



HAL
open science

Modélisation des mécanismes émotionnels pour un robot autonome : perspective développementale et sociale

Cyril Hasson

► **To cite this version:**

Cyril Hasson. Modélisation des mécanismes émotionnels pour un robot autonome : perspective développementale et sociale. Informatique. Université de Cergy Pontoise, 2011. Français. NNT : . tel-00904481

HAL Id: tel-00904481

<https://theses.hal.science/tel-00904481>

Submitted on 14 Nov 2013

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 DE CERGY-PONTOISE

Ecole Doctorale STIC
SCIENCES ET TECHNOLOGIES DE L'INFORMATION
ET DE LA COMMUNICATION

THÈSE

pour obtenir le titre de

Docteur en Sciences

de l'Université de Cergy-Pontoise

Mention : INFORMATIQUE

Présentée et soutenue par

Cyril HASSON

**Modélisation des mécanismes émotionnels pour un
robot autonome : perspective développementale et
sociale**

Thèse dirigée par Philippe GAUSSIER

préparée au laboratoire ETIS équipe neurocybernétique

dans le cadre du projet Européen FEELIX GROWING

soutenue le 22 septembre 2011

<i>Directeur :</i>	Philippe GAUSSIER	-	Professeur des Universités, ETIS Université de Cergy-Pontoise
<i>Rapporteurs :</i>	Frédéric ALEXANDRE	-	Directeur de recherche, LORIA INRIA Lorraine
	Philippe TARROUX	-	Professeur des Universités, LIMSI Université Paris Sud
<i>Examineurs :</i>	Benoît GIRARD	-	Chargé de recherche, ISIR UPMC
	Pierre-Yves OUDEYER	-	Chargé de recherche, FLOWERS INRIA Bordeaux Sud-Ouest
	Catherine PELACHAUD	-	Directeur de recherche, LTCI TELECOM Paristech CNRS

Remerciements

Todo

Résumé : L'objectif de cette thèse est de s'inspirer de la neurobiologie pour modéliser les mécanismes émotionnels de bas niveau sur un robot évoluant en environnement réel. Ce travail présente un modèle des émotions cohérent avec les données expérimentales décrivant le fonctionnement des structures cérébrales principales impliquées dans les mécanismes émotionnels. Les émotions jouent un rôle capital aussi bien pour la régulation du comportement des êtres humains que des animaux. En accord avec la vision darwinienne, les émotions sont vues comme des mécanismes adaptatifs favorisant la survie. Cependant, leur organisation autour de signaux essentiellement positifs et négatifs leur donne un caractère dimensionnel. Notre modèle considère les émotions comme le résultat de la dynamique d'interactions entre deux systèmes permettant l'évaluation des interactions avec l'environnement physique d'une part et l'environnement social d'autre part. Cette approche bioinspirée des émotions permet de donner aux robots une mécanique de base pour construire leur autonomie comportementale et leurs capacités de communication. Dans cette thèse, nous montrons qu'elles permettent autant de s'adapter aux caractéristiques de l'environnement que de servir de support à une communication non verbale. L'approche biomimétique de notre travail se traduit en termes méthodologiques par l'utilisation de réseaux de neurones formels pour les architectures de contrôle du robot mais aussi en termes fonctionnels par l'organisation de ces réseaux comme modèles de différentes structures du cerveau et de leurs interactions (amygdale, accumbens, hippocampe et cortex préfrontal). Suivant le courant *animat*, le robot est vu comme un animal aux besoins vitaux satisfaits par les ressources de son environnement. Les expérimentations seront illustrées sur des comportements de navigation reposant sur les apprentissages de conditionnements visuo-moteurs (stratégie visuelle) et sur l'intégration de chemin (stratégie proprioceptive).

Les conditionnements associant les signaux nocicepteurs et hédoniques aux autres informations sensorielles ou aux actions du robot sont à la base des régulations émotionnelles. Les prédictions que forme le robot lui permettent d'apprendre des comportements aversifs ou appétitifs en réponse à ses anticipations de "douleur" ou de "plaisir". Il peut aussi monitorer ses prédictions afin d'évaluer l'efficacité de ses comportements. C'est ce qui lui permet de réguler ses motivations et de sélectionner ses stratégies (navigation visuelle ou proprioceptive) et ses buts (ressources de l'environnement) de façon à satisfaire au mieux son équilibre interne en fonction de son environnement. Cette utilisation de signaux bas niveau positifs et négatifs permet de construire un modèle émotionnel minimal assurant au robot une autonomie comportementale.

Dans un deuxième temps, nous utilisons l'expressivité émotionnelle comme base à une communication avec le robot. Une tête mécanique permet au robot d'exprimer ses émotions grâce à ses expressions faciales. Cette communication consiste à donner au robot des signaux de récompense et de punition. Nous avons développé un modèle permettant de construire de manière autonome ces signaux d'interaction en leur donnant leur valeur émotionnelle. Cet échange d'informations avec le robot lui permet d'apprendre à valuer son environnement ou son comportement et ainsi d'apprendre interactivement à résoudre ses problèmes de navigation.

La modélisation des mécanismes émotionnels présentées dans cette thèse permet d'aborder aussi bien les questions de robotique autonome que d'interactions Homme-Machine. Plus largement, cette approche illustre l'intérêt de placer la robotique au coeur des sciences cognitives grâce à l'éclairage que permet l'analyse des comportements rendus possibles par des architectures neuronales relativement simples.

Mots clés : Robotique autonome, modélisation des émotions, inspiration neurobiologique.

Modeling emotional mechanisms for an autonomous robot : a developmental and social perspective

Abstract : The objective of this thesis is to draw inspiration from the neurobiology to model low level emotional mechanisms on a robot evolving in real environment. This work presents an emotional model coherent with experimental data describing the functioning of the cerebral structures involved in emotional mechanisms. Emotions play a central part in the regulation of behavior of humans as well as animals. In agreement with the darwinian view, emotions are seen as adaptive mechanisms enhancing survival. However, their organization around essential positive and negative signals gives them a dimensional flavor. Our model considers emotions as the result of the interaction dynamics between two systems. These systems allow the evaluation of the interactions with the physical and the social environment. This bio-inspired approach of emotions gives robots a basic framework to construct their behavioral autonomy and their communication skills. In this thesis, we show that they allow the robot to adapt itself to the characteristics of the environment as well as they underlie non verbal communication. The bio-mimetic approach of this thesis is reflected in methodological terms by the use of artificial neural networks for robot control architectures but also in functional terms by the organization of these networks as models of different brain structures and their interactions (amygdala, accumbens, periaqueductal grey, hippocampus, prefrontal cortex). Following the animat paradigm, the robot is seen as an animal which vital needs are satisfied by the resources of the environment. Experimentation are conducted on navigation behaviors relying on visuo-motor conditionings (visual strategy) and on path integration (proprioceptive strategy).

Conditionings between nociceptive or hedonic signals and other sensory information or actions of the robot are at the basis of emotional regulation. The robot predictions allow it to learn aversive or appetitive behavior in response to its "pain" or "pleasure" expectations. The robot can also monitor its predictions to assess the effectiveness of its behaviors. This enables it to regulate its motivations and select its strategies (visual navigation or proprioceptive) and goals (environmental resources) in order to best meet its internal balance depending on its environment. This use of low level positive and negative signals allows to build a minimal emotional model providing autonomy to the robot behavior.

In a second step, we use the emotional expressiveness as the basis for communication with the robot. A mechanical head enables it to express its emotions through its facial expressions. This communication consists in giving the robot reward and punishment signals. This exchange of information with the robot allows it to learn to value its environment or its behavior and thus to learn interactively to solve its navigation tasks.

The model of emotional mechanisms presented in this work allows to investigate issues of autonomous robotics as well as issues of Human-Robot interactions. Moreover, this approach shows the interest of putting robotics at the heart of cognitive sciences due to the perspective given by the analysis of robot's behaviors supported by relatively simple neuronal architectures.

Keywords : Autonomous robotics, emotional modeling, neurobiological inspiration.

Table des matières

1	Introduction	1
1	Robotique comportementale	2
2	Inspiration neurobiologique et systèmes dynamiques	3
3	Les émotions et la robotique	5
3.1	Les émotions pour l'autonomie	5
3.2	Les émotions dans les interactions Homme-Robot	7
4	Le robot et son environnement	9
5	Plan de la thèse	10
2	Emotions : le point de vue des psychologues	13
1	Spinoza, Spencer, Wundt et les modèles dimensionnels	13
2	Descartes, Darwin et les émotions basiques discrètes	17
3	De la question de la séquence à la question de l'évaluation cognitive	22
4	Discussion	26
3	Neurobiologie des émotions	29
1	Historique	29
1.1	Cannon et Bard : l'interaction entre le cortex et l'hypothalamus	29
1.2	Le circuit de Jakob-Papez	30
1.3	Les trois cerveaux de MacLean	31
2	Les principales structures cérébrales	33
2.1	Structures du tronc cérébral	34
2.2	Les structures sous-corticales	36
2.3	Le cortex préfrontal	43
3	Neurobiologie de la douleur et du plaisir	46
3.1	Douleur	46
3.2	Plaisir	48
4	Discussion	48
4	Modélisation neuronale et navigation	53
1	Introduction	53
2	Modélisation neuronale	54
2.1	Neurones et synapses	54
2.2	Apprentissage associatif hebbien	56
2.3	Apprentissage de conditionnements : Widrow et Hoff et méthode des moindres carrés	58
2.4	Processus de catégorisation et compétition	59
3	Navigation visuelle sensori-motrice	62
3.1	Cellules de lieu visuelles	62
3.2	Conditionnements Lieu-Action	65
4	Intégration de chemin et navigation proprioceptive	67

5	Douleur, plaisir, motivations et conditionnements	73
1	Introduction	73
2	Physiologie artificielle, douleur, plaisir et motivation	74
3	La peur comme un mécanisme de conditionnement négatif	78
3.1	Principe et architecture	78
3.2	Expérimentation robotique	80
4	Hédonisme et motivations	83
4.1	Motivations et navigation sensori-motrice	83
4.2	Motivations et intégration de chemin	87
5	Discussion	96
6	Automonitoring et frustration	99
1	Introduction	99
2	Prédictions des buts	100
2.1	Prédiction visuelle du but	101
2.2	Prédiction proprioceptive du but	101
3	Détection de l'échec	102
4	Inhibition et frustration	105
5	Expérimentations robotiques	111
6	Discussion	115
7	Surprise et nouveauté	121
1	Introduction	121
2	Modèle neuronal de l'analyse d'un prédicteur	123
3	Tests unimodaux	127
3.1	Perturbations statiques de la vision	128
3.2	Navigation et changement d'environnement	128
4	Interprédiction et détection de conflits sensoriels	130
4.1	Les conflits sensoriels chez l'homme	130
4.2	Les conflits sensoriels pour le robot	132
4.3	Test du modèle d'interpredictibilité	134
5	Discussion	138
8	Les Interactions émotionnelles sources de sens	141
1	Introduction	141
2	Toute stimulation peut prendre un sens émotionnel	143
2.1	Principe et fonctionnement	143
2.2	Expérimentation robotique	146
2.3	Point de vue neurobiologie	149
3	L'expressivité faciale et la reconnaissance d'émotion	151
4	Vers un référencement social	152
4.1	Couplage comportemental	156
4.2	Couplage environnemental	157
5	Discussion	161

9 Conclusion	163
1 Résumé des travaux	163
2 Discussion et Perspectives	165
Bibliographie	169

Introduction

La science-fiction fait souvent le portrait des robots comme étant des machines dotées de capacités intellectuelles mais dénuées de la moindre trace d'émotion. Jusque récemment, cette description était tout à fait pertinente si l'on s'en réfère aux robots existants. Les robots développés depuis le début des années 60 correspondent à cette vision. Principalement utilisés dans le monde industriel, ils ont été conçus pour effectuer de manière pré-programmée des tâches répétitives dans des environnements parfaitement contrôlés. Il s'agit de l'âge d'or de l'intelligence artificielle classique issue de la logique formelle. Cette conception des robots peut être illustrée par la position intellectuelle traditionnelle qui consiste à faire l'éloge de la raison sur les émotions. Le raisonnement est l'outil permettant d'accéder à la vérité et les émotions ne peuvent que perturber la neutralité et donc l'exactitude de celui-ci. Cependant, ces techniques classiques se révèlent inadaptées dans les environnements non contrôlés et dynamiques de la vie quotidienne. Les techniques classiques ne permettent pas aux robots d'adapter leurs comportements à un environnement inconnu et changeant. C'est pourquoi, depuis le milieu des années 90, cette vision dominante de la robotique est bousculée une approche plus comportementale et dynamique de la programmation des robots (robotique comportementale [Brooks 1986, Arkin 1998], approche Animat [Wilson 1991]) et d'autre part par un changement de point de vue vis-à-vis du rôle des émotions. Les psychologues, les éthologues et les biologistes ont permis de mieux comprendre le fonctionnement et les bases neurobiologiques des émotions ainsi que leur rôle dans le bon fonctionnement de la cognition. Au lieu d'être considérées comme des obstacles à surmonter pour permettre un raisonnement efficace et une action optimale, les émotions sont placées au coeur même de la régulation du comportement. Le creuset des sciences cognitives a favorisé la circulation de ces idées et tout un domaine de recherche est aujourd'hui axé sur l'étude des émotions dans des systèmes artificiels. Ce domaine est connu sous le nom d'**informatique affective** (*affective computing*). Comme Rosalind Picard le décrit dans son ouvrage du même nom [Picard 1997] :

«L'informatique affective est différente de la présentation d'une théorie des émotions ; cette dernière se concentre généralement sur la nature des émotions humaines, comment et quand sont-elles générées, et quels sont leurs effets. L'informatique affective s'étend à l'implémentation des émotions et peut donc aider au développement et au test de nouvelles mais aussi d'anciennes théories des émotions. Cependant, l'informatique affective comprend aussi bien d'autres choses, comme le fait de donner à un ordinateur la capacité de reconnaître et d'exprimer des émotions, de développer ses capacités à répondre intelligemment aux émotions humaines et lui permettre de réguler et d'utiliser ses émotions.»

Les travaux de Reeves et Nass présentés dans [Reeves 1996] illustrent la tendance humaine à anthropomorphiser des objets même très rudimentaires et à percevoir le monde comme teinté d'émotions. C'est alors sans surprise que les machines "affectives" se révèlent tout à fait efficaces pour créer des interactions suscitant des réactions émotionnelles. Les robots autonomes sont donc particulièrement intéressants pour l'étude des émotions dans la mesure où ils sont des dispositifs physiques qui se déplacent, prennent des décisions et résolvent des problèmes mais aussi parce qu'ils peuvent interagir avec leur environnement social.

Tout au long de cette thèse, nous aborderons donc la modélisation des mécanismes émotionnels pour un robot autonome avec un double objectif. D'un côté, nous explorerons le rôle des émotions en tant que mécanismes de régulation du comportement. Nous montrerons comment elles peuvent constituer une base à partir de laquelle un robot peut construire son autonomie d'action, son adaptabilité à l'environnement et ses capacités de communication. D'un autre côté, les robots, en tant que modèles simplifiés de systèmes cognitifs, nous serviront d'outils de test pour l'étude du fonctionnement neurobiologique des émotions et de leur émergence à partir de mécanismes de bas niveau comme l'apprentissage de conditionnements.

1 Robotique comportementale

Notre travail se situe dans le courant de la robotique comportementale [Brooks 1991, Arkin 1998]. Dans ce paradigme, issu de la biologie, de l'éthologie et des neurosciences, l'architecture de contrôle du robot est divisée en plusieurs boucles de processus fonctionnant en parallèle. Ces processus correspondent aux différents comportements que le robot doit maîtriser. Par exemple, les informations provenant de la vision peuvent être traitées et utilisées pour contrôler des comportements complexes comme la navigation ou la préhension d'objets. Il est cependant important d'introduire un mécanisme de réaction rapide permettant aux robots d'éviter les obstacles lorsqu'ils se déplacent. Ces différents processus exécutés en parallèle aboutissent à diverses commandes motrices dont certaines peuvent se révéler incompatibles. Leurs exécutions sur un même actuateur entrent alors en compétition. La fusion de ce flot asynchrone d'informations parallèles peut être effectuée grâce à un mécanisme de gestion des priorités. Dans l'architecture de subsomption par exemple (*subsumption architecture* [Brooks 1986]), une des premières architectures de l'approche comportementale, les comportements sont organisés en couches fonctionnant en parallèle. Chacune d'entre elles possède ses propres capacités de perception et d'action. Une couche peut ainsi inhiber les entrées sensorielles ou la sortie motrice d'une autre couche. La hiérarchie de priorité de ces mécanismes est déterminée par l'importance de chaque couche (comportement) du point de vue de la survie du robot dans un environnement donné. Par exemple, si le robot se rend à un emplacement déterminé de son environnement, la présence d'un danger à proximité prend le pas et un comportement d'évitement est produit. Dans ce cas, l'alerte liée à la présence d'un danger supprime l'activation du système de navigation. Le processus de prise de décision se caractérise alors par sa décentralisation et donc par une robustesse bien supérieure à celle des processus clas-

siques de type "perception-raisonnement-action". Si la nécessité se présente, n'importe quel processus de bas niveau peut prendre le contrôle du comportement du robot.

Un grand nombre d'architectures de contrôle du comportement et de stratégies de prise de décisions peuvent être utilisées (voir [Arkin 1998, Pfeifer 1999] pour un état de l'art). Certaines fonctionnent en gérant des objectifs internes ou des motivations permettant de contrôler le comportement du robot. Elles reposent toutes sur les principes suivants : des boucles sensation-action fonctionnant en parallèle permettant de faire émerger une perception orientée vers l'action, une prise de décision décentralisée et l'absence de représentations (ou alors l'usage de représentations minimales, grossières et distribuées). Cette philosophie de conception des architectures de contrôle robotique permet notamment d'éviter le problème de l'ancrage des symboles (*symbol grounding problem*) [Harnad 1990].

2 Inspiration neurobiologique et systèmes dynamiques

L'approche biomimétique de notre travail se traduit en termes méthodologiques par l'utilisation de réseaux de neurones formels pour les architectures de contrôle du robot mais aussi en termes fonctionnels par l'organisation de ces réseaux comme modèles de différentes structures du cerveau et de leurs interactions (amygdale, accumbens, substance grise périaqueducule, hippocampe et cortex préfrontal). Le but est de proposer un modèle des émotions cohérent avec notre compréhension de leurs bases neurobiologiques. Ainsi, il n'est pas question d'implémenter un modèle sophistiqué et rigide d'émotions complexes. Le but est au contraire de déterminer le matériel minimal permettant l'émergence des comportements émotionnels les plus simples (mais aussi les plus fondamentaux). Nous proposerons un modèle des émotions totalement en accord avec la vision du neurobiologiste Antonio Damasio. Damasio décrit les émotions comme émergeant depuis les signaux physiologiques les plus simples. Il utilise pour cela la métaphore d'un arbre conceptuel. Dans cet arbre, la verticalité est un gradient de complexité. Chaque niveau est construit sur le précédent. La physiologie forme le tronc : les réflexes basiques, la régulation métabolique et même les réponses immunitaires. A ce niveau cependant, ces mécanismes sont indifférenciés d'un point de vue émotionnel. Le schisme fondamental, la première apparition d'une valuation, intervient à partir des sensations de douleur et de plaisir. C'est sur ces sensations que sont construits les différentes motivations et les *drives* (motivations de bas-niveau comme la faim ou la soif). Les émotions à proprement parler seraient construites sur la base de ces drives et motivations. Ce point de vue permet de décrire les émotions comme des propriétés émergentes de l'abstraction progressive de traitement des informations corporelles. La figure 1.1 présente l'arbre des émotions de Damasio.

Les robots autonomes nous permettent d'étudier les émotions en tant que propriétés émergentes d'un système dynamique, sans avoir à modéliser explicitement un système émotionnel. L'absence d'un système émotionnel explicite n'implique pas que les émotions soient considérées comme un phénomène n'ayant d'existence que dans l'œil du spectateur. Mandler considère par exemple que les émotions peuvent être comprises comme des processus dynamiques intégrant de manière causale des processus

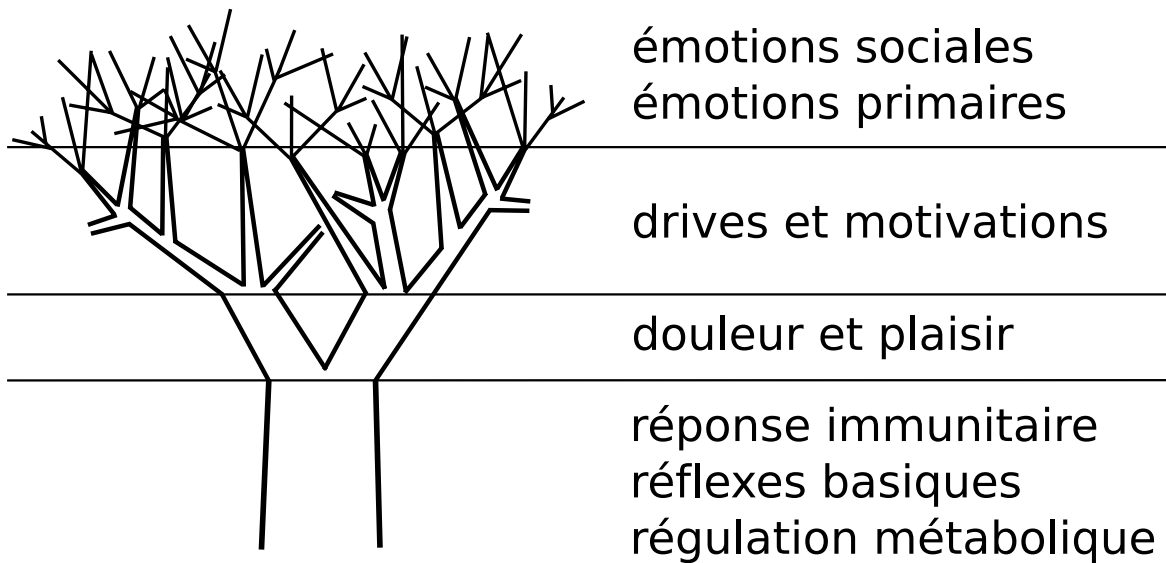


FIGURE 1.1 – A la base de cet arbre conceptuel, la physiologie du corps est à l'oeuvre. Une différenciation se fait à partir des sensations de douleur et de plaisir. Les drives et motivations reposent sur ces stimulations fondamentales. Enfin, les émotions primaires puis les émotions plus complexes sont construites à partir de ce niveau motivationnel. Schéma adapté depuis [Damasio 2003])

de plusieurs sous-systèmes liés entre eux [Mandler 1985] : la physiologie, la cognition, la communication et l'expérience subjective. Suivant Pfeifer, les émotions pourraient alors être vues comme «des constructions permettant de caractériser et d'identifier simultanément les états dynamiques complexes et causalement liés des (quatre) sous-systèmes» [Pfeifer 1991]. Dans ce travail nous ne parlerons pas d'expérience subjective ni de conscience chez les robots, mais nous nous inscrirons dans cette vision "système dynamique" des processus émotionnels défendu notamment par [Lewis 2005] afin de concilier les approches psychologiques et biologiques de la compréhension des émotions. L'état émotionnel d'un agent est caractérisé par un état spécifique de la dynamique de ce système en interaction avec le monde physique et social (voir figure 1.2)

Cette approche de la robotique donne une importance particulière aux robots de type créature artificielle, ou animat [Wilson 1991], qui entretiennent de vraies interactions en boucle fermée avec leur environnement (l'environnement influence leurs comportements et ceux-ci modifient l'environnement). Le robot étant conçu sur le modèle d'une créature vivante. Il est équipé d'une physiologie artificielle (e.g. taux de glucose, hydratation) créant des besoins (faim et soif) pouvant être comblés par les ressources simulées de son environnement (patchs de couleur au sol représentant nourriture et eau). Il doit être capable d'utiliser les informations provenant de la dynamique du fonctionnement de son corps. C'est via son corps que le robot interagit avec le monde physique et social [Cañamero 2001a]. Ces systèmes robotiques peuvent également apporter des informations précieuses sur les méthodes de modélisation utilisées : les propriétés émergentes induites par les effets de la dynamique de l'interaction

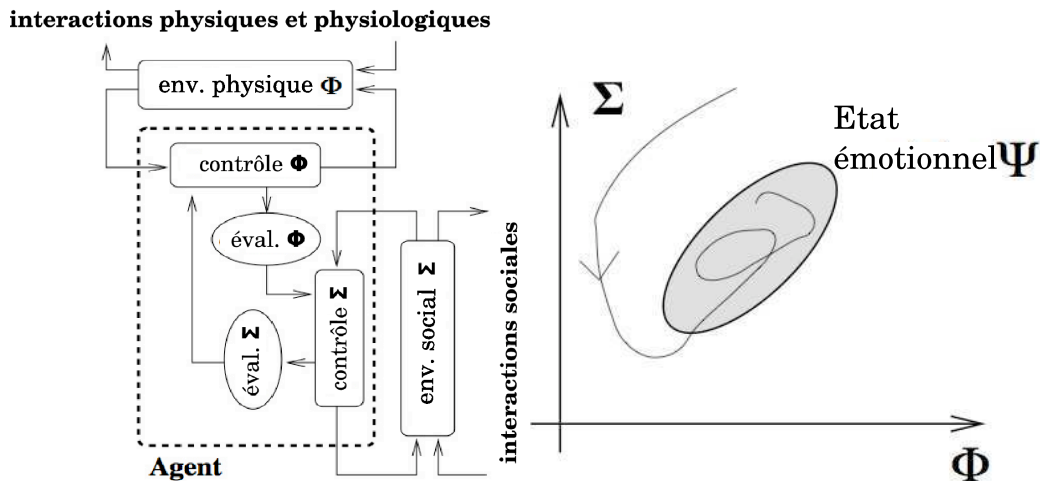


FIGURE 1.2 – On peut voir un agent comme étant constitué de deux contrôleurs, l'un dédié à réguler les interactions avec l'environnement physique et physiologique et l'autre aux interactions sociales. Un état émotionnel correspondrait à une dynamique particulière d'interaction entre ces deux systèmes. Φ : physique ; Σ : social ; Ψ : psychologique.

peuvent grandement simplifier le processus de conception mais aussi le modèle émotionnel lui-même. L'élaboration de tels systèmes émotionnels est un exemple de psychologie synthétique pour laquelle les robots apparaissent comme des outils efficaces pour simuler et étudier les effets dynamiques des modèles du cerveau.

3 Les émotions et la robotique

Les émotions peuvent intéresser les roboticiens à plusieurs titres grâce aux fonctions qu'elles remplissent en tant que mécanismes de régulation du comportement des animaux et des hommes confrontés à des problèmes d'autonomie, d'adaptation et de survie. Nous allons montrer dans ce travail que les émotions peuvent remplir principalement deux rôles majeurs. D'une part les émotions peuvent servir à amorcer l'autonomie d'un robot. Elles permettent de guider l'apprentissage, d'évaluer la qualité des comportements produits, d'orienter la sélection de l'action, de valuer les informations apprises en fonction de leur importance... D'autre part, les émotions jouent un rôle capital dans les interactions et plus précisément dans la communication. Elles permettent d'échanger des informations, mais aussi de susciter une réponse émotionnelle. Les humains sont habitués à utiliser le registre émotionnel. En communiquant de cette manière, les interactions Homme-Robot peuvent alors être améliorées.

3.1 Les émotions pour l'autonomie

Apprentissage autonome et valeur émotionnelle : l'apprentissage pour les robots autonomes repose habituellement sur des modèles de renforcement ou d'association et utilise des signaux externes en guise de récompense ou de punition (venant de l'environnement ou fournies par un "critique", suivant les algorithmes). D'autres

modèles utilisent l'apprentissage par imitation pour apprendre l'action réalisée par un autre agent (humain ou robot) généralement appelé le "démonstrateur". Une des difficultés principales sous-jacentes à tous ces modèles est la question de la signification de ces signaux pour le robot afin que le processus d'apprentissage puisse être plus autonome et mieux ancré dans l'architecture du robot. En d'autres termes, comment un robot peut-il donner un sens aux signaux qu'il perçoit plutôt que d'utiliser un signal de renforcement fourni par un "professeur?". Comment peut-il décider ce qu'il faut et ce qu'il ne faut pas apprendre? Un mécanisme ancré dans un "système de valeurs" interne est indispensable pour fournir les signaux internes quant aux qualités positives et négatives des actions et des sensations du robot. Ces signaux internes leur donnent un sens par rapport à ses valeurs, ses besoins et ses buts. Comme le décrit Wehrle [Wehrle 2001] : «l'ancrage implique d'une façon ou d'une autre que nous permettions au robot d'établir sa propre catégorisation émotionnelle en référence à ses propriétés physiques, à la tâche, aux propriétés de l'environnement ainsi qu'à l'interaction en cours avec son environnement». Au niveau le plus simple, les émotions permettent de catégoriser des stimuli comme étant plaisants ou déplaisants, bénéfiques ou nocifs.

Motivations et sélection de l'action : La catégorisation émotionnelle transforme des stimuli neutres en stimuli valués à poursuivre ou à éviter. Les robots dotés d'un système motivationnel qui établit et gère les priorités des "buts internes" afin de contrôler les processus de prise de décision disposent alors d'une bien meilleure autonomie que les robots ne faisant que réagir à la présence de stimuli externes. Cependant, dans certains environnements dynamiques changeant rapidement et pouvant présenter des menaces pour la survie du robot, la priorité des buts doit pouvoir être modifiée afin de répondre à un problème urgent (par exemple, pour éviter un danger imminent comme la proximité d'un prédateur) tout en inhibant le besoin courant (par exemple, trouver de l'eau). Selon ce point de vue, les émotions sont considérées comme des "mécanismes d'interruption" [Simon 1967]. Frijda illustre bien cette idée [Frijda 1995] : "Les émotions portent sur l'action (ou la motivation pour l'action) et le contrôle de l'action."

Comportements répétitifs et inefficaces : à cause de nombreux paramètres comme la prise en compte d'informations locales ou l'utilisation de senseurs et d'actuateurs bruités ou de faible qualité, les robots autonomes peuvent produire des comportements répétitifs aboutissant à des situations de "cul-de-sac" (*deadlocks*). Ces situations critiques empêchent aux robots l'accomplissement de leurs buts. Ce type de comportement a un coût pour le robot aussi bien en termes de temps que d'énergie. Le robot peut par exemple essayer de manière répétée d'attraper un objet alors que celui-ci est hors de portée. Il peut aussi chercher à rejoindre une position de l'environnement occupée par un obstacle. Un système de raisonnement lourd et coûteux nécessitant un modèle du robot, un modèle de l'environnement et un modèle de la tâche est rarement une solution satisfaisante. Une fois encore, un tel système serait très spécialisé et ne répondrait pas à la variabilité intrin-

sèque des environnements naturels particulièrement dynamiques. Un mécanisme de méta-contrôle (contrôle d'ordre supérieur) surveillant les performances du robot par rapport à ses buts et pouvant donc détecter les comportements inefficaces peut apporter une solution à ce type de problème. Un état interne comme l'ennui ou la frustration, émanant des situations d'échec persistant, doit provoquer une inhibition du but ou du comportement courant, permettant ainsi de sortir des *deadlocks*.

Surcharge cognitive : les robots, comme les humains, disposent de capacités de traitement de l'information limitées. Le dépassement de ces limites provoque une surcharge ou saturation cognitive pouvant entraîner des difficultés d'attention ou de mémoire, qui devient trop vaste et lente en termes de temps d'accès à des informations pertinentes. Les émotions peuvent fournir un critère permettant de sélectionner l'information pertinente aussi bien pour les mécanismes attentionnels que mnésiques. Les émotions peuvent servir à orienter la sélectivité des processus attentionnels : attention accrue pour les stimuli à forte valence émotionnelle [Burke 1992, Hulse 2006]. La gestion de la mémoire est aussi un problème majeur en robotique autonome. Si le robot ne possède pas les critères appropriés pour filtrer l'information, sa mémoire n'est alors pas assez spécifique et ceci entraîne des temps élevés pour le rappel d'informations. Un certain nombre de travaux en neurosciences montrent comment, chez l'homme, les émotions interagissent avec la mémoire aux différents stades du traitement de l'information : depuis l'encodage et la consolidation initiale jusqu'au rappel à long terme [LaBar 2006]. Le rôle de l'hippocampe ainsi que de l'amygdale dans ce cadre est notamment décrit dans [Richter-Levin 2004a]. Des mécanismes de mémoire sélective inspirés de la mémoire émotionnelle chez l'homme (par exemple, meilleur rappel de souvenirs liés à des émotions) ainsi que la notion apparentée de mémoire autobiographique pourraient résoudre certains de ces problèmes [Gadanh 2003], mais aussi apporter des réponses cohérentes à un large éventail de situations. Cette problématique de la surcharge cognitive se pose aussi pour l'apprentissage. Il faut sélectionner ce qui doit être appris pour ne pas trop apprendre et ainsi arriver à saturation des capacités cognitives.

3.2 Les émotions dans les interactions Homme-Robot

Tandis que la modélisation des émotions pour un robot seul est centrée sur l'aspect "interne" des émotions, la conception d'émotions pour des robots interagissant socialement est principalement basée sur les caractéristiques "externes" des émotions. L'expressivité émotionnelle est une composante clé de la communication et de l'interaction sociale dans la mesure où les manifestations externes des émotions peuvent jouer, à plusieurs niveaux, un rôle majeur en tant que mécanismes de "signalisation" [Arkin 2003, Rani 2006, Scheutz 2006]. Par exemple, l'expression émotionnelle d'un individu peut être utilisée par un autre en tant que référence sociale pour évaluer le type de situation à laquelle il est confronté ainsi que la réaction à adopter. Elle peut être contrôlée dans une certaine mesure et être utilisée intentionnellement pour permettre aux autres de connaître notre état émotionnel, de construire des attentes concernant notre propre comportement et ainsi d'adapter le leur en fonction. Les travaux sur les robots expressifs destinés à l'interaction avec les humains ont reçu une

attention croissante depuis le début des années 2000 [Cañamero 2001b, Breazeal 2002, Fukuda 2004, Gaussier 2004, Nadel 2006, Gaussier 2007b, Cañamero 2008]. Cela vient du fait que l'expressivité émotionnelle joue différents rôles majeurs dans les interactions sociales entre humains et robots.

Echanger des informations entre l'homme et le robot : nous comprenons les comportements que nous observons chez nos partenaires sociaux (humains ou artificiels) comme résultant de causes ou d'intentions déterminées. Nous pouvons alors élaborer des explications cohérentes pour interpréter leurs comportements passés mais aussi pour faire des prédictions et construire des attentes à propos de leurs comportements futurs. Les émotions sont interprétées comme étant à la sources de ces intentions. Les robots autonomes peuvent donc utiliser leur expressivité émotionnelle pour communiquer leurs intentions ou leurs besoins. Réciproquement, les robots pourraient interpréter les subtilités de notre expressivité pour obtenir des informations pertinentes concernant leur environnement ou leur comportement (expressivité faciale, prosodie, gestes, posture, ...). Nous pourrions alors attendre d'eux qu'ils "comprennent" ce que nous voulons dire et pas seulement ce que nous disons. A un niveau plus fondamental, cette communication permet aux humains d'accéder indirectement à l'état interne du robot. Il ne s'agit pas alors de comprendre les intentions du robot mais plus simplement de savoir comment celui-ci évalue sa situation. De son côté, le robot peut utiliser les informations venant de l'expressivité de l'humain pour valuer son comportement ou son environnement (l'humain jouant le rôle d'un référent social).

Susciter une réponse émotionnelle : nous réagissons habituellement aux émotions des autres par une réaction émotionnelle. Les émotions des robots peuvent donc servir à susciter une réponse émotionnelle. Par exemple, un robot exprimant de la tristesse en raison de son incapacité à accomplir une tâche donnée peut recevoir de l'aide de la part d'un humain "touché" par son état émotionnel. Plus pragmatiquement, les robots peuvent obtenir des informations sur leur environnement ou leur comportement depuis la reconnaissance de l'expressivité émotionnelle de leurs partenaires humains. La reconnaissance d'expressions émotionnelles positives (e.g. joie) ou négative (e.g. colère) permet alors au robot d'évaluer son propre comportement ou son environnement.

Améliorer le confort de l'interaction avec le robot : nous pouvons attendre des robots capables d'exprimer des émotions et d'adapter leurs comportement à l'état émotionnel de leurs partenaires sociaux qu'ils permettent aux humains de se sentir bien plus à l'aise en interagissant avec eux. Une raison évidente à cela est que cette interaction emprunte les canaux naturels de communication entre humains. Une autre raison importante tient à notre perception des robots. S'ils montrent une expressivité et des comportements émotionnels, les robots nous apparaîtront comme plus crédibles et plus proches de nous. Donner l'impression du vivant, la crédibilité et une interaction adaptée aux humains semblent être les caractéristiques principales que les robots expressifs devraient montrer pour que nous les acceptions en tant que "partenaires

sociaux".

4 Le robot et son environnement

Le travaux réalisés dans cette thèse ont été menés sur un robot mobile équipé de différents capteurs et doté d'une tête mécanique expressive. La figure 1.3 est une illustration schématique du robot et de ses composants principaux.

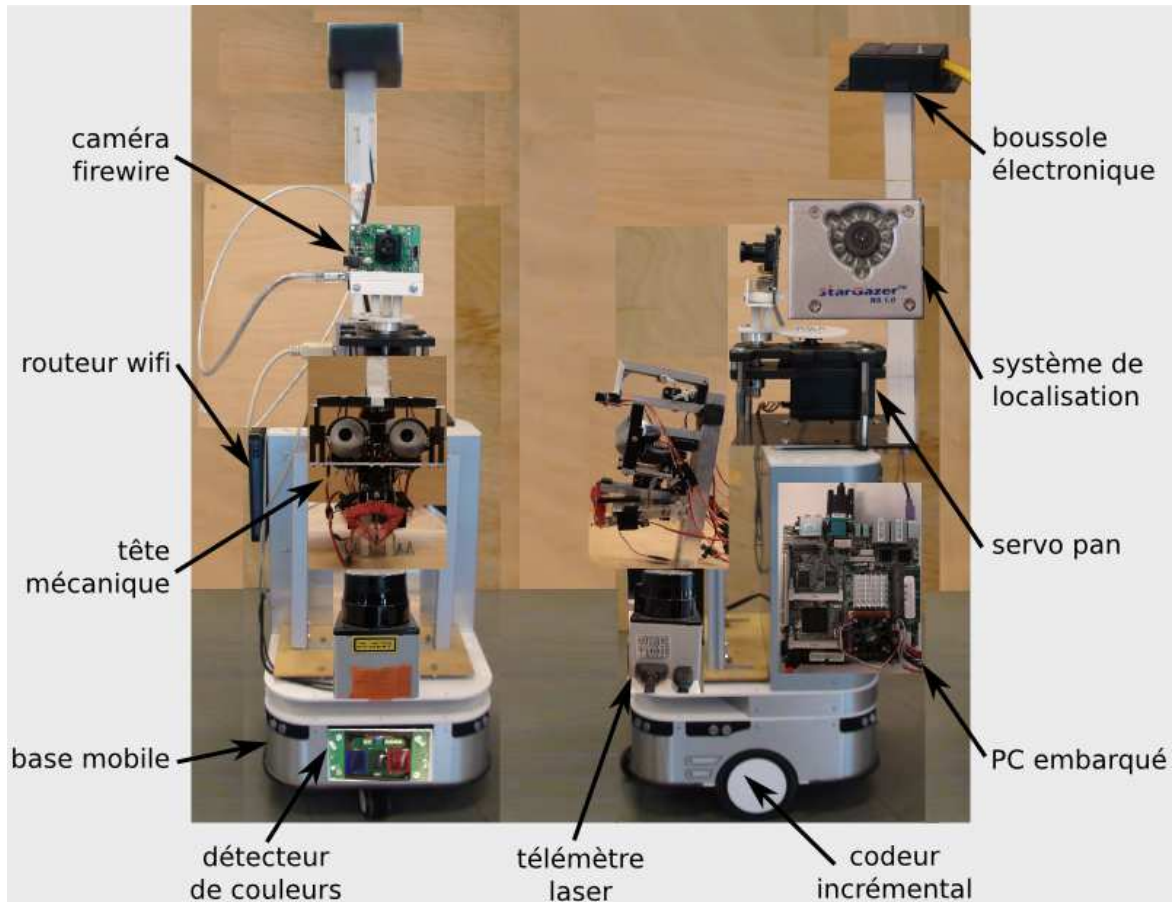


FIGURE 1.3 – Equipement embarqué sur le robot. Le détecteur de couleur est en réalité accroché sous le châssis et le PC embarqué est fixé dans le coffre. Le système de localisation n'est jamais utilisé par le robot. Il permet de récupérer avec précision les trajectoires du robot.

La base mobile permet au robot de se déplacer. Il s'agit de sa principale capacité d'action. C'est pourquoi tout au long de cette thèse, le robot est soumis à différentes tâches de navigation. Les moteurs sont équipés de codeurs incrémentaux permettant d'assurer une proprioception. Le robot embarque aussi un système de vision pan composé d'une caméra et d'un servomoteur permettant de balayer le panorama. La boussole électronique donne au robot une mesure d'orientation et remplace à faible coût une centrale inertielle. Le télémètre laser permet une détection efficace des obstacles et

constitue donc la "peau" du robot. Le capteur de couleur placé sous le robot lui permet de détecter les zones spécifiquement marquées de son environnement (les ressources simulées). La tête mécanique permet non seulement de produire des expressions faciales émotionnelles mais elle est aussi équipée d'une caméra permettant au robot de voir le visage de la personne interagissant avec lui. Le robot embarque un PC (carte Mini ITX équipé d'un processeur Dual Core cadencé à 2,2 GHz avec 2 Go de mémoire RAM) lui permettant de réaliser une partie des calculs de manière totalement autonome. Cependant, étant donné la complexité des architectures globales faisant fonctionner le robot, il est nécessaire de pouvoir répartir les calculs sur plusieurs PC fonctionnant en parallèle. Les réseaux de neurones faisant fonctionner le robot sont donc distribués et sont exécutés en temps réel. C'est pourquoi le robot embarque un routeur wifi permettant la communication entre le PC embarqué et les PC servant aux calculs délocalisés. Les expériences réalisées avec le robot se déroulent dans l'environnement "naturel" de la salle de travail du laboratoire. Le robot construit ses comportements sur la base de stimuli réels et donc bruités et changeants. La figure 1.4 présente cette environnement.



FIGURE 1.4 – Le robot dans son environnement. Les repères au sol ne sont pas visibles depuis la caméra qui ne perçoit que ce qui est plus haut que le niveau du sol.

5 Plan de la thèse

Le chapitre 2 décrit le point de vue psychologique de la compréhension des émotions. Les principales théories ainsi que leurs racines philosophiques y sont développées. Le chapitre 3 s'arrête sur les bases neurobiologiques des processus émotionnels. Le chapitre 4 décrit les outils qui seront utilisés tout au long de ce travail :

réseaux de neurones formels, algorithmes d'apprentissage, méthodes de navigation. Le chapitre 5 décrit le fonctionnement de la physiologie artificielle du robot sur laquelle sont basées les motivations et les sensations de douleur et de plaisir. Ce chapitre décrit le premier niveau de conditionnement permettant de construire les comportements d'attraction et de répulsion. Le chapitre 6 introduit le notion de frustration et décrit le système d'autorégulation du comportement basé sur le monitoring des différentes prédictions produites par le robot. Le chapitre 7 s'intéresse à la notion de surprise et décrit les mécanismes de traitement des informations reçues par le robot pour détecter la nouveauté. Le chapitre 8 est centré sur l'interaction expressive Homme-Robot et les différents moyens dont le robot peut en tirer parti. Enfin, le chapitre 9 conclut ce travail en faisant la synthèse des modèles et des expériences menées pour proposer un modèle biologiquement plausible de l'émergence des comportements émotionnels.

Emotions : le point de vue des psychologues

Les théories actuelles des émotions ayant été développées au cours du 20^{ème} siècle perpétuent et raffinent les mouvements de pensée initiés par les grands précurseurs du 19^{ème} siècle tels que James, Wundt, Darwin... La psychologie a pris son indépendance de la philosophie et est rentrée dans le domaine scientifique en proposant des théories et des modèles testables empiriquement, et de fait, réfutables, et en introduisant une méthodologie expérimentale rigoureuse. Cependant, si l'étude psychologique des émotions est née il y a 150 ans environ, comprendre le fonctionnement des émotions est une quête bien plus ancienne. Depuis Aristote qui a compilé des connaissances déjà bien connues à son époque, près de 2500 ans d'héritage philosophique s'est intensément intéressé à la question. De ce fait, la plupart des modèles précurseurs des émotions s'inscrivent dans la droite lignée de thèses soutenues par des philosophes plus anciens. Et même si le point de vue philosophique n'est qu'indirectement le sujet de ce travail, nous nous y attarderons brièvement tant certaines de ces thèses contiennent l'essence des modèles psychologiques des émotions.

1 Spinoza, Spencer, Wundt et les modèles dimensionnels

L'invention du concept d'espace des émotions trouve ses racines dans les théories que Spinoza (1632-1677) a publié en 1677 dans *Ethique, Partie III : Concernant la Nature et l'Origine des Emotions* [Spinoza 1677]. D'après Spinoza, trois éléments permettent de décrire toutes les expériences émotionnelles : la joie, le désir et la tristesse. Et même si ce modèle utilise pour dimensions ce que les conceptions modernes définissent comme étant des émotions et non des caractéristiques transversales de plus bas niveau, il présente déjà les attributs majeurs qui caractérisent les approches dimensionnelles modernes. Plusieurs dimensions définissent un espace émotionnel. Chaque émotion est définie par sa position le long de chaque dimension et donc par un point (ou une zone) dans cet espace émotionnel. Plus récemment, le philosophe et sociologue anglais Herbert Spencer (1820-1903) reprend cette conception et décrit les émotions comme étant les dimension de la conscience [Spencer 1890]. Mais c'est Wilhem Wundt (1832-1920), un physiologiste, philosophe et psychologue allemand, qui est considéré comme le père fondateur de l'approche dimensionnelle dans le domaine de la psychologie des émotions. Wundt a étudié les émotions en travaillant à partir de son introspection d'expériences simples. Dans son ouvrage *Grundriss der Psychologie* (les Contours de la Psychologie) publié en 1896 [Wundt 896], Wundt décrit son approche des émotions à partir d'un exemple délibérément simple. Wundt s'est intéressé au

ressenti subjectif que lui procurait l'écoute de séries de clicks sonores de différents rythmes. Il distingue que certaines séries sont plus agréables à entendre que d'autres. Il conclut qu'une partie de son expérience subjective est une sensation de plaisir ou de déplaisir qui peut être située le long d'une dimension bipolaire plaisir-désagrément. Mais en écoutant ses séries de clicks, il détecte un autre type de sensation. Alors qu'il s'attend au prochain click, il ressent une certaine tension tandis que lorsque le click attendu arrive, il ressent un soulagement. Cette succession de tensions et de relâchements est particulièrement claire quand le rythme est lent. Wundt en conclut donc que les émotions ont une composante dimensionnelle tension-relâchement. Enfin, quand la série de clicks est rapide, il se sent plus excité que quand la série est lente. Cette caractéristique définit une troisième dimension de l'expérience subjective des émotions, la dimension calme-excitation. Wundt décrit ainsi un espace caractérisé par trois dimensions. Il défend l'idée que les émotions sont définies par des combinaisons ou patterns d'activations de ressentis émotionnels primaires qui peuvent tous être caractérisés par une position dans cet espace tri-dimensionnel. Le figure 2.1 est une illustration de la théorie tri-dimensionnelle des émotions de Wundt.

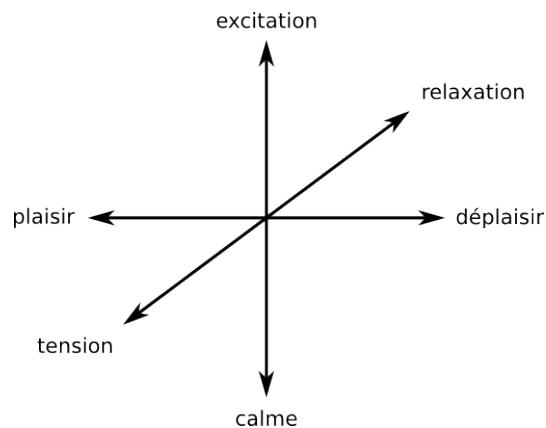


FIGURE 2.1 – La théorie tri-dimensionnelle de Wundt. Chaque émotion peut être comprise comme un pattern d'activation de ressentis émotionnels pouvant chacun être situé dans cet espace.

Cette conception dimensionnelle des émotions a inspiré un grand nombre de psychologues et, depuis le milieu des années 1930, un nombre croissant de variations et d'arguments sur ce thème ont été proposés. Elizabeth Duffy illustre cette tendance avec ces deux articles : *Emotion : an example of the need for reorientation in psychology* [Duffy 1934] et *An explanation of "emotional" phenomena without the use of the concept "emotion"* [Duffy 1941]. Elle y défend la nécessité d'identifier les aspects irréductibles communs à tous les comportements émotionnels : leurs dimensions fondamentales. L'idée centrale de cette approche est d'arriver à une vue fractionnée d'une réaction émotionnelle en composantes fondamentales définies *a priori* comme l'intensité de la réaction, la désorganisation du comportement, les mécanismes physiologiques impliqués, ... Les différences entre réactions émotionnelles peuvent être décrites par des variations d'intensité selon les dimensions définies. Comme il n'y a pas nécessité de recourir à un ensemble propre de dimensions pour chaque réaction émotionnelle, Duffy

soutient que l'étude des émotions en tant que telle devrait être abandonnée et remplacée par l'étude des variations des aspects simples et irréductibles du comportement en général. Cette théorie s'oppose donc naturellement à toute conception décrivant les émotions comme des mécanismes précisément catégorisables. Donald Lindsley, un neurophysiologiste ayant particulièrement étudié le système activateur du tronc cérébral, a proposé une "théorie de l'activation" [Lindsley 1951] plaçant les émotions comme le degré maximum d'une échelle d'activation du comportement (le sommeil étant le degré 0). Harold Schlosberg a développé une technique permettant d'observer et d'étudier cette activation : l'analyse de la conductance de la peau. Mais l'activation à elle seule ne permet pas de caractériser les différentes émotions de manière complète. C'est pourquoi il a aussi défini un modèle bi-dimensionnel de l'expressivité faciale [Schlosberg 1952]. Selon ce modèle, les expressions faciales peuvent être décrites (ou du moins classifiées) selon les dimensions agrément-désagrément et attention-rejet. Il propose donc un modèle à trois dimensions pour décrire les émotions [Schlosberg 1954]. S'il reprend le concept d'activation, il ajoute les dimensions qu'il utilise pour décrire les expressions faciales. Ainsi, l'étude de la conductance de la peau et l'analyse des expressions faciales sont les meilleurs moyens pour un psychologue d'étudier ces trois dimensions. James Russell a d'abord proposé un modèle tri-dimensionnel lui aussi mais pour lequel la dimension Attention-Rejet est remplacée par la dimension dominance-soumissions [Russell 1977]. Cela provient probablement du fait qu'au lieu d'étudier l'expressivité faciale, il a d'abord étudié les descriptions verbales et les jugements sur échelles différentielles (inspiré de Osgood [Osgood 1952, Osgood 1957]). Mais l'on retient plus de Russell son modèle "Circumplex" [Russell 1980]. Russell propose que les émotions sont organisées selon une structure circulaire dans un espace plan défini par les dimensions agrément-désagrément et degré d'activation (ou *arousal*). Russell a testé son modèle pour qualifier des termes émotionnels aussi bien que pour rapporter une expérience subjective émotionnelle propre. Les sujets doivent situer le mot ou l'état émotionnel comme un point dans cet espace bi-dimensionnel. La figure 2.2 illustre le modèle de Russell. L'uniformité interindividuelle des réponses des sujets sert d'argument pour défendre la validité de ces dimensions. Il est remarquable cependant que si les réponses des sujets sont placées sur un cercle dans cet espace bidimensionnel, alors elles ne sont pas réellement définies par deux coordonnées dans le plan mais par une seule dimension : l'angle dans le cercle. Russell attaque aussi l'idée que les expressions faciales sont universellement reconnues [Russell 1994]. D'après lui, les biais expérimentaux sont trop nombreux pour pouvoir accepter les résultats des études qui cherchent à le prouver. Les réponses à choix forcé, l'utilisation de protocoles *within subject*¹ ou la présélection des photographies d'expressions faciales utilisées posent des problèmes de biais méthodologiques limitant l'acceptation des conclusions de ces études.

Plus récemment, dans les années 90, P.J. Lang utilise la méthodologie *startle probe*

1. Les protocoles *within subject* consistent à utiliser la totalité des sujets d'une expérience dans chaque condition expérimentale à tester. Ils diffèrent des protocoles *between subject* qui séparent les sujets en autant de groupes qu'ils testent de conditions expérimentales. Si les protocoles *within subject* permettent d'assurer un plus grand nombre de sujets et donc une meilleure significativité statistique, ils induisent la possibilité de biais expérimentaux liés aux effets de répétition (apprentissage, fatigue, ...).

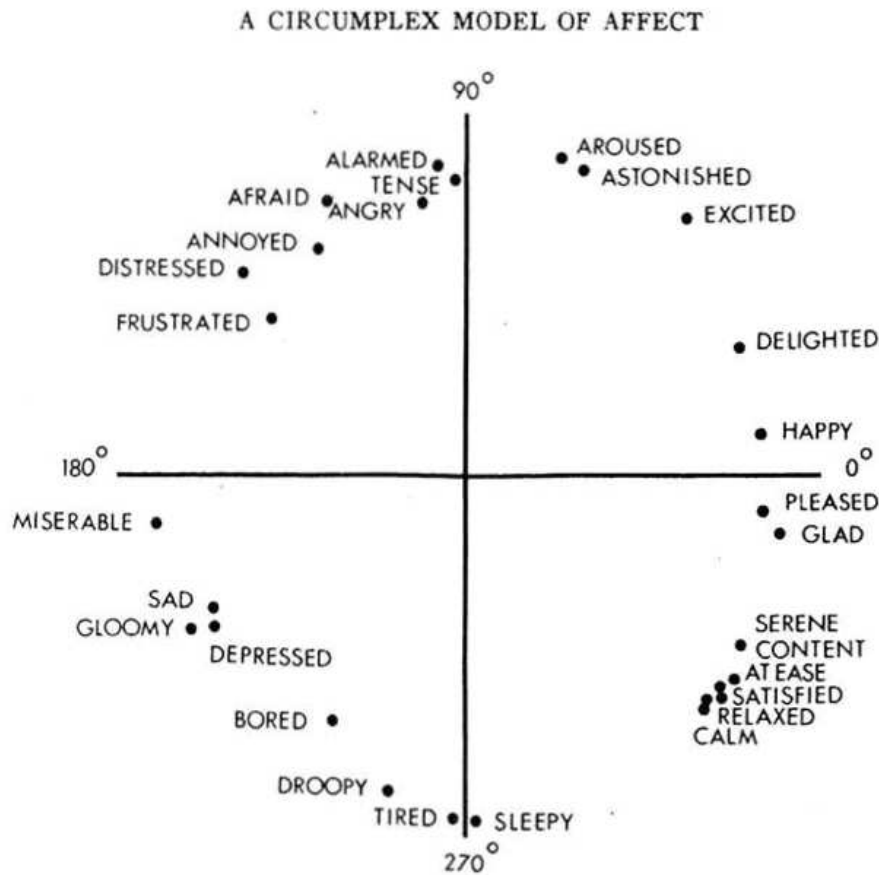


FIGURE 2.2 – Modèle *Circumplex* de James Russell. 28 termes émotionnels sont placés le long des deux dimensions agrément-désagrément (axe horizontal) et degré d'activation (axe vertical). Dans l'espace qu'elles définissent, un cercle permet de caractériser de manière relativement uniforme des termes émotionnels.

(sonde par la surprise) pour différencier les émotions qui sur-activent la réaction de surprise des émotions qui l'inhibent. Il distingue ainsi les émotions appétitives, comme la joie, qui diminuent les effets de la sonde, des émotions aversives, comme la peur, qui au contraire augmentent ses effets [Lang 1990]. Lang montre aussi qu'il existe une covariance importante entre, d'une part, l'expressivité faciale et la dimension de valence émotionnelle, et d'autre part, entre la conductance dermique et la dimension d'activation [Lang 1993, Lang 1994].

S'il n'existe pas de consensus exact sur les dimensions qui caractérisent les expériences émotionnelles, les dimensions d'activation et de valence sont les plus fréquemment utilisées. Le point de vue dimensionnel s'oppose directement au point de vue basique et discret. Le point de vue dimensionnel affirme que ce qui distingue les émotions les unes des autres repose sur des différences quantitatives continues des éléments composants les réactions engendrées lors d'un épisode émotionnel. La variabilité inter et intra-individuelle des expériences émotionnelles est essentielle aux modèles dimensionnels. Le point de vue discret au contraire, défend l'idée que les émotions se distinguent par des différences qualitatives discrètes impliquant des réactions différentes par nature.

2 Descartes, Darwin et les émotions basiques discrètes

Si toute une tradition de philosophe s'est intensément penchée sur la question des émotions et de leur nature, c'est généralement à Descartes (1596-1650) que l'on se réfère pour situer les racines des théories modernes des émotions basiques. Dans son ouvrage, *Les Passions de l'Ame* [Descartes 1649], il décrit un modèle du fonctionnement des émotions humaines. Ce modèle repose sur 6 primitives émotionnelles : l'émerveillement, l'amour, le désir, la joie, la haine et la tristesse. Les humains ressentent le plus souvent ces différentes primitives non pas de manière pure et isolée mais simultanément, sous forme d'états émotionnels complexes. C'est la variété des configurations possibles de la forme d'activation des primitives émotionnelles qui produit la diversité du répertoire émotionnel humain. Un grand nombre de raffinements ont été apportés à cette vision des émotions. Peu de personnes sont d'accord quant à une liste exhaustive des émotions basiques, et certainement pas avec exactement les 6 primitives de Descartes. Cependant, sa proposition est à la base des théories modernes des émotions basiques.

Charles Darwin (1809-1882), le co-fondateur avec Alfred Russel Wallace (1823-1913) du mouvement évolutionniste, a eu un impact retentissant sur le monde intellectuel de son époque et depuis lors, tant dans les milieux scientifiques, que philosophiques ou spirituels. Il s'est aussi beaucoup intéressé aux émotions dont traite son livre *L'expression des émotions chez l'homme et l'animal* [Darwin 1872]. Durant l'écriture de ce livre, Darwin correspond avec Guillaume-Benjamin-Amand Duchenne de Boulogne (1806-1875), un neurologue français s'intéressant lui aussi de près aux émotions et à leurs expressions faciales. Duchenne pense qu'il existe au moins 60 émotions discrètes. Il se base pour cela sur ses études du fonctionnement des muscles faciaux et du potentiel de leurs capacités expressives [Duchenne 1876]. En stimulant électriquement le visage de sujets afin de produire des contractions musculaires toniques et contrôlées, Duchenne crée artificiellement des expressions faciales émotionnelles (voir la figure 2.3). Sa capacité à en produire plusieurs dizaines correspond selon lui au grand nombre des émotions distinctes que peut ressentir un être humain. Allant plus loin même, Duchenne suggère que l'expressivité faciale peut induire l'état émotionnel. Cette idée peut être résumée par l'aphorisme : "Souris et tu seras heureux". Au contraire, Darwin pense qu'il n'existe qu'un nombre restreint d'émotions réellement universelles. En testant la manière dont les gens reconnaissent les expressions de Duchenne, Darwin fait la distinction entre celles dont la reconnaissance est consensuelle (les émotions universelles) de celles qui sont ambiguës (les simulations ratées). Dans *L'expression des émotions chez l'homme et l'animal*, Darwin décrit avec détails l'expressivité émotionnelle des humains mais aussi de chats, de chiens, de chevaux et de singes. Il réfute dans ce domaine aussi une frontière imaginaire de plus entre les hommes et les autres animaux. Etant les fruits d'une sélection adaptative et dynamique, les émotions sont décrites comme un ensemble restreint de mécanismes innés et universels permettant une adaptation rapide aux changements des conditions environnementales afin d'optimiser les chances de survie, de reproduction et donc de perpétuation. Les émotions sont vues comme des mécanismes discrets remplissant



FIGURE 2.3 – Duchenne a produit artificiellement des expressions faciales émotionnelles contrôlées par stimulations électriques des muscles du visage.

chacun une fonction adaptative propre et déclenchant une expressivité spécifique. L'étude de l'expressivité plutôt que du ressenti émotionnel permet, selon Darwin, une approche plus scientifique de la compréhension de ces phénomènes. Il défend par exemple l'hypothèse du caractère inné des émotions en faisant remarquer que les aveugles de naissance ont des expressions émotionnelles similaires aux expressions des personnes voyant normalement. Darwin insiste aussi sur le lien entre les émotions et le système nerveux. Il décrit le principe de l'action directe du système nerveux sur le corps, indépendante de la volonté et en partie des habitudes (rougissement, tremblements, transpiration, changement de couleur des cheveux, ...).

La synthèse des points de vue de Descartes et Darwin forme le socle du paradigme des émotions basiques. Les émotions basiques sont des processus distincts, résultant d'une adaptation sélective et qui définissent une sorte de vocabulaire émotionnel dont la combinatoire caractérise les états émotionnels complexes. Les théories modernes

diffèrent quant au nombre ou à la liste exacte des émotions basiques. De plus, elles apportent chacune un éclairage particulier sur un aspect de ce type de modèle. Cependant, elles ont toutes en commun ce paradigme de pensée.

Les psychologues américains Sylvan Tomkins et Robert Plutchik sont à l'origine des premières théories modernes des émotions basiques. Ils ont tous les deux publié en 1962 les premiers articles décrivant leur théories : l'*Affect Theory* de Tomkins [Tomkins 1962, Tomkins 1963, Tomkins 1980, Tomkins 1991] et la *General Psychoevolutionary Theory of Emotion* de Plutchik [Plutchik 1962, Plutchik 1980, Plutchik 2002]. Selon Tomkins, les émotions sont composées d'affects basiques au nombre de 9. Les affects sont la part biologique de l'émotion, c'est à dire «*les mécanismes pré-câblés, pré-programmés et transmis génétiquement qui existent en chacun de nous*» [Nathanson 1992]. Ces affects s'organisent en trois catégories, positif, neutre ou négatif et sont représentés par des paires qui décrivent leurs manifestations normales et extrêmes :

- Affects positifs : intérêt-excitation, satisfaction-joie
- Affect neutre : surprise-effarouchement
- Affects négatifs : détresse-angoisse, colère-rage, peur-terreur, honte-humiliation, dégoût gustatif et dégoût olfactif.

Tomkins défend l'idée que ces 9 affects sont distincts, qu'ils partagent un héritage biologique commun avec ce que l'on qualifie d'émotion chez les animaux et que leurs patterns d'activation définit la complexité des émotions. Il propose donc une théorie résolument orientée vers l'approche évolutionniste des émotions basiques. De plus, cette théorie affine la description des émotions en tenant compte de leur intensité.

Partant du même paradigme scientifique, Plutchik a produit une théorie assez proche de celle de Tomkins. Cependant, il ajoute un certain nombre de concepts qu'il illustre élégamment avec le concept de roue des émotions. La théorie de Plutchik repose sur 10 postulats. 7 d'entre eux ne sont que la réaffirmation des principes de l'approche basique évolutionniste. Nous ne présenterons donc que les 3 postulats qui font la spécificité de la théorie de Plutchik :

- Les émotions primaires peuvent être conceptualisées en termes de paires d'opposés.
- Toutes les émotions varient dans leur degré de similitude réciproque.
- Chaque émotion peut exister à différents degrés d'intensité.

Plutchik décrit 8 émotions primaires : la joie, la tristesse, la confiance, le dégoût, la peur, la colère, la surprise et l'anticipation. Ces émotions primaires recourent en partie les affects basiques de Tomkins. Les deux psychologues considéraient comme essentielle la notion d'intensité de l'émotion, cependant Plutchik va plus loin en postulant pour l'existence de relations spécifiques entre émotions discrètes. Il illustre cette idée avec la représentation de la roue des émotions, une modélisation conceptuelle des émotions suivant la métaphore des couleurs (voir figure 2.4). Comme les couleurs pouvant être plus ou moins pales, les émotions primaires peuvent s'exprimer à différentes intensités. Par exemple, l'appréhension, la peur et la terreur correspondraient à la même émotion primaire, mais exprimée à des degrés d'intensité croissants. Deux couleurs opposées

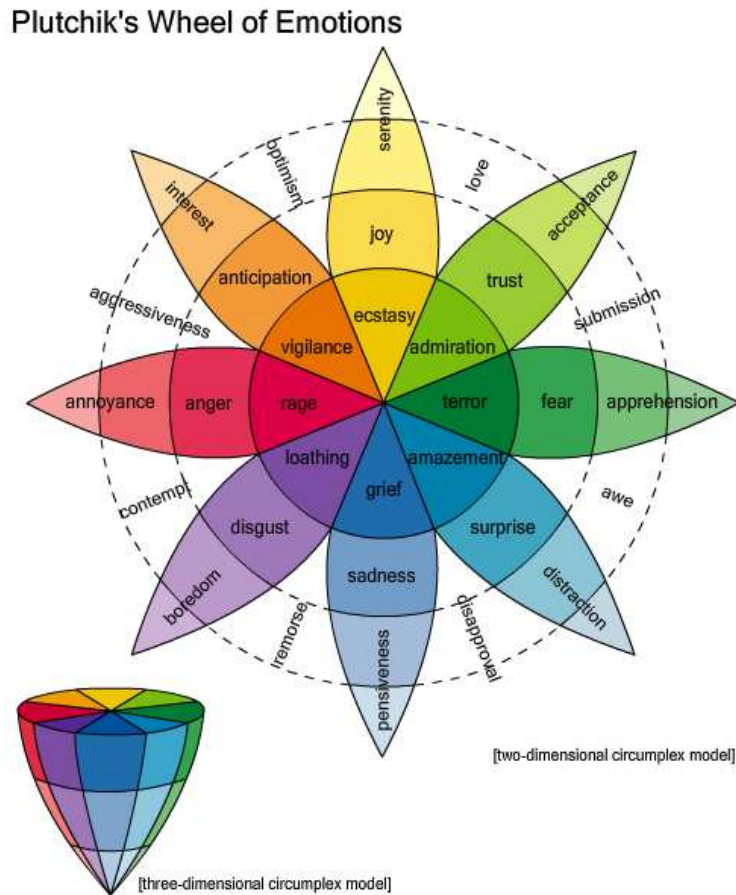


FIGURE 2.4 – Roue des émotions de Plutchik. Les émotions primaires, par paires d'opposés : joie-tristesse, confiance-dégoût, peur-colère et surprise-anticipation. Chaque émotion primaire peut varier en intensité (e.g. appréhension < peur < terreur). Les mélanges d'émotions primaires donnent des émotions secondaires plus complexes (e.g. agressivité = anticipation + colère).

sur la roue correspondraient à une paire d'émotions opposées. Ainsi, la peur serait l'opposée de la colère, la tristesse l'opposée de la joie, ... Et comme deux couleurs primaires créent une nouvelle couleur en se mélangeant, la simultanéité de deux émotions primaires produirait une nouvelle émotion, secondaire celle-ci. Par exemple, le mépris est compris comme étant un état simultané de colère et de dégoût, l'amour comme un état simultané de joie et de confiance.

Les deux premiers tiers du XXème siècle ont été dominés par les modèles dimensionnels. Cependant, les théories de Tomkins et Plutchik ont permis de lancer le mouvement moderne des émotions basiques et ainsi de remettre à égalité l'influence de ces deux types d'approches. Et si Tomkins lui-même n'est pas très connu aujourd'hui, ses deux étudiants Paul Ekman et Carroll Izard sont devenus des noms incontournables de ce courant de pensée.

