lui suggérant des indices numériques que l'informaticien considère comme facilitant la convergence. Par ailleurs, aucune de ces mesures n'a d'impact sur le calcul de la *fitness*.

Chacune de ces valeurs est reportée sur une zone mémoire particulière au niveau des registres d'entrée (voir figure 12.7). Elles sont répétées à l'identique sur un ensemble de registres afin d'augmenter leur visibilité par rapport au programme interne du robot, lequel aura statistiquement plus de chances de percevoir leur valeur sur un ensemble de registres contigus plutôt que sur une case isolée au milieu de dizaines d'autres registres.

Le programme interne ayant directement accès à ces valeurs par l'intermédiaire des registres d'entrée, il ne lui sera donc pas nécessaire de calculer lui-même des valeurs internes pertinentes s'apparentant à l'évaluation du danger ou du profit.

Nous verrons dans le chapitre suivant si l'utilisation d'heuristiques permet d'améliorer les résultats obtenus.

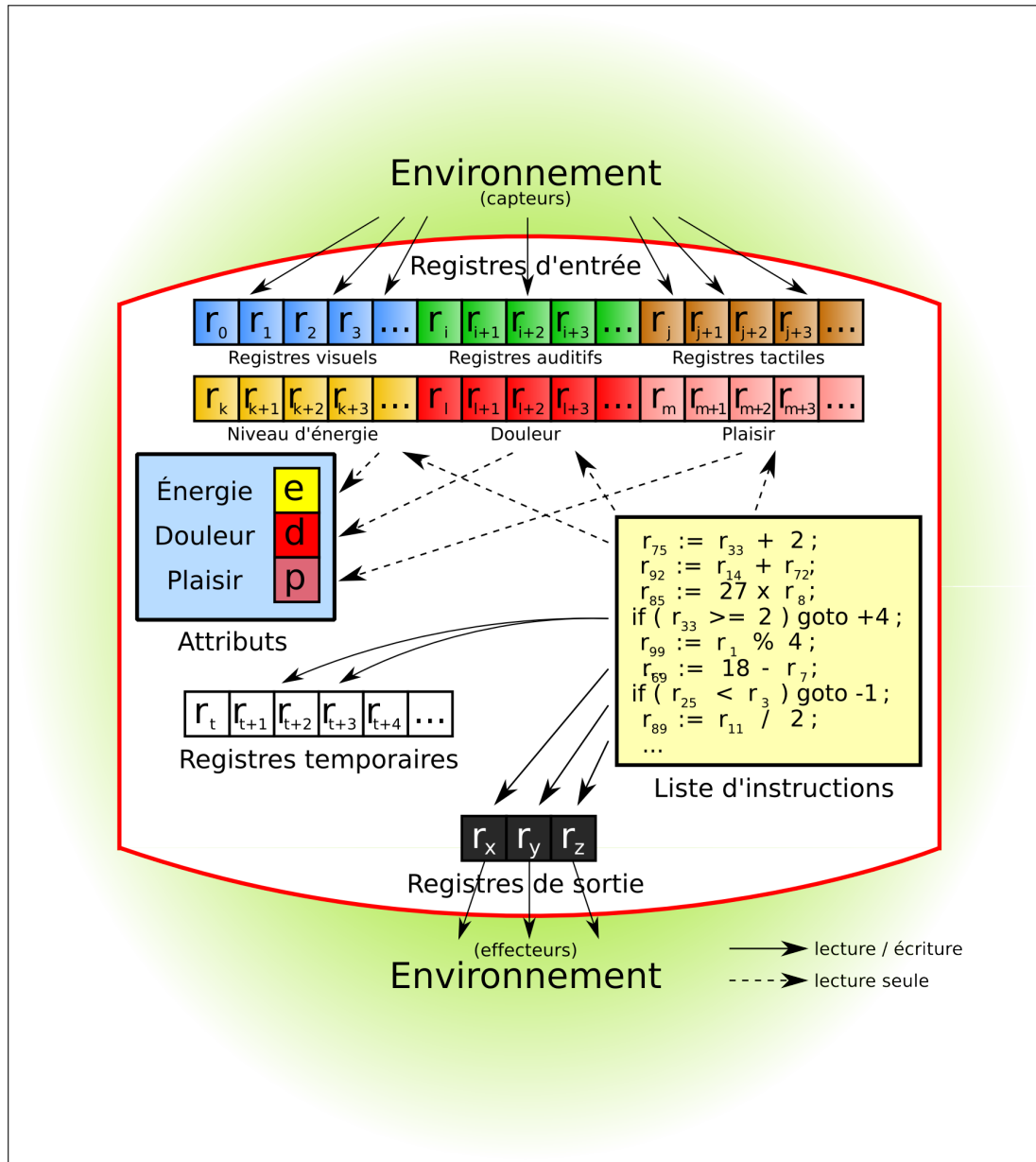


FIG. 12.7 – ORGANISATION GÉNÉRALE DU ROBOT (AVEC HEURISTIQUES)

SOMMAIRE

13.1	INTERFACE GRAPHIQUE	192
13.1.1	Création d'une simulation	192
13.1.2	Chargement d'une simulation	194
13.1.3	Manipulation d'une simulation	194
13.1.4	Environnement 3D	196
13.2	PREMIERS RÉSULTATS	199
13.3	RÔLE DES HEURISTIQUES	203

*La vieillesse nous attache plus de rides en l'esprit
qu'au visage.*

— MONTAIGNE

Ce chapitre a pour fonction de présenter les résultats obtenus à l'issue du développement de GAEA. Dans un premier temps, *l'interface graphique*, qui est la partie visible du moteur génétique, nous permet de suivre l'évolution de la population des individus-programmes. Elle offre en outre la possibilité de visualiser le déroulement d'une simulation proies-prédateurs en trois dimensions, et d'observer le comportement du robot dans son environnement.

Dans un second temps, nous verrons si l'exécution du moteur génétique permet d'obtenir des individus mieux adaptés, grâce à l'analyse des résultats de l'algorithme génétique sur plusieurs générations en suivant l'évolution globale de la *fitness* et de la longueur des programmes.

Enfin, nous comparerons ces résultats avec une version de GAEA utilisant les heuristiques décrites au chapitre précédent.

13.1 INTERFACE GRAPHIQUE

13.1.1 Création d'une simulation

Parallèlement au moteur génétique, l'application développée se compose d'une interface graphique permettant de créer une simulation et de suivre son évolution. La fenêtre de création, séparée en onglets, se présente comme un tableau de bord récapitulant les principaux paramètres que l'on peut manipuler, tels que les réglages associés au moteur génétique (voir figure 13.1), ou encore les valeurs de paramètres des programmes génétiques (voir figure 13.2).

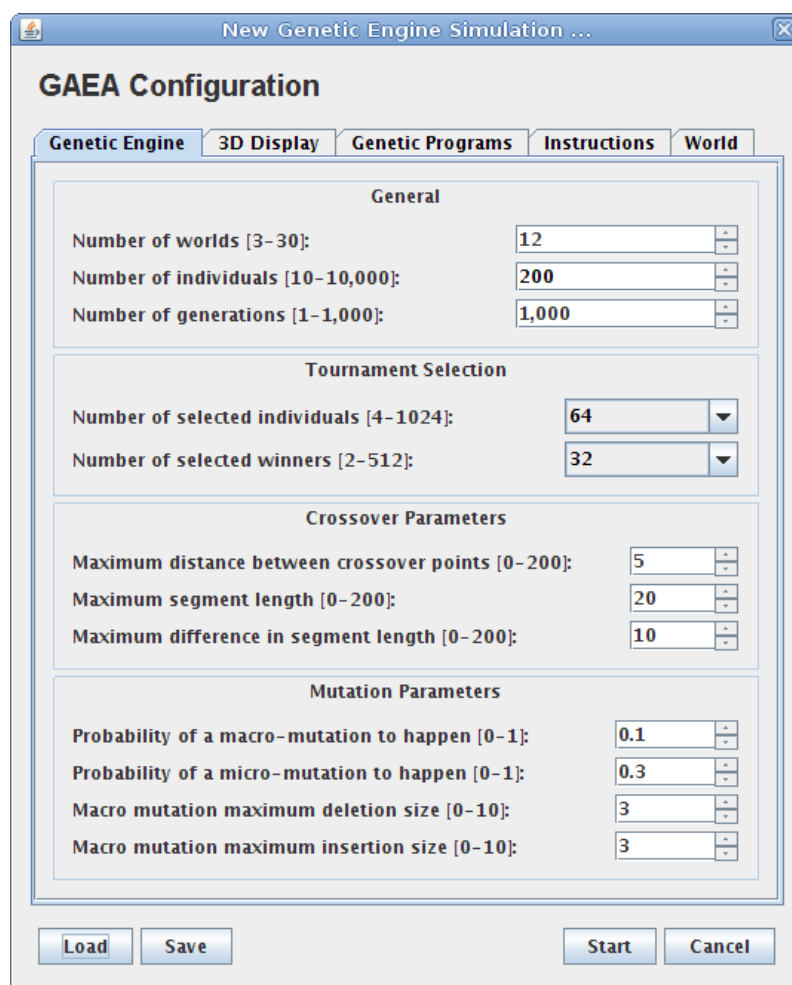


FIG. 13.1 – FENÊTRE DE CRÉATION D'UNE NOUVELLE SIMULATION – PARAMÈTRES DU MOTEUR GÉNÉTIQUE

La fenêtre de droite, correspondant aux paramètres du moteur génétique, permet de modifier les valeurs liées à la simulation telles que le nombre de mondes générés à chaque étape, ou le nombre de proies et de prédateurs dans

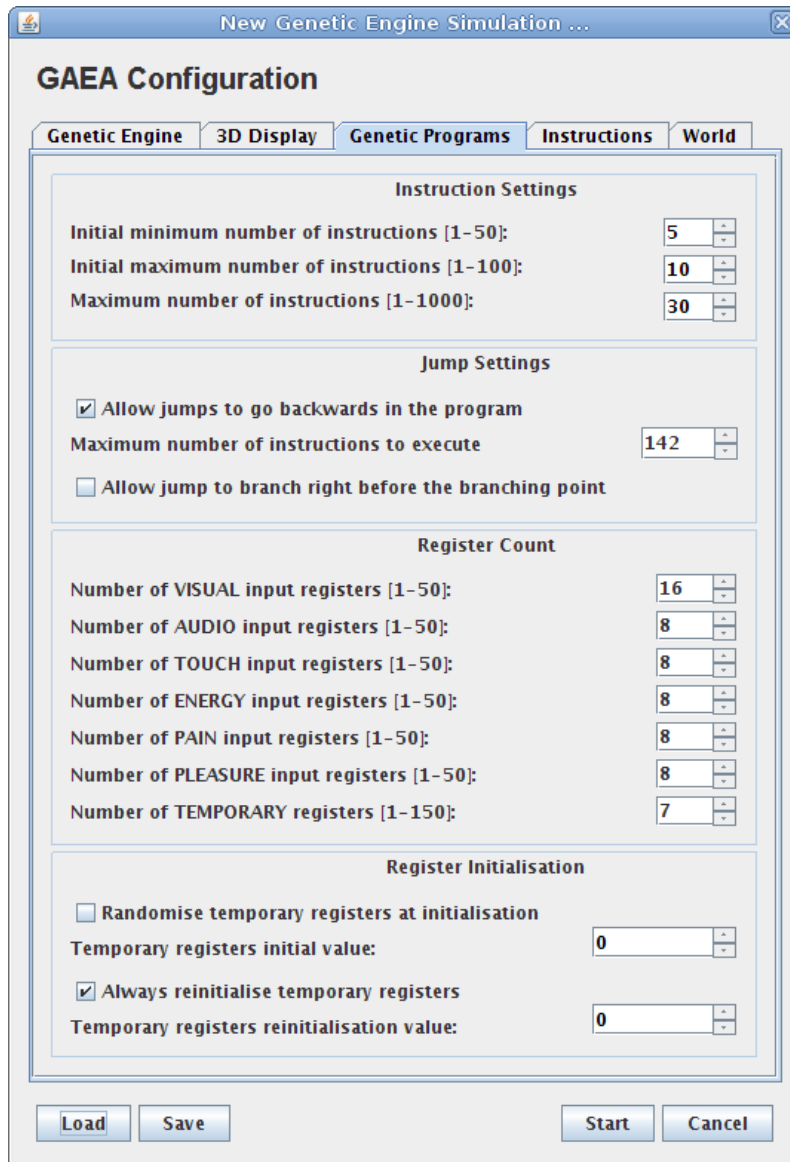


FIG. 13.2 – FENÊTRE DE CRÉATION D'UNE NOUVELLE SIMULATION – PARAMÈTRES DES PROGRAMMES GÉNÉTIQUES

chaque monde, ou encore les réglages associés aux phases de tournoi, de croisement et de mutation.

La fenêtre de gauche décrit les paramètres des individus-programmes. En plus de pouvoir modifier les réglages concernant les instructions et les sauts conditionnels, il est possible de déterminer précisément le nombre de registres de chaque sorte, ainsi que la manière dont ils sont initialisés.

En outre, les deux boutons situés en bas de la fenêtre permettent de charger et de sauvegarder l'ensemble des paramètres à l'aide de fichiers de configuration, ce qui s'avèrera très pratique étant donné le nombre de valeurs de paramètres à manipuler — plus d'une cinquantaine tout de même !

13.1.2 Chargement d'une simulation

À l'issue du paramétrage de la simulation, le moteur génétique s'occupe de créer un dossier spécifique contenant l'ensemble des mondes et programmes générés ainsi que plusieurs fichiers de journalisation. Le dossier créé est par conséquent responsable du stockage de l'état du système et peut donc être chargé ultérieurement.

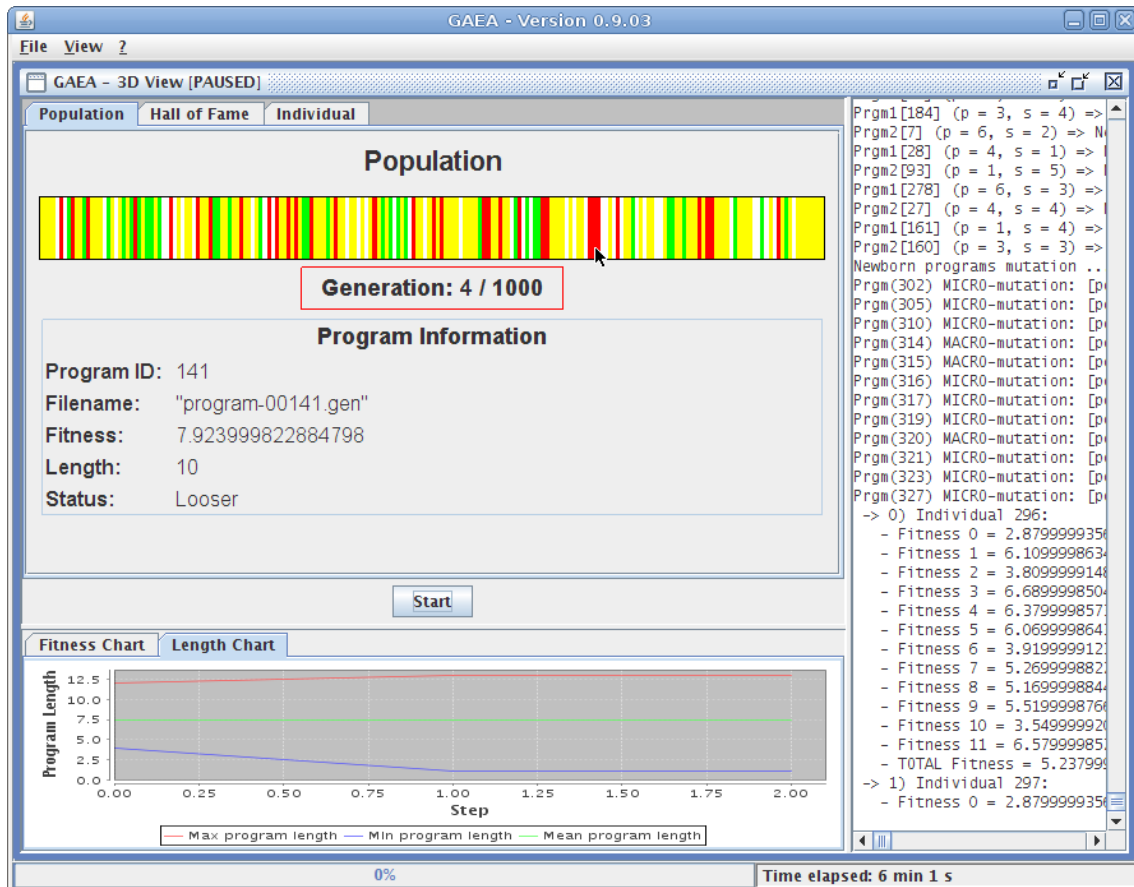
Il s'agit en tout cas d'une des possibilités qu'offre GAEA en permettant, en spécifiant le nom du répertoire, de charger une simulation au niveau de la génération dans laquelle le programme s'était arrêté précédemment.

13.1.3 Manipulation d'une simulation

Une fois la simulation créée ou chargée, la fenêtre principale de GAEA permet l'affichage de l'état général de la population et des différents programmes la composant (voir figure 13.3). La fenêtre se compose d'une zone de texte à droite décrivant les actions effectuées par le moteur génétique. En bas se trouve un graphique représentant l'évolution des *fitnesses* et longueurs de programmes en fonction des générations qui se succèdent. Enfin, dans la zone principale en haut à gauche se trouve l'affichage des différentes informations issues de la simulation.

L'onglet *Population* (figure 13.3) décrit l'état des programmes de la génération courante à l'aide d'un code de couleur lié au statut du programme : **jaune** si celui-ci a déjà fait l'objet d'une sélection, **vert** s'il vient tout juste de remporter un tournoi, **rouge** lorsqu'il vient de perdre un tournoi et **blanc** s'il n'a jamais été sélectionné. En outre sont indiqués le numéro d'identifiant, le nom, la *fitness* et la longueur du programme surligné par le pointeur de la souris.

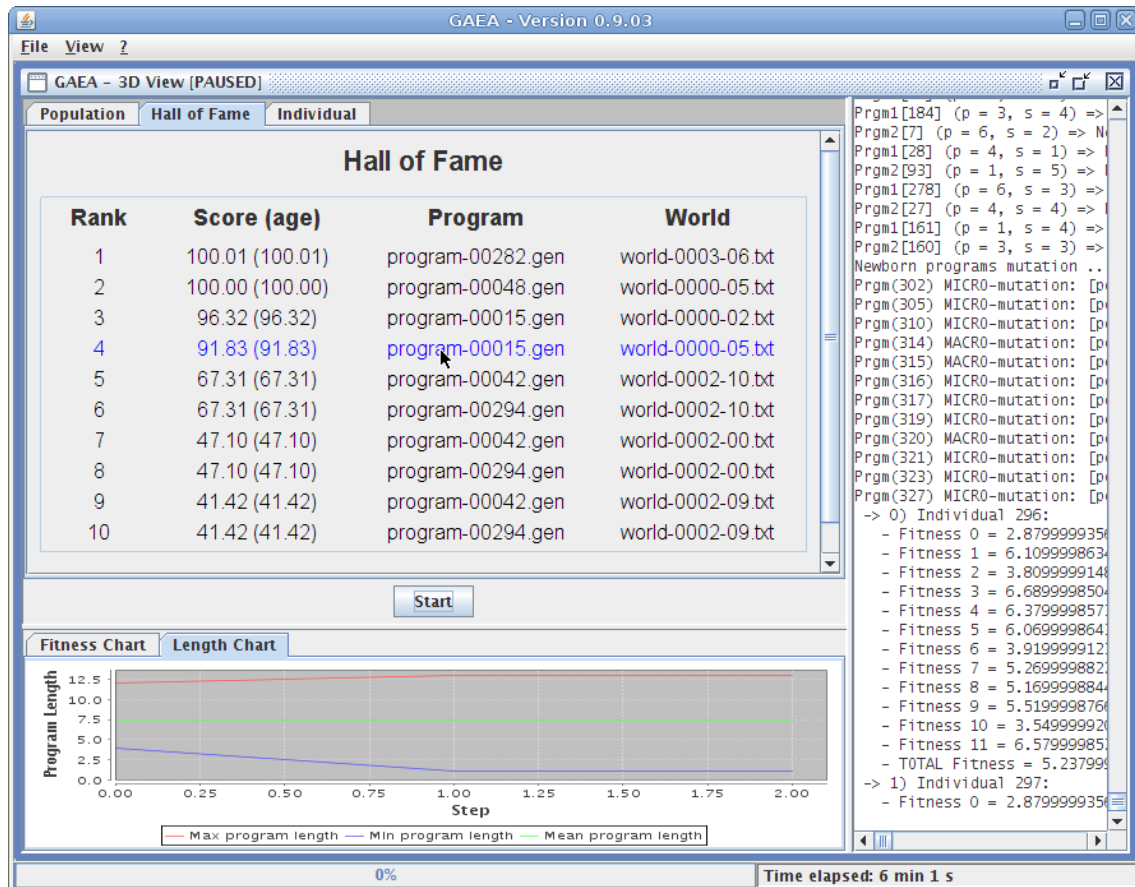
L'onglet intitulé *Hall of Fame* (figure 13.4), que l'on peut traduire par « panthéon », permet d'afficher les meilleurs programmes rencontrés jusqu'à présent, toutes générations confondues. En effet, les dix individus ayant obtenu les scores

FIG. 13.3 – FENÊTRE PRINCIPALE DE GAEA — ONGLET *Population*

les plus élevés auront le privilège d'apparaître dans cette page. Chaque ligne indiquant le rang, le score, ainsi que les noms du programme et du monde simulé, a la particularité de pouvoir être sélectionnée à l'aide de la souris. Un environnement 3D s'ouvre alors, permettant de visualiser le robot en pleine action, au milieu de proies et de prédateurs dont la position est initialisée grâce au fichier du monde. Nous décrirons cet environnement plus en profondeur ci-après.

Le troisième onglet, intitulé *Individual* (pour « individu »), résume la composition du robot en termes de registres d'entrée-sortie — liés aux capteurs et effecteurs — et de registres internes (figure 13.5), à l'aide d'un code de couleur :

- **bleu** pour les registres visuels ;
- **vert** pour les registres auditifs ;
- **marron** pour les registres tactiles ;
- **jaune** pour les registres associés à l'énergie ;
- **rouge** pour les registres associés à la douleur ;
- **rose** pour les registres associés au plaisir ;

FIG. 13.4 – FENÊTRE PRINCIPALE DE GAEA — ONGLET *Hall of Fame*

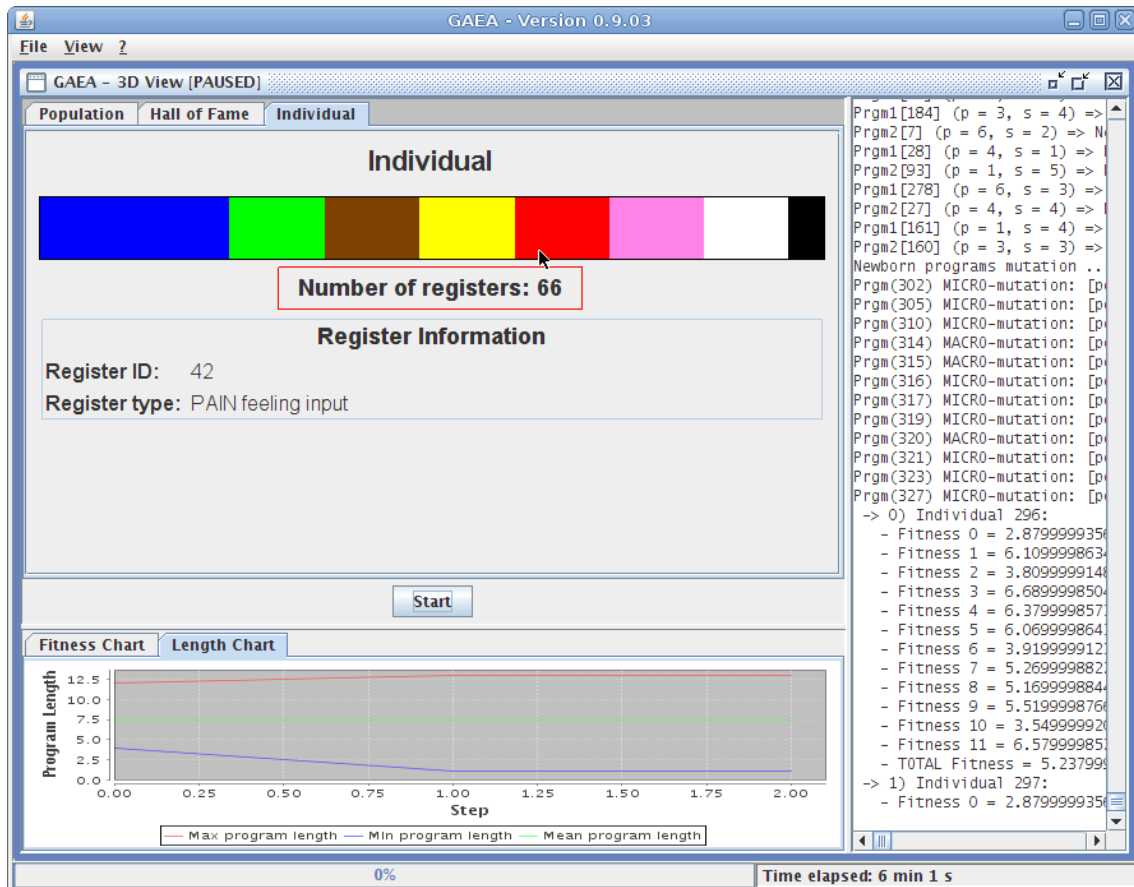
- **blanc** pour les registres temporaires ;
- **noir** pour les registres de sortie.

De la même manière que pour l'onglet *Population*, les informations relatives au registre surligné par le pointeur de la souris apparaissent.

13.1.4 Environnement 3D

Les fenêtres décrites jusqu'ici représentent les informations générales relatives au moteur génétique et à la population des individus-programmes. Il est en outre possible de visualiser *a posteriori* la simulation proies-prédateurs de chaque individu, et ce dans chaque monde auquel il a été confronté. Cette visualisation peut s'effectuer de deux manières. Premièrement, l'onglet *Hall of Fame* expliqué précédemment permet de sélectionner une simulation parmi les meilleurs scores obtenus et de charger conséquemment la visualisation 3D correspondante.

Une deuxième possibilité consiste à ouvrir une fenêtre spécifique de char-

FIG. 13.5 – FENÊTRE PRINCIPALE DE GAEA — ONGLET *Individual*

gement d'une simulation (figure 13.6), permettant de charger le programme, le monde, ainsi que les paramètres de la simulation associée.

Un programme se présente sous la forme d'un fichier d'extension .gen dont voici un exemple :

```

----- Instructions -----
#NBINSTR=9
-----
R57 := -30 + 14
R63 := R33 DIV R39
if (R58 <= -45 ) goto +2
R57 := R33 - R53
R58 := R5 DIV 5
R64 := R38 DIV -18
if (R28 <= -19 ) goto -1
if (R29 <= R12 ) goto +0
R64 := R21 + R9
-----

```

Il se compose naturellement des instructions du programme, de type affectation.

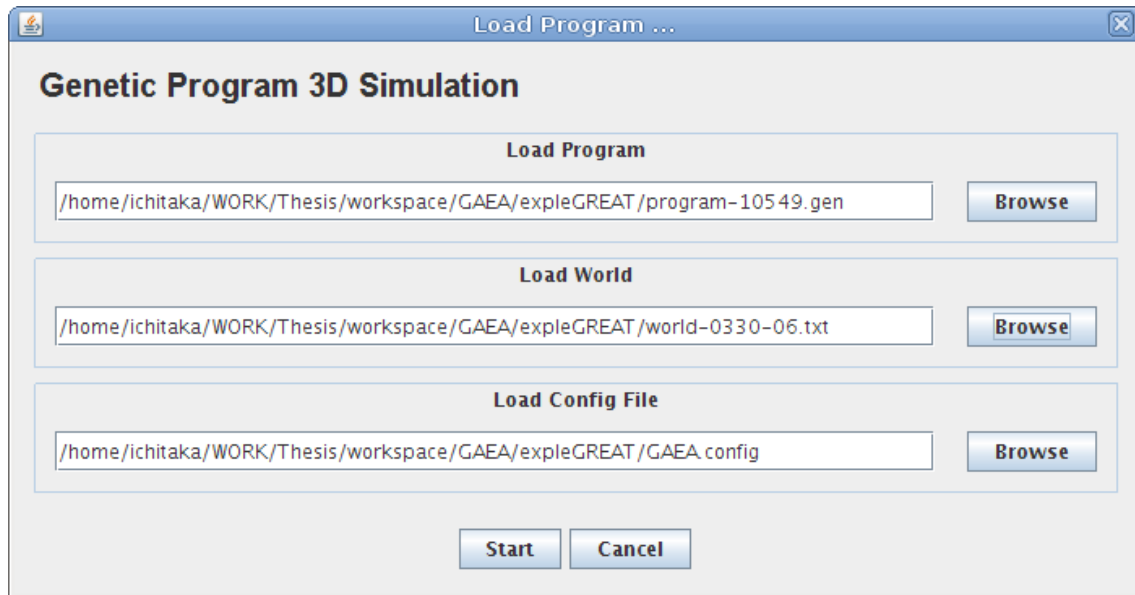


FIG. 13.6 – FENÊTRE DE CHARGEMENT D’UN ENVIRONNEMENT 3D

tation ou condition, précédées du nombre de ces instructions afin de faciliter la lecture du fichier.

En parallèle, un fichier constitué du monde modélisé doit être chargé par la fenêtre, et ce afin de connaître la position des proies et prédateurs au sein de l’environnement. Enfin, les paramètres sont lus dans un troisième fichier, sous la forme PARAMÈTRE=vaieur, qui contient un ensemble de variables régissant le fonctionnement du moteur génétique et de la simulation proies-prédateurs.

La fenêtre de visualisation (figure 13.7) se compose au centre d’un environnement 3D, réalisé en JOGL ([111]), représentant le robot (en rouge sur l’image), les proies (en jaune) et les prédateurs (en noir), dans une zone carrée limitée par quatre murs (symbolisés par une ligne rosée). En haut à droite se trouve une zone représentant l’état des capteurs du robot, déjà décrite précédemment (voir figure 12.1). Juste en dessous, trois jauges indiquant les niveaux d’énergie (en jaune), de douleur (en rouge) et de plaisir (en rose).

En bas à droite se trouve une carte bidimensionnelle de l’environnement et des positions des créatures. Il faut noter que les zones de détection visuelle et auditive sont indiquées par un halo lumineux respectivement orange et bleu. La zone de détection tactile n’est pas représentée car elle correspond au corps du robot.

Enfin, un graphique décrivant les valeurs des registres du robot est dessiné tout en bas de la fenêtre, en suivant le même code couleur que l’onglet *Individual* de la section précédente. Il est assorti d’une zone de texte indiquant les informations du registre localisé à l’aide de la souris, à savoir son identifiant, son type et sa valeur.

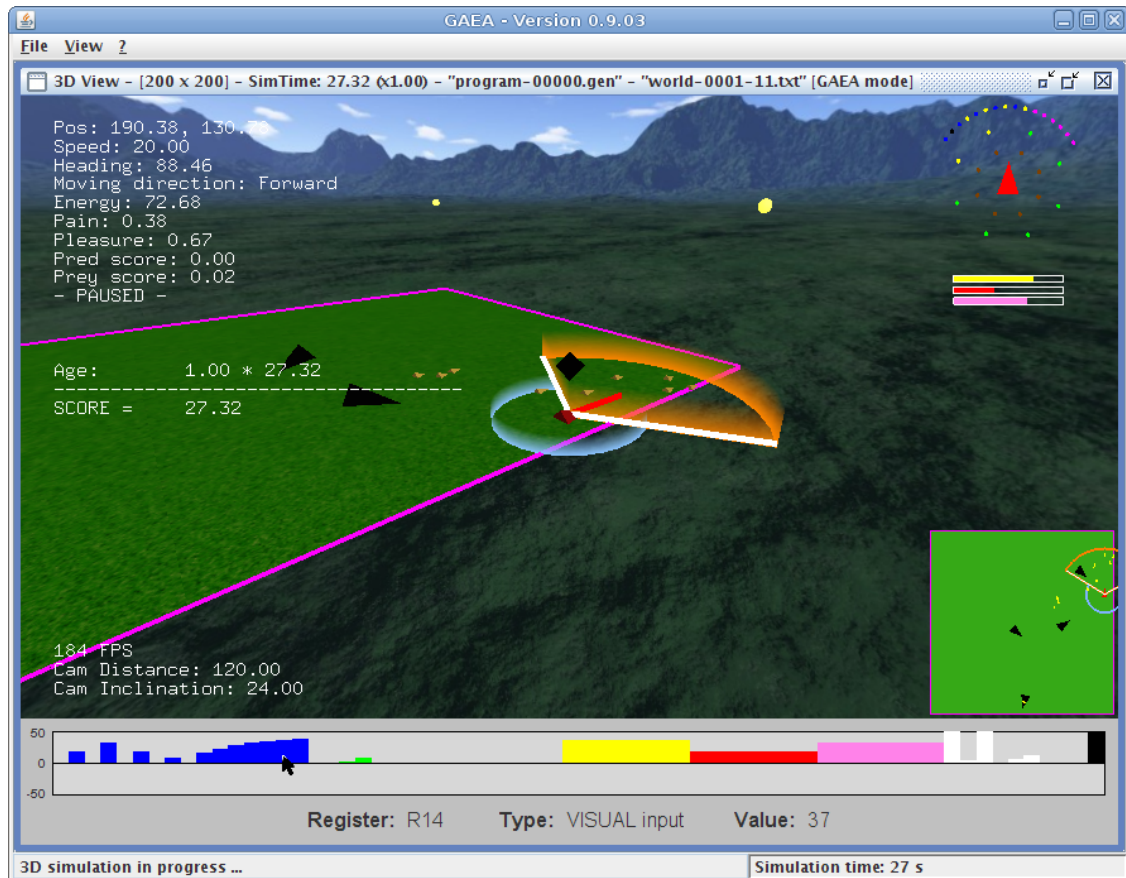


FIG. 13.7 – ENVIRONNEMENT DE SIMULATION PROIES-PRÉDATEURS EN 3D

Bien entendu, l'ensemble des paramètres de la visualisation est paramétrable, et il suffit d'utiliser certains raccourcis clavier pour faire apparaître ou disparaître certaines des informations représentées. De même, la souris permet de changer l'angle et la distance de la caméra. Enfin, l'affichage des textures est optionnel et peut être modifié.

Après avoir décrit dans les grandes lignes l'application et l'interface réalisées, il nous faut parler des résultats obtenus grâce à elles. C'est l'objet de la section suivante.

13.2 PREMIERS RÉSULTATS

Une des particularités des algorithmes de programmation génétique linéaire est la difficulté à obtenir une convergence rapide de la population de programmes, et ce pour une raison principale : la propension des opérateurs génétiques à détruire facilement les meilleures solutions, du fait du lien extrêmement fort qu'entretiennent les instructions d'un programme en termes de signification

d'un ensemble de lignes de code. En effet, un groupe d'instructions ne fait sens que s'il est exécuté dans son ensemble, et les éventuels croisements et mutations ont alors un rôle primordial à jouer dans la conservation ou la destruction de la pertinence du code. De la même façon, l'ordre des instructions est capital et permet de préserver une certaine logique dans la modification des variables du programme. En somme, la potentielle importance de la proximité et de l'ordre des instructions dénote les nombreuses interactions existant entre les lignes de code d'un programme, révélant une certaine fragilité.

De plus, au manque de stabilité effective du code lors de ses nombreuses modifications liées aux croisements et autres mutations prévus dans l'algorithme doivent s'ajouter les erreurs apparaissant à l'issue de ces transformations, qui impliquent souvent des corrections et occasionnent par conséquent de profonds changements comportementaux du programme. Ainsi par exemple, les branchements conditionnels, dont la valeur de saut se révélerait incorrecte à l'issue d'un croisement génétique, se verront attribuer une nouvelle valeur.

Fort heureusement, malgré ces limitations, la redondance de code ainsi qu'un paramétrage judicieux du moteur génétique permettent d'obtenir une convergence dans la population de programmes. Les tests effectués sur GA EA ont été réalisés à l'aide des paramètres suivants :

Paramètre	Valeur
Nombre de mondes à générer à chaque génération :	12
Nombre d'individus dans la population :	200
Nombre de générations testées :	500
Nombre d'individus sélectionnés :	64
Nombre d'individus vainqueurs du tournoi :	32
Taux de macro-mutations :	10%
Taux de micro-mutations :	30%
Nombre maximal d'instructions dans un programme :	30

Par ailleurs, les sauts conditionnels négatifs sont autorisés et les branchements résultants se font systématiquement après les conditions-objectifs. Voici maintenant les réglages du robot et de l'environnement :

Paramètre	Valeur
Nombre de registres visuels :	16
Nombre de registres auditifs :	8
Nombre de registres tactiles :	8
Nombre de registres temporaires :	7
Intervalle de valeurs de registres :	$[-50; 50]$
Nombre de prédateurs :	4
Nombre de proies :	10

Les courbes de la figure 13.8 ont été obtenues par la mesure des *fitness* maximale, moyenne et minimale de chaque programme à toutes les générations. Il faut rappeler au passage que la *fitness* d'un programme correspond globalement au score moyen obtenu dans chaque monde, et que le critère d'arrêt du moteur génétique est l'accès à la dernière génération — la 500^{ème} en l'occurrence.

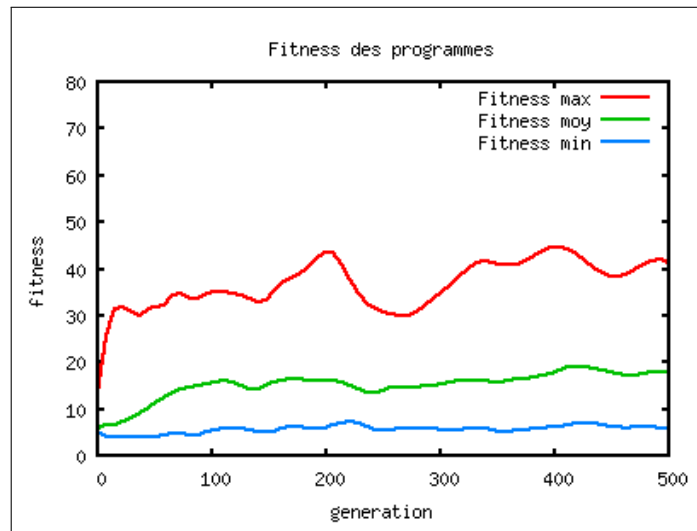


FIG. 13.8 – ÉVOLUTION DE LA *fitness* MAXIMALE, MOYENNE ET MINIMALE AU SEIN DE LA POPULATION EN FONCTION DE LA GÉNÉRATION

Sur le graphique, on constate tout d'abord une certaine instabilité générale quant aux résultats obtenus, notamment en ce qui concerne la courbe représentative de la *fitness* maximale, celle-ci étant amenée à changer régulièrement à cause du renouvellement assez rapide de la population. Après un démarrage très favorable sur les 20 premières générations, la courbe tend à stagner, avec des *fitness* entre 30 et 40, puis aux alentours de 40 en fin de simulation. La *fitness* moyenne, elle aussi, démarre bien, avec une croissance correcte durant les 100 premières générations. Puis, elle stagne aux environs de 15, avec une valeur maximale de 20 vers la 420^{ème} génération. Quant à la courbe décrivant la *fitness* minimale, elle reste aux alentours de 5 du début à la fin de la simulation.