D'après Ekman, les émotions basiques possèdent des patterns d'activation spéci-

riques permettant de les différencier sur la base de l'activation du système nerveux autonome [Ekman 1983]. Mais si l'induction d'états émotionnels forts chez des sujets expérimentaux afin de mesurer cette réaction du système nerveux autonome pose des problèmes éthiques évidents et limite l'expérimentation humaine, Ekman rejoint Darwin sur l'importance donnée à l'étude de l'expressivité faciale émotionnelle. Il a donc conduit un grand nombre d'études interculturelles sur la reconnaissance des expressions faciales émotionnelles. Dans une première série d'études, Ekman a montré des photos d'expressions faciales à des membres de différentes cultures et tous ont réussi à les reconnaître. Mais afin de garantir que les sujets testés n'avaient pas précédemment été en contact avec d'autres cultures (notamment via le cinéma), il a mené ses expériences sur des tribus totalement isolées du reste du monde. Il a étudié en particulier les tribus Fore de Papouasie-Nouvelle-Guinée. Ekman a observé que les membres de cultures isolées n'avaient aucun mal à reconnaître les expressions faciales de membres de cultures différentes qu'ils n'avaient jamais rencontré [Ekman 1969, Ekman 1971, Ekman 1972]. Ces études lui ont permis de solidifier sa théorie de l'universalité et du caractère inné des expressions faciales liées aux émotions basiques. Cela lui a valu de nombreuses réactions de l'anthropologue Margareth Mead, grande défenseur de la thèse du déterminisme culturelle de l'expressivité émotionnelle. Ekman a centré son étude des émotions basiques sur leur composantes expressives non verbales. C'est dans ce sens qu'il a conçu le système FACS (*Facial Action Coding System*) [Ekman 1977, Ekman 1978] afin de pouvoir décrire et catégoriser n'importe quelle expression faciale. Selon ces critères, Ekman a tout d'abord défini une liste de 6 émotions basiques [Ekman 1972] : la colère, la peur, la tristesse, le dégoût, la joie et la surprise. La figure 2.5 représente les 6 émotions basiques et leurs expressions faciales associées.

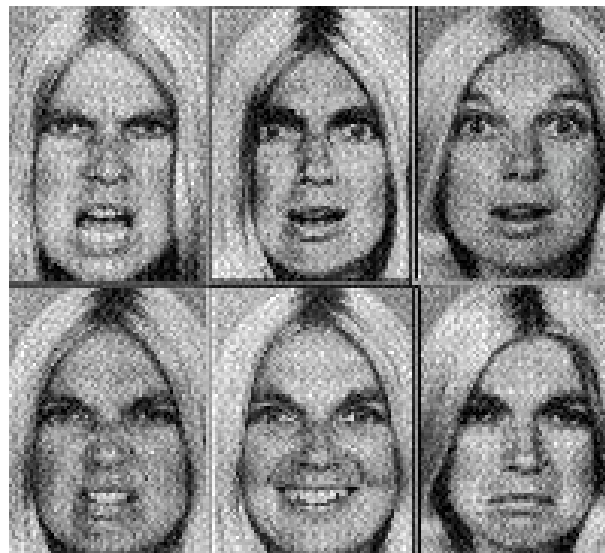


FIGURE 2.5 – Les expressions faciales associées aux émotions basiques définies par Ekman [Ekman 1972]. De gauche à droite et de haut en bas : colère, peur, surprise, dégoût, joie, tristesse (si cette légende est réellement nécessaire, alors ces expressions ne sont pas si universelles...)

Cependant, dans les années 1990, Ekman a augmenté la taille de sa liste initiale en ajoutant des émotions positives et négatives dont certaines ne sont même pas caractérisées par une expression faciale [Ekman 1999]. Il inclut ainsi l'amusement, le mépris, le contentement, l'embarras, l'excitation, la culpabilité, la fierté, le soulagement, la satisfaction, le plaisir sensoriel et la honte.

Izard a lui aussi beaucoup contribué à l'approche basique des émotions. S'il a lui aussi conçu un système de codage des expressions faciales, le *Maximally Discriminative Affect Coding System* (MAX) [Izard 1979], il est plus reconnu pour son travail sur le développement affectif et l'expressivité émotionnelle. Izard défend l'approche basique des émotions [Izard 1971, Izard 1972, Izard 1977, Izard 1993], en particulier à l'aide d'arguments développementaux [Izard 1980, Izard 1994, Cicchetti 1995] notamment concernant l'apparition durant les premiers mois de la vie des expressions faciales des émotions basiques sans précurseurs de mouvements faciaux (*facial movement precursors*) ou "proto expressions faciales". Izard défend aussi l'idée de la congruence de l'expressivité et de l'expérience subjective émotionnelle. Il propose l'hypothèse du *facial feedback* selon laquelle les émotions ayant différentes fonctions produisent des expressions faciales distinctes et fournissent ainsi des indices pour déterminer ce que quelqu'un ressent [Izard 2004].

3 De la question de la séquence à la question de l'évaluation cognitive

La naissance de l'étude scientifique moderne des émotions est souvent associée à la théorie James-Lange proposée durant les années 1880 de manière indépendante par le psychologue américain William James (1842-1910) et le psychologue danois Carl Lange (1834-1900) [James 1884, James 1890, Lange 1885]. Cette théorie qui tranche radicalement avec la conception commune de l'époque propose que les émotions sont des conséquences directes des modifications physiologiques (hausse du rythme cardiaque, transpiration, tension musculaire, ...) dues au système nerveux autonome en réaction aux interactions avec l'environnement. James et Lange définissent ainsi une vision «périphéraliste» des émotions. Les émotions ne seraient pas des états internes fondamentaux provoquant des réactions périphériques, mais au contraire seraient le résultat de ces réactions périphériques. Lange illustre ce point de vue par la question suivante : «*Si je commence à trembler parce que je suis menacé par un pistolet chargé, est-ce que tout d'abord un processus psychique se produit, la terreur apparaît, et c'est cela qui cause mes tremblements, mes palpitations du coeur, et la confusion de la pensée ; or alors, est-ce que ces phénomènes corporels sont produits directement par la cause terrifiante de telle sorte que l'émotion consiste exclusivement en une modification fonctionnelle dans mon corps ?*». James, lui, explique sa théorie comme suit : «*Le sens commun nous dit que si nous perdons notre fortune, nous sommes désolés et nous pleurons ; si nous rencontrons un ours, nous sommes effrayés et nous courrons ; si nous sommes insultés par un rival, nous sommes en colère et nous frappons. L'hypothèse défendu ici dit que l'ordre de cette séquence est incorrect ... et*

3. De la question de la séquence à la question de l'évaluation cognitive

que la déclaration la plus rationnelle est que nous nous sentons désolé parce que nous pleurons, en colère parce que nous frappons, effrayés parce que nous tremblons ...». La position de James et Lange est donc que les émotions sont des conséquences de nos changements corporels et non leurs causes. Chaque émotion serait ainsi caractérisée par son propre pattern de changements physiologiques (idée déjà très ancienne). Lange, qui prônait une position sensiblement plus extrême que James, a été jusqu'à définir les émotions comme étant des changements vasomoteurs (accroissement ou diminution du diamètre des vaisseaux sanguins).

Vers la fin des années 1920, un mouvement de pensée concurrent apporte le débat. La théorie Cannon-Bard, proposée par les américains Walter Cannon (1871-1945) [Cannon 1920, Cannon 1927] et Philip Bard (1898-1977) [Bard 1928], prend le contre-pied de la théorie James-Lange. Cannon et Bard défendent une conception centraliste des émotions pour laquelle le déclenchement d'une émotion spécifique dépend du traitement d'un stimulus par le système nerveux central, les changements physiologiques ou pattern d'activation périphérique étant non spécifiques. Cette théorie met en avant l'importance du système nerveux central, et en particulier du thalamus, dans le déclenchement d'une émotion donnée. Les changements physiologiques sont donc vus comme des conséquences des émotions et non comme leurs causes. D'après la théorie Cannon-Bard, c'est bien la peur qui fait courir, la colère qui fait frapper et la tristesse qui fait pleurer. Cannon a d'ailleurs proposé un certain nombre d'arguments pour réfuter la position de la théorie James-Lange : des réactions viscérales diffuses semblables se produisent pour toutes les émotions, mais aussi pour des états non émotionnels comme la digestion ou la fièvre, la suppression des afférences viscérales ne supprime pas les ressentis émotionnels, les viscères réagissent relativement lentement et le ressenti émotionnel peut se produire avant un quelconque changement physiologique, ... Si Cannon propose un bon nombre d'arguments pour s'opposer à la théorie James-Lange, la théorie Cannon-Bard n'est pas non plus dénuée de défauts. Le principal étant l'absence d'explication du mécanisme sous-tendant les processus émotionnels.

Ces théories sont toutes deux fondées sur une approche physiologique des émotions mais diffèrent radicalement quant à leur conception de la séquence temporelle des émotions. Ce débat scientifique a eu un impact considérable sur les travaux ultérieurs liés aux émotions, tant s'agissant de la relation causale entre changements physiologiques et émotions que concernant le rôle de la cognition dans les mécanismes émotionnels. Le besoin de clarification nécessaire pour comprendre le mécanisme de traitement (d'évaluation) que la théorie Cannon-Bard attribue au système nerveux central peut être considéré comme l'un des points de départ des théories de l'évaluation cognitive ou théorie de l'*appraisal*.

Les théories de l'*appraisal* mettent l'accent sur le processus d'évaluation d'un événement, qu'il soit externe ou interne (corporel), ou d'une situation, qui déclenche un épisode émotionnel. Depuis le début des années 50, ces théories ont connu un succès grandissant. En 1950, Magda Arnold propose l'*excitatory theory of emotion* et introduit le terme *appraisal* pour expliquer la production d'émotions différenciées

[Arnold 1950]. Une évaluation cognitive y est définie comme étant un processus cognitif, rapide, automatique, inconscient, dont la fonction est d'évaluer les stimuli perçus sur la base de critères particuliers [Arnold 1960]. Les émotions y sont vues comme une tendance d'attraction ressentie face à tout ce qui est intuitivement jugé comme bénéfique ou une tendance à l'évitement face à ce qui est intuitivement jugé nocif. Cette attirance ou cette aversion est accompagnée d'un *pattern* de changements physiologiques orientés respectivement vers l'approche ou l'évitement. Ces patrons de changements physiologiques varient en fonction des différentes émotions.

Schachter et Singer proposent une approche tenant compte à la fois des déterminants cognitifs, sociaux et physiologiques des émotions. Les états émotionnels peuvent être considérés comme une fonction d'un état physiologique d'activation (*arousal*) et d'une cognition qui y est approprié [Schachter 1962]. Ce modèle prédit qu'un individu dans un état physiologique d'*arousal* pour lequel il n'a pas d'explication immédiate décrit son état et son ressenti à l'aide de la cognition à laquelle il a accès. Si au contraire, il est dans un état d'*arousal* pour lequel il a une explication parfaitement appropriée, aucune évaluation ne sera nécessaire et il est très improbable qu'il fasse une évaluation cognitive alternative mais plausible de la situation. Enfin, il ne réagit émotionnellement ou ne décrit ce qu'il ressent comme étant des émotions que dans la mesure où il est dans un état d'*arousal*. Schachter ranime ainsi l'ancien débat des périphéralistes et des centralistes [Schachter 1975].

Depuis le milieu des années 60, Richard Lazarus a aussi profondément influencé le champ des théories de l'*appraisal*. Il suggère que le stress et les émotions sont le fruit d'un processus d'*appraisal* en deux temps [Lazarus 1966]. L'*appraisal* primaire consiste à évaluer la signification positive ou négative d'un événement ou d'une situation pour le bien être subjectif. L'*appraisal* secondaire est la capacité à faire face aux conséquences de cet événement ou de cette situation. De plus, l'*appraisal* n'est pas un processus rigide et sa nature dynamique lui permet d'être réactivé sur la base de nouvelles informations [Lazarus 1968]. Lazarus défend l'idée que la pensée, ou la cognition, est une condition nécessaire des émotions [Lazarus 1982, Lazarus 1984, Lazarus 1991] s'opposant ainsi sur cette question à Zajonc qui défend la primauté de l'affect sur la cognition [Zajonc 1984a].

Depuis le début des années 80, de nombreux psychologues ont développé et raffiné ces modèles. Parmi les incontournables de ce mouvement, citons notamment Ira Roseman [Roseman 1984, Roseman 1990, Roseman 1991, Roseman 1996], Craig Smith [Smith 1985, Smith 1987, Smith 1989, Smith 1993, Smith 1996], Nico Frijda [Frijda 1986, Frijda 1987, Frijda 1993], Andrew Ortony [Ortony 1988, Ortony 1990], Phoebe Ellsworth [Ellsworth 1988, Ellsworth 1991, Ellsworth 1994], Bernard Weiner [Weiner 1979, Weiner 1982, Weiner 1986, Weiner 1987] ou Robert Zajonc [Zajonc 1980, Zajonc 1984a, Zajonc 1984b]. Les nuances qui distinguent les modèles proposés reflètent les divergences de postulats ou de croyances concernant la nature des émotions. Qu'il s'agisse des questions de nature discrète ou dimensionnelle des émotions ou des questions de séquences temporelles des mécanismes impliqués.

Afin d'illustrer plus en détail cette approche, voici une description du modèle des

3. De la question de la séquence à la question de l'évaluation cognitive

composantes proposé par Klaus Scherer [Scherer 1982, Scherer 1984b, Scherer 1984a, Scherer 1986, Scherer 1988, Scherer 1993]. Scherer définit une émotion comme étant une séquence de changements d'état ayant lieu dans cinq composantes :

- La composante cognitive réalise le traitement de l'information et donc la fonction d'*appraisal* (ou d'évaluation de la stimulation).
- La composante neurophysiologique sert de support et permet la régulation des différents sous-systèmes.
- La composante motivationnelle assure la fonction exécutive. C'est elle qui prépare et dirige l'action.
- La composante expressive correspond à la communication de la réaction et des intentions face à l'événement déclencheur (mouvements, expressions faciales, vocalisations).
- Le ressenti subjectif a le rôle de monitorer ce processus. Il dirige le focus attentionnel, permet la réflexion et participe au sentiment subjectif de soi.

Ces changements d'états des 5 composantes ont lieu de manière interdépendante et synchronisée en réponse à l'évaluation d'un stimulus environnemental ou corporel. La figure 2.6 illustre le fonctionnement de ces composantes. L'*appraisal* consiste en véri-

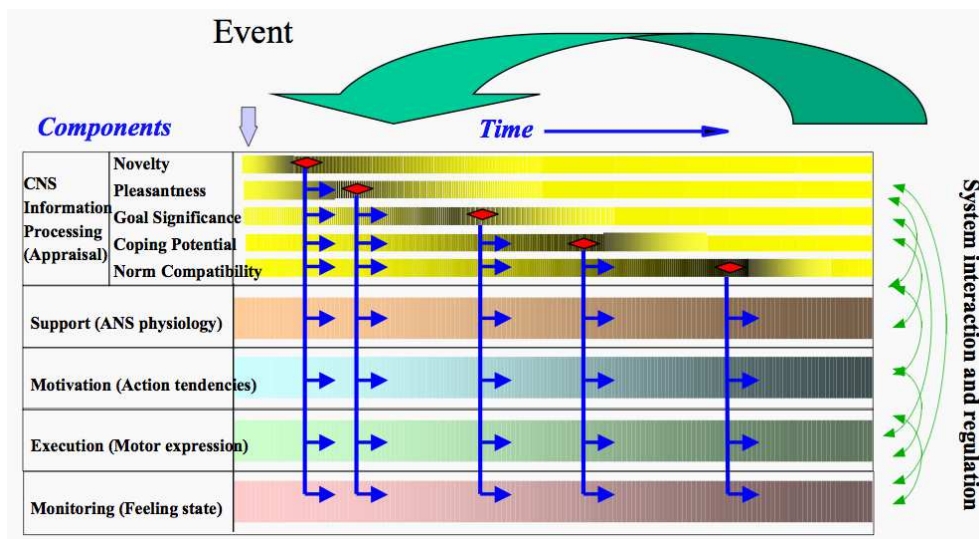


FIGURE 2.6 – Modèle componentiel de Scherer [Scherer 2001] : les différentes composantes (cognitive, physiologique, motivationnelle, expressive et subjective) fonctionnent en interaction pour permettre de s'adapter à un événement de la manière la plus adéquate (en fonction des buts et des besoins). Depuis l'évaluation de la situation à l'exécution de la réponse appropriée.

fication d'un ensemble de critères permettant d'évaluer un événement, ce que Scherer appelle les *Stimulus Evaluation Checks* (SEC). Le résultat de cette évaluation permet de prédire le type et l'intensité de l'émotion élicite par l'événement. Les SEC sont organisés autour de quatre intérêts principaux :

- Pertinence : est-ce que cet événement est pertinent pour moi ? Est-ce qu'il affecte directement ma personne ou mon groupe social ?

- Implications : Quelles sont les implications ou les conséquences de cet évènement et à quel point vont-elles affecter mon bien-être ou mes buts à court et long terme ?
- Potentiel de coping : à quel point suis-je capable de faire face à ces conséquences ?
- Significativité normative : Quelle significativité a cet évènement par rapport à mes convictions personnelles ainsi que face aux normes et valeurs sociales ?

La figure 2.7 illustre le processus d'*appraisal* et ses liens avec les fonctions de haut niveau d'une part (attention, mémoire, motivation, raisonnement et conscience de soi) et les fonctions physiologiques (système endocrinien, physiologie autonome, réponse somatique, système nerveux). Les SEC correspondent aux informations

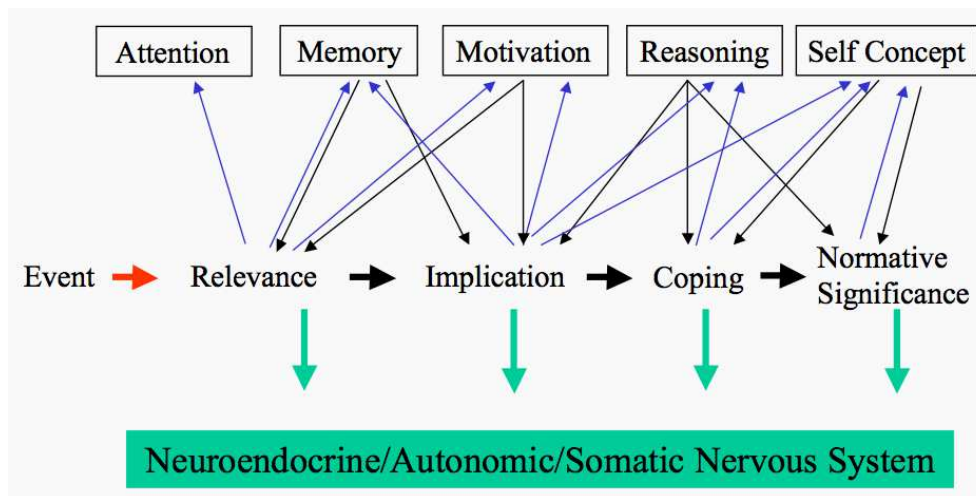


FIGURE 2.7 – Processus de l'*appraisal* selon Scherer consiste en l'évaluation d'un évènement du point de vue de sa pertinence, des ses implications, de son potentiel de coping et de sa significativité normative. Ces évaluations reposent sur des fonctions de hauts niveaux (attention, mémoire, motivation, raisonnement et concept de soi) et impliquent des réactions physiologiques de bas niveau.

les plus importantes dont a besoin l'organisme pour avoir une réaction appropriée à la situation. L'évaluation des SEC se fait toujours de manière subjective. Elle dépend donc de la perception ainsi que des inférences qu'un individu peut faire d'une situation. De plus, comme déjà suggéré par Lazarus, l'évaluation n'a pas lieu qu'une seule fois, mais elle se répète dans un processus de *reappraisal* permanent qui permet de rester adapté à l'évènement alors qu'il évolue. Contrairement aux théories des émotions discrètes, le modèle des composants de Scherer ne requiert pas un nombre restreint d'émotions basiques. Au contraire, la fluctuation constante du fonctionnement des sous-systèmes présenté plus haut permet de faire ressortir un très large spectre d'états émotionnels. Ceci dit, Scherer ne rejette pas l'idée de l'universalité de certaines réponses émotionnelles comme la peur (associée à la fuite) et la colère (associée à l'affrontement). D'après son modèle, il paraît très vraisemblable que des combinaisons différentes de résultats aux SEC puisse aboutir à des patterns réguliers de changements d'états spécifiques. C'est pourquoi Scherer parle d'émotions modales pour décrire ces réponses prédominantes qui sont dues

selon lui aux conditions générales de vie, aux contraintes de l'environnement social ou encore aux similarités génétiques.

4 Discussion

Le débat fondamental sur la nature des émotions semble se jouer entre les modèles dimensionnels et les modèles discrets. Les modèles de l'évaluation cognitive portent leur attention sur la précision de la description de la mécanique des processus émotionnels. Ils se répartissent d'ailleurs dans ces deux grandes familles de modèles émotionnels. Mais cette opposition apparemment diamétrale n'est peut être pas si nette que ce que suggèrent les partisans du débat. La différence majeure entre modèles dimensionnels et discrets semble être le niveau de description. Tandis que les modèles dimensionnels comparent les émotions en fonction de ce qui les rapproche, les modèles discrets les comparent en fonction de ce qui les distingue. A ce titre, un point de vue purement dimensionnel semble s'apparenter à une grille de lecture, un modèle descriptif, plus qu'à une explication du fonctionnement des émotions. Scherer a proposé par exemple que les modèles dimensionnels décrivent la variabilité (notamment d'intensité mais aussi de style) d'émotions distinctes et clairement définies. S'il existe une faible variabilité interindividuelle dans les tâches consistant à situer des termes émotionnels dans un espace défini par des dimensions fondamentales, on peut se poser la question de la convergence des résultats pour la tâche opposée. Imaginons une expérience dans laquelle les sujets doivent dénommer les émotions correspondant à des points situés aléatoirement dans ces espaces émotionnels ? La tâche semble beaucoup moins évidente. D'un autre côté, il est remarquable de noter que la plupart des modèles discrets admettent une dimension d'intensité. Difficile alors de pleinement rejeter le point de vue dimensionnel. Les modèles discrets se revendiquent d'une tradition darwinienne et à ce titre semblent s'approprier la primauté de l'argument évolutionniste. Darwin a effectivement participé à la genèse de l'approche discrètes des émotions mais pour autant, les modèles dimensionnels ne sont pas en contradiction avec le point de vue évolutionniste. Ce n'est pas parce que l'on décrit les émotions comme un ensemble d'états ne variant qu'en degré et pas en nature que l'on postule pour la nécessité d'un système complet afin d'avoir le moindre avantage sélectif. En d'autres termes, l'approche dimensionnelle ne constitue pas plus un argument contre l'évolution que pour le créationnisme.

Un modèle comme le *Component Process Model* de Scherer décrit particulièrement en détails les processus émotionnels. Cependant, cette description est axée sur l'aspect fonctionnel. C'est pour cela qu'un même modèle permet de représenter l'ancrage corporel des émotions et le jugement normatif qu'elles impliquent. Scherer ne donne pas directement de réponse précise quant à la manière de concevoir le plus bas niveau d'un modèle émotionnel. Cette barrière peut être franchie si l'on s'intéresse aux émotions du point de vue de leurs bases neurobiologiques. En comprenant mieux les mécanismes cérébraux sous-tendant les émotions, certaines des questions posées par les psychologues peuvent trouver une réponse. De plus, les notions de plaisirs et de

douleurs et de construction des signaux émotionnels sont très peu discutées dans les modèles psychologiques des émotions. C'est pourquoi le prochain chapitre a pour propos de faire un état de l'art des connaissances neurobiologiques sur le fonctionnement des émotions. C'est ce regard biologique qui nous servira principalement d'inspiration pour concevoir un modèle émotionnel plausible. En retour, l'expérimentation robotique viendra confirmer ou sanctionner la validité des modèles proposés.

Neurobiologie des émotions

Dès 1929, le physiologiste allemand Hans Berger a utilisé l'électroencéphalographie (EEG) pour observer l'activité d'un cerveau humain. Des électrodes placées à la surface du cuir chevelu perçoivent l'activité électrique des neurones. Si cette méthode a le mérite de mesurer l'activité d'un cerveau vivant, elle ne permet pas comme l'autopsie d'étudier les différentes structures du cerveau. Mais depuis les années 80, les progrès impressionnants en matière d'imagerie cérébrale ont permis deux avancées majeures. De nouvelles techniques ont permis d'affiner de manière significative la résolution spatiale et temporelle de l'enregistrement des signaux cérébraux (imagerie fonctionnelle : imagerie par résonance magnétique fonctionnelle (IRMf), tomographie par émission de positrons (TEP), magnétoencéphalographie (MEG)). Mais de nouvelles méthodes ont aussi permis de visualiser en 3 dimensions la structure du cerveau (imagerie anatomique : tomodensitométrie (CT-scan), imagerie par résonance magnétique anatomique (IRM), tomographie à émission mono-photonique (TEMP)). Ces deux familles de techniques ont permis un nouvel essor des neurosciences. Un très grand nombre d'études ont permis d'améliorer la compréhension du rôle des différentes structures du cerveaux impliquées dans les réponses émotionnelles. Les différentes méthodes d'imageries cérébrales se sont progressivement imposées comme des outils indispensables de la neuropsychologie et dans l'étude des émotions en particulier. Elles ne remplacent pas mais complètent au contraire les méthodes traditionnelles ayant permis depuis le début du 20^{ème} siècle d'étudier le fonctionnement du cerveau : chirurgie, analyse histologique, stimulation intracrânienne électrique ou chimique, autopsie. Dans premier temps, nous ferons un bref historique de la neurobiologie des émotions avant de présenter les principales structures cérébrales impliquées ainsi que les circuits de la douleur et du plaisir. Enfin, nous présenterons notre modèle de travail des structures impliquées dans le fonctionnement des émotions.

1 Historique

1.1 Cannon et Bard : l'interaction entre le cortex et l'hypothalamus

Cannon et Bard ont apporté une contribution très importante pour l'approche psychologique des émotions (cf. "Le point de vue des psychologues"). Cependant, ils ont aussi été des précurseurs dans la compréhension des bases neurobiologiques des émotions. Ils défendent une vision centraliste selon laquelle une stimulation est traitée et évaluée par le système nerveux central avant de produire une réponse physiologique et comportementale non spécifique. Cette vision s'oppose à la vision périphéraliste dont James et Lange sont les porte-étendard qui défend au contraire que les émotions sont

issues du traitement par le système nerveux central des réponses corporelles. Leurs critiques à propos de la théorie James-Lange sont soutenues par des études sur les effets de lésions corticales sur le comportement émotionnel de chats. L'ablation du cortex a notamment pour effet de provoquer des comportements agressifs soudains et inappropriés qu'ils ont appelé *sham rage* (rage factice). Cependant, cette réaction disparaît si la lésion touche également la région postérieure de l'hypothalamus. Cannon et Bard soutenaient que si les émotions étaient en réalité la perception des changements corporels, elles devraient nécessiter que les cortex sensoriels et moteurs soient intacts. Les comportements de *sham rage* prouvaient donc selon eux que James et Lange se trompaient.

Cannon et Bard ont donc proposé la première théorie moderne des mécanismes cérébraux impliqués dans le fonctionnement des émotions [Bard 1928, Bard 1937]. Cette théorie décrit l'hypothalamus, une structure profonde, comme étant la région du cerveau qui produit les réponses émotionnelles en réaction aux stimulations de l'environnement. L'hypothalamus est une petite structure sous-corticale de la taille d'une noisette qui garantit l'équilibre biochimique du corps : l'homéostasie. L'hypothalamus contrôle la sécrétion de toutes les hormones du corps et supervise le fonctionnement de la plupart des organes internes. Nous savons aujourd'hui que cette structure participe à la régulation des principaux comportements de survie : faim, soif, reproduction, allaitement et agressivité. Le néocortex, organe plus récent du point de vue évolutionniste, permet d'inhiber la réponse de l'hypothalamus. L'ablation du cortex libère donc l'hypothalamus de son contrôle descendant et permet l'apparition des comportements incontrôlés de *sham rage*.

Les travaux de Cannon et Bard sont précurseurs aussi du point de vue des méthodologies employées. L'utilisation des animaux (suivant le point de vue darwiniste) et l'étude des lésions cérébrales qui leurs sont pratiquées chirurgicalement sont des méthodologies utilisées abondamment aujourd'hui encore dans le domaine des neurosciences affective.

1.2 Le circuit de Jakob-Papez

Le neurobiologiste argentin, Christofredo Jakob (1866-1956) a proposé l'idée qu'il existe un "cerveau viscéral" au sein du système nerveux central dont la fonction est de réguler les émotions [Jakob 1907]. Il en fit une description poussée, mais comme Jakob n'a publié qu'en allemand et en espagnol, son travail s'est très peu diffusé. C'est pourquoi ce n'est que trente ans plus tard que le neuroanatomiste américain James Papez (1883-1958) rendit célèbre ce circuit en publiant sa description anatomo-fonctionnelle en 1937 [Papez 1995].

Le circuit de Jakob-Papez est une boucle de régulation reliant le thalamus, l'hypothalamus, l'hippocampe et le cortex cingulaire :

- Le faisceau mammillo-thalamique relie les corps mamillaires de l'hypothalamus aux noyaux antérieurs du thalamus.
- Le genou de la capsule interne relie les noyaux antérieurs du thalamus au cortex cingulaire antérieur.
- Le cingulum relie le cortex cingulaire antérieur au gyrus parahippocampique.

- Le fornix relie l'hippocampe (plus particulièrement le subiculum) aux corps mamillaires.

Dans cette description initiale, le circuit n'est pas fermé puisqu'il manque la liaison entre le gyrus parahippocampique et l'hippocampe lui-même. Celle-ci est effectuée par le cortex entorhinal via la voie perforante.

Papez a proposé que les entrées sensorielles du thalamus divergent en deux voies : la voie "de la pensée" (*thought stream*) et la voie "du ressenti" (*feeling stream*). La voie "de la pensée" passe du thalamus au cortex cingulaire via les cortex sensoriels. Papez propose que cette voie continue au-delà du cortex cingulaire via le cingulum vers l'hippocampe, puis via le fornix vers le corps mamillaires de l'hypothalamus. Le long de cette voie, les sensations seraient transformées en perceptions, en pensées puis en souvenirs. La voie du ressenti passe directement du thalamus vers les corps mamillaires dont les projections descendantes régulent les systèmes corporels permettant ainsi la génération de la réponse corporelle émotionnelle. Cette voie se poursuit en remontant vers le cortex cingulaire en passant par les noyaux antérieurs du thalamus. Selon Papez, l'expérience émotionnelle est fonction de l'activité du cortex cingulaire qui peut être générée et régulée par ces deux voies. Les projections descendantes allant vers l'hypothalamus depuis le cortex cingulaire (voie de la pensée) permettent une régulation *top down* de la réponse émotionnelle (contrôle des structures profondes par le cortex). Elles sont cependant bien minoritaires par rapport aux connexions ascendantes allant vers le cortex cingulaire depuis l'hypothalamus (contrôle du cortex par les structures profondes). La figure 3.1 permet de visualiser cette boucle de contrôle entre le cortex et les structures profondes.

Ce circuit, ce cycle plutôt, formé par les voies du ressenti et de la pensée peut être compris comme un modèle de la base neurobiologique de la dualité de nature des émotions. D'un côté la composante corporelle et de l'autre l'expérience phénoménologique (i.e expérience de conscience).

Les hypothèses de Papez sont partiellement confirmées par la description du syndrome de Klüver-Bucy (ou cécité psychique) qui suit une lésion bilatérale des lobes temporaux (surtout s'il y a lésion hippocampale) [Klüver 1938, Klüver 1939]. Ce syndrome se caractérise par une perte des réactions émotionnelles (peur et colère essentiellement), une perte des interactions sociales, une agnosie visuelle et une hypersexualité. Il est intéressant de noter que les travaux de Klüver et Bucy ont très probablement été inspirés par les travaux publiés en 1888 de Brown et Shafer sur les fonctions des lobes occipitaux et temporaux [Brown 1888]. L'étude de ce syndrome a été un moteur du développement de la neuropsychologie des émotions, mais il a fallu attendre la fin de la deuxième guerre mondiale pour que de nouvelles avancées significatives fassent progresser le domaine.

1.3 Les trois cerveaux de MacLean

En 1949, l'américain Paul MacLean propose sa théorie des trois cerveaux ou *Triune Brain Theory* [MacLean 1949]. Au cours de l'évolution, trois cerveaux se sont superposés chez les mammifères. Le plus primitif, et donc le plus profond, est appelé le complexe (ou cerveau) reptilien. Il s'agit en fait essentiellement des ganglions de la

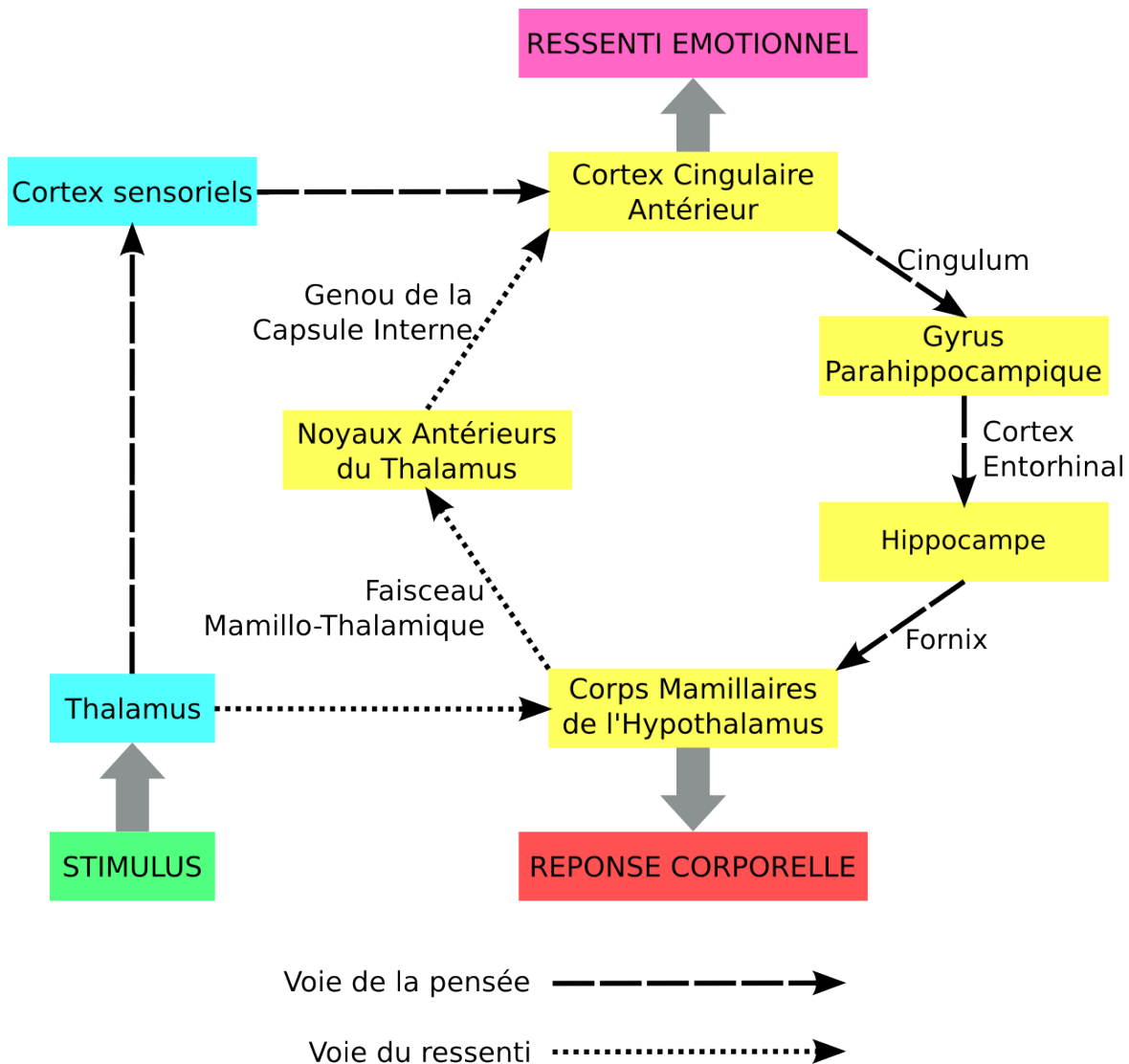


FIGURE 3.1 – Circuit de Jakob-Papez. La voie du ressenti va directement vers l'hypothalamus avant de remonter vers le cortex cingulaire. La voie de la pensée passe par les cortex sensoriels et le cortex cingulaire avant de redescendre vers l'hippocampe et l'hypothalamus.

base. Le terme "reptilien" vient du fait que les neuro-anatomistes pensaient que le pro-sencéphale des oiseaux et des reptiles est dominé par ces structures. Selon MacLean, le complexe reptilien est responsable des comportements instinctifs d'agression, de dominance ou de territorialité. Vient ensuite le complexe paléomammalien ou cerveau mammifère ancien (que MacLean appelait aussi *visceral brain*). Il s'agit du circuit de Jakob-Papez, thalamus, hypothalamus, hippocampe et cortex cingulaire mais aussi de l'amygdale, du cortex préfrontal, du septum... Dans un article publié en 1953, MacLean a introduit le terme de "système limbique" pour désigner cet ensemble de structures interconnectées [MacLean 1953]. La figure 3.2 illustre cette théorie. Le système limbique peut amplifier les réactions du cerveau reptilien mais il est aussi responsable des

motivations et des émotions impliquées dans les comportements alimentaires, sexuels et parentaux. Le néocortex ou nouveau cerveau mammalien permet de faire l'interface entre émotion et cognition et assure le contrôle descendant (*top-down*) sur les réponses émotionnelles des cerveaux plus anciens. L'idée essentielle de MacLean est qu'une expérience émotionnelle résulte de l'intégration des sensations du monde extérieur avec les informations venant du corps. Les événements du monde extérieur mènent à des changements corporels et les signaux sur ces changements sont intégrés par le cerveau à la perception courante. MacLean propose que cette fonction d'intégration soit le rôle du système limbique. Quel que soit le mérite de sa théorie des trois cerveaux, c'est pour le concept de système limbique comme système fonctionnel majeur que MacLean est reconnu. Même si aujourd'hui, le concept de système limbique en tant qu'unité fonctionnelle est attaqué, MacLean peut être considéré comme étant à la base d'une immense quantité de travaux sur les structures du système limbique. Il a donc permis des progrès importants dans la compréhension du fonctionnement et du rôle des structures profondes du cerveau.

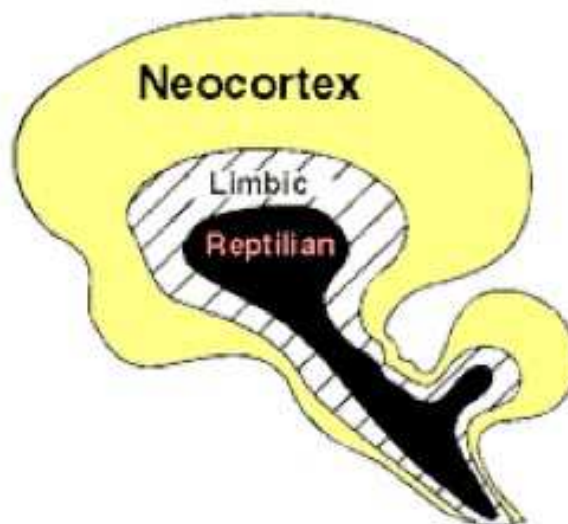


FIGURE 3.2 – Les trois cerveaux de MacLean. Le système nerveux central garde dans sa structure fonctionnelle aussi bien qu'anatomique les étapes principales de son évolution au travers de l'évolution des espèces. Ainsi, le cerveau reptilien assure la régulation corporelle et les comportements de territorialité, de dominance et d'agression. Le système limbique, propre aux mammifères, permet de contrôler le cerveau reptilien et assure le fonctionnement motivationnel. Enfin, le néocortex propre aux grands primates et aux cétacés permet une régulation descendante des cerveaux plus primitifs. D'après MacLean, il permet de faire le lien entre les émotions et la cognition.

2 Les principales structures cérébrales

Les soixante ans de travaux réalisés depuis MacLean ont permis de révéler l'implication d'un très grand nombre de structures cérébrales dans le fonctionnement des

émotions. Devant leur nombre mais surtout devant la diversité de leur localisation dans le cerveau, le concept de système limbique a progressivement perdu en pertinence [LeDoux 2003b]. Il est cependant encore fréquemment employé dans la littérature. Soit pour désigner le cerveau limbique de MacLean, soit sa version étendue qui inclut des parties du cerveau reptilien et du néocortex qui se sont révélées tenir des rôles clés. La figure 3.3 donne une bonne représentation de ces structures et de leur disposition dans le cerveau.

Le tableau que dépeint l'état de l'art des neurosciences affective est un système complexe et étendu de structures cérébrales organisé en sous-réseaux fonctionnels. Devant la multiplicité des structures impliquées et plus encore de leurs connexions réciproques, cet état de l'art ne prétend pas être exhaustif. Il comprend une description des principales structures et de leurs relations. Cette description est faite en suivant le développement phylogénétique du système nerveux central. Les structures du tronc cérébral (substance grise periaqueducale, aire tegmentale ventrale), puis sous-corticales (nucleus accumbens, amygdale, hippocampe) et enfin corticales (cortex préfrontal).

2.1 Structures du tronc cérébral

Le tronc cérébral est la partie la plus primitive du cerveau. Cela ne veut pas dire qu'il ne sert que de relais des afférences sensorielles et des commandes motrices et viscérales comme le décrit la vision traditionnelle. Deux zones notamment, l'aire tegmentale ventrale et la substance grise periaqueducale, jouent un rôle majeur dans le fonctionnement des sensations de plaisir et de douleur qui servent d'information de base pour les structures plus "hautes" ainsi que dans les comportements agressifs, défensifs et appétitifs basiques.

2.1.1 La substance grise periaqueducale

La substance grise periaqueducale est située le long de l'aqueduc de Sylvius qui traverse verticalement le tronc cérébral. Elle joue un rôle majeur dans la transmission mais aussi la modulation du signal de douleur [Tracey 2002, Rea 2009, Mor 2010]. Elle est aussi impliquée dans les comportements défensifs [Depaulis 1989, Carrive 1993, Pobbe 2010]. [Fanselow 1995] montre par exemple comment la stimulation de sa partie dorsale et latérale chez le rat peut provoquer des réponses comportementales défensives : immobilité subite, course, saut, tachycardie, hausse de la pression sanguine et du tonus musculaire. Au contraire, une lésion de cette zone rend le rat incapable de produire de tels comportements défensifs.

2.1.2 L'aire tegmentale ventrale

L'aire tegmentale ventrale est située dans le mésencéphale (zone médiane haute du tronc cérébral) à proximité de la partie postérieure de l'hippocampe et des corps mamillaires. Cette aire ne se caractérise pas parce qu'elle est un noyau anatomique bien défini mais parce qu'elle est la zone qui regroupe la plupart des corps cellulaires des neurones dopaminergiques du tronc cérébral. Ses neurones dopaminer-

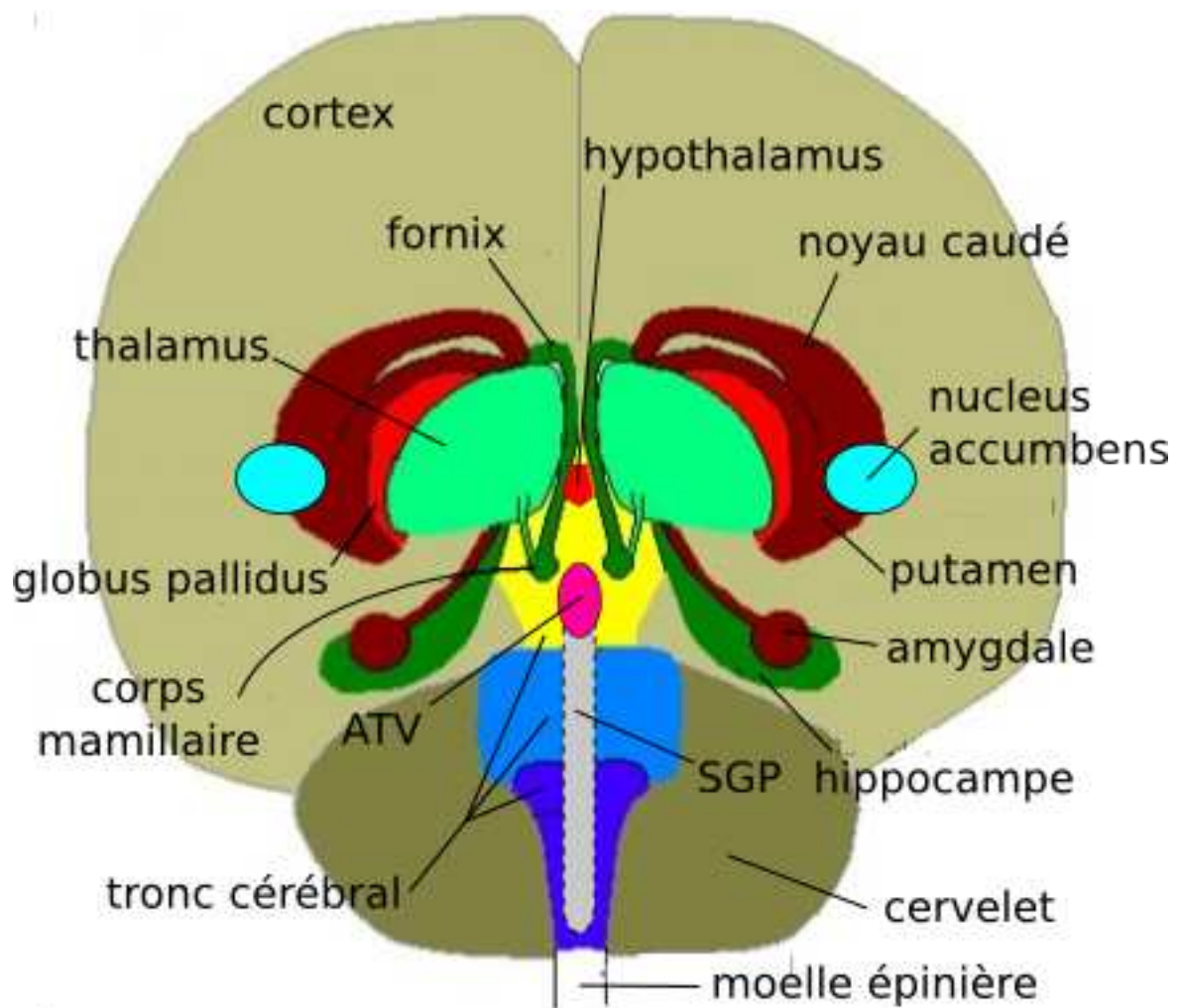


FIGURE 3.3 – Les principales structures impliquées dans le fonctionnement émotionnel du cerveau. ATV : aire tegmentale ventrale, SGP : substance grise périaqueducale.

giques en font une zone clé pour le fonctionnement du système de récompense et des motivations [Bals-Kubik 1993, Schultz 1997, Wise 2002, Chen 2008, Jalabert 2009]. C'est pourquoi l'aire tegmentale ventrale est impliquée dans les troubles addictifs [Hyman 2001, Nestler 2005]. L'aire tegmentale ventrale peut être considérée comme le départ des voies dopaminergiques mésolimbique et mésocorticale. Ces voies lui permettent de projeter vers un grand nombre de structures. Principalement le nucleus accumbens mais aussi le septum, l'hippocampe, l'amygdale et les cortex préfrontaux, orbito-frontaux et cingulaires. Ces connexions sont le plus souvent réciproques (contrôle et rétrocontrôle des circuits dopaminergiques). Des enregistrements cellulaires ont montré que les neurones de l'aire tegmentale ventrale répondent aux stimuli nouveaux [Lisman 2001, Bunzeck 2006], aux récompenses inattendues [Lisman 2005, D'Ardenne 2008] ainsi qu'aux indices perceptifs de prédiction de récompense [Schultz 1997, Nicola 2005, Fields 2007].

2.2 Les structures sous-corticales

2.2.1 L'hippocampe

L'hippocampe est une structure particulièrement complexe capitale pour un certain nombre de fonctions cognitives essentielles. Du fait de l'importance qui lui a précocement été reconnue, une littérature importante permet d'éclairer ses différents rôles et ses liens avec les autres structures permettant de réguler les émotions. Ses entrées sensorielles passent essentiellement par le cortex entorhinal mais aussi par le thalamus et le bulbe olfactif. L'hippocampe reçoit des connexions depuis une grande variété d'autres structures. Notamment l'amygdale, le nucleus accumbens et le cortex préfrontal. L'hippocampe projette en retour surtout vers l'amygdale, l'aire septale latérale, l'accumbens et le cortex préfrontal. La figure 3.4 schématise les entrées et sorties principales de l'hippocampe.

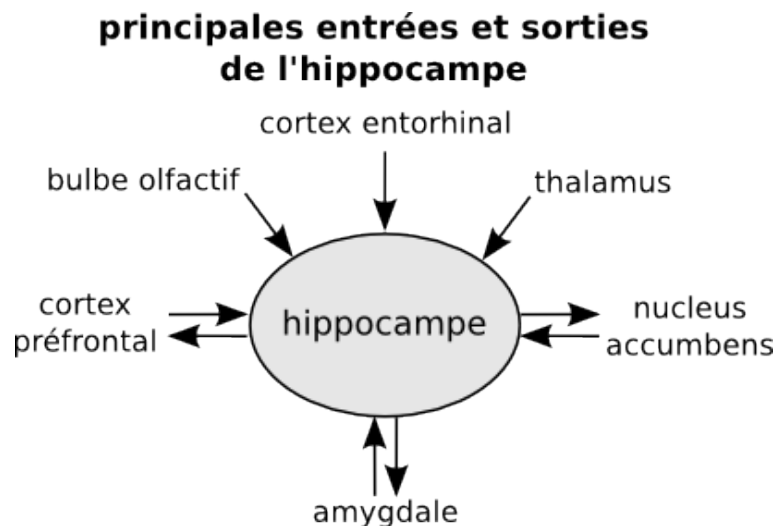


FIGURE 3.4 – Entrées et sorties principales de l'hippocampe.

Le système hippocampique est probablement majoritairement reconnu pour son

rôle dans la mémoire à long terme [LeDoux 1993, LaBar 2006] et la mémoire déclarative, contrairement à l'amygdale qui apprend des conditionnements [Bechara 1995, Zola-Morgan 1991]. Ses connexions réciproques avec l'amygdale permettent une modulation de la fonction de mémoire par la valence émotionnelle de l'activité de l'amygdale [Cahilla 1998, Richter-Levin 2004b, Richardson 2004, Phelps 2004]. C'est ce qui permet aux informations les plus pertinentes d'être mémorisées le plus facilement. En retour, un signal venant de l'hippocampe peut être évalué par l'amygdale en terme de valence émotionnelle positive et négative [Smith 2006]. Elle est aussi le siège de la mémoire spatiales [Bohbot 1998, Bird 2008]. Cette mémoire spatiale de la perception de l'environnement est fortement influencée par les informations visuelles mais aussi par les informations venant des autres sens et notamment du système vestibulaire [Stackman 2002, Russell 2003]. Cette fonction de l'hippocampe est nécessaire pour la navigation spatiale [O'Keefe 1971, Terrazas 2005, Morris 1982]. C'est pourquoi la maladie d'Alzheimer qui endommage progressivement l'hippocampe provoque des symptômes de perte de mémoire et de désorientation. L'hippocampe est aussi impliquée dans la détection de nouveauté [Crusio 1989, Knight 1996, Grunwald 1998, Lisman 2001, Shaomin 2003, Kumaran 2007, Kumaran 2009]. Nous décrirons plus loin le lien que nous faisons avec la réaction de surprise. L'hippocampe joue aussi un rôle central pour la navigation en tant que siège de la mémoire spatiale. Nous décrirons donc aussi le modèle computationnel développé dans notre laboratoire pour l'apprentissage de la navigation.

2.2.2 L'amygdale

La définition anatomique précise de l'amygdale et des noyaux qui la constitue a beaucoup évolué depuis sa découverte au début du 19^{ème} siècle. Aujourd'hui encore, l'amygdale est régulièrement redéfinie. Par exemple, Heimer [Heimer 2003] défend le concept d'une amygdale étendue qui comprend le groupe de noyaux situés à l'extrémité du noyau caudé ainsi qu'une part de la strie terminale, la bande de fibres nerveuses qui se situe sur le bord latéral de la surface ventriculaire du thalamus. Swanson et Petrovich [Swanson 1998] ont même proposé que l'amygdale n'existe pas en tant qu'unité structurelle puisqu'elle n'est constituée que de régions appartenant déjà à d'autres régions du cerveau. Selon ce point de vue, les zones latérales et basales sont considérées comme étant des projections du cortex, les zones centrales et médiales comme l'extension ventrale du striatum et le noyau cortical comme associé au système olfactif. Selon ce point de vue radical, il n'est nul besoin de parler d'amygdale. Cependant, pour des raisons pratiques, nous rejoindrons la vision la plus répandue qui définit l'amygdale comme l'agglomérat de noyaux qui constitue le renflement de la corne ventrale du striatum (cf. figure 3.5).

Les différents noyaux de l'amygdale reçoivent et émettent des projections vers un très grand nombre de structures différentes du cerveau. La figure 3.6 illustre schématiquement les principales entrées et sorties des différents noyaux (exceptions faites des connexions entre noyaux). Le noyau latéral reçoit des projections venant d'une part du thalamus et des cortex sensoriels et d'autre de l'hippocampe et du cortex entorhinal, la porte d'entrée corticale de l'hippocampe. Les noyaux basolatéraux et

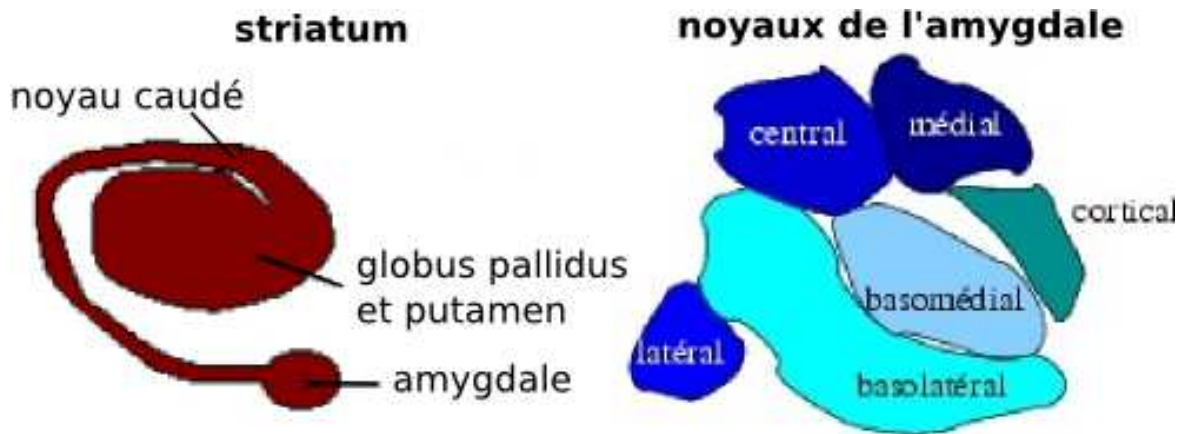


FIGURE 3.5 – A gauche : l'amygdale est le renflement qui termine la corne ventrale du striatum. A droite : une vue en coupe de l'amygdale révèle une structure interne. Les différents noyaux de l'amygdale ont des connectivités particulières et donc des fonctions spécifiques.

basomédians reçoivent des projections de l'hippocampe et du cortex entorhinal, mais ils reçoivent aussi des projections venant des cortex associatifs (cortex temporal). Le noyau central reçoit des entrées depuis le cortex préfrontal (orbitofrontal essentiellement), le cortex cingulaire, le cortex insulaire (ou insula), l'hypothalamus et les zones sensorielles du tronc cérébral (substance noire, aire tegmentale ventrale, raphe, locus coeruleus, noyau solitaire...). Les noyaux médiaux et corticaux reçoivent des projections du bulbe olfactif et du cortex olfactif. Les principales efférences de l'amygdale partent depuis les noyaux basaux et le noyau central. Les noyaux basaux projettent vers le cortex préfrontal, les cortex associatifs et le reste du striatum. Le noyau central projette lui essentiellement vers la substance grise périaqueducule et les structures régulant les systèmes de neuromodulation, notamment l'hypothalamus (aire septo-préoptico-hypothalamique) et le tronc cérébral (la substance noire, le noyau du raphé et la formation réticulée). Les systèmes de neuromodulation sont des voies chimiques permettant de moduler le fonctionnement des neurones. Il existe quatre neuromodulateurs principaux : la dopamine (DA), la sérotonine (5HT), la norépinéphrine (NE), l'acétylcholine (ACh). La dopamine est impliquée dans le fonctionnement des mouvements volontaires, des motivations, des récompenses et des punitions, du plaisir, du sommeil, de l'humeur, de l'attention, de la mémoire de travail, de l'apprentissage. Afin d'unifier ces différents rôles et d'arriver à une compréhension plus fondamentale de la dopamine, [Schultz 2007] propose qu'elle assure la médiation de la réactivité de l'organisme à l'environnement à différentes échelles de temps. La sérotonine joue un rôle dans la régulation des comportements alimentaires et sexuels, du cycle veille-sommeil, de la douleur, du contrôle moteur, de l'anxiété. La norépinéphrine est impliquée dans la régulation des processus attentionnels ainsi que la régulation de l'humeur. L'acétylcholine joue un rôle important dans les fonctions d'*arousal* et de récompense. Elle est aussi impliquée dans le contrôle attentionnel.

Vers la fin des années 30, les recherches de Klüver et Bucy ont souligné l'importance

principales entrées et sorties de l'amygdale

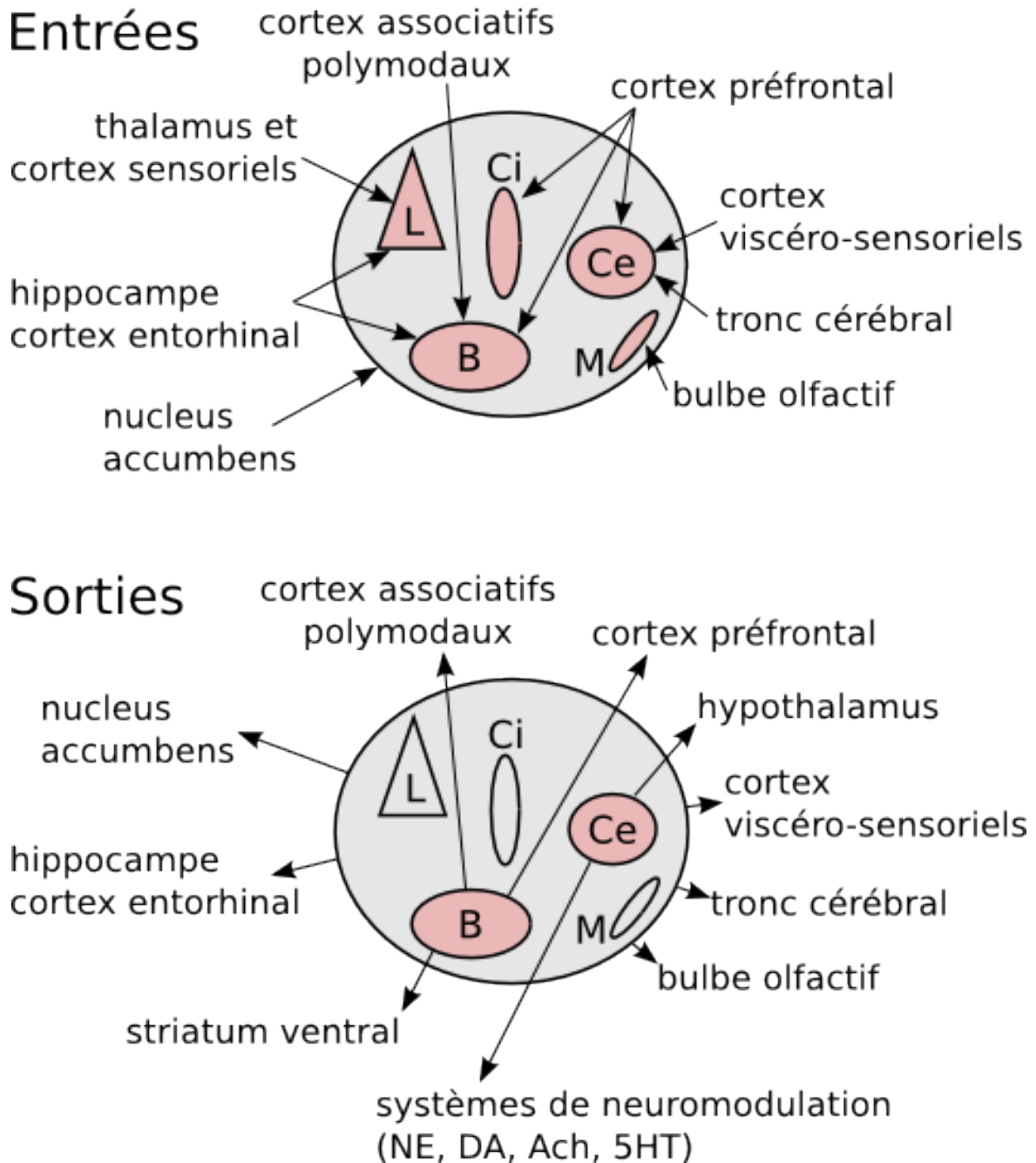


FIGURE 3.6 – Les principales afférences (entrées) et efférences (sorties) des noyaux de l'amygdale. L'amygdale reçoit des entrées depuis des zones perceptives extérieures (vision, odorat, audition, ...) et intérieures (viscères, douleur, plaisir ...), depuis le cortex préfrontal (régulation top-down) et depuis l'hippocampe (mémoire spatiale). L'amygdale projette en retour vers la plupart de ses structures d'entrée mais aussi vers le reste du striatum ventral et vers les systèmes de neuromodulation. L : noyau latéral, Ci : cellules intercalées, Ce : noyau central, M : noyau médial, B : noyau basal, NE : norepinéphrine, DA : dopamine, ACh : acétylcholine, 5HT : sérotonine.

de lésions des lobes temporaux (comprenant les amygdales) dans le fonctionnement de la peur et des comportements sexuels et alimentaires. Depuis le milieu des années 50, il existe un relatif consensus quant à l'implication de l'amygdale dans ses modifications de comportements et sur l'implication de l'amygdale dans l'ensemble des structures cérébrales permettant la régulation émotionnelle. Un très grand nombre d'études ont donc été menées sur les propriétés de l'amygdale et sur son implication dans le fonctionnement de la peur et **l'apprentissage de conditionnements émotionnels**. C'est sûrement la popularité de ce sujet d'étude qui a fait de la peur une des émotions les mieux comprises en termes de mécanismes cérébraux. Les connexions entrantes et sortantes de l'amygdale ont été brièvement décrites plus haut, mais il existe aussi des connexions entre les noyaux de l'amygdale. Celles-ci convergent plutôt vers le noyau central et les noyaux basaux [Fudge 2009]. Le noyau latéral et les noyaux basaux sont les portes d'entrée sensorielles extérieures de l'amygdale (exception faite de l'olfaction). Ils projettent vers le noyau central qui reçoit aussi des entrées sensorielles internes (informations viscérales, douleur) et des projections du cortex préfrontal. Les travaux présentés dans [Cardinal 2002] soutiennent l'idée que les conditionnements "stimulus-stimulus" (conditionnements classiques) ont lieu dans le noyau basolatéral mais que les conditionnements "stimulus-réponse motrice" (conditionnements instrumentaux) ont lieu dans le noyau central. Le noyau central serait le noyau recevant les informations nécessaires pour apprendre le conditionnement et produisant la réponse de l'amygdale (en particulier pour la peur) [Saha 2005].

Cet intérêt particulier pour l'amygdale a permis de constituer un vaste corpus d'études montrant l'implication de l'amygdale dans le fonctionnement des conditionnements émotionnels [Grossberg 2008], aussi bien pour les conditionnements de comportements aversifs [Dunsmoor 2009, Blair 2005, Knight 2005, Maren 2005] que de comportements appétitifs [Fudge 2000, Baxter 2002, Cardinal 2002, Balleine 2003, Balleine 2006]. En dehors des conditionnement lié à la peur, l'amygdale est donc aussi impliquée dans les mécanismes de récompense et de motivation. Cependant, moins de travaux permettent d'éclairer les détails de fonctionnement de ces mécanismes. Les conclusions de ces études sur les rôles multiples de l'amygdale sont supportées par des études anatomiques montrant que les afférences de l'amygdale transportent l'information concernant aussi bien les événements positifs que négatifs [McDonald 1998, Pitkanen 2000]. Ces études montrent que ce regroupement de projections recouvrant une part substantielle du cerveau donne à l'amygdale un rôle central dans le fonctionnement des émotions : la capacité à donner une valeur émotionnelle positive ou négative aux stimulations entrantes [Paton 2006].

L'exemple suivant est une illustration du processus de conditionnement de la peur tel que décrit par J.E. Ledoux au travers de ses nombreux articles sur la question [Sigurdsson 2007, LeDoux 1990, LeDoux 2003a, LeDoux 2007]. L'expérience menée sur des souris est tout à fait classique. Le dispositif expérimental permet de délivrer un choc électrique à une souris ainsi que de faire retentir un son spécifique. Quand le choc électrique est délivré, la souris s'immobilise (réaction de peur). Le son est le stimulus neutre (stimulus conditionnel) qui doit être conditionné par la perception de douleur (stimulus inconditionnel). Si le son est émit à chaque fois que l'expérimentateur délivre

un choc électrique, la souris apprend à produire une réaction d'immobilisation alors que seul le son est émis. La figure 3.7 décrit ce mécanisme de conditionnement de la peur.

conditionnement de la peur par l'audition

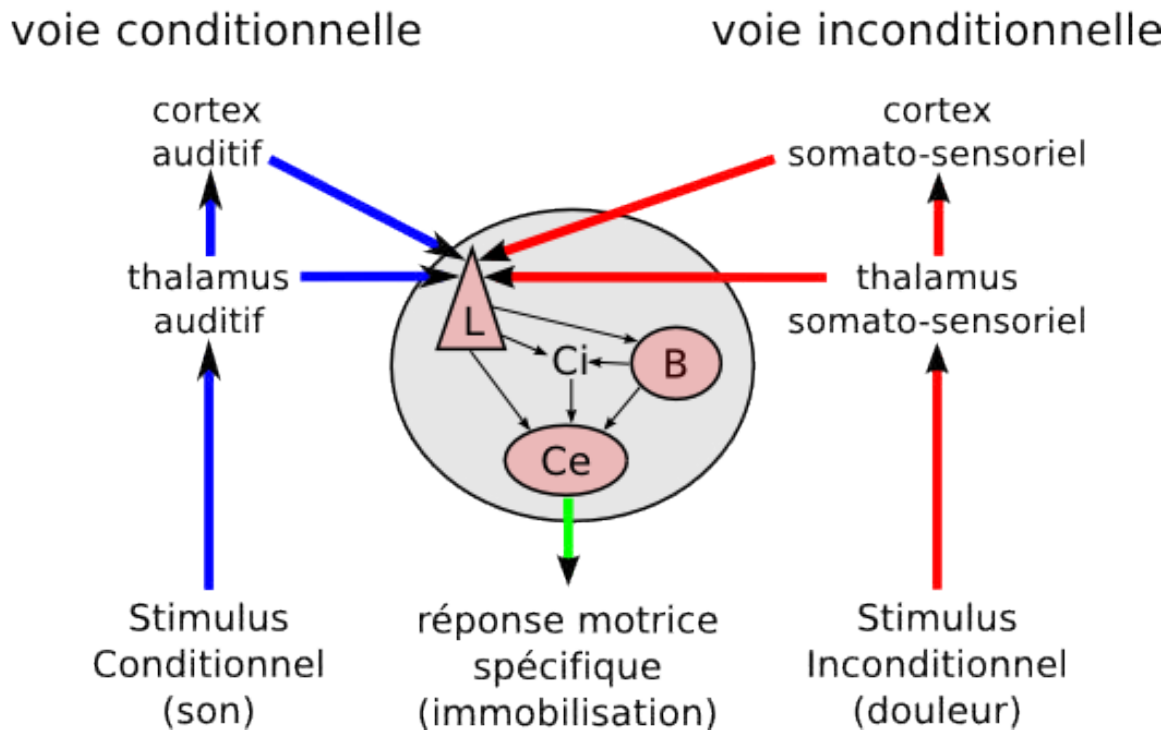


FIGURE 3.7 – Apprentissage d'un conditionnement de la peur à l'aide d'une stimulation auditive : si un son spécifique (stimulus conditionnel) est perçue de manière répétée en même qu'un stimulation douloureuse (stimulus inconditionnel), une association peut alors être apprise (conditionnement). Le son produit alors la même réaction que la douleur (immobilisation).

Cette capacité à apprendre des conditionnements entre les informations sensorielles de l'environnement et celles du corps fait de l'amygdale une structure centrale du fonctionnement des émotions. Elle est impliquée dans les états émotionnels associés aux comportements agressifs, maternels, sexuels et alimentaire. Etant donné que l'amygdale apprend et stocke des informations relatives aux événements émotionnels, on dit couramment qu'elle participe à la mémoire émotionnelle. Cette mémoire émotionnelle diffère de la mémoire explicite ou déclarative sous-tendue par l'hippocampe en ce sens qu'elle est implicite et inconsciente. En plus de cette fonction, l'amygdale est impliquée dans la régulation de fonctions cognitives comme la perception, l'attention ou la mémoire explicite. Le point de vue le plus couramment admis est que l'amygdale module ces fonctions cognitives en réponse à l'évaluation de la valeur émotionnelle des stimulations venant de l'environnement. Les sorties de l'amygdale mènent à la libéra-

tion d'hormones et de neurotransmetteurs permettant de moduler le fonctionnement de processus cognitifs assurés par le cortex.

Un grand nombre des troubles psychiatriques sont corrélés avec des changements aussi bien structuraux que fonctionnels de l'amygdale. Notamment, les troubles anxieux (phobies [Stein 2002, Furmark 2004, Phan 2006, Birbaumer 1998], le stress post traumatique [Shekhar 2006, Hull 2002, Yehuda 2002, Rauch 2000]), la dépression [Pezawas 2005, Sheline 1998, Drevets 1999, Mervaala 2000], la schizophrénie [Bogerts 1993, Schneider 1998, Lawrie 2006, Williams 2004], l'autisme [Baron-Cohen 2000, Aylward 1999, Schultz 2005, Howard 2000], ... Cette corrélation ne fait pas toujours des anomalies de l'amygdale la cause du trouble psychiatrique concerné. Cependant, elle montre clairement l'implication de l'amygdale dans la régulation émotionnelle du comportement.

2.2.3 Le nucleus accumbens

Le nucleus accumbens reçoit des entrées essentiellement depuis le cortex préfrontal, l'amygdale baso-latérale et l'aire tegmentale ventrale mais aussi depuis l'hippocampe. Ses sorties vont d'une part vers le pallidum ventral, l'hippocampe et l'amygdale (voie mésolimbique) ainsi que la substance noire et de manière réciproque vers l'aire tegmentale ventrale [Walaas 1980] et d'autre part vers le cortex préfrontal (voie mésocorticale). La figure 3.8 présente les différentes entrées et sorties du nucleus accumbens.

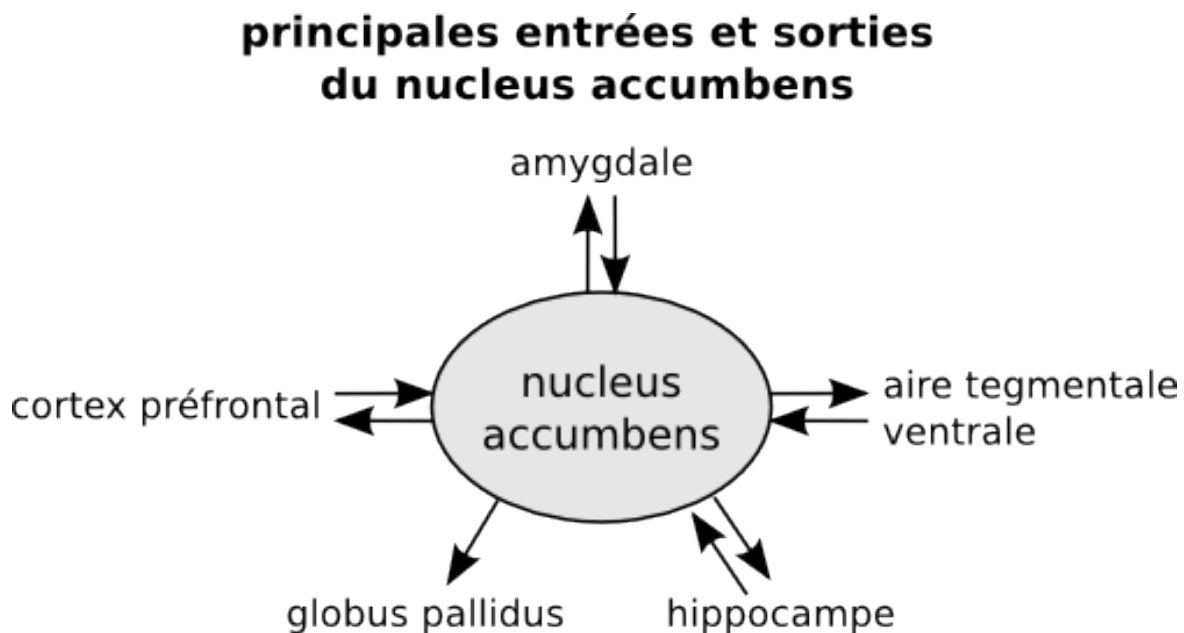


FIGURE 3.8 – Le nucleus accumbens reçoit des entrées depuis l'aire tegmentale ventrale, l'amygdale et le cortex préfrontal et projette d'une part vers le cortex préfrontal le long de la voie mésocorticale et vers l'hippocampe, l'amygdale et le pallidum ventral le long de la voie mésolimbique.

Sur le plan fonctionnel, le nucleus accumbens est le plus souvent associé à

la notion de circuit de récompense. Il est impliqué dans la sensation de plaisir [Berridge 2003, Burgdorf 2006], le rire [Osaka 2005, Okun 2007] et les comportements addictifs [Chiara 2002, Chiara 2004]. L'accumbens ne permet pas seulement de diffuser le signal de récompense. Il est aussi impliqué dans les conditionnements instrumentaux et les comportements motivés par la recherche de plaisir [Salamone 1994, Corbit 2001]. Les projections qu'il reçoit depuis les neurones dopaminergiques de l'aire tegmentale ventrale en font une structure clé de la transmission des messages de récompense ainsi que de la sensation de plaisir. C'est à partir du nucleus accumbens que les voies dopaminergiques mésolimbique et mésocorticale se séparent. D'un côté vers le pallidum ventral, l'amygdale et l'hippocampe et de l'autre vers le cortex préfrontal. Il semble cependant que le nucleus accumbens soit aussi impliqué dans les comportements agressifs [Lee 1983, Ferrari 2003] et certains comportements de peur [Levita 2002, Schwienbacher 2004]. Le nucleus accumbens et l'aire tegmentale ventrale sont capitales dans le contrôle de la sensation de plaisir mais aussi dans l'action contre la sensation de douleur. Ces deux structures permettent au corps de produire ses propres antalgiques. Par exemple, l'altération chimique du fonctionnement du nucleus accumbens (avec un antagoniste de la dopamine) permet de moduler l'effet antalgique d'une stimulation de l'aire tegmentale ventrale [Altier 1998]. Nous reviendrons un peu plus loin sur les circuits du plaisir et les voies mésolimbique et mésocorticale.

2.2.4 L'hypothalamus

L'hypothalamus est une glande cérébrale située sous le thalamus et au niveau du sommet du tronc cérébral. Son rôle principal est de permettre l'interface entre le système nerveux et le système endocrinien. L'hypothalamus reçoit de multiples connexions via le fornix, le thalamus, les corps mamillaires, le tronc cérébral, ... [Nieuwenhuys 2008]. C'est ce qui lui permet de répondre à un grand nombre de sensations externes, comme la lumière ou la présence de phéromones, et internes, comme l'intéroception viscérale, les paramètres sanguins (insuline, glycémie, ...), le stress, les attaques microbiennes, ... Le relais vers le système endocrinien à proprement parler se fait par le biais de l'hypophyse, une glande impliquée dans la régulation du système nerveux autonome ainsi que des fonctions endocrines. Il intervient ainsi dans la régulation d'un ensemble varié de fonctions corporelles et comportementales dont : le contrôle du rythme circadien (réponse à la lumière), les comportements reproductifs (réponse aux phéromones), la faim et la soif (réponses aux sensations viscérales et aux paramètres sanguins), la thermorégulation (réponse immunitaire). Pour ces motivations de bas niveau, nous parlons de drives.

2.3 Le cortex préfrontal

Le cortex préfrontal est la zone du cortex situé à l'avant du cerveau (lobe frontal). C'est la partie du cerveau qui s'est le plus développée lors de l'évolution du cerveau des hominidés pour occuper aujourd'hui environ un tiers du cortex. Son importance a longtemps été largement sous-estimée et c'est ce qui a permis l'essor des leucotomies, lobotomies et lobectomie frontale (de la section de certaines connexions du cortex

préfrontal à son ablation) pour traiter les symptômes de certaines maladies mentales. Aujourd'hui, les procédures médicales sont les plus légères possibles et ne sont utilisées qu'en ultime recours. L'étude du cas célèbre de Phineas Gage permet d'éclairer le rôle du cortex préfrontal. Même si elle n'a eu que très peu de conséquences à son époque dans le monde de la neurologie, elle demeure néanmoins la première à établir un lien entre une lésion cérébrale et un changement de personnalité. Gage était un jeune et brillant chef d'équipe dans la construction de chemin de fer. Lors d'un accident sur un chantier, une barre métallique fut projetée à travers sa tête, entrant par sa joue gauche, traversant la partie avant de son cerveau et ressortant par le haut du crâne. Gage survécut à son accident et se remit en moins de deux mois. Cependant son médecin constata que la personnalité de Gage avait radicalement changé : *«Cet homme, très équilibré avant son accident, considéré comme très fin et habile en affaires, capable d'énergie et de persévérance dans l'exécution de tous ses plans d'action, était devenu d'humeur changeante, irrévérencieux ; proférait parfois les plus grossiers jurons (ce qu'il ne faisait jamais auparavant), ne manifestait que peu de respect pour ses amis, supportait difficilement les contraintes ou les conseils lorsqu'ils venaient entraver ses désirs ; s'obstinait parfois de façon persistante tout en étant capricieux et inconstant ; formait quantité de projets aussitôt abandonnés ; se comportant comme un enfant, il avait néanmoins les pulsions animales d'un homme vigoureux ; il employait un langage tellement grossier qu'on avertissait les dames de ne pas rester longtemps en sa présence si elles ne voulaient pas être choquées...C'était un homme si différent que ses employeurs n'ont pu le garder...La chute du statut social s'est alors enclenchée...Il ne semblait réussir qu'à trouver des emplois qui ne lui convenaient pas...Il a fait partie du spectacle présenté par le cirque Barnum à New York où il montrait orgueilleusement ses blessures ainsi que la barre de fer»* (citation retranscrite depuis [Lotstra 2002]). Bien que Harlow, le médecin de Gage, fut bien conscient des implications d'une telle découverte, son étude publiée en 1848 [Harlow 1848] n'a pas eu d'impact particulier à l'époque. Peut-être parce que l'idée que notre sens moral pourrait être dépendant d'une partie précise de notre cerveau était alors inacceptable. Peut-être aussi parce que Harlow, n'avait pas l'aura scientifique d'un Broca ou d'un Wernicke. Quoiqu'il en soit, ce cas n'est devenu célèbre que lorsqu'il a été redécouvert par Damasio [Damasio 1994] et décrit dans son best-seller *L'erreur de Descartes* [Damasio 1995]. Le cortex préfrontal est aujourd'hui considéré comme responsable de l'orchestration des pensées et des actions en fonction des buts internes. C'est pourquoi on dit qu'il réalise les fonctions exécutives. Cependant une zone aussi grande remplit un grand nombre de fonctions différentes. La figure 3.9 situe le cortex préfrontal dans le cerveau et illustre la manière dont il se subdivise en sous-région. Une reconstruction 3D de la blessure de Gage réalisée à partir de son crâne semble indiquer une lésion de la zone orbitofrontale gauche.

Si les patients souffrant d'une lésion du cortex préfrontal peuvent répondre verbalement de manière normale à des questions liées aux comportements sociaux et que leur intelligence n'est pas atteinte, leur comportement réel n'en est pas moins affecté par une tendance forte à la recherche de gratification immédiate et l'incapacité à évaluer un profit supérieur plus lointain. Cela a des conséquences désastreuses

les différentes zones du cortex préfrontal

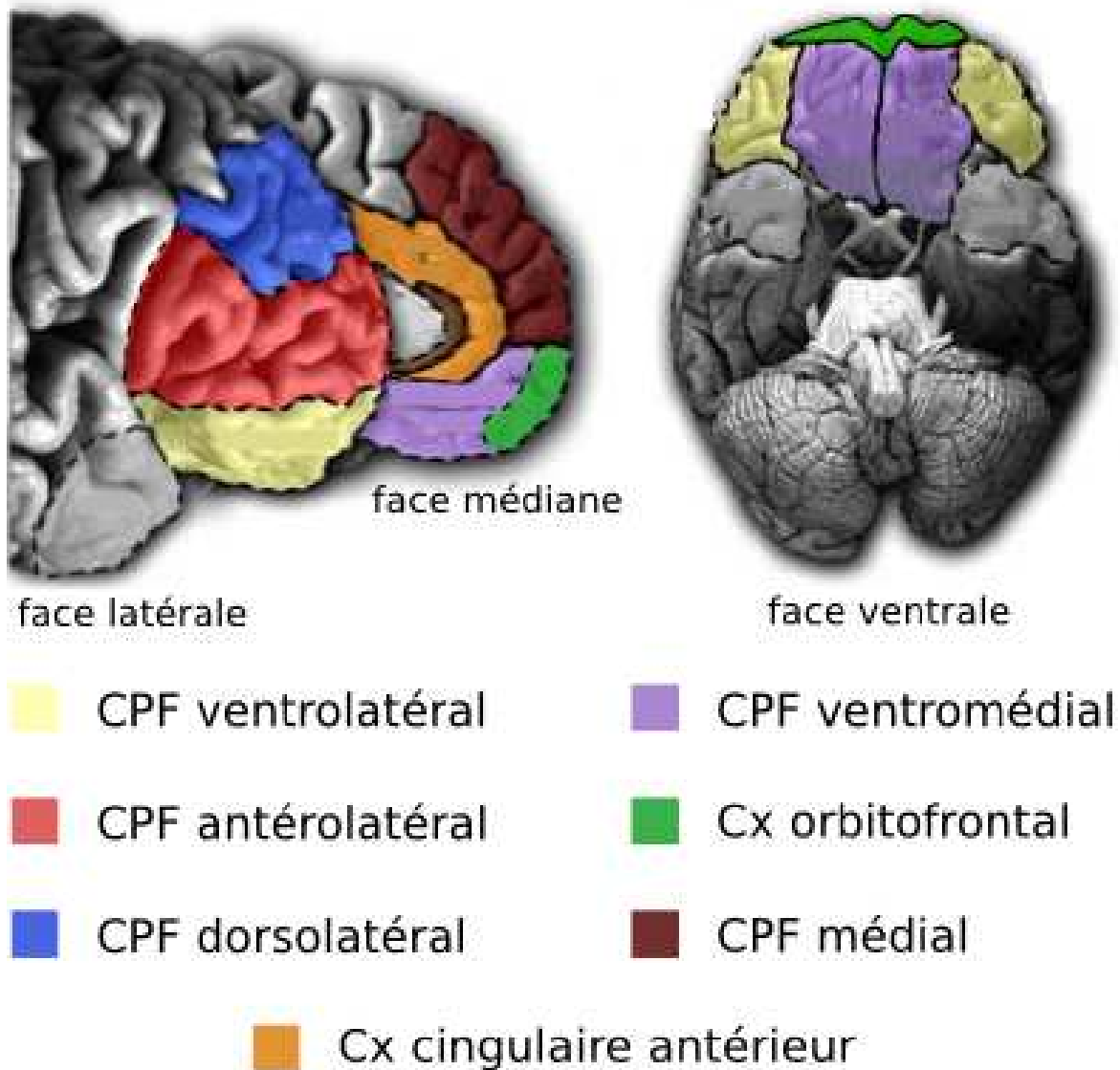


FIGURE 3.9 – Le cortex préfrontal et ses sub-divisiones : zone ventrolatérale, ventromédiale, orbitofrontale, médiale, dorso-latérale et le cortex cingulaire antérieur. Chaque zone est impliquée plus particulièrement dans un type de traitement cognitif. Globalement, le cortex préfrontal est considéré comme assurant les fonctions exécutives (raisonnement, prise de décision, inhibition de comportement, ...)

puisque cette capacité est un des éléments clés permettant de contrôler le comportement. [Shimamura 2000] propose que le cortex préfrontal agit comme un filtre qui sur-active les activations dirigées vers les buts et inhibe les activations non pertinentes. La littérature sur le cortex préfrontal est pléthorique mais les fonctions essentielles qu'il remplit sont centrées sur la prise de décision et la planification en fonctions des récompenses et punitions potentielles [Bechara 2000, Koechlin 2000]. Aussi bien pour des stimulations sensorielles basiques que pour des récompenses et punitions plus abstraites comme un gain ou une perte d'argent [O'Doherty 2001]. Il s'agit de la représentation du potentiel de récompense des stimuli et de l'apprentissage rapide de l'association entre stimuli et issues positives ou négatives [Hornak 2004, Rolls 2004]. Le cortex préfrontal est aussi impliqué dans la détection de nouveauté [Tulving 1994, Knight 1995, Persson 2002, Dias 2002]. C'est pourquoi il est nécessaire pour s'adapter rapidement à des conditions changeantes. Ces fonctions sont réalisées par la régulation des structures sous-corticales [Etkin 2006, Diorio 1993, Vertes 2006]. Un certain nombre de modèles computationnels du cortex préfrontal ont été proposés depuis la fin des années 80. Je citerai notamment les modèles de Stanislas Dehaene, Jean Pierre Changeux et Lionel Naccache [Dehaene 1989, Dehaene 1991, Dehaene 1995, Dehaene 2006] ainsi que le modèle d'Edelman et Tononi [Tononi 1998, Edelman 2001].

3 Neurobiologie de la douleur et du plaisir

Pour compléter ce tableau de structures impliquées dans les émotions, nous allons nous intéresser à la neurobiologie de la douleur et du plaisir. Douleur et plaisir sont des composantes essentielles des émotions et qui ont le mérite d'avoir été beaucoup étudié

3.1 Douleur

Il y a un peu plus de 100 ans, le neurophysiologiste anglais Charles Sherrington proposait que la sensation de douleur est générée à partir de capteurs spécifiques : les nocicepteurs [Sherrington 1906]. Au moins trois types de nocicepteurs ont été identifiés en fonction de la nature de la sensation à laquelle ils répondent : chaleur intense, forte pression mécanique ou agents chimiques irritants [Julius 2001]. L'information remonte par la voie afférente depuis la capteur jusqu'à la moelle épinière. De là, une bifurcation intervient : le signal de douleur déclenche une réaction réflexe directement depuis la moelle épinière vers le muscle via la voie afférente, mais l'information continue pour autant son trajet vers le cerveau afin d'y être traitée extensivement [Purves 2008]. D'un côté il y a l'arc réflexe qui permet une réponse très rapide de retrait. De l'autre côté, la voie "longue" permet à la douleur d'être ressentie, identifiée et évaluée. Le signal de douleur peut alors intervenir à de multiples niveaux (perception, attention, mémoire, apprentissage, réaction motrice plus élaborée, ...). Le but n'est pas d'entrer en profondeur dans le traitement de la douleur par le cerveau. Il est cependant remarquable de voir la diversité des structures cérébrales impliquées. Depuis la substance grise périaqueducale (dans le tronc cérébral), les cortex sensoriels primaires et secon-

dares, le thalamus, l'insula, le cortex pré-moteur et le cortex pariétal inférieur, jusqu'à l'hippocampe, le cortex entorhinal, le cortex cingulaire, le cortex orbito-frontal et les cortex pré-frontaux dorso-latéraux et médians. Cette diversité reflète la multiplicité des niveaux auxquels la douleur est traitée par le cerveau. Elle permet un traitement de la douleur aussi bien sensoriel, qu'affectif et cognitif. C'est ce qui distingue une réponse rapide et basée sur une bonne reconnaissance d'une réponse plus lente et prolongée, dépendante d'une évaluation plus profonde de la situation et dépendante de l'expérience d'interactions passées.

La figure 3.10 adaptée depuis [Casey 2006] est le résultat de la compilation des résultats d'un grand nombre d'études d'imagerie cérébrale. Elle représente les structures activées par la douleur et les catégorise selon leur implication supposée aux différents niveaux de traitement de l'information par le cerveau. Les fibres sensorielles portant le signal sensoriel de douleur arrivent par la moelle épinière via le tronc cérébral en activant notamment la substance grise périaqueducale. De là, le signal se répand à travers le cerveau. Dans un premier temps vers le thalamus, l'insula, les cortex sensoriels et le cortex cingulaire (traitements sensoriels et identification précoce). Vers l'amygdale, le cortex pariétal et le cortex pré-moteur ensuite (reconnaissance et réponse motrice immédiate). Et enfin vers l'hippocampe et le cortex préfrontal (évaluation, mémorisation et comportements maintenus).

3.2 Plaisir

Du point de vue du fonctionnement du cerveau, la notion de plaisir est étroitement liée aux effets de la dopamine [Wise 1980, Bressan 2005, Ikemoto 2007]. Ce neurotransmetteur est donc associé au fonctionnement du système de récompense. Il existe huit voies dopaminergiques, mais concernant la sensation de plaisir, les deux principales sont la voie mésolimbique et la voie mésocorticale. Elles assurent la diffusion de la dopamine depuis les neurones émetteurs situés dans la partie supérieure du tronc cérébral (mésencéphale) vers les structures limbiques (voie mésolimbique) et le cortex préfrontal (voie mésocorticale). Les premières études ayant permis de mettre en évidence ce circuit du plaisir ont été menées dans les années 50. James Olds et Peter Milner ont utilisé la stimulation électrique directe de la zone septale du cerveau de rats pour renforcer des conditionnements instrumentaux [Olds 1954]. Depuis, un grand nombre d'études ont permis de préciser nos connaissances sur le fonctionnement du plaisir.

La voie mésocorticale et la voie mésolimbique partent de l'aire tegmentale ventrale où résident les corps cellulaires de la plupart des neurones dopaminergiques du tronc cérébral. Ces neurones projettent leurs axones vers le nucleus accumbens. La voie mésolimbique distribue ensuite la dopamine vers les structures sous corticales. Principalement l'amygdale, le pallidum ventral et l'hippocampe. Une controverse existe encore aujourd'hui quant au rôle de cette voie mésolimbique dans le système de récompense. Les trois hypothèses principales attribuent à cette voie un rôle pour l'apprentissage, la sensation de plaisir ou la motivation (*incentive value*) [Berridge 1998, Schultz 1998]. La voie mésocorticale projette quant à elle vers les cortex perceptifs, associatifs, moteurs et limbiques. Cette voie est impliquée dans le fonctionnement des émotions et

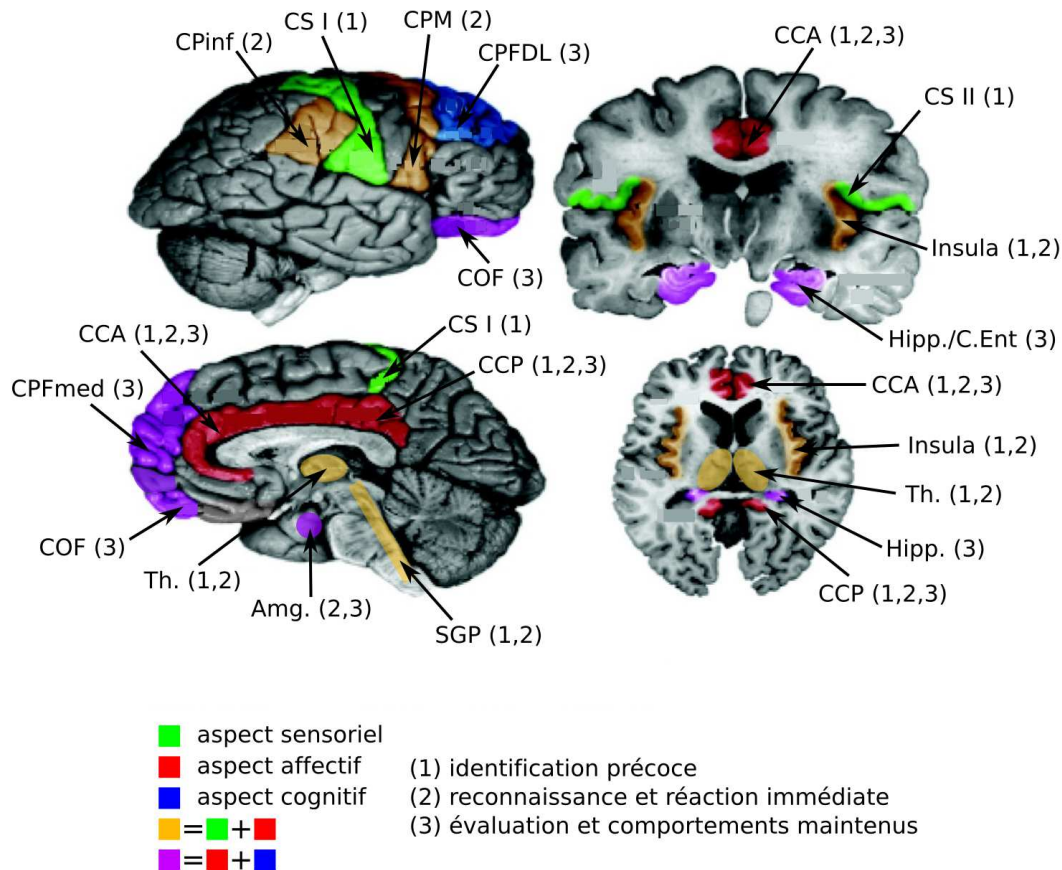


FIGURE 3.10 – Les aires cérébrales impliquées dans le traitement de la douleur. Le code couleur reflète le rôle supposé de chacune de ces zones dans le traitement des différentes dimensions de la douleur. Le nombre entre parenthèse indique l'implication relative de ces aires durant les différentes étapes de l'expérience douloureuse. Les aires impliquées sont : la substance grise périacqueducale (SGP), le thalamus (Th.), l'insula, l'amygdale (Amg.), l'hippocampe (Hipp.), le cortex entorhinal (C.Ent), le cortex cingulaire antérieur (CCA) et postérieur (CCP), le cortex sensoriel primaire (CS I) et secondaire (CS II), le cortex pariétal inférieur (CPinf), le cortex préfrontal dorso-latéral (CPFDL), le cortex pré-moteur (CPM), le cortex orbito-frontal (COF) et le cortex préfrontal médian (CPFmed). Schéma adapté et complété depuis [Casey 2006]

des motivations mais à un plus haut niveau de traitement de l'information. La figure 3.11 situe ces deux voies dans le cerveau.

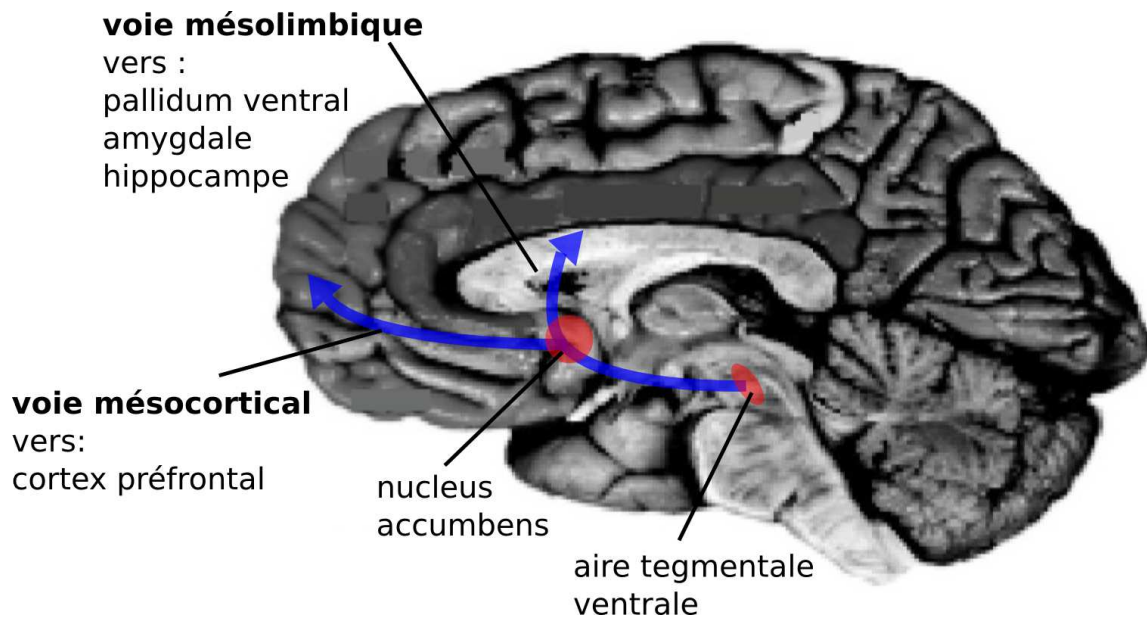


FIGURE 3.11 – Circuits du plaisir. Les voies dopaminergiques mésolimbique et mésocorticale prennent naissance dans l'aire tegmentale ventrale (partie supérieure du tronc cérébral) et passent par le nucleus accumbens. Tandis que la voie mésolimbique se poursuit vers les structures limbiques (pallidum ventral, amygdale et hippocampe essentiellement), la voie mésocorticale se poursuit vers les lobes frontaux (notamment le cortex préfrontal). Ces deux voies assurent la distribution des neuromodulations dopaminergiques liées au plaisir.

4 Discussion

La synthèse des descriptions anatomiques et fonctionnelles des principales structures cérébrales impliquées dans les processus émotionnels permet l'élaboration d'un modèle de travail (figure 3.12). Le propos de ce modèle n'est en aucun cas d'apporter une représentation exhaustive et définitive des structures cérébrales impliquées dans le fonctionnement des émotions et de leurs interactions. Nous n'attendons pas non plus de lui qu'il nous permette de comprendre finement des émotions complexes. Néanmoins, il décrit les bases neurobiologiques des processus émotionnels de bas niveau et permet de mieux comprendre l'émergence de signaux émotionnels.

Essentiellement, ce modèle de travail consiste à décrire un sous-ensemble de structures cérébrales formant un système permettant de réguler le comportement par l'émergence de mécanismes émotionnels de bas niveau. Afin d'arriver à un modèle fonctionnel et intelligible, nous simplifions le rôle fonctionnel de chacune de ces structures en le limitant à son aspect le plus important où le plus pertinent dans notre contexte.

Un certain nombre de signaux forment les entrées de ce système. D'un côté, il y a les afférences corporelles arrivant via la moelle épinière et le tronc cérébral : toucher, proprioception, intéroception viscérale, paramètres sanguins (hypothalamus) ainsi que les signaux de douleur (SGP) et de plaisir (ATV). De l'autre, il y a les perceptions externes (vue, odorat, ouïe, goût) toutes relayées par le thalamus.

Nous faisons intervenir ensuite trois structures sous-corticales majeures : le complexe hippocampique, l'amygdale et le nucleus accumbens. Parmi le très grand nombre de fonctions réalisées par l'hippocampe, nous retenons ici son rôle dans l'intégration des informations sensorielles externes (via le thalamus) permettant de construire un contexte perceptif (i.e. mémoire spatiale). Bien que peu développé dans ce travail, nous avons aussi considéré le rôle de l'hippocampe dans la détection de nouveauté. L'amygdale permet d'apprendre des conditionnements entre des perceptions externes (simples via le thalamus et complexes via l'hippocampe) et les signaux internes de douleur et de plaisir. Cette association permet de donner une valeur subjective aux différentes perceptions de l'extérieur. C'est grâce à l'amygdale qu'il est par exemple possible d'apprendre qu'un lieu spécifique de l'environnement est source de douleur. Le nucleus accumbens permet quant à lui d'apprendre les potentiels de récompense ou valeurs incitatives des différentes actions possibles dans un contexte perceptif donné. Autrement dit, le nucleus accumbens donne aux différents comportements une valeur endogène ancrée dans le système de valeur plaisir/douleur et contextualisée par le complexe hippocampique. Il s'agit aussi dans ce cas d'apprentissage de conditionnements, mais à la différence des conditionnements amygdaliens, les conditionnements de l'accumbens sont instrumentaux (c'est l'action et non la sensation qui est conditionnée). Le noyau accumbens permet donc de donner un potentiel de récompense important à une action qui rapproche d'un but tandis qu'une action qui rapproche d'un danger aura un potentiel de récompense faible voire nul. C'est ce qui permet de choisir l'action la plus profitable dans un contexte donné. Comme nous le verrons plus loin, l'accumbens peut même donner un potentiel de punition permettant ainsi de choisir l'action la moins dommageable dans un contexte donné. Les signaux de plaisir et de douleur venant depuis le tronc cérébral apparaissent comme les sensations fondamentales permettant l'émergence de critères de valeur endogènes.

Le cortex préfrontal enfin assure la régulation de ces différents systèmes sous-corticaux. Notre modélisation est très simple et très limitée mais elle permet de rendre compte d'un rôle majeur du cortex préfrontal : la supervision et la régulation du comportement. Le cortex préfrontal permet d'apprendre à évaluer l'efficacité du comportement en fonction du système de valeur endogène douleur/plaisir. Cette auto-évaluation est critique puisqu'elle permet de déterminer qu'un comportement en cours est un succès (s'éloigner des sources de punitions, s'approcher des sources de récompenses) ou un échec (s'approcher des sources de punitions, s'éloigner des sources de récompenses et surtout stagner). Dans notre modèle, le rôle du cortex préfrontal ne se limite pas à cette auto-évaluation. Il a aussi un rôle de régulateur du comportement. En utilisant des signaux d'inhibitions, le cortex préfrontal module le fonctionnement des différentes structures présentées précédemment.

Le tableau des émotions dépeint par ce chapitre neurobiologique est celui de l'interaction de processus bas niveau de traitement de l'information. Les signaux de plaisir et de douleur venant depuis le tronc cérébral apparaissent comme les sensations fondamentales permettant l'émergence de critères de valeur endogènes. C'est à l'aune de ces critères intrinsèques que sont évaluées les différentes stimulations et actions. La neurobiologie permet ainsi d'apporter quelques éclaircissements aux questions laissées en suspend au chapitre précédent. Tout d'abord, il n'y a pas de centre des émotions dans le cerveau mais plutôt un réseau complexe de structures anatomo-fonctionnelles interconnectées. Ces structures sont aussi bien phylogénétiquement anciennes (e.g. tronc cérébral) que récentes (e.g. cortex préfrontal). Ceci suggère que le fonctionnement des processus émotionnels s'est complexifié au cours du développement des espèces mais en se construisant sur les structures déjà présentes et bien fonctionnelles. De même, il n'y a pas de structure spécifique à une émotion particulière. Chaque structure remplit un ou plusieurs (le plus souvent) rôles fonctionnels spécifiques dans le traitement des informations sensorielles et motrices.

L'importance du système de valeur basé sur la douleur et le plaisir donne échos aux modèles dimensionnels puisqu'ils défendent l'idée d'une dimension de valence permettant de caractériser les émotions. Cependant, une émotion comme la surprise dépend d'une compétence plus spécifique comme la détection de nouveauté qui semble être dénuée de valence émotionnelle. La neurobiologie ne permet pas de formuler un modèle définissant les émotions de manière uniforme comme aurait pu le souhaiter l'approche dimensionnelle. En revanche, il apparaît clairement que les émotions ne sont pas chacune le résultat de processus spécifiques et exclusifs. L'idée même d'une nature essentielle commune des émotions, bien que séduisante, ne correspond pas aux résultats venant des neurosciences. Mais la mécanique décrite dans ce modèle de travail servira d'inspiration pour le développement des modèles et architectures de contrôle décrits plus loin.

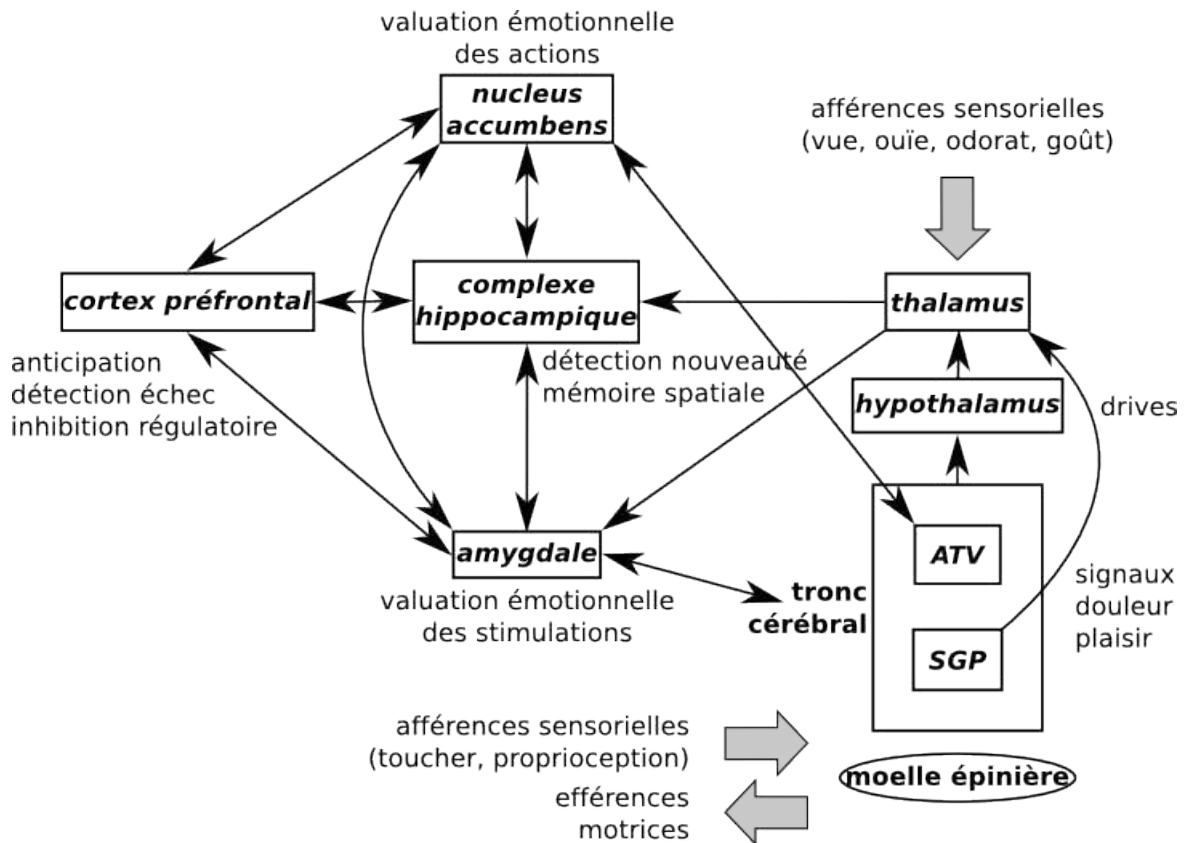


FIGURE 3.12 – Modèle de travail de bases neurobiologiques des processus émotionnels de bas niveau. Les sensations extérieures relayées par le thalamus sont traitées par l'hippocampe pour construire des contextes perceptifs multimodaux ainsi que pour la détection de nouveauté. Les sensation internes accédant au cerveau par le tronc cérébral permettent notamment la naissance des sensations de douleur (SGP) et de plaisir (ATV). L'hypothalamus permet d'assurer les motivations fondamentales de bas niveau, ou drives (faim, soif, ...). L'amygdale permet de donner une valeur endogène aux différentes stimulations grâce aux sensations de douleur et de plaisir. Le nucleus accumbens permet lui de donner un potentiel de récompense ou de punition aux différentes actions possibles dans un contexte donné. Le cortex préfrontal permet d'évaluer l'efficacité comportementale grâce à l'anticipation et la détection de l'échec (ces points seront développés au chapitre 6). L'inhibition frontale permet aussi de réguler le fonctionnement des structures sous-corticales en agissant à différents niveau du modèle pour modifier le comportement.

Modélisation neuronale et navigation

1 Introduction

Le propos de ce chapitre est de fournir les outils nécessaires à la compréhension des méthodes et des modèles utilisés tout au long de ces travaux pour instancier les modèles de structures cérébrales présentées précédemment. Dans un premier temps, nous aborderons le formalisme neuronal et les différents apprentissages dont nous doterons les architectures de contrôle du robot. La modélisation neuronale est un modèle de calcul inspiré du fonctionnement des neurones biologiques. Il permet de produire des modèles fonctionnels simples des structures cérébrales qui nous intéressent. Très schématiquement, les neurones formels sont des petites unités de calcul, des intégrateurs. Ces neurones fortement interconnectés reçoivent des informations par leurs connexions entrantes, opèrent un calcul et renvoient leur résultat par leurs connexions sortantes. Si individuellement, chaque neurone est particulièrement simple, les réseaux de neurones formels sont des modèles efficaces de calculs, d'opérations logiques et d'apprentissages. On espère des propriétés émergentes liées aux interactions dynamiques entre les neurones et l'environnement. Nous décrirons le fonctionnement des neurones et des différents types de connexions qu'ils peuvent former. L'apprentissage dans un réseau de neurone correspond à l'adaptation des poids synaptiques. Le réseau doit apprendre à produire une sortie spécifique en fonction de ses entrées. Nous détaillerons donc les différents algorithmes d'apprentissage utilisés. L'apprentissage hebbien et la réduction d'erreur de Widrow et Hoff permettront d'opérer des conditionnements. Nous détaillerons ensuite le processus de catégorisation.

Dans la seconde partie de ce chapitre, nous présenterons les deux stratégies de navigation utilisées par le robot. Ces deux stratégies d'inspiration biologiques reposent sur des informations sensorielles et des procédés totalement dissemblables. Nous utiliserons l'architecture PerAc développée au laboratoire ETIS [Gaussier 1995, Gaussier 1997, Gaussier 2000]. Cette architecture permet au robot de naviguer en apprenant des couples perception-action. Nous décrirons un modèle biologiquement plausible de cellule de lieux. Ces neurones apprennent à réagir sélectivement à la localisation dans l'environnement. L'apprentissage sensori-moteur consiste à associer les cellules de lieux à des actions permettant ainsi au robot de se diriger. L'intégration de chemin repose quand à elle sur les informations proprioceptives relatives aux déplacements effectués par le robot. Cette stratégie ne nécessite pas d'apprentissage. Elle réalise un calcul de mouvement effectué en fonction des informations odométriques (mouvements des roues du robot).

2 Modélisation neuronale

2.1 Neurones et synapses

Les réseaux de neurones sont un ensemble de modèles computationnels inspirés de la biologie et utilisés pour réaliser différents types de calculs et de traitements. Cette idée fondamentale a d'abord été proposée par McCulloch et Pitts dans [McCulloch 1943]. Une de leurs premières applications ayant grandement participé à leur popularisation est le *Perceptron* de Rosenblatt [Rosenblatt 1958] composé d'une rétine artificielle et d'un processus d'apprentissage d'une catégorisation. Pour comprendre ce qu'est un réseau de neurones, il convient de décrire le fonctionnement des neurones ainsi que de leurs synapses. Chaque neurone est une unité de calcul simple faisant la somme de ses entrées venant de ses connexions entrantes depuis d'autres neurones. Les connexions entre les neurones ont un poids w qui permet de pondérer l'importance de chacune des entrées dans le comportement de sortie. Ce poids est appelé poids synaptique ou efficacité synaptique. Différents types de connexions peuvent aussi être définis, mais nous y reviendrons plus tard. Le potentiel d'entrée P d'un neurone i est la somme de ses entrées x_n pondérées par leur poids synaptique w . Il se calcule comme suit :

$$P_i = \sum_1^n x_n \cdot w_{ni} \quad (4.1)$$

Ce potentiel d'entrée est ensuite transformé par une fonction de sortie Φ puis, selon les conditions, il déclenche un potentiel de sortie transmis par ses connexions sortantes vers d'autres neurones. Différents neurones sont donc définis par différents types de fonctions Φ . Il existe une grande variété de fonctions Φ mais nous n'en utiliserons que deux : la fonction Identité seuillée et la fonction Heaviside.

$$\text{fonction Identité seuillée} : I_S(x) = \begin{cases} x - S & \text{si } x - S > 0 \\ 0 & \text{sinon} \end{cases} \quad (4.2)$$

$$\text{fonction Heaviside} : H_S(x) = \begin{cases} 1 & \text{si } x > S \\ 0 & \text{sinon} \end{cases} \quad (4.3)$$

En raison d'une utilisation prépondérante d'un nombre restreint de types différents de neurones et ainsi pour simplifier les descriptions, nous définissons deux neurones types :

- le neurone seuillé : sa fonction d'activation est la fonction Identité seuillée et son seuil S est compris entre 0 et 1. Sa sortie est la somme pondérée de ses entrées moins le seuil S . Si S vaut 0, le neurone est linéaire.
- le neurone binaire : sa fonction d'activation est la fonction Heaviside et son seuil S est compris entre 0 et 1. Sa sortie est 1 ou 0 selon que la somme pondérée de ses entrées est supérieure à S ou non.

Il existe deux contraintes sur le potentiel de sortie d'un neurone. Il ne peut pas être négatif et il ne peut pas dépasser 1 (valeur de saturation). Ces deux règles de fonctionnement traduisent le fait que l'activité d'un neurone réel correspond en réalité à

une fréquence de décharge de potentiels d'action. Cette fréquence ne peut pas évidemment pas être négative et il existe une fréquence maximale de décharge du neurone. Le fonctionnement du neurone ainsi que les différents types de neurones définis sont illustré par la figure 4.1.

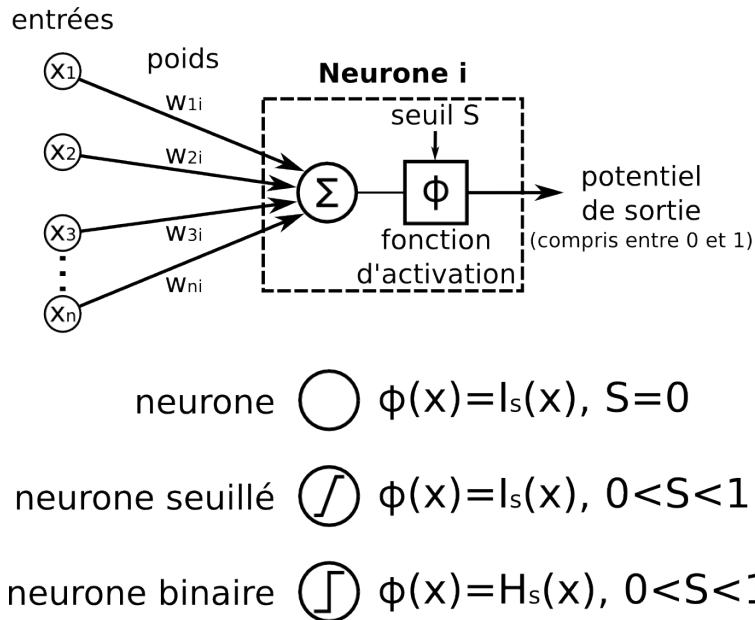


FIGURE 4.1 – Modèle simple de neurone formel. Le neurone calcule son potentiel d'entrée en faisant la somme de ses entrées pondérées par le poids de leur connexion. Si ce potentiel d'entrée est supérieur à un seuil donné, le neurone déclenche un potentiel de sortie. Deux types de neurones sont utilisés dans nos travaux. Le neurone seuillé soustrait la valeur du seuil à son potentiel d'entrée (la fonction d'activation est $I_S(x)$ avec $S \in [0;1]$). Le neurone linéaire fait simplement la somme pondérée de ses entrées. Il s'agit d'un cas particulier du neurone seuillé (la fonction d'activation est $I_S(x)$ avec $S = 0$). Le neurone binaire renvoie 0 en dessous du seuil et 1 au delà (la fonction d'activation est $H_S(x)$ avec $S \in [0;1]$).

Différents types de synapses peuvent être utilisées. Tout d'abord, elles peuvent être plastiques ou non. Une connexion plastique est une connexion dont le poids peut varier. Au contraire, une connexion non plastique a un poids fixe. L'apprentissage dans un réseau de neurones correspond à la modification des poids synaptiques. Dans la suite de ce travail, nous utiliserons deux types d'apprentissage : les conditionnements et la catégorisation. Une description détaillée en sera donnée dans la suite de ce chapitre. Les connexions non plastiques quand à elles peuvent être soit excitatrices (poids positif), soit inhibitrices (poids négatif). Les connexions excitatrices augmentent donc le potentiel d'entrée d'un neurone tandis qu'une connexion inhibitrice le diminue. Enfin, dans le cas non pas de deux neurones mais de deux groupes de neurones, les connexions peuvent soit être de "un vers un" soit de "un vers tous". Les connexions de un vers un relient chaque neurone de premier groupe à un unique neurone du second groupe. Les connexions de un vers tous relient donc chaque neurone du premier groupe à tous les neurones du second groupe. Dans ce second cas de figure, chaque neurone du

second groupe est donc également connecté à tous les neurones du premier groupe. Les connexions inconditionnelles nécessaires pour l'apprentissage de conditionnements seront décrites dans la prochaine section de ce chapitre. Enfin, un dernier type de connexion est utilisé mais il ne faut pas le confondre avec les connexions précédentes. Il s'agit du lien de neuromodulation permettant de contrôler les modifications du poids des synapses plastiques. Il s'agit d'un lien vers des synapses et non vers un neurone. Son rôle sera expliqué plus en détails lors de la description des capacités d'apprentissage des réseaux de neurones.

2.2 Apprentissage associatif hebbien

Le neuropsychologue Donald Olding Hebb s'est particulièrement illustré par ses recherches sur le fonctionnement des neurones. Le but de ses travaux était de comprendre le rôle des neurones dans le fonctionnement des processus psychologiques. Hebb s'est particulièrement intéressé au lien entre les capacités d'adaptation d'un réseau de neurones et le comportement d'apprentissage d'un être humain. Dans son ouvrage majeur, *The organization of behavior* [Hebb 1949], Hebb fait le lien entre son expérience en tant que neurochirurgien et ses travaux en psychologie comportementale. Il y décrit le cerveau comme un réseau de neurones hautement interconnectés. Toute information est représentée par l'activité des neurones et par la qualité ou la force de leurs connexions. L'apprentissage hebbien ou loi de Hebb est le mécanisme de base permettant de renforcer la connexion entre deux neurones. L'idée est très simple : quand un neurone A excite souvent un neurone B, la connexion entre A et B devient de plus en plus efficace. Autrement dit, le poids de la connexion A-B est renforcé quand A excite B. Il en résulte que l'activation de B par A est facilitée par leur co-activation. Hebb l'explique en ces mots : "*When an axon of cell A is near enough to excite B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.*" (*The Organization of Behavior*, p. 62). Depuis les travaux d'Eric Kandel dans les années 60 sur la mise en évidence du phénomène de potentialisation à long terme [Kandel 1961, Kandel 1968], un grand nombre de travaux en neurobiologie ont permis d'explorer la variété des phénomènes de plasticités synaptiques. Si ces résultats apportent des mises en évidences biologiques du modèle de Hebb, ils montrent aussi que le cerveau est bien plus complexe qu'un simple réseau de neurones capables d'apprentissages hebbiens. Cependant, cette simplicité permet une approximation facile à modéliser et à mettre en oeuvre. De plus, suivant une approche constructiviste, il semble préférable de partir de modèles simples et d'explorer leurs capacités une fois mis en réseau plutôt que d'élaborer les modèles de neurones les plus sophistiqués et réalistes possibles.

L'apprentissage hebbien peut être modélisé très simplement. La figure 4.2 illustre son principe. Il peut être décrit plus formellement de la manière suivante. Si deux neurones i et j sont reliés par une connexion (synapse) dont le poids est noté w_{ij} et que x_i et x_j représente le potentiel d'activation des neurones i et j , la loi de Hebb décrivant l'évolution du poids w_{ij} dans le temps peut être modélisée par l'équation 4.4 :

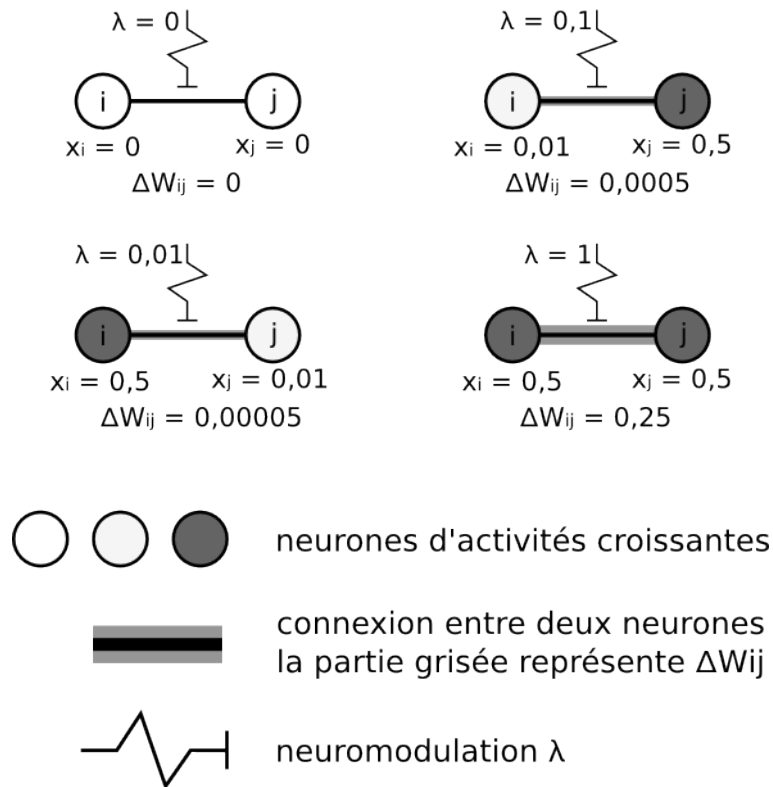


FIGURE 4.2 – Apprentissage hebbien : le poids w_{ij} de la connexion entre les neurones i et j est renforcé par la co-activation de i et j conjointement avec la neuromodulation λ .

$$\Delta w_{ij} = \lambda x_i x_j - \varepsilon w_{ij} \quad (4.4)$$

λ est le taux d'apprentissage. Nous utiliserons de la même manière les expressions "taux d'apprentissage" et "neuromodulation". La neuromodulation peut être comprise comme la facilitation plus ou moins forte des modifications de poids synaptiques. Pour le neurone réel, le neuromodulateur est une neuro-hormone qui modifie le fonctionnement de la cellule. Ce paramètre d'apprentissage est reçu par le lien de neuromodulation. Le produit $x_i x_j$ représente la co-activation des neurones i et j . Une fois que w_{ij} est suffisamment important, l'activité du neurone i conditionne l'activité du neurone j . Cependant, si les poids w_{ij} ne sont modifiés qu'en fonction de la co-activation des neurones i et j , la règle d'apprentissage n'est pas stable : les poids w_{ij} peuvent augmenter indéfiniment. C'est pourquoi le terme $-\varepsilon w_{ij}$ est ajouté à cette équation. Il correspond à un oubli passif du lien entre les neurones x et y . A la place de cet oubli passif, nous aurions pu utiliser une normalisation des poids synaptiques, certainement plus fiable, mais pas nécessaire dans notre cas.

2.3 Apprentissage de conditionnements : Widrow et Hoff et méthode des moindres carrés

Bernard Widrow, un chercheur en électronique et son étudiant Ted Hoff sont les inventeurs de l'algorithme qui permet à un réseau de neurones de mettre en oeuvre la méthode de réduction de erreurs quadratiques (*Least Mean Square* ou LMS) [Widrow 1960]. Ce type d'algorithme d'apprentissage d'un réseau de neurone constitue une classe de filtres adaptatifs utilisés pour apprendre une forme désirée en trouvant les coefficients (les poids synaptiques) permettant de produire la plus faible erreur quadratique moyenne entre le signal désiré et le signal effectif. Il s'agit d'une méthode stochastique de descente de gradient dans la mesure où le réseau ne s'adapte à un moment donné qu'en fonction de l'erreur produite à ce moment. La figure 4.3 est une illustration simplifiée du principe de fonctionnement de cet algorithme. Au lieu d'être appliqué à une forme représenté par le codage d'une population de neurones, l'apprentissage est effectué pour trois neurones seulement. Dans cet exemple, un neurone code pour le signal à apprendre, la sortie désirée Sd_j , un neurone code pour le signal d'entrée E_i et un neurone représente la sortie S_j : le résultat de l'apprentissage. Un lien inconditionnel relie la sortie désirée Sd_j au neurone de sortie S_j . Ce lien n'influe pas sur le potentiel du neurone de sortie mais il permet de comparer leur potentiels respectifs afin de déterminer le signe et l'importance de l'apprentissage. Cet apprentissage est effectué par la pondération du lien conditionnel reliant le neurone d'entrée E_i et le neurone de sortie S_j . L'adaptation du poids de ce lien dépend de l'écart ou l'erreur entre la sortie S_j et la sortie désirée Sd_j . De même que pour l'apprentissage hebbien, le pas d'apprentissage λ permet de contrôler la vitesse d'apprentissage. Le lien plastique w_{ij} relie l'entrée E_i et la sortie S_j . L'équation 4.5 décrit formellement le fonctionnement de l'apprentissage par la règle de Widrow et Hoff :

$$\Delta w_{ij} = \lambda E_i (Sd_j - S_j) \quad (4.5)$$

Pavlov a introduit la notion de conditionnement pour ce que l'on appelle aujourd'hui **conditionnement classique, répondant, de type I** ou simplement **conditionnement pavlovien** [Pavlov 1927]. Il s'agit de l'apprentissage du lien entre une perception neutre (le signal conditionnel) et une perception de plaisir ou de douleur (le signal inconditionnel). L'image souvent utilisée pour illustrer ce type de conditionnement est la fameuse expérience de Pavlov et de son chien. Avant de lui servir à manger, Pavlov fait retentir une cloche. En répétant cette procédure à chaque fois qu'il le nourrit, le chien ne commence plus à saliver quand la nourriture est servie mais quand la cloche retentit. Un autre type de conditionnement est toutefois possible. Un modèle computationnel de ce conditionnement classique a aussi été proposé par Rescorla et Wagner [Rescorla 1972]. Au lieu de conditionner une perception, Thorndike puis Skinner se sont intéressés à la possibilité de conditionner une action : apprendre à renforcer ou inhiber un comportement pour obtenir une récompense ou éviter une punition. C'est le **conditionnement instrumental, opérant, de type II** ou encore **conditionnement skinnérien** [Skinner 1938]. Ces deux types de conditionnements ont été largement étudiés, et modélisés par l'appren-

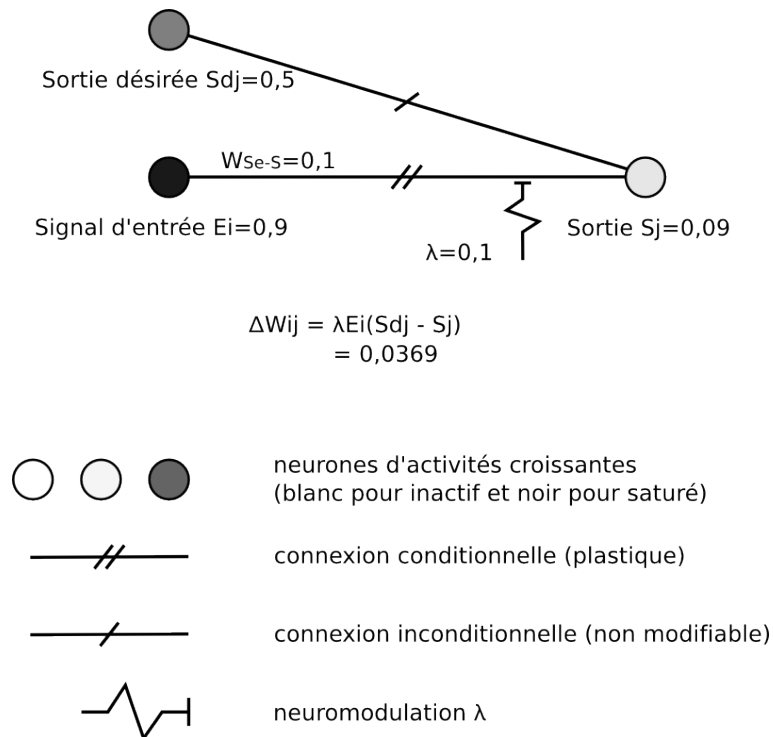


FIGURE 4.3 – Apprentissage de Widrow et Hoff : la connexion conditionnelle (plastique) entre l'entrée E_i et la sortie S_j s'adapte en fonction de l'erreur entre la sortie S_j et la sortie désirée S_{dj} venant depuis la connexion inconditionnelle. Dans cet exemple, le poids de la connexion plastique w_{ij} est renforcé puisque la sortie est inférieure à la sortie désirée. Dans le cas contraire, la connexion aura été affaiblie. La neuromodulation λ permet de contrôler la vitesse d'apprentissage.

tissage de réseaux de neurones [Grossberg 1971, Grossberg 1974, Grossberg 1987b, Grossberg 1987a, Rescorla 1967a, Rescorla 1967b, Rescorla 1972, Rescorla 1988]. Ces conditionnements ont aussi été étudiés du point de vue des structures neuronales impliquées [Schmajuk 1992, Schmajuk 1986, Hasselmo 1999].

2.4 Processus de catégorisation et compétition

Le processus de catégorisation correspond à la capacité à construire et reconnaître différentes catégories dans une forme évolutive. En raison de son aspect fondamental, cette capacité est nécessaire pour pratiquement toutes les fonctions cognitives. Une sensation est codée comme un pattern d'activité spécifique d'une population de neurones. Par exemple, les photorécepteurs de la rétine ou les cellules cillées de la cochlée produisent une réponse spécifique à une image ou un son donné. Le processus de catégorisation repose sur la capacité à caractériser et reconnaître ces réponses spécifiques codées en population. Les problèmes de catégorisation ont notamment été abordés par Rumelhart [Rumelhart 1985b] et par Carpenter et Grossberg avec leur architecture ART (*Adaptive Resonance Theory*) [Carpenter 1987, Carpenter 1988, Carpenter 1990].

Dans le groupe de neurones de catégorisation, chaque neurone peut représenter une catégorie. Lorsque le groupe reçoit un signal de neuromodulation, un neurone peut être recruté pour représenter la catégorie correspondant à la forme perçue. Le calcul de l'activité des neurones de catégorisation est décrit par l'équation 4.6 :

$$CAT_j = 1 - \frac{1}{N} \sum_{i=1}^N |w_{ij} - E_i| \quad (4.6)$$

CAT_j est l'activité du neurone j du groupe de catégorisation. E est l'entrée (de taille N). w_{ij} est le poids du lien entre l'entrée E_i et le neurone de catégorisation CAT_j . Tous les poids synaptiques sont initialisés à 0. L'apprentissage à proprement parlé est réalisé par l'adaptation des poids w_{ij} entre les neurones i de l'entrée et les neurones j du groupe de catégorisation. L'équation 4.7 décrit cette modification des poids w_{ij} :

$$\Delta w_{ij} = \delta_j^k (a_j(t) E_i) \quad (4.7)$$

avec $k = \text{ArgMax}(CAT_j)$, δ_j^k est le symbole de Kronecker¹ utilisé pour ne modifier les poids des liens que vers la catégorie gagnante. $a_j(t)=1$ seulement quand le neurone j est recruté et 0 sinon. Ce terme permet de modifier les poids synaptiques vers la catégorie gagnante uniquement quand un signal de recrutement est présent. Pour le recrutement d'un neurone de catégorie, un neurone parmi les j neurones est sélectionné. Ce neurone sélectionné doit simplement ne pas avoir déjà été utilisé pour coder une autre catégorie. Cela peut être n'importe quel neurone du groupe. Au moment du recrutement à proprement parler, l'activité du neurone sélectionné est fixé à 1. C'est ce qui permet de modifier les poids synaptiques uniquement pour les connexions entre les neurones codant le pattern d'entrée et le neurone de catégorie nouvellement recruté. Les poids de ses connexions entrantes sont modifiés pour correspondre à l'entrée (terme E_i). Ainsi, plus la forme perçue est proche de la forme apprise et plus le neurone de catégorie correspondant à cette forme est actif.

Un mécanisme de compétition sur les sorties du groupe de catégorisation permet alors de déterminer quelle catégorie est la mieux reconnue. Les activités des neurones du groupe de catégorisation traduisent la plus ou moins grande proximité entre la forme perçue et les différentes formes apprises. Le neurone le plus actif correspond à la forme la mieux reconnue. Il faut cependant bien se rendre compte que cette reconnaissance est définie en termes relatifs. Si par exemple une forme d'entrée est très différent de toutes les formes apprises, la compétition produit toujours un gagnant. C'est pourquoi il convient de définir un seuil de vigilance sur l'activité maximale de sortie du groupe de catégorisation. Sous le seuil de vigilance, i.e. quand les formes apprises sont trop éloignées de la forme perçue, un neurone est recruté pour représenter cette catégorie nouvelle. Concrètement, la sortie du groupe de compétition vient inhiber un neurone

1. fonction de Kronecker :

$$\delta_j^k(x) = \begin{cases} x & \text{si } j = k \\ 0 & \text{sinon} \end{cases}$$

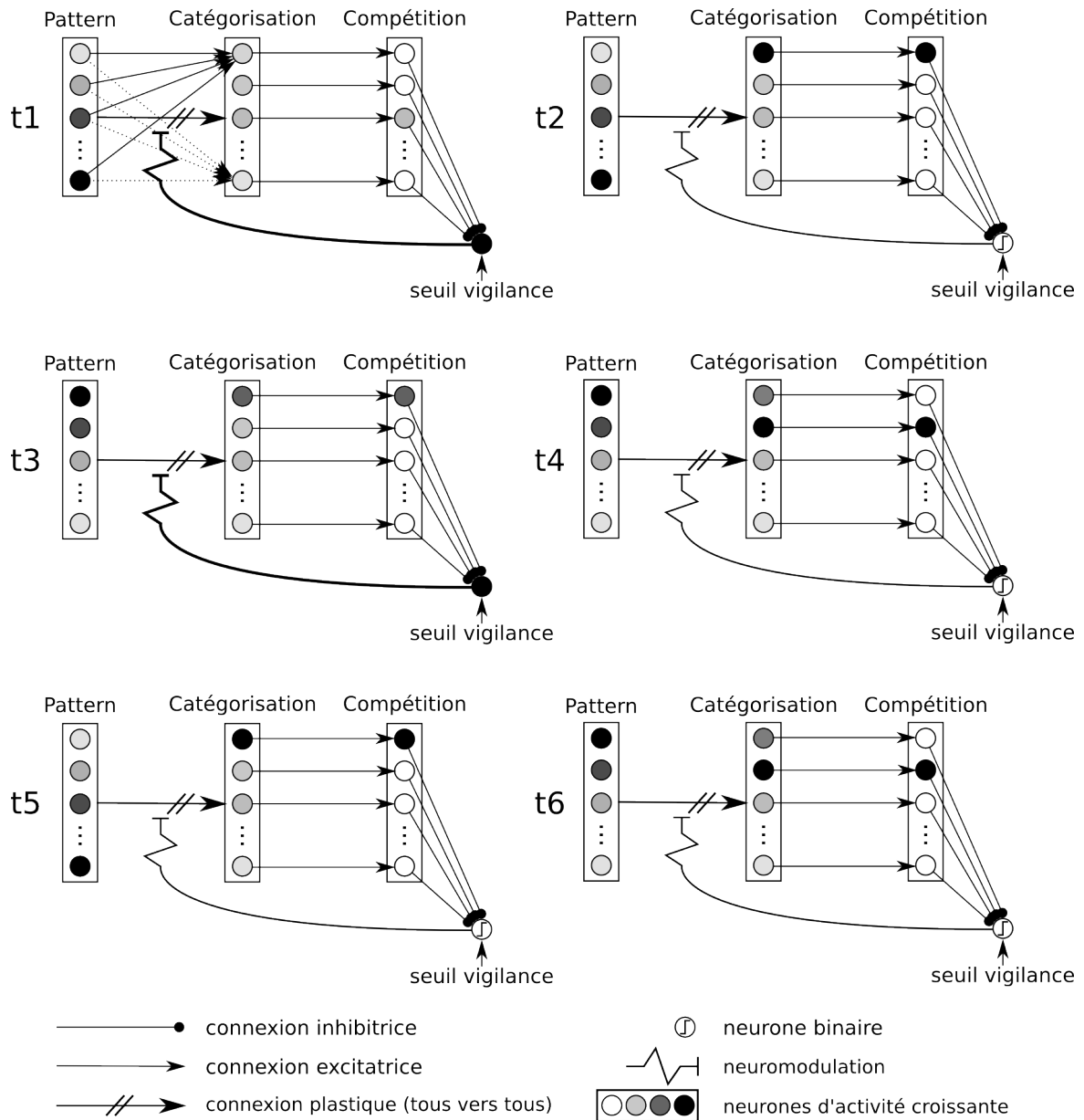


FIGURE 4.4 – Le processus de catégorisation formalisé comme l'apprentissage d'un réseau de neurones suivi d'un processus de compétition. Un seuil de vigilance sur le niveau de reconnaissance du gagnant de la compétition permet de contrôler l'apprentissage des catégories. Les connexions modifiables connectent chaque neurone du groupe de catégorisation à chaque neurone du groupe représentant la forme d'entrée. En t1, aucune catégorie n'a été apprise donc aucun neurone de catégorisation ne correspond à la forme perçue. Le potentiel du neurone gagnant de la compétition est en dessous du seuil de vigilance. Un signal d'apprentissage est déclenché. En t2, après avoir reçu ce signal, un neurone est recruté et les poids de ses connexions avec les neurones représentant la forme d'entrée sont adaptés afin que ce neurone ait une activité maximale lorsque cette forme spécifique est perçue (forme 1). En t3, la forme perçue a changé (forme 2). Aucun neurone de catégorie ne code pour cette nouvelle forme : le gagnant de la compétition est en dessous du seuil de vigilance et un signal de recrutement est généré. En t4, un second neurone est recruté correspondant à une nouvelle catégorie apprise. En t5 et t6, le réseau perçoit à nouveau les formes 1 puis 2. Ces formes ayant déjà été apprises, la compétition permet de déterminer quelle forme est reconnue (et le recrutement est inhibé).

binaire recevant aussi une activation représentant le seuil de vigilance. Si le gagnant a un potentiel supérieur au seuil de vigilance, le neurone binaire est inhibé. Dans le cas contraire, un signal de recrutement est envoyé vers le groupe de catégorisation. D'avantage de détails sur le contrôle de ce seuil seront présentés dans les chapitres suivants. La figure 4.4 est une illustration du processus de catégorisation. Tout au long de ces travaux, nous utiliserons une convention d'écriture. Afin de simplifier les schémas représentant les réseaux de neurones, nous ne représenterons pas toujours tous les liens existants. Quand deux groupes de neurones sont reliés par un lien simple (excitateur ou inhibiteur) ou inconditionnel, chaque neurone du premier groupe projette un lien vers un neurone exclusif du second groupe. Si le lien est une connexion plastique (signalé par une double barre), alors chaque neurone du premier groupe projette des connexions vers tous les neurones du second groupe.