Le second graphique (figure 13.9) mesure l'évolution de la longueur maximale, moyenne et minimale des programmes de la population. On découvre alors que toutes les courbes croissent relativement vite, la longueur maximale (30 instructions) étant atteinte à la 100^{ème} génération, la courbe des longueurs minimales finissant entre les valeurs 15 et 20 en fin de simulation, et la longueur moyenne se terminant aux alentours de 27 instructions.

La croissance relativement rapide de la longueur des programmes est révélatrice de la nécessité qu'ont ces programmes de développer le nombre de leurs instructions afin d'augmenter leur efficacité potentielle. En effet, au dé-

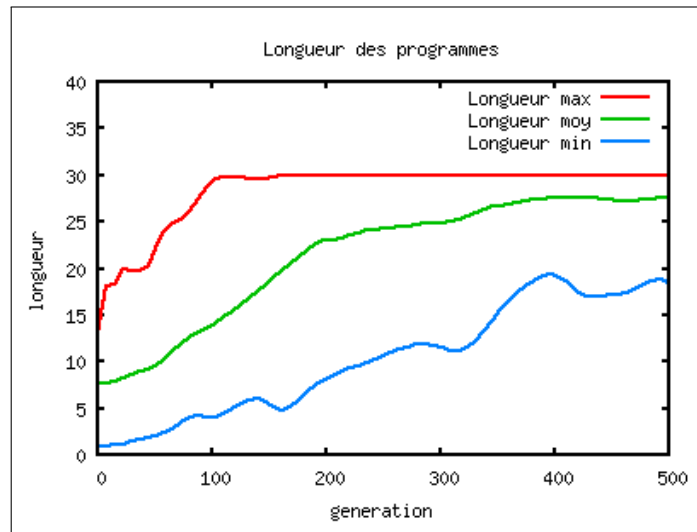


FIG. 13.9 – ÉVOLUTION DE LA LONGUEUR DE PROGRAMME MAXIMALE, MOYENNE ET MINIMALE AU SEIN DE LA POPULATION EN FONCTION DE LA GÉNÉRATION

but de la simulation, l'initialisation de la population consiste en la création de programmes d'une taille située entre 5 et 10. Bien entendu, il est impossible d'obtenir de bons résultats avec si peu d'instructions. Cependant, ce mécanisme permet une meilleure croissance initiale de la population.

De ces deux graphiques, on peut conclure que la convergence des *fitnesses* vers de meilleures valeurs se révèle bien timide. En effet, malgré le nombre de tests effectués sur différents paramètres, il est assez difficile d'obtenir une évolution importante de la population, la faute sans doute à un manque de lien clair et direct entre d'un côté les informations détectées par le robot au moyen de ses capteurs et de ses registres d'entrée, et d'un autre côté, une décision efficace devant être construite en accord avec la situation. Dans la théorie de l'évolution, ce lien est généralement le fruit d'un apprentissage long et sélectif, qui mène habituellement à la distinction primitive entre le bien et le mal, entre ce qui est avantageux pour l'individu et ce qui représente un danger potentiel.

C'est en tout cas ce que prônent de nombreuses théories de l'émotion, dont celle de LAZARUS (section 3.4.4) et d'ORTONY, CLORE et COLLINS (section 3.3.2), lesquels utilisent une dualité entre émotions positives et négatives qui peuvent favoriser ou retarder la réalisation d'objectifs propres à un individu. Il nous a donc paru nécessaire d'introduire un ensemble d'heuristiques, abordées précédemment, et dont le but sera de définir une opposition entre douleur et plaisir qui guidera plus efficacement les décisions du robot.

13.3 RÔLE DES HEURISTIQUES

Les heuristiques introduites en entrée des programmes permettent de catalyser la convergence du système sans influencer directement le comportement du robot. En effet, bien que ces heuristiques semblent être l'incarnation de réels états émotionnels par des notions telles que le plaisir ou la douleur, en réalité, il ne s'agit que d'informations objectives parfaitement dénuées de sens. En somme, ce ne sont que des valeurs numériques indépendantes de tout contexte, bien qu'elles soient le résultat de calculs bien réels et surtout bien utiles (la douleur, par exemple, est liée au danger représenté par la proximité des prédateurs). Il appartient donc à chaque individu-programme du système d'interpréter ces valeurs d'entrée et d'en faire usage d'une manière optimale.

L'introduction de ces nouveaux registres d'entrée nécessite la définition de quelques paramètres supplémentaires :

Paramètre	Valeur
Nombre de registres pour l'énergie :	4
Nombre de registres pour la douleur :	4
Nombre de registres pour le plaisir :	4

Le graphique de la figure 13.10, à l'instar de la figure 13.8, mesure les *fitness* maximale, moyenne et minimale des individus-programmes de la population pour chaque génération. À première vue, nous obtenons une bien meilleure croissance enregistrée par les trois courbes, ainsi qu'un niveau général beaucoup plus élevé — de l'ordre de 20 points supplémentaires.

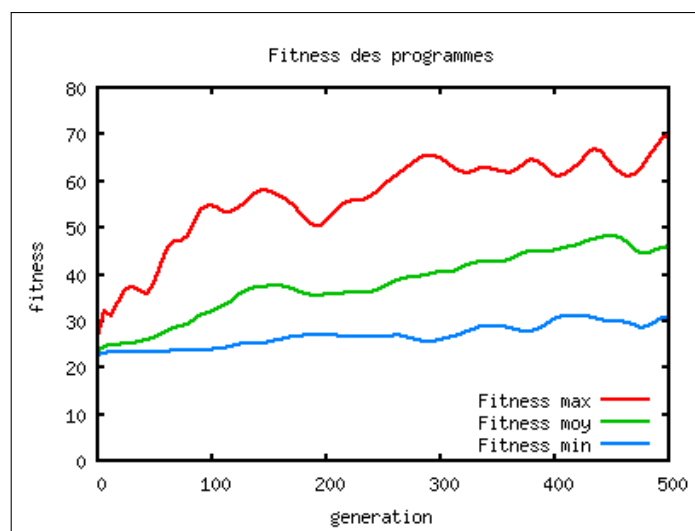


FIG. 13.10 – ÉVOLUTION DE LA *fitness* MAXIMALE, MINIMALE ET MOYENNE AU SEIN DE LA POPULATION EN FONCTION DE LA GÉNÉRATION (AVEC HEURISTIQUES)

En outre, la croissance initiale des courbes se déroule sur un intervalle plus important — environ 150 générations dans cet exemple — et permet une évolution plus saine de la population. Elle se termine sur une *fitness* maximale d'une valeur de 70, soit 66% d'augmentation par rapport à la première version, une *fitness* moyenne finale égale à 46, soit une augmentation de 155%, et une *fitness* minimale qui atteint 31, soit 416% d'augmentation. Avec de tels résultats, on ne peut que constater que la population se porte mieux qu'auparavant.

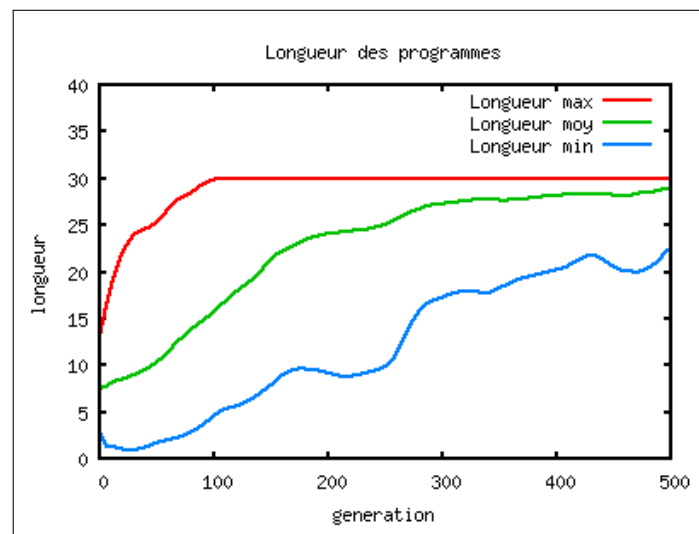


FIG. 13.11 – ÉVOLUTION DE LA LONGUEUR DE PROGRAMME MAXIMALE, MINIMALE ET MOYENNE AU SEIN DE LA POPULATION EN FONCTION DE LA GÉNÉRATION (AVEC HEURISTIQUES)

L'évolution de la longueur des programmes (figure 13.11) est analogue à la précédente pour une raison simple : les paramètres responsables des opérations de croisement et de mutation sont identiques. Toutefois, on peut constater que l'utilisation des heuristiques accentue sensiblement la croissance des programmes. Bien que la longueur maximale plafonne à 30, qui est le nombre d'instructions maximal autorisé, la longueur moyenne finit sa course avec une valeur de 28, soit 4% d'augmentation, et la longueur minimale termine à 22 instructions, soit 22% d'augmentation. Cette accentuation confirme l'hypothèse selon laquelle les meilleurs programmes sont statistiquement ceux composés du plus grand nombre d'instructions. Les heuristiques ne font que faciliter l'émergence des programmes les plus efficaces, et donc les plus longs.

SOMMAIRE

14.1 CONCLUSION	205
14.2 PERSPECTIVES	206

One day at a time — this is enough. Do not look back and grieve over the past, for it is gone, and do not be troubled about the future, for it has not yet come. Live in the present, and make it so beautiful that it will be worth remembering.^a

— Ida Scott TAYLOR

a. « Chaque jour en son temps — c’est suffisant. Ne regardez pas en arrière et ne pleurez pas sur le passé, car il est passé, et ne vous en faites pas pour le futur, car il n’est pas encore arrivé. Vivez dans le présent et faites-en quelque chose de si magnifique qu’il vaudra la peine que l’on s’en souviennne. » (NDR)

14.1 CONCLUSION

Les deux versions du modèle GAEA exposées auparavant décrivent l’évolution d’une population de programmes de plus en plus adaptés au problème posé. Les meilleurs d’entre eux ont alors pu développer des stratégies réactives qui contraignent l’exploration du domaine de recherche, phénomène que l’on considère comme le résultat de réactions émotionnelles progressivement assimilées. L’émotion « émerge » du programme et n’est donc pas le résultat d’un calcul élaboré au préalable.

Par ailleurs, les heuristiques introduites dans un second temps permettent la redondance ainsi que l'ajout d'informations numériques quantitatives sur la situation dans laquelle se trouve le robot. Cependant, il ne s'agit aucunement de données qualitatives car il n'est nullement décrit le rôle que peuvent jouer la douleur, le plaisir, ou même l'énergie sur le comportement de l'individu. Ainsi, seule l'émergence comportementale est à l'origine de l'étiquetage cognitif des différents attributs.

14.2 PERSPECTIVES

Les résultats de GAEA sont prometteurs mais restent encore quelque peu timides quant aux capacités réelles apportées par les algorithmes de programmation génétique. Quelques améliorations peuvent néanmoins être apportées afin d'accroître le potentiel de ce modèle.

Tout d'abord, concernant le score, il serait parfaitement envisageable d'en redéfinir le calcul par l'introduction de nouveaux paramètres. En effet, pour le moment, le score considère uniquement l'âge du robot en fin de simulation. Il s'agit dans tous les cas de la mesure la plus logique compte tenu de l'aspect évolutionniste du modèle. En réalité, la difficulté majeure du calcul du score d'un individu est de déterminer, hormis l'âge, quelles peuvent être ses composantes et surtout quelles sont les valeurs de pondération de chacune de ces composantes. Quelques tests ont déjà été réalisés, lesquels intègrent au calcul la position du robot par rapport aux autres créatures, ainsi que la capacité de maintien d'un haut niveau d'énergie. Malheureusement, les résultats ne sont pour le moment pas concluants. Bien qu'utiliser des informations provenant de l'environnement puisse paraître judicieux, le calcul d'un score ne doit pas dénaturer l'esprit auto-adaptatif du modèle en imposant un cadre évolutif de manière trop directive.

Cette difficulté met en avant le problème plus général du réglage des paramètres. En effet, chaque variable utilisée — par exemple le nombre d'individus, la taille maximale des programmes ou le pourcentage de mutations — a un impact particulier, plus ou moins important, sur l'évolution de la population. Il serait tout à fait envisageable d'effectuer des plans d'expérience détaillés en modifiant les paramètres successivement afin d'étudier leur éventuelle influence. Néanmoins, le nombre important de paramètres, ainsi que les conséquences imprévisibles et complexes inhérentes à leur modification, rendent le problème très difficile à analyser. Une des pistes à emprunter serait alors la capacité de caractériser la population à l'aide de valeurs sémantiquement plus riches que de simples *fitnesses*, afin de déterminer non plus les besoins propres à chaque individu mais les aspirations de la population entière. Ainsi, il serait possible de calculer dynamiquement l'évolution des paramètres pendant la simulation, et permettre au système dans son ensemble de s'adapter aux aléas naturels pour

pouvoir converger sans se heurter à une rigidité numérique qui risquerait justement de freiner sa convergence.

De la même manière que l'évolution des programmes peut être endiguée par un réglage de paramètres trop strict, l'environnement peut lui aussi se révéler trop limité pour permettre à une population de pouvoir s'exprimer pleinement. L'introduction de nouvelles contraintes spatiales, par l'ajout d'obstacles naturels, par la création d'un relief varié, ou par l'introduction de créatures, de matériaux ou d'objets divers par exemple, permettrait de bouleverser les repères comportementaux du robot en lui proposant un environnement plus complet, certes plus difficile à appréhender, mais plus riche en possibilités.

Enfin, une des hypothèses à plus long terme qu'il serait particulièrement intéressant d'évoquer est la création d'un procédé permettant de faire évoluer des programmes durant leur exécution, aussi bien structurellement qu'au niveau de l'organisation de leurs registres. L'idée est de simuler une dynamique structurelle comparable aux processus ontogénétiques et à l'expérience d'un individu de manière générale. Pour ce faire, une approche de type *linear-tree GP* (KANTSCHIK [60]) serait réalisable, mais elle implique un ensemble de difficultés supplémentaires en termes de manipulation de programmes telles que la gestion des croisements, les corrections d'erreurs, ou tout simplement la mise en place des règles de dynamique évolutive au sein d'un tel modèle.

CONCLUSION

For those who believe, no proof is necessary. For those who don't believe, no proof is possible.^a

— Stuart CHASE

a. « Pour ceux qui croient, aucune preuve n'est nécessaire. Pour ceux qui ne croient pas, aucune preuve n'est possible. »

Ce travail de thèse a permis le développement de deux modèles de l'émotion que sont le problème des Cascades et GAEA, illustrant à eux deux des points de vue radicalement différents, et mettant en œuvre des mécanismes que l'on pourrait qualifier d'opposés, le premier modèle décrivant concrètement l'émotion numériquement et le second construisant une émergence émotionnelle *ex nihilo*.

À l'origine, le problème des Cascades se veut cognitif, construisant un ensemble de micro-stratégies reliées entre elles au sein d'un graphe marqué de phéromones numériques représentant l'émotion. L'objectif étant de remplir une grille de calcul de manière déterministe, le processus responsable de la construction de la stratégie globale fait évoluer le graphe à l'aide d'un algorithme fourni et se stabilise plus ou moins vite vers une solution potentielle. Les résultats démontrent que la dynamique introduite par les fourmis sur les arêtes du graphe par les dépôts répétés de phéromones est non seulement indispensable à la bonne marche de la résolution, mais possède de surcroît une grande faculté de résilience.

Par ailleurs, une plateforme de simulation proies-prédateurs mettant en scène un robot doté de capteurs et effecteurs a été développée. Ce système a initialement prouvé son efficacité lors de son intégration au modèle à base de logique floue de William BLEWITT au sein duquel il permet d'illustrer la pertinence émotionnelle du robot vis-à-vis de la situation dans laquelle il se trouve.

Ce même environnement a ensuite été réutilisé dans le cadre du modèle GAEA, dont l'objectif est de favoriser une construction émergente de l'émotion en milieu hostile, à partir d'un processus de programmation génétique linéaire. L'algorithme principal manipule une population de programmes composés d'une part d'un ensemble d'instructions croisées génétiquement, et d'autre

part de registres mémoire dont certains sont directement reliés aux capteurs et effecteurs du robot. En outre, l'intégration d'heuristiques par l'ajout d'informations supplémentaires de nature émotionnelle en entrée du système permet de catalyser l'évolution de la population vers une meilleure adaptation à l'environnement.

De manière générale, étudier l'émotion sous plusieurs angles permet de s'en faire une idée plus complète et précise. Ces trois modèles ont confirmé, chacun à leur manière, ce que les études neurologiques ont démontré il y a quelques années, à savoir que l'émotion est une aide à la décision, en ce sens qu'elle forme un compromis avantageux entre efficacité et rapidité. Les émotions conduisent donc à une meilleure adaptabilité dans le cadre de la résolution de problèmes.

Mais aussi et surtout, les résultats de ce travail de thèse permettent de nous conforter dans l'idée que l'émotion représente désormais un champ d'investigation utile à la création de modèles artificiels intelligents. En conséquence, je souhaiterais profiter de cette conclusion pour répondre à celles et ceux qui considèrent qu'il est impossible d'implanter l'émotion au sein d'une machine, car son apparente désincarnation ainsi que les mystères qui l'entourent semblent propres à l'être humain, ou plus généralement, à l'être vivant.

En réalité, je pense qu'il n'en est rien. La complexité caractérisant les processus émotionnels est certes extrêmement difficile à appréhender, mais cet imbroglio de circuits formidablement nombreux et interconnectés qu'est le cerveau reste finalement une entité mécanique. Bien sûr, ses différents niveaux d'organisation ainsi que sa dynamique particulièrement développée, fruit de millions d'années d'évolution, font du cerveau une machine encore majoritairement indéchiffrable de nos jours. Cependant, il ne tient qu'à la communauté scientifique de comprendre les mécanismes sous-jacents à la construction cérébrale, aussi compliqués soient-ils, et tenter d'en appliquer les règles au sein de modèles comportementaux.

Ainsi, il va sans dire qu'une machine faite de mémoire et d'instructions ne devrait pas nécessairement s'essayer à reproduire une dynamique neuronale en occultant totalement ses caractéristiques propres. En effet, contrairement au cerveau, un ordinateur possède de gigantesques capacités de calcul et de stockage, mais au détriment de facultés de parallélisation plus réduites, surtout dans le cas de robots mobiles. Bien que les progrès de la miniaturisation et l'augmentation de la puissance de calcul des microprocesseurs sauront sans doute un jour vaincre les cent milliards de neurones qui composent notre cerveau, il n'en reste pas moins que cette puissance ne sera pas suffisante.

En effet, les limitations imposées par les approches classiques dans la construction d'agents cognitifs nous informent que de nouvelles perspectives doivent être envisagées dans l'élaboration de modèles de l'intelligence à l'image du fonctionnement humain. L'expérience unique que vit chaque individu au cours de sa vie crée des conformations cognitivo-émotionnelles si différentes que

deux organisations cérébrales, hormis leurs structures fondamentales, programmées génétiquement, et qui font que deux cerveaux sont composés des mêmes cellules et circuits de base, ne peuvent se construire à l'identique. Chacun développe ainsi sa propre sensibilité et ses propres connaissances du monde selon son parcours spécifique. Bien que ces différences interindividuelles ne posent aucun problème — la communication entre deux individus par exemple est tout à fait possible —, l'analyse scientifique des régularités de fonctionnement cérébral, en revanche, constitue un véritable défi, tant les modes d'action et de représentation sont variables d'une personne à une autre. Là encore, il faut rendre compte de cette particularité dans nos modèles, car la solution se trouve peut-être dans la compréhension de cette mécanique.