3 Navigation visuelle sensori-motrice

L'architecture PerAc (Perception-Action) est un modèle d'apprentissages sensori-moteurs développé par l'équipe neurocybernétique du laboratoire ETIS depuis le milieu des années 90 et utilisé pour permettre une navigation visuelle [Gaussier 1995, Gaussier 1997, Gaussier 2000, Gaussier 2001, Giovannangeli 2006, Giovannangeli 2007, Giovannangeli 2008b]. Il permet l'apprentissage des conditionnements entre des perceptions (reconnaissance visuelle des lieux de l'environnement) et des actions à effectuer (directions vers lesquelles aller). Cette méthode de navigation de bas niveau ne nécessite pas de représentation symbolique du but à atteindre. Le robot repose sur l'arrangement des associations sensori-motrices apprises lui permettant d'effectuer des actions le faisant converger vers son but. Dans un premier temps, nous décrivons le modèle que nous utilisons pour simuler les cellules de lieux servant de base aux apprentissages sensori-moteurs. Ces cellules de lieux, mise en évidence chez le rat dans [O'Keefe 1971] et aussi présentes chez les primates sont des neurones dont l'activité est corrélée à la localisation dans l'environnement. C'est pourquoi elles portent le nom de cellules de lieux. Ensuite, nous décrivons le fonctionnement de l'architecture PerAc basée sur la *counterpropagation* décrite par Hecht-Nielsen [Hecht-Nielsen 1987, Hecht-Nielsen 1989] et permettant d'apprendre des comportements de navigation.

3.1 Cellules de lieu visuelles

La stratégie de navigation sensori-motrice est basée sur la capacité à apprendre des cellules de lieux visuelles. Cette section contient une description du modèle utilisé tout au long de cette thèse. Il s'agit d'un modèle biologiquement plausible de cellules de lieux pré-hippocampiques. Ce modèle a été testé sur différents types de robots mobiles (Koala, Labo3, Pioneer, Robulab et Roburoc), évoluant en environnement ouvert, intérieur et extérieur. La figure 4.5 résume la chaîne de traitements, depuis le flux de la caméra (image brute) jusqu'à la caractérisation des cellules de lieu.

La caméra est montée sur un système pan (contrôlé par un servomoteur) permettant

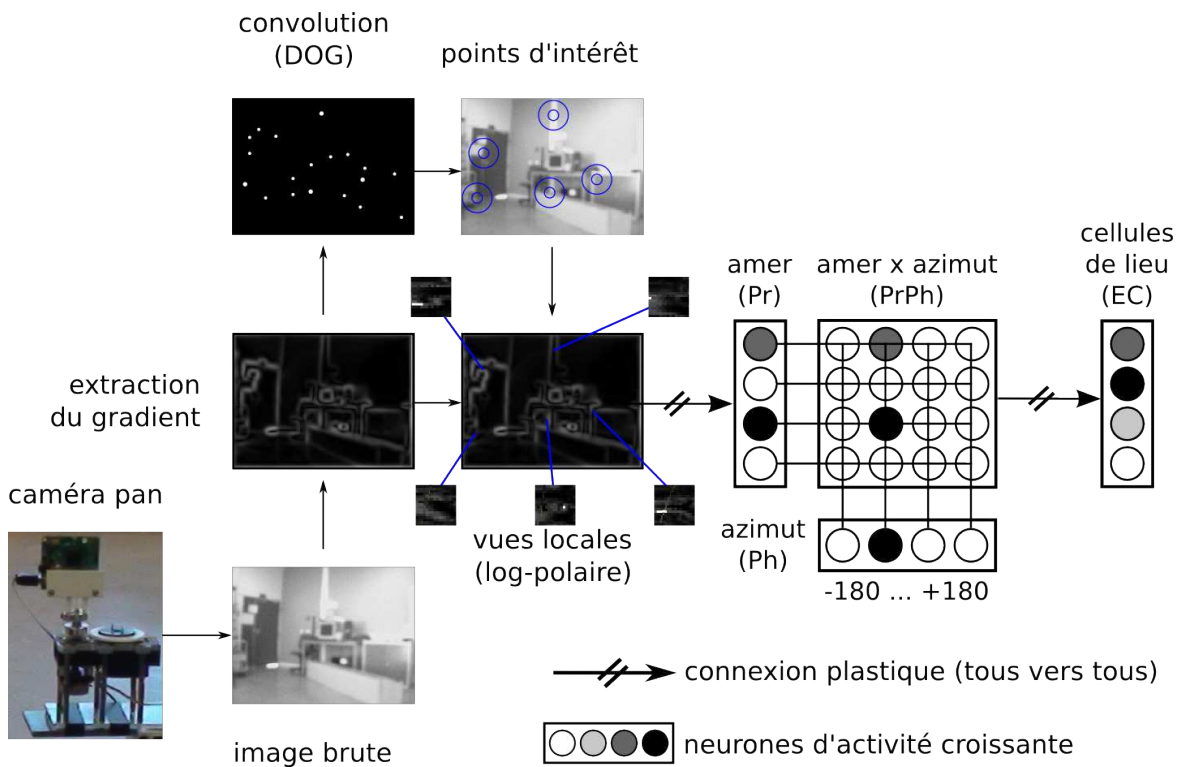


FIGURE 4.5 – Architecture d'apprentissage de cellules de lieux. Le système visuel extrait des vues locales en coordonnées log-polaires centrées sur les points à forte courbure de l'image de gradient. Une étape de fusion (PrPh) de l'identité des vues locales (Pr) et de leur azimuth (Ph) permet de combiner les informations "où" et "quoi". Une catégorisation de la matrice PrPh permet de caractériser les cellules de lieux (EC).

au robot de voir dans toutes les directions. En pratique, le robot perçoit son environnement visuel en balayant son panorama horizontalement. La caméra a un angle de vue de 40° , mais afin d'éviter de perdre des informations à cause du traitement des bordures d'images elle capture une image tout les 24° (15 positions de caméra pour un panorama complet). Afin de minimiser l'impact des changements d'éclairage, notre système n'utilise pas l'image brute mais son gradient. L'image du gradient est ensuite convoluée avec un filtre de type DOG (différence de gaussiennes), dans le but de détecter des points de forte courbure (points d'intérêt ou points de focalisation) en basse résolution (voir figure 4.5). Un mécanisme de compétition entre les points de focalisation permet de sélectionner les points les plus intenses (en termes de contraste et de rayon de courbure). Enfin, une transformation log-polaire [Schwartz 1980] des vues locales extraites autour de chaque point focal est calculée sur l'image du gradient. Cela permet une représentation des amers robuste à de petites rotations et de petits changements d'échelle. La figure 4.5 montre les points de focalisation et la zone utile des vues locales sur l'image brute et sur l'image de gradient.

Ce modèle du système visuel repose sur deux flux d'information distincts. Le "quoi", la reconnaissance d'une vue locale ou amer visuel en coordonnées log-polaires de 32

x 32 points est codé sur un vecteur de neurones de taille L . Il s'agit de la voie Pr (pour cortex périrhinal). Le "où", l'azimut absolu du point de focalisation, est codé sur un vecteur de neurones de taille Θ . Il s'agit de la voie Ph (pour cortex parahippocampique). Dans notre cas, cette orientation absolue est obtenu grâce à une boussole électronique mais tout autre moyen est équivalent (boussole visuelle notamment [Leprêtre 2000, Labrosse 2004, Montiel 2006, Sturm 2009]). Ces deux flux sont ensuite fusionnés dans un espace produit, au moyen d'une matrice de neurones. Celle-ci comprend $L \times \Theta$ unités sigma-pi [Rumelhart 1985a]. Tous les neurones d'une même ligne sont liés à un des L neurones de reconnaissance des amers. Tous les neurones d'une même colonne sont liés à un des Θ neurones de Ph. C'est pourquoi cette matrice des informations d'amer et d'azimut est appelée PrPh. Une mémoire à court terme est utilisée au niveau de PrPh mémorisant les entrées passées. Avant chaque exploration séquentielle du panorama, toutes les activités des neurones de PrPh sont remis à 0. La mémoire à court terme permet de mémoriser au cours d'une exploration de panorama l'activité de l'ensemble des couples quoi et où. L'activité globale du PrPh est ensuite catégorisée comme un pattern codant une représentation invariante du lieu. L'activité de chaque neurone de PrPh est le produit $S_{ij}^{PrPh} = S^{L_i} \times S^{\Theta_j}$, avec S_{ij}^{PrPh} l'activité du neurone ij de PrPh. Ce neurone est lié à la sortie S^{L_i} du neurone codant pour la vue locale i , et à la sortie S^{Θ_j} du neurone codant pour son azimut j . Une version optimisée de PrPh est utilisée pour réduire le nombre de neurones nécessaires. j dans PrPh est fixé à une valeur faible (e.g. 5 neurones) comprimant ainsi l'information angulaire. Le pattern d'activité de la matrice PrPh permet une caractérisation visuelle du lieu dans lequel se trouve le robot. Ce pattern peut alors être appris par un groupe de neurones représentant EC (cortex entorhinal). Un signal de neuromodulation reçu par EC permet le recrutement d'un neurone pour apprendre à caractériser le lieu visuel courant. Dans notre modèle, ce neurone est une cellule de lieu. L'activité d'une cellule de lieu découle directement de la distance entre le pattern d'activité du PrPh au moment de l'apprentissage et le pattern courant. La variation de cette distance traduit une plus ou moins bonne reconnaissance d'un lieu appris. L'activité P_k d'une cellule de lieu k s'exprime de la manière suivante :

$$P_k = \frac{1}{l_k} \sum_{i=1}^L \omega_{ik} \cdot f_s(L_i) \cdot (1 - g_d(\theta_{ik}^L - \theta_i)) \quad (4.8)$$

$l_k = \sum_{i=1}^l \omega_{ik}$. $\omega_{ik} = 0, 1$ exprime le fait que la vue locale i a servi à coder le lieu k . l_k est le nombre d'amers utilisé pour coder la cellule k . L est le nombre de vues locales apprises. L_i est le niveau de reconnaissance de la vue locale i . $f_s(x)$ est la fonction d'activation du groupe de neurones Pr de reconnaissance des amers. θ_{ik}^L est l'azimut appris par le groupe de neurones codant pour l'azimut (cortex pariétal postérieur (périrhinal) pour les rats et les primates [Kolb 1983]) de la i^{eme} vue locale pour le k^{eme} lieu. θ_i est l'azimut courant de la vue locale i . d est un paramètre de diffusion de l'information angulaire définissant la pente de la fonction $g_d(x)$. Le rôle de $f_s(x)$ et $g_d(x)$ est d'adapter la dynamique des neurones codant le quoi et le où. Leur définition est :

$$g_d(x) = \begin{cases} \frac{|x|}{d \cdot \pi} \text{ si } |x| > d \cdot \pi \\ 1 \text{ sinon} \end{cases} \quad (4.9)$$

$$f_s(x) = \frac{1}{1-s} [x-s]^+ \quad (4.10)$$

où $[x]^+ = x$ si $x > 0$, et 0 sinon. La figure 4.6 montre l'activité de 5 x 5 cellules apprises de manière supervisée dans une expérience robotique afin de montrer l'étendue et l'homogénéité du champs d'activité de ces neurones.

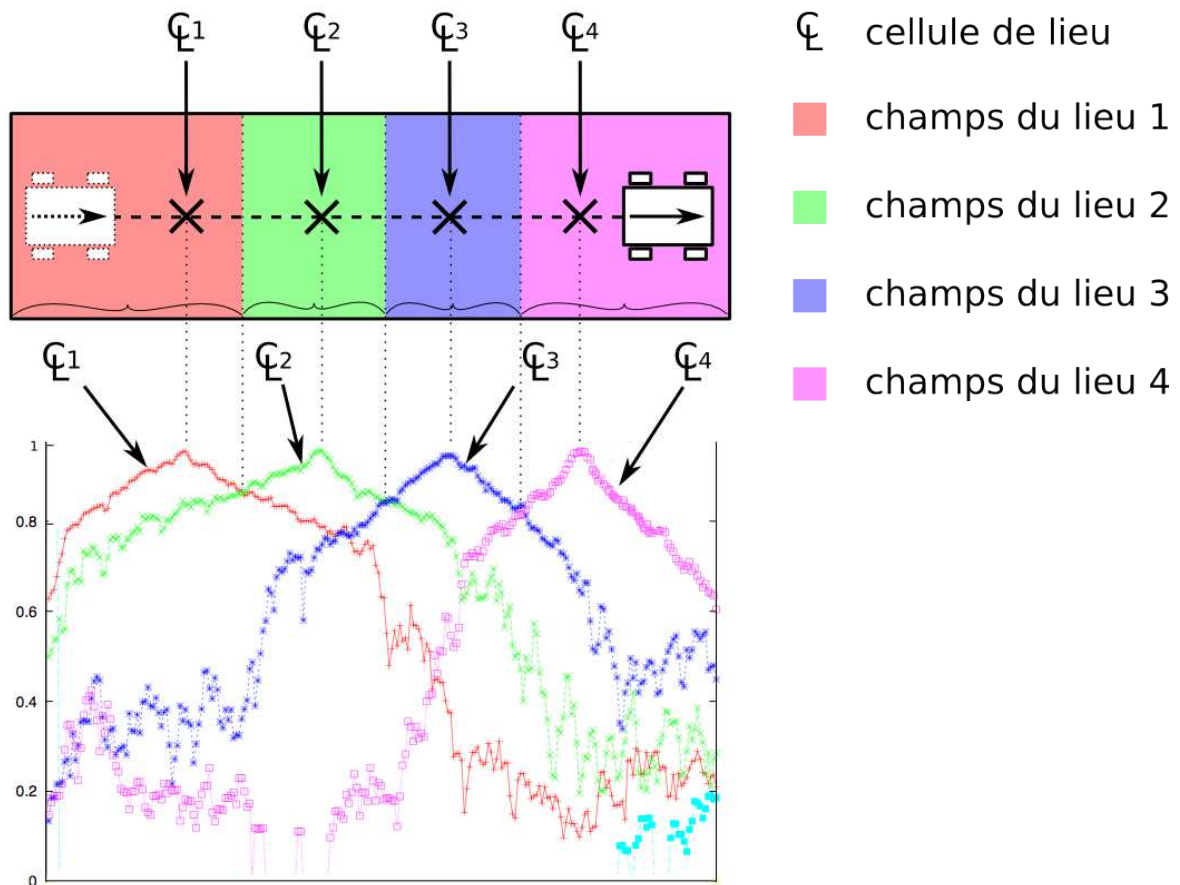


FIGURE 4.6 – Activités de 4 cellules de lieu préalablement apprises en ligne. Lorsque le robot refait le parcours, il reconnaît successivement les lieux 1 à 4. L'activité de chaque cellule de lieu dépend de la distance du robot au lieu appris. La zone dans laquelle une cellule de lieu a l'activité la plus forte est son champs de lieu.

3.2 Conditionnements Lieu-Action

L'architecture PerAc permet d'apprendre des conditionnements sensori-moteurs en utilisant la reconnaissance de cellules de lieu comme information sensorielle et

une direction à prendre comme action. Ces conditionnements peuvent être compris comme l'apprentissage du potentiel de récompense des différentes actions possibles dans un contexte perceptif donné. En ce sens, les conditionnements lieu-action sont des représentations simplifiées du rôle du nucleus accumbens (cet aspect sera traité plus en détails au chapitre 8). La figure 4.7 présente cette architecture. Dans un premier temps, nous décrirons son fonctionnement. Nous montrerons ensuite comment elle permet d'apprendre des comportements de navigation.

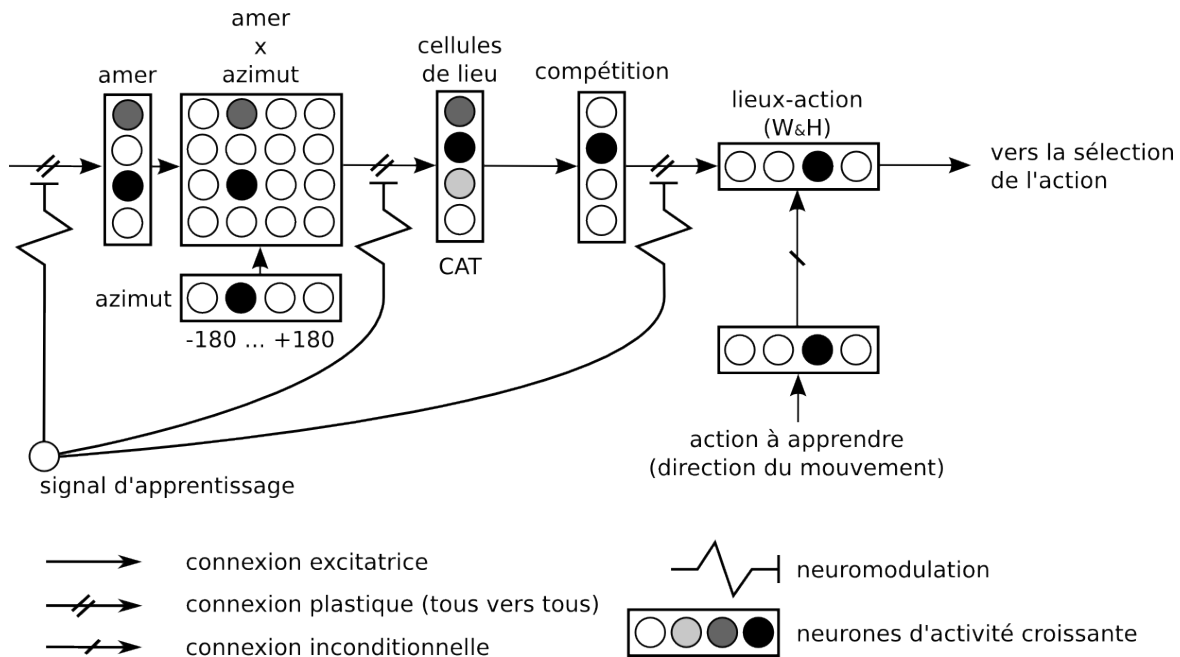


FIGURE 4.7 – L'architecture PerAc permet l'apprentissage de conditionnements sensori-moteurs entre la reconnaissance d'une perception (cellule de lieu) et d'une action (direction à prendre). Ces conditionnements permettent au robot de naviguer dans son environnement. Ils peuvent être compris comme le mécanisme de fonctionnement du nucleus accumbens qui permet d'apprendre le potentiel de récompense des actions possibles (propension à l'action).

Toute la partie gauche de la figure 4.7 correspond à l'apprentissage de cellules de lieu. Le signal de neuromodulation permet d'une part d'apprendre les amers visuels caractérisant un lieu et d'autre part d'apprendre à catégoriser les différents patterns d'activité du PrPh correspondant aux différents lieux de l'environnement. Une compétition permet ensuite de déterminer le lieu courant correspondant à la cellule de lieu la plus active. Quand un lieu vient d'être appris, son activité est proche de 1 et il gagne donc la compétition. L'apprentissage sensori-moteur à proprement parler est réalisé par un groupe de neurones apprenant les conditionnements entre les cellules de lieu et les actions à effectuer. Dans notre cas, ces actions correspondent à des directions que le robot doit prendre. Il s'agit donc d'actions liées au comportement de navigation. Ces actions sont donc représentées sur un champs de neurones sans connexion latérale qui peut être compris comme un vecteur de neurones. Chaque neurone code pour une orientation. La précision en orientation dépend de la taille de ce champs. Pour

un champs de taille n , la précision de la discrétisation angulaire est donc simplement de $\frac{360^\circ}{n}$. Le groupe de neurones réalisant l'apprentissage du conditionnement sensori-moteur, le champs lieu-action, a la même taille que le champs de neurones des actions à apprendre. Ses neurones utilisent la règle d'apprentissage de Widrow et Hoff. Chaque neurone reçoit une connexion inconditionnelles depuis le neurone correspondant du champs de l'action à apprendre (e.g. la direction courante du robot). Chaque neurone reçoit aussi des connexions conditionnelles (plastiques) depuis chaque cellule de lieu. L'apprentissage consiste à modifier le poids des liaisons entre les cellules de lieu et les neurones du champs lieu-action afin de permettre à la cellule de lieu gagnante d'activer dans le champs lieu-action l'action qui y est associée. La reconnaissance d'un lieu active donc une direction à prendre. Cette commande motrice peut alors être exécutée pour le robot.

Pour comprendre comment cette architecture permet à un robot de naviguer, nous prendrons l'exemple d'une tâche de retour au nid [Gallistel 1996] (*homing*). Pour rejoindre son but, le robot apprend des couples lieu-action convergeant vers le but. Pour que ce mécanisme soit opérant, il faut que le robot puisse mettre en compétition les différents lieux qu'il perçoit. C'est ce qui lui permet de définir sa perception d'un lieu : définir dans quel lieu il se trouve. La conjonction de ces conditionnements sensori-moteur mis en compétition crée un bassin d'attraction qui définit dynamiquement l'attraction vers le but sans nécessiter d'apprendre ou même de définir explicitement celui-ci. Une fois appris un minimum de trois lieux-action convergeant vers le but, le robot n'a plus qu'à réagir en prenant la direction correspondant au lieu dans lequel il se trouve. Itération après itération, la distance du robot au but se trouve ainsi réduite : le robot est attiré par le but. Cette compétition permet une généralisation des apprentissages i.e. la convergence vers le but est possible depuis toute une zone autour de ce dernier. Cette généralisation vient du fait que la décision réelle n'est prise qu'au niveau moteur et doit être comprise en fonction de la dynamique globale du système. Cependant, elle a ses limites. Au delà d'une certaine distance au lieu d'apprentissage, le signal d'une cellule de lieu n'est plus que du bruit. Le robot ne peut plus alors déterminer quel lieu est le plus actif, donc le plus proche. A une distance suffisamment grande d'un bassin d'attraction, les actions prises par le robot ne permettent alors plus de converger vers le but. Le bruit sur la perception visuelle est trop important et le robot navigue aléatoirement. La figure 4.8 montre les trajectoires du robot depuis différents points de départ et après avoir appris quatre lieu-action convergeant vers un lieu but. Tant que le robot part depuis l'intérieur du bassin d'attraction ainsi formé, il converge vers le but. Depuis des positions de départ plus lointaines, le robot navigue aléatoirement. Cette navigation aléatoire n'exclue pas que le robot finisse pas "tomber" dans le bassin d'attraction, cependant, le bassin d'attraction définit les limites de la convergence assurée vers le but.

4 Intégration de chemin et navigation proprioceptive

L'intégration de chemin est la capacité d'utiliser les informations proprioceptives correspondants aux mouvements effectués afin de construire une estimation ou

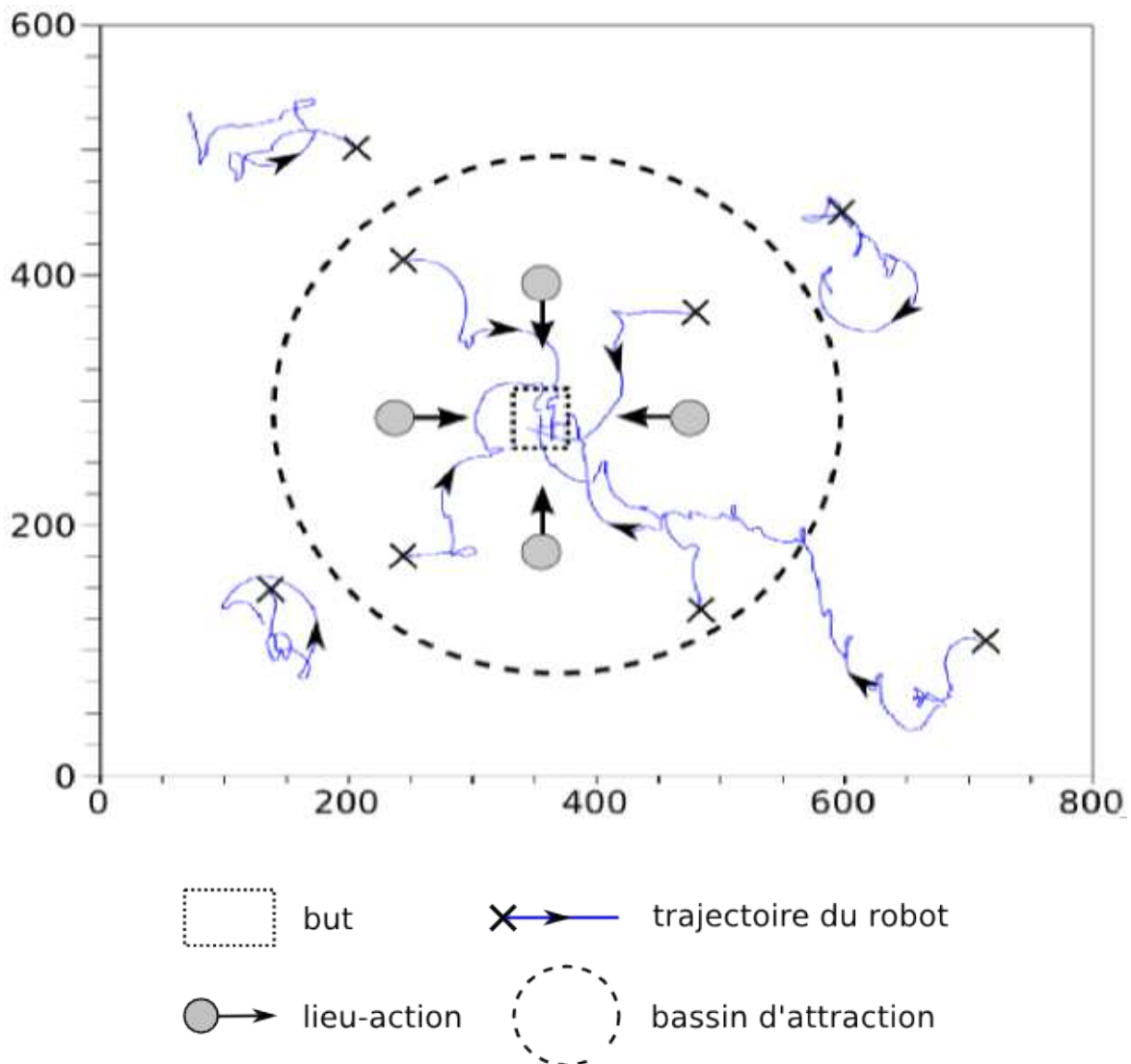


FIGURE 4.8 – Trajectoires du robot ayant appris un bassin d'attraction autour d'un but. Le robot a appris quatre couples lieu-action convergeant vers un but. Quand il est dans le bassin d'attraction, ses actions le rapprochent donc de son but. En dehors du bassin d'attraction, le robot navigue aléatoirement. C'est pourquoi il est possible que le robot rejoigne le bassin d'attraction mais cela repose sur le hasard et non sur un apprentissage.

représentation d'un trajet parcouru et de pouvoir ainsi rejoindre un point de sa trajectoire. Le principe de l'intégration de chemin repose sur la capacité à "sommer" les mouvements effectués depuis un point de départ en gardant la trace des distances parcourues et des angles effectués pour déterminer le vecteur de mouvement global (du point de départ vers la position courante) et son opposé, le vecteur de retour (de la position courante vers le point de départ). Cette capacité purement motrice peut être utilisée pour naviguer lorsque les repères visuels ne peuvent être utilisés. Cette technique de navigation utilisée par les animaux [Collett 2006, Gallistel 1996,

[Gallistel 1998, Mittelstadt 1980] a été appliquée à la navigation robotique bio-inspirés [Vickerstaff 2005, Gaussier 2007a, Haferlach 2007, Mudra 2003]. Pour un robot mobile, on parle le plus souvent d'odométrie. Plusieurs formalismes ont été mis au point pour implémenter l'intégration de chemin avec des réseaux de neurones. Le formalisme utilisé dans ces travaux est celui présenté dans [Gaussier 2007a].

Le principe de fonctionnement du calcul de l'intégration de chemin est illustré par la figure 4.9. L'intégration de chemin est un calcul itératif qui consiste à intégrer dans le temps les déplacements ayant lieu entre chaque itération. Ces déplacements sont caractérisés par une vitesse et une orientation particulière. Chacune de ces entrées à sommer est codée par une courbe en cloche centrée sur la direction du mouvement et dont l'amplitude correspond à la vitesse de ce mouvement. En théorie, pour retomber sur les résultats de la trigonométrie, cette courbe doit être un cosinus. En pratique, n'importe quelle courbe en cloche peut faire l'affaire si l'on n'a pas besoin de trop de précision. La courbe résultant de la somme de ces entrées caractérise le mouvement global depuis le point de départ. La position du maximum de cette courbe globale correspond à l'orientation du vecteur de mouvement global et la valeur de ce maximum correspond à sa norme (voir [Gaussier 2007a] pour une démonstration).

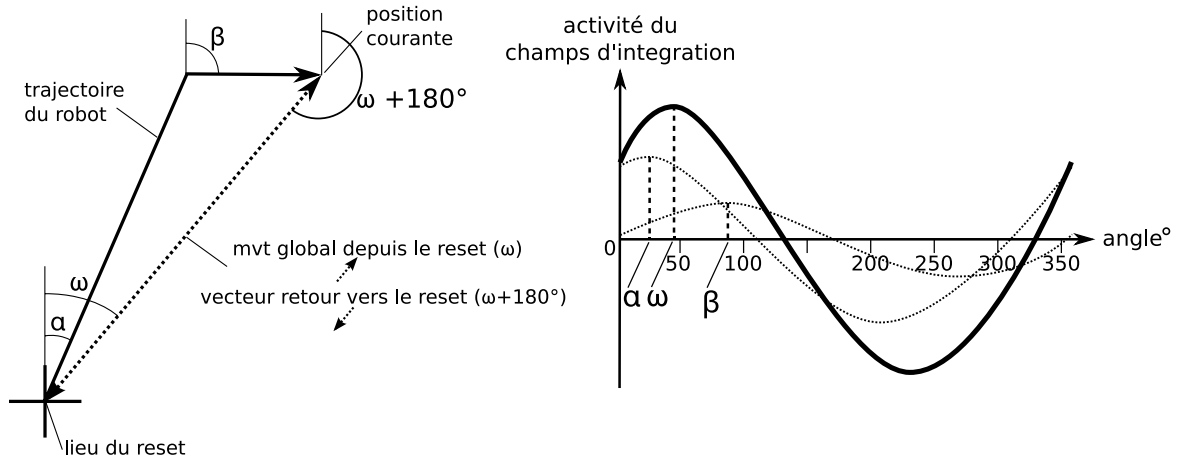


FIGURE 4.9 – Illustration du calcul d'intégration de chemin. La figure de gauche montre la trajectoire suivie par le robot. Il s'agit de deux segments de différentes longueurs (le premier est trois fois plus long que le second) et orientation (25° et 90° par rapport à une direction absolue arbitraire). La figure de droite montre ces deux segments sous la forme de deux courbes en cloche en pointillés centrées sur les directions absolues des mouvements correspondants (α et β). La courbe en gras est la somme de ces deux courbes. L'orientation du vecteur de mouvement global (ω) correspond à la position du maximum de cette courbe et sa norme est codée par la valeur de ce maximum.

La figure 4.10 montre comment l'intégration de chemin est implémentée en pratique à l'aide d'un réseau de neurones. La proprioception du robot permet de récupérer les informations d'orientation et de vitesse de ses mouvements. La vitesse est codée par

l'intensité de l'activité d'un neurone. Quand son activité est nulle, le robot ne se déplace pas. Ce neurone ne peut coder de vitesse au delà de sa valeur de saturation (valeur jamais atteinte en pratique). L'orientation du mouvement est codée par un neurone actif dans un champs de neurones. Ici, un champs de neurones signifie simplement un groupe de neurones (pas de connexions entre les neurones du champs). Cependant, la topologie du champs est importante : le champs complet représente 360° et chaque neurone code pour une direction particulière. Il en résulte que la taille du champs, i.e. le nombre de neurones qui le composent, définit sa précision ou la finesse de sa discrétisation. Pour un champs de taille N , chaque neurone code pour $\frac{360^\circ}{N}$. Cette orientation est convoluée par un masque cosinus et est ensuite multipliée par l'activité du neurone de vitesse. Par commodité d'écriture, les champs représentés possèdent des neurones dont l'activité peut être négative. L'utilisation d'une voie codant les activités positives et une autre voie pour les activités positives est tout à fait possible. Cependant, cette précision dans l'écriture alourdirait grandement tout les schémas explicatifs. A chaque itération, les mouvements du robots sont donc caractérisés par l'activité du champs de neurones "vitesse-direction". L'activité $V_i(t)$ de chacun des neurones de ce champs (vitesse du neurone de direction i au temps t) est calculée par l'équation 4.11 :

$$V_i(t) = S(t) \cdot \cos\left(\frac{d_w(t) - \theta_i}{N}\right) \quad (4.11)$$

N est le nombre de neurones de ce champs, $i \in [1 : N]$, $S(t)$ est l'activité du neurone vitesse, $d_w(t)$ est la position du neurone actif dans le champs de direction et $\theta_i = -2\pi \frac{i}{N}$.

L'intégration à proprement parler consiste à sommer dans le temps ces entrées instantanées vitesse-direction dans un champs de neurones de même taille via des connexions excitatrices un vers un (chaque neurone reçoit une connexion depuis le neurone correspondant dans le champs vitesse-direction). Ce champs d'intégration possède une réentrance (connexions récursives de chaque neurone vers lui-même). Il réalise donc simplement une somme cumulée. Un signal de reset vient inhiber toute activité sur le champs d'intégration. Le reset permet donc de définir un nouveau point de référence pour le calcul du mouvement global. L'activité $IC_i(t)$ de chacun de neurones de ce champs (activité du neurones i d'intégration de chemin au temps t) est calculée par l'équation 4.12 :

$$IC_i(t) = IC_i(t - dt) + \alpha V_i(t) - r(t) \quad (4.12)$$

N est le nombre de neurones de ce champs, $i \in [1 : N]$. $r(t)$ vaut 1 au moment du reset et 0 sinon. α est un facteur de gain permettant d'éviter la saturation des neurones du champs IC.

Une compétition sur ce champs d'intégration permet de détecter le neurone le plus actif et indique donc la direction du mouvement global. Puisque le vecteur retour est l'opposé du vecteur de mouvement global, un décalage de 180° suffit à passer de l'un à l'autre. En effectuant un reset de son champs d'intégration sur un lieu d'intérêt,

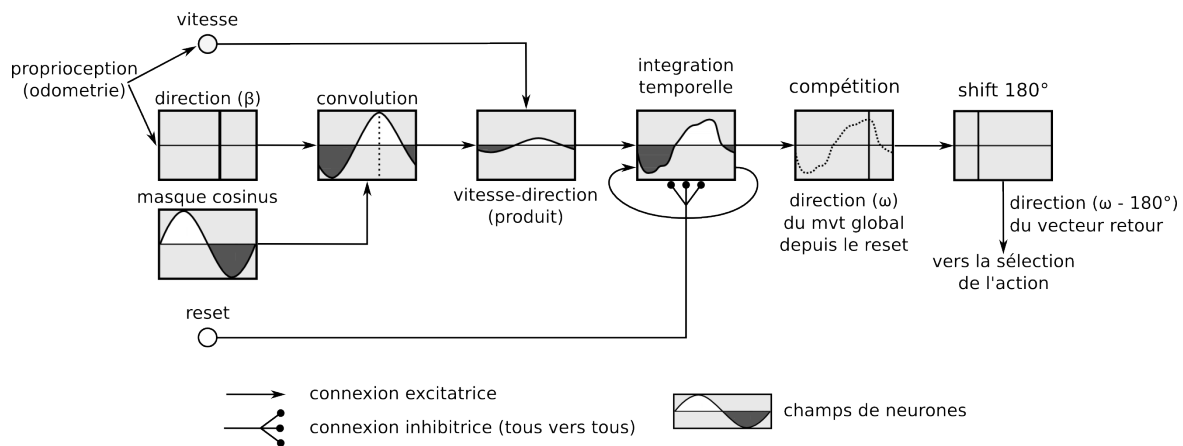


FIGURE 4.10 – La proprioception du robot permet de récupérer la vitesse et l’orientation du déplacement entre chaque itération. La direction est codée par un neurone actif dans un champs de neurones et la vitesse par l’activité d’un neurone. Cette direction convoluée par un cosinus puis multipliée par la valeur du neurone de vitesse donne le champs vitesse-direction. Les valeurs successives de ce champs sont intégrées par un champs de neurones récursifs depuis le dernier reset. Une compétition permet de déterminer la direction absolue du mouvement global. Un décalage de 180° donne la direction du vecteur retour. Un signal de reset permet d’inhiber toute activité sur le champs d’intégration et définit ainsi un nouveau point de référence pour le calcul du mouvement global.

le robot est capable à tout moment de suivre son vecteur retour en s’orientant dans sa direction et ainsi de rejoindre ce lieu particulier. L’intégration de chemin dérive dans le temps. A chaque pas de temps, des incertitudes sont cumulées. Il s’agit d’une limitation structurelle. Pour éviter qu’elle soit trop handicapante, il faut utiliser une discrétisation angulaire suffisamment fine et surtout, il faut périodiquement recalibrer le champs d’intégration sur le lieu but. Cependant, pour tirer pleinement parti de cette compétence, la limitation à un seul champs d’intégration doit être dépassée. Le robot doit pouvoir stocker plusieurs champs en parallèle, mis à jour dynamiquement, et sélectionner le champs d’intégration correspondant au lieu à rejoindre.

Les outils présentés dans ce chapitre sont utilisés dans l’ensemble des travaux réalisés. La modélisation neuronale est le contexte global et la navigation est le domaine d’application principal. Tout au long de ce travail, nous proposerons et testerons un certain nombre de modèles neuronaux utilisés comme des simulations très simplifiées de différentes structures du cerveau. De plus, les méthodes de navigation présentées seront ajustées et améliorées afin de permettre une plus grande flexibilité et une meilleure adaptation du robot à son environnement. Les stratégies de navigation présentées dans ce chapitre sont certes puissantes et flexibles (surtout si elles sont utilisées ensembles) mais à elles-seules, elles ne permettent pas de produire des comportements autonomes. Un point essentiel abordé tout au long de ce travail est la nécessité d’un système émotionnelle de bas-niveau pour amorcer l’autonomie du comportement. Dans

les chapitres suivants, nous décrirons comment un système de valeur endogène ancré dans les sensations de douleur et de plaisir permet d'amorcer l'autonomie comportementale d'un robot tant en termes d'autonomie d'apprentissage que d'auto-évaluation et d'auto-régulation.

Douleur, plaisir, motivations et conditionnements

1 Introduction

Le plaisir et la douleur sont des sensations fondamentales, mais il convient néanmoins de définir précisément ce qu'elles veulent dire dans le contexte de l'expérimentation robotique et au sens large dans cette étude. Bien entendu, ces deux termes peuvent être pris dans un sens plus large pour les humains que pour les autres animaux. Le plaisir peut être une sensation physique (alimentation, comportements sexuels, ...) autant que le résultat d'une évaluation cognitive impliquant les fonctions de haut niveau de l'intellect (plaisir intellectuel, professionnel, parental, moral, ...). La douleur est comprise assez naturellement comme la sensation désagréable signalant une atteinte à l'intégrité du corps. Mais, à l'instar du plaisir, la douleur peut aussi être plus cognitive. C'est le cas par exemple de la douleur d'avoir perdu un être cher. Le fil directeur de cette thèse est d'essayer de comprendre les émotions comme des phénomènes émergents de mécanismes simples et de bas niveau. Il ne s'agit pas de modéliser des émotions sophistiquées comme la honte ou la fierté, qui demandent une conception élaborée d'autrui. En prenant un point de vue commun entre le comportement humain et animal, le plaisir peut être défini comme la satisfaction d'un besoin physiologique et la douleur comme le signal avertissant d'un dommage corporel. Ce premier niveau d'information sur les interactions d'un être sensible avec son environnement va nous servir de base pour les capacités fondamentales de régulation du comportement. Le neuro-psychologue Antonio Damasio voit le plaisir et la douleur comme les soubassements du fonctionnement des motivations et des drives (i.e. motivation de bas niveau comme la faim) qui à leur tour forment la base des émotions (figure 1.1) [Damasio 2003]. Selon [Craig 2003], la douleur est non seulement une sensation, mais aussi une émotion homéostatique au même titre que la faim et la soif. Un nombre grandissant d'études se sont penché sur le lien entre douleur et émotion [Keefe 2001, Vogt 2005], soulignant l'importance de cette stimulation fondamentale dans le fonctionnement des comportements aversifs et des émotions négatives. Les liens entre la sensation de plaisir et le fonctionnement des émotions a été sensiblement moins étudié. Mais son importance dans l'origine des motivations était déjà soulignée par [Duncker 1941] publié en 1941. De même que pour l'étude de la douleur, les progrès en imagerie cérébrales ont permis de préciser les structures cérébrales impliquées ainsi que les différents circuits d'information liée à la sensation de plaisir et ses liens avec les émotions positives ou source de récompenses [Burgdorf 2006, Kringelbach 2009].

Les mécanismes permettant d'apprendre à associer ces signaux fondamentaux aux

différentes perceptions ou actions reposent sur les deux types de conditionnements présentés précédemment. Les comportements d'appétence et d'aversion appris par conditionnement avec les stimulations source de douleur ou de plaisir peuvent être vu comme le résultat d'une régulation émotionnelle de bas niveau. Le propos de ce chapitre est de décrire un premier niveau de régulation émotionnelle du comportement reposant sur de simples mécanismes de conditionnement au plaisir et à la douleur. Dans un premier temps nous décrirons l'implémentation de la physiologie artificielle du robot à partir de laquelle sont construits les signaux de douleur et de plaisir. Nous présenterons ensuite notre modèle simplifié de l'hypothalamus qui permet de générer des *drives* ou motivations de bas niveau comme la faim ou la soif. Ces drives seront nécessaires pour constituer une régulation de base (besoins satisfaits/insatisfaits). Ensuite, nous montrerons comment des apprentissages de conditionnements utilisant ces signaux permettent au robot de développer des comportements appétitifs ou aversifs.

2 Physiologie artificielle, douleur, plaisir et motivation

Inspiré par l'approche *animat* [Wilson 1991] selon laquelle le robot est vu comme la simulation d'un animal, une physiologie artificielle et un modèle simple d'hypothalamus ont été programmés pour notre robot. Les sources de douleur et de plaisir sont donc transposées aux contraintes du robot (et sont volontairement très simplifiées). L'état physiologique du robot est qualifié par des variables simulant le taux d'hydratation ou le taux de glucose. Le niveau de ces variables physiologiques essentielles décroît avec le temps puisque le robot utilise ses ressources internes pour son fonctionnement. En pratique, une ressource est simulée par un rectangle de couleur sur le sol. A chaque ressource correspond une couleur spécifique (bleu pour l'eau, rouge pour la nourriture). Sous le châssis du robot, un détecteur de couleur lui permet de percevoir les ressources. Il est important de noter que le robot ne peut pas "voir" les ressources au sol directement avec sa caméra. Quand le robot détecte et récolte une ressource simulée (e.g. eau), cela produit logiquement l'augmentation du niveau de la variable physiologique correspondante (taux d'hydratation). La survie du robot n'est possible que s'il récolte périodiquement les ressources dont il a besoin afin que ses variables physiologiques ne descendent pas en dessous d'un seuil critique provoquant une mort simulée. Les équations 5.1 et 5.2 décrivent l'évolution temporelle des variables physiologiques :

$$N(dt) = \alpha_n(N_{max} - N(t))n - \beta_n N(t)(1 - n) \quad (5.1)$$

$$E(dt) = \alpha_e(E_{max} - E(t))e - \beta_e E(t)(1 - e) \quad (5.2)$$

$N(t)$ et $E(t)$ sont les niveaux de nourriture et d'eau, N_{max} et E_{max} valent 1 (saturation), n et e valent 1 quand, respectivement, de la nourriture ou de l'eau sont ingérés. α_n et α_e sont des facteurs de vitesse d'ingestion des ressources, β_n et β_e sont des facteurs de vitesse de consommation. Sachant que fréquence d'exécution de la simulation est de 10Hz, les valeurs suivantes ont été utilisées : $\alpha_n = \alpha_e = 0.3$ (l'ingestion

est rapide), $\beta_n = \beta_e = 0.03$ (la consommation est lente). La figure 5.1 représente le réseau de neurones utilisé pour modéliser ce système de drives.

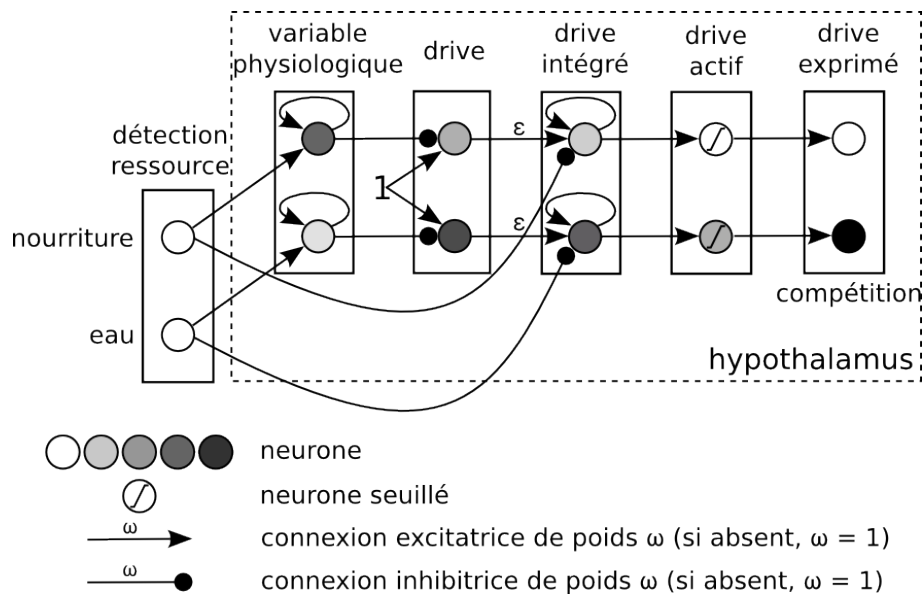


FIGURE 5.1 – Drives ou motivations de bas niveau générées dans un modèle simple de l’hypothalamus. Les drives internes sont calculés directement à partir des variables physiologiques. Les drives intégrés peuvent être manipulés sans changer pour autant l’état physiologique réel du robot. Les drives sont actifs s’il dépassent un seuil d’expression. Le drive gagnant est le résultat de la compétition entre les drives.

Pour réagir aux niveaux des différentes variables physiologiques, nous avons développé un modèle très simplifié de l’hypothalamus. Un système de motivations de bas niveau, ou système de *drives*, réagit à l’état physiologique du robot. Quand le niveau d’une variable physiologique descend (e.g. taux de glucose), le drive correspondant augmente (faim). C’est ce système rudimentaire des variables physiologiques et de drives générés dans l’hypothalamus qui donne un but au robot : la survie. Cependant, les drives ne se réduisent pas à l’expression directe du niveau des variables physiologiques. Les drives réagissent directement à la physiologie du robot (biais activateur et inhibition depuis le niveau de la variable physiologique). Plus le niveau d’une variable est faible, plus le drive est fort. Les drives intégrés cumulent dans le temps la valeur des drives internes (multiplié par un facteur ϵ pour ne pas saturer trop rapidement). Les drives sont actifs s’ils dépassent un seuil d’activation (valeur fixe définie manuellement). Plus ce seuil est bas et plus le robot déclenche rapidement le comportement associé (chercher de la nourriture ou de l’eau). Enfin, le drive gagnant est le drive le plus actif (compétition). Il est intéressant de noter qu’une compétition souple entre les drives permet l’apparition de comportements opportunistes (e.g. rejoindre une ressource même peu convoitée du moment qu’elle est proche). Mais afin de simplifier notre système ainsi que le comportement du robot, nous avons opter pour une compétition stricte. Si cela limite la possibilité de faire émerger un comportement intéressant, cela permet de comprendre plus simplement le

comportement du robot. La distinction entre drive interne, intégré, actif et gagnant permet une plus grande souplesse de ce système de drive bas niveau. Les drives intégrés peuvent être manipulés par des influences de plus haut niveau sans nécessiter pour autant de manipuler l'état réel de la physiologie du robot. Si les drives intégrés sont inhibés, c'est donc bien le drive interne le plus important qui deviendra le plus rapidement actif et donc gagnant. Enfin, la différence entre les drives actifs et le drive gagnant permet de distinguer un besoin présent, d'un besoin qui guide le comportement. A tout moment le robot peut avoir plusieurs besoins. Mais un besoin prioritaire oriente son comportement. Quand une ressource est détectée, le niveau de la variable physiologique correspondante augmente et le drive intégré est remis à zéro.

Nous définissons le plaisir comme la satisfaction d'un besoin physiologique et la douleur comme le signal avertissant d'un dommage corporel. Satisfaire sa faim et sa soif doit donc être source de plaisir pour le robot. Un signal de plaisir doit donc être généré dans un modèle de l'aire tegmentale ventrale. A contrario, un niveau très bas en eau ou en glucose dans l'organisme pose des risques de dommages corporels et peut générer un signal de douleur dans un modèle de la substance grise périaqueducule. Si ce ne sont pas directement ces niveaux physiologiques mais plutôt leur conséquences qui peuvent générer de la douleur, il n'a pas semblé pertinent dans ce travail de pousser le réalisme sur ces aspects du fonctionnement biologique d'un être vivant. La manière dont les signaux de plaisir et de douleur dérivent de la physiologie simulée du robot est décrit dans la figure 5.2.

Quand le niveau d'une variable physiologique est suffisamment bas (e.g. niveau d'hydratation), le drive correspondant devient actif (soif). Si de l'eau est détectée et "consommée", le niveau d'hydratation remonte et le drive est alors satisfait. Un signal de plaisir est alors produit (équation 5.4). Au contraire, un seuil absolu sur le niveau des variables physiologiques définit l'apparition de la sensation de douleur (équation 5.5). Ce seuil permet de détecter qu'une variable atteint un niveau critique. Contrairement au niveau d'activation du drive, ce signal n'indique pas un besoin mais l'arrêt prochain des fonctions vitales.

$$Douleur = \begin{cases} 1 & \text{si } 1 - V_n(t) > S_d \\ 0 & \text{sinon} \end{cases} \quad (5.3)$$

$$Plaisir = \begin{cases} 1 & \text{si } \omega_r R_{detect} + \omega_d D_R > S_p \\ 0 & \text{sinon} \end{cases} \quad (5.4)$$

$V_n(t)$ est le niveau de la variable physiologique n à l'instant t . S_d , le seuil de douleur, est une valeur fixée *a priori*. Plus cette valeur est proche de 1 plus il faut que $V_n(t)$ soit petit pour déclencher le signal de douleur. R_{detect} vaut 1 quand la ressource R est détectée et D_R vaut 1 quand le drive correspondant à la ressource R est actif. S_p , le seuil de plaisir est supérieur à ω_d (poids de la connexion depuis le drive actif) et ω_r (poids depuis la détection de ressource) agissant comme un opérateur "ET". Nous avons opter pour des signaux de douleur et de plaisir binaires afin de simplifier notre système. Des niveaux de douleur et de plaisir continus pourraient permettre de faire émerger de comportements intéressants mais ici aussi nous avons favorisé la simplicité.

douleur.

3 La peur comme un mécanisme de conditionnement négatif

3.1 Principe et architecture

Chez l'homme ou l'animal, quand une stimulation douloureuse est ressentie, une réaction réflexe de retrait ou d'évitement se produit. Elle permet de faire cesser le plus rapidement possible tout dommage ayant pu être subi. Cette réaction se distingue par sa rapidité, ce qui la rend particulièrement utile pour préserver l'organisme. Cette rapidité est atteinte par le fait que le message nerveux court-circuite les structures supérieures du système nerveux central. L'arc réflexe décrit la circulation du message nerveux. Il part de l'afférence (le capteur) ayant généré le signal de douleur, remonte vers la colonne vertébrale, puis repart vers l'effecteur (le muscle) correspondant pour provoquer une contraction. Ce circuit n'implique pas la partie supérieure du système nerveux central. Le mouvement de réaction est donc incontrôlé. Cependant, ce n'est pas parce qu'une réaction réflexe a lieu qu'aucune information ne parvient au cerveau. Bien au contraire, une fois que la douleur est perçue, l'organisme est bien en alerte. Nous avons développé un modèle simple permettant d'apprendre à anticiper la douleur en s'inspirant des conditionnements appris par l'amygdale entre des stimuli divers et les sensations de douleur et de plaisir. Au plus bas niveau d'apprentissage, les conditionnements apprennent des associations entre les informations sensorielles corporelles et environnementales. L'ensemble des perceptions du monde extérieur peuvent être conditionnées à la douleur. Quand la perception de l'environnement est proche des conditions présentes lors de l'apprentissage (cf. chapitre 3, navigation visuelle et proprioceptive), une réaction de douleur conditionnée est déclenchée au niveau de l'amygdale. Plus cette proximité est grande et plus la douleur conditionnée est importante. Ces conditionnements peuvent donc servir de base à la constitution d'une attente de stimulation douloureuse. Il suffit alors que la douleur conditionnée provoque la même réaction que la douleur réelle pour que la réaction d'évitement anticipe la sensation effective de douleur. La figure 5.3 est un schéma conceptuel de ce conditionnement amygdalien, de la voie réflexe et de la voie conditionnée. Cette voie conditionnée est très simple dans son fonctionnement. Elle présente cependant des capacités particulièrement intéressantes. Ce mécanisme pourrait servir à réguler le comportement en interrompant l'action en cours pour imposer une réaction d'évitement anticipant un événement déplaisant. La description qui est faite ici correspond bien à notre définition de travail de la peur.

Afin de mettre en oeuvre cette idée sur un robot mobile, un réseau de neurones a été conçu pour tester cette voie d'évitement conditionné (voir figure 5.4). C'est assez naturellement que vient l'idée de tester ce mécanisme pour apprendre à éviter les obstacles. Un grand nombre de techniques ont été élaborées pour réaliser l'évitement des obstacles et le but n'est pas de proposer un nouveau mécanisme plus performant mais d'illustrer le potentiel de ce mécanisme de conditionnement. Donc, comme décrit précédemment, en dessous d'une certaine distance entre le robot et un obstacle, un

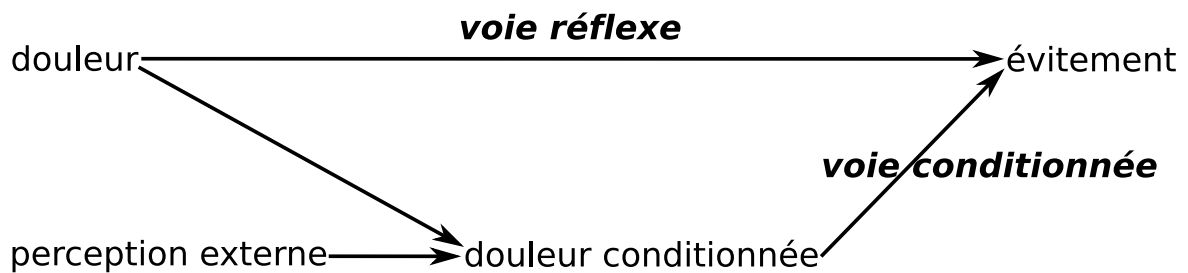


FIGURE 5.3 – Conditionnement à la douleur réalisé par l’amygdale. La douleur provoque une réaction de l’organisme via la voie réflexe (une douleur localisée provoque un mouvement dans la direction opposée). L’association de ce signal douloureux avec la perception externe permet une réaction de douleur conditionnée quand des conditions perceptives semblables sont rencontrées. Si la douleur conditionnée provoque la même réaction que la douleur elle-même, la voie conditionnée pourrait permettre d’anticiper la douleur en provoquant une réaction d’évitement.

signal de douleur est généré, simulant la rencontre avec l’obstacle et le robot. Ce signal produit la réaction réflexe qui consiste à aller dans la direction opposée à la direction dans laquelle est perçue la douleur. Bien sûr, l’histoire s’arrête ici si l’on considère qu’il suffit de fixer le seuil de douleur sur la réponse du capteur de proximité à une valeur plus basse. Cependant, il s’agit alors de répondre au problème par l’ajout d’un capteur spécifique et d’un seuil déterminé manuellement pour répondre aux besoins de l’expérimentateur. Le but était simplement de simuler un capteur de pression. Or un capteur de pression ne peut pas directement permettre d’anticiper un obstacle. Donc le but pour le robot est d’apprendre l’association entre la perception d’une stimulation douloureuse et sa perception de l’environnement. Cette perception peut être réalisée grâce à son système de vision (caméra sur système pan permettant de regarder à 360°). Cette perception pourrait aussi être assurée par un télémètre laser. En réalité, par tout système de perception à distance dépeignant une image stable de l’environnement. L’audition pourrait aussi être utilisée si elle permettait de caractériser des lieux. Par exemple si différentes sources de son permettaient une sorte de triangulation. Quoi qu’il en soit, le processus reste le même. Une perception caractérisée par un flux d’informations codées sur une population de neurones doit être catégorisée afin qu’un état particulier puisse être reconnu. Pour la vision, les cellules de lieu visuelles du cortex entorhinal et du gyrus dentelé (hippocampe) sont le résultat de cette catégorisation. Le contrôle de son apprentissage a été introduit dans le chapitre 3, lors de la description du processus de catégorisation. Cependant, son utilisation dans ce modèle de conditionnements à la douleur nécessite une régulation plus fine. Le recrutement d’une nouvelle catégorie dépend toujours d’un seuil de vigilance, mais d’autres facteurs rentrent en jeu. D’abord, la perception de douleur vient augmenter la vigilance. Ceci correspond au fait que la douleur étant une information hautement pertinente, sa détection pousse le robot à analyser plus précisément son environnement et assure ainsi que l’information douloureuse est associée à un contexte perceptif suffisamment précis pour être utile et exploitable (e.g. ne pas éviter une zone neutre). Pour

contre-balancer l'effet de la douleur sur le contrôle de la catégorisation, nous avons modélisé un mécanisme d'habituation. Sans ce mécanisme, nous observerions un effet du sur-apprentissage dans les zones où un signal douloureux est détecté (densité très élevée de lieux appris sans gain en précision pour la navigation). Ce mécanisme est donc nécessaire mais il ne s'agit pas d'un élément spécifique de notre modèle de la peur. Il s'agit plutôt d'un élément de régulation de tout système afin d'éviter le sur-apprentissage. Quand un signal de recrutement est déclenché, son potentiel est stocké dans une mémoire dynamique qui décroît lentement (neurone avec une connexion récurrente de poids $\omega = 1 - \varepsilon$). Ce potentiel qui décroît inhibe en retour le déclenchement d'un nouveau signal de recrutement. Les équations 5.5 et 5.6 décrivent le contrôle du recrutement.

$$R(t) = \begin{cases} 1 & \text{si } V + D(t) - C_{max} - H(t) > 0 \\ 0 & \text{sinon} \end{cases} \quad (5.5)$$

$$H(t) = (1 - \varepsilon)^{t-t_R} \quad (5.6)$$

$R(t)$ est le signal de recrutement, V est le biais de vigilance (une constante comprise entre 0 et 1), $D(t)$ est le signal de douleur, C_{max} est l'activité de la catégorie gagnante et $H(t)$ est l'inhibition venant de l'habituation. Pour le calcul de $H(t)$, $(1 - \varepsilon)$ correspond au poids de la connexion récurrente du potentiel d'habituation et t_R correspond à t au dernier recrutement.

Une compétition permet de déterminer quelle catégorie est la mieux reconnue. Le conditionnement de cette catégorie au signal de douleur est réalisé par un apprentissage hebbien. Les neurones servant à donner la réponse de douleur conditionnée reçoivent des entrées excitatrices depuis les signaux de douleurs et des connexions plastiques depuis la compétition sur les catégories perceptives. La douleur est représentée sur un champs de neurones qui caractérise les différentes douleurs (venant des différentes directions). L'apprentissage de ce conditionnement est contrôlé par le signal de douleur lui-même. Ainsi, quand une douleur survient, la connexion entre la douleur correspondante dans le groupe du conditionnement et la catégorie gagnante est renforcée. Quand cette catégorie perceptive sera à nouveau gagnante, un signal de douleur conditionné sera généré au même endroit dans le champs de neurones. En utilisant cette réponse comme le signal de douleur authentique, le même comportement d'évitement sera produit et le robot évitera le lieu en question avant de subir des dommages.

3.2 Expérimentation robotique

Ce réseau de neurones a été implémenté et testé sur robot mobile. Dans un premier temps, la voie conditionnée est désactivée : tout apprentissage de conditionnement est inhibé. Le robot produit le réflexe d'évitement quand il perçoit de la douleur (proximité d'un obstacle supérieur au seuil de douleur) mais il est incapable d'apprendre à anticiper cette douleur. Quand un lieu est appris, aucune action ne lui est associée. C'est donc le bruit dans l'activité du champs de neurone de douleur conditionnée qui

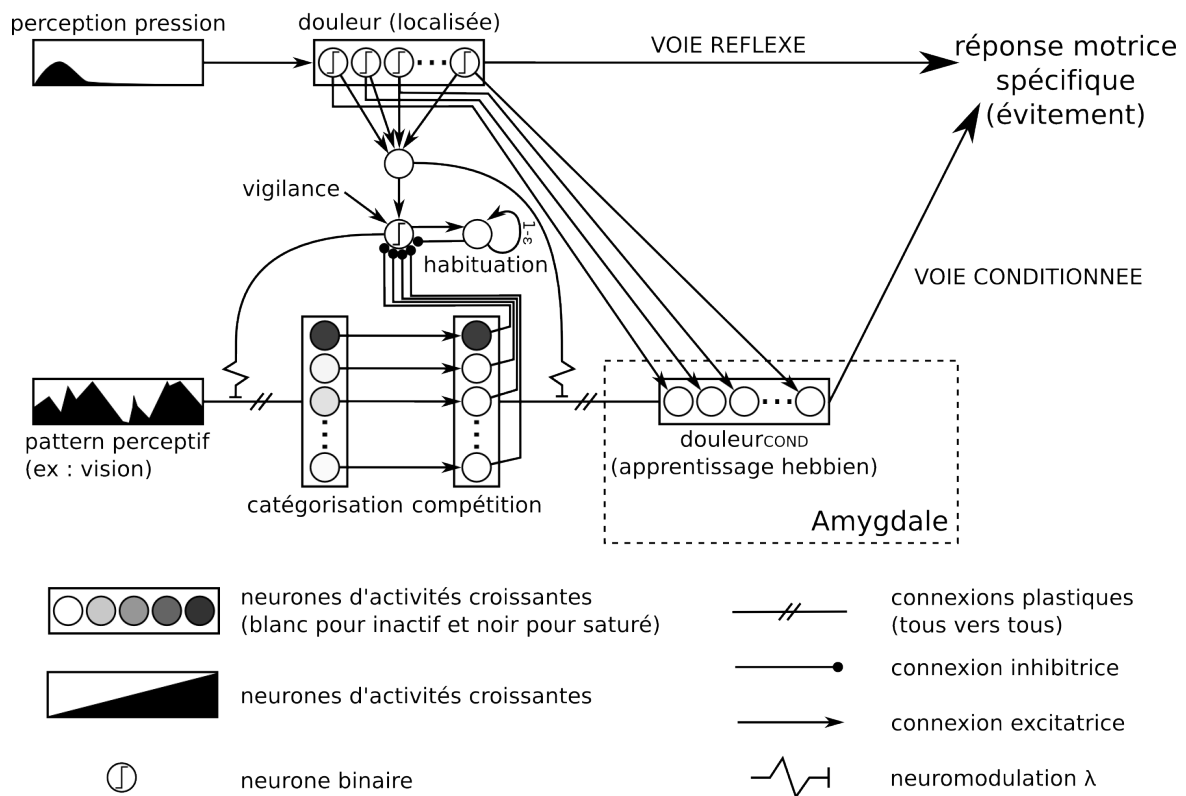


FIGURE 5.4 – Réseau de neurones permettant le conditionnement amygdalien d'un stimulus perceptif à la douleur pour produire un comportement d'évitement d'obstacle. Au delà d'un seuil sur la perception de pression sur le robot, un signal de douleur est produit. Un conditionnement est appris entre ce signal et la catégorie perceptive gagnante au niveau de l'amygdale. La douleur incite le recrutement d'une nouvelle catégorie, ce qui assure une bonne représentativité des lieux impliquant une stimulation douloureuse. Tandis que la voie réflexe déclenche automatiquement un comportement de retrait ou d'évitement, la voie conditionnée permet d'anticiper le signal douloureux et produit l'évitement avant que les dommages ne soient subits.

est utilisé comme commande motrice et le robot navigue donc aléatoirement. Le robot peut revenir vers un obstacle qu'il a déjà rencontré et pour autant continuer dans sa direction jusqu'à atteindre le seuil de douleur avant de réagir. La figure 5.5 montre la trajectoire aléatoire du robot dans cette condition sans apprentissage.

Ensuite, l'apprentissage dans la voie conditionnée est rétabli et le robot peut apprendre à associer la reconnaissance d'un lieu à une perception douloureuse et donc à l'action réflexe d'évitement. Le robot explore ainsi son environnement en apprenant périodiquement des lieux visuels (c.f. contrôle de la neuromodulation de l'apprentissage de lieu). Lorsqu'une stimulation douloureuse est ressentie et que la réaction réflexe d'évitement est produite, l'action produite est associée au lieu le mieux reconnu. Si le lieu le mieux reconnu est en dessous du seuil de neuromodulation, un nouveau lieu appris et est associé à la direction à prendre pour l'évitement. La figure 5.6 montre la

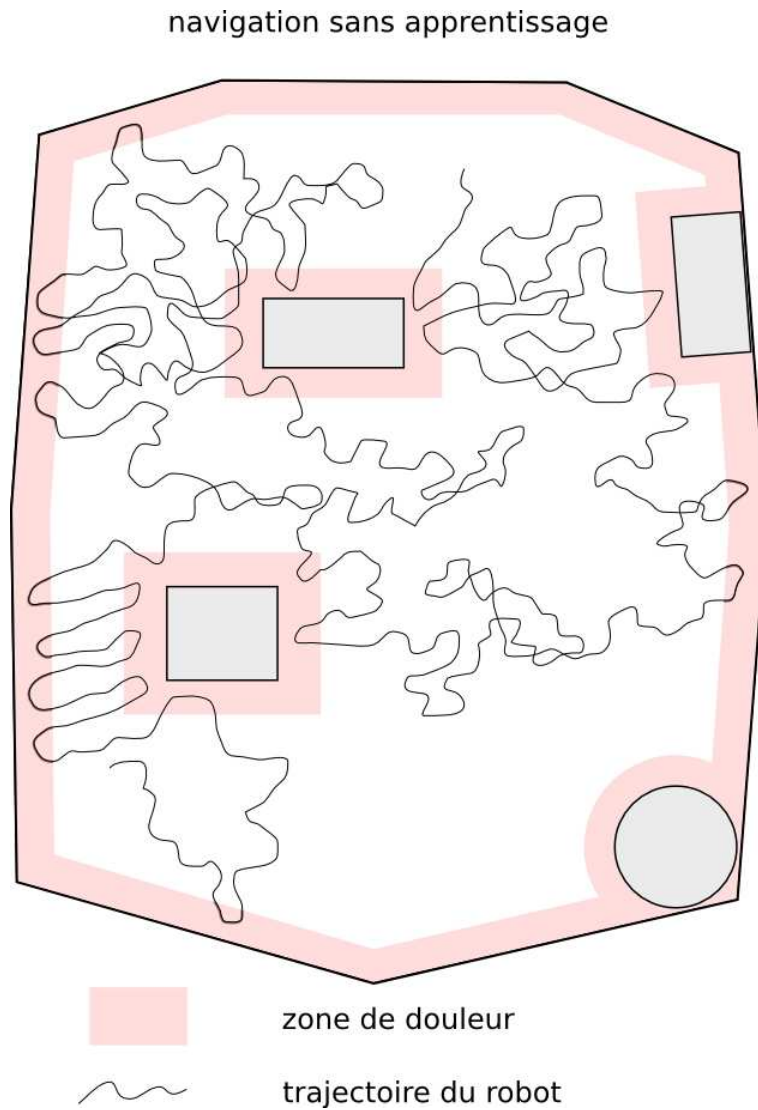


FIGURE 5.5 – Trajectoire du robot sans apprentissage de conditionnements à la douleur. La douleur ne déclenche ni l'apprentissage d'une cellule de lieu, ni l'apprentissage d'un conditionnement sensorimoteur à la douleur. Le robot navigue donc aléatoirement et produit simplement un réflexe d'évitement quand la douleur est ressentie.

trajectoire du robot durant les trois premières minutes d'apprentissage. L'effet de la douleur sur le seuil de neuromodulation de l'apprentissage de lieu concentre les lieux appris dans les zones où la douleur est ressentie.

Après apprentissage (environ dix minutes), le robot navigue aléatoirement quand il reconnaît un lieu neutre et il produit une réaction d'évitement anticipé quand il reconnaît un lieu conditionné à la douleur. Cette anticipation de la douleur lui permet d'éviter les obstacles. Il est à noter cependant qu'une grande concentration de lieux n'est pas un avantage. Trop de lieux appris proches les uns des autres provoque une confusion dans la compétition entre les lieux. Les actions conditionnées à ces lieux sont alors exécutées de manière chaotique et ne permettent pas de produire le comportement d'évitement souhaité.

4 Hédonisme et motivations

Le réflexe de retrait lié à la perception de douleur ainsi que les conditionnements amygdalien permettant d'éviter la douleur en anticipant la réaction d'évitement illustrent les mécanismes d'aversion et la tendance à l'éloignement ou la répulsion ("s'éloigner de"). Ces mécanismes reposent sur l'apprentissage de conditionnements de type stimulus-stimulus (e.g. vision-douleur) au niveau de l'amygdale. Au contraire, le plaisir est à la base des comportements d'appétence et la tendance à se rapprocher ("aller vers"). Cependant, s'éloigner d'une stimulation déplaisante est plus simple que d'aller vers une stimulation plaisante (ce phénomène est décrit et discuté dans le chapitre 8). Les comportements appétitifs ne peuvent reposer sur un équivalent positif du réflexe d'évitement. C'est pourquoi les mécanismes qui sont présentés ici font le lien entre, d'un côté la physiologie simulée du robot, la sensation de plaisir et les motivations qui en découlent et d'un autre côté les stratégies de navigation dont il dispose.

Le plaisir a été défini plus tôt comme résultant de la satisfaction d'un besoin physiologique. Cependant, ce n'est pas directement le plaisir qui oriente le comportement mais plutôt le besoin. C'est parce que l'on a soif que l'on boit et c'est parce que l'on boit que l'on ressent du plaisir. La soif peut donc être comprise comme la force qui attire vers le plaisir de boire. Nous avons vu précédemment que la navigation visuelle (sensori-motrice) et la navigation proprioceptive (intégration de chemin) permettent des comportements de convergence vers un lieu précis de l'environnement. Afin de pouvoir utiliser cette compétence pour satisfaire ses différents besoins, le robot doit apprendre les liens entre son système de drives et ses apprentissages sensori-moteurs ou son intégration de chemin. Ce lien peut être appris grâce à des conditionnements. C'est ce que nous montrerons dans cette section.

4.1 Motivations et navigation sensori-motrice

Pour converger vers un lieu but, la navigation sensori-motrice repose sur l'apprentissage d'un bassin d'attraction visuel. Un nombre suffisant d'associations lieu-action convergentes sont apprises autour du but. Cet apprentissage définit un bassin d'attraction permettant au robot de converger vers le but. S'il existe plusieurs buts satisfaisant

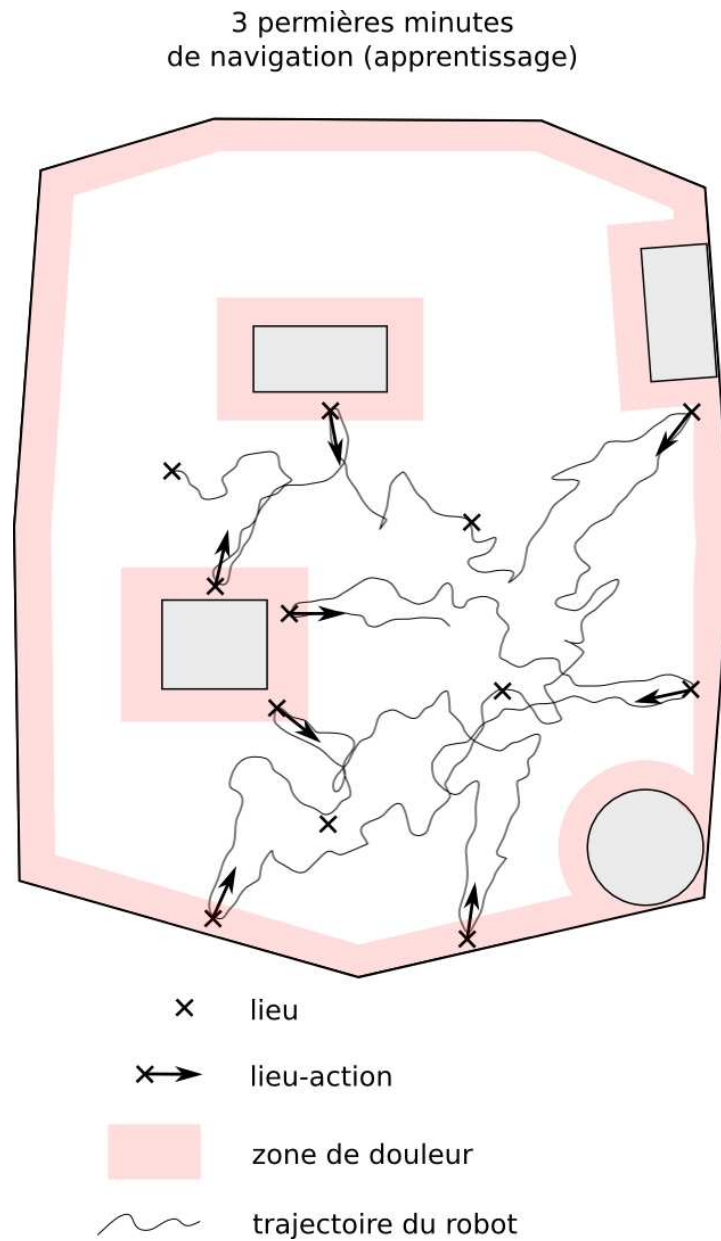


FIGURE 5.6 – Trajectoire du robot pendant les trois premières minutes de navigation avec apprentissage des conditionnements à la douleur. La perception de douleur déclenche le plus souvent l'apprentissage d'un lieu (rehaussement du seuil de neuromodulation de l'apprentissage des lieux). Dans tous les cas, l'action réflexe d'évitement est conditionnée à la reconnaissance du lieu courant. Le robot apprend progressivement à anticiper les situation provoquant de la douleur.

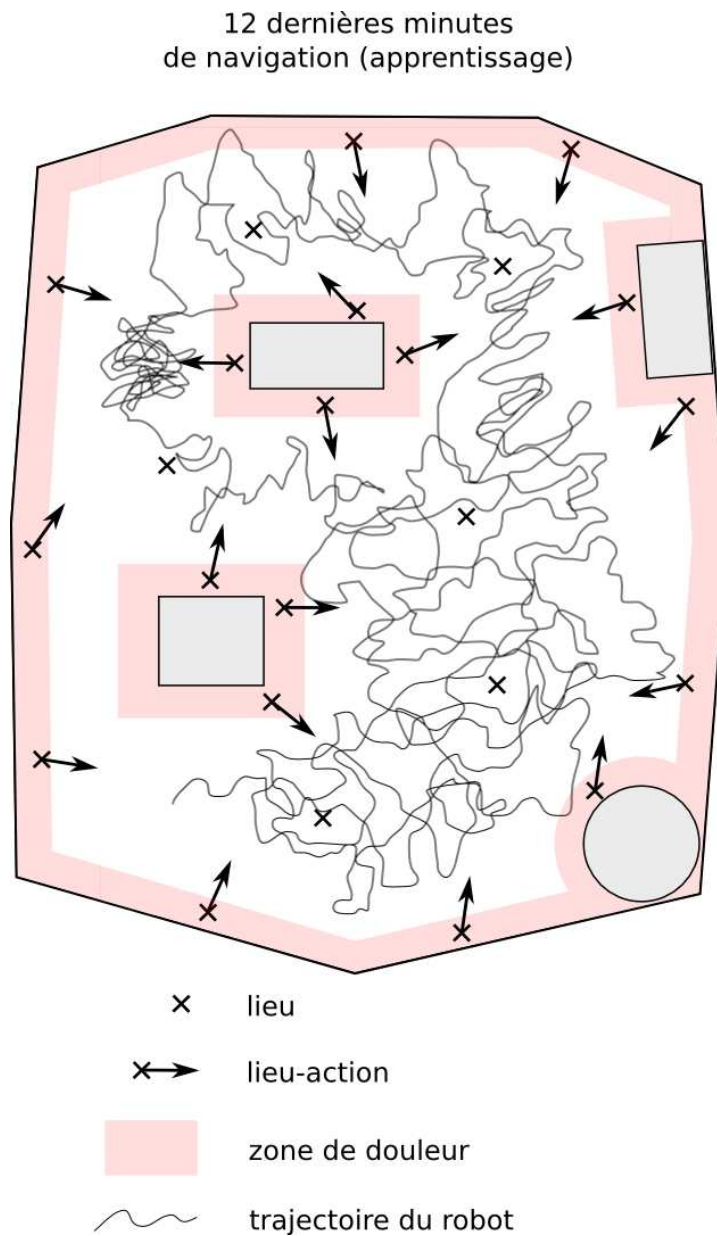


FIGURE 5.7 – Trajectoire du robot pour les douze dernières minutes de navigation avec apprentissage des conditionnements à la douleur. Il s'agit de la suite de la trajectoire présentée dans la figure 5.6 (après apprentissage donc). Le robot navigue aléatoirement lorsqu'il n'est pas dans un lieu conditionné à la sensation douloureuse. Par contre il anticipe le réflexe d'évitement quand il reconnaît un lieu associé à la perception de la douleur. Après apprentissage, le robot évite la zone douloureuse avant de détecter la sensation de douleur.

des drives différents, le robot doit apprendre un bassin d'attraction pour chacun des buts. Cependant, cet apprentissage ne garantit pas au robot de rejoindre le but désiré. En réalité, le robot convergera inexorablement vers le bassin d'attraction le plus proche (dont les lieux appris sont les mieux reconnus). Un conditionnement peut être appris entre la reconnaissance de lieu et les motivations. Un biais motivationnel doit permettre de faire pencher la compétition du côté des lieux composant les bassins d'attraction qui correspondent au but désiré. En conditionnant les lieux d'un bassin d'attraction avec le drive correspondant au but qu'il contient, l'activation de ce drive rehausse l'activité de ces lieux. Le drive gagnant oriente ainsi le robot vers le but désiré en biaisant la compétition entre les lieux. Un simple conditionnement hebbien entre les lieux et les drives permet d'apprendre ce biais.

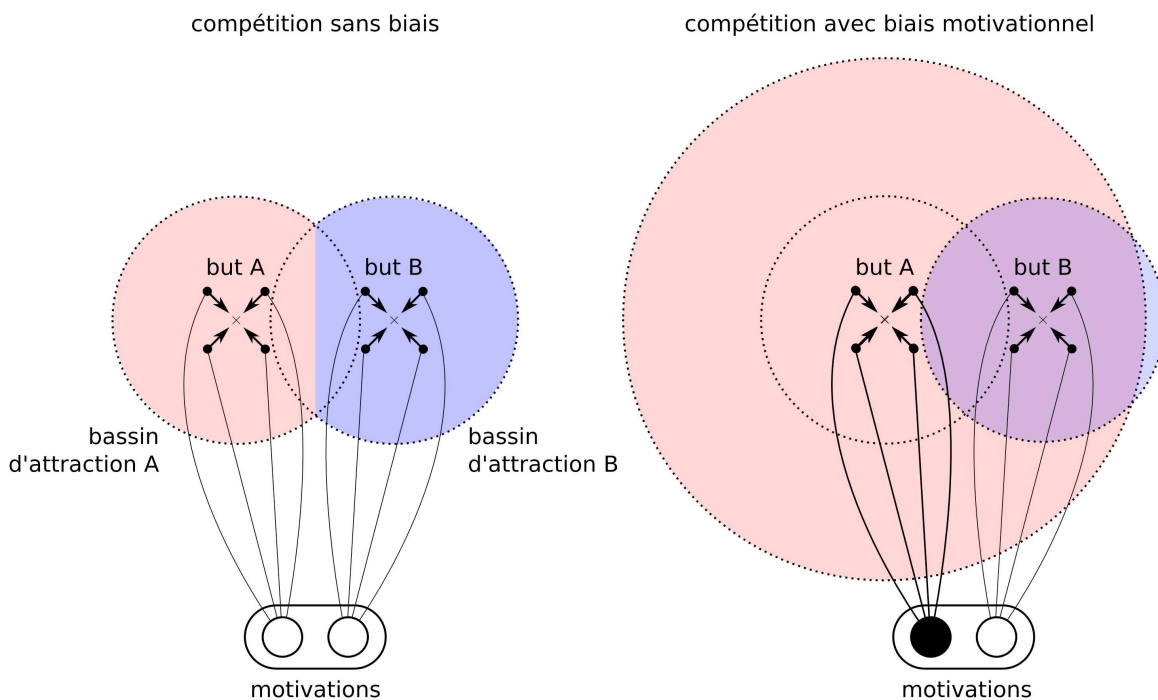


FIGURE 5.8 – L'introduction d'un simple biais lié aux drives appris grâce à un conditionnement doit permettre d'agrandir ou de réduire un bassin d'attraction en fonction du niveau du drive correspondant. Cela permet au robot d'utiliser la dynamique d'attraction qui correspond au drive gagnant. Les frontières changent en fonction de la modulation introduite par les drives au niveau de la compétition entre les cellules de lieux.

Il n'y a donc que très peu de modifications à apporter à l'architecture PerAc présentée précédemment. Un groupe de neurones appelés neurones "lieux-drive" est rajouté avant la compétition qui détermine quel lieu est le mieux reconnu. Ce groupe reçoit des connexions activatrices simple depuis les lieux (un vers un) et des connexions plastiques depuis les drives (tous vers tous). En stimulant le drive gagnant afin qu'il corresponde au but découvert, les lieux appris à proximité apprennent un conditionnement. La compétition se fait alors sur la base de la reconnaissance visuelle des lieux mais aussi sur le drive gagnant. La figure 5.9 décrit cette architecture PerAc motivée

(cf. [Gaussier 1994]).

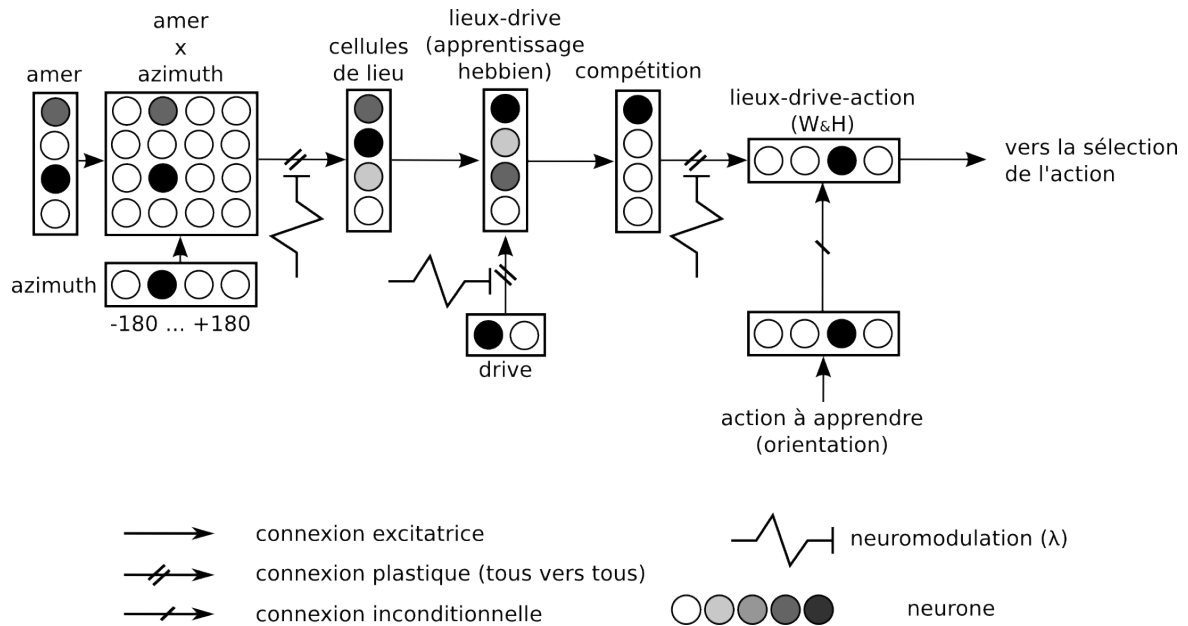


FIGURE 5.9 – Architecture PerAc motivée. Un conditionnement peut être appris entre les lieux et les drives. Cela permet de sur-activer le niveau de reconnaissance des lieux qui sont liés au drive gagnant. Le lieu le mieux reconnu dépend donc des besoins du robot. L'action exécutée permet ainsi au robot de se rapprocher du but qui satisfait son drive gagnant.

4.2 Motivations et intégration de chemin

L'intégration de chemin décrite précédemment peut aussi permettre de rejoindre un lieu but. Mais pour être autonome (du point de vue de sa survie), le robot doit être capable de rejoindre à intervalles réguliers les différents lieux contenant ses ressources vitales. Du point de vue du réseau de neurones utilisé, cela implique un certain nombre de modifications [Hasson 2010d]. Puisqu'un champs d'intégration ne peut stocker qu'un seul vecteur de mouvement global, il faut pouvoir gérer plusieurs champs d'intégration en parallèle. Le principe de base de l'intégration de chemin multi-buts repose sur le fait qu'à chaque itération, le champs vitesse direction est cumulé simultanément sur un réservoir de champs d'intégration. Cette idée pourrait correspondre à ce que les insectes comme les fourmis sont capable de faire [Wehner 2003]. Le comportement observé pourrait ressembler à de la planification car l'animal semble capable de se rappeler les chemins permettant de revenir vers plusieurs lieux but en procédant à une sélection opportuniste. La figure 5.10 illustre cette idée. Cependant, la gestion de plusieurs champs d'intégration en parallèle nécessite un certain nombre de capacités supplémentaires. Il faut pouvoir recruter un champs lorsqu'un nouveau but est détecté (reset de recrutement). Le nombre de mémoires dynamiques (champs) disponible devrait être inférieur aux nombre de buts que le robot peut rencontrer. Il doit procéder à une affectation dynamique des champs aux drives. Il faut donc pouvoir distinguer

un nouveau but d'un but connu et possédant déjà son propre champs d'intégration (auquel cas un reset de reconnaissance permet d'éliminer les erreurs résiduelles). Enfin, il faut pouvoir sélectionner le champs qui correspond au but qui satisfera le mieux les besoins du robot (e.g. de l'eau quand il a soif, et si possible la plus proche). Les trois sous-sections suivantes décrivent en détails ces aspects de l'architecture multi-buts présentée dans la quatrième sous-section.

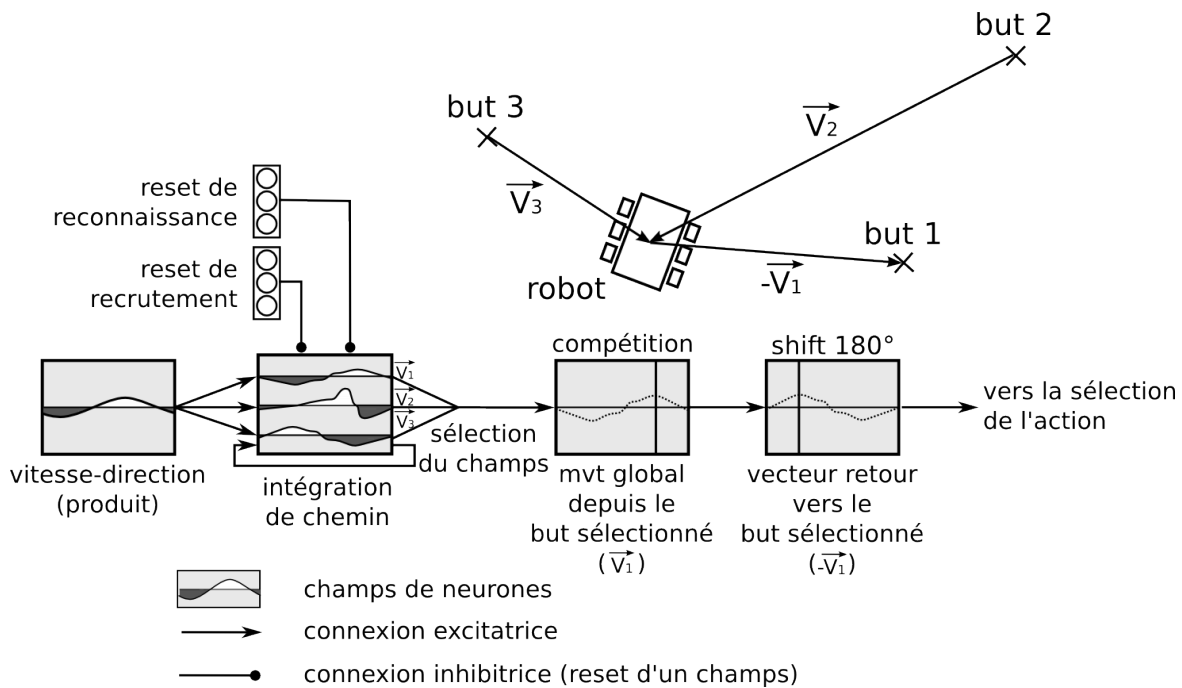


FIGURE 5.10 – Intégration de plusieurs champs en parallèle permettant de retourner à plusieurs buts. Pour pouvoir tirer parti de cette architecture, il est nécessaire de distinguer un but connu d'un but nouveau (remettre à zéro ou recruter un nouveau champs) ainsi que de pouvoir sélectionner le champs d'intégration correspondant au but désiré. Ces différents aspects seront décrits en détails par la suite.

4.2.1 Différencier un nouveau but d'un but connu

Puisqu'il faut pouvoir recruter un nouveau champs d'intégration quand un nouveau but est détecté et simplement le remettre à zéro s'il s'agit d'un but connu, il est primordiale pour le robot de savoir distinguer nouveaux buts et buts connus. Fort heureusement, les champs d'intégration contiennent l'information permettant de faire cette distinction. Plus un but est lointain et plus les neurones du champs associé auront une activité forte. Au contraire, le champs correspondant à un but très proche aura une très faible activité. La figure 5.11 représente le réseau de neurones utilisé pour assurer la distinction entre buts nouveaux et buts connus.

La quantité d'activité dans le champs d'intégration dépend donc de la distance au but. Si cette activité est utilisée pour inhiber un neurone recevant aussi un biais

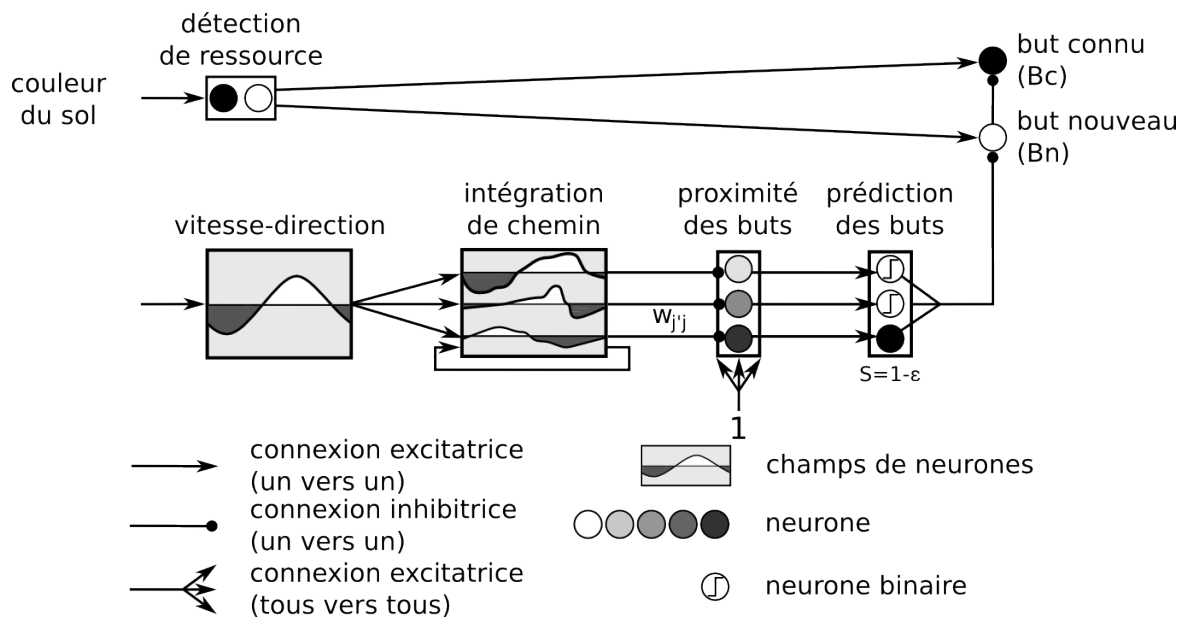


FIGURE 5.11 – Distinguer un but connu d'un but nouveau. A l'approche d'un but connu, le robot peut utiliser le fait que le champs d'intégration correspondant doit être quasiment nul. En seuillant l'activité de proximité des buts, le robot peut se forger une attente de détection d'un but. La détection sera alors interprétée comme un but connu plutôt que comme un nouveau but.

activateur (une entrée constante), son activité sera d'autant plus forte que le but correspondant est proche. Ce neurone contient donc l'information de proximité du but. Il est donc nécessaire de disposer d'un neurone de proximité pour chaque champs d'intégration. Les neurones binaires de prédiction des buts seuillent ces activités de proximité à une valeur proche de 1 (seuil de type $1 - \varepsilon$). Une réponse d'un neurone de proximité ne produira une activité de prédiction du but que si le but correspondant est très proche. Si ε est choisi trop petit, le robot risque cependant de ne pas reconnaître un but connu simplement parce que l'emplacement d'un but n'est pas ponctuel. Il existe alors la possibilité que le robot détecte le but (quel que soit le capteur utilisé) tandis que l'activité de son champs d'intégration correspondant n'est pas encore nulle. Quand le champs n'inhibe plus suffisamment le neurone de proximité et que le seuil est franchi, l'activité résultante peut être interprétée comme une attente ou une prédiction de l'imminence du but s'il n'est pas encore détecté, et comme une confirmation dans le cas contraire. La détection d'une ressource active aussi bien les neurones représentant un nouveau but (B_n) qu'un but connu (B_c). Cependant, un but nouveau inhibe un but connu. Donc par défaut, une ressource détectée sera perçue comme un nouveau but (B_n actif). Le potentiel de prédiction de but inhibe B_n et provoque ainsi la reconnaissance d'un but (B_c actif). L'équation 5.7 montre comment sont calculés ces prédictions-but $P_j^B(t)$:

$$P_j^B(t) = \begin{cases} 1 & \text{si } (1 - w_{j'j} \sum_{i=1}^{i=n} |IC_{ij'}(t)|) > S \\ 0 & \text{sinon} \end{cases} \quad (5.7)$$

Le 1 auquel est soustrait le reste de la formule est l'activité amenée par le biais. $w_{j'j}$ est le poids de la connexion entre les neurones du champs j' et le neurone de prédiction-but j . j et $j' \in [1; nb_{but}]$. n est la taille des champs d'intégration. $IC_{ij'}(t)$ est le potentiel du neurone i du champs d'intégration j' à l'instant t . S est un seuil de la forme $(1 - \varepsilon)$.

Les équations 5.8 et 5.9 détaillent le calcul de B_n et B_c , les potentiels des neurones représentant la détection d'un but nouveau et d'un but connu :

$$B_n(t) = \sum_{i=1}^{i=k} R_k - \sum_{i=1}^{i=j} P_j^B(t) \quad (5.8)$$

$$B_c(t) = \sum_{i=1}^{i=k} R_k - B_n(t) \quad (5.9)$$

R_k vaut 1 quand la ressource k est détectée et 0 sinon. Dans notre exemple ainsi que le reste de ces travaux, k vaut toujours 2 (eau et nourriture). Les poids des liens sont négligés puisqu'ils sont tous fixés à 1. Une limitation de ce système est la difficulté à prendre en compte des buts très proches les uns des autres.

4.2.2 Recrutement de champs d'intégration et reconnaissance de but

Quand un nouveau but est détecté, un champs d'intégration doit être recruté. Ceci est réalisé simplement en faisant le reset (mise à zéro du potentiel des neurones) d'un des champs d'intégration. La partie supérieure de la figure 5.12 montre comment s'opère cette sélection. L'idée consiste à utiliser un champs inutilisé ou du moins le champs le moins utilisé (correspondant donc à l'information en mémoire la moins précise). Le groupe de neurone représentant l'utilisation des buts (taille = nb_{but}) reçoit des connexions excitatrices depuis les groupes faisant les reset de recrutement et de reconnaissance (même taille) et possède des connexions récurrentes ayant un poids légèrement inférieur à 1. Ainsi, quand un but est recruté ou reconnu, le neurone correspondant dans le groupe d'utilisation des buts est excité et son activité monte à 1. Elle décroît ensuite lentement du fait de la connexion récurrente. Mais afin de pouvoir faire une compétition sur ces neurones pour sélectionner un champs libre ou pas utilisé depuis longtemps, il faut utiliser leurs compléments par rapport à 1 puisque qu'un champs disponible aura un potentiel nul et un champs peu utilisé un potentiel faible. Il faut donc utiliser leur activité pour inhiber un neurone recevant un biais excitateur. Le vainqueur de la compétition sur ces compléments correspond donc au champs qui est le moins utilisé et donc le plus préférable à recruter. Le reset de recrutement n'est actif que quand il reçoit une activation depuis la détection d'un nouveau but en plus de l'activité du gagnant de la compétition des buts les moins utilisés. Les deux entrées ont une pondération de 0,5 et les neurones sont seuillés à $0,5 + \varepsilon$. Ce seuillage fonctionne donc comme un opérateur "ET". Un seul neurone du groupe produisant les

resets de recrutement peut être actif à la fois et seulement quand un nouveau but est détecté. Chacun de ces neurones permet d'inhiber l'intégralité du champs d'intégration correspondant.

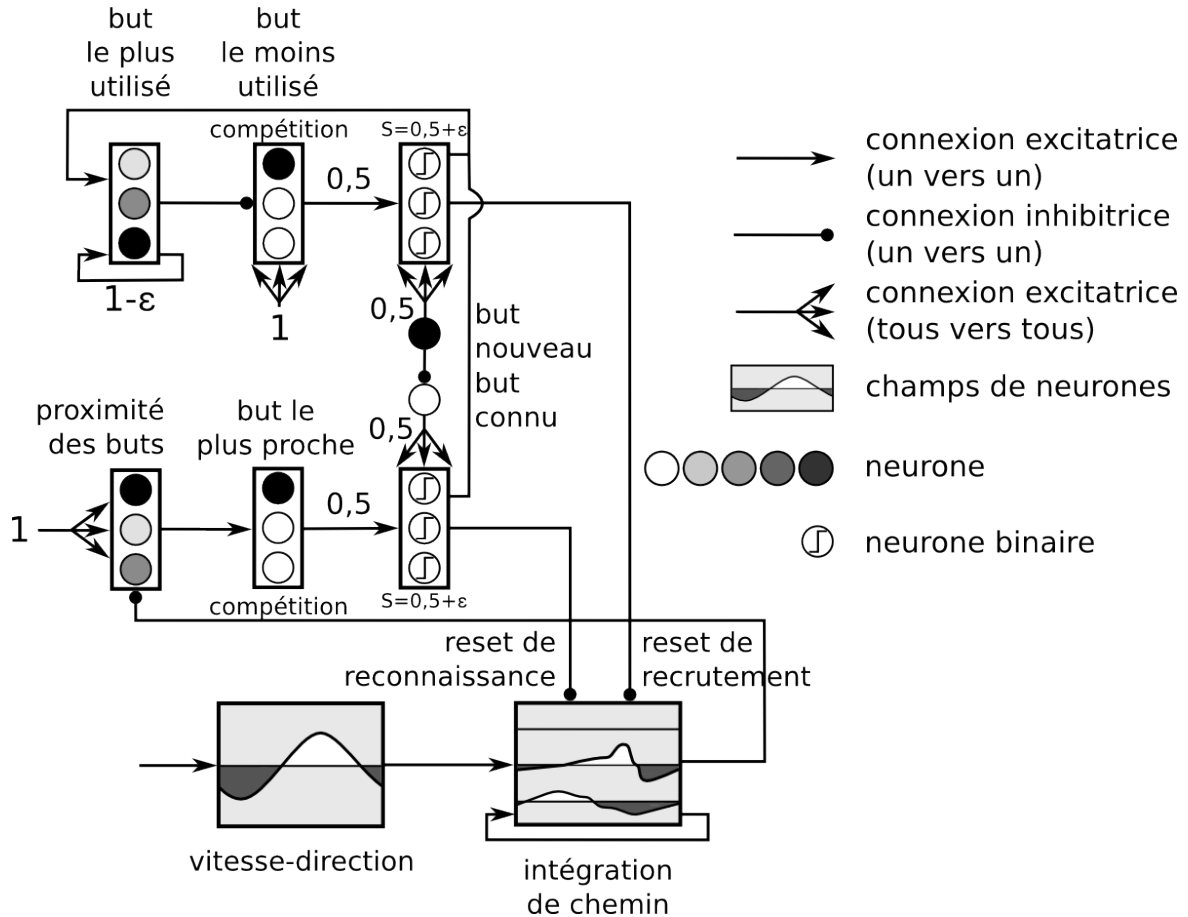


FIGURE 5.12 – Recrutement de champs d'intégration et reconnaissance de but. Comme le montre cet exemple, quand un nouveau but est détecté, le champs d'intégration correspondant au but le moins utilisé est entièrement inhibé (recrutement). Quand un but connu est détecté, le champs d'intégration correspondant au but connu le plus proche est entièrement inhibé (reconnaissance).

Théoriquement, si l'on fait abstraction des erreurs d'intégration venant de la discrétisation angulaire (taille des champs utilisés) et des possibles glissements du robot sur le sol lors de ses déplacements, de retour sur un but connu, le champs correspondant devrait avoir une activité nulle. Le reset est utile pour éviter l'effet cumulatif de cette source d'erreur. Il fonctionne de la même manière que l'effet de recalibration présenté dans [Arleo 2001]. De plus, comme les ressources ne sont pas représentées par des points mais par des surfaces, le robot peut détecter une ressource depuis une position légèrement différente du précédent point de reset, ce qui laisse une activité résiduelle sur le champs d'intégration. La partie inférieure de la figure 5.12 (au dessus des champs de neurones) montre comment est sélectionné le champs d'intégration à

inhiber suite à la reconnaissance d'un but. Le but le plus proche est sélectionné par une compétition sur la proximité des différents buts. Seul le gagnant de la compétition est actif. Le reset de reconnaissance fonctionne de la même manière que le reset de recrutement. Ses neurones sont seuillés (fonction heavyside) à $0,5 + \varepsilon$ et chacune de leurs deux entrées a une pondération à 0,5. Ils fonctionnent eux aussi comme un opérateur "ET". Le reset de reconnaissance inhibe l'intégralité du champs d'intégration correspondant.

4.2.3 Sélection du champs d'intégration approprié

Une fois que plusieurs champs d'intégration correspondant à plusieurs buts ont été recrutés, il est capital de pouvoir sélectionner le champs lié au but désiré. Si le robot a soif, il doit pouvoir sélectionner parmi ses champs d'intégration celui qui correspond à la source d'eau la plus proche. Cette architecture repose donc sur la capacité du robot à apprendre l'association entre les buts (et leur champs d'intégration correspondant) et le drive qu'ils satisfont. Comme une même ressource peut être présente à plusieurs endroits dans l'environnement, mieux vaut sélectionner la plus proche. La figure 5.13 représente le réseau de neurones utilisé.

Le groupe faisant l'association but-drive reçoit des connexions de un vers un depuis les groupes de reset de recrutement et de reconnaissance. Il reçoit aussi des connexions plastiques de tous vers tous depuis le groupe du drive exprimé. Quand une ressource est détectée, les connexions plastiques apprennent à adapter leur poids de tel sorte que l'association entre cette ressource et le drive correspondant est apprise. Ensuite, quand un drive est actif, il produit un potentiel dans les neurones correspondant aux buts qui satisfont ce drive. A chaque fois qu'un but est détecté, l'association but-drive correspondante est renforcée. Suivant la règle d'apprentissage hebbien, les équations 5.10 et 5.11 montre comment est calculée l'activité de chaque neurone but-drive $BD_i(t)$ et comment les poids des connexions plastiques (w_{ji}) sont mis à jour :

$$BD_i(t) = \sum_1^j D_j(t) \cdot w_{ji}(t) + (recrutR_i(t) + recoR_i(t)) \quad (5.10)$$

adaptation des poids :

$$\Delta w_{ji}(t) = \lambda(t) \cdot (recrutR_i(t) + recoR_i(t)) \cdot D_j(t) \lambda(t) = recrutR_i(t) + recoR_i(t) + \varepsilon \quad (5.11)$$

i est le nombre de buts mémorisables, j est le nombre de drives, $D_j(t)$ le neurone j du groupe "drive à associer au but", $w_{ji}(t)$ est le poids de la connexion entre D_j et BD_i , $\lambda(t)$ est le taux d'apprentissage (contrôlé par la neuromodulation), $recrutR_i(t)$ et $recoR_i(t)$ sont les signaux de recrutement et de reconnaissance pour le but i . La neuromodulation de cet apprentissage hebbien est forte quand un but est détecté afin d'apprendre rapidement les associations but-drive. Par contre, elle est très faible quand Le groupe de neurones codant pour le but motivé le plus proche reçoit des connexion excitatrices de un vers un depuis l'association but-drive ainsi que du groupe de la proximité des buts. Grâce à un mécanisme de compétition, il sélectionne un but de

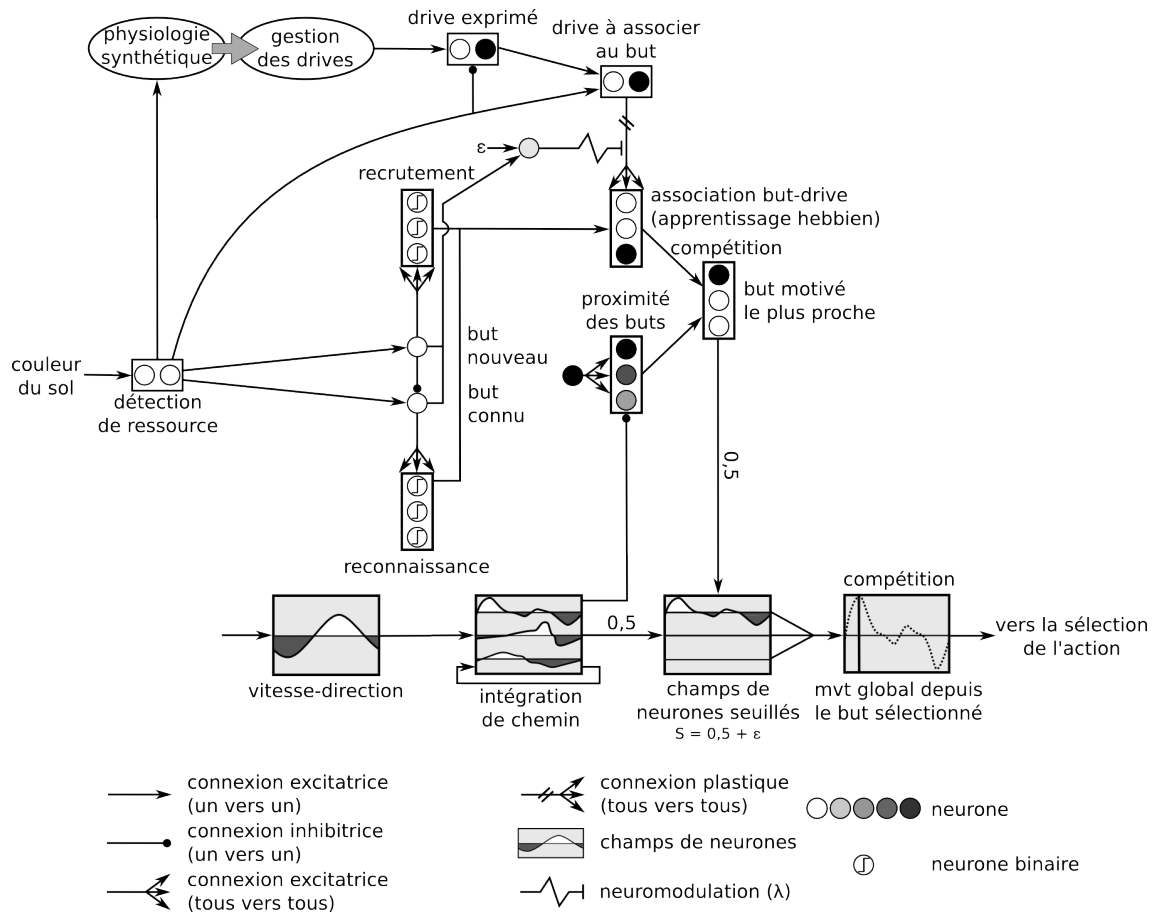


FIGURE 5.13 – Compétition entre les buts et sélection du champs d'intégration. L'association but-drive est apprise grâce à un apprentissage hebbien. Une compétition est menée sur un groupe de neurones recevant les activités venant de cette association but-drive ainsi que du groupe représentant la proximité des buts. Ainsi, c'est le but le plus proche satisfaisant le drive exprimé qui est sélectionné. Le champs d'intégration correspondant est sélectionné par produit matriciel.

telle sorte qu'il satisfasse le drive exprimé, mais aussi qu'il soit le plus proche. C'est donc la conjonction de ces deux informations qui sélectionne le but le plus proche associé au drive exprimé. La sélection du champs d'intégration à utiliser se fait par produit matriciel entre le groupe du but sélectionné et les champs d'intégration. Ce produit matriciel est obtenu en deux étapes. Les champs de neurones seuillés linéaires récupèrent l'activité des champs d'intégration et chaque champs seuillé reçoit aussi l'activité de son neurone correspondant dans le groupe de sélection du but. Ces entrées ont un poids de 0,5 et le seuil S est de 0,5 lui aussi. Seul le champs correspondant au but sélectionné aura donc une activité. Les champs sont ensuite tous sommés vers un champs simple qui ne fait que recopier le champs sélectionné. Cette sortie peut alors être utilisée pour calculer le vecteur retour et prendre la direction du but.

4.2.4 architecture globale

Le robot n'a besoin de détecter une ressource qu'une seule fois pour être capable de stocker et calculer dynamiquement son vecteur retour. Les actions menant au but n'ont donc pas besoin d'être apprises. Le seul apprentissage nécessaire est l'association but-drive permettant de déterminer quel drive chaque but est capable de satisfaire. Quand le robot n'est pas motivé, il peut explorer son environnement (e.g. aléatoirement). Quand un drive est suffisamment fort pour être exprimé, le robot est alors capable de rejoindre les emplacements des ressources correspondantes déjà découvertes. La figure 5.14 montre l'architecture neuronale complète permettant de réaliser cette intégration de chemin multi-buts.

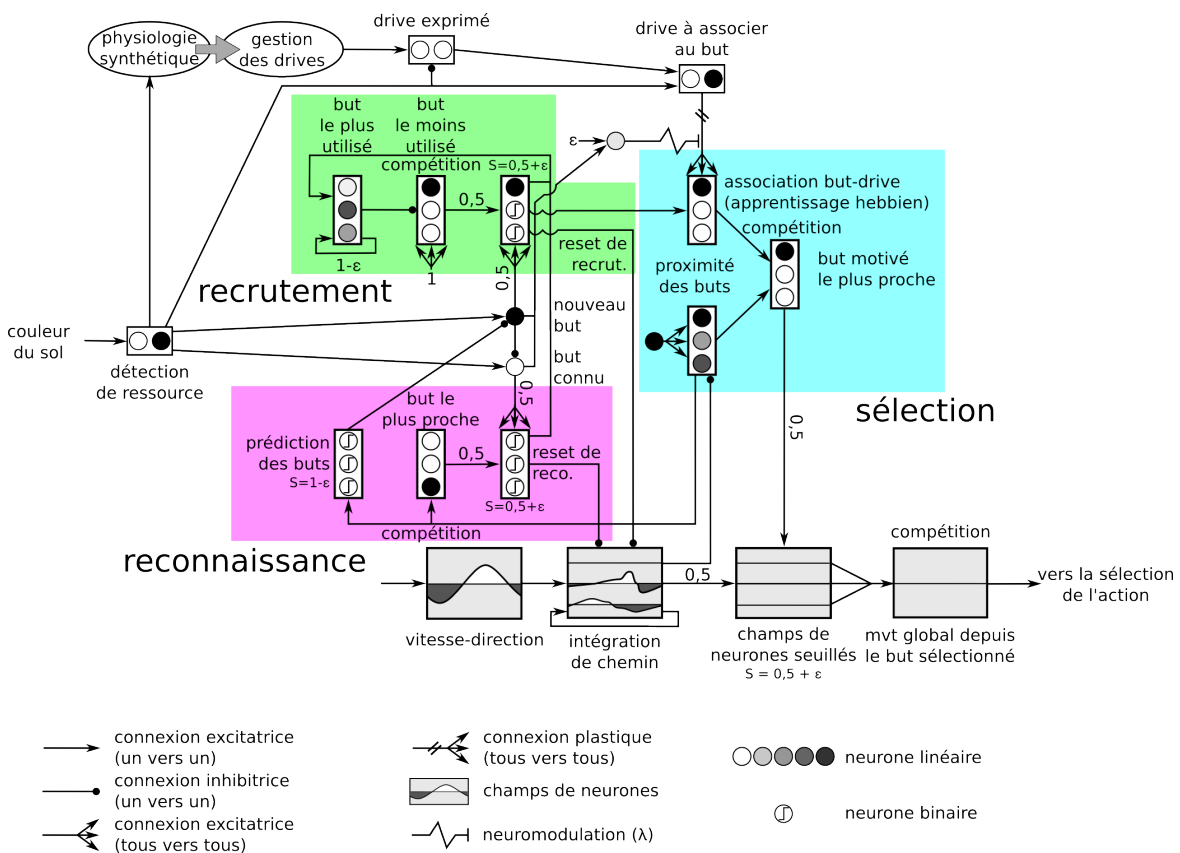


FIGURE 5.14 – Architecture globale pour l'intégration de chemin multi-buts. Trois fonctions principales assurent le fonctionnement de cette architecture gérant plusieurs but en parallèle : le recrutement d'un champs pour un but nouveau, la reconnaissance d'un but déjà connu et la sélection du champs d'intégration qui répond au mieux aux besoins du robot.

Les limitations de cette architecture sont directement liées aux limitations de l'intégration de chemin. Cette méthode de navigation perd en précision avec le temps (chaque déplacement implique une dérive du champs d'intégration). Des études sur les capacités d'intégration de chemin de différents animaux [Etienne 1996] ont montré que ces erreurs cumulatives sont des limitations structurelles. Afin de ne pas trop

dérivée, il faut donc que le robot retourne périodiquement visiter les différents buts. La figure 5.15 montre la trajectoire du robot dans un environnement contenant quatre ressources (deux sources d'eau et deux de nourriture). Lorsque qu'aucun drive n'est actif, le robot navigue aléatoirement. Quand un besoin est suffisamment fort pour qu'un drive s'exprime, le robot se dirige vers la ressource correspondante la plus proche.

Cette architecture de navigation proprioceptive illustre la possibilité d'écrire un algorithme utilisant des mémoires dynamiques et des associations à court terme dans un réseau de neurones non supervisés n'utilisant que des mécanismes élémentaires. C'est pourquoi cette architecture est plausible aux vues de ce que l'on sait du cerveau. Elle illustre le compromis pré-câblage/apprentissage. Tandis que l'architecture de navigation sensori-motrice repose plutôt sur l'apprentissage, l'architecture proprioceptive présentée ici repose essentiellement sur le pré-câblage.

5 Discussion

La douleur et le plaisir sont les sensations fondamentales permettant un ancrage physiologique de la notion de valence émotionnelle. Étant donné leur nature intrinsèque, elles servent de référentiels subjectifs permettant une perception de la nature positive ou négative d'une stimulation ou d'une action. Cette nécessité d'un système de valeur ancré dans la physiologie est un argument majeur de la démarche *animat*. C'est pourquoi nous avons modélisé une physiologie synthétique à partir de laquelle les signaux de douleur et de plaisir trouvent leur origine. Cette construction des signaux de douleur et de plaisir correspond pour le robot à une modélisation simple du rôle de la substance grise périaqueducule et de l'aire tegmentale ventrale. Il ne s'agit pas ici de modéliser finement la génération des signaux de douleur et de plaisir depuis le fonctionnement du corps du robot. Cependant, ces deux structures cérébrales sont à la base des sensations de douleur et de plaisir telles qu'elles sont construites par le cerveau.

Dans ce contexte, l'apprentissage de conditionnements se révèle être la capacité fondamentale permettant de donner une valeur émotionnelle, ou valence, aux différentes perceptions et actions du robot. Les conditionnements avec les sensations de douleur et de plaisir constituent un premier niveau d'informations apprises sur l'interaction du robot avec son environnement. C'est pourquoi nous parlons de signaux émotionnels de bas niveau. L'amygdale et le nucleus accumbens sont des structures clés dans le fonctionnement de ce premier niveau d'apprentissage émotionnel. L'amygdale permet l'apprentissage du conditionnement des stimuli avec les sensations de douleur et de plaisir. La reconnaissance de ces stimuli déclenche ainsi des signaux positifs ou négatifs conditionnés (appris). Le nucleus accumbens quant à lui est impliqué dans la représentation du potentiel de récompense des actions possibles et permet lui aussi d'orienter le comportement. Le couple "amygdale - nucleus accumbens" assure donc la valuation interne des différentes perceptions et actions. Des modèles de ces deux structures sont donc utilisés par le robot pour assurer cet apprentissage de la valuation des informations perceptives et motrices. Il est important de comprendre que l'amygdale n'est pas strictement consacrée au conditionnement à la douleur et le nucleus accumbens

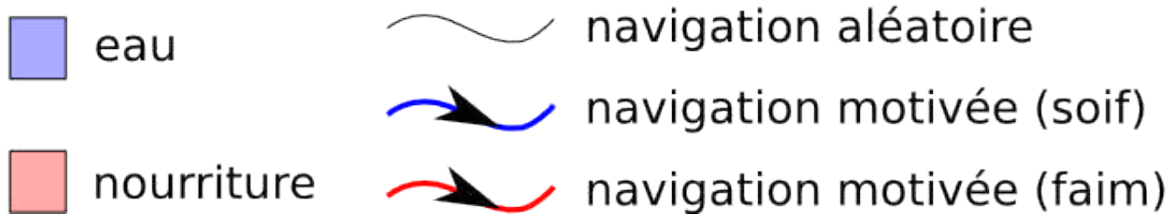
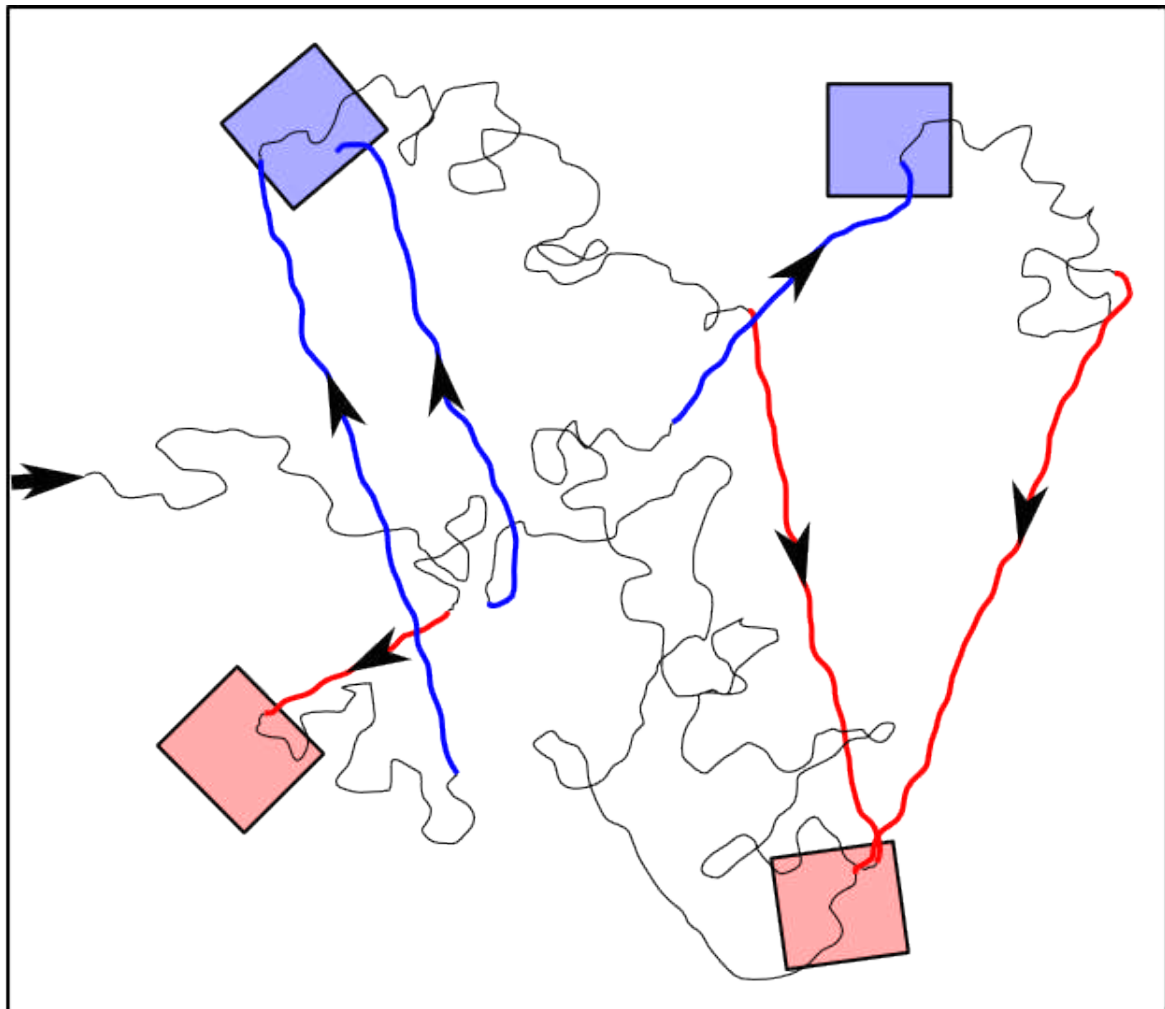


FIGURE 5.15 – Après avoir navigué aléatoirement dans l’environnement afin de détecter au moins une fois chaque ressource, le robot alterne entre navigation aléatoire quand il n’est pas motivé et navigation dirigée vers le but le plus proche satisfaisant son drive exprimé. L’environnement fait environ 7m x 7m.

consacré au plaisir. Nous développerons plus en détails au chapitre 8 notre modèle du nucleus accumbens et nous montrerons des exemple d'apprentissage d'évitement avec utilisant le nucleus accumbens. S'il existe un récompense associée à une action qui rapproche d'une source de plaisir, il existe de même une récompense associée à une action qui éloigne d'une source de douleur.

Ce système de valuation interne particulièrement simple permet au robot d'organiser son comportement autours de ce qui l'attire et ce qui le repousse. Nous avons montré comment des conditionnements permettent d'utiliser la navigation visuelle aussi bien que la navigation proprioceptive pour résoudre des problèmes de survie (plusieurs buts). Le plaisir et les motivations qui en découlent sont à la base de l'apprentissage de comportements attractifs (convergents). Notamment, les associations lieu-action apprises par le robot dans notre modèle du nucleus accumbens correspondent à la représentation du potentiel de récompense des différentes actions. A l'opposé, le conditionnement d'une stimulation à la sensation de douleur permet à cette stimulation de provoquer une douleur conditionnée. C'est par exemple ce qui permet au robot d'anticiper la perception effective de douleur. Nous avons illustré ce principe simple par la construction d'une stratégie d'évitement d'obstacle basée sur les conditionnements à la douleur. Notre propos n'était pas d'améliorer les méthodes d'évitement d'obstacle existantes mais bien d'illustrer le potentiel d'organisation du comportement que permet un simple modèle de conditionnement à la douleur.

Dans le prochain chapitre, nous montrerons comment l'analyse de l'évolution des différents signaux conditionnés permet au robot de construire des capacités d'auto-régulation de son comportement inspiré par le rôle du cortex préfrontal.

Automonitoring et frustration

1 Introduction

Jusqu'à présent, nous avons décrit comment une construction autonome des comportements les plus élémentaires peut être rendue possible par l'apprentissage de conditionnements avec des sensations de douleur et de plaisir liées à la physiologie simulée du robot. Cependant, maintenir dans le temps un comportement efficace dans un environnement dynamique et complexe reste une tâche difficile. Par opposition aux environnements simulés dont tous les paramètres peuvent être contrôlés, l'utilisation d'un robot autonome dans un environnement réel présente des défis supplémentaires. Dans un environnement réel, les conditions dans lesquelles le robot évolue sont susceptibles de changer et d'affecter le comportement du robot. Par exemple, les conditions d'éclairage peuvent changer au long de la journée, affectant ainsi la perception visuelle de l'environnement. Les conditionnements visuels permettant la navigation sensori-motrice peuvent alors être incapables d'assurer des comportements efficaces (convergence, évitement). Un lieu but (l'emplacement d'une ressource simulée) auparavant accessible peut être obstrué par un obstacle. Une ressource peut aussi être retirée de l'environnement afin de simuler son épuisement. Ces changements de conditions environnementales peuvent produire des situations où les comportements appris par le robot ne sont pas suffisants pour lui permettre de satisfaire ses motivations. En particulier, le robot peut être bloqué dans un cycle d'actions contre-productives : une alternance d'évitements et de rapprochements de l'obstacle recouvrant le but. C'est la situation de *deadlock*. Afin de gagner en autonomie et ainsi de pouvoir réagir aux situations où ce qu'il a appris ne lui permet pas de satisfaire ses objectifs, le robot nécessite la capacité à évaluer ses comportements. Traditionnellement, cette capacité d'auto-évaluation fait défaut aux robots qui ne tirent pas de satisfaction de leurs actions "productives" ni ne sont frustrés par leurs actions "improductives". Bien sûr, il existe des techniques d'apprentissage permettant aux robots de produire de manière autonome des signaux de récompense et de punition (apprentissage par renforcement). C'est le cas par exemple du Q-Learning [Watkins 1992] ou du TD(λ) [Sutton 1990]. Cependant, l'auto-évaluation dont il est question ici ne correspond pas strictement à l'utilisation de ces méthodes. Il y a une différence entre la capacité à apprendre un comportement et la capacité à évaluer l'efficacité de ce comportement. C'est pourquoi les robots font généralement preuve d'une grande rigidité comportementale lorsqu'ils sont confrontés à des situations imprévues.

Le propos de ce chapitre est de décrire un système générique d'auto-régulation du comportement basé sur l'auto-évaluation de l'efficacité comportementale [Hasson 2010c]. Cette auto-évaluation est rendue possible par la construction et l'utilisation de signaux

de prédictions des buts. L'objectif de ces signaux de prédiction est de permettre au robot une évaluation endogène et autonome de sa progression par rapport à ses buts. Ces prédictions peuvent être construites à partir des différentes informations sensorielles à la disposition du robot, notamment la vision et la proprioception. Le détail de ces prédictions est donné plus loin mais ce qui est capital, c'est la notion d'information endogène. Une prédiction de but basée sur la vision consiste en un conditionnement d'une cellule de lieu visuelle à la sensation de plaisir provoqué par la consommation de la ressource que ce lieu contient. Les variations du niveau de reconnaissance de ce lieu but sont répercutées sur le niveau de la réponse de plaisir conditionné. Une prédiction de but basée sur la proprioception repose quant à elle sur l'information de distance codée dans les champs d'intégration des mouvements. L'activité maximale d'un champ d'intégration représente la distance au but associé à ce champ. A mesure que le robot se rapproche du lieu but, l'activité maximale du champ d'intégration associé diminue. Bien que ces deux mécanismes de prédiction des buts fonctionnent dans des sens opposés (i.e. augmentation versus diminution d'un potentiel pour traduire la progression vers le but), l'information produite est la même. Ce qui compte, ce n'est pas le niveau absolu d'une prédiction visuelle ou proprioceptive du but, mais la manière dont ces prédictions évoluent dans le temps.

Notre système d'auto-évaluation comportementale repose donc sur l'analyse de l'évolution temporelle des différentes prédictions des buts. Un comportement inefficace peut ainsi être détecté comme étant une incapacité prolongée à réduire la distance au but. Nous avons développé un modèle simple de régulation des stratégies, des buts et des motivations du robot inspiré du rôle inhibiteur du cortex préfrontal [Shimamura 2000, Rubia 2003, Knight 1999, Constantinidis 2002, Dehaene 1989, Dehaene 1995]. Ce modèle consiste à reproduire la fonction inhibitrice du cortex préfrontal permettant de changer de stratégie, de but ou de motivation courante. Cette inhibition déclenchée par l'inefficacité du comportement du robot peut être apparentée à une réaction de frustration. Elle permet au robot de contrôler son comportement de manière plus souple et plus adaptée aux conditions de l'environnement.

Dans un premier temps, nous décrirons le fonctionnement des prédictions de but basées sur la vision et la proprioception. Nous détaillerons ensuite le mode d'action du système de régulation à proprement parler : détection des situations d'échec et signal d'inhibition sur la stratégie, le but ou la motivation. Enfin, avant de conclure, nous présenterons les résultats expérimentaux obtenus grâce à ce mécanisme de frustration.

2 Prédictions des buts

La prédiction des buts est une compétence capitale pour l'autonomie comportementale du robot. Dans le contexte de nos travaux, la prédiction d'un but est une évaluation de la distance d'un but construite à partir d'informations sensorielles. L'intérêt de ces prédictions ne réside pas réellement dans l'évaluation absolue de la distance d'un but mais plutôt dans la capacité du robot à déterminer s'il se rapproche ou s'éloigne d'un but appris. C'est à partir de l'évolution de ces prédictions que le robot peut évaluer la qualité de son comportement. La prédiction des buts peut être construite à partir

des différentes perceptions à sa disposition. Dans notre cas, nous pouvons construire deux types de prédictions de but basées respectivement sur les informations visuelles et proprioceptives. Ces deux sources de sensations sont suffisamment riches et adaptées à l'environnement pour permettre au robot de s'y référer et de se situer relativement à ses buts.

2.1 Prédiction visuelle du but

La prédiction visuelle d'un but repose sur l'apprentissage d'un conditionnement entre les cellules de lieux visuelles et la sensation de plaisir liée à la consommation d'une ressource. La figure 6.1 décrit le fonctionnement de cette prédiction. Le neurone de prédiction visuelle du but reçoit des connexions plastiques depuis les cellules de lieux et une connexion inconditionnelle depuis le neurone transmettant le signal de plaisir. De plus, le signal de plaisir sert aussi à contrôler la neuromodulation de l'apprentissage des connexions plastiques. Quand un signal de plaisir est perçu, le neurone de prédiction apprend à renforcer ses connexions avec les cellules de lieux. Etant donné que les variations de poids synaptiques sont proportionnelles aux potentiels d'entrée (i.e. l'activité des cellules de lieux), cette modification de poids est d'autant plus forte que les cellules de lieux correspondent à des lieux proches du but. Un seuillage ou une normalisation de l'activité des cellules de lieux permet d'isoler mieux encore les lieux les plus proches du but. Un conditionnement entre la sensation de plaisir et les lieux servant de base à la construction du bassin d'attraction autour du but permet d'apprendre une évaluation de la distance au but. Lorsque le robot se rapproche du but, les cellules de lieux correspondant au bassin d'attraction sont de plus en plus actives et le plaisir conditionné est de plus en plus important. Plus ce plaisir conditionné est important et plus le but est proche.

Une autre manière de considérer la prédiction visuelle du but consiste à recruter une cellule de lieu pour caractériser directement le lieu-but. Nous ne voulons cependant pas que la reconnaissance du lieu-but entre en compétition avec les lieux-action permettant la navigation visuelle. Les lieux-but peuvent donc être appris de la même manière que les lieux associés aux actions (catégorisation du produit tensoriel amer-azimuth), mais dans un groupe de neurones distinct. Il suffit alors de monitorer directement le niveau de reconnaissance de ces lieux-but pour évaluer la distance des différents buts. Ce type de prédiction visuelle du but sera décrit un peu plus loin lors de la description de l'inhibition d'un but visuel.