Tout compte fait, je pense que la conception de l'émotion qu'a façonnée la majeure partie de la communauté scientifique cristallise à elle seule les difficultés qu'ont les théoriciens à comprendre les phénomènes complexes tels que les processus cérébraux. En tout état de cause, mon postulat est de dire que l'émotion n'est pas une entité présente dans le système mais fait simplement l'objet d'une émergence comportementale issue de l'évolution non seulement phylogénétique mais aussi et surtout ontogénétique du réseau neuronal. Ce mécanisme a permis le développement d'une forme de comportement mêlant de concert les aspects structurels (les neurones contenant l'information) et les aspects fonctionnels (l'émergence activationnelle résultant de la stimulation du réseau nerveux en entrée qui produit en sortie une réponse comportementale ainsi qu'une rétroaction sur l'ensemble du système). Pour ma part, je pense donc que l'émotion se place à la base de l'organisation cérébrale. Elle est directement issue de l'activité du substrat neuro-anatomique et des motivations du système.

Curieusement, l'intitulé de cette thèse ne mentionne nullement la modélisation de l'émotion mais des processus émotionnels, et cela change tout ! Premièrement, il faut bien comprendre que considérer l'émotion en tant que mécanisme clairement délimité, voire isolé, n'a absolument aucun intérêt. En effet, seuls comptent les impacts de cette émotion sur la prise de décision, et cela nous mène à la deuxième chose importante à retenir. Le but principal de ce genre de modèle étant « décision-centré », il ne faut chercher aucun fonctionnement émotionnel standard qui serait présent chez tout individu, même si, à l'instar des chemins de la peur ou du plaisir, il existe des similitudes génétiques communes à une même espèce animale impactant directement l'activité émotionnelle. Ces similitudes me semblent en grande partie structurelles et non pas fonctionnelles.

Finalement, l'émotion doit être observée non pas comme un ensemble de processus éventuellement isolés mais comme un véritable système complexe, déterminé à plusieurs échelles d'organisation, et en constante auto-adaptation ontogénétique.

De façon quelque peu ambitieuse, mon objectif, dans un avenir très lointain (peut-être trop), serait de concevoir un modèle de l'émotion de type hié-

chique, qui puisse tenir compte des processus de différenciation émotionnelle, en partant d'émotions dites fondamentales et allant jusqu'aux émotions composites plus complexes, telles que les décrivent les théories évolutionnistes de l'émotion. Étant donné le peu d'informations que l'on possède, il me paraît peu probable d'arriver un jour à obtenir un fonctionnement émotionnel anthropomorphe réaliste. Toutefois, il ne faut pas oublier que l'on parle d'ordinateurs et d'émotions dites artificielles. Si l'informaticien prend le parti de concevoir une émotionnalité spécifique couplée à la corporéité même d'un robot, rien ne semble empêcher l'émergence d'une forme d'affect différente, et sans doute tout aussi adaptative.

Par ailleurs, si l'on considère que l'émotion se situe en bas de l'échelle des processus cérébraux, pourquoi ne pas envisager les processus cognitifs comme des productions évolutives organisées émergeant des émotions, au même titre que les différentes zones du cortex cérébral, responsables de nos capacités mentales les plus évoluées — telles que la planification ou le langage —, sont des productions émergeant du système limbique, liées à nos fonctions cérébrales les plus basiques — telles que la respiration ou les pulsations cardiaques — et, paraît-il, des émotions fondamentales telles que la soif ou la peur ...

Se pose alors un dilemme inextricable : bien que l'émotion ne puisse se résumer à un ensemble de sous-processus indépendants communiquant entre eux (déclenchement, expression, influence avec la cognition), mais plutôt comme un processus diffus, qui est le fruit d'une activité globale du système, la difficulté reste de choisir entre un modèle très componentiel, facilement contrôlable mais assez rigide, ou bien un système plus dynamique et adaptatif, mais doué d'une activité diffuse et permettant un contrôle par conséquent plus que limité.

À l'instar des systèmes complexes cérébraux qui permettent de générer une activité intellectuelle rapide, efficace, résiliente et auto-adaptative, comment construire des systèmes complexes informatisés qui permettraient de générer une activité artificielle auto-organisée ?

BIBLIOGRAPHIE

- [1] Magda B. Arnold. An exitatory theory of emotion. In M.L. Reymert, editor, *Emotions and personality*. New York Academic Press, New York, 1950.
- [2] Atilf. Trésor de la langue française informatisé. Analyse et traitement informatique de la langue française, Université Nancy II, <http://atilf.atilf.fr/>.
- [3] R.C. Atkinson, D.J. Herrmann, and K.T. Wescourt. Search processes in recognition memory. In R.L. Solso, editor, *Theories in cognitive psychology: The Loyola symposium*. Erlbaum, Hillsdale, NJ, 1974.
- [4] James R. Averill. A constructivist view of emotion. In Robert Plutchik and H. Kellerman, editors, *Emotion, Theory, Research and Experience: Theories of Emotion*, volume 1, pages 305–340. New York: Academic Press, 1980.
- [5] Paolo Baggia, Felix Burkhardt, Catherine Pelachaud, Christian Peter, and Enrico Zovato. Emotion markup language, 2011. <http://www.w3.org/TR/emotionml/>.
- [6] Wolfgang Banzhaf, Markus Brameier, Martin Stautner, and Klaus Weinert. *Advances in Computational Intelligence*, chapter Genetic Programming and its application in Machining Technology, pages 194–241. Springer, Berlin, 2002.
- [7] Wolfgang Banzhaf, Peter Nordin, Robert E. Keller, and Frank D. Francone. *Genetic Programming — An Introduction On the Automatic Evolution of Computer Programs and its Applications*. dpunkt, Heidelberg and Morgan Kaufmann, San Francisco, 1998.
- [8] Philip A. Bard. The neuro-humoral basis of emotional reaction. In C. Murchinson, editor, *Handbook of general experimental psychology*, pages 264–311. Clark University Press, Worcester, MA, 1934.
- [9] Sophie Baudic and Gérard H.E. Duchamp. Emotions: theoretical models and clinical implications. In *European Simulation and Modelling Conference '2006*, Toulouse, France, 2006.
- [10] Antoine Bechara, Hanna Damasio, Daniel Tranel, and Antonio R. Damasio. Deciding advantageously before knowing the advantageous strategy. *Science*, 275:1293–1295, 1997.
- [11] Antoine Bechara, Hannah Damasio, and Antonio R. Damasio. Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, 10:295–307, 2000. <http://cercor.oxfordjournals.org/cgi/content/full/10/3/295>.

- [12] Cyrille Bertelle, Heidi Charvin, Evelyne Clément, Joël Colloc, Véronique Jay, Karim Mahboub, Damien Olivier, Thierry Piot, Pierrick Tranouez, and Bruno Vivicorsi. Développement des dispositifs diagnostics et de remédiation : intégration et modélisation des dimensions cognitives, émotionnelles et psychomotrices dans une tâche de résolution de problème. Rapport final, Convention de soutien de programme CPER 2006, septembre 2008.
- [13] William Blewitt, Aladdin Ayesch, Cyrille Bertelle, and Karim Mahboub. Psychologically grounded emotion model for a npc fuzzy controller. *International Journal of Gaming and Computer-Mediated Simulations '2010*, in progress, 2010.
- [14] Eric Bonabeau, Marco Dorigo, and Guy Theraulaz. *Swarm Intelligence: From Natural to Artificial Systems*. Santa Fe Institute Studies in the Sciences of Complexity. Oxford University Press, New York (NY), 1999.
- [15] Markus Brameier. *On Linear Genetic Programming*. PhD thesis, Université de Dortmund, 2004.
- [16] C. Breazeal. Emotion and sociable humanoid robots. *International Journal of Human-Computer Studies*, 59:119—155, 2003.
- [17] Walter Cannon. The james-lange theory of emotions: A critical examination and an alternative theory. *The American Journal of Psychology*, 39:106—124, 1927.
- [18] Alain Cardon. *Modéliser et concevoir une machine pensante*. Vuibert, Paris, 2004.
- [19] Alain Cardon. Artificial consciousness, artificial emotions, and autonomous robots. *Cognitive Processing*, 7:245—267, 2006. <http://dx.doi.org/10.1007/s10339-006-0154-7>.
- [20] Teresa Chandler. “kinds” of emotion. *Biology and Philosophy*, 16:109—115, 2001.
- [21] Ahmed Channouf and Georges Rouan. *Émotions et cognitions*. Neurosciences & cognition. De Boek, 2002.
- [22] Antonio R. Damasio. *Descartes' error: emotion, reason and the human brain*. Avon Books, 1994.
- [23] Antonio R. Damasio. *L'erreur de Descartes : la raison des émotions*. Odile Jacob, 1995.
- [24] Antonio R. Damasio. *Le sentiment même de soi*. Odile Jacob, Paris, 1999.
- [25] Antonio R. Damasio. Les émotions, source de la conscience. *Sciences Humaines*, 119:44—47, 2001.
- [26] Hanna Damasio, T. Grabowski, R. Frank, A.M. Galaburda, and Antonio R. Damasio. The return of phineas gage: clues about the brain from a skull of a famous patient. *Science*, 264:1102—1105, 1994.

-
- [27] Charles Darwin. *The expression of the emotions in man and animals*. Charles Darwin, London, 1872.
- [28] Charles Darwin. *L'expression des émotions chez l'homme et les animaux*. Charles Darwin, 1874.
- [29] Derek Denton. *The Primordial Emotions: The Dawning of Consciousness*. Oxford University Press, USA, 2006.
- [30] René Descartes. *Les passions de l'âme*. Henry Le Gras, Paris, 1649.
- [31] Marco Dorigo. *Optimization, learning and natural algorithms*. PhD thesis, Politecnico di Milano, Italy, 1992.
- [32] Antoine Dutot. Graphstream: A dynamic graph library., 2011. <http://graphstream-project.org/>.
- [33] Antoine Dutot, Frédéric Guinand, Damien Olivier, and Yoann Pigné. Graphstream: A tool for bridging the gap between complex systems and dynamic graphs. In *EPNACS: Emergent Properties in Natural and Artificial Complex Systems*, 2007.
- [34] Paul Ekman. *Emotion in the human face*. Cambridge University Press, New York, 1982.
- [35] Paul Ekman. Expression and the nature of emotion. In Klaus R. Scherer and Paul Ekman, editors, *Approaches to emotion*. Erlbaum, Hillsdale, NJ, 1984.
- [36] Paul Ekman. An argument for basic emotions. *Cognition and Emotion*, 6:169–200, 1992.
- [37] Paul Ekman and W. Friesen. *Unmasking the face. A guide to recognizing emotions from facial clues*. Prentice Hall Trade, 1975.
- [38] Paul Ekman and W. Friesen. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press, Palo Alto, 1978.
- [39] Clark Elliott. *The Affective Reasoner. A process model of emotions in a multi-agent system*. PhD thesis, Northwestern University. Institute for the Learning Sciences., Evanston, IL, 1992.
- [40] Clark Elliott. Research problems in the use of a shallow artificial intelligence model of personality and emotion. In *Proceedings of the 12th National Conference on Artificial Intelligence*, volume volume 1, page 9–15, Seattle, Washington, United States., 1994.
- [41] George W. Ernst and Allen Newell. *GPS: a case study in generality and problem solving*. Academic Press, New York, 1969.
- [42] Beverley Fehr and James A. Russell. Concept of emotion viewed from a prototype perspective. *Journal of experimental psychology*, 113:464–486, 1984.

- [43] Lawrence J. Fogel, Alvin J. Owens, and Michael J. Walsh. *Artificial intelligence through simulated evolution*. Wiley, Chichester, WS, UK, 1966.
- [44] Sigmund Freud. *Das Unbehagen in der Kultur*. Internationaler Psychoanalytischer Verlag, 1930.
- [45] Sigmund Freud. *Malaise dans la civilisation*. Presses Universitaires de France, 1979.
- [46] Nico Frijda. *The Emotions*. Cambridge University Press, 1986.
- [47] Félix Gaffiot. *Dictionnaire latin-français*. Hachette, 1934.
- [48] Joseph C. Giarratano and Gary D. Riley. *Expert Systems: Principles and Programming*. Course Technology, 4th edition, 2004.
- [49] Daniel Goleman. *L'intelligence Emotionnelle*. Robert Laffont, 1997.
- [50] Pierre-Paul Grassé. La reconstruction du nid et les coordinations interindividuelles chez *Bellicositermes natalensis* et *Cubitermes sp.* la theorie de la stigmergie : essai d'interpretation du comportement des termites constructeurs. *Insectes Sociaux*, 6(1):41—81, 1959.
- [51] Johnathan Gratch. Why you should buy an emotional planner. In *Proceedings of the Agents'99 Workshop and Emotion-based Agent Architectures (EBAA'99)*., 1999.
- [52] Johnathan Gratch. Émile: Marshalling passions in training and education. In *Proceedings 4th International Conference on Autonomous Agents (Agents'2000)*, page 325–332, New York, 2000. ACM Press.
- [53] Johnathan Gratch and Stacy Marsella. A domain-independent framework for modeling emotion. *Cognitive Systems Research*, 5:269–306, 2004.
- [54] Johnathan Gratch and Stacy Marsella. Computational models of emotion: Designs and evaluation. Tutorial ACII'2007, 2007.
- [55] Paul Griffiths. *What emotions really are: The problem of psychological categories*. Science & Its Conceptual Foundations. University of Chicago Press, Chicago, 1997.
- [56] John H. Holland. *Adaptation In Natural And Artificial Systems*. University of Michigan Press, 1975.
- [57] C.E. Izard. *Patterns of emotions*. Academic Press, San Diego, CA, 1972.
- [58] L.L. Jacoby. Perceptual enhancement: Persistent effects of an experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9:21–38, 1983.
- [59] William James. What is an emotion? *Mind*, 9(34):188–205, 1884.
- [60] Wolfgang Kantschik and Wolfgang Banzhaf. Linear-tree gp and its comparison with other gp structures. In J. Miller et al., editor, *Proceedings of 4th EuroGP Conference, Como 2001, LNCS 2038*, pages 302 – 312, Berlin Heidelberg, 2001. Springer-Verlag.

- [61] Paul R. Kleinginna and Anne M. Kleinginna. A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion*, 5(4):345–359, 1981.
- [62] H. Klüver and P.C. Bucy. An analysis of certain effects of bilateral temporal lobectomy in the rhesus monkey with special reference to « psychic blindness ». *Journal of Psychology*, 5:33–54, 1938.
- [63] H. Klüver and P.C. Bucy. Preliminary analysis of functioning of the temporal lobes in monkeys. *Archives of Neurology and Psychiatry*, 42:979–1000, 1939.
- [64] John R. Koza. *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press, 1992.
- [65] John R. Koza. *Genetic Programming II: Automatic Discovery of Reusable Programs*. MIT Press, 1994.
- [66] W.R. Kunst-Wilson and Robert B. Zajonc. Affective discrimination of stimuli that cannot be recognized. *Science*, 207:557–558, 1980.
- [67] Carl Georg Lange. *Om sindsbevaegelser. Et Psyko-fysiologisk Studie*. Kjobenhavn: Jacobs Lunds Forlag, 1885.
- [68] Richard S. Lazarus. Thoughts on relations between emotion and cognition. *The American Psychologist*, 37(9):1019–1024, Sep 1982.
- [69] Richard S. Lazarus. *Emotion and Adaptation*. Oxford University Press, New York, 1991.
- [70] Richard S. Lazarus, A.D. Kanner, and S. Folkman. Emotion : a cognitive-phenomenological analysis. In R. Plutchick and H. Kellerman, editors, *Emotion, theory, research, and experience: Theories of emotions*, volume 1, pages 189–217. New York : Academic Press, 1980.
- [71] Richard S. Lazarus and R. Launier. Stress-related transactions between person and environment. In L.A. Pervin and M. Lewis, editors, *Perspectives in interactional psychology*, pages 287–327. New York : Plenum, 1978.
- [72] Joseph E. LeDoux. Émotion, mémoire et cerveau. *Pour La Science*, 202:50–57, 1994.
- [73] Joseph E. LeDoux. *The emotional brain*. Weidenfeld & Nicolson, 1998.
- [74] Joseph E. LeDoux and Jeff Muller. Emotional memory and psychopathology. *Philosophical transaction of the Royal Society of London. Series B, Biological sciences.*, 352:1719–1726, 1997.
- [75] R.W. Levenson, Paul Ekman, and W.V. Friesen. Voluntary facial action generates emotion-specific autonomous nervous system activity. *Psychophysiology*, 27:363–384, 1990.

- [76] Françoise Lotstra. Le cerveau émotionnel ou la neuroanatomie des émotions. *Cahiers critiques de thérapie familiale et de pratiques de réseaux*, 29(2):73–86, 2002.
- [77] Paul D. MacLean. Culminating developments in the evolution of the limbic system: the thalamocingulate division. In B.K. Doane and K.F. Livingston, editors, *The limbic system*, pages 1–28. Raven Press, New York, 1986.
- [78] Karim Mahboub. Modélisation des processus émotionnels dans la prise de décision. Master's thesis, Université du Havre, Septembre 2006.
- [79] Karim Mahboub. Simulation of emotional processes in decision making. In *European Simulation and Modelling Conference '2006*, Toulouse, France, 2006.
- [80] Karim Mahboub, Cyrille Bertelle, Véronique Jay, and Evelyne Clément. Emotion : modèle d'appraisal-coping pour le problème des Cascades. In *Association pour la Recherche Cognitive '2007, Communications affichées*, pages 25–27, 2007.
- [81] Karim Mahboub, Evelyne Clément, Cyrille Bertelle, and Véronique Jay. Emotion: Appraisal-coping model for the "Cascades" problem. In *European Simulation and Modelling Conference '2007*, Malta, 2007.
- [82] Karim Mahboub, Evelyne Clément, Cyrille Bertelle, and Véronique Jay. Emotion: Appraisal-coping model for the "Cascades" problem. In M.A. Aziz-Alaoui and Cyrille Bertelle, editors, *From System Complexity to Emergent Properties*, pages 355–363. Springer Berlin / Heidelberg, 2009.
- [83] George Mandler. *Mind and emotion*. W.W. Norton and Company, New-York, London, 1975.
- [84] George Mandler. Recognizing: The judgment of previous occurrence. *Psychological Review*, 87:252–271, 1980.
- [85] George Mandler, Y. Nakamura, and B.J. Shebo-Van Zandt. Non-specific effects of exposure on stimuli that cannot be recognized. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*, 13:646–648, 1987.
- [86] Gregorio Marañón. Contribution à l'étude de l'action émotive de l'adrénaline. *Revue Française d'Endocrinologie*, 2:301–325, 1924.
- [87] MATLAB. *version 7.10.0 (R2010a)*. The MathWorks Inc., Natick, Massachusetts, 2010.
- [88] J.D. Mayer, P. Salovey, and D. Caruso. Models of emotional intelligence. *Handbook of human intelligence*, 2nd edition:396–420, 2000.
- [89] John McCarthy. Making robots conscious of their mental states, 1995.
- [90] Warren S. McCulloch and Walter H. Pitts. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:115–133, 1943. Reproduit dans Boden, M.A. (Ed.). (1990) - [17], p22-39.