2.2 Prédiction proprioceptive du but

La prédiction proprioceptive d'un but repose sur la propriété de codage de l'information de distance des champs d'intégration des mouvements. Comme nous l'avons décrit dans le chapitre sur les méthodes de navigation bio-inspirée, un champ d'intégration des mouvements permet de coder deux informations capitales : la direction et la distance d'un point de référence (i.e. d'un but). L'activité maximale d'un champ d'intégration est proportionnelle à la distance du point de reset. Elle est donc un bon indicateur de la distance au but. Cette activité maximale du champs d'intégration peut

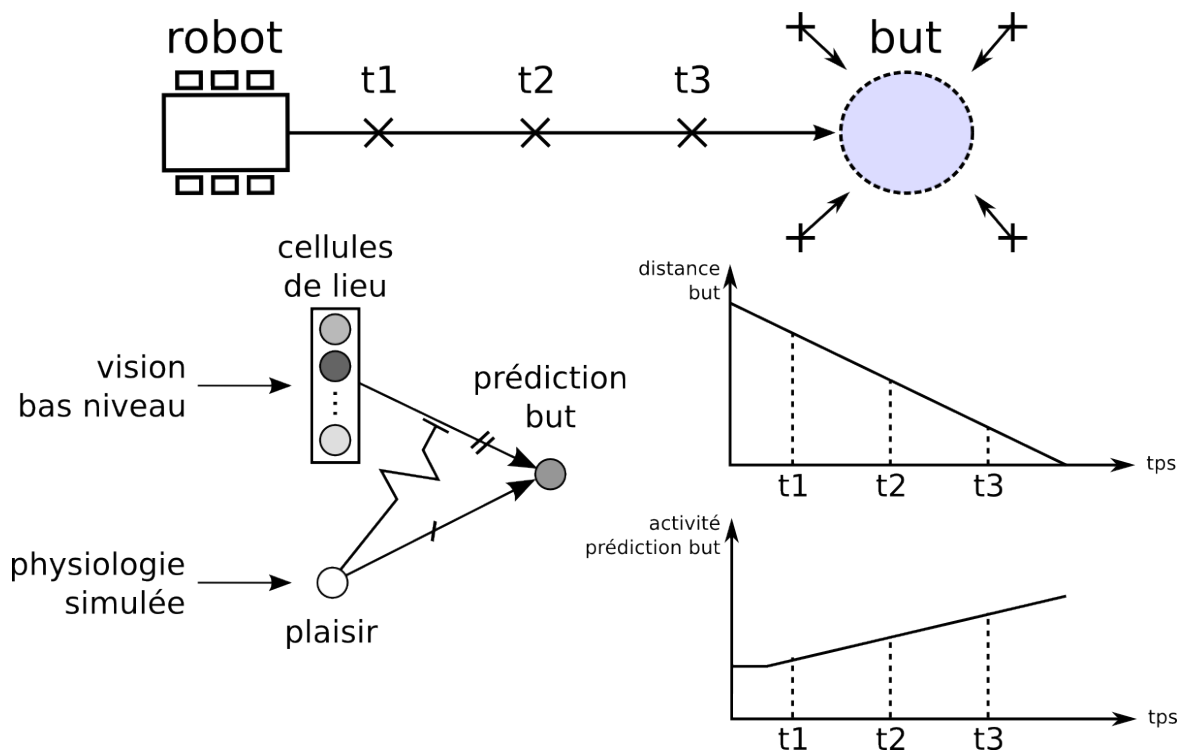


FIGURE 6.1 – Prédiction visuelle du but. Un conditionnement est appris entre la sensation de plaisir et les lieux visuels situés à proximité du but. L'évolution de ce plaisir conditionné en fonction des déplacements du robot représente la proximité du but.

être représentée par l'activité d'un neurone. Une compétition est opérée sur tous les neurones du champ d'intégration. Le résultat est un champ dont tous les neurones ont une activité nulle à l'exception d'un seul neurone dont l'activité est le maximum du champ d'intégration. Tous les neurones de ce champ de compétition projettent ensuite une connexion vers un même neurone qui ne reçoit donc que l'activité du neurone gagnant. Comme le montre la figure 6.2, à mesure que le robot se rapproche du but, l'activité contenue dans le champ d'intégration diminue. Le maximum d'activité diminue jusqu'à devenir nul lorsque le robot atteint le point du précédent reset du champ d'intégration. A la différence de la prédiction visuelle qui code la proximité du but, la prédiction proprioceptive code la distance au but. La différence réside uniquement dans une symétrie de l'interprétation de ce que représente une hausse ou une baisse de cette activité.

3 Détection de l'échec

Notre objectif est de construire un mécanisme permettant de détecter les situations d'échec ou *deadlock*. Un *deadlock* est défini comme une situation dans laquelle le robot produit un comportement inefficace et répétitif. Autrement dit, il s'agit d'une situation où le robot ne progresse pas vers son but. Le monitoring de l'évolution de la prédiction

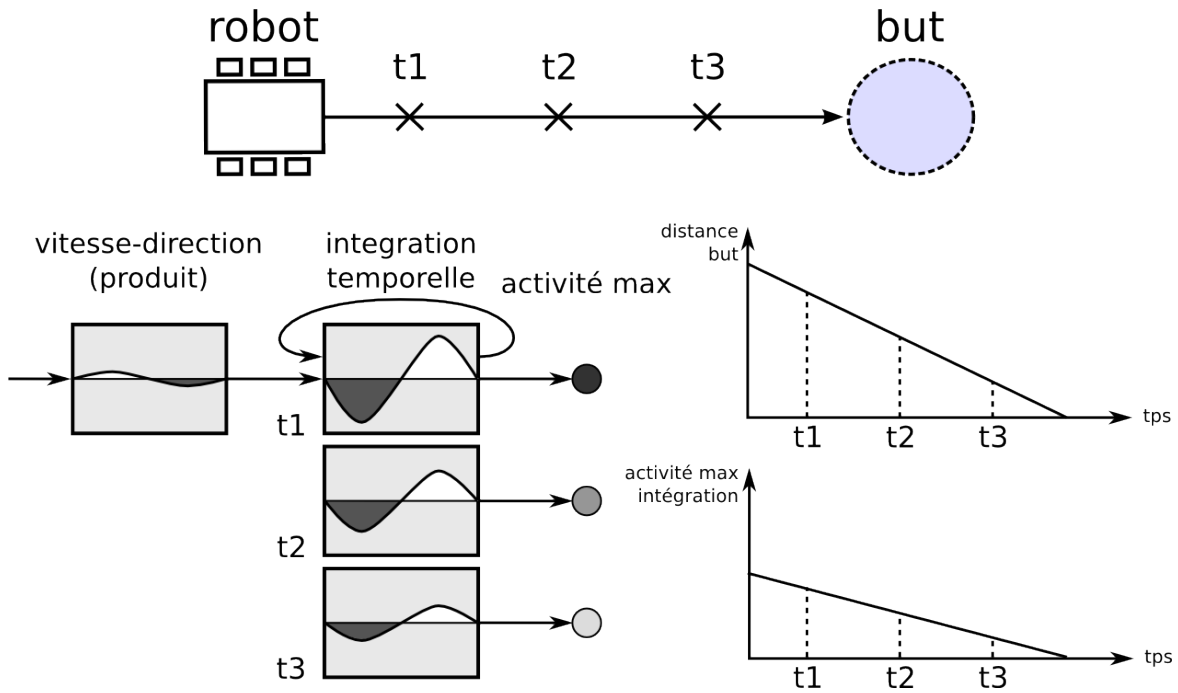


FIGURE 6.2 – Prédiction proprioceptive du but. Le champ d'intégration des mouvements contient l'information de distance par rapport au but.

de la distance au but dans le temps est donc un moyen simple et efficace pour le robot d'évaluer ses progrès. Lorsque le robot a un drive actif, e.g. quand il a faim, avant que de la nourriture ne soit trouvée, tant que la distance prédite de la nourriture diminue dans le temps, le robot peut considérer que tout va bien. Mais si par contre la distance au but ne diminue pas, cela traduit le fait que le comportement du robot est inefficace. Si cette inefficacité perdure, cela signifie que le robot est dans un *deadlock*. La figure 6.3 décrit le réseau de neurone utilisé pour réaliser cette détection de l'échec.

Pour déterminer l'évolution du signal de prédiction du but P_b , les neurones P_i et R_i font la différence entre le niveau de ce signal à l'instant t et à l'instant $t - \Delta t$ (un délai d'une itération sur P_b permet la comparaison de sa valeur à deux instants successifs). P_i soustrait $P_b(t - \Delta t)$ à $P_b(t)$ et correspond donc au progrès instantané vers le but. R_i soustrait $P_b(t)$ à $P_b(t - \Delta t)$ et correspond donc au recul instantané par rapport au but. Les équations 6.1 et 6.2 décrivent le calcul du progrès et du recul instantané :

$$\text{Recul} : R_i(t) = P_b(t) - P_b(t - \Delta t) \quad (6.1)$$

$$\text{Progres} : P_i(t) = P_b(t - \Delta t) - P_b(t) \quad (6.2)$$

Selon cette définition le progrès et l'échec sont des quantités opposées ($P_i(t) = -R_i(t)$). Le principe de notre détection de l'échec consiste à intégrer dans le neurone $f(t)$ une faible quantité ε (habituellement 0,001) à laquelle est sommé le recul instantané $R_i(t)$ et soustrait le progrès instantané $P_i(t)$. Lorsque le robot progresse vers

son but, le progrès instantané est suffisant pour que $f(t)$ n'accumule pas d'activité (voir qu'il diminue un potentiel d'échec déjà accumulé). Au contraire, lorsque le robot s'éloigne de son but, le recul instantané s'additionne avec la constante ε et le potentiel d'échec $f(t)$ croît. La constante ε permet de au robot de finir par considérer comme un échec le fait de ni se rapprocher ni s'éloigner de son but. Le neurone F est un simple neurone binaire recevant l'activité de f . Le seuil S du neurone F définit la tolérance du robot à l'échec. Plus ce seuil est bas et plus le robot se considère rapidement en situation d'échec. La valeur de S doit cependant respecter deux contraintes. Si S est trop bas, le robot se "frustre" très rapidement et le moindre détection d'absence de progrès, aussi brève soit-elle, entraîne un changement de comportement. Si au contraire S est trop grand, le temps que le robot détecte qu'il est en situation d'échec, une de ses variables physiologiques aura atteint un niveau critique et le robot cessera de fonctionner. Une valeur raisonnable de S se situant entre ces deux extrêmes peut aisément être déterminée expérimentalement en fonction de la taille de l'environnement. Une gestion adaptative de ce seuil en fonction du niveau des motivations ou de la fréquence d'apparition du succès et de l'échec est discutée en conclusion. $f(t)$ est inhibé par $F(t-1)$. Cette rétro-inhibition correspond au fait qu'une fois que le signal de détection d'échec est déclenché et que le comportement du robot change, il ne faut pas que les nouveaux comportements soient aussi évalués comme des sources immédiates d'échec. Un potentiel R vient totalement inhiber l'intégration de l'échec lorsque le but est satisfait (afin d'éliminer tout résidu éventuel). L'équation 6.3 décrit l'intégration du potentiel d'échec $f(t)$:

$$f(t) = [f(t - \Delta t) + R_i(t) - P_i(t) + \varepsilon - R - F(t - 1)]^+ \quad (6.3)$$

avec $[x]^+$ vaut x si $x > 0$ et vaut 0 sinon.

$f(t)$ est l'intégration temporelle de l'échec à se rapprocher du but, Δt est la durée entre chaque pas de temps, $R_i(t)$ et $P_i(t)$ sont respectivement le recul et le progrès instantané du robot vers son but, ε est une petite constante (habituellement 0,001), R est un signal de reset valant 1 quand le but est satisfait (quand le ressource nécessaire est détectée) et 0 sinon et $F(t-1)$ est le signal de détection d'échec lors de la précédente itération.

La détection de l'échec à proprement parler est réalisée par le neurone binaire $F(t)$. Si $f(t)$ est supérieur au seuil S , $F(t)$ vaut 1 et 0 sinon.

$$F(t) = \begin{cases} 1 & \text{si } [f(t - \Delta t) + R_i(t) - P_i(t) + \varepsilon - R]^+ > S \\ 0 & \text{sinon} \end{cases} \quad (6.4)$$

Ce mécanisme diffère d'un simple système de *watchdog* parce que c'est l'échec et non simplement le temps qui font monter le potentiel de détection de l'échec. Selon cette conception, résoudre un problème long ne provoque pas de détection d'échec tant que des progrès sont perçus. Ce n'est pas la durée d'une tâche mais bien l'estimation des progrès du robot qui est à la base de la détection de l'échec. C'est ce qui permet à ce mécanisme une plus grande souplesse d'utilisation qu'un simple compteur nécessitant une connaissance a priori de la durée de la tâche à accomplir. La détection de l'échec

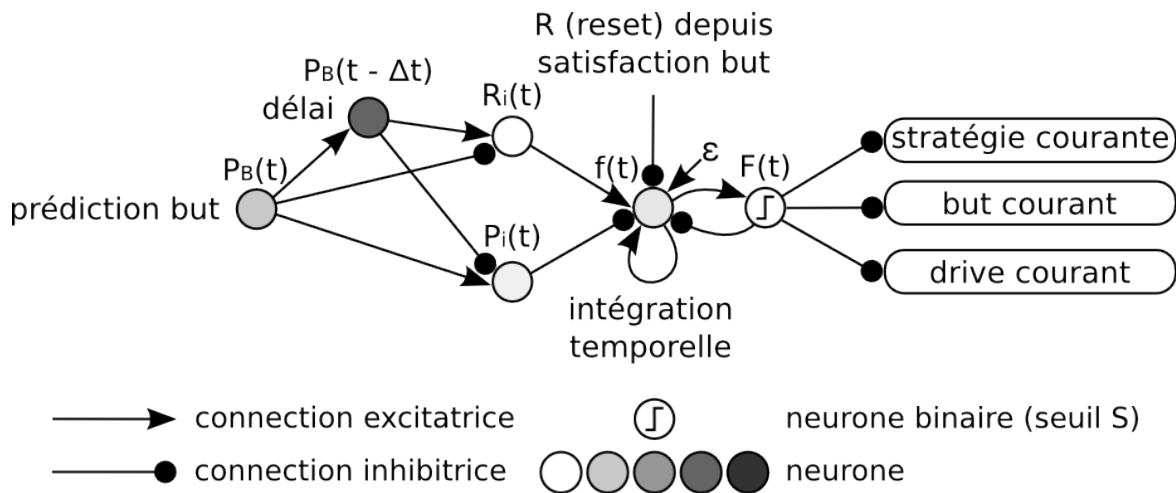


FIGURE 6.3 – Mécanisme de frustration : intégration non linéaire de la dérivée dans le temps de la distance au but. Une petite entrée constante (ε) est sommée avec cette intégration afin d'assurer que bien que la distance au but reste constante, la réaction de frustration est possible. Au delà d'un seuil S , la stratégie, le but ou le drive actif est inhibé. Les dérivées des prédictions visuelles et proprioceptives du but sont calculées en sens opposé. En contrôlant le poids des liens inhibiteurs vers les stratégies, drives et buts, nous pouvons décider quelle inhibition nous voulons voir fonctionner.

donne au robot un moyen d'échapper aux comportements répétitifs et inefficaces. Le moyen le plus simple d'échapper à un *deadlock* est d'utiliser la détection de l'échec pour inhiber le comportement sous-jacent. Mais il existe de nombreux moyens d'altérer le comportement du robot. La détection de l'échec peut par exemple être utilisée pour inhiber la stratégie de navigation courante e.g. en passant de la navigation visuelle à la navigation proprioceptive. Mais cette détection peut aussi inhiber le but actif afin de changer de lieu but tout en gardant le même type de but à satisfaire (e.g. une autre source de nourriture). Enfin, cette détection peut aussi inhiber le drive actif afin de poursuivre un autre but et tenter de satisfaire le drive actif ultérieurement.

4 Inhibition et frustration

Dans une situation d'échec persistant, le comportement produit ne permet plus au robot de réduire la distance au but et entraîne donc une accumulation de détection d'échec (f croît en moyenne). Lorsque f atteint le seuil critique S , un potentiel inhibiteur F est produit. C'est cette inhibition ainsi que le changement de comportement qui y est lié que nous appelons une réaction de frustration. Cependant, l'architecture de contrôle du robot permet un certain nombre de possibilités de régulation. Le signal inhibiteur produit par l'accumulation d'échec peut intervenir au niveau des drives, des buts et des stratégies de navigation du robot. Ces différentes alternatives vont être développées dans la suite de cette section.

Au niveau des drives :

L'inhibition du drive exprimé ne permet de résoudre directement le problème rencontré par le robot. En changeant de drive, le robot change d'objectif et donc de comportement. Alors qu'il cherche à satisfaire ses autres buts, ce qui l'a empêché de rejoindre son but ne sera peut être plus présent quelques minutes plus tard. De plus, en s'acharnant pas sur une ressource inaccessible, le robot peut trouver une nouvelle ressource satisfaisant son besoin frustré. Qu'elle qu'en soit l'effet bénéfique possible, inhiber le drive actif a un impact important sur le comportement. La figure 6.4 décrit le système d'inhibition du drive exprimé.

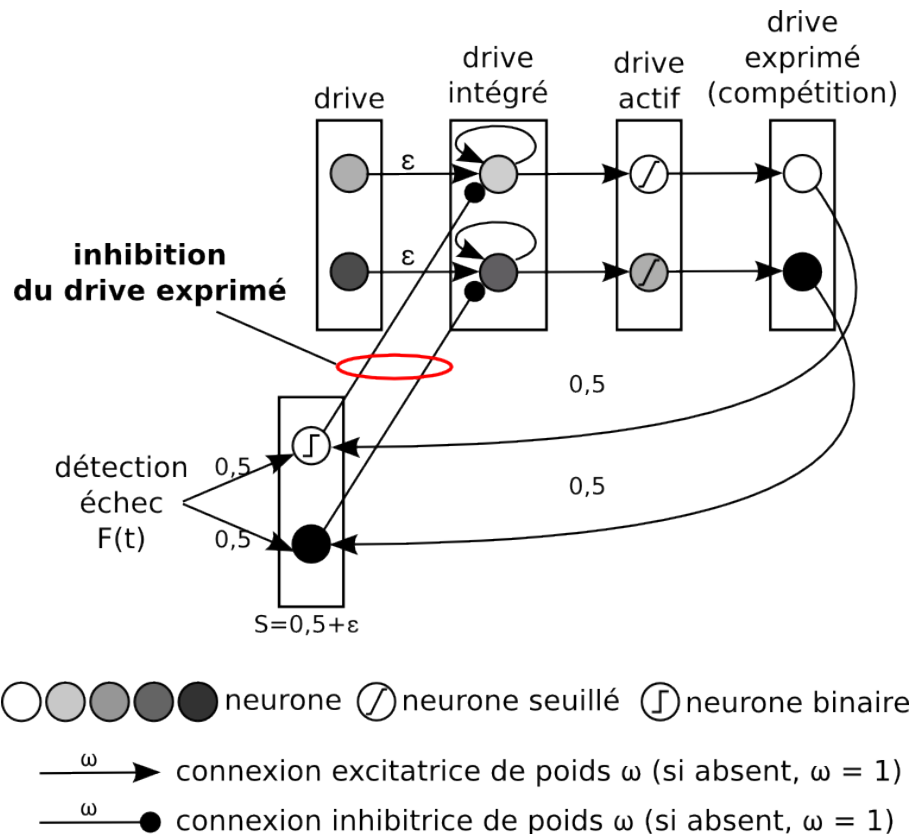


FIGURE 6.4 – Inhibition du drive exprimé : En arrivant au système des drives, le signal d'inhibition est filtré pour n'affecter que le drive exprimé. Après inhibition, le drive frustré perd inévitablement la compétition et le changement de drive entraîne le changement de comportement.

C'est dans ce cadre que l'intérêt du drive intégré prend son sens. Nous avons décrit précédemment le système de drive simulé que nous utilisons. Les drives sont calculés directement sur les niveaux des variables physiologiques et les drives intégrés somment dans le temps la valeur de chaque drive. Nous ne voulons pas inhiber directement les drives puisque ceux-ci reflètent l'état physiologique du robot et cette information doit rester valable. Les drives intégrés permettent donc de recevoir l'inhibition venant du système de frustration sans fausser la physiologie du robot. Le seul élément nécessaire qui n'a pas encore été expliqué est la manière de sélectionner le drive

à inhiber. Pour inhiber le drive exprimé, responsable du comportement source de frustration, il faut simplement moduler le signal d'inhibition par l'activité des drives exprimés (1 si le drive est exprimé et 0 sinon). Le but est d'inhiber uniquement le drive intégré correspondant au drive exprimé. Ce produit logique peut être réalisé pour chaque drive par un neurone binaire prenant deux entrées : le signal d'inhibition (1 ou 0) et l'expression du drive correspondant (1 ou 0). En utilisant des poids synaptiques w à 0,5 pour chaque entrée et un seuil S légèrement supérieur à 0,5 ($0,5+\varepsilon$ (habituellement ε vaut 0,001)), le neurone ne répond que si les deux entrées sont actives simultanément. Quand un signal inhibiteur est envoyé vers les drives, seul le drive exprimé est donc inhibé.

Au niveau des buts :

Qu'il s'agisse de la navigation visuelle ou proprioceptive, le robot est capable d'apprendre à retourner à plusieurs emplacements distincts de son environnement. Alors qu'un but en particulier peut être inaccessible, un autre but peut contenir la même ressource et offrir ainsi une solution alternative. De même qu'il fallait inhiber le drive exprimé causant le comportement inefficace, il faut inhiber le but actif pour poursuivre ensuite le but alternatif.

Pour la stratégie proprioceptive, cette inhibition ne demande pratiquement pas de modification à l'architecture présentée précédemment (navigation proprioceptive multi-buts motivée). La figure 6.5 montre comment le signal d'inhibition est dirigé vers le but actif.

L'objectif est de pouvoir inhiber sélectivement le but actif. Le principe de sélection est proche de celui utilisé pour l'inhibition des drives. L'inhibition est le résultat du ET logique entre le signal de détection d'échec et le signal du but motivé le plus proche. C'est ce dernier signal qui sélectionne le champs d'intégration à utiliser pour naviguer. Cependant, le signal d'inhibition est phasique i.e. il n'est présent qu'au moment de la détection de l'échec. Mais à la différence de l'inhibition des drives, les neurones inhibés ne réalisent pas d'intégration mais une compétition. Sans "mémoire" de l'inhibition, le changement de but ne serait effectif que pour une seule itération puisque le but le plus proche et l'association but-drive ne changent pas. Nous avons donc utilisé une simple mémoire à court terme. Le signal d'inhibition est relayé par des neurones à connexions récurrentes de poids légèrement inférieurs à 1 (de type $1 - \varepsilon$). L'inhibition est ainsi totale après la détection de l'échec et diminue progressivement avec le temps. Le poids de ces connexions récurrentes permet de contrôler la durée de l'inhibition d'un but. Plus le poids est proche de 1 (ε proche de 0) et plus cette inhibition dure longtemps.

En ce qui concerne la navigation visuelle, l'inhibition des buts s'opère plus simplement sur une architecture PerAc légèrement modifiée. Dans sa version habituelle, il n'est pas aisé d'inhiber sélectivement un but. La navigation repose sur la dynamique d'attraction apprise par les lieux-action et le but n'est pas codé explicitement. C'est pourquoi nous utiliserons ici un modèle PerAc avec conditionnement des lieux-action par les lieux-but. La figure 6.6 décrit cette architecture. Les lieux-but sont appris de la

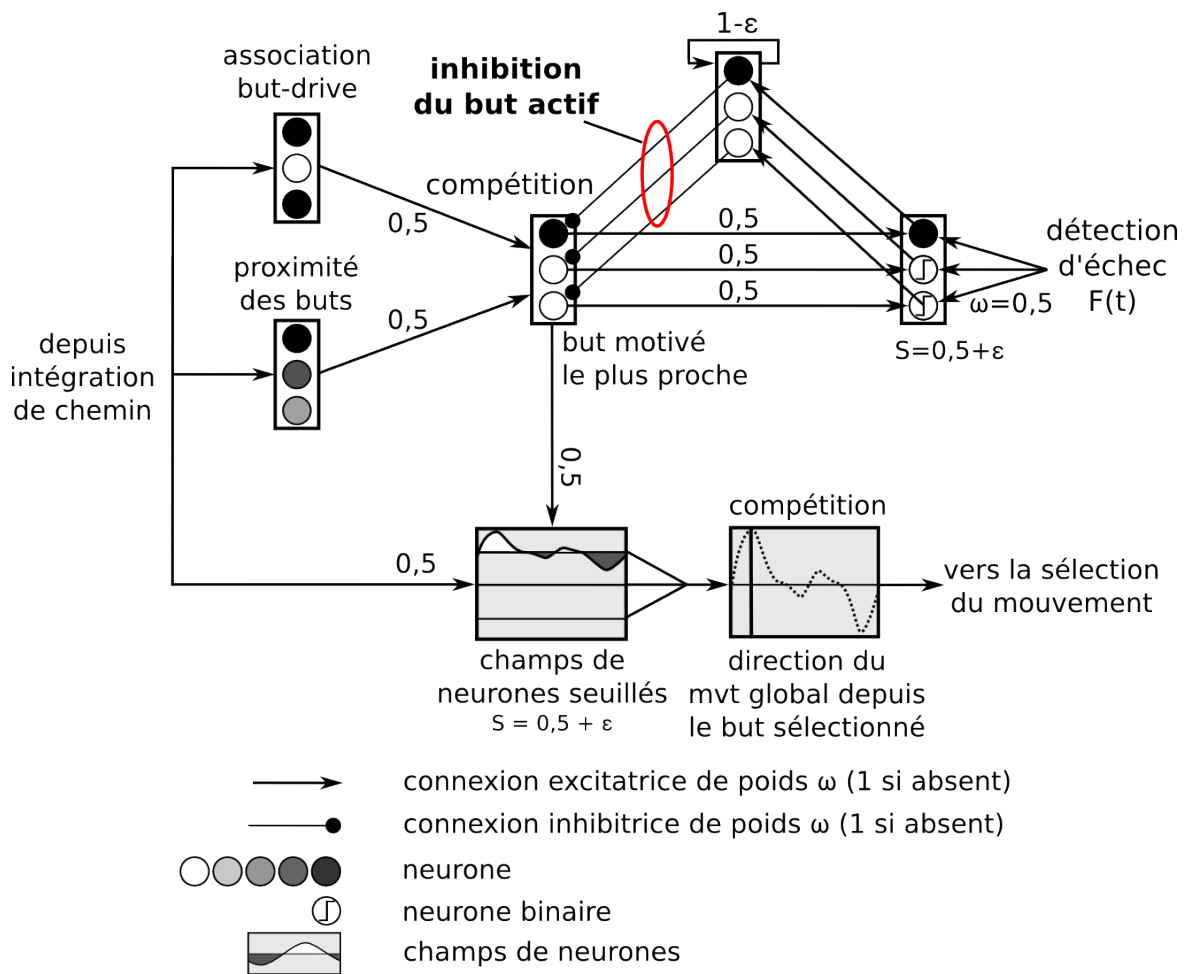


FIGURE 6.5 – Inhibition du but actif pour la stratégie de navigation proprioceptive. Quand un signal de détection d'échec est généré il est filtré pour n'inhiber que le but actif. Cette inhibition est un potentiel qui diminue progressivement. Une fois le but actif inhibé, le second but motivé le plus proche gagne la compétition et le robot se dirige vers son nouveau but.

même manière que les lieux classiques mais dans un groupe de neurones séparés. C'est la détection d'une ressource qui déclenche cet apprentissage. En retour, un seuillage est effectué sur l'activité de ces lieux-but. Lorsqu'un lieu-but a une activité suffisamment forte (le seuil est déterminé expérimentalement pour correspondre à quelques dizaines de centimètres) un potentiel inhibiteur empêche le recrutement d'une nouvelle cellule de lieu pour le même but. Ces lieux-but sont ensuite conditionnés aux drives auxquels ils correspondent (lieu-but motivé). Ce conditionnement hebbien fonctionne de la même manière que la conditionnement des différents buts aux différents drives dans l'architecture de navigation proprioceptive. Une compétition permet alors de déterminer le lieu-but le plus proche correspondant au drive exprimé. Au lieu de conditionner les cellules de lieux par les drives, il faut alors les conditionner par le lieu-but motivé vainqueur de la compétition. Cette caractérisation supplémentaire des lieux reconnus permet de sélectionner les lieux correspondants à un but et pas seulement à un drive.

Cette association lieu-lieu but peut aussi être interprétée comme une transition depuis le lieu courant vers le lieu but. Si cette distinction par rapport à l'architecture classique n'implique pas de différence comportementale lors de la navigation, elle permet cependant d'agir spécifiquement sur un sous-ensemble de cellules de lieux : les cellules de lieux associées à un but particulier.

Le mécanisme utilisé pour inhiber un but visuel particulier est exactement le même que celui utilisé pour inhiber un but proprioceptif. Le signal de détection d'échec est filtré de manière à n'affecter que le but actif i.e. la cellule de lieu-but motivé gagnant la compétition.

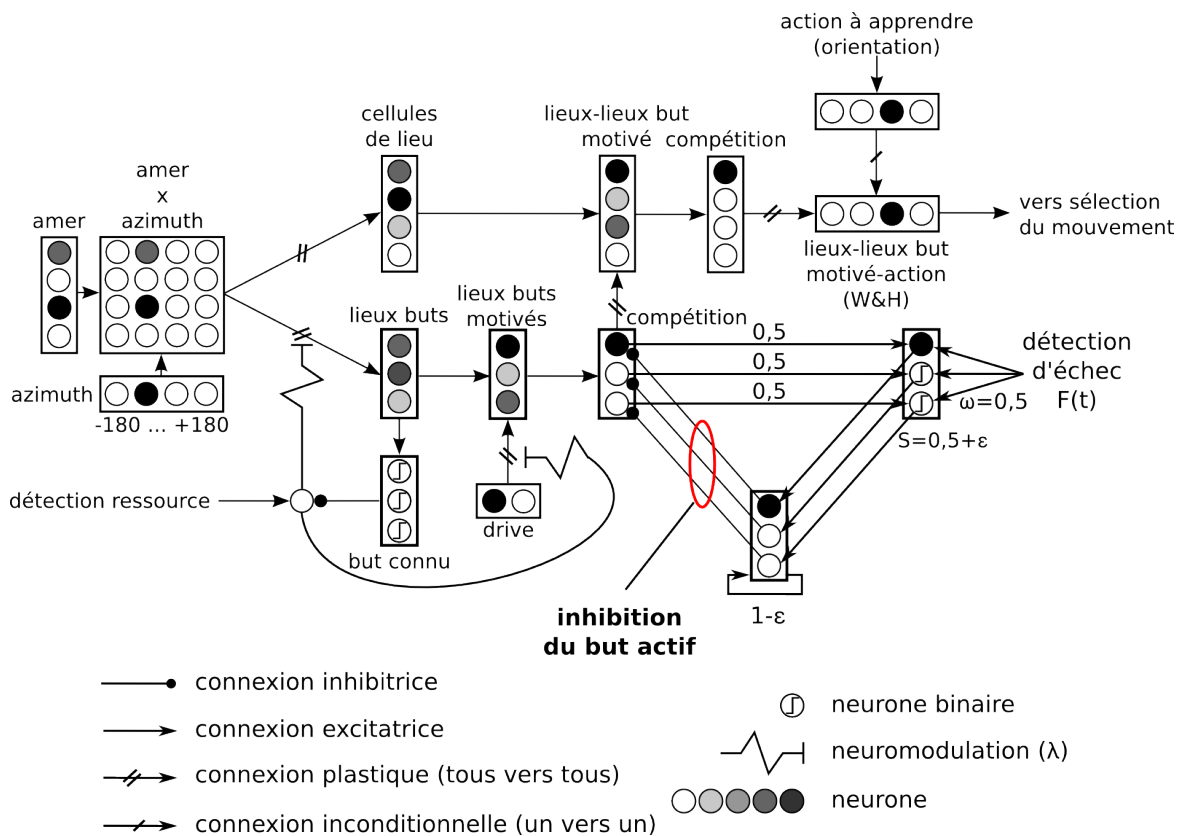


FIGURE 6.6 – Inhibition du but actif pour la stratégie de navigation visuelle. Quand un signal de détection d'échec est généré il est filtré pour n'inhiber que le but actif. Cette inhibition est un potentiel qui diminue progressivement. Une fois le but actif inhibé, le second but motivé le plus proche gagne la compétition et le robot se dirige vers son nouveau but.

Au niveau des stratégies de navigation :

Comme nous l'avons décrit précédemment, le robot peut utiliser la navigation visuelle ou la navigation proprioceptive. Ces deux stratégies de navigation ont chacune leurs avantages et leurs inconvénients. Chacune peut être mise en défaut dans un contexte particulier. Par exemple, si l'environnement du robot est plongé dans le noir, la vision ne sera d'aucun secours au robot pour se repérer et naviguer. Si le robot est déplacé manuellement i.e. si l'expérimentateur soulève le robot et le dépose un peu plus loin, le

mouvement effectué n'est pas intégré puisqu'il n'est pas détecté par la proprioception du robot (limitée à la perception des actions motrices) et les vecteurs-retour calculés deviennent faux. Dans ces deux situations, la perturbation affecte une stratégie mais pas l'autre, étant donné que chaque stratégie de navigation repose sur des informations distinctes. Tandis que le robot évolue dans son environnement, la mise en échec d'une stratégie produit nécessairement un potentiel inhibiteur. Celui-ci peut être dirigé vers la stratégie active. Dans notre cas, le robot ne dispose que de deux stratégies. Mais même si nous voulons un système simple, ce système doit pouvoir gérer un nombre quelconque de stratégies. La figure 6.7 présente notre mécanisme de sélection et d'inhibition des stratégies de navigation.

Chaque stratégie produit une sortie codée sur un groupe de neurones. Cette sortie décrit la direction que le robot doit prendre. Nous souhaitons que le robot n'utilise qu'une seule stratégie de navigation à la fois. C'est pourquoi la sortie de chaque stratégie est filtrée par un "sélecteur" de stratégie. Ce sélecteur est un groupe de neurones comportant autant de neurones que le robot dispose de stratégies de navigation. Comme ce groupe fait une compétition entre ses neurones, seule un neurone y est actif à un moment donné. Ce neurone actif permet de laisser passer l'activité de sortie de la stratégie de navigation correspondante. Ce filtrage est réalisé pour chaque stratégie par un groupe de neurones seuillés ($S=0,5$) de même taille que le signal de sortie de la stratégie. La sortie de la stratégie de navigation est transmise par des connexions de poids $w = 0,5$. Le signal venant du neurone sélecteur est lui aussi pondéré par une connexion de poids $w = 0,5$. La sortie est donc copiée si le neurone sélecteur est actif. Elle est supprimée sinon. Les sorties pondérées de toutes les stratégies sont ensuite sommées, ne laissant que l'activité de la stratégie sélectionnée. Cette activité est ensuite envoyée vers l'exécution motrice.

Le fonctionnement du sélecteur de stratégie lui-même est très similaire à la gestion du recrutement des champs d'intégration dans la navigation proprioceptive. Avant qu'un signal de détection d'échec ne soit produit, une stratégie est sélectionnée aléatoirement. Quand celle-ci mène à une situation d'échec, le signal d'inhibition est filtré de la même manière que pour l'inhibition des buts ou des drives. Mais au lieu d'inhiber directement le neurone correspondant à la stratégie mise en défaut, ce potentiel inhibiteur est stocké dans une mémoire (connexion récurrente de poids légèrement inférieur à 1). Ainsi, le groupe de neurones "mémoire stratégie inhibée" possède autant de neurones qu'il y a de stratégies de navigation. Si une stratégie n'a pas été utilisée ou n'a pas été inhibée, le potentiel de son neurone correspondant est de 0. Si au contraire elle mène à la détection d'échec, son neurone correspondant reçoit un signal de 1 qui fait croître son potentiel. Sinon le potentiel décroît lentement avec le temps. En cas de détection d'échec nous voulons changer de stratégie pour en choisir une dont le neurone correspondant a une activité faible ou nulle. Il suffit donc de mener une compétition sur le complément par rapport à 1 de ces potentiels inhibiteurs mémorisés. Une stratégie jamais utilisée correspond alors à un potentiel de 1 et une stratégie qui vient d'être inhibée correspond à un potentiel de 0. Si toutes les stratégies ont déjà été utilisées et mises en échec, c'est alors la stratégie qui n'a pas été utilisée depuis le plus longtemps qui sera sélectionnée après une détection d'échec.

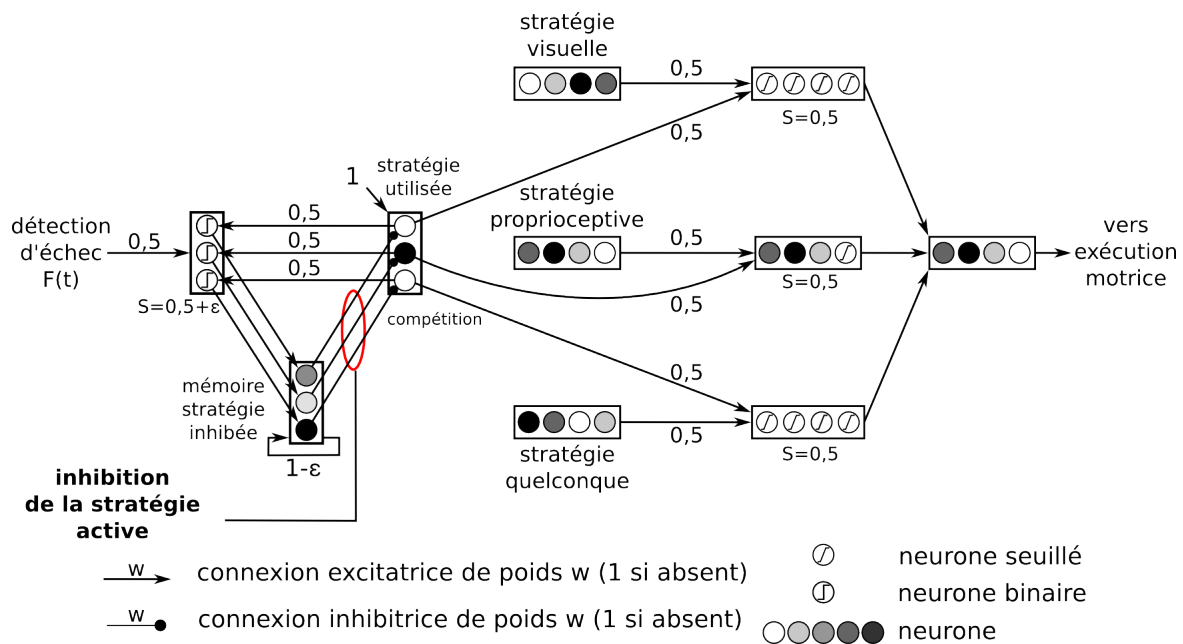


FIGURE 6.7 – Inhibition de la stratégie active. Quand un signal de détection d'échec est généré il est filtré pour n'inhiber que la stratégie active.

Dans la prochaine section, nous illustrerons cette détection de l'échec et les différents mécanismes d'inhibition régulant le comportement du robot lors d'expériences comportementales classiques de navigation motivée.

5 Expérimentations robotiques

Pour expérimenter notre système de régulation du comportement par la frustration, nous avons testé en situation réelle son fonctionnement sur les drives, les stratégies et les buts. La figure 6.8 présente l'environnement expérimental utilisé. Chaque ressource est présente à deux endroits différents de l'environnement. Cette multiplicité des buts n'est utilisée que dans l'expérience concernant l'inhibition des buts.

Inhibition des drives :

Lors de la première expérience, notre but est de tester le bon fonctionnement de l'inhibition des drives du robot. Plus précisément, il s'agit de confronter le robot à une situation d'échec et de tester le mécanisme d'inhibition sur le drive actif. Si le robot ne dispose que de ce mécanisme de régulation, celui-ci peut être compris comme la mise en attente d'un besoin devant l'incapacité à le satisfaire. Les autres besoins du robot passent alors en priorité. Lorsque que ceux-ci sont satisfaits le robot peut de nouveau tenter de satisfaire son but initial. Cette inhibition est de bas niveau en ce sens qu'elle intervient en amont de toute notion de navigation ou même de but. Elle ne peut résoudre que les situations où la cause de l'échec initial disparaît avec le temps.



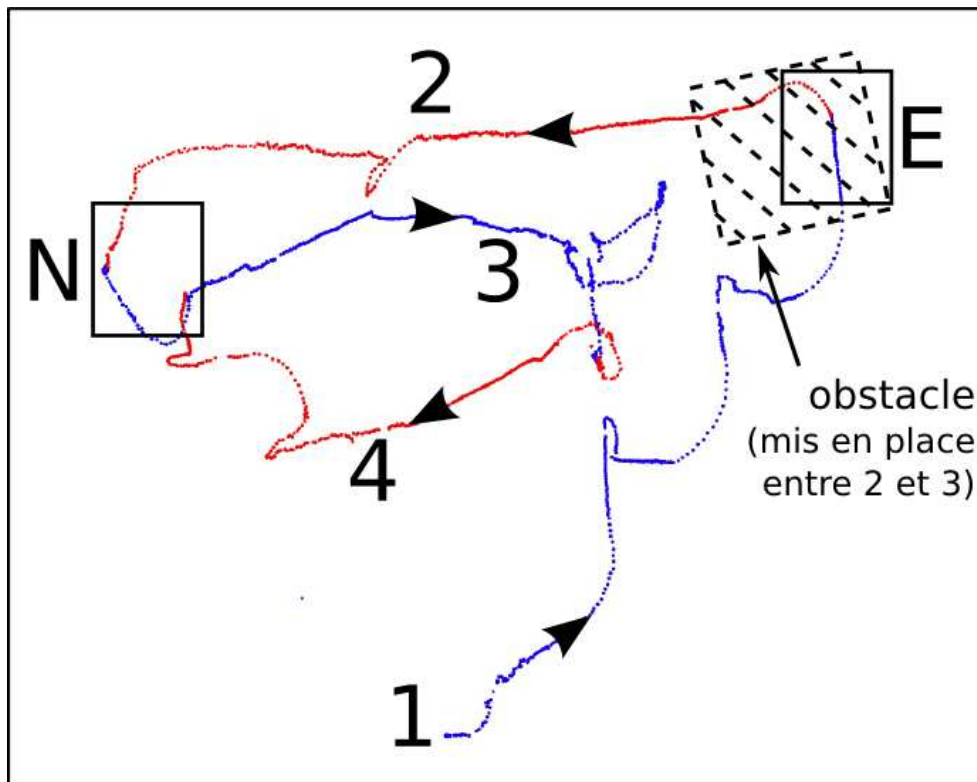
FIGURE 6.8 – L'environnement expérimental contient deux types de but chacun présent en deux lieux différents. Au démarrage de l'expérience, le robot a appris les associations sensori-motrices et calculé les vecteurs retour lui permettant de rejoindre chacun des buts avec les deux stratégies.

Dans cette expérience, le robot navigue en utilisant la stratégie visuelle.

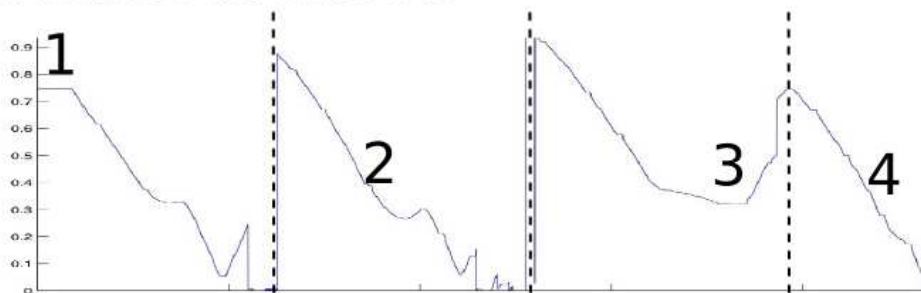
La figure 6.9 décrit la trajectoire du robot, son estimation de sa distance au but, le niveau de frustration $f(t)$ ainsi que le drive actif. Dans un premier temps, le drive actif est la soif. Le robot se dirige alors vers la source d'eau (étape 1 sur la figure). La distance au but diminue et très peu de frustration est accumulée (essentiellement à cause du bruit dans l'estimation de la distance au but). Ayant satisfait ce drive, la frustration résiduelle est inhibée et le robot se dirige alors vers la source de nourriture (étape 2 sur la figure). Ce drive est satisfait à son tour. Mais lorsque le robot a de nouveau soif et qu'il retourne vers la source d'eau, celle-ci est obstruée par un obstacle (étape 3 sur la figure). Incapable de rejoindre la source d'eau, la distance au but ne peut diminuer. Elle augmente même lorsque le robot évite l'obstacle. Lorsque suffisamment de frustration est accumulée, un potentiel inhibiteur permet de passer du drive "soif" au drive "faim" alors que la soif du robot n'a pas été satisfaite. Le robot se dirige donc de nouveau vers la source de nourriture (étape 4 sur la figure).

Inhibition des buts :

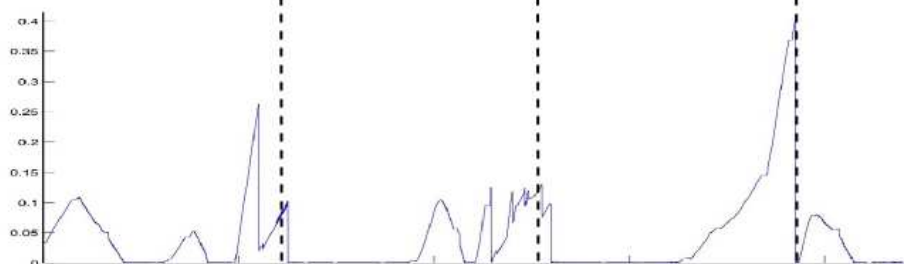
Dans cette seconde expérience, il s'agit de tester l'inhibition des buts du robot. C'est dans cette situation que nous utiliserons deux exemplaires de chaque ressource dans l'environnement. Une fois encore, le robot est confronté à une situation d'échec mais cette fois-ci, étant donné que chaque ressource est présente plusieurs fois, un même drive peut être satisfait par plusieurs buts. Le but visé par le robot est obstrué comme



estimation distance but



frustration



drive

soif faim soif faim

FIGURE 6.9 – Trajectoire et signaux internes du robot avec inhibition du drive actif. 1 : le robot a soif et se dirige vers l'eau. 2 : le robot a maintenant faim et se dirige vers la nourriture. 3 : le robot a de nouveau soif et se dirige vers l'eau mais celle-ci est inaccessible. 4 : ayant accumulé suffisamment d'échec, un signal de frustration inhibe la soif et le robot se dirige vers la nourriture.

lors de la première expérience, mais dans ce contexte plus complexe, changer de but sans changer de drive suffit à résoudre le problème. Cette inhibition est plus fine que celle décrite précédemment puisque si elle permet de modifier le comportement du robot, elle ne consiste pas à "ignorer" le problème en "espérant" qu'il se résolve de lui-même. Dans cette expérience, le robot navigue en utilisant la stratégie proprioceptive.

La figure 6.10 décrit la trajectoire du robot, son estimation de sa distance au but, le niveau de frustration $f(t)$ ainsi que le but actif. Le drive actif n'est pas indiqué mais il correspond directement à la nature du but poursuivi. L'environnement contient deux sources d'eau E1 et E2 ainsi que deux sources de nourriture N1 et N2. Dans un premier temps, le drive actif est la faim. Le robot se dirige alors vers la source de nourriture la plus proche : N1 (étape 1 sur la figure). La distance au but diminue et très peu de frustration est accumulée. Ayant satisfait son drive, la frustration résiduelle est inhibée et le robot se dirige alors vers la source d'eau la plus proche : E1 (étape 2 sur la figure). Ce drive est satisfait à son tour (étape 3 sur la figure). Mais lorsque le robot a de nouveau faim et qu'il retourne vers N1, celle-ci est obstruée par un obstacle. Incapable de la rejoindre, la distance au but ne peut diminuer et le potentiel d'échec augmente. Lorsque suffisamment de frustration est accumulée, un potentiel inhibiteur permet de passer du but E1 au but E2 alors que la soif du robot n'a pas été satisfaite (étape 4 sur la figure). Le robot se dirige donc vers E2 même si elle est plus loin que E1 (étape 5 sur la figure).

Inhibition des stratégies :

Lors de la troisième expérience, le but est de tester l'inhibition des stratégies du robot. Une fois encore, le robot est confronté à une situation d'échec mais cette fois-ci, il ne s'agit pas d'obstruer le but par un obstacle. L'objectif est de créer une situation dans laquelle la perturbation empêchant le robot d'atteindre son but n'affecte pas la but lui-même mais la stratégie de navigation utilisée pour le rejoindre. Alors que le robot utilise la navigation proprioceptive, l'expérimentateur soulève et déplace le robot avant de le reposer. Ce mouvement ne pouvant pas être intégré par le robot, le vecteur retour suivi pour atteindre son but sera nécessairement incorrect (en angle, en norme ou les deux). C'est ce que nous appelons le kidnapping du robot.

La figure 6.11 montre la trajectoire du robot, son estimation de sa distance au but, le niveau de frustration $f(t)$ ainsi que la stratégie active. Le drive actif n'est pas indiqué mais il correspond directement à la nature du but poursuivi. Dans un premier temps, le robot navigue grâce à sa proprioception. Ayant soif, il se dirige d'abord vers la source d'eau (étape 1 sur la figure). La distance au but diminue et très peu de frustration est accumulée. Ayant satisfait son drive, la frustration résiduelle est inhibée et le robot se dirige alors vers la source de nourriture (étape 2 sur la figure). Ce drive est satisfait à son tour et le robot retourne vers la source d'eau. Mais sur la trajectoire, celui-ci est kidnappé et déplacé par l'expérimentateur (étape 3 sur la figure). Son vecteur retour étant maintenant biaisé, le robot n'est plus capable d'atteindre la source d'eau et décrit des cercles autour de ce qu'il estime être l'emplacement de l'eau (étape 4 sur la figure). Incapable de rejoindre la source d'eau, la distance au but ne peut diminuer et le potentiel d'échec augmente. Lorsque suffisamment de frustration est accumulée, un potentiel inhibiteur permet de changer de stratégie de navigation. La navigation

proprioceptive inefficace est inhibée au profit de la navigation visuelle. Celle-ci n'ayant pas subi de perturbation, elle permet au robot de rejoindre la source d'eau (étape 5 sur la figure).

6 Discussion

Nous avons décrit dans ce chapitre un mécanisme de régulation du comportement inspiré de la réaction de frustration et basé sur l'utilisation d'un signal inhibiteur produit par la détection d'une situation d'échec. Il ne faut pas confondre la détection de l'échec et la réaction de frustration qui en résulte. La détection de l'échec déclenche un potentiel inhibiteur. La réaction de frustration est le changement de comportement du robot. Ce mécanisme est inspiré par le fonctionnement du rôle de régulateur du cortex pré-frontal [Dehaene 1991, Dehaene 1995, Konishi 1998]. Nous avons souligné l'importance des capacités d'auto-évaluation pour la construction de l'autonomie d'un robot. La base de cette auto-évaluation repose sur la capacité à estimer la distance au but. Nous avons montré que cette estimation de la distance au but peut être construite à partir des différentes sources d'informations à la disposition du robot. De même que le robot peut apprendre à naviguer grâce à sa vision ou sa proprioception, l'estimation de la distance au but peut provenir de la proprioception (maximum d'activité d'un champ d'intégration) aussi bien que de la vision (conditionnement d'une cellule de lieu par la perception d'un but). Si d'autres capteurs équipaient le robot, il pourrait certainement construire d'autres estimations de la distance au but. Nous pourrions par exemple imaginer de doter le robot d'un micro et de disposer une source de son particulière sur chaque but. Le même type de mécanisme que pour la vision pourrait alors être utilisé pour naviguer et pour estimer la distance d'un but. De plus, nous avons utilisé une seule évaluation de la distance au but quand le robot est capable d'en produire plusieurs (au moins une par système perceptif). Un progrès intéressant de ce système consisterait à utiliser simultanément toutes les évaluations de distance au but. Cela pourrait nous permettre de construire une estimation multi-modale de la distance au but plus robuste aux perturbations. Plus intéressant encore serait l'utilisation d'un conditionnement de chaque estimation de la distance au but à un contexte perceptif donné. Si par exemple dans le noir complet l'estimation visuelle de la distance au but devient fautive, l'estimation proprioceptive reste correcte. La détection du contexte perceptif "pas d'information visuelle" pourrait alors sélectionner l'utilisation exclusive de l'estimation proprioceptive.

Nous avons développé un système de détection d'échec primordial dans notre système de régulation du comportement du robot. Bien que ce système soit très simple, il est supérieur à un simple système de *watchdog* en ce sens qu'il n'y a pas de durée déterminée pour qualifier une situation d'échec. Ce qui compte, ce n'est pas simplement le temps écoulé pour atteindre le but, mais bien l'estimation de la progression vers le but. Si le robot met longtemps à atteindre son but mais que sa progression est constante, aucun échec ne sera détecté. Cependant, le seuil de frustration définit bien la composante temporelle de la perception d'échec. Nous avons travaillé avec un seuil fixe déterminé expérimentalement en fonction de la taille de l'environnement, de la dis-

tance respective des différents buts ainsi que de la vitesse d'utilisation des ressources par le robot. Un ajustement adaptatif de ce seuil serait bien plus pertinent. En partant d'une valeur quelconque ce seuil pourrait être augmenté à chaque détection d'échec et diminuer à chaque détection du but. Ainsi, le robot apprendrait à être très exigeant dans un environnement d'abondance et très tolérant dans un environnement pauvre en ressource. On peut aussi se poser la question de l'intérêt de la modulation de ce seuil de frustration par le niveau des drives. Autrement dit, il pourrait sembler pertinent de permettre au robot une plus grande souplesse face à l'échec lorsqu'aucun besoin n'est trop fort. Au contraire, lorsque le niveau d'un variable physiologique devient critique, il pourrait être utile d'abandonner plus rapidement un comportement inefficace.

Ce système d'auto-évaluation peut sembler poser problème lorsque le chemin permettant de rejoindre le but n'est pas une ligne droite. S'il est nécessaire de faire un détour pour atteindre le but, l'absence de progrès perçu durant la phase d'éloignement devrait déclencher la perception d'une situation d'échec. Cependant, il ne s'agit pas ici d'une limitation de notre système d'auto-évaluation mais bien d'une limitation des stratégies de navigation utilisées. La construction d'un bassin d'attraction visuel de même que l'intégration de chemin ne permettent pas de construire telle quelle des solutions de navigation à des problèmes complexes. Il existe cependant des stratégies de navigation plus sophistiquées et permettant de résoudre ce type de problème de navigation. La navigation par carte cognitive décrites par Tolman [Tolman 1948] et notamment implémentée au laboratoire ETIS [Giovannangeli 2008b] en est un bon exemple. Sans rentrer dans les détails du fonctionnement de cette stratégie de navigation, la carte cognitive permet d'apprendre les actions de transition entre les lieux. La navigation vers le but correspond donc à la remonter d'un gradient diffusé depuis le lieu-but à travers les différentes transitions existantes. Le chemin le plus court vers le but ne peut donc emprunter que les transitions existantes. Ainsi, même avec un détour le robot remonte continuellement le gradient de prédiction du but diffusé dans la carte cognitive. En utilisant notre système d'auto-monitoring avec ce gradient plutôt que directement avec les signaux perceptifs, nous dépasserions donc le problème du détour. Nous avons aussi illustré le fonctionnement du signal d'inhibition préfrontale à différents niveaux de l'architecture de contrôle du robot : les drives, les buts et les stratégies de navigation. Nous n'avons exploré qu'une toute petite partie du rôle du cortex préfrontal. Notre modèle se limite ici au monitoring des prédictions de buts et à la régulation par l'inhibition. Notre modèle se limite à un contrôle comportemental réactif en cas de détection d'échec. Nous pourrions affecter à chaque but, drive et stratégie une estimation apprise du potentiel de satisfaction et d'échec. A l'instar des potentiels de récompense des actions appris par le nucleus accumbens, le cortex préfrontal pourrait apprendre à évaluer les drives, buts et stratégies. Il s'agirait d'apprendre des conditionnements adaptatifs (apprentissage continu) entre drives, buts et stratégies d'une part et satisfaction du but et détection d'échec d'autre part. Un but, un drive ou une stratégie aurait ainsi un potentiel intrinsèque de satisfaction et d'échec. Cette évaluation apprise pourrait s'ajouter à la sélection réactive proposée ici. Elle permettrait par exemple d'apprendre au robot à ne plus retourner vers un but définitivement inaccessible. De même que pour une utilisation contextuelle des différentes estimations

de distance au but, la sélection d'une stratégie de navigation pourrait être associée au contexte perceptif. Par exemple, le robot pourrait apprendre que la navigation visuelle mène toujours à une situation d'échec lorsque la luminosité de l'environnement est trop faible. Cette utilisation du contexte permettrait au robot d'anticiper certaines situations d'échec.

Dans ce chapitre nous avons présenté trois niveaux possibles pour recevoir le signal d'inhibition permettant de modifier le comportement du robot. Ces différents types d'inhibition ont été testés expérimentalement avec succès. Mais pour un système pouvant utiliser ces différentes possibilités de contrôle, il n'y a pas en l'état de système permettant d'apprendre à favoriser une inhibition par rapport aux autres. Cependant, de même que nous envisageons d'apprendre des potentiels de satisfaction et d'échec pour les différents drives, buts et stratégies, nous pourrions apprendre à conditionner chaque inhibition par son potentiel à produire de la satisfaction ou de l'échec. Par exemple, dans un environnement plongé dans le noir, le robot pourrait apprendre qu'inhiber le but actif ou le drive exprimé n'est pas efficace, au contraire de l'inhibition de la stratégie active. Si par contre un but est définitivement inaccessible au robot, celui-ci pourrait alors apprendre que l'inhibition de la stratégie n'est pas efficace à la différence de l'inhibition du but ou du drive. Si cette apprentissage n'est pas contextualisé, le robot apprendra à réguler efficacement son comportement mais uniquement dans son environnement. En changeant d'environnement, ou en changeant les caractéristiques de l'environnement, les inhibitions efficaces peuvent ne plus être les mêmes. Si la priorité des différentes stratégies ou des différents buts peut être apprise par l'expérience de l'environnement, nous devrions pouvoir concevoir un protocole expérimental permettant de tester la réelle capacité des animaux à changer la priorité de leurs méthodes et de leur buts.

La frustration n'est généralement pas considérée comme une émotion basique. Il est d'ailleurs très peu fait état de la frustration dans l'étude des émotions et de leur rôle dans la régulation du comportement. Mais nous avons voulu souligner au contraire l'aspect primordial de cette émotion. Nous avons montré que la douleur permet d'apprendre la peur et que le plaisir est à la source des motivations. C'est à partir de ces fondamentaux qu'il est possible d'apprendre à orienter le comportement comme le ferait un être vivant i.e. pour assurer sa survie. Mais une régulation active et dynamique du comportement demande un niveau d'information un peu plus complexe sur l'environnement et sur l'impact du comportement du robot. C'est pourquoi les capacités d'auto-évaluation du comportement sont si importantes.

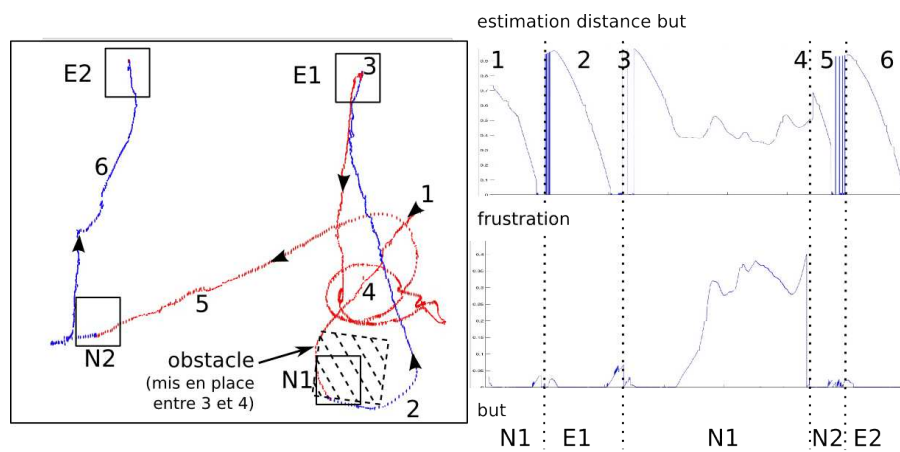


FIGURE 6.10 – Trajectoire et signaux internes du robot avec inhibition du but actif. 1 : le robot a faim et se dirige vers N1. 2 : le robot a maintenant soif et se dirige vers la E1. 3 : le robot a de nouveau faim et se dirige vers N1. 4 : N1 étant inaccessible, le potentiel d'échec s'accumule jusqu'à la libération du potentiel inhibiteur sur N1. 5 : le robot se dirige maintenant vers N2.

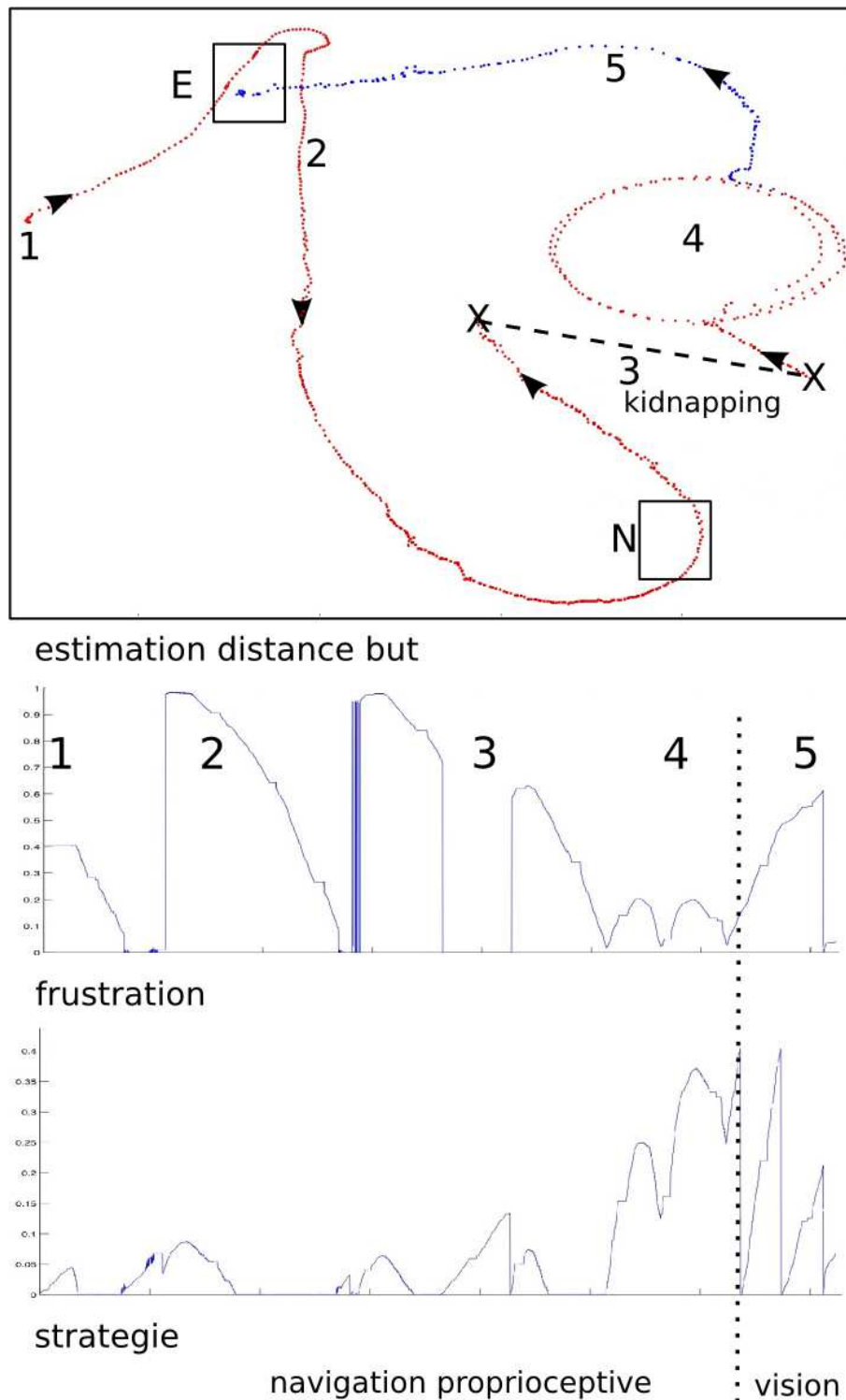


FIGURE 6.11 – Trajectoire et signaux internes du robot avec inhibition de la stratégie. Le robot utilise la navigation proprioceptive. 1 : le robot a soif et se dirige vers l'eau. 2 : le robot a faim et se dirige vers la nourriture. 3 : en allant vers l'eau, le robot est kidnappé et déposé ailleurs. 4 : le vecteur retour est biaisé, le robot tourne autour de ce qu'il considère comme étant l'emplacement de son but. Le potentiel d'échec s'accumule jusqu'à la libération du potentiel inhibiteur sur la stratégie active. 5 : le robot navigue maintenant visuellement et se dirige vers l'eau.

Surprise et nouveauté

1 Introduction

La surprise est habituellement définie comme la réaction à une situation inattendue. Si cette définition est consensuelle, il existe cependant une controverse quand à sa nature émotionnelle ou purement cognitive. D'un côté, la surprise est universelle en ce sens qu'elle est intrinsèquement associée à une expressivité spécifique comme la peur ou la colère [Ekman 1969, Ekman 1971, Ekman 1972]. De plus, elle est vue comme un processus adaptatif fondamental de régulation du comportement [Darwin 1872]. Cependant, ses détracteurs en tant que processus émotionnel soulignent sa neutralité en termes de valence affective [Ortony 1988]. La surprise peut aussi bien être neutre que plaisante ou déplaisante. Selon ce point de vue, la surprise peut éventuellement engendrer une réaction émotionnelle positive ou négative, mais elle-même est neutre. Quoi qu'il en soit, la surprise définie comme la réaction à l'inattendue repose sur deux capacités fondamentales. Tout d'abord, il faut pouvoir construire des attentes sur les régularités de fonctionnement du monde. Ensuite, il faut pouvoir détecter quand ces attentes ne sont pas respectées. Autrement dit, il faut pouvoir apprendre ce qui est normal pour détecter ce qui est nouveau. Nous avons présenté précédemment l'hippocampe, cette structure cérébrale assurant tant de fonctions cognitives majeures. Entre autres, l'hippocampe semble fortement impliquée dans la détection de nouveauté (c.f. chapitre 3). Le propos de ce chapitre est donc notamment de proposer une modèle minimal et générique de la circuiterie neuronale de l'hippocampe permettant d'assurer la détection de nouveauté. La surprise comme réaction à la détection de nouveauté est aussi bien adaptée au sens commun de la surprise qu'au sens minimal que nous avons l'intention d'en retirer pour notre utilisation dans l'architecture de contrôle du robot. Concrètement, un neurone utilisant la règle d'apprentissage du *Least Mean Square* (LMS) apprend à prédire un signal d'entrée grâce à un signal prédicteur. L'algorithme d'adaptation des poids des connexions entre le signal prédicteur et le neurone prédicteur lui permet de reproduire le signal à prédire. La surprise telle qu'elle a été définie s'applique particulièrement bien à ce niveau fondamental de traitement de l'information. Elle peut alors être comprise comme étant le résultat de la détection d'une erreur significative d'un neurone prédicteur. Cette utilisation de l'analyse du signal d'un prédicteur s'apparente au modèle proposé par Kaplan et Oudeyer [Oudeyer 2004, Oudeyer 2005, Oudeyer 2007, Kaplan 2003, Kaplan 2004] dans leur étude de la curiosité comme motivation intrinsèque. Leur modèle est constituée d'un module prédicteur M qui apprend à prédire les signaux sensori-moteurs en entrée et d'un module méta-prédicteur $metaM$, qui apprend à prédire les erreurs faites par le prédicteur M . Le progrès est défini comme la différence entre l'erreur prédite et

l'erreur courante. De façon à doter le robot d'une forme de curiosité, ils associent à ces prédicteurs un système de valeurs qui incite le robot à éviter les situations trop familières (plus de progrès possible) et les situations trop chaotiques (pas de progrès possible). Le robot privilégie ainsi les situations où le progrès d'apprentissage des prédicteurs est maximal. Au lieu cependant de nous intéresser à la curiosité comme recherche de nouveauté, nous nous intéresserons à la surprise comme détecteur de nouveauté et plus particulièrement comme détecteur de conflits sensoriels. Autrement dit, le système de détection de nouveauté et donc génération de la réaction de surprise fonctionne dans l'état stabilisation de l'apprentissage du système de curiosité décrit par Kaplan et Oudeyer.