-
- [91] J.R. Millenson. *The Psychology of Emotion: Theories of Emotion Perspective*, chapter 4, pages 35—36. John Wiley & Sons, 1967.
- [92] Marvin Minsky. *The Society of Mind*. Simon & Schuster, 1988.
- [93] P.M. Niedenthal. Embodying emotion. *Science*, 316:1002–1005, 2007.
- [94] Peter Nordin. A compiling genetic programming system that directly manipulates the machine code. In Kenneth E. Kinneer, Jr., editor, *Advances in Genetic Programming*, chapter 14, pages 311–331. MIT Press, 1994.
- [95] Armelle Nugier. Histoire et grands courants de recherche sur les émotions. *Revue électronique de Psychologie Sociale*, 4:8–14, 2009. Disponible à l'adresse suivante : <http://RePS.psychologie-sociale.org/>.
- [96] Keith Oatley and Jennifer M. Jenkins. Human emotions: Function and dysfunction. *Annual Review of Psychology*, 43:55–85, 1992.
- [97] Andrew Ortony. On making believable emotional agents believable. In R. Trappl, P. Petta, and S. Payr, editors, *Emotions in Humans and Artifacts*, pages 189–211. MIT Press, 2002.
- [98] Andrew Ortony, Gerald L. Clore, and Allan Collins. *The cognitive structure of emotions*. Cambridge University Press, 1988.
- [99] Andrew Ortony and T.J. Turner. What's basic about basic emotions? *Psychological Review*, 97:315–331, 1990.
- [100] James P. Papez. A proposed mechanism of emotion. *Archives of Neurology and Psychiatry*, 38:725–743, 1937.
- [101] Rosalind Picard. *Affective Computing*. MIT Press, 1997.
- [102] Rosalind W. Picard. Affective computing: Challenges. *International Journal of Human-Computer Studies*, 59 (1–2):55—64, 2003.
- [103] Robert Plutchik. A general psychoevolutionary theory of emotion. In Robert Plutchik and H. Kellerman, editors, *Emotion: Theory, research, and experience: Vol. 1. Theories of emotion*. New York: Academic, 1980.
- [104] Robert Plutchik. The nature of emotions. *American Scientist*, 89:344–350, 2001.
- [105] Ingo Rechenberg. Cybernetic solution path of an experimental problem. *Royal Aircraft Establishment*, 1122, 1965.
- [106] Ingo Rechenberg. *Evolutionsstrategie – Optimierung technischer systeme nach prinzipien der biologischen evolution*. PhD thesis, Technical University of Berlin, 1971. Reprinted by Fromman-Holzboog (1973).
- [107] Craig W. Reynolds. Flocks, herds, and schools: a distributed behavioral model. *Computer Graphics*, 21(4):25–34, 1987. SIGGRAPH '87 Conference (1987).

- [108] Ira J. Roseman. Appraisal determinants of emotions: Constructing a more accurate and comprehensive theory. *Cognition & Emotion*, 10(3):241–278, 1996.
- [109] Ira J. Roseman, P.E. Jose, and M.S. Spindel. Appraisals of emotion-eliciting events: testing a theory of discrete emotions. *Journal of Personality and Social Psychology*, 59(Issue 5):899–915, 1990.
- [110] James A. Russell. Culture and the categorisation of emotion. *Psychological Bulletin*, 110:426–450, 1991.
- [111] Kenneth Bradley Russell and Christopher John Kline. Jogl: Java binding for the opengl api.
- [112] P. Salovey, B. Bedell, J. Detweiler, and J. Mayer. Current directions in emotional intelligence research. In M. Lewis and J. Haviland-Jones, editors, *Handbook of Emotions*, page 504–520. Guilford Press, 2000.
- [113] S. Schachter. The interaction of cognitive and physiological determinants of emotiona state. In L. Berkowitz, editor, *Advances in experimental social psychology*, volume 1, pages 49–80. Academic Press, New York, 1964.
- [114] S. Schachter and J.E. Singer. Cognitive, social, and psychological determinants of emotional state. *Psychological Review*, 69(5):379–399, 1962.
- [115] Klaus R. Scherer. On the nature and function of emotion. a component process. In Klaus R. Scherer and Paul Ekman, editors, *Approaches to Emotion*, pages 293–317. Erlbaum, Hillsdale, NJ, 1984.
- [116] Klaus R. Scherer. Studying the emotion-antecedent appraisal process: An expert system approach. *Cognition and Emotion*, 7 (3/4):325–355, 1993.
- [117] Klaus R. Scherer. Appraisal theories. In T. Dalgeish and M. Power, editors, *Handbook of cognition and emotion*. Wiley, Chichester, 1999.
- [118] Klaus R. Scherer. Emotion. In M. Hewstone and W. Stroebe, editors, *Introduction to Social Psychology : A European perspective*, page 151–191. Oxford Blackwell Publishers, 2000.
- [119] Klaus R. Scherer. Appraisal considered as a process of multilevel sequential checking. In Klaus R. Scherer, A. Schorr, and T. Johnstone, editors, *Appraisal Processes in Emotion: Theory, Methods, Research*, pages 92–119. Oxford University Press, 2001.
- [120] Klaus R. Scherer. Emotions are emergent processes: they require a dynamic computational architecture. *Philosophical transaction of the Royal Society of London. Series B, Biological sciences.*, 364:3459–3474, 2009.
- [121] Klaus R. Scherer and M. Peper. Psychological theories of emotion and neuropsychology research. In F. Boller and J. Grafman, editors, *Handbook of Neuropsychology*, volume 5: Emotional behavior and its disorders, pages 17–48. Elsevier, Amsterdam, f. boller & j. grafman edition, 2001.

- [122] J.G. Seamon, N. Brody, and D.M. Kauff. Affective discrimination of stimuli that are not recognized: II. effect of delay between study and test. *Bulletin of the Psychonomic Society*, 21:187–189, 1983.
- [123] E.H. Shortliffe, F.S. Rhame, S.G. Axline, S.N. Cohen, B.G. Buchanan, R. Davis, A.C. Scott, R. Chavez-Pardo, and W.J. van Melle. Mycin: A computer program providing antimicrobial therapy recommendations. *Clinical Medicine*, 34, 1975.
- [124] Aaron Sloman. *Human Cognition and Social Agent Technology*, volume 19 of *Advances in Consciousness Research Series*, chapter Architectural Requirements for Human-like Agents Both Natural and Artificial. (What sorts of machines can love?), pages 163–195. John Benjamins Publishing Company, 2000.
- [125] Aaron Sloman et al. The architectural basis of affective states and processes. In M. Arbib and J.-M. Fellous, editors, *Who Needs Emotion? The Brain Meets the Robot*. Oxford University Press, 2003.
- [126] A. Teller and M. Veloso. Pado: A new learning architecture for object recognition. In *Symbolic Visual Learning.*, pages 81–116. Oxford University Press, 1996.
- [127] S. Tomkins. *Affect, imagery and consciousness: The positive affects (vol.1)*. Springer, New York, NY, 1962.
- [128] S. Tomkins. *Affect, imagery and consciousness: The negative affects (vol.2)*. Springer, New York, NY, 1963.
- [129] J.D. Velásquez. Cathexis : A computational model for the generation of emotions and their influence in the behavior of autonomous agents. Master’s thesis, MIT, 1996.
- [130] J.D. Velásquez. Modeling emotions and other motivations in synthetic agents. In *Fourteenth national conference on artificial intelligence and ninth innovative applications of artificial intelligence conference*, Menlo Park, 1997.
- [131] J.D. Velásquez. Modeling emotionbased decisionmaking. In *Proceedings of the 1998 AAAI Fall Symposium Emotional and Intelligent: The Tangled Knot of Cognition*, Orlando, FL, 1998. AAAI Press.
- [132] J.D. Velásquez. When robots weep: Emotional memories and decision-making. In *Proceedings 15th National Conference on Artificial Intelligence (AAAI-98)*, page 70–75, 1998.
- [133] Anna Wierzbicka. Talking about emotions: Semantics, culture, and cognition. *Cognition and Emotion*, 6(3/4):285–319, 1992.
- [134] S.A.K. Wilson. Some problems in neurology II: Pathological laughing and crying. *Journal of Neurology and Psychopathology*, 6:299–333, 1924.
- [135] Lotfi A. Zadeh. Fuzzy sets. *Information and Control*, 8(3):338–353, 1965.

- [136] Robert B. Zajonc. Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35:151–175, 1980.
- [137] D.H. Zald and S.W. Kim. Anatomy and function of the orbitofrontal cortex ii : Function and relevance to obsessive compulsive disorder. *Journal of Neuropsychiatry and Clinical Neurosciences*, 8:249–261, 1996.

ANNEXES

LE MODÈLE OCC

A

		+	-
Consequences of events	For others	Happy for	Resentment
		Gloating	Pity
	For self	Hope	Fear
		Joy	Distress
Action of agents	Self agent	Pride	Shame
		Gratification	Remorse
		Gratitude	Anger
	Other agent	Admiration	Reproach
		Gratification	Remorse
		Gratitude	Anger
Aspects of objects		Love	Hate

FIG. A.1 – ATTRIBUTS ÉMOTIONNELS SELON LE MODÈLE OCC

LE MODÈLE DE MILLENSON

B

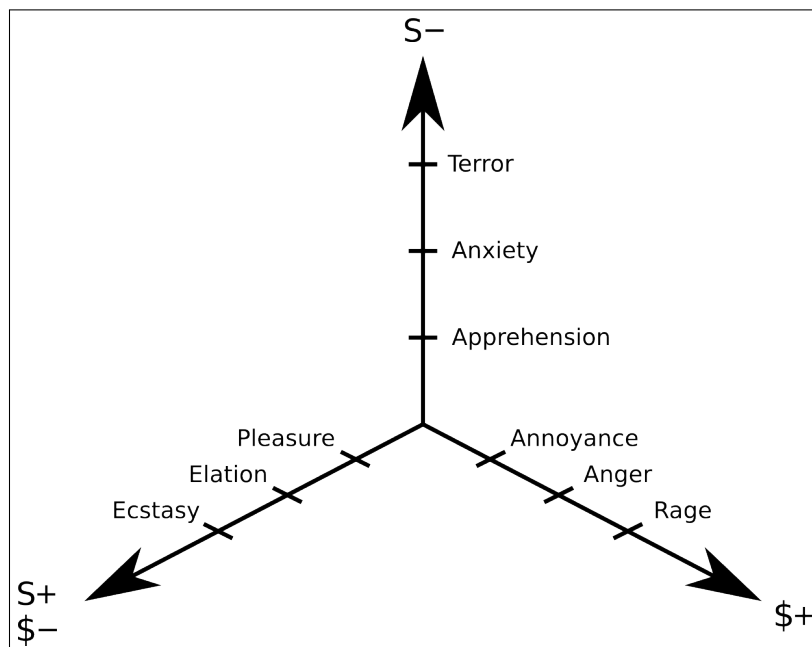


FIG. B.1 – MODÈLE TRIDIMENSIONNEL DE L'INTENSITÉ ÉMOTIONNELLE DE MILLENSON

*Psychologically Grounded
Emotion Model for a NPC Fuzzy
Controller*

C

Voici l'article publié en collaboration avec William BLEWITT, Aladdin AYESH et Cyrille BERTELLE. Cet article est en cours de révision pour être présenté dans une revue internationale.

Psychologically Grounded Emotion Model for a NPC Fuzzy Controller

William Blewitt, Aladdin Ayesch, Cyrille Bertelle, Karim Mahboub



Abstract—Emotions as a function of the complex role they play within cognition can have a profound impact on perceptions and decision-making. The abstract representation of emotional state within an artificial agent is therefore a highly active field of investigation. The work presented here centres on the implementation of a psychologically grounded, CER-based model of emotion as a function of stimulus. The model, first proposed by Millenson and informed by the earlier works of Watson, is represented through fuzzy inferencing, and the pedagogue example is implemented through an OpenGL game environment.

Index Terms—Index Terms— emotion, fuzzy logic, artificial intelligence, emotion representation, Geneva Emotion Wheel.

1 INTRODUCTION

THE advent of advanced interactive games, in the context of the ever-advancing field of computer and video-gaming in particular, has provided a largely untapped outlet for the exploration and mimicry of human decision-making. An important factor within this arena that currently enjoys significant attention is the representation of emotional state.

The accurate modelling of the emotional state of any agent, and changes to that emotional state, has long been the subject of extensive research [1], [2], [3], [4], [5]. This work has varied in scope and purpose across an array of disparate topics. From philosophical questions about why to make an emotional agent, to sociological questions about what impact emotional agents have on human-computer interactions [6], [7]. Within computing sciences, much debate has tended to revolve around the

practical questions regarding the mimicry of emotional responses and internal modelling of an agent's emotional state.

Within psychology, there are many schools of thought regarding the nature of emotions and how they are best modelled, although such models are often abstract in nature and thus difficult to quantify computationally. And, importantly, many fail to address the question of why computing scientists should concern themselves with the emotional state of an agent as a computational construct in its own right.

Recent publications in psychology have discussed a growing view that, rather than emotion being impedance to human creativity and decision-making, it is indeed an asset, if not a fundamental basis to the cognitive process [8], [9], [10], [11]. As this viewpoint begins to garner wider acceptance, it is expected that a renewed interest in emotion modelling and representation, from an academic standpoint, will follow, along with an exploration of alternative emotion models connecting environment to feeling and, consequently, to action.

The psychological model implemented in this paper was first published in 1967, and explicitly connects perceived stimulus with emotion [12]. This connection is of particular consequence when considered from the perspective of an autonomous agent in a complex game environment. A brief overview of the preliminary work associated with this research was published in [13]; a more exhaustive analysis, discussing extensively the psychological background and outlining the mathematics in more explicit terms is included here, along with a revised form of the model implementation.

*W. Blewitt and A. Ayesch are with De Montfort University, Leicester, UK. C. Bertelle and K. Mahboub are with l'Université du Havre, Le Havre, France.
Email: aayesh@dmu.ac.uk*

This paper shall present a brief overview of the psychological background of the implemented emotion model, which is dubbed the Millenson Model after its originator, in the context of alternative models. It shall then discuss in detail the mathematical construction of the emotion model's fuzzy representation. Implementation of the system shall be presented in particular detail, both in the context of implementation of the fuzzy system, and the translation from that implementation into the game environment. This shall be followed by discussion of the testing methodology applied to the system, and critical analysis of the testing results, before the paper is drawn to a conclusion.

2 BACKGROUND

Possibly more than any other multidisciplinary research area, the computational modelling of emotion requires some significant measure of preface, if only to outline the standpoint from which the researchers approach the issue. It is hoped that by outlining both the basic psychology of the work, including the historical research context in which the work fits, and the mathematics upon which the work is based, the overall aims and goals of the research might be better understood.

2.1 Psychology of Emotions

Within the various fields of psychological research, two schools of thought appear to dominate the debate regarding the nature of emotions, and how they are best modelled [14], [15]. From a philosophical perspective, the nature of their divergence and their theoretical differences are of great importance; from a computing sciences perspective, however, their differences lie entirely in the nature of the models they propose.

The view of emotions as an evolutionary construct was initially proposed by Darwin in 1872 [16]. It was this work which postulated the idea of basic emotions, differing combined intensities of which might give rise to an overall emotional state.

Over the past century, this has ultimately given rise to swathes of psychological research dedicated to determining both the number and nature of these basic emotions. The exact number of fundamental emotions given widely varies. Plutchik first proposed his system of emotion classification in 1980 [17], containing eight fundamental emotions. In contrast, Ekman proposed a system consisting of six fundamental, or basic, emotions in 1982 [14].

The maximum number of basic emotions is generally thought to be fourteen [18]. Following on from the definition of basic emotions comes the definition of more complex emotions. Occasionally, these categories are divided using nomenclature indicating primary and secondary emotions as in the structure proposed by Parrott [15]. Oftentimes, however, these more complex emotions are simply defined by the relative intensities of their parent emotions. In general terms, however, it is the view of this school of thought that the sum of human emotional experience can be defined as a function, or construct, of less than a dozen named emotions [19], [20].

An alternative to this view proposed by Wundt in 1904 suggested that emotions could be better defined in the context of experience than crisp linguistics [21]. Research based on this principle has, as with the Darwinian view, given rise to many varied schools of thought following the same fundamental idea.

In Wundt's original model, emotional state was represented in terms of three facets of experience which he labeled pleasantness, approach and arousal. He asserted that any individual emotion would be better modelled in the context of relative magnitudes of these facets of the emotional experience than through verbal labels.

Subsequent to Wundt's original work, significant research has been performed regarding this idea of a dimensional emotion model. In many cases it is common for the third axis to be ignored and, instead, for proponents of this view to model emotions in the context of valence, which might be seen as a clearer definition of pleasantness, and arousal.

More recently, however, it has been suggested that these views are not necessarily

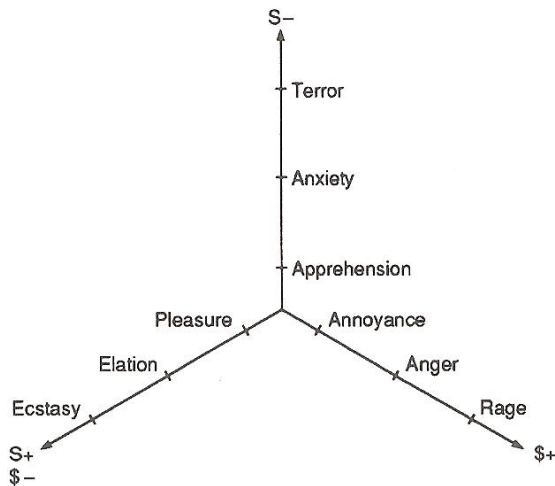


Fig. 1. Millenson's 3-Dimensional Model of Emotional Intensity [12]

mutually exclusive. Russell produced a circular model of emotions outlining the position of what he believed to be fundamental emotions in terms of relative values of what were effectively arousal and valence [22].

It is upon this idea of hybridised conceptual models that this work pitches its focus. Such models permit us to consider the emotional state as an output, or resultant, of disparate and seemingly unrelated contextual inputs. In particular, attention has been directed towards the model proposed by Millenson in 1967 [12].

Millenson's model of emotion, which has been discussed in terms of computational simulation previously [23], was defined as a stand-alone model of emotion and not presented with its own psychological theory. Rather, it was built upon Watson's three-factor theory [24], [25]. Often considered the father of behaviourism [26], Watson proposed the connection between applied and withdrawn stimulus, and resultant emotional response.

Millenson's model took this idea and derived a three-axis system that associated certain applied and removed stimulus with different facets of emotional experience. Figure 1 shows an interpretation of his modular structure, where $S+$ represents an applied positive stimulus, $\$+$ represents a removed positive stimulus, $S-$ represents an applied negative stimulus, and $\$-$ represents a removed negative stimulus.

Along each axis, Millenson places an emo-

tional archetype. He associates the x -axis with anger, the y -axis with anxiety, and the z -axis with pleasure. He acknowledges that three emotions do not account for the sum total of emotional experience, and accounts for this in two ways.

First, he posits that some emotions vary from each other only in terms of their intensity. Given the structure of his model, this is a linguistically ambiguous statement, with one of two meanings. The first possible meaning is that along a given axis, all named emotions are essentially the same emotion at varying levels of intensity. The second possible meaning is semantic inasmuch as it may be interpreted that his statement meant that emotions along a particular axis are only triggered by a more intensely felt application or removal of their associated stimulus. Ultimately, it is this latter interpretation which is pursued here; that there are three discrete emotions associated with each of his axes.

Second, he suggests that some emotions are simply complex compounds of the basic emotions his model acknowledges. In this, his theory is consistent with many subsequent works suggesting the existence of basic emotions; that being that complex emotions, which might be linguistically recognised in popular language, are complex compounds of two or more basic emotions [20].

In terms of an applied example of Millenson's theory in the context of compound emotional responses to stimulus, his own example was that of a child taking a cookie from a jar [12]. In this example, there is an associated applied positive stimulus with the action, that being to eat the cookie, and an associated applied negative stimulus, that being the fear of being caught. We can define this compound as guilt. We can also consider other combinations of more dramatically conflicting stimulus, such as behaviours that arise from neuroses.

As Figure 1 outlines, his nine basic emotions [12] are divided into three groups: those associated with the removal of positive stimulus; those associated with the application of negative stimulus; and, those associated with the conceptual combination of applied positive stimulus and removed negative stimulus.

Listing these respectively, and subsequently in order of implied intensity, these emotions are explicitly shown in Figure 1 as being

- Annoyance
- Anger
- Rage
- Apprehension
- Anxiety
- Terror
- Pleasure
- Elation
- Ecstasy

Looking at Millenson’s model in a process sense, fundamentally is associates a given event with a composite of application or removal of two stimuli. From this, it defines an emotional response associated with the event, represented by nine basic emotions. These emotional responses to stimuli, over time, can be used to define an adaptive emotional state.

It is, in particular, the model’s reliance upon stimuli that makes it valuable from a computational sense, and worth exploring alongside more well-known multi-dimensional emotion representations. We are compelled to compare Millenson’s approach to that of Mehrabian’s Pleasure-Arousal-Dominance (PAD) model [27].

In his work, Mehrabian interconnects the idea of basic emotions to a three-dimensional construct. He describes “nearly orthogonal” scales, Pleasure, Arousal and Dominance, creating a temperament space providing a method of comparing personality measures. Interestingly, the three scales described are analogous to those of Wundt [21], but the idea of this emotional ‘state-space’ draws natural comparison with the model produced by Millenson.

The difference between the two lies in the psychological basis. Built upon Conditioned Emotional Response theory, Millenson’s model provides explicit connection between emotion and stimulus in a simple and finite form - the application, or removal, of positive or negative stimulus. As such, from a computational perspective it is far easier to apply while maintaining psychological consistency.

PAD, while a more exhaustive and comprehensive system for generating emotional state from environmental experience, requires a far

greater level of abstraction in implementation. This necessitates a level of abandonment of the psychology as the model is computationally transliterated.

2.2 Fuzzy Logic

First postulated by Zadeh in 1965 [28], fuzzy logic, or multivalued logic, is an adaption of set theory which permits variable degrees of membership. Essentially, rather than binary logic, whereby a statement is either true or false, fuzzy logic permits that a statement have a given degree of truth, dependent upon functions that define that degree, or membership grade.

Emotional linguistics has been described by psychology researchers as ‘fuzzy’ [18], and the application of fuzzy logic to the solution of problems that are defined linguistically is well documented [29], [30]. Indeed, prior precedent would suggest that the only way to attempt to model concepts as vague as emotion would be through multivalued logic systems. Here we present a brief overview to the theory of fuzzy logic.

2.2.1 Fuzzy Sets

In traditional set theory, an object k has a binary relation, defined by a bivalent condition, with a set A . If k is a member of the set A , we denoted this

$$k \in A \quad (1)$$

In fuzzy set theory, set A is actually represented as a pair, (A, μ) where A is a set, and

$$\mu: A \rightarrow [0, 1] \quad (2)$$

Membership of k in the set (A, μ) is now defined as $\mu(k)$; we call this the membership grade of k .

k is considered excluded from the set (A, μ) if $\mu(k) = 0$; k is considered fully included in the set (A, μ) if $\mu(k) = 1$; and, k is considered a fuzzy member of (A, μ) if $0 < \mu(k) < 1$. For a finite set $A = k_1, k_2, \dots, k_n$, we denote the fuzzy set (A, μ) as $\mu(k_1)/k_1, \mu(k_2)/k_2, \dots, \mu(k_n)/k_n$.

Fuzzy logic, the extension of fuzzy set theory, utilises degrees of truth to determine the nature

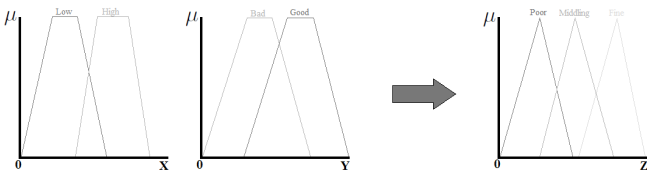


Fig. 2. Two Fuzzy Inputs leading to One Fuzzy Output

of a system. In particular, while mathematical variables often take crisp, numerical values, fuzzy logic permits the use of linguistic variables [31]. These variables may be associated with qualifying terms such as short, warm or poor.

2.2.2 Mamdani Fuzzy Inferencing

In his original work regarding the application of fuzzy set theory to control systems [29], Mamdani described in detail an experiment regarding "linguistic" synthesis of a controller for a steam engine. He applied fuzzy logic to convert heuristic control rules as stated by a human operator into an automatic control strategy. His work was based on a paper published by Zadeh in 1973 discussing complex systems and decision making [32].

The structure of a Mamdani FIS is comparatively simple, and built upon rules of the form

If X is ' x ', then Y is ' y '

where x and y are linguistic values determined by fuzzy sets along the ranges X and Y . The 'if' statement we define as the premise, or antecedent, while the 'then' statement is defined as the conclusion or consequent [33].

Applying fuzzy logic to these statements, we introduce the idea of gradated truth, or membership grades, of individual statements. For example, let us consider two input ranges: X possessed of two linguistic variables, *High* and *Low*, and Y , possessed of two linguistic variables, *Good* and *Bad*. These two input ranges are used in the determination of an output variable Z , with linguistic variables *Poor*, *Middling* and *Fine*. These are shown in Figure 2.

Let us generate three rules for our example system.

Rule 1:

If X is *High*, and Y is *Good*, then Z is *Fair*

Rule 2:

If X is *Low*, and Y is *Good*, then Z is *Middling*

Rule 3:

If X is *Low*, and Y is *Bad*, then Z is *Poor*

Our rules are using the And (min) operator. There are two operators we commonly use in Mamdani fuzzy inferencing systems: And (min), and Or (max). Our systems focus particularly upon the And case. Numerically, these are represented through simple equations relating to the membership function of a linguistic variable associated with an input. Let us consider the specifics of these relations here.

Let us consider input range X , with linguistic variables *High* and *Low*. Following on from fuzzy set theory, we formalise this as

$$(X, \mu) = \{\mu(\text{Low})/\text{Low}, \mu(\text{High})/\text{High}\}$$

Likewise, in considering our input range Y and output Z , we formalise them as

$$(Y, \mu) = \{\mu(\text{Bad})/\text{Bad}, \mu(\text{Good})/\text{Good}\}$$

$$(Z, \mu) = (\{\mu(\text{Poor})/\text{Poor}, \mu(\text{Middling})/\text{Middling}, \mu(\text{Fair})/\text{Fair}\})$$

For any given values applied to the ranges X and Y , let us call these discrete values X' and Y' , values of m associated with specific linguistic variables are produced, and applied to the rules as follows. Let us consider Rule 1 in our hypothetical system. Rule 1 is an And rule, meaning that the membership of the linguistic output variable is determined by the minimum membership of the associated linguist input variables. Thus, for any values of X and Y , Rule 1 can be represented as

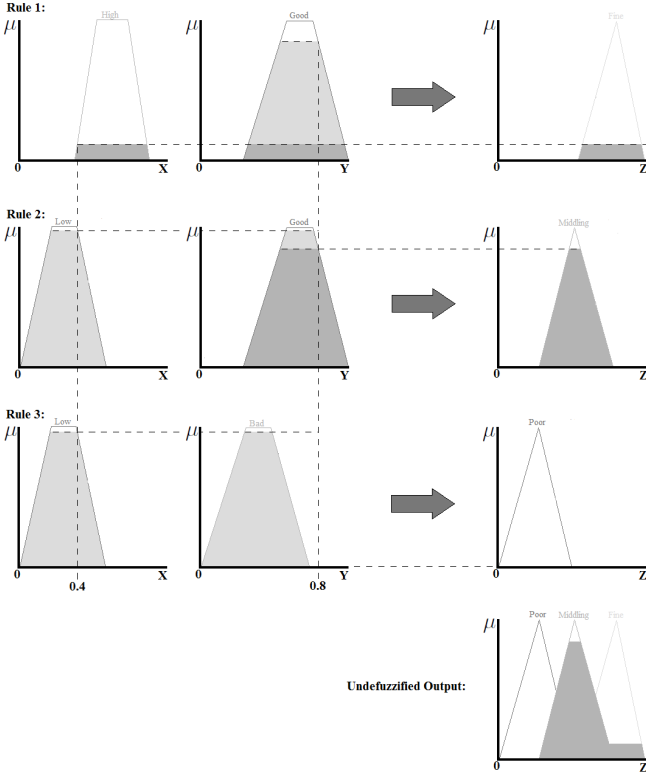


Fig. 3. Fuzzy Rules Triggered by a Value X'

$$\left\{ \frac{\mu_Z(\text{Fair})}{\text{Fair}} \right\} = \min\left\{ \frac{\mu_X(\text{High})}{\text{High}}, \frac{\mu_Y(\text{Good})}{\text{Good}} \right\}$$

In a standard Mamdani fuzzy inferencing system, each of these rules is considered for a given input to the system. That is to say that for any value X and Y received by the system, all three rules shall be evaluated to determine the resultant membership grades of the linguistic variables within Z . These membership grades determine the contribution from each linguistic variable to the pre-defuzzified output of Z . Figure 3 illustrates this process by assigning the value 0.4 to X , and the value 0.8 to Y .

Note that, as previously stated, the output shown in Figure 3 has not yet been defuzzified. Indeed, it is important to realise that the fuzzy inferencing system does not return memberships for the three linguistic variables that define Z . Rather, the inferencing system should return to us a value within the limits of Z by which we can define our output.

As shall be discussed in the context of the first Millenson implementation featured within

this paper, this crisp output can then be used to define other variables, dependent upon the needs and design of the system. For the purpose of this illustration, however, it is simply a desired, crisp value of Z .

In order to obtain this crisp output, it is required that the output shown in Figure 3 be defuzzified. The most common form of defuzzification associated with the Mamdani fuzzy inferencing system is known as centroid defuzzification.

Centroid defuzzification calculates the total area defined by the membership grades of the linguistic variables, with respect to their membership functions. In Figure 3 this area is represented by the dark grey region of the defuzzified output. Having determined the total area under the curve, the defuzzifier then calculates the value of Z representing the centre of the area. Explicitly, the point along the Z -axis where the area under the curve to the left of the output value is equal to the area under the curve to the right.

The description, in fuzzy terms, of the emotion models outlined in this paper, and how they may be used to generate a conceptual emotion model, is the dominant feature of Section 3 of this paper. The specifics of the construction of the models' respective inferencing systems is covered in Section 4, *Implementation*.

3 APPLICATION OF FUZZY LOGIC TO THE MILLENSON MODEL

3.1 Fuzzifying the Millenson Model

Millenson provides connective links between stimuli of differing valence and specific facets of emotional experience. As this link is contextual, we must first represent each axis as a conceptual relativistic sum of respective stimuli, with a crisp value between 0 and 1.

$$X = \sum\{\$+\}[0, 1] \quad (3)$$

$$Y = \sum\{S-\}[0, 1] \quad (4)$$

$$Z = \sum\{S+, \$-\}[0, 1] \quad (5)$$

Note that the three dimensional model presented in Figure 1 is now segregated into three,

independent variables. While the makeup of the model is heavily informed by Figure 1, the mathematics of representation are somewhat independent.

The nature of these variables, and the manner in which they are normalised into quantifiable values between 0 and 1, is naturally dependent upon the context in which they are applied and the setting in which they are being implemented; the Implementation section of this paper discusses this in explicit detail. As an example, however, let us consider a mobile agent within a universe shared with two other objects: an item of food, and a predator.

Conceptually speaking, at a time t , the distance between the agent and the food we define as r_t , and the distance between the agent and the predator we define as s_t . The stimulus the agent receives between time t and time $t + 1$ may be derived through changes to these two variables.

Thus we might consider the agent to have received application of positive stimulus ($S+$) if $r_{t+1} < r_t$, informing the variable Z . Conversely, if $r_{t+1} > r_t$, we might consider this a removal of positive stimulus ($S-$), informing the variable X .

Similarly, we might consider the application of negative stimulus ($S-$), associated with the variable Y , to be influenced should $s_{t+1} < s_t$. Likewise, if $s_{t+1} > s_t$ we could consider this the removal of negative stimulus ($S-$), and thus define an impact on the variable Z .

Thus we define the concept of an Stimulus Event \mathbf{J} . \mathbf{J} is a column vector, defined by the variables X , Y and Z such that

$$\mathbf{J} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (6)$$

At this point, the manner in which we associate the stimulus event with emotional output conceptually bifurcates, as was briefly discussed during the psychological outline of the model. The inference explored in this paper is that the nine discrete emotions provided in Figure 1 are distinct, but distinct emotions along a given axis are associated by the nature of the stimulus, removed or applied, that triggers

them, and differentiated by the intensity of said stimulus that is required to impact them.

Millenson's model presents nine emotions that he describes as basic, whatever the interpretation of how they might be related to stimulus. It stands to reason, then, that in consideration of Millenson's emotion model, the emotional state include representation of all nine of these elements.

In explicit terms, this work defines the emotional state of an agent governed by the Millenon Model,

$$\mathbf{E}_M = \begin{bmatrix} \mu x_i[0, 1] \\ \mu y_i[0, 1] \\ \mu z_i[0, 1] \end{bmatrix} \quad (7)$$

where $i = 1, 2, 3$, as an array of nine elements, and where μ in all cases represents the membership grade of the named emotion the variable is associated with, and where; x_1 represents *Annoyance*, x_2 represents *Anger*, and x_3 represents *Rage*; y_1 represents *Apprehension*, y_2 represents *Anxiety*, and y_3 represents *Terror*; z_1 represents *Pleasure*, z_2 represents *Elation*, and z_3 represents *Ecstasy*.

These relative magnitudes indicate the level to which the agent is experiencing each individual emotion at a given instant. Naturally, these magnitudes are expected to vary as a function of stimulus events, and are informed by the system subsequently outlined.

3.2 Changing the Emotional State

Let us assume a given interpretation of Millenson's model that connects the intensity of applied and removed stimuli, of differing valence, with nine emotional components. Let us explicitly define these connections of the form

$$\begin{aligned} X &\rightarrow \text{Annoyance, Anger, Rage} \\ Y &\rightarrow \text{Apprehension, Anxiety, Terror} \\ Z &\rightarrow \text{Pleasure, Elation, Ecstasy} \end{aligned}$$

where X , Y and Z are defined in equations (4.1), (4.1) and (4.1), respectively, and their associated axes are shown in Figure 1.

We assign variables to each of these named emotions, as outlined in equation 7. Mirroring

our merging of stimulus components into a single variable, J , we define the emotional component of a stimulus event J as the nine element vector E_J . This is given by

$$E_J = \begin{bmatrix} \mu_J x_i [0, 1] \\ \mu_J y_i [0, 1] \\ \mu_J z_i [0, 1] \end{bmatrix} \quad (8)$$

where $i = 1, 2, 3$; and where μ_J of a specific discrete emotion is a crisp value between 0 and 1, representing the relative intensity with which that emotion is triggered by the stimulus event J ; and where x_{1-3} , y_{1-3} and z_{1-3} represent the discrete emotions with which they were associated previously. The distinction between E_J and E_M is conceptually fundamental. E_M is the emotional state of the agent. E_J is the emotional impact of a given stimulus event J ; as such, E_J informs E_M , but the two are not equivalent, as the distinctions in their notation emphasise.

Obtaining the emotional impact E_J from the stimulus event J is a conceptual problem which this paper approaches from the perspective of fuzzy inferencing. Let us consider an individual component of J , X .

As has been previously stated, a given value assigned to X is a number between 0 and 1, which conceptually represents the significance of removed positive stimulus from the system. Understanding that such a quantifier cannot crisply represent the nature of stimuli affecting the system, it is proposed that fuzzy linguistics be implemented in its interpretation.

Thus let us define X as an input to a Mamdani fuzzy inferencing system, with linguistic variables describing "No Intensity", "Some Intensity" and "Particular Intensity" and "Extreme Intensity". Thus we define X in fuzzy terms as

$$(X, \mu) = \begin{cases} \mu(\text{No Intensity})/\text{No Int.}, \\ \mu(\text{Some Int.})/\text{Some Int.}, \\ \mu(\text{Par. Int.})/\text{Par. Int.}, \\ \mu(\text{Ext. Int.})/\text{Ext. Int.} \end{cases}$$

Figure 4 illustrates a hypothetical structure for X as the input to a fuzzy inferencing system.

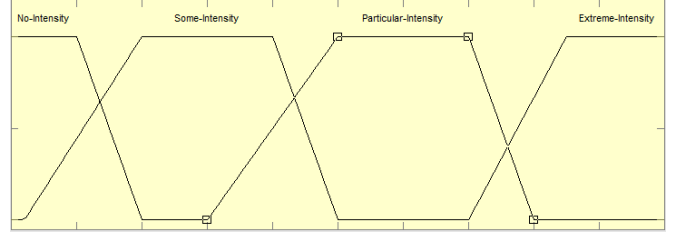


Fig. 4. Potential Implementation of X as an Input Variable of a FIS

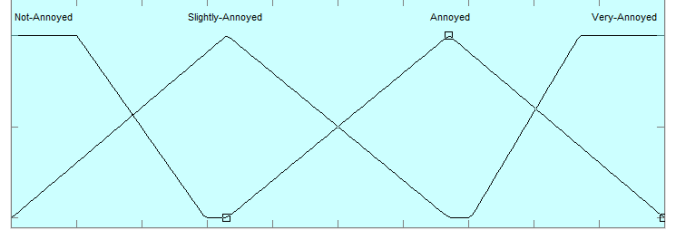


Fig. 5. Potential Implementation of x_1 as an Output Variable in a FIS

Millenson stipulates in conceptual terms the correlation between stimulus and emotional output, and thus we define x_1 , Annoyance, as an output of a Mamdani fuzzy inferencing system, with linguistic variables describing "Not Annoyed", "Slightly Annoyed", "Annoyed" and "Very Annoyed" describing the agent's reaction to given behavioural stimuli in emotional terms. In mathematical terms, this redefines x as

$$(x_1, \mu) = \begin{cases} \mu(\text{Not Ann.})/\text{Not Ann.}, \\ \mu(\text{Slight Ann.})/\text{Slight Ann.}, \\ \mu(\text{Annoyed})/\text{Annoyed}, \\ \mu(\text{Very Ann.})/\text{Very Ann.} \end{cases}$$

Figure 5 illustrates a potential interpretation of x_1 as an output for a fuzzy inferencing system.

Using comparable linguistic variables, we define the other two output variables associated with X , these being x_2 and x_3 .

$$(x_2, \mu) = \begin{cases} \mu(\text{Not Angry})/\text{Not Ang.}, \\ \mu(\text{Slightly Ang.})/\text{Slight. Ang.}, \\ \mu(\text{Angry})/\text{Angry}, \\ \mu(\text{Very Angry})/\text{Very Ang.} \end{cases}$$

$$(x_3, \mu) = \{ \mu(\text{Not Enraged})/\text{Not Enr.}, \\ \mu(\text{Slightly Enr.})/\text{Slight. Enr.}, \\ \mu(\text{Enraged})/\text{Enraged}, \\ \mu(\text{Very Enraged})/\text{Very Enr.} \}$$

The fuzzy inferencing system, by necessity, must have an array of rules associating inputs and outputs, as discussed in Section 2.2.2. In the case of X , we create a list of rules of the form:

Rule 1:

If X is No Intensity, then x_1 is Slightly Annoyed, and x_2 is Not Angry, and x_3 is Not Enraged

Rule 2:

If X is Some Intensity, then x_1 is Very Annoyed, and x_2 is Slightly Angry, and x_3 is Not Enraged

Rule 3:

If X is Particular Intensity, then x_1 is Annoyed, and x_2 is Very Angry, and x_3 is Slightly Enraged

Rule 4:

If X is Extreme Intensity, then x_1 is Slightly Annoyed, and x_2 is Angry, and x_3 is Very Enraged

It should be clarified that these linguistic terms defining the level of the named emotions refer specifically to its level relative to the other emotions. One might say that to be Enraged one must be Very Angry by definition, but that would be the wrong context. Rather, when the agent is Very Enraged, but only Angry rather than Very Angry, this is because Anger and Rage are treated as two distinct basic emotions by Millenson, and the system is experiencing Rage to a greater degree than it is experiencing Anger.