Nous avons défini plus haut la surprise comme la violation d'une attente sur la régularité de fonctionnement du monde. Mais qu'est-ce que le robot peut forger comme attente sur le fonctionnement du monde? Et bien dans un premier temps, il peut apprendre les liens qui existent entre ses différentes perceptions. Comment la vision du robot (caméra) permet-elle de prédire sa proprioception motrice (odométrie)? En retour bien sûr, comment la proprioception motrice permet-elle de prédire la vision? De même, chacune de ces perceptions peut apprendre à prédire les autres. Ces prédictions intermodales permettent d'apprendre la cohérence fondamentale des processus perceptifs. La détection d'une violation de cette cohérence correspond donc à une situation nouvelle : une perturbation du robot ou de ses conditions environnementales. C'est donc toujours avec cette approche minimale que nous approcherons le fonctionnement du processus de surprise. Non seulement parce que nous le modélisons à un niveau élémentaire du traitement de l'information, mais aussi parce que nous l'appliquons à un contexte de bas niveau : la construction de la cohérence perceptive.

Du fait de son absence de valence propre, il est habituel de considérer que la surprise n'a pas de contrepartie. Cependant, si l'on se focalise plutôt sur ce qui caractérise la surprise, i.e. sa composante cognitive (la détection de nouveauté), il semble sensé de considérer que son opposé est l'anticipation. Mais l'anticipation est précisément ce qui permet la détection de nouveauté (i.e. la violation d'une attente). De ce point de vue, l'anticipation participe à la surprise.

Dans un premier temps nous décrirons le modèle neuronal générique utilisé pour analyser le signal d'un prédicteur en termes d'erreurs et de progrès. Ce modèle est autonome en ce sens qu'il assure par lui même la régulation de ses apprentissages. Après avoir testé ce réseau dans des situations simplifiées, nous l'utiliserons de manière massivement redondante pour analyser les différentes prédictions intermodales apprises par le robot dans un contexte normal. Cet apprentissage permet alors au robot de détecter les situations de conflits sensoriels (incohérence entre plusieurs modalités sensorielles) survenant lorsque le fonctionnement du robot est perturbé (aveuglement, éblouissement, kidnapping, capteur endommagé, ...). Non seulement le conflit peut être détecté, mais ce système de monitoring des prédictions intermodales permet aussi de déterminer quelle modalité sensorielle en est à la base. Cette surprise de bas niveau peut donc être comprise comme un premier niveau de traitement des attentes sur le monde.

2 Modèle neuronal de l'analyse d'un prédicteur

Tout les prédicteurs utilisent la règle d'apprentissage de Widrow et Hoff. C'est pourquoi ils sont noté *lms* pour *Least Mean Square* (réduction des erreurs quadratiques). La réponse d'un prédicteur initial est analysée en termes d'erreurs et de progrès de prédiction. La différence entre le signal à prédire et la réponse du prédicteur initial est l'erreur de prédiction. Un prédicteur approxime l'erreur moyenne de prédiction en prédisant le signal d'erreur à partir d'une constante. La différence entre une erreur de prédiction et l'erreur moyenne de prédiction correspond soit à un progrès soit à une régression de l'apprentissage du prédicteur initial (réduction ou augmentation de l'erreur). Cependant, au lieu de chercher à guider le robot vers les domaines de fonctionnement qui maximise son progrès d'apprentissage (notion de curiosité telle que définie par Kaplan et Oudeyer), nous voulons utiliser ce traitement du signal de prédiction pour déterminer lorsqu'une erreur de prédiction peut être considérée comme anormale et devrait donc initier un signal de surprise. De plus cette chaîne de traitement nécessitant elle-même plusieurs prédicteurs, il faut que ses apprentissages soient contrôlés de manière autonome. La figure 7.1 représente le réseau de neurones utilisé pour réaliser ce traitement du signal de prédiction.

A l'entrée du système se trouvent deux signaux A et B et le prédicteur initial, $Pred_A^B$ qui apprend à prédire B (lien inconditionnel) à partir de A (lien conditionnel). L'erreur de prédiction courante E est simplement calculée en faisant la différence entre le signal à prédire B et la sortie du prédicteur initial $Pred_A^B$. Deux neurones calculent cette différence dans chaque sens (surestimation et sous-estimation du signal à prédire) et leur somme permet d'obtenir l'erreur de prédiction absolue. C'est ce qui permet de ne pas moyennner des erreurs positives et négatives.

$$E = |Pred_A^B - B| \quad (7.1)$$

En répétant la même structure neuronale mais en considérant l'erreur E comme un nouveau signal d'entrée, notre réseau permet d'estimer l'erreur moyenne de prédiction (prédiction de l'erreur à partir d'une constante), le progrès dans la prédiction initiale (diminution de l'erreur de prédiction par rapport à l'erreur de prédiction moyenne), mais aussi la régression (le contraire). La somme du progrès et de la régression correspond donc à la variation de l'erreur de prédiction $var E$.

$$var E = |E - \bar{E}| \quad (7.2)$$

Fondamentalement, une erreur de prédiction anormale est une erreur de prédiction fortement supérieure à l'erreur moyenne. Pour que ce réseau soit fonctionnel, il lui faut donc un critère pour déterminer quand une erreur de prédiction est suffisamment importante pour être considérée comme anormale. Pour cela, il convient d'utiliser un critère relatif aux données traitées, aucun critère absolu ne permettant d'obtenir une sensibilité aux erreurs équivalente pour des signaux d'entrée différents. De la même manière que l'erreur moyenne \bar{E} est évaluée à partir des erreurs de surestimation et de sous-estimation, la variation moyenne de l'erreur $var \bar{E}$ est évaluée à partir du progrès

et de la régression recul. Cette variation moyenne de l'erreur de prédiction est une estimation de l'écart-type de l'erreur (écart moyen à la moyenne). Cette valeur peut donc servir de critère interne aux données. Une erreur anormale est une erreur de prédiction supérieure à l'erreur moyenne d'au moins n fois la variation moyenne de l'erreur. Ce critère n est codé par le poids w_c du lien inhibiteur entre la variation moyenne de l'erreur et la détection d'erreur anormale. Un poids w_c valant 3 signifie qu'une erreur de prédiction doit être supérieure à l'erreur moyenne d'au moins 3 fois l'écart-type pour être considérée comme anormale. De plus, pour être autonome, ce réseau doit contrôler l'apprentissage des différents prédicteurs utilisés. C'est pourquoi un prédicteur de cette chaîne de traitement a pour rôle d'approximer le progrès moyen \overline{Prog} . Lorsque \overline{Prog} est proche de 0, c'est que le prédicteur ne peut pas améliorer son apprentissage. On peut alors dire qu'il est stable. Le neurone utilisé pour représenter cette stabilité reçoit une entrée constante (biais) activatrice avec un poids arbitraire w_b et une entrée inhibitrice depuis le progrès moyen. Ce neurone de stabilité *stab* (neurone binaire, seuil $S = w_b - \varepsilon$) devient actif quand le prédicteur ne progresse plus i.e. quand le progrès moyen \overline{Prog} est proche de 0. Il est utilisé pour deux fonctions. D'une part, il inhibe la détection d'erreur anormale. Cela correspond au fait que tant que le prédicteur initiale n'est pas stable, il n'est pas pertinent de considérer ses réponses comme fiables. Une erreur anormale dans ce contexte ne traduit pas alors nécessairement une irrégularité significative dans le flux de données. D'autre part, l'activité du neurone de fiabilité est utilisée pour inhiber le neurone de nouveauté. Ce neurone de nouveauté, un neurone binaire (seuil $S = w_b - \varepsilon$), reçoit une entrée activatrice constante avec un poids w_b . Il n'est donc actif que lorsque la prédiction est considérée comme "non stable". Son activité est utilisée en rebouclage pour contrôler la neuromodulation de l'apprentissage des prédicteurs du réseau. λ_1 , λ_2 , λ_3 et λ_4 contrôlent l'apprentissage du prédicteur initial, de l'erreur moyenne, du progrès moyen et de la variation moyenne de l'erreur. Ainsi, ce réseau apprend tant qu'il n'est pas stable. Enfin, la détection d'un erreur anormale inhibe le neurone de stabilité ce qui active le neurone de nouveauté afin que le réseau puisse de nouveau apprendre à prédire la nouvelle situation. Il convient donc d'utiliser des vitesses d'apprentissage λ qui ne soient pas trop élevées pour que le robot n'apprenne pas qu'une situation de perturbation transitoire est normale. Les erreurs anormales E_a et la stabilité *stab* se calculent donc de la manière suivante :

$$E_a = (E - \overline{E}) - w_c \cdot \overline{varE} - stab \quad (7.3)$$

A l'opposé des prédictions anormalement fausses, que peut-on alors dire des prédictions anormalement bonnes ? Dans les faits, celles-ci ne sont observées que lors des phases d'apprentissage, alors que les prédicteurs ne sont pas encore exploitables (le réseau n'est pas encore stable). Il est cependant possible d'imaginer que des prédictions anormalement bonnes traduisent une simplification brusque des caractéristiques du signal à prédire. Des prédictions anormalement bonnes devraient alors aussi déclencher une réaction de surprise. Cependant, en pratique, cette situation ne se produit pas si les signaux à prédire sont les perceptions d'un environnement réel, complexe par nature.

$$stab = \begin{cases} 1 & \text{si } (w_b - \overline{Prog} - E_a) > w_b - \varepsilon \\ 0 & \text{sinon} \end{cases} \quad (7.4)$$

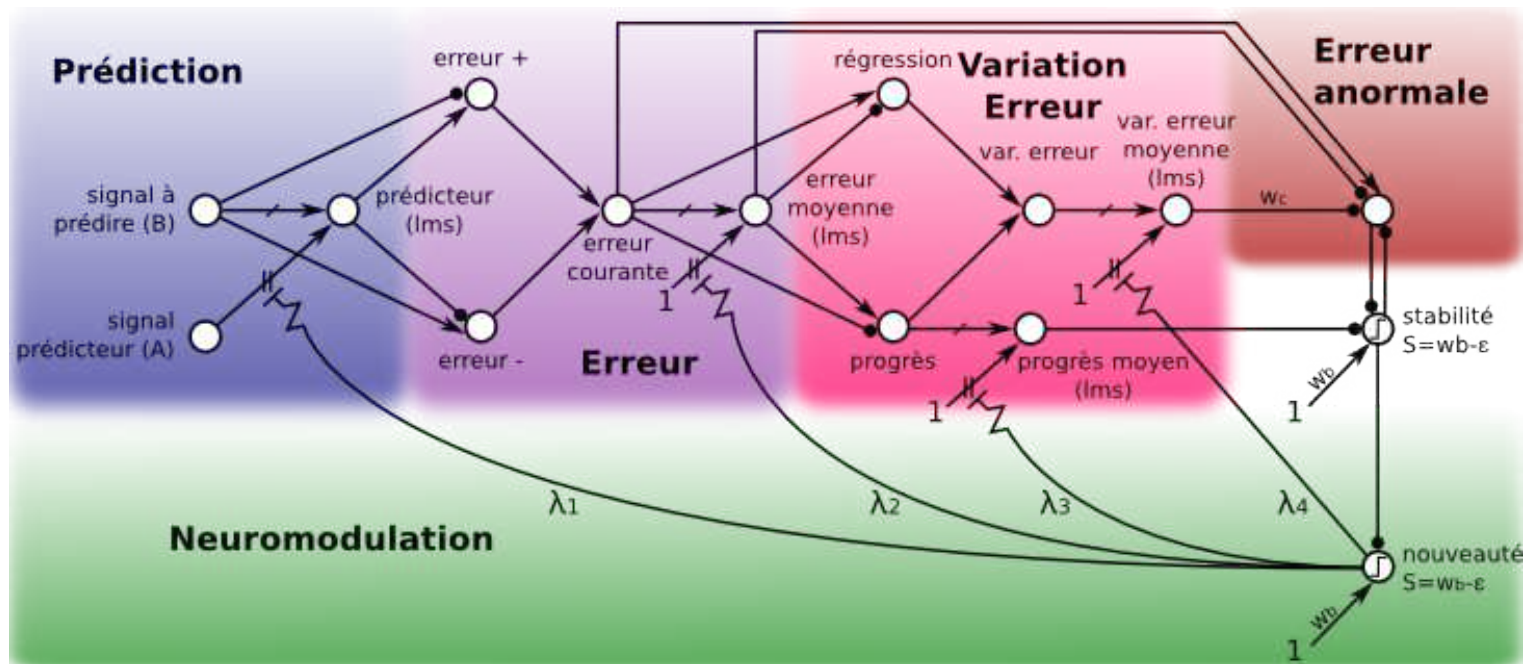


FIGURE 7.1 – Réseau de neurones utilisé pour analyser la réponse d’un prédicteur. Après évaluation de l’erreur, l’erreur moyenne peut être approximée par un second prédicteur. L’estimation de la variation moyenne de l’erreur permet de déterminer quand une erreur est considérée comme anormale. Le progrès est la diminution de l’erreur. L’approximation du progrès moyen permet de contrôler l’apprentissage de ce réseau. Quand le progrès moyen est nul (i.e. quand le prédicteur n’est plus capable de réduire son erreur), l’apprentissage est suspendu.

3 Tests unimodaux

Ce réseau permet d'analyser la prédiction d'un signal sensoriel par un autre signal. Nous comptons l'utiliser pour permettre aux différentes modalités sensorielles de s'interpréter. Mais dans un premier temps, le but est de l'évaluer dans un cas plus simple : la cas unimodal. Un seul signal sensoriel d'entrée est utilisé. Lors de ces tests unimodaux, nous avons pris la vision. En particulier, le signal à prédire est la luminance moyenne de l'image perçue. Le signal prédicteur est une simple constante valant 1. Bien sûr cette limitation ne permet pas de tirer pleinement parti du potentiel de ce réseau. cela permet néanmoins d'éclairer son fonctionnement élémentaire. En pratique, utiliser une constante comme signal prédicteur avec l'algorithme LMS revient à faire une approximation de la valeur moyenne du signal [Giovannangeli 2008a]. Le prédicteur initial approxime donc la luminance moyenne. Le signal d'entrée est le flux visuel d'une caméra (640x480) codé dans une matrice de neurones réalisant un sous-échantillonnage de l'image (32x24 = 768 neurones). L'activité de chaque neurone de cette matrice correspond à la luminance moyenne observée dans le carré de 20x20 pixels correspondant dans l'image brute. Une activité nulle sur un neurone d'entrée correspond à une zone locale noire dans l'image, et inversement une forte activité traduit une zone de forte luminance (blanche). La figure 7.2 montre le résultat de ce codage pour une image prise dans l'environnement du robot.

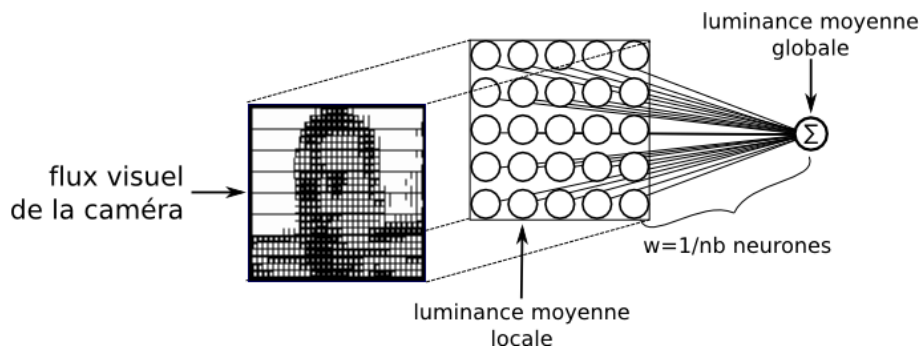


FIGURE 7.2 – Le flux visuel venant de la caméra est codé sur une matrice de neurones (32x24). L'activité de chaque neurone représente la luminance moyenne de la zone de l'image qu'il perçoit (en quelque sorte son champs récepteur). La convergence des activités de cette "rétine" via des connexions normalisées donne la luminance moyenne globale perçue.

L'activité de chaque neurone de cette "rétine" permet donc de coder la luminance moyenne locale de l'image de la caméra. Une luminance moyenne globale peut donc être calculée très simplement en faisant converger les activités de ces 768 neurones vers un seul neurone le long de connexions activatrice dont le poids est de 1/768. Il s'agit de l'équivalent neuronal d'une somme normalisée ou calcul de moyenne. Ce premier système va donc étudier et prédire les variations de luminance moyenne pour détecter les situations nouvelles.

3.1 Perturbations statiques de la vision

Le premier test réalisé sur la modalité visuelle montre le fonctionnement de l'analyse des prédicteurs dans une situation relativement simple. La caméra est statique pour toute la durée de l'expérience. Dans un premier temps, le signal se stabilise sur l'image perçue après sa mise en fonctionnement. Puis la caméra est ensuite obstruée par un obstacle placé devant l'objectif. Ensuite, l'obstacle est retiré et le système voit à nouveau. Après cela, la caméra est éblouie à l'aide d'une lampe électrique braquée sur l'objectif. Enfin, la lampe est éteinte, laissant le système percevoir normalement à nouveau. La figure 7.3 présente le déroulement temporel des différents signaux du réseau de neurones d'analyse du prédicteur : le signal à prédire (luminance moyenne), sa moyenne prédite, l'erreur de prédiction, l'erreur de prédiction moyenne, les erreurs de prédiction anormales et le progrès de prédiction.

Le premier graphique présente le signal de luminance moyenne de l'image qui varie en fonction des perturbations subites par le système. Le signal plonge quand la caméra est obstruée et sature quand elle est éblouie. Il est sensible aux changements de conditions avant d'adapter et de stabiliser rapidement sa réponse. Ce comportement est dû au fonctionnement logiciel bas niveau de la caméra qui cherche à optimiser la balance des blancs en fonction de la luminance moyenne. Le signal prédit, i.e. l'approximation de la valeur moyenne du signal, suit le signal à prédire d'autant mieux qu'il est stable. Dans le deuxième graphique, l'erreur de prédiction fait un pic lors des phases de transition entre les différentes situations subies par le système de vision. Elle se réduit progressivement tandis que le prédicteur apprend à s'adapter à cette nouvelle situation. Dans le troisième graphique, le progrès est présent après chaque perturbation et correspond à une phase pendant laquelle le prédicteur s'adapte à la nouvelle situation en minimisant l'erreur de prédiction. Les erreurs anormales sont détectées lors des phases de transition, quand le prédicteur n'a pas encore eu le temps de s'adapter. Dans cet exemple, ce système permet de détecter quand le système produit des réponses inhabituelles ou nouvelles. Son fonctionnement ressemble à celui que l'on aurait pu obtenir en mettant un filtre passe bas et un dérivateur.

3.2 Navigation et changement d'environnement

Le second test, plus dynamique, consiste à étudier le comportement du prédicteur lors d'un enchaînement d'actions du robot. La caméra est installée sur le robot qui dans un premier temps navigue dans un couloir étroit en ligne droite, avant d'entrer dans une pièce aux caractéristiques visuelles très différentes du couloir. Ensuite le robot fait un tour complet en rotation sur la gauche puis sur la droite. Le but est de voir dans quelle mesure l'analyse du signal de prédiction, en particulier des ses erreurs, permet de détecter ces changements aux plus bas niveau de traitement du signal visuel. La figure 7.4 présente le signal prédit, l'erreur de prédiction et les erreurs anormales.

Après une première phase d'adaptation du prédicteur, le graphique de la figure 7.4 montre une phase de navigation dans le couloir durant laquelle les changements dans le flux visuel sont faibles et progressives. Durant cette phase, l'erreur de prédiction reste faible. Le système parvient à prédire la luminance moyenne du flux visuel correc-

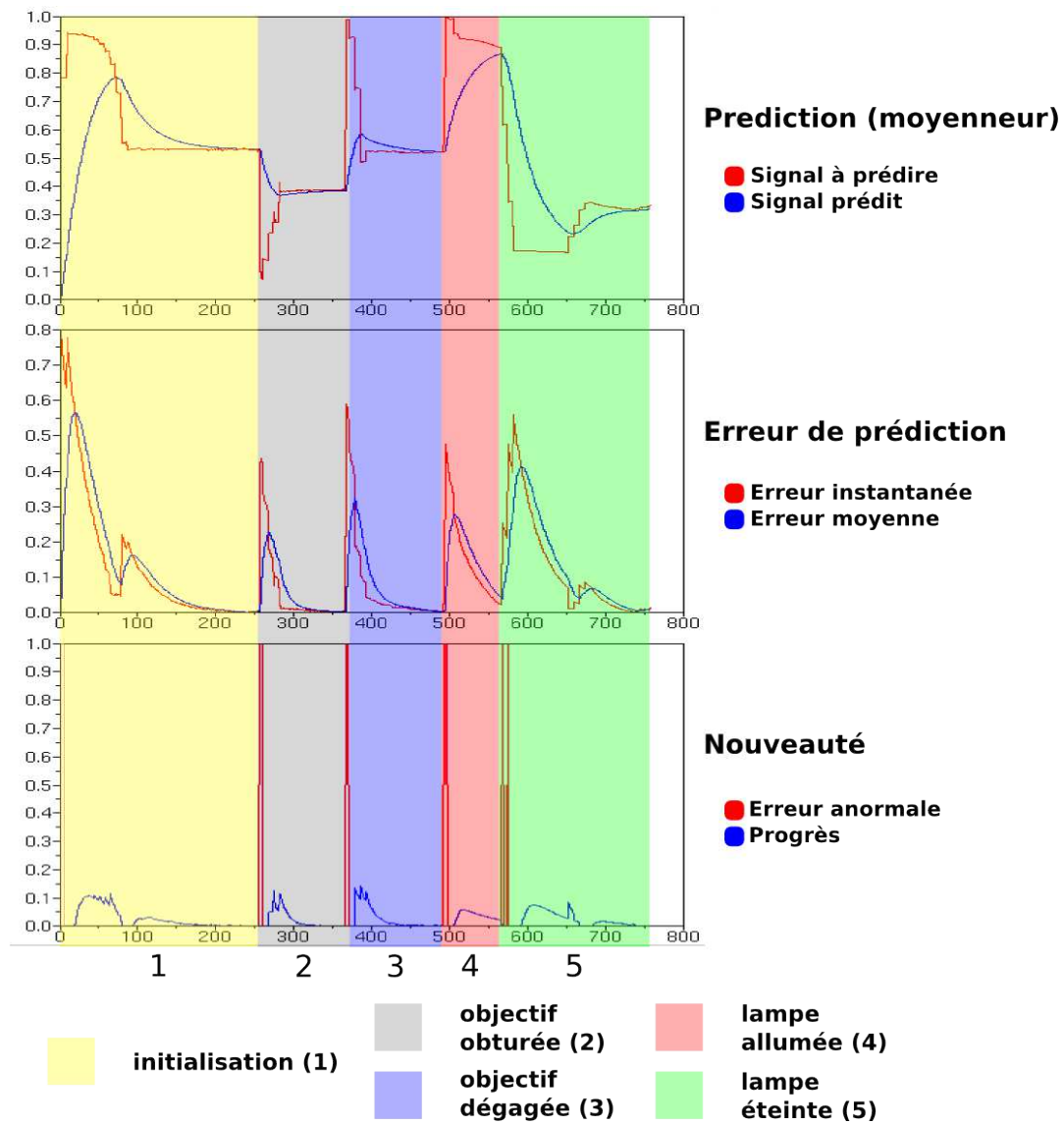


FIGURE 7.3 – Lors du premier test du réseau d'analyse de la prédiction, la caméra est obstruée puis éblouie. Le réseau produit des erreurs anormales lors des phases de transition. Il est donc capable de détecter les changements inhabituels dus à une situation nouvelle.

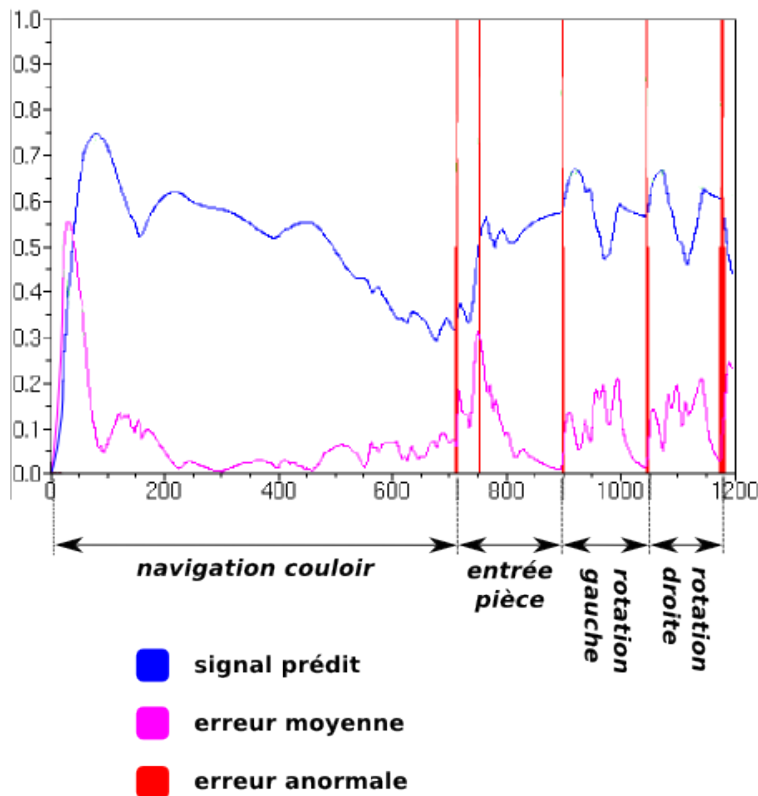


FIGURE 7.4 – Le robot navigue depuis un couloir vers une pièce puis fait un tour complet vers la gauche et vers la droite. Simplement sur la base du signal visuel, l'analyse des erreurs de prédiction permet de détecter les phases de transition vers des conditions nouvelles.

tement puisque la caméra avance en ligne droite dans un environnement symétrique et longiligne. La luminance évolue donc peut et progressivement. Aucune erreur anormale n'est détectée. Le robot entre ensuite dans un espace ouvert, une grande pièce dont tout un mur est constitué de baies-vitrées. Cela provoque une augmentation brutale de la luminance moyenne de l'image perçue et donc une erreur de prédiction suffisamment forte pour être anormale. Le système s'habitue peu à peu à ce nouvel environnement et parvient de nouveau à réduire l'erreur. Enfin, le robot effectue une rotation de 360 degrés à gauche, puis de 360 degrés à droite. Le changement rapide luminance provoqué par ces rotations déclenche lui aussi des erreurs anormales. Avec cette architecture rudimentaire, le robot est donc capable de distinguer des irrégularités dans son flux visuel et ainsi de percevoir un changement de pièce (couloir versus pièce) ou de type d'action (mouvement rectiligne versus rotation).

4 Interprédiction et détection de conflits sensoriels

4.1 Les conflits sensoriels chez l'homme

Un conflit sensoriel se produit lorsque le cerveau reçoit des informations contradictoires depuis différentes modalités sensorielles. Le corps humain est constitué de sens

externes tels que la vision, l'ouïe, l'odorat, le toucher ou le goût, mais également de sens internes comme la proprioception qui nous renseigne sur la position relative de nos membres ou encore le système vestibulaire, situé dans l'oreille interne, qui permet de percevoir la gravité et l'accélération. Il est important de faire la différence entre ces deux classes sensorielles dont une distinction fondamentale réside dans le fait qu'une modalité externe peut facilement être perturbée, tandis que les modalités internes sont, par nature, rarement faussées dans la vie de tous les jours. Ainsi, le cerveau gère relativement bien les situations de conflit entre modalités externes, mais est bien souvent incapable de résoudre, de manière rationnelle, les conflits qui font intervenir une modalité interne. Il en résulte la création d'une nouvelle cohérence irrationnelle perturbante (e.g. illusion, désorientation) voir, dans les cas plus extrêmes, d'une rupture perceptive désagréable (e.g. douleurs fantômes, cinétose). Le cerveau est parfaitement adapté pour résoudre les conflits sensoriels mineurs qui se produisent constamment dans la vie de tous les jours. Ces conflits se produisent uniquement à partir d'une modalité sensorielle externe, et font référence à des situations dans lesquelles un sens est décorrélié des autres. Un bon exemple de ce type de conflit est la privation d'un des sens par une perturbation spécifique. L'obscurité ou à l'inverse, l'éblouissement sont des conditions provoquant un conflit entre la vision devenue non informative et les autres perceptions qui elles ne sont pas affectées. Dans cette situation, il paraît légitime de ne pas prendre en compte la vision et de privilégier les autres sens. Cette disposition temporaire dans le cas normal se traduit par des stratégies de substitution pour le cas d'une lésion permanente. Ainsi, les personnes non-voyantes utilisent généralement une canne pour se déplacer, c'est à dire une extension directe du sens du toucher. D'autres préféreront un système auditif directif renseignant l'utilisateur des obstacles proches, ou même des couleurs présentes dans l'environnement. Un conflit avec un sens interne a des conséquences plus importantes puisqu'il concerne un sens qui n'est habituellement pas perturbable. C'est le cas notamment de la cinétose, plus connu sous le nom de "mal de mer", qui désigne en réalité le mal du mouvement dans sa globalité. Le phénomène se produit lorsqu'une personne est en mouvement dans un véhicule quelconque, abandonnant ainsi le mode de déplacement naturel. Les accélérations et décélérations répétées perçues par l'oreille interne entre en conflit avec les signaux visuels provenant de l'intérieur du véhicule et qui décrivent un environnement immobile. C'est pourquoi la meilleure solution consiste alors à regarder à l'extérieur (en voiture par exemple) pour percevoir visuellement les mouvements du véhicule et rétablir la cohérence intermodale. Une utilisation particulièrement impressionnante de ces situations de conflits sensoriels est illustrée par la stimulation du système vestibulaire (*Galvanic Vestibular Stimulation*) [Fitzpatrick 2004]. Une situation d'incohérence perceptive est obtenue en stimulant, par impulsions électriques, le système vestibulaire. Ce processus affecte sensiblement l'équilibre du sujet de telle sorte qu'il ait l'impression de tomber dans une direction particulière. Pour compenser cette sensation de chute et retrouver son équilibre, ce procédé pousse le sujet à faire quelques pas dans la direction en question. Ainsi, il est possible de "contrôler" la locomotion d'une personne en excitant son oreille interne. L'illusion de Pinocchio est un autre exemple de conflit sensoriel qui permet d'illustrer la notion de reconstruction rationnelle [Lackner 1988, Ehrsson 1988]. En ap-

pliquant sur les muscles du bras des vibreurs électriques, il est possible de déclencher des signaux factices d'élongation ou de contraction de ces muscles, créant chez le sujet une illusion de mouvement. Le sujet place son doigt en contact avec le bout de son nez de sorte qu'il puisse sentir ce contact et le percevoir visuellement. Lorsque l'on simule un mouvement du bras par la vibration, on génère un conflit sensoriel entre la vision et le toucher d'une part et la proprioception d'autre part. Le toucher et la vision affirme au sujet qu'il y a bien contact avec le nez, tandis que la proprioception indique un mouvement du bras vers l'avant ou l'arrière. Le résultat de cette expérience est la création d'une nouvelle cohérence permettant de lever la contradiction. Puisque le contact n'est pas interrompu, le sujet sent son nez se raccourcir si le mouvement illusoire rapproche l'index du visage et, au contraire, s'allonger si l'index s'en éloigne (à l'instar du nez de Pinocchio lorsqu'il ment).

4.2 Les conflits sensoriels pour le robot

Le but du modèle de détection de conflits sensoriels présenté dans cette section est de permettre au robot de détecter une irrégularité dans ses perceptions et ainsi de déterminer quand il est dans une situation où il ne peut pas se fier à une de ses modalités sensorielles. Il s'agit d'une application bas niveau du modèle de la surprise définie comme la détection d'une erreur de prédiction inattendue. Ce choix se justifie par la volonté de comprendre le fonctionnement des émotions d'un point de vue constructiviste. Il convient donc de tester notre modèle de détection de la nouveauté directement sur les signaux sensoriels. De plus, la question des conflits sensoriels est primordiale pour un robot qui navigue dans la mesure où il utilise ses différentes perceptions pour construire ses différentes stratégies de navigation. La navigation sensori-motrice repose sur la vision et l'intégration de chemin sur la proprioception. Si l'une de ces modalités est perturbée, la stratégie de navigation correspondante sera alors inefficace. Il est donc nécessaire de détecter la présence d'un conflit mais aussi de déterminer quelle perception est à la source de ce conflit. C'est le rôle du modèle de prédictions inter-modales décrit dans cette section. Ce modèle repose sur une utilisation systématique du réseau de neurones présenté dans la figure 7.1 et utilisé de manière simplifiée sur le signal visuel lors des tests unimodaux. Au lieu de prédire une caractéristique moyenne d'un signal sensoriel, ce réseau est utilisé pour prédire chaque signal sensoriel à partir de chacun des autres. Dans notre cas, le robot dispose d'une vision (caméra), d'une proprioception (odométrie) et d'un modèle simplifié de système vestibulaire (boussole électronique). Possédant trois modalités sensorielles et chaque modalité prédisant les deux autres, notre modèle repose sur six prédicteurs prenant chacun un signal prédicteur (connexion conditionnelle) et un signal à prédire (connexion inconditionnelle). Nous aurions pu considérer aussi les prédictions de chaque modalité par la concomitance des deux autres (prédicteurs prenant chacun deux signaux prédicteurs) mais notre objectif est d'arriver à une architecture minimale (soucis de parcimonie). Voilà donc les prédicteurs utilisés :

- VpO : Vision prédit Odométrie
- VpA : Vision prédit Accélération
- OpV : Odométrie prédit Vision

- OpA : Odométrie prédit Accélération
- ApV : Accélération prédit Vision
- ApO : Accélération prédit Odométrie

Chacun de ces prédicteurs est suivi du réseau de neurones permettant l'analyse du signal de prédiction. Ainsi sont évalués son erreur de prédiction, son erreur de prédiction moyenne, son progrès et surtout ses erreurs de prédictions anormales. Cependant, si une modalité sensorielle est perturbée, par exemple si la caméra du robot est obstruée, la vision prédira aussi mal les autres perceptions que les autres perceptions la prédiront. Des erreurs anormales seront donc détectées sur plusieurs des prédicteurs. C'est donc grâce à un mécanisme de compétition que ce modèle permet de déterminer quelle est la modalité sensorielle à la cause du conflit. Les erreurs anormales des différents prédicteurs sont sommées en fonction de leur modalité sensorielle prédictive. Les erreurs anormales de VpO sont donc sommées avec celles de VpA, OpV avec OpA et ApV avec ApO. Le but est de calculer un potentiel d'erreurs anormales pour chacune des modalités. La valeur de ce potentiel n'est pas intrinsèquement importante. Seul compte le rapport entre les potentiels d'erreurs anormales des différentes modalités. Les poids des connexions entre les erreurs anormales et les potentiels d'erreurs anormales nécessitent donc d'être égaux et pas trop élevés. Si les neurones codant pour ces potentiels d'erreurs anormales saturent (i.e. reçoivent une activation cumulée supérieure à 1), l'information est perdue et la compétition ne désignera pas nécessairement la modalité au potentiel d'erreurs anormales le plus important. Dans notre cas, avec trois modalités sensorielles utilisées, chaque potentiel d'erreurs anormales a deux entrées. Avec quatre modalités, cela aurait été trois entrées, et ainsi de suite. Le signal d'erreur anormale est binaire donc le poids de ces connexions excitatrices entre les erreurs anormales et les potentiels d'erreurs anormales peut être pris à n'importe quelle valeur comprise entre 0 (exclu) et $1/(\text{nombre modalité} - 1)$. Si par exemple, un robot qui navigue a sa caméra qui tombe en panne, il détectera des erreurs anormales sur la vision (VpO et VpA), sur l'odométrie (OpV) et sur le système vestibulaire (accélération) (ApV). Le potentiel d'erreurs anormales le plus important correspond donc à la modalité à la cause des conflits sensoriels. Ce fonctionnement compétitif de la détection d'une modalité sensorielle problématique ne peut donc être mis en oeuvre que si le robot dispose d'au moins trois entrées perceptives. C'est ce qui permet une asymétrie entre les potentiels d'erreurs anormales des différentes modalités. Avec deux entrées seulement, chaque perception prédit l'autre et une perturbation sur l'une d'entre elle provoque une erreur anormale pour les deux prédicteurs. La figure 7.5 présente le modèle d'interprédiction utilisé sur le robot. Afin de tester ce modèle sur des signaux simples, les entrées sensorielles sont les dérivées temporelles des différents flux. Pour la vision il s'agit des variations de luminosité moyenne, pour l'odométrie il s'agit des variations d'activité motrice et pour l'accélération (système vestibulaire) il s'agit des variations d'activité de la boussole électronique (ou de vrais accéléromètres quand ils sont disponibles).

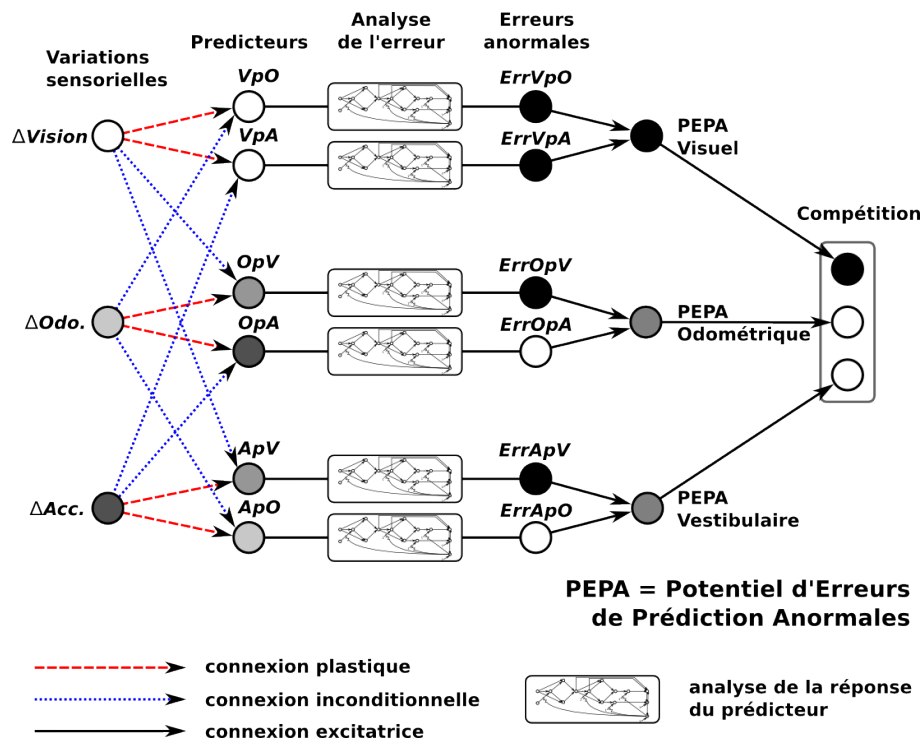


FIGURE 7.5 – Modèle de détection de conflit sensoriel par l’analyse des interprédictions perceptives. Chaque modalité tente de prédire les autres. Lors d’un conflit sensoriel, la modalité provoquant le plus d’erreurs anormales est considérée comme responsable. Dans cet exemple, la caméra est obstruée alors que le robot se déplace. Aucune modalité ne prédit la vision (une erreur anormale chacune) et le vision ne prédit plus aucune modalité (deux erreurs anormales). La vision est donc détectée comme étant à la cause du conflit sensoriel.

4.3 Test du modèle d’interprédictibilité

Afin de valider ce modèle de manière expérimentale, il est indispensable de mettre volontairement le robot dans des situations de conflits sensoriels. Ces situations sont très faciles à produire, même dans les situations les plus simples. Le robot tourne sur lui même à vitesse constante et va tenter d’apprendre les corrélations existantes entre ses différentes variations sensorielles. Dans un premier temps, toutes les perceptions du robot sont fonctionnelles. Cette première phase est maintenue le temps nécessaire à la convergence des prédicteurs afin de laisser le système apprendre correctement les interprédictions lorsque tout va bien (pas de conflits sensoriels). Deux situations de perturbation sont alors testées. La première est la privation d’une modalité sensorielle du robot. En pratique, et parce que cela peut correspondre à la situation concrète où le robot se retrouve dans le noir, c’est la vision qui est perturbée. La caméra du robot est obstruée par un carton opaque alors que le robot continue de se déplacer. Aucune variation n’est perçue sur la vision alors que les deux autres modalités détectent toujours les déplacements du robot. La seconde situation est une sorte de kidnapping. Alors que la robot est à l’arrêt, il s’agit de le faire tourner sur lui même à l’aide d’une planche préalablement placée sous ses roues. Faire tourner la planche permet

d'appliquer une rotation au robot sans activation de ses moteurs. En le déplaçant "manuellement" plutôt qu'en le laissant se déplacer de lui-même, le conflit est alors causé par la proprioception qui ne perçoit pas de variation alors que les signaux de vision et d'accélération détectent bien le déplacement du robot. La figure 7.6 présente les différents signaux sensoriels durant le déroulement de cette expérience.

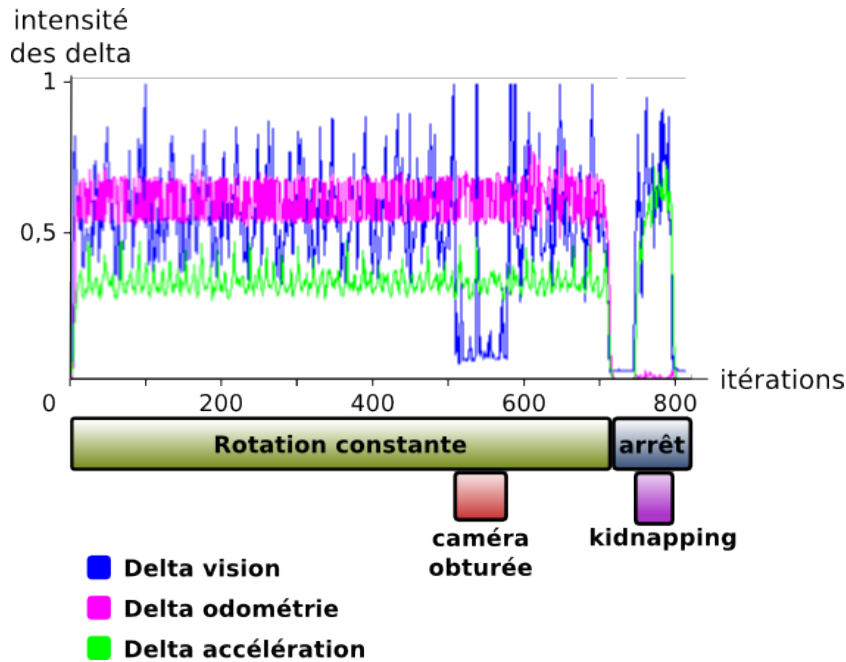


FIGURE 7.6 – Dans cette expérience test, le système de prédictions intermodales utilise les variations perçues sur les flux sensoriels de la vision (caméra), de la proprioception (odométrie) et du système vestibulaire (boussole électronique). Tandis que le robot est en rotation constante, ces trois entrées sensorielles perçoivent des changements. Dans un second temps, la caméra est obturée tandis que le robot continue de tourner (odométrie et accélération toujours perçues). Enfin, le robot arrête sa rotation mais l'expérimentateur déplace manuellement le robot (vision et accélération toujours perçues).

Le delta visuel (en bleu) est le plus bruyé des trois, et ce parce que la modalité visuelle, étant une modalité externe, est la plus sujette aux "bruits" dus à la diversité de l'environnement. Cependant, cela ne doit pas avoir d'impact puisque l'évaluation du caractère anormal d'une erreur de prédiction dépend de sa distance à l'erreur de prédiction moyenne. Une erreur de prédiction anormale concernant la vision devra donc nécessairement être de forte amplitude. Les trois signaux sensoriels restent cohérents lors de la rotation jusqu'à ce qu'on obstrue la caméra. La vision ne perçoit plus qu'une image pratiquement noire dont les variations sont quasiment nulles, tandis que les autres modalités perçoivent toujours les variations dues aux mouvements du robot. La vision est ensuite rétabli et retrouve instantanément ses variations habituelles. Ensuite, on stoppe le mouvement de rotation du robot. En conséquence, aucun signal sensoriel ne varie plus. Enfin, durant la phase de kidnapping, l'odométrie ne perçoit aucune variation de mouvement puisque les moteurs ne sont pas activés mais la caméra et la boussole révèlent la perception du mouvement forcé du robot. On peut consta-

ter que la variation d'orientation est légèrement supérieure à la normale puisque le mouvement de rotation forcée réalisé manuellement est un peu plus brusque que la commande motrice exécutée par le robot.

La figure 7.7 présente les prédictions de la vision par l'odométrie (à gauche) et l'accélération (à droite). Sur ces deux graphes on visualise les activités des deux prédicteurs (en rose) qui tentent de prédire la vision (en bleu). Les poids des liens conditionnels sont initialisés à zéro et s'adaptent au fur et à mesure que les prédicteurs apprennent. Tel que prévu, l'odométrie et l'accélération cessent de prédire la vision quand la caméra est obstruée. Et lors du kidnapping, l'odométrie ne permet pas de prédire la vision au contraire de la perception de l'accélération qui n'est pas perturbée.

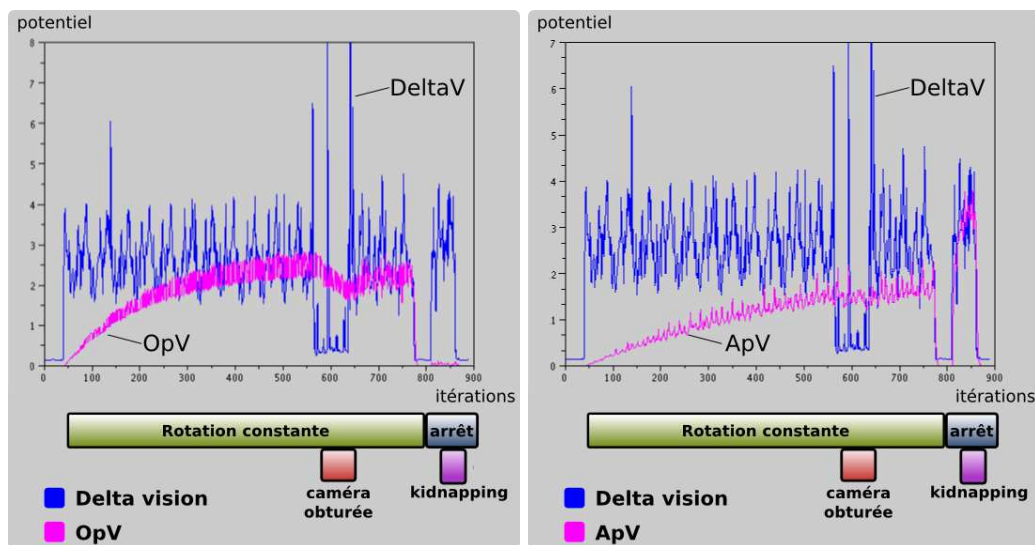


FIGURE 7.7 – Prédiction du signal visuel à partir des signaux d'odométrie et d'accélération.

La figure 7.8 présente les prédictions de l'odométrie par la vision (à gauche) et l'accélération (à droite). Comme on s'y attend la vision n'est plus capable de faire de prédictions correctes lorsque l'on voile la caméra. Durant la phase de kidnapping, les prédicteurs utilisant la vision et l'accélération sont actifs puisqu'ils s'attendent à percevoir une variation d'odométrie (qui ne survient pas). Ces deux prédicteurs commettent donc une erreur de prédiction qui sera utilisée pour en déduire par la suite un problème concernant l'odométrie.

La figure 7.9 présente les prédictions de l'accélération par la vision (à gauche) et l'odométrie (à droite). Comme pour la prédiction de l'odométrie, la vision ne prédit plus rien lors de l'obturation de caméra. Par contre, si le kidnapping ne permet pas à l'odométrie de prédire l'accélération, il n'empêche pas à la vision de prédire l'accélération.

Etant donné leur nombre, il n'est pas pertinent de présenter les courbes d'erreur et d'erreurs moyennes de tout les prédicteurs. La figure 7.10 prend deux exemples. Les erreurs pour la prédiction de l'accélération par l'odométrie et pour la prédiction de l'odométrie par la vision. Pour les deux prédicteurs, l'erreur converge vers zéro au démarrage. Cela correspondant à la phase d'apprentissage. Ensuite, le prédicteur uti-

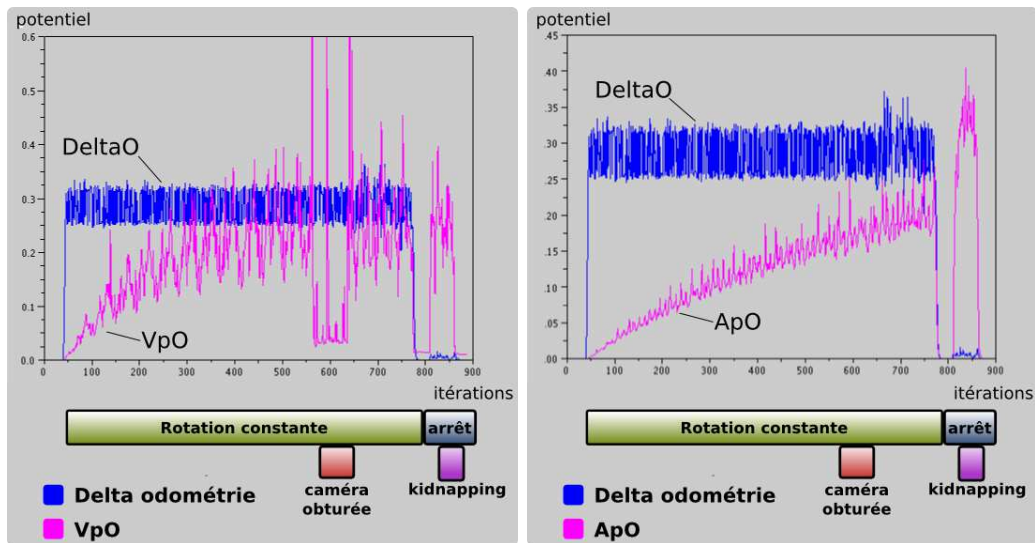


FIGURE 7.8 – Prédiction du signal odométrique à partir des signaux de vision et d'accélération.

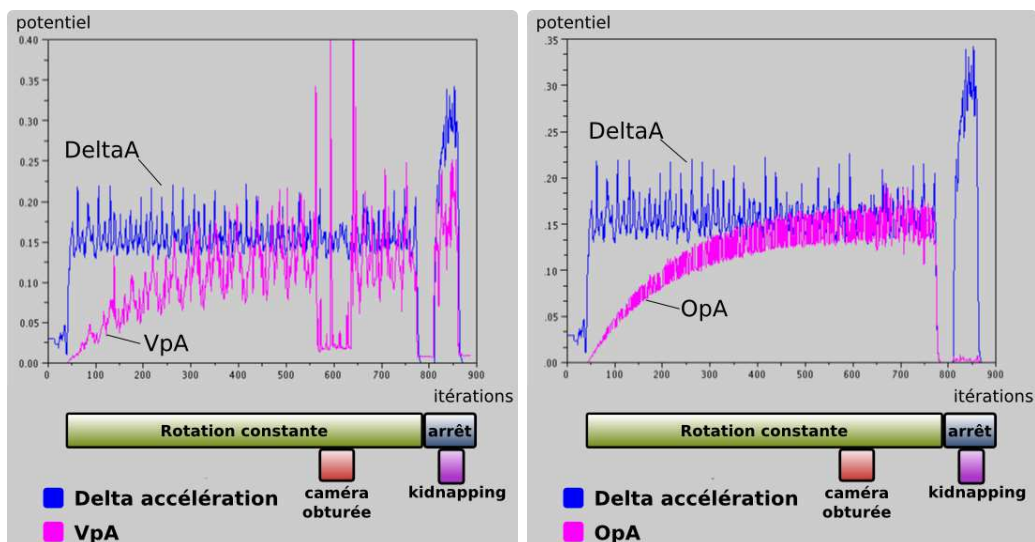


FIGURE 7.9 – Prédiction du signal d'accélération à partir des signaux de vision et d'odométrie.

lisant la vision, VpO , produit une erreur importante au moment où la caméra est obstruée. L'odométrie et l'accélération ne subissant aucune perturbation, OpA progresse normalement et son erreur continue de diminuer. Enfin, durant le kidnapping, les deux prédicteurs produisent une erreur. La variation sur la vision prédit une variation de l'odométrie qui n'a pas lieu. L'odométrie ne prédit donc plus l'accélération du robot.

La figure 7.11 permet de visualiser l'évaluation des potentiels d'erreurs de prédiction anormales de chaque modalité. Aucun problème important n'est détecté durant la première phase puisque le robot tourne à vitesse constante sans perturbation et que tant que les prédicteurs n'ont pas convergé, la détection d'erreur anormale est inhi-

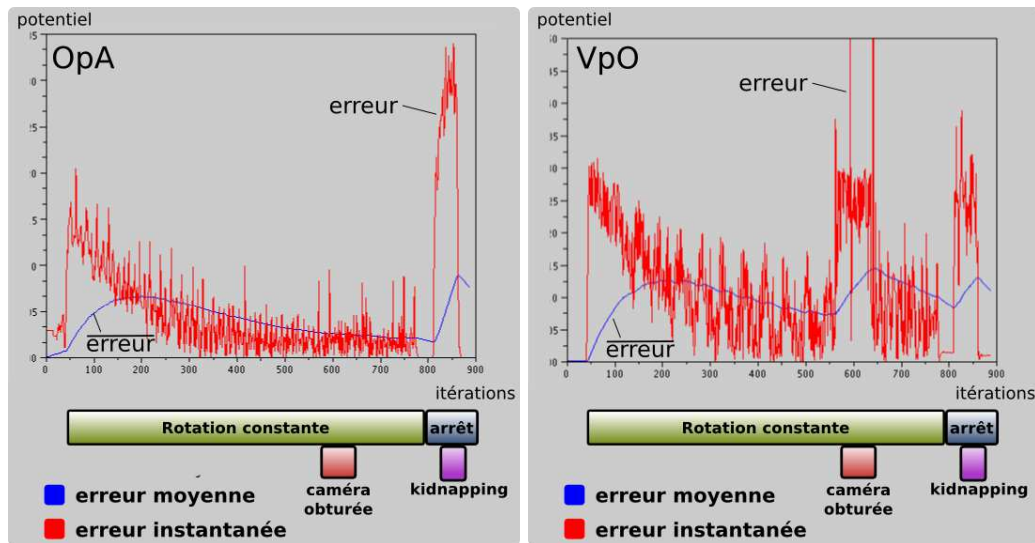


FIGURE 7.10 – Erreurs et erreurs moyennes des prédicteur OpA et VpO

bée. L’obturation de la caméra génère par contre un potentiel d’erreurs de prédiction anormales majeur sur la vision, plus léger sur l’odométrie et presque nul sur la boussole. De même, le kidnapping engendre un potentiel d’erreurs de prédiction anormales important associé à l’odométrie, et plus faible sur les autres modalités. Le système est donc capable de caractériser la fiabilité de chacune de ses entrées en fonction des perturbations de l’environnement.

5 Discussion

Cette analyse des signaux d’interprédiction sensorielles permet d’extraire des informations temporelles particulièrement intéressantes pour la construction de la cohérence perceptive intermodale ainsi que pour la détection de nouveauté et la réaction de surprise qui en résulte. Le robot apprend les régularités de sa perception, ou comment ses différentes perceptions évoluent dans le temps en fonction de ses actions. Nous avons utilisé ce modèle afin de permettre au robot de détecter les situations de conflits sensoriels dues à des perturbations que nous lui avons fait subir. Malgré son apparence lourde et complexe, le système d’interprédiction sensoriel repose sur la répétition d’une unité fonctionnelle relativement simple : le réseau de neurones analysant le comportement d’un prédicteur. Ce réseau lui même consiste essentiellement en une répétition de l’analyse de l’erreur de prédiction permettant de calculer l’erreur, puis l’erreur de l’erreur. Les tâches effectuées à l’aide de notre système de détection de la nouveauté ne sont que des illustrations basiques et incomplètes du potentiel de ce système. Un premier niveau d’utilisation de ce modèle consisterai à moduler l’utilisation des différentes stratégies de navigation du robot en fonction de son évaluation de la fiabilité respective de ses différentes modalités sensorielles. Chaque stratégie est construite sur la base d’informations venant préférentiellement d’une modalité sensorielle. La navigation visuelle est inefficace dans le noir puisque le robot ne peut reconnaître aucun

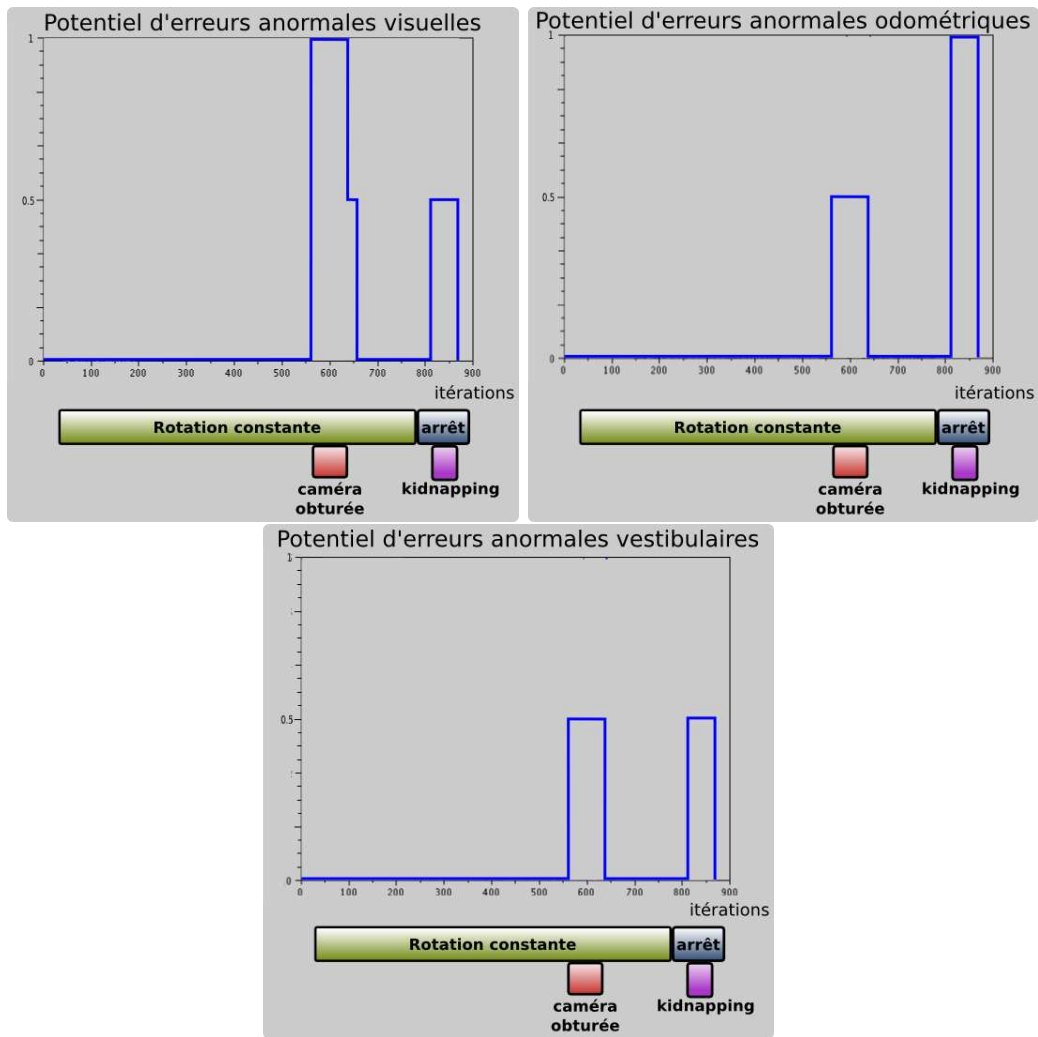


FIGURE 7.11 – Potentiel d’erreurs de prédiction anormales pour chaque modalité. Lors de l’obturation de la caméra, c’est la vision qui produit le plus d’erreurs anormales. Lors du kidnapping, c’est la proprioception qui produit le plus d’erreur anormales. Lorsqu’une modalité sensorielle est perturbée, c’est bien elle qui produit le plus d’erreur de prédiction anormales.

des lieux visuels qu’il a appris. De même, la navigation proprioceptive est inutilisable après un kidnapping puisque le robot n’intègre pas les déplacements qu’il ne contrôle pas et ses vecteurs retours pointent alors vers de mauvaises destinations. L’évaluation de la fiabilité des différentes modalités sensorielles pourrait donc permettre de moduler les comportements associés à telle ou telle modalité. Bien sûr, il serait tout à fait possible de mettre au point un système bien plus simple pour détecter que la caméra du robot ne reçoit plus de lumière et qu’il ne faut donc pas utiliser la navigation visuelle. Cependant, le but de notre modèle est double. D’une part, nous ne voulons pas traiter séparément chaque modalité sensorielle avec un mécanisme *ad hoc* et spécifique. Le but est de proposer un modèle biologiquement plausible de la détection de nouveauté telle qu’elle est effectuée par l’hippocampe. En plus du formalisme employé inspiré des neurones biologiques, c’est surtout sa généralité qui fait de notre modèle un modèle

plausible biologiquement. C'est de la répétition d'une même architecture de base relativement simple que notre modèle produit sa fonctionnalité principale : la détection de nouveauté. D'autre part, notre modèle permet plus qu'un traitement en parallèle des différentes modalités sensorielles. L'information qu'il permet d'obtenir est bien plus riche. La cohérence intermodale est à la base de l'apprentissage des invariants de la perception et des interactions du robot avec son environnement. Elle pourrait donc être à l'origine de la construction d'un modèle de son propre fonctionnement en interaction avec le monde physique.

Tout au long de ce chapitre, nous avons utilisé de manière quasi interchangeable les notions de détection de nouveauté et de surprise. Cela provient de la simplicité de notre modèle et des possibilités qu'il offre dans nos expériences robotiques. Cependant, la surprise est bien plus que la détection de nouveauté. La surprise est déclenchée par la détection de nouveauté mais elle se définit par un ensemble de modifications comportementales que nous n'avons pas mis en oeuvre dans nos travaux. La surprise, oriente l'attention et mobilise des ressources physiques pour une éventuelle réaction. Enfin la surprise doit être exprimée. C'est ce qui lui donne, comme pour les autres émotions, une dimension sociale puisqu'elle peut alors être utilisée pour communiquer de l'information.

Les Interactions émotionnelles sources de sens

1 Introduction

Les émotions sont des mécanismes indispensables pour acquérir une certaine autonomie comportementale. Cependant, il serait incomplet d'étudier le fonctionnement des émotions uniquement du point de vue d'un être ou d'un robot seul dans son environnement. Les animaux communiquent entre eux par leur expressivité émotionnelle. Pour la plupart, grâce à des vocalisations et leur expressivité faciale. Darwin avait déjà produit un certain nombre de travaux sur ces questions et il défendait la proximité du fonctionnement de cette expressivité émotionnelle entre l'homme et les grands primates. Des recherches ont depuis été menées sur l'expressivité faciale des chimpanzés. Pour éclairer son rôle social [Kraut 1979, Fernández-Carriba 2002, Parr 2003, Parr 2006] mais aussi pour mieux comprendre le développement affectif, cognitif et social des jeunes chimpanzés [Bard 1996, Bard 1999]. L'expressivité d'autres espèces ont aussi été étudiées tels que les chiens [Walker 2010] et les renards [Moe 2006]. Il a fallu attendre une étude étonnamment récente pour mettre précisément en évidence les expressions faciales de douleur chez la souris [Langford 2010]. D'autres mammifères ont été étudiées, mais globalement le nombre d'animaux non étudiés est écrasant. Des liens étroits entre les hommes et les animaux existent depuis plusieurs milliers d'années. Et que ce soit par l'élevage, la domestication, la chasse ou plus généralement par le besoin de connaître son environnement, l'homme a appris à comprendre certains signaux de la communication animale. Si l'on sait qu'un éléphant battant des oreilles va charger, c'est d'abord parce cela a été observé de nombreuses fois mais surtout parce que cette information, dans le contexte, est tout à fait pertinente pour la survie. Cette expressivité non verbale permet aux animaux de communiquer leur état émotionnel entre congénères (construction des liens sociaux) mais aussi entre espèces (e.g. posture et expression agressive pour décourager un prédateur). Chez l'homme, elle est bien entendu un élément crucial de la communication. Sorti du débat sur les émotions primaires et les émotions complexes, il est bien naturel pour notre visage d'exprimer d'innombrables possibilités de nuances et mélanges d'expressions faciales. Cette capacité est pourtant particulièrement complexe. puisqu'elle requiert le contrôle simultané d'un grand nombre de muscles.

Ce chapitre porte sur les interactions émotionnelles Homme-Robot. Aussi bien dans le but de comprendre le fonctionnement et le rôle des émotions que pour permettre à un robot d'interagir non seulement physiquement avec son environnement (perception, action) mais aussi socialement (échange d'informations). Il est capital pour un robot

autonome de pouvoir apprendre par lui-même à extraire les informations pertinentes de son environnement. Il doit cependant pouvoir aussi acquérir des informations grâce aux interactions inter-individuelles. Quelque soit leur niveau de complexité ou de sophistication, ces interactions sont porteuses de sens et constituent donc une forme de communication dont le robot et l'humain peuvent tirer parti. Il existe une grande variété de vecteurs possibles d'interactions émotionnelles : le langage (le contenu autant que la prosodie), les expressions faciales, la posture, ... Un grand nombre d'études ont exploré ces différents modes d'interaction [Bruce 2002, Arkin 2003, Kozima 2005, Breazeal 2000, Breazeal 2001, Gockley 2006, Picard 2002, Takanori 1999, Duffy 2003]. Cependant, quand il est question d'interactions émotionnelles homme-robot, l'accent est souvent mis sur la forme de ces interactions plutôt que sur leur mécanismes de fonctionnement intrinsèques. L'effort est porté sur la production de comportements émotionnels les plus proches d'une situation naturelle de la part du robot pour susciter une réponse émotionnelle de la part de l'homme. Cette démarche joue sur la crédibilité du robot en tant que simulation d'un être vivant. Elle permet d'étudier les comportements des hommes participants à ce genre d'interactions. Certains robots expressifs ont notamment été utilisés dans le cadre d'expérimentations avec des enfants atteints d'autisme, notamment dans l'espoir de mettre au point de nouveaux moyens thérapeutiques [Nadel 2004, Nadel 2006, Robins 2005, Dautenhahn 2004]. Mais la démarche constructiviste qui qualifie ce travail fait approcher la question des interactions émotionnelles homme-robot par les situations et les modèles les plus simples avant de s'intéresser à des interactions émotionnelles plus sophistiquées. Alors qu'est-ce qu'une interaction émotionnelle minimale ? C'est justement le propos de la première partie de ce chapitre de montrer que n'importe quelle stimulation, aussi simple soit-elle, aussi neutre soit-elle, peut servir de base à une interaction émotionnelle (e.g. appuyer sur un capteur tactile du robot ou plus simplement un bouton). Ce qui est primordial, c'est la capacité à donner un sens à cette stimulation. Cet apprentissage peut venir de l'interaction entre l'homme et le robot. Le robot doit exprimer son état interne de douleur ou de plaisir (réflexe). L'homme qui interagit avec lui peut donc donner un sens à une stimulation neutre en la répétant en même tant que le robot exprime sa douleur ou son plaisir. Encore une fois, cette acquisition de sens peut venir d'un simple mécanisme de conditionnement. Cependant, ce conditionnement se distingue du conditionnement pavlovien classique. Nous approfondirons ce point dans la suite de ce chapitre. Une fois conditionnée, cette stimulation provoque alors une douleur ou un plaisir conditionné qui peut servir de signal respectivement négatif ou positif. Le même raisonnement est bien sûr similaire pour un conditionnement avec la sensation de plaisir. Dès lors qu'une stimulation neutre a acquis un sens, l'homme peut utiliser cette stimulation pour donner au robot des informations sur son environnement ou son comportement. Il s'agit ici d'un type d'interaction dans lequel l'homme a de meilleures capacités de perception que le robot et possède donc une connaissance de l'environnement qui lui serait utile. En stimulant positivement ou négativement le robot, l'homme peut donc lui permettre d'en apprendre sur son environnement ou son comportement. Cette communication très simple possède des similarités avec la méthode de la carotte et du bâton. Deux stimulations de sens opposées permettent d'apprendre à adapter

son comportement. Dans notre cas cependant, un conditionnement permet de donner le sens de la carotte ou du bâton à une stimulation neutre. La situation est donc mieux décrite par un conditionnement classique : la réaction d'un chien à la vue du journal roulé (associé à un coup) ou de la laisse (associée à la promenade). Ces objets neutres prennent leur sens par associations répétées.