As we have with all other numerical variables, we apply limits of $[0,1]$ to x_1 , x_2 and x_3 in this context. That is to say that upon the conclusion of centroid defuzzification, as discussed in a previous section, the crisp output associated with each of the three named emotions shall

be a value between 0 and 1. We explicitly label these outputs as μx_1 , μx_2 , and μx_3 explicitly as, for our purposes, they define membership of their associated emotion within the emotional component of the stimulus vector \mathbf{J} , \mathbf{E}_J .

Expanding this system to include the other two input variables, and their associated six basic emotions, permits us to generate a Mamdani fuzzy inferencing system which will absorb the numerical contents of the stimulus vector \mathbf{J} and produce a value, between 0 and 1, for each of the nine named emotions in Millenson's model.

As these nine values represent the experiential level of the individual emotions, they are defined as the elements necessary to determine \mathbf{E}_J .

At this stage, the system can now inform the emotional state of the agent. How this emotional state is informed depends specifically upon the level of emotional memory it is desired that the agent experience; that meaning, how strongly the current emotional state mutes the impact of the stimulus event \mathbf{J} .

Let us return to the emotional state \mathbf{E}_M as defined by equation 7. In merging $\mathbf{E}_M(t)$ with \mathbf{E}_J , to obtain $\mathbf{E}_M(t+1)$, the simplest operator would be to take the mean of the two vectors, such that

$$\mathbf{E}_M(t+1) = \frac{(\mathbf{E}_M(t) + \mathbf{E}_J)}{2} \quad (9)$$

which is more exhaustively written as

$$\mathbf{E}_M(t+1) = \begin{bmatrix} (\mu x_i(t) + \mu_J x_i)/2 \\ (\mu y_i(t) + \mu_J y_i)/2 \\ (\mu z_i(t) + \mu_J z_i)/2 \end{bmatrix} \quad (10)$$

where $i = 1, 2, 3$. However, this notation is inherently limiting and precludes study into the area of emotional memory and its impact on learning systems. As such, it is preferred to use alternative notation of the form

$$\mathbf{E}_M(t+1) = \frac{(u\mathbf{E}_M(t) + v\mathbf{E}_J)}{u+v} \quad (11)$$

where u and v are constants introduced at the point of implementation, permitting the system to adjust the weightings of import between a

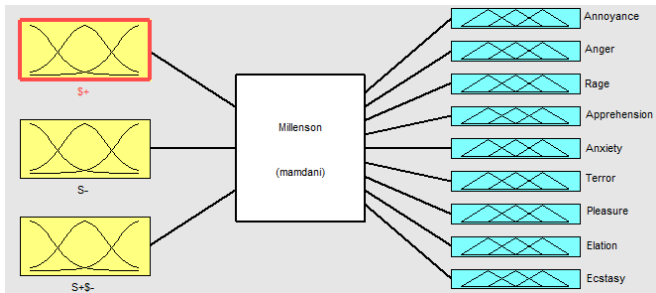


Fig. 6. Millenson: FIS Structure

new event and an established emotional state. Those wishing to investigate systems where emotional memory makes up a large component have the freedom to do so with this notation, as do those who wish to study systems with a limited memory component where emotional state is based predominantly upon instantaneous stimulus. The testing section of this Paper clarifies the implemented weightings.

4 IMPLEMENTATION

The Matlab Fuzzy Inferencing System Editor (FISE) provided a stable, well-integrated platform through which to generate the fuzzy implementation of the Millenson model of emotion.

Subsequently, we outline in specific detail the Matlab implementation of the mathematical mechanism previously outlined, and provide particular detail regarding the game environment to which the system was applied, using a pedagogue example involving an agent, predators and prey [34]. Also presented are the specific mechanisms implemented reflecting changes to the emotional state, and conceptual weightings.

4.1 Millenson in Matlab

Within Matlab's FISE, a Mamdani fuzzy inferencing system was generated, three inputs to nine outputs. This fuzzy inferencing system utilised the minimum 'And' operator as discussed in Section 2.2.2., and the centroid method of defuzzification. This structure is illustrated by Figure 6.

TABLE 1
Millenson: Input MFs

No Intensity	–	0.0	0.2
Some Int.	0.0	0.2	0.4
Particular Int.	0.4	0.6	0.8
Extreme Int.	0.8	1.0	–

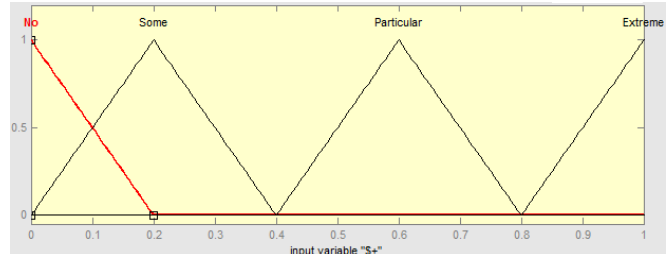


Fig. 7. Millenson: MFs of Input Variable X

As indicated in Section 3.1, these inputs were defined so as to represent the elements of a stimulus event, J , which was previously declared

$$J = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

where

$$X = \sum \{ \$+ \} [0, 1]$$

$$Y = \sum \{ S- \} [0, 1]$$

$$Z = \sum \{ S+, \$- \} [0, 1]$$

and where the symbols have previously declared contextual meanings. Each of the input elements X , Y and Z was, as a fuzzy construct, defined as having four associate fuzzy membership functions, representing quantified levels of associated stimulus in the context of intensity. Within this implementation, those fuzzy membership functions were represented as triangular functions, with their geometric vertices given by table 1.

Figure 7 illustrates the membership functions described by Table 1, as applied to input variable X .

These membership functions were applied uniformly to the three input variables X , Y and Z . The nine output variables, mathematically denoted as x_{1-3} , y_{1-3} , and z_{1-3} , were each

TABLE 2
Millenson: Output MFs

Not Annoyed	–	0.0	0.33
Slightly Ann.	0.0	0.33	0.67
Annoyed	0.33	0.67	1.0
Very Ann.	0.67	1.0	–

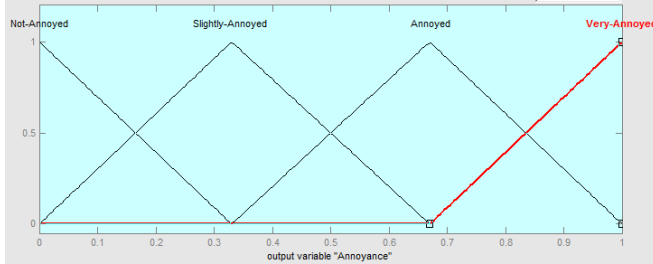


Fig. 8. Millenson: MFs of Output Variable x_1 - Annoyance

TABLE 3
Rules associating X with x_1 , x_2 and x_3

X	x_1	x_2	x_3
1	2	1	1
2	4	2	1
3	3	4	2
4	2	3	4

associated with four membership functions as outlined in 3.1.2. In implementing them, these membership functions were represented as triangular functions; their vertices are given in Table 2, in the context of the output x_1 , or the 'Annoyed' emotion. They are illustrated, again in the context of output x_1 , in Figure 8.

These membership functions were applied uniformly across all nine output emotions to maintain consistency across the target resultant, vector E_J , elements.

Rules were input into the system reflecting those outlined in Section 3.1.2., where the mathematics of this model were discussed in depth. Tables 3, 4 and 5 explicitly clarify the rules connections made within the FIS in terms of numerical hierarchy of coordinates.

Figure 9 illustrates Matlab's internal results calculation mechanism included within the FISE. It should be noted, however, that the actual testing environment utilised Matlab's command line functions, and the diagram is included for the sake of completion only.

TABLE 4
Rules associating Y with y_1 , y_2 and y_3

Y	y_1	y_2	y_3
1	2	1	1
2	4	2	1
3	3	4	2
4	2	3	4

TABLE 5
Rules associating Z with z_1 , z_2 and z_3

Z	z_1	z_2	z_3
1	2	1	1
2	4	2	1
3	3	4	2
4	2	3	4

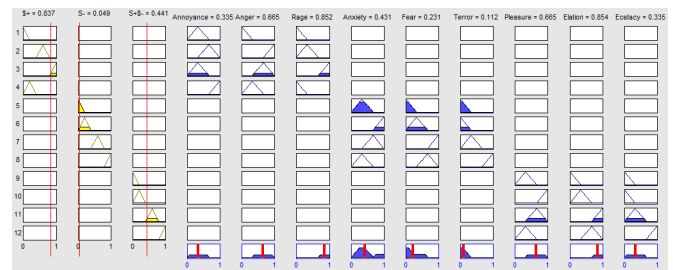


Fig. 9. Millenson: Matlab Rules Screen

An M-File was written which periodically obtained values for X , Y and Z from a delimited text file "Input.txt" generated by the agent simulation programme discussed in Section 5.2; at the same time, it obtained the previous emotional state, E_M from a delimited text file "Output.txt".

These discrete values were processed by the FIS, generating a resultant vector E_J . The mean of this vector E_J and E_M was then calculated, and saved to the delimited text file "Output.txt". The agent simulation programme then used the nine elements contained in "Output.txt" to amend its behaviour, before generating a new set of values saved to "Input.txt".

The specifics of how this output affected the agent's behaviour, and how the values for "Input.txt" were created, are included in Section 5.2.

4.2 Java implementation

The Millenson emotional model was adapted to a Java application representing a predator-prey

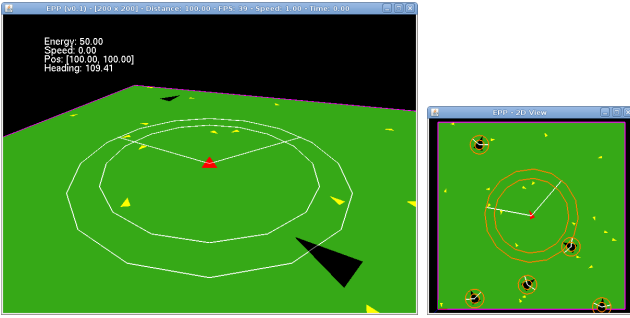


Fig. 10. Java environment

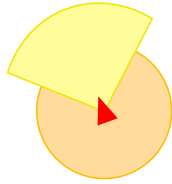


Fig. 11. Agent's viewing (yellow) and hearing radius (orange)

environment. In this environment, the main agent has to eat prey in order to maintain its energy at a sufficient level, while avoiding predators, whose goal is to eat our agent.

In the figure 10, the main agent is represented by the red triangle, while prey are the yellow triangles and predators are the black ones. To initiate the prey and predator displacements we need to implement moving strategies. Prey have been simulated as boids (see Craig Reynolds, [35]). They move around space with a tendency of keeping the same trajectory by staying together while fleeing in the presence of the agent. As for the predators, they wander until they detect the agent and hence move towards him.

In terms of detection, each species has not only its own viewing angle and viewing radius, but also a hearing radius, used for 'sensing' individuals coming from the back (see figure 11).

As a result, when an individual is approaching the agent, the latter has to decide what to do considering the individual type (prey or predator) and its distance. To do so, we first read the "output.txt" file which gives us the vector E_J produced by the Matlab module, in terms of the nine output variables x_{1-3} , y_{1-3} , and z_{1-3} .

In relating these nine distinct emotions to behaviour, we associate each triplet of emotions with a facet of behaviour. Let us consider x_{1-3} , or, in explicit terms, Annoyance, Anger and Rage. In this implementation, following from Millenson's psychology, these three emotions are associated with the removal of positive stimulus; as such, we associate these three emotions with the increasing desire of the agent to obtain food. As these distinct emotions represent differing intensities of removed positive stimulus, the fashion in which we apply them weights each emotion differently, and with increasing magnitude in line with increasing numeric identifier. Thus the weighting of Anger is greater than the weighting of Annoyance, and the weighting of Rage is greater than the weighting of Anger.

Ultimately, these concepts must be merged into an overall weighting value that can be applied to the attraction of prey. Let us call this weighting value X_e , and the weighting values of each distinct emotion w_{1-3} . Thus X_e is defined in explicit terms as

$$X_e = \frac{w_1x_1 + w_2x_2 + w_3x_3}{w_1 + w_2 + w_3}$$

or

$$X_e = \frac{\sum_{n=1}^3 w_n x_n}{\sum_{n=1}^3 w_n}$$

We project this representation across all nine distinct emotions in order to generate similar variables Y_e and Z_e , which are of the form

$$Y_e = \frac{\sum_{n=1}^3 w_n y_n}{\sum_{n=1}^3 w_n}$$

$$Z_e = \frac{\sum_{n=1}^3 w_n z_n}{\sum_{n=1}^3 w_n}$$

thus defining the emotionally-informed behavioural weighting vector B_e

$$B_e = \begin{bmatrix} X_e \\ Y_e \\ Z_e \end{bmatrix} \quad (12)$$

As this new vector is representing the behavioural impact of the new emotional state coming from the Matlab module, it will be used

as a weighting value for the agent displacement.

The application of X_e , and its psychological justification, has already been explicitly clarified. Y_e , being informed by the proximity of predators, is applied to the repulsive quality of predators, prompting the agent to 'run away' from danger. Lastly, the variable Z_e , representing the current levels of pleasure, elation and ecstasy felt by the agent, acts as a retardant force against the overall output vector - the more pleasurable the agent's current situation, the less likely it is to want to change that situation. These are mathematically indicated in the context of the dir direction vector.

The dir direction vector is calculated from each of the prey and predator positions, prey consisting of an attraction att compound and predators having a repulsing effect rep on the dir vector:

$$\vec{dir} = (X_e \vec{att} + Y_e \vec{rep}) \times (1 - Z_e)$$

with:

$$\vec{att} = \sum_{f \in F} (\vec{Pos}_f - \vec{Pos}_{agent})$$

$$\vec{rep} = \sum_{p \in P} (\vec{Pos}_{agent} - \vec{Pos}_p)$$

F being the set of prey (food) and P the set of predators. This gives us the new displacement vector for the agent and allows us to update each of the individuals position in the world.

The next step is to compute the future stimulus input vector which is to be sent back to Matlab as a new "input.txt" file. As has been said, emotion is directly linked to the agent's local situational stimuli, *i.e.* the positions of local prey and predators. Considering a local situation with two prey (food individuals) f_1 and f_2 , and two predators p_1 and p_2 , we can compute the new emotional output vector compounds, X representing the removal of positive stimuli ($\sum\{\$+\}$), Y the addition of negative stimuli ($\sum\{S-\}$), and Z both the removal of negative stimuli and addition of positive stimuli ($\sum\{S+, \$-\}$), in line with figure 1 and equations given in section 3.

In figure 12, the agent has two individuals in his viewing range, one of each type, and

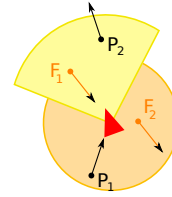


Fig. 12. Example of an agent local situation

TABLE 6
Representation of the local situation

	p_1	p_2	f_1	f_2	Total
$\delta dist$	-3	+2	-3	+2	
weight	0.5	1	1	0.5	
$X = \sum\{\$+\}$	-	-	-	1	1
$Y = \sum\{S-\}$	1.5	-	-	-	1.5
$Z = \sum\{S+, \$-\}$	-	2	3	-	5

two other individuals in his hearing radius. For each of these individuals, we first calculate the $\delta dist$ value which corresponds to the difference between the current distance and the previous one. Therefore, if this value is negative, it means that the corresponding individual is coming closer to the agent, and *vice versa*.

Before computing the three stimulus compounds, we need to consider using weighting values corresponding to the difference between the individuals detection areas. In this example, 1 is given for individuals in the viewing area while 0.5 is for prey and predators located only in a hearing range. This means that a viewed predator for instance is considered twice as important as a heard one. These weightings, of course, are situationally subjective and can be amended depending upon the importance one wishes to place upon specific sensory stimuli. The weightings included in table 6 are provided for illustrative purposes only.

Finally, we compute the three emotional compounds by taking the two last values for each individual, with respect to these rules:

$$X = \sum_{f \in F} \delta dist_f \quad \text{if } \delta dist_f > 0$$

$$Y = \sum_{p \in P} \delta dist_p \quad \text{if } \delta dist_p < 0$$

$$Z = \begin{cases} \sum_{f \in F} \delta dist_f & \text{if } \delta dist_f < 0 \\ \sum_{p \in P} \delta dist_p & \text{if } \delta dist_p > 0 \end{cases}$$

5 TESTING AND ANALYSIS

The system was tested several times for stability. Generally, stability tests were simple, single-predator, single-prey environments. Such environments being explored at length in the context of Millenson already [13], however, this work directed its attention specifically towards more complex environments with multiple predators and multiple prey. Two such tests are discussed and analysed at length.

5.1 Testing

The testing environment for Tests 1 and 2 had several uniformities. The agent began each simulation with a given amount of energy, which decreased at a constant rate over time; the maximum speed of predators and prey was consistent across both experiments; the maximum speed of the agent was consistent across both experiments; each simulation began with five predators and twenty prey present within the environment; the relative size of the environment was consistent across both experiments; the starting position and orientation of each of the predators and prey were determined randomly for both simulations; the maximum rate of change of orientation of the agent was consistent for both simulations.

Log files of the elements of stimulus vector J were recorded, along with the elements of the emotional state vector E_M and the emotionally-influenced behaviour vector B_e . In addition, the agent's orientation and displacement were recorded. In order to grant some measure of understanding regarding the agent's behaviour, graphs of J , B_e , and a graph of the agent's rate of change of orientation $\delta\theta$, are provided in the subsequent behavioural analyses in an effort to demonstrate the urgency the agent felt at any given moment to alter course.

5.2 Analysis

5.2.1 Experiment 1

Figure 13 illustrates the positions and orientations of the agent, predators and prey at the start of this experiment; figure 14 indicates the stimulus vector J ; figure 15 indicates the emotional behaviour vector B_e ; and, figure 16

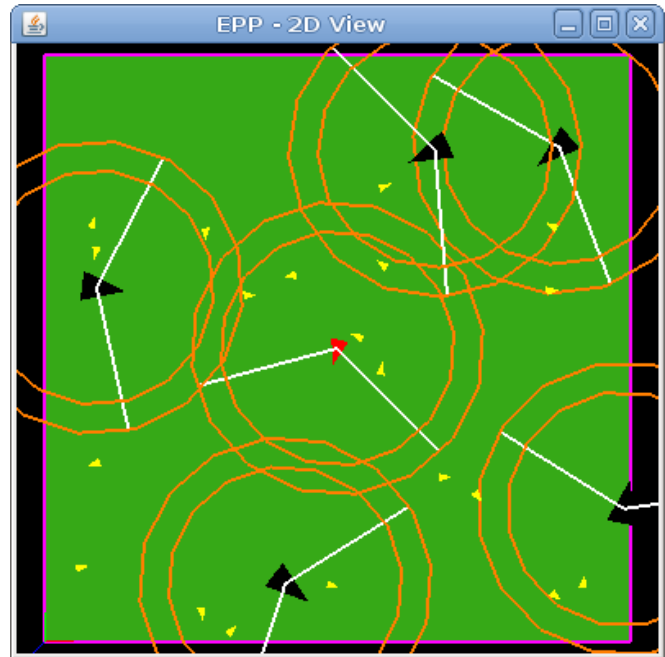


Fig. 13. Experiment 1: Starting Positions and Orientations

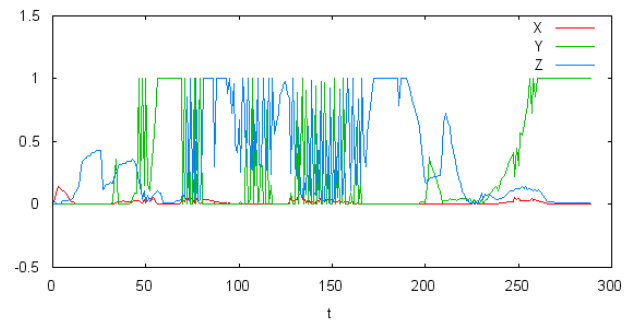


Fig. 14. Experiment 1: Stimulus Vector J over time

indicates the rate of change of orientation, $\delta\theta$, all over time. Note that each unit of time equates to 0.2s, the testing refresh rate of the system.

The agent begins this run with nothing present in its field of vision, and five prey within alternative sensory detection range. This prompts a hard about-face from the agent, at maximum magnitude of $\delta\theta$, shown in figure 16 as the plateau from $t = 0$ to $t = 20$. It is also noted that during this process, the fact that the agent is initially forced to move away from the prey in order to make a pursuit turn increases his anger, which then trails off and pleasure increases as he begins to close the gap and

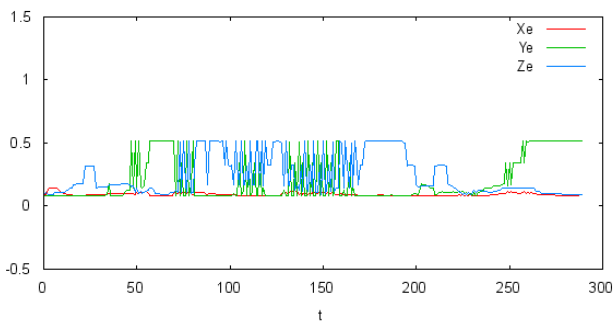


Fig. 15. Experiment 1: Emotional Behaviour Vector B_e over time

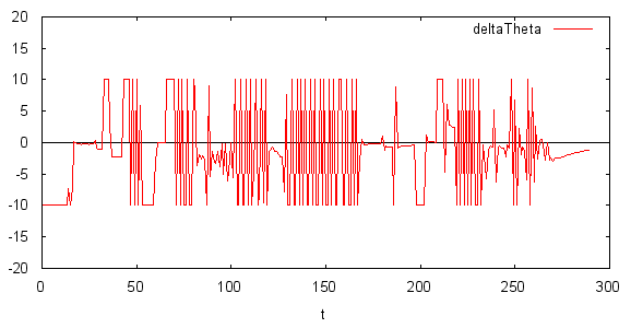


Fig. 16. Experiment 1: Rate of Change of Orientation $\delta\theta$ over time

briefly spikes as he consumes one of his prey.

The initial spike in Y , applied negative stimulus shown in figure 14, indicates the pursuit path of the agent leading to detection by a predator. This forces the agent to alter course slightly, pursuing two prey but abandoning the third; this causes an increase in all emotion-informing stimuli, again, in line with expectations.

The next major spike in applied negative stimulus occurs when a second predator encounters the agent, at $t = 46$. As figure 14 illustrates, through the highly variable nature of the anxiety-inducing experience, two predators are both hovering in the edge of the agent's vision. Figure 15 indicates the impact this has on agent behaviour, increasing the repulsive weighting of predators along similar lines, and figure 16 indicates the agent's path darting from side to side, each time attempting to escape a predator only for the other to move into its field of vision. This could be considered as simulating terror-induced confusion.

As both predators fall in to pursue the agent,

it reacts by running between them, outpacing their advance. This path also decreases the distance, albeit slightly, between the agent and his prey, indicated by the increased application of positive stimulus shown in figure 14, which is then reinforced by the removal of negative stimulus caused by his outpacing the predators.

The noisy, but comparatively steady turn indicated just prior to $t = 100$ on figure 16 marks the agent's pursuit course of his prey, moderated by the proximity of the nearest predator. The agent takes a sweeping path, slightly increasing the removal of positive stimulus, as shown by the increase in X in figure 14, the result of which increases the attraction of the prey, leading the agent to make an increase in rate of change of orientation to make a run towards them. This tactic succeeds, allowing the agent to devour two more prey by $t = 160$, and a third by $t = 190$. This, combined with the fact his path has let him make ground on the pursuing predator, gives significant reinforcement to the Z stimulus component, as shown in figure 14.

Shortly after $t = 200$, another predator begins pursuing the agent. The introduction of this predator on the agent's flank, outside his field of vision, prompts the slight increase in Y shown in figure 14 at this time. The sharp indecision in angle seen in the spiking oscillations in figure 16 circa $t = 220$ to $t = 240$ are caused by conflicting desires to escape the predator, and consume the one remaining prey of the small party the agent began the simulation chasing.

The increase in pleasure seen prior to $t = 250$, however, is caused by another entire school of prey swarming into the agent's field of vision. These five prey, which soon merge with another four, distract the agent entirely from the lone prey he had been pursuing, allowing it to escape.

Unfortunately for our agent, following this flock of prey leads him directly into the path of another predator, travelling in exactly the opposite direction. This causes the Y spike and plateau shown in figure 14 at $t = 260$. Unfortunately for our agent, the number of prey beyond the predator is too great a temptation,

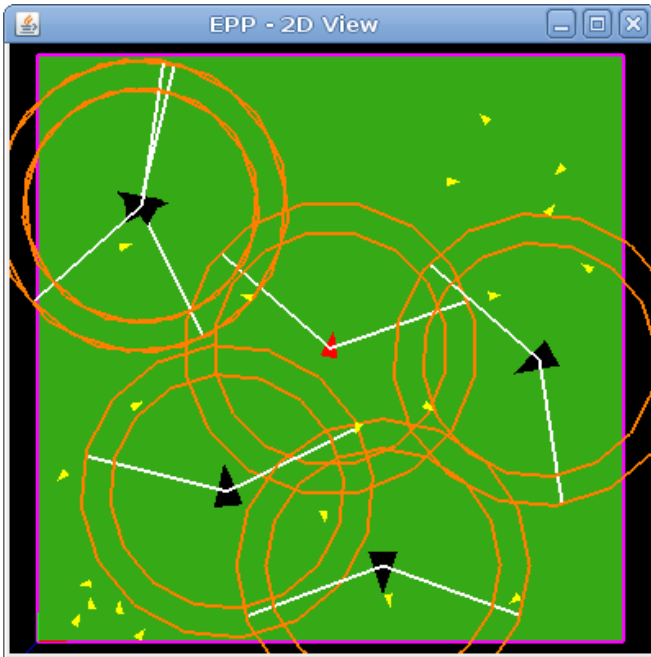


Fig. 17. Experiment 2: Starting Positions and Orientations

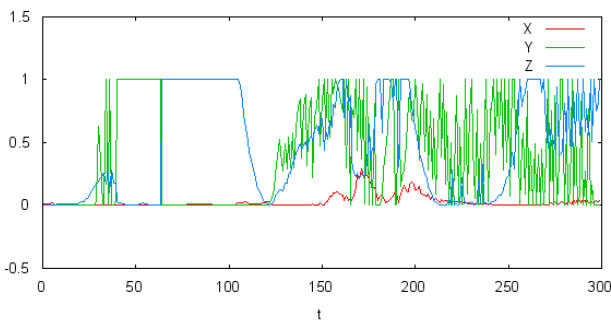


Fig. 18. Experiment 2: Stimulus Vector J over time

and while he makes some attempt to skim past the predator, he fails and is devoured, ending the simulation.

5.2.2 Experiment 2

Figure 17 illustrates the positions and orientations of the agent, predators and prey at the start of this experiment; figure 18 indicates the stimulus vector J ; figure 19 indicates the emotional behaviour vector B_e ; and, figure 20 indicates the rate of change of orientation, $\delta\theta$, all over time. Again, it should be noted that each unit of time equates to 0.2s, the testing refresh rate of the system.

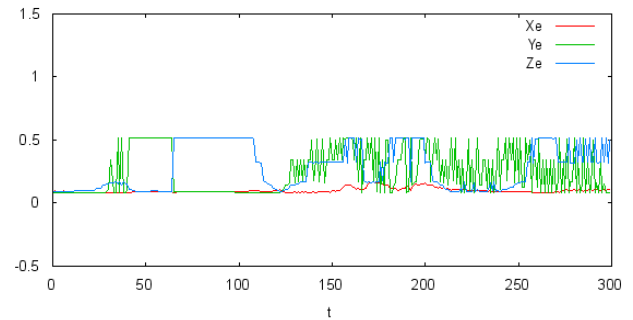


Fig. 19. Experiment 2: Emotional Behaviour Vector B_e over time

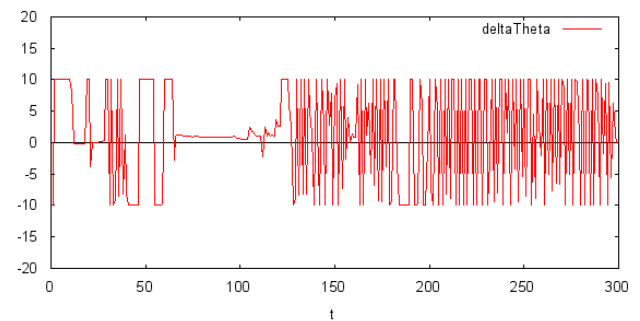


Fig. 20. Experiment 2: Rate of Change of Orientation $\delta\theta$ over time

In this simulation, the agent begins with no predators in detectable range, and two prey, one nearer than the other. This leads the agent to make a sharp turn until $t = 10$, as shown in figure 20. The abandonment of one prey in favour of another causes a slight increase in the X variable of J that we consider to be mild annoyance, but it quickly tails off as the agent focusses on its pursuit.

At $t = 20$, the agent begins to close distance with the prey, leading to an expected increase in Z as shown in figure 18. After $t = 25$, however, figure 18 displays a spike in Y , followed in quick succession by another two. This is caused by two predators entering the agent's field of vision near-simultaneously, leading to the same terror-induced confusion response discussed in the previous experiment.

In this state of confusion, the agent turns away from one predator towards another, then veers in a swooping escape trajectory, managing to come up on the tail of the prey he was chasing, which by now has been joined by a second. This daring manoeuvre explains

the fear plateau from $t = 40$ to $t = 60$. The subsequent Z spike and plateau, caused by the urgent reduction of negative stimulus (escape from the predator) and application of positive stimulus (closing on the prey) might be considered ecstatic relief at the gambit paying off. Also, we note the steady, carefully modulated course the agent follows during this pursuit, as shown in figure 20.

This strong elation-related response tails off by $t = 120$, the prey altering their course to cross the path of the predators, thus forcing the agent to take a wider arc and increase distance from them, leading to the wavering X increase shown in figure 18; again, mild annoyance.

We witness, however, a steady increase in Z from $t = 125$ to $t = 155$ brought about by a school of new prey entering the agent's field of vision and closing with them. Parallel to this, we see an increase in Y , as the path of pursuit of the prey also leads to a decrease in range from the remaining pursuing predator.

By $t = 170$, the complexity of the environment, and the sheer number of potential prey, combined with the ever-present threat of the predator forcing course alterations, has prompted a decrease in Z as the agent is unable to close distance with any significant number of prey, nor open up distance with the pursuing predator, and an increase in X as the agent is forced to ignore an increasing number of prey just to avoid the predator.

Circa $t = 190$, the agent moves to an extremely close distance with the single prey it has managed to keep tracking, prompting a Z response, only to be denied its meal by the arrival of yet another predator which generates a spike in Y and forces the agent to abandon its meal by turning away, generating an increase in X , as shown in figure 18.

His new course, however, brings the agent to track another two prey. Opening distance on the predator and closing with the prey, remembering that the agent has not eaten since the simulation began, increases Z . However, the path ultimately requires the agent to cut across the path of the predator, prompting spikes in Y .

At $t = 250$, our agent's perseverance is apparently rewarded, as a school of prey move

into view, on a trajectory that leads away from the predator, accounting for the increase in Z and sudden decrease in Y shown in figure 18 for this period. At the last, however, just Z makes a final peak and the agent is on the verge of consuming three prey at once, his energy reaches zero, and he dies of starvation.

6 CONCLUSION AND FUTURE WORK

In this article we have presented a CER-based model of emotions as applied to a motivational controller for a non-player agent in a game environment. The model has been shown to be psychologically grounded, and compatible with the general principles of the applied fuzzy inferencing system. We have outlined in detail the manner of the testing implementation, both in terms of representation and environment, and provided detailed accounts of two complex tests of the system.

The system we have presented, when applied, demonstrated informed, path-finding behaviour driven by an emotional component. While both tests included here ultimately ended in the termination of the agent, the purpose of the implementation was not to create an infallible controller; it was to create an emotionally informed controller capable of accounting for emotion within strategy. In that, the sometimes straightforward, sometimes bold, and sometimes conflicted nature of the agent's behaviour indicated in figures 13 through 20 is an identifier of success, rather than otherwise.

We note in particular the fact that both experiments clearly showed confusion prompted by an unwelcome decision - in the pedagogue case, the conflict caused upon the arrival of multiple predators from opposing angles, and the similar conflict caused by the decision between pursuing food or escaping danger.

In future work, it is planned that more detailed and complex experimentation be undertaken, both in terms of the length of the experiments and the complexity of the environment. In addition, a wider range of emotionally-informed controllers are expected to be explored and compared in terms of efficiency and behaviour.

Further to this, higher order fuzzy logic is to be explored, for effectiveness-comparison purposes and behavioural exploration. Lastly, it is expected that experimentation shall be conducted which guides multiple agents within the same environment, both comparative and antagonistic, using the emotionally-governed behavioural controller.

REFERENCES

- [1] H. Ahn and R. W. Picard, "Affective cognitive learning and decision making: The role of emotions." in *In The 18th European Meeting on Cybernetics and Systems Research*, 2006.
- [2] T. Archevapanich, B. Purahong, M. Klingajay, and P. Sooraksa, "Facial visualization for robotic indicator by using fuzzy emotional system," in *SICE-ICASE, 2006. International Joint Conference*, Oct. 2006, pp. 5634–5637.
- [3] W. F. Blewitt and A. Ayesh, "Modeling the emotional state of an agent through fuzzy logic with reference to the geneva emotion wheel," in *European Simulation and Modelling (ESM'2008) Conference, Le Havre, France*, 2008.
- [4] C. Gershenson, "Modelling emotions with multidimensional logic," in *Fuzzy Information Processing Society, 1999. NAFIPS. 18th International Conference of the North American*, 10-12 June 1999, pp. 42–46.
- [5] T. Shibata, T. Tashima, and K. Tanie, "Emergence of emotional behavior through physical interaction between human and robot," in *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on*, vol. 4, 10-15 May 1999, pp. 2868–2873vol.4.
- [6] K. Boehner, R. DePaula, P. Dourish, and P. Sengers, "How emotion is made and measured," *International Journal of Human-Computer Studies*, vol. 65, no. 4, pp. 275–291, Apr. 2007. [Online]. Available: <http://www.sciencedirect.com/science/article/B6WGR-4MVN0CB-1/2/b40616bd63b521d1270d6807171fc7e1>
- [7] G. Wang, Z. Wang, S. Teng, Y. Xie, and Y. Wang, "Emotion model of interactive virtual humans on the basis of mdp." *Frontiers of Electrical and Electronic Engineering in China*, vol. 2(2), pp. 156–160, 2006.
- [8] R. Muramatsu and Y. Hanoch, "Emotions as a mechanism for boundedly rational agents: The fast and frugal way," *Journal of Economic Psychology*, vol. 26, no. 2, pp. 201–221, Apr. 2005. [Online]. Available: <http://www.sciencedirect.com/science/article/B6V8H-4CCFBGP-1/2/804af79699585d758d6977f99da634f1>
- [9] N. Naqvi, B. Shiv, and A. Bechara, "The role of emotion in decision making: a cognitive neuroscience perspective." *Current Directions in Psychological Science*, vol. 15(5), pp. 260–264, 2006.
- [10] L. Perlovsky, "Integrated emotions, cognition, and language," in *Neural Networks, 2006. IJCNN '06. International Joint Conference on*, 16-21 July 2006, pp. 1570–1575.
- [11] M. Zeelenberg, R. M. A. Nelissen, S. M. Breugelmans, and R. Pieters, "On emotion specificity in decision making: Why feeling is for doing." *Judgement and Decision Making*, vol. 3(1), pp. 18–27, 2008.
- [12] J. R. Millenson, *The Psychology of Emotion: Theories of Emotion Perspective*. John Wiley & Sons, 1967, ch. 4, pp. 35–36.
- [13] W. Blewitt and A. Ayesh, "Implementation of millensons model of emotions in a game environment," in *AISB'09 Symposium: AI & Games, Edinburgh, UK*, 2009.
- [14] P. Ekman, "Are there basic emotions?" *Psychological Review*, vol. 99, pp. 550–553, 1992.
- [15] W. G. Parrott, *Emotions in Social Psychology*. Psychology Press, 2001.
- [16] C. Darwin, *The Expression of the Emotions in Man and Animals*, A. Introduction and C. by Paul Ekman, Eds. Harper Collins/Oxford University Press, 1872/1998.
- [17] R. Plutchik, *Emotion: Theory, research, and experience: Vol. 1. Theories of emotion*. New York: Academic, 1980, ch. A general psychoevolutionary theory of emotion, pp. 3–33.
- [18] K. R. Scherer, "What are emotions? and how can they be measured?" *Social Science Information*, vol. 44, no. 4, p. 695729, 2005.
- [19] P. Ekman, "An argument for basic emotions," *Cognition and Emotion*, vol. 6, pp. 169–200, 1992.
- [20] P. Ekman, *Handbook of Cognition and Emotion*. John Wiley & Sons, Ltd., 1999, ch. 3 - Basic Emotions.
- [21] W. Wundt, *Principles of physiological psychology.*, E. Titchener, Ed. New York, NY: Macmillan., 1904.
- [22] J. A. Russell, "A circumplex model of affect," *Journal of Personality and Social Psychology*, vol. 39, pp. 1161–1178, 1980.
- [23] A. Ayesh, J. Stokes, and R. Edwards, "Fuzzy individual model (fim) for realistic crowd simulation: Preliminary results," in *Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007. IEEE International, London, 2007*, pp. 1–5.
- [24] J. B. Watson, *Psychology. From the Standpoint of a Behaviourist*. Philadelphia: Lippincott, 1929.
- [25] J. B. Watson, *Behaviorism*. Chicago: University of Chicago Press, 1930.
- [26] J. B. Watson and W. MacDougall, *The Battle of Behaviorism: An Exposition and an Exposure*. W. W. Norton & Co, 1929.
- [27] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," *Current Psychology*, vol. Vol. 14, No. 2, pp. 261–292, 1996.
- [28] L. A. Zadeh, "Fuzzy Sets," *Information and Control*, vol. 8, pp. 338–353, 1965.
- [29] E. Mamdani and S. Assilian, "An experiment in linguistic synthesis with a fuzzy logic controller," *International Journal of Man-Machine Studies*, vol. Vol. 7, No. 1, pp. 1–13, 1975.
- [30] E. H. Mamdani, "Applications of fuzzy logic to approximate reasoning using linguistic synthesis," *IEEE Transactions on Computers*, vol. Vol. 26, No. 12, pp. 1182–1191, 1977.
- [31] L. A. Zadeh, *Fuzzy Sets, Fuzzy Logic and Fuzzy Systems*, G. J. Klir, Ed. World Scientific Press, 1996.
- [32] L. A. Zadeh, "Outline of a new approach to the analysis of complex systems and decision processes," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. Vol. 3, No. 1, pp. 28–44, 1973.
- [33] Sivarao and P. Brevern, "Mamdani fuzzy inference system modeling to predict surface roughness in laser machining," *International Journal of Intelligent Information Technology Application*, vol. 2(1), pp. 12–18, 2009.
- [34] A. Ayesh, "Perception and emotion-based reasoning: A connectionist approach," in *Informatica*, vol. 27, 2003, pp. 119–126.
- [35] C. W. Reynolds, "Flocks, herds and schools: A distributed behavioral model," in *SIGGRAPH '87: Proceedings of the*

14th annual conference on Computer graphics and interactive techniques, 1987, pp. 25-34.

Ce document a été réalisé en L^AT_EX, à l'aide des logiciels
Kile, Okular, Inkscape et Gimp

Dessin en couverture disponible à l'adresse :
<http://fav.me/d15t3b>
(Traduction : « *Je peux toujours te faire sourire* »)

© Karim MAHBOUB
2012, Université du Havre