Dans la suite de ce chapitre, le robot est équipé d'une tête mécanique permettant de produire des expressions faciales. Une caméra vidéo installée dans un œil permet à la tête de voir ce qui se trouve en face d'elle. Un modèle neuronal développé au laboratoire Etis permet la reconnaissance d'expressions faciales émotionnelles [Gaussier 2004, Gaussier 2007b, Prepin 2008, Boucenna 2008, Andry 2008]. Les travaux utilisant la reconnaissance d'expressions faciales présentés dans cette étude ont donc été menés conjointement avec Sofiane Boucenna, un autre doctorant du laboratoire. Il s'agit d'un modèle de l'apprentissage de la reconnaissance d'expressions faciales lors du développement des jeunes humains. Cet apprentissage repose sur le mécanisme de résonance émotionnelle. Il s'agit de l'imitation involontaire d'une expression faciale émotionnelle reconnue. Cette résonance émotionnelle est particulièrement prononcée lors des interactions entre le bébé et ses parents. Quand le bébé est content ou se sent bien, et qu'il sourit à sa mère, elle lui sourit en retour. Un conditionnement peut alors être appris entre la perception de l'état interne qui a provoqué son expression faciale et la perception visuelle correspondant à l'expression faciale de la personne qui interagit avec lui. Ce modèle a été utilisé pour permettre au robot d'apprendre de manière interactive à reconnaître des expressions faciales émotionnelles. Avec ce moyen d'interaction plus sophistiqué et surtout plus naturel, la question qui sera approfondie est la question de la signification pour le robot des expressions qu'il reconnaît. Fondamentalement, cette question revient à déterminer une forme de référencement social pour le robot. Ce problème complexe est traité du point de vue d'un robot dont l'univers des possibles est contraint à la locomotion et l'expressivité.

2 Toute stimulation peut prendre un sens émotionnel

2.1 Principe et fonctionnement

Qu'est-ce qu'une interaction émotionnelle minimale et comment le robot peut-il en tirer un bénéfice ? Si le robot peut exprimer et donc communiquer son état interne de douleur ou de plaisir, l'homme qui interagit avec lui peut stimuler le robot de manière sélective afin de conditionner cette stimulation avec l'état de douleur ou de plaisir du robot. Cette stimulation initialement non spécifique acquiert donc un sens. Elle peut alors être utilisée pour qualifier le comportement du robot [Hasson 2010b]. C'est ce qu'illustre la figure 8.1. Toutefois, le type de conditionnement impliqué dans le mécanisme que nous allons décrire diffère du cadre de description du conditionnement classique (pavlovien). Dans le cadre pavlovien, un stimulus inconditionnel (injection de nourriture dans la bouche) provoque la réponse inconditionnelle (salivation). Un stimulus neutre (le son de la clochette) est présenté en même temps que le stimulus inconditionnel et devient un stimulus conditionné. Il produit alors la réponse condi-

tionnée (salivation) identique à la réponse inconditionnelle. Dans notre cas, le stimulus neutre est l'activation du capteur non spécifique. Le stimulus inconditionnel est le signal de plaisir ou de douleur et la réponse inconditionnelle est la génération d'un potentiel de récompense ou de punition. En présentant simultanément le stimulus neutre et le stimulus inconditionnel, le conditionnement est appris, l'activation du capteur non spécifique devient un stimulus conditionné qui déclenche alors la réponse conditionnée : une récompense ou une punition conditionnée. Au lieu de provoquer une réponse comportementale directe (comme la salivation) nous considérons que la réponse conditionnée peut être un potentiel de récompense ou de punition.

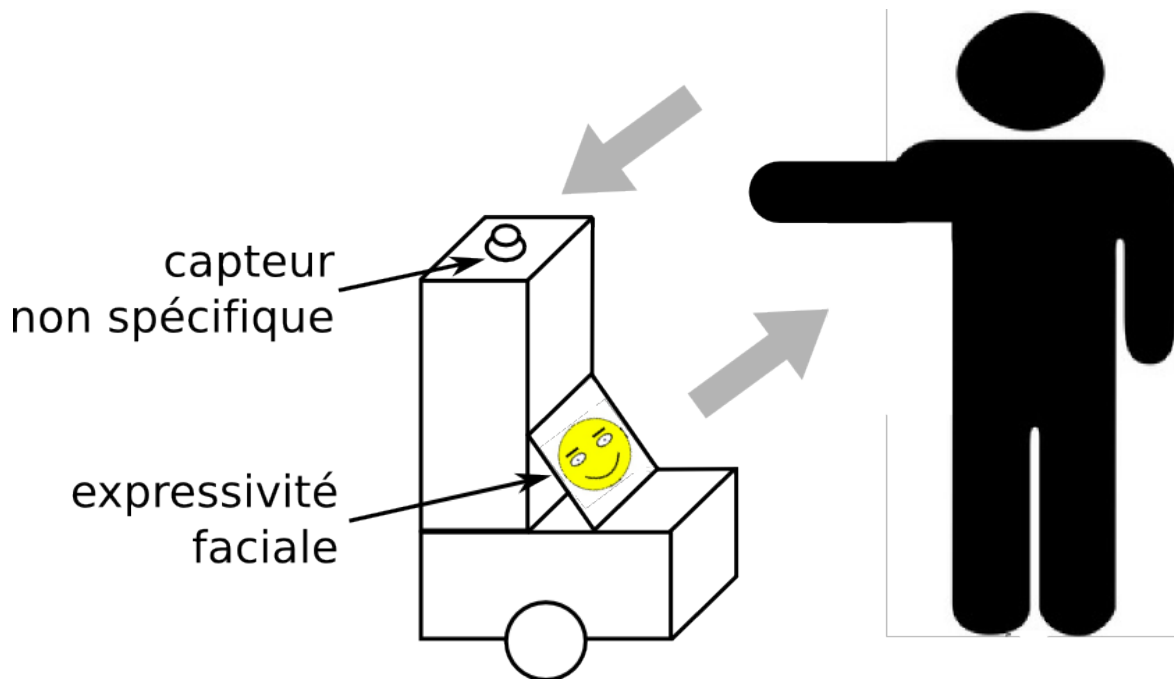


FIGURE 8.1 – Interaction émotionnelle minimale. Le robot dispose de la capacité d'exprimer son état interne de douleur ou de plaisir et dispose d'un capteur non spécifique (un bouton). Si l'homme active sélectivement le capteur quand le robot exprime un état interne particulier, un conditionnement peut être appris pour associer l'activation du capteur avec le signal de douleur ou de plaisir. Une fois conditionnée, l'activation de ce capteur par l'homme peut alors servir à donner au robot un signal positif (récompense) ou négatif (punition).

Pour être minimale, il ne faut pas que la stimulation utilisée soit spécifique. Par exemple, si le robot possède deux capteurs prédéfinis de "punition" et de "récompense", une grande part du problème est alors ignorée. Au plus bas niveau, il faut comprendre les questions d'acquisition du sens des signaux d'interaction. C'est pourquoi un seul capteur est utilisé. Ce capteur ne dispose pas de sens prédéfini. C'est pourquoi il s'agit d'un capteur non spécifique. Il peut être stimulé sélectivement par l'homme afin de lui donner un sens spécifique pour le robot. La capacité du robot à exprimer son état interne est primordiale. C'est le signal qui permettra à l'expérimentateur de choisir le sens qu'il souhaite donner à la stimulation du capteur. La

stimulation du capteur est alors le stimulus conditionnel et les signaux de douleur et de plaisir les stimuli inconditionnels. La figure 8.2 montre le réseau de neurones utilisé pour apprendre à conditionner le capteur non spécifique.

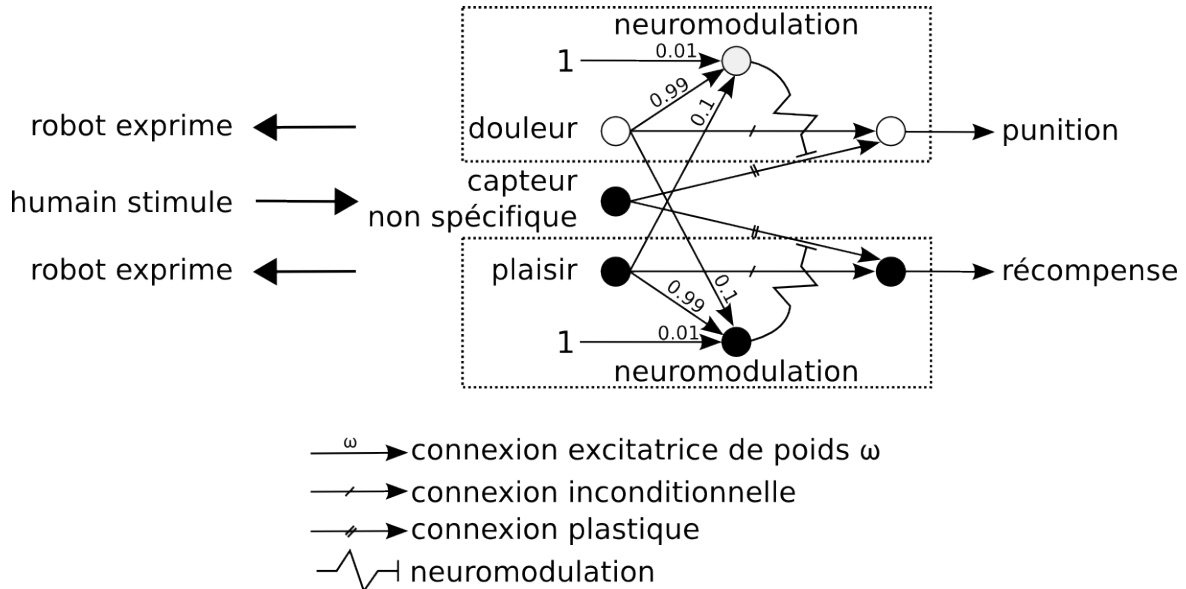


FIGURE 8.2 – Réseau de neurones utilisé pour conditionner le capteur non spécifique. Le conditionnement du capteur peut être appris aussi bien avec le signal de douleur que de plaisir. Ce conditionnement peut aussi être modifié dynamiquement. Le conditionnement initial est alors oublié pour être remplacé par le nouveau conditionnement.

Ce réseau permet de conditionner le capteur aussi bien avec la douleur qu’avec le plaisir. De plus, il permet un apprentissage souple pouvant être modifié dynamiquement en fonction des interactions entre l’homme et le robot. La stimulation du capteur non spécifique est représentée par l’activation d’un neurone. Ce neurone est relié à deux neurones via de connexions plastiques (signal conditionnel) utilisant la règle d’apprentissage de Widrow et Hoff. Ces deux neurones reçoivent respectivement l’activation d’un neurone représentant le douleur et d’un neurone représentant le plaisir (signaux inconditionnels). La neuromodulation λ de l’apprentissage de ces conditionnements est conçue de tel sorte que la spécification du capteur spécifique est rapide (pour un apprentissage initial comme pour un changement de conditionnement du capteur), tandis que l’oubli est lent. C’est le différentiel de vitesse d’apprentissage des conditionnements quand la douleur et le plaisir sont présents ou non qui permet de moduler le comportement de ce réseau. En présence de douleur ou de plaisir, le réseau apprend vite. L’oubli correspond à l’apprentissage de l’absence de lien systématique entre le capteur et la douleur ou le plaisir. Mais si cet oubli n’est pas suffisamment lent, le conditionnement n’est utilisable que très peu de temps.

La figure 8.3 montre les différents signaux de ce réseau lors de son fonctionnement. Le capteur est d’abord stimulé alors que le robot exprime son état interne de plaisir.

Le conditionnement est appris et la stimulation du capteur produit une réaction de plaisir conditionné qui peut être utilisé comme un signal de récompense. Ce plaisir conditionné est oublié lentement (neuromodulation λ faible). Quand le capteur est de nouveau activé lorsque le robot exprime du plaisir, le conditionnement est renforcé. Si maintenant le capteur est activé quand le robot exprime de la douleur, le conditionnement avec le plaisir est oublié rapidement et le conditionnement avec la douleur est appris rapidement lui aussi (neuromodulation λ forte). La stimulation du capteur produit alors un signal de douleur conditionnée pouvant être utilisé comme un signal de punition.

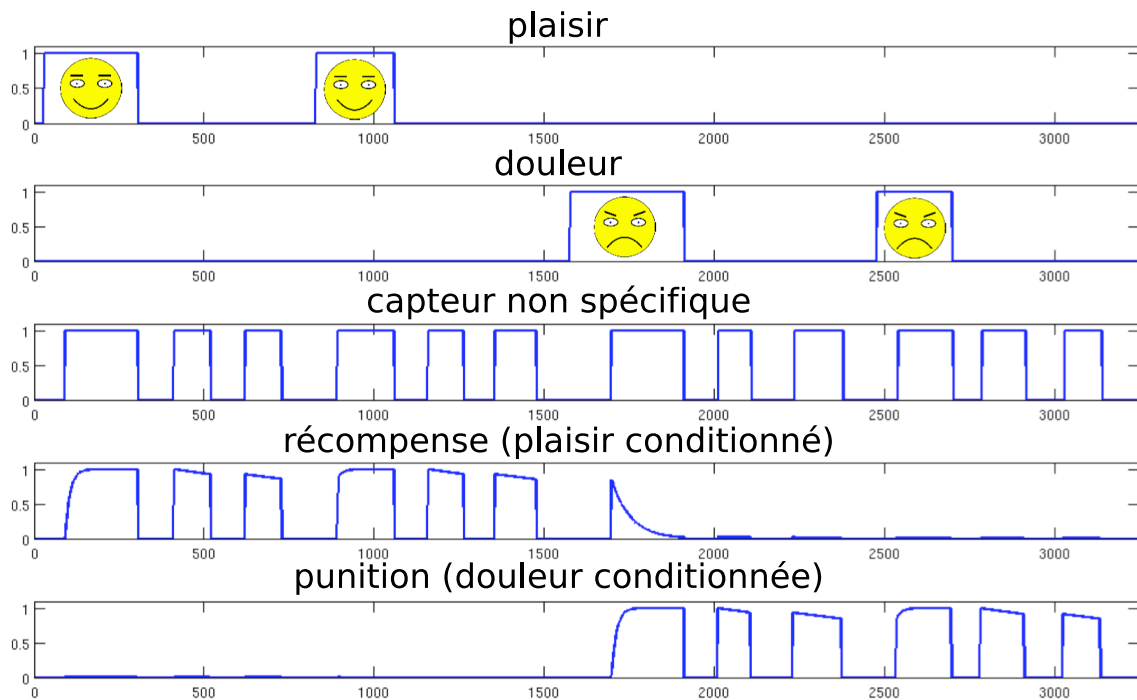


FIGURE 8.3 – Signaux du réseau de neurone utilisé pour l'apprentissage du conditionnement du capteur non spécifique. Quand l'homme stimule la capteur alors que le robot exprime le plaisir qu'il ressent, le robot apprend le conditionnement de ces deux stimulations et la stimulation du capteur provoque alors un plaisir conditionné (récompense). Mais le sens pris par la stimulation du capteur peut être modifié dynamiquement. Quand l'homme stimule le capteur en même temps que le robot exprime la douleur qu'il ressent, le précédent conditionnement est oublié et la stimulation du capteur provoque maintenant une douleur conditionnée (punition).

2.2 Expérimentation robotique

Pour l'expérimentation robotique de ce set-up, un écran a été disposé sur le robot pour lui permettre d'exprimer avec des expressions faciales stéréotypées son état interne de plaisir ou de douleur. Le capteur non spécifique correspond à l'activation

d'un bouton installé lui aussi sur le robot. La figure 8.4 montre le robot utilisé pour cette expérience dans son environnement.



FIGURE 8.4 – Le robot dans son environnement.

Pour l'expérimentateur, le but de cette expérience est de conditionner le capteur en interagissant avec le robot. Ensuite, l'expérimentateur doit se servir de ce capteur pour apprendre au robot à rejoindre un lieu particulier de son environnement. Cette expérience repose donc sur le couplage entre les signaux de récompense ou de punition venant du capteur et les réseaux de neurones dédiés à la navigation visuelle. Cependant, le modèle de navigation sensori-motrice présenté précédemment ne repose que sur un seul type d'apprentissage : l'association entre une perception visuelle (cellule de lieu) et une action (direction à prendre). En reconnaissant le lieu en question, le robot prend la direction conditionnée. Pour déclencher l'apprentissage de ce comportement, le robot peut utiliser le signal de récompense. Ce signal de récompense fait alors fonction de renforcement positif. Quand dans un lieu donné le robot prend une direction considérée comme bonne par l'expérimentateur, le signal de récompense lui permet d'apprendre à reprendre cette direction la prochaine fois qu'il revient dans ce lieu. Mais le capteur peut aussi être conditionné pour donner des signaux de punition. Cette punition peut être utilisée pour apprendre à inhiber la direction courante. Pour apprendre à naviguer à base d'associations sensori-motrices, le robot pourrait alors utiliser aussi bien des associations positives (activation) que négatives (inhibition). Au lieu d'une voie de conditionnement simple, le système de navigation doit reposer sur deux voies de conditionnement : une voie activatrice (champs lieu-action +) et une voie inhibitrice (champs lieu-action -). La sortie de ce réseau, la direction à prendre, est un champs recevant des connexions activatrices depuis le champs lieu-action + et des connexions inhibitrices depuis le champs lieu-action -. La figure 8.5 montre cette architecture sensori-motrice.

La tâche de navigation que le robot doit apprendre consiste à rejoindre un lieu de l'environnement : le lieu but. Un bassin d'attraction a été appris pour permettre au robot de converger vers le but. Cependant, dès que le robot sort du bassin d'attraction, il est perdu. La figure 8.6 montre les trajectoires du robot lorsqu'il est dans le bassin d'attraction (en direction du but) et quand il en est en dehors (mouvements aléatoires).

En interagissant avec le robot, l'expérimentateur doit permettre au robot d'ap-

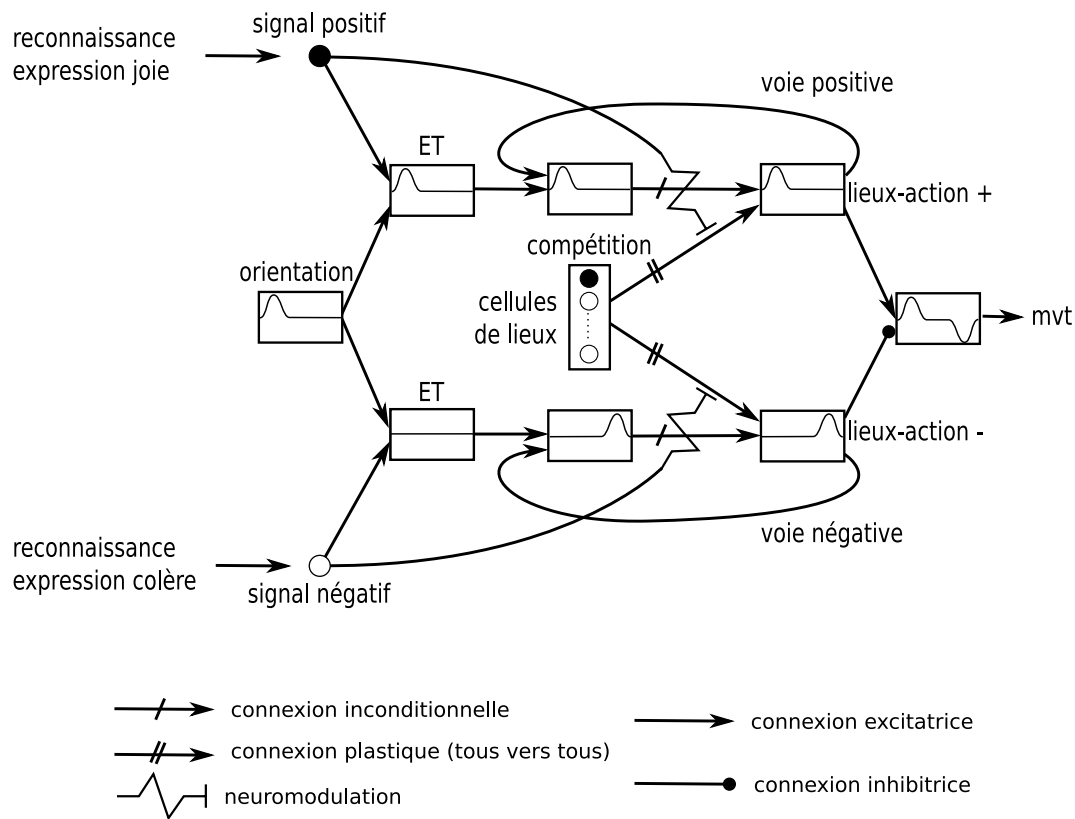


FIGURE 8.5 – Architecture sensori-motrice assurant l'apprentissage de conditionnements positifs et négatifs. La voie positive permet d'apprendre à activer une action (champs lieu-action +) tandis que la voie négative permet d'apprendre à inhiber une action (champs lieu-action -). Dans cet exemple, le lieu reconnu a été précédemment associé à l'inhibition d'une direction. Le robot est cette fois-ci dans une autre direction et un signal de récompense est reçu. Ce lieu est désormais aussi associé à l'activation de cette nouvelle direction. Ce couplage entre les signaux émotionnels et les informations sensorielles du robot sera par la suite décrit comme étant le couplage comportemental.

prendre à étendre ce bassin d'attraction primaire pour pouvoir rejoindre le but depuis de plus grandes distances. Les figures 8.7 et 8.8 montrent les trajectoires du robot après que les interactions avec l'expérimentateur lui ait permis d'apprendre à étendre sa capacité à converger vers le but.

Pour la première expérience, le capteur non spécifique a été conditionné avec le signal de plaisir et l'expérimentateur donne ainsi des signaux de renforcement positif. Au contraire, dans la seconde expérience, le capteur est conditionné avec le signal de douleur. Ce n'est donc qu'en apprenant ce qu'il ne faut pas faire que le robot doit rejoindre le but. Bien que cela soit moins efficace, ces conditionnements permettent tout de même au robot de converger vers le but. Mais ce processus tient de la prise de décision par élimination. C'est pourquoi il est moins efficace pour ce type de tâche de navigation. De manière générale, rejoindre un but demande de savoir ce qu'il faut faire. Si par contre, la tâche de navigation consiste à être capable d'éviter une zone de l'environnement, cela demande de savoir ce qu'il ne faut pas faire. Cette situation

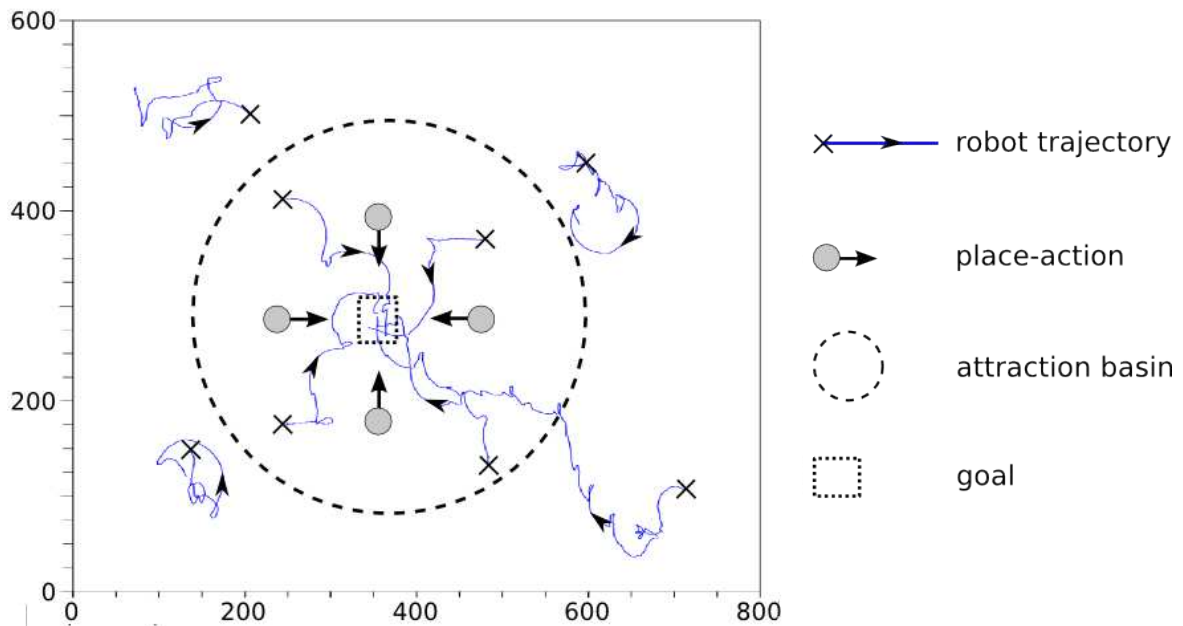


FIGURE 8.6 – Des associations sensori-motrices sont apprises autour du but créant un bassin d'attraction. Depuis l'intérieur du bassin d'attraction, le robot converge vers le but. En dehors, il navigue aléatoirement.

représente bien le cas où une zone de l'environnement contient un danger. Par exemple, s'il y a un trou dans le sol. A proximité du trou, il n'y a pas nécessairement de "bonne" direction à prendre. Par contre il y a bien une direction à ne pas prendre : celle du trou. Pour que l'on puisse parler d'interactions émotionnelles homme-robot, aussi minimales soient-elles, il faut que certaines informations essentielles soient accessibles pour permettre d'apprendre une base commune de communication. Ceci demande à l'une des parties au moins de pouvoir exprimer son état émotionnel (dans notre cas le robot) afin que l'autre puisse s'en servir pour donner un sens à une stimulation neutre.

2.3 Point de vue neurobiologie

La capacité à donner un sens à une stimulation non spécifique peut être utilisée par le robot comme une source d'information pour améliorer son comportement. Ce mécanisme d'apprentissage par interaction est compatible avec les études neurobiologiques du conditionnement émotionnel. Un grand nombre d'étude ont montré l'implication de l'amygdale pour les conditionnements émotionnels [Grossberg 2008]. Aussi bien pour les conditionnements de comportement aversifs [Dunsmoor 2009, Blair 2005, Knight 2005, Maren 2005] qu'appétitifs [Baxter 2002, Cardinal 2002, Balleine 2003, Balleine 2006]. Des études anatomiques ont aussi montré que l'amygdale reçoit des connexions transportant l'information de stimulations aversives comme appétitives [McDonald 1998, Pitkanen 2000]. Le rôle principal de l'amygdale est de donner une valeur positive ou négative aux stimulations entrantes [Paton 2006]. Des modèles de

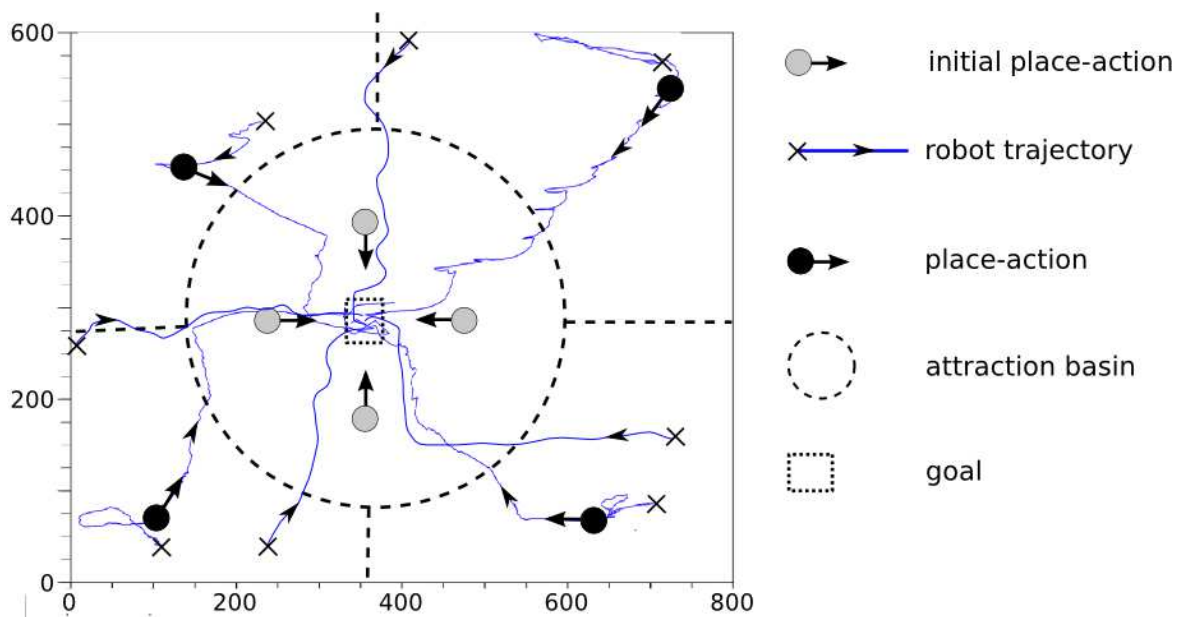


FIGURE 8.7 – Trajectoires du robot quand il est perdu mais apprend via ses interactions avec l'expérimentateur à rejoindre le but. Le capteur est associé avec l'état de plaisir. Le robot est placé en dehors du bassin d'attraction et navigue donc aléatoirement. Quand le comportement du robot est considéré comme "bon" (en direction du but), l'expérimentateur stimule la capteur et le robot apprend à renforcer la direction courante.

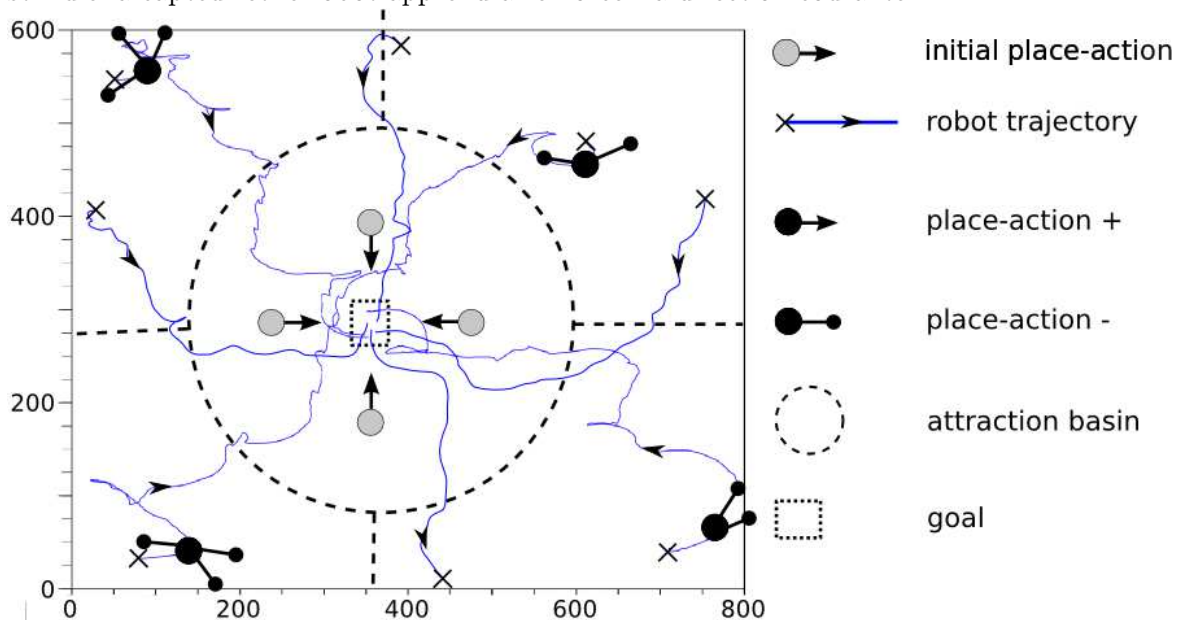


FIGURE 8.8 – Cette fois-ci, le capteur est associé avec l'état de douleur. Le robot est placé en dehors du bassin d'attraction et navigue donc aléatoirement. Quand le comportement du robot est considéré comme "mauvais" (pas en direction du but), l'expérimentateur stimule la capteur et le robot apprend à inhiber la direction courante.

conditionnements émotionnels réalisés par l'amygdale ont été simulés sur robot virtuel en environnement simulé [Balkenius 2009, Mannella 2008, Moren 2000]. Notre modèle de conditionnement émotionnel utilisant un capteur non spécifique est inspiré de ces travaux. Il correspondrait au fonctionnement de l'amygdale baso-latérale qui permet l'apprentissage de conditionnements stimulus-stimulus (capteur non spécifique et douleur ou plaisir dans notre cas). La figure 8.9 montre comment l'architecture de contrôle du robot peut être comprise en termes de réseau de structures cérébrales. L'amygdale reçoit des connexions depuis différents centres perceptifs et projette des connexions vers l'aire tegmentale ventrale qui projette à son tour vers l'amygdale (apprentissage de conditionnements), le para-hippocampe (caractérisation de la scène visuelle en combinaison d'amers visuels à des azimuts spécifiques), le cortex entorhinal (catégorisation de la perception visuelle en cellule de lieu) et le nucleus accumbens (apprentissage sensori-moteur).

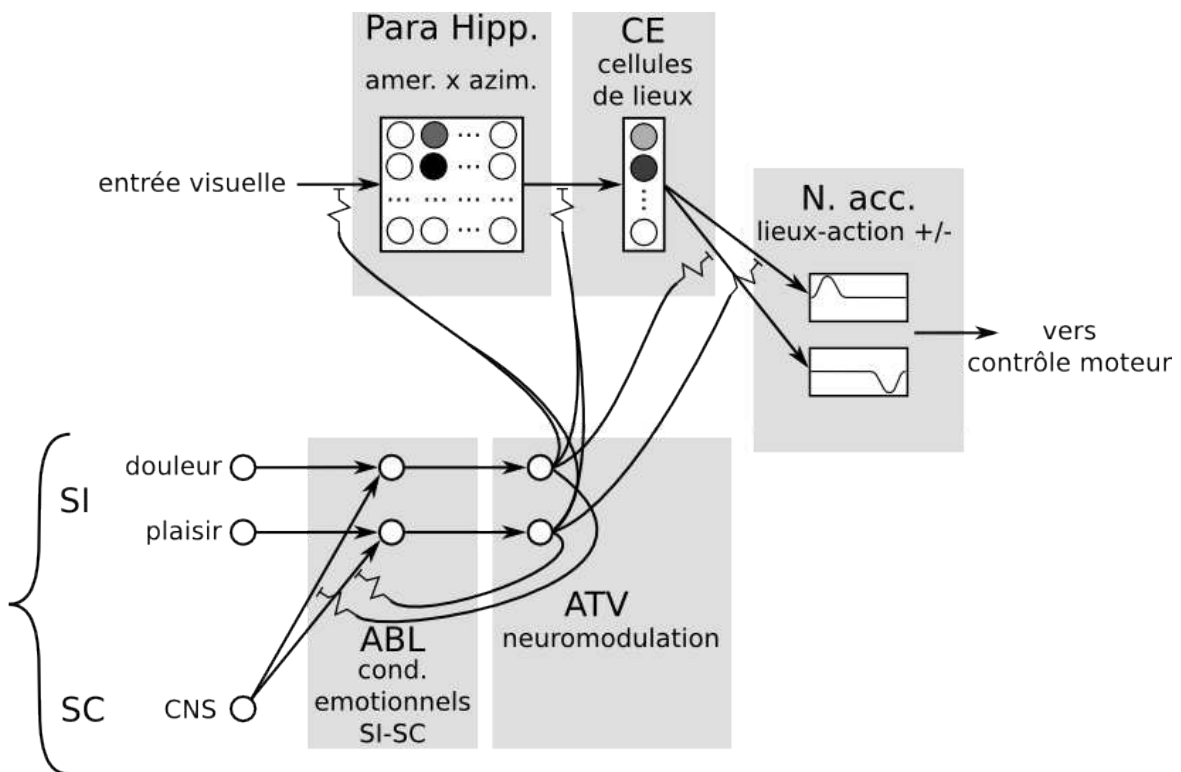


FIGURE 8.9 – L'architecture de contrôle du robot peut être comprise comme un réseau impliquant différentes structures cérébrales. L'amygdale baso-latérale apprend les différents conditionnements du capteur non spécifique avec la perception de plaisir ou de douleur. L'aire tegmentale ventrale neuromodule l'apprentissage de ce conditionnement ainsi que l'apprentissage du conditionnement sensori-moteur. Dans le para-hippocampe (vision bas niveau), puis le cortex entorhinal (cellules de lieu) et enfin dans le nucleus accumbens (aspect prédictif de récompense : associations sensori-motrices positives et négatives). La sortie du nucleus accumbens est l'action à exécuter. Elle est transmise au contrôle moteur.

3 L'expressivité faciale et la reconnaissance d'émotion

Pour remplacer l'écran qui était placé sur le robot, une tête mécanique a été installée. Cette tête équipée de 15 servomoteurs présente les caractéristiques minimales nécessaires pour produire des expressions faciales émotionnelles humaines. La bouche peut se déformer, s'ouvrir et se fermer et les sourcils peuvent s'orienter, se plier et se hausser. Une caméra vidéo installée dans un des yeux permet à la tête de voir ce qui se trouve en face d'elle. La figure 8.10 montre la tête de robot produisant les expressions de joie, de colère, de tristesse et de surprise. Un modèle neuronal développé au laboratoire [Gaussier 2004, Gaussier 2007b, Boucenna 2008] permet au robot de reconnaître des expressions faciales émotionnelles (figure 8.11). Il s'agit d'un modèle de l'apprentissage de cette compétence lors du développement des jeunes humains. Ces travaux ont été développés conjointement avec le groupe de Jacqueline Nadel du Centre Emotion de la Pitié-Salpêtrière. L'apprentissage de la reconnaissance d'expressions faciales repose sur le mécanisme de résonance émotionnelle. Il s'agit de l'imitation involontaire d'une expression faciale émotionnelle. Cette résonance émotionnelle est particulièrement prononcées lors des interactions entre le bébé et ses parents. En résonnant avec le bébé, les parents se comportent essentiellement comme des miroirs, permettant de refléter au bébé ses propres expressions faciales. Il peut donc apprendre par conditionnement à associer la perception de l'état interne qui a provoqué son expression faciale et la perception visuelle correspondant à l'expression faciale de la personne qui interagit avec lui. En revoyant cette expression faciale, cet apprentissage produit un état interne conditionné spécifique pour chaque expression faciale. Cet état interne conditionné peut être interprété comme une résonance de bas niveau cognitif (ou pseudo-empathie). Bien sûr il ne faut pas comprendre ici empathie comme la capacité à "comprendre" les sentiments et les émotions d'un autre. Encore moins dans sa vision plus cognitive concernant les états mentaux non-émotionnels comme les croyances. Il s'agit plus d'une résonance de bas niveau : le bébé ressent un état interne apparenté à l'état qui produit chez lui l'expression qu'il voit chez l'autre. C'est cette réaction empathique conditionnée qui permet de parler de reconnaissance. Ce modèle a été utilisé pour permettre au robot d'apprendre de manière interactive à reconnaître et à imiter des expressions faciales émotionnelles.

4 Vers un référencement social

L'architecture de contrôle est conçue de manière à permettre au robot d'apprendre des comportements d'attraction et de répulsion vis-à-vis des sources de récompense et de punition présentes dans son environnement. Sa tête lui permet d'interagir avec un homme en exprimant et en reconnaissant des expressions faciales émotionnelles. La situation expérimentale mise en place part du principe que lors de ces interactions, l'homme doit aider le robot. En exprimant sa joie ou sa colère vis-à-vis de la situation, l'homme transmet au robot des signaux positifs ou négatifs [Hasson 2010a]. Nous utilisons précédemment les signaux de douleur et de plaisir pour donner une valence émotionnelle aux conditionnements amygdaliens. A présent, nous utiliserons

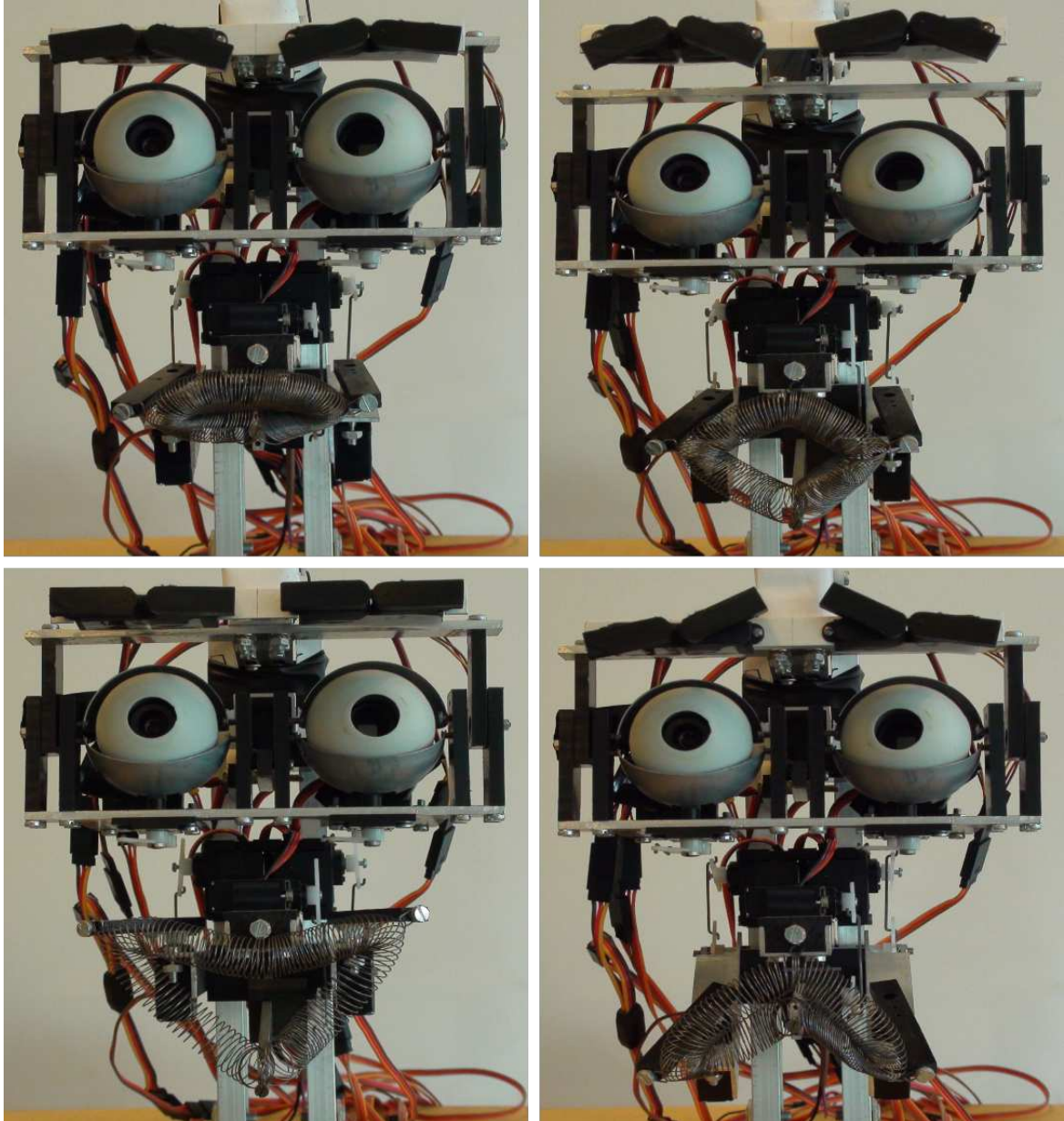


FIGURE 8.10 – Expressions faciales de la tête du robot. De gauche à droite et de haut en bas : expression neutre, surprise, joie et tristesse. Seules les caractéristiques expressives principales sont utilisées : la forme de la bouche ainsi que le haussement et le froncement des sourcils.

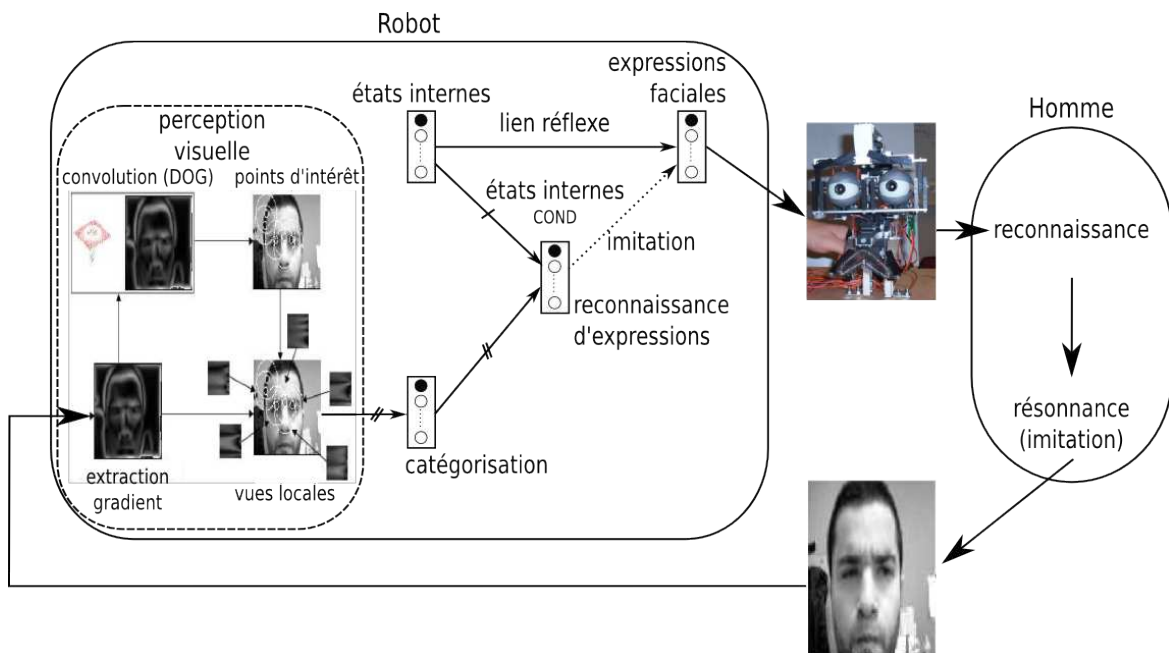


FIGURE 8.11 – Modèle neuronal permettant au robot d'apprendre à reconnaître des expressions faciales émotionnelles en interagissant avec un homme qui l'imitate. Quand un état interne du robot produit de manière réflexe une expression faciale, l'homme l'imitate et le robot perçoit visuellement un pattern de points d'intérêt qui caractérise l'expression faciale de l'homme. En associant la reconnaissance de ce pattern visuel à l'état interne qui a provoqué l'expression, le robot apprend un conditionnement lui permettant désormais de donner une valeur émotionnelle aux expressions faciales qu'il reconnaît. Si l'état interne conditionné par la reconnaissance de l'expression de l'homme est utilisé pour produire l'expression faciale correspondante à l'état interne authentique, le robot peut alors à son tour imiter les expressions de l'homme.

les signaux conditionnés produit par le reconnaissance d'une expression faciale positive ou négative. Ce protocole expérimental est inspiré des recherches menées sur le référencement social. Le référencement social est un concept issu de la psychologie développementale qui décrit la capacité à reconnaître, comprendre et modifier son comportement en réaction à l'expressivité émotionnelle d'un partenaire social [Klinnert 1983, Walden 1988, Russell 1997]. Le référencement social est étudié chez les humains aussi bien que chez les animaux (notamment chez les chimpanzés). Etant essentiellement basées sur des informations non verbales et ne demandant pas de compétences cognitives haut niveau, ces interactions semblent être un moyen rapide et efficace pour apprendre les caractéristiques de son environnement lors des étapes précoces du développement cognitif. La question n'a pas été abordée plus tôt mais, du point de vue du robot, il existe une réelle ambiguïté quand à la signification ou destination de ces signaux d'interactions : à quoi ce réfère ce référencement social ?

Nous proposons que dans un tel contexte, le signal d'interaction peut être utilisé par l'amygdale pour donner une valeur émotionnelle au contexte perceptif (i.e. condition-

nement d'une cellule de lieu à un signal positif ou négatif). Mais il peut aussi être utilisé par le *nucleus accumbens* pour apprendre le potentiel de récompense d'une action dans un contexte donné (i.e. conditionnement d'une association lieu-action à un signal positif ou négatif). Ces deux possibilités d'apprentissage peuvent produire un comportement similaire mais via des méthodes différentes et sous-tendues par des structures cérébrales distinctes. Selon notre modèle, l'expression de l'homme peut concerner soit l'environnement du robot (conditionnement de l'amygdale), soit son comportement (conditionnement du *nucleus accumbens*).

Considérons dans un premier temps la cas du référencement à l'environnement du robot. Le robot peut par exemple être sur l'emplacement d'une ressource ou au contraire à proximité d'une source de danger. Le signal d'interaction vient dans ce cas plutôt comme un avertissement. Il s'agit alors du couplage environnemental supporté par l'amygdale. Les sensations visuelles sont interprétées sous la forme d'une cellule de lieu gagnante. Celle-ci est associée par conditionnement à une valeur émotionnelle positive ou négative. En s'approchant du lieu conditionné, une réponse conditionnée (positive ou négative selon le cas) sera alors générée. Nous développerons par la suite comment cette réponse conditionnée permet au robot d'orienter son comportement de navigation.

Plutôt que de qualifier l'environnement du robot, l'interaction émotionnelle peut aussi se référer à son comportement. L'homme peut se servir de son expressivité faciale pour renforcer un comportement du robot. Au contraire, il peut aussi apprendre au robot qu'un comportement est mauvais et qu'il faut l'éviter. En somme, le robot fait ce qu'il faut ou se trompe et l'interaction expressive vient alors comme une punition ou comme une récompense. Il s'agit du couplage comportemental permis par le *nucleus accumbens*. C'est ce type d'interaction qui a été utilisée dans la première partie de ce chapitre lors de l'utilisation du conditionnement du capteur non spécifique. Le *nucleus accumbens* permet d'apprendre le potentiel de récompense des différentes actions possibles dans un contexte perceptif donné. C'est donc lui qui associe un potentiel de récompense à l'action en cours dans le contexte du moment. Autrement dit, il s'agit du modèle présenté dans la figure 8.5. Les lieux-actions prennent soit une valeur positive soit une valeur négative.

Pour interagir avec le robot pour des tâches de navigation, la tête a été montée sur le robot. La figure 8.12 montre ce set-up expérimental.

4.1 Couplage comportemental

Le couplage comportemental consiste à utiliser les signaux venant de l'interaction pour apprendre à renforcer ou inhiber les comportements du robot. Plus précisément, cela veut dire que ces signaux interviennent lors de l'apprentissage des conditionnements sensori-moteurs qui orientent les comportements de navigation. Comme décrit précédemment, puisque la joie est l'expression d'un état interne positif, la reconnaissance d'une expression faciale de joie provoque un signal positif. De même, une expression de colère provoquera un signal négatif. Le modèle présenté dans la figure 8.5 et déjà expliqué précédemment décrit le couplage comportemental et son fonctionnement. Un signal positif permet d'apprendre à renforcer l'association d'une information senso-

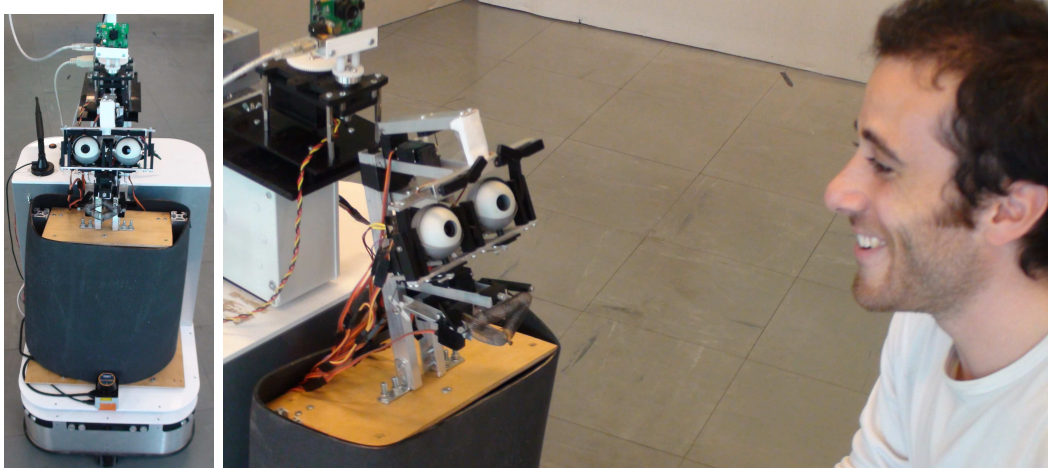


FIGURE 8.12 – La plate-forme mobile est équipée de la tête expressive. Le robot peut ainsi naviguer mais aussi interagir avec l'expérimentateur en produisant et en reconnaissant des expressions faciales émotionnelles.

rielle et d'une action. Ce couplage peut donc être utilisé pour apprendre au robot des tâches de navigation. En apprenant par cette interaction à renforcer des lieux-action convergeant vers un but, le robot peut construire un bassin d'attraction sensori-moteur autour du but. De même, en apprenant à inhiber ces mêmes lieux-action, le robot construit un bassin de répulsion. Ce couplage offre donc des possibilités de guidage des apprentissages du robot pour ses comportements d'attraction et de répulsion vers les différentes particularités de son environnement.

Plus concrètement, lorsque le robot va dans une direction qui est considérée comme "bonne" par l'homme (e.g. en direction d'un but), en faisant une expression de joie, l'homme permet au robot d'apprendre à associer de manière activatrice cette direction avec le lieu dans lequel il se trouve. En faisant au contraire une expression de colère, le robot apprend une association inhibitrice entre cette direction et le lieu dans lequel il se trouve. La figure 8.13 précise le fonctionnement de ce couplage entre reconnaissance d'expressions et système de navigation.

Les signaux d'interaction font office de signaux de neuromodulation des apprentissages des conditionnements sensori-moteurs. En ce sens, on pourrait aussi parler de couplage direct entre les capacités d'interactions émotionnelles et la capacités de navigation. Puisqu'il permet d'associer une stimulation à une action source de récompense (obtention d'un but ou évitement d'un danger), ce conditionnement est assimilable à un conditionnement opérant (ou instrumental) tel que décrit par Thorndike [Thorndike 1901] puis Skinner [Skinner 1953]. Cependant, il subsiste une différence importante. Pour initier cet apprentissage, ce ne sont pas les conséquences directes de l'action à renforcer ou à inhiber qui fournissent le signal de récompense ou de punition, mais les signaux d'interactions émotionnelles venant du partenaire social du robot : l'expérimentateur humain. Cette boucle d'apprentissage entre le robot, son environnement et l'humain bienveillant aide le robot à apprendre plus rapidement. L'expérimentateur prend la position d'un professeur qui récompense et punit le robot

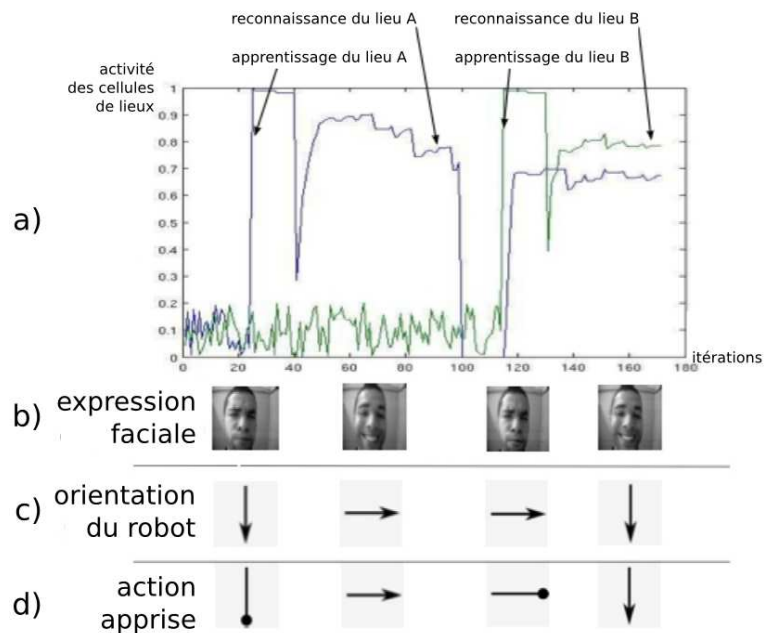


FIGURE 8.13 – a) signal des cellules de lieux. b) expression faciales de l’humain reconnue par le robot. c) orientation courante du robot. d) action apprise par le robot (une flèche signifie que la direction est activée, une boule que la direction est inhibée). Les expressions faciales de l’humain permettent au robot d’apprendre à associer une perception (reconnaissance visuelle d’un lieu) à une action à activer ou inhiber (direction à prendre ou à éviter).

en fonction de son comportement. La figure 8.14 montre pour ces deux cas de figure les lieu-action positifs et négatifs appris par le robot ainsi que sa trajectoire convergente ou divergente depuis plusieurs points de départ différents.

4.2 Couplage environnemental

Le couplage environnemental consiste à utiliser les signaux de l’interaction expressive pour apprendre au robot à qualifier les différents lieux de son environnement. Il s’agit donc de donner une valeur émotionnelle directement à la reconnaissance d’une perception. Quand l’humain présente une expression de joie au robot, cela veut dire que le lieu dans lequel se trouve le robot est "bon". Autrement dit, le lieu dans lequel se trouve le robot contient un but (e.g. une ressource nécessaire au robot). Au contraire, une expression de colère signale un "mauvais" lieu. Cet apprentissage permet de donner une valeur intrinsèque à un lieu de l’environnement. En reconnaissant une expression faciale émotionnelle, le robot apprend à caractériser le lieu courant et apprend un conditionnement entre la cellule de lieu correspondante et le signal d’interaction (expression de joie ou de colère). Ainsi, la reconnaissance de ce lieu provoque un signal conditionné positif ou négatif en fonction du conditionnement appris. Mais pour orienter son comportement de navigation, le robot doit utiliser non pas le fait que cette réponse conditionnée soit présente ou non mais la manière dont elle évolue en fonction des actions du robot. Tandis que le robot se rapproche du lieu conditionné, la

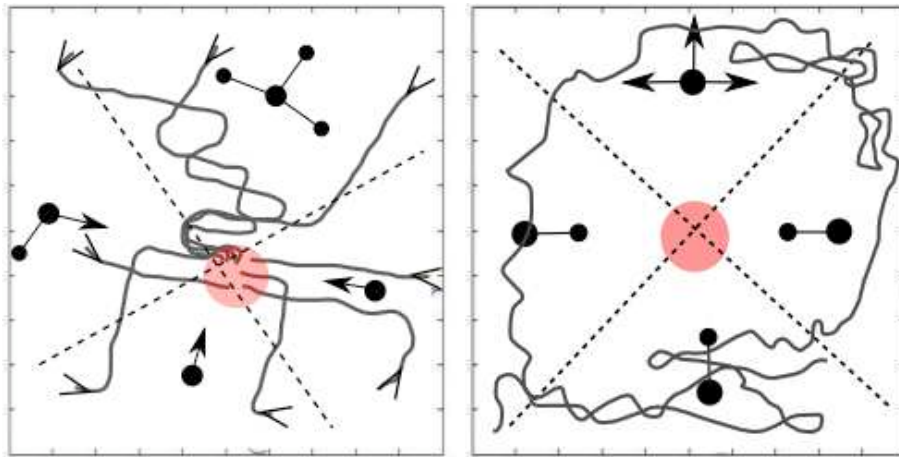


FIGURE 8.14 – A gauche, trajectoires du robot depuis différents points de départ et convergent vers le but (bassin d’attraction). A droite, trajectoires du robot depuis différents points de départ et évitant la zone dangereuse (bassin de répulsion).

réponse conditionnée augmente. Elle diminue au contraire quand le robot s’éloigne du lieu conditionné. Ceci est dû au fait que le potentiel de la réponse conditionnée dépend du niveau de reconnaissance du lieu conditionné. L’évolution de ces signaux de réponse émotionnelle conditionnée aux lieux de l’environnement peut donc servir à qualifier le comportement du robot. Lorsque le robot se rapproche d’un lieu conditionné positivement ou lorsqu’il s’éloigne d’un lieu conditionné négativement, la dérivée temporelle du signal de conditionnement peut servir de signal de récompense. De même, lorsque le robot se rapproche d’un lieu conditionné négativement ou lorsqu’il s’éloigne d’un lieu conditionné positivement, la dérivée temporelle du signal de conditionnement peut servir de signal de punition. Les équations 8.1 et 8.2 formalisent le calcul des signaux de récompense et de punition.

$$\text{Récompense} = \frac{dCOND_{pos}^+}{dt} + \frac{dCOND_{neg}^-}{dt} \quad (8.1)$$

$$\text{Punition} = \frac{dCOND_{neg}^+}{dt} + \frac{dCOND_{pos}^-}{dt} \quad (8.2)$$

$\frac{dCOND_{pos}^+}{dt}$ est la composante positive de la dérivée temporelle de la réponse du conditionnement positif avec un lieu ($\frac{dCOND_{pos}^-}{dt}$ est la composante négative). $\frac{dCOND_{neg}^+}{dt}$ est la composante positive de la dérivée temporelle de la réponse du conditionnement négatif ($\frac{dCOND_{neg}^-}{dt}$ est la composante négative). La figure 8.15 décrit le réseau de neurone utilisé pour réaliser ce coupage environnemental.

Ces signaux de punition et de récompense que le robot reconstruit à partir des valeurs données aux lieux de son environnement peuvent être utilisés de différentes façons. Une version minimaliste consisterait à utiliser ces signaux pour modifier les poids des connexions entre un neurone constant (i.e. un biais) et un champ de neurones à utiliser pour le contrôle moteur (chaque neurone représente une direction). Cette situation est bien adaptée pour l’utilisation d’une version neuronale simple d’apprentissage

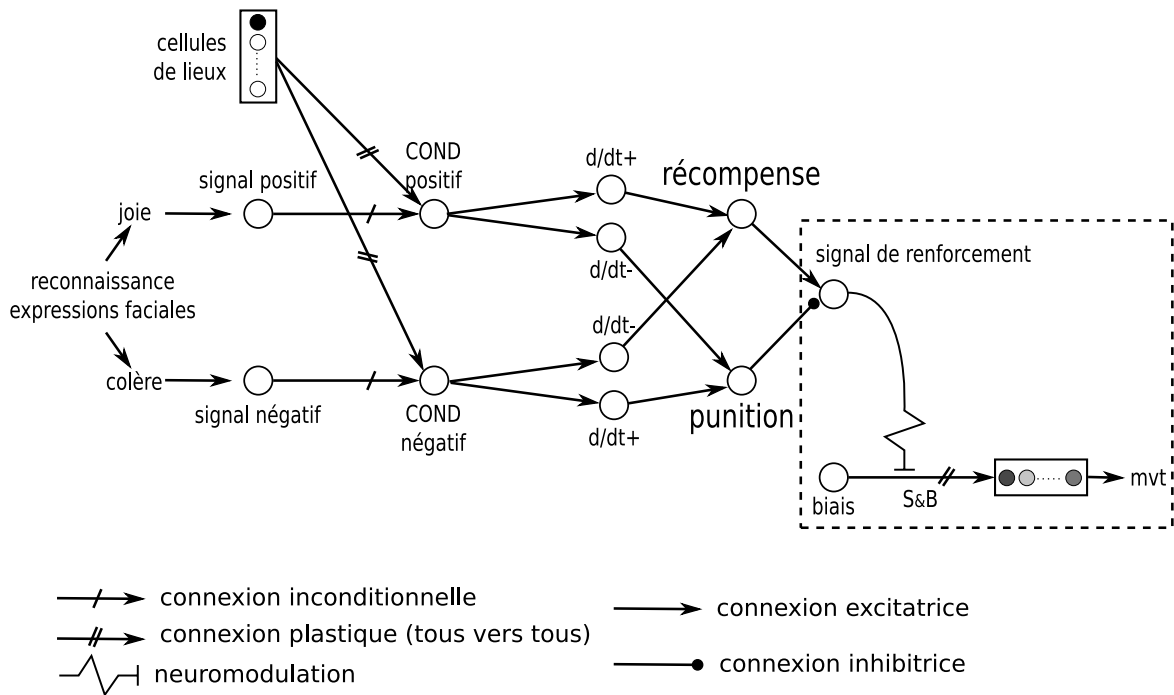


FIGURE 8.15 – Le couplage environnemental consiste à apprendre un conditionnement entre les signaux venant de l’interaction avec l’homme (reconnaissance d’une expression faciale émotionnelle positive ou négative) et la perception visuelle de l’environnement (cellule de lieu). Alors que le robot s’approche ou s’éloigne d’un lieu ayant appris une association via l’interaction, la réponse conditionnée augmente ou diminue. La dérivée temporelle de cette réponse peut alors servir de signal de récompense si le robot se rapproche d’un lieu positif ou s’éloigne d’un lieu négatif et de signal de punition dans le cas contraire. Ce signal de renforcement est ici simplement utilisé pour neuromoduler les connexions entre un neurone constant (biais) et un groupe de neurones moteurs (contrôle de l’orientation du robot). Si une direction prise par le robot produit un signal de récompense, cette direction est renforcée (inhibée dans le cas contraire).

par renforcement avec la règle d’apprentissage de Sutton et Barto [Barto 1981]. Les connexions sont initialisées aléatoirement avec des poids faibles. Si aucune interaction n’a encore eu lieu, quand le robot se déplace, il ne perçoit aucun renforcement conditionné. S’il se déplace loin de tout lieu conditionné avec une interaction expressive, il reçoit très peu de renforcement de manière aléatoire. Ceci est causé par le bruit dans le fonctionnement de la perception visuelle qui provoque des petites variations chaotiques de la réponse des cellules de lieu. Dans ce cas, le robot a un comportement de navigation aléatoire. Si par contre le robot se rapproche d’un lieu positif ou s’éloigne d’un lieu négatif, le poids de la connexion entre le neurone biais et le neurone gagnant du champs moteur est augmenté et le robot devrait conserver sa direction. Dans le cas contraire, un signal de punition ou renforcement négatif diminue le poids de cette connexion et le robot change donc de direction. Les équations 8.3 et 8.4 décrit l’adaptation des poids synaptiques suivant la règle d’apprentissage par renforcement

de Sutton et Barto :

$$R = \text{Récompense} - \text{Punition} \quad (8.3)$$

$$\Delta w = \frac{dR}{dt} * \frac{dS_{max}}{dt} * E_i \quad (8.4)$$

R est le signal de renforcement et w est le poids de la connexion entre l'entrée E_i (ici uniquement le neurone biais) et la sortie S_{max} (neurone gagnant du champs moteur). Quand $\frac{dR}{dt}$ est positif, le poids de la connexion entre le biais et le neurone de direction gagnant est augmenté. Il est diminué quand $\frac{dR}{dt}$ est négatif. La figure 8.16 met en parallèle les différents signaux utilisés dans ce réseau de neurones.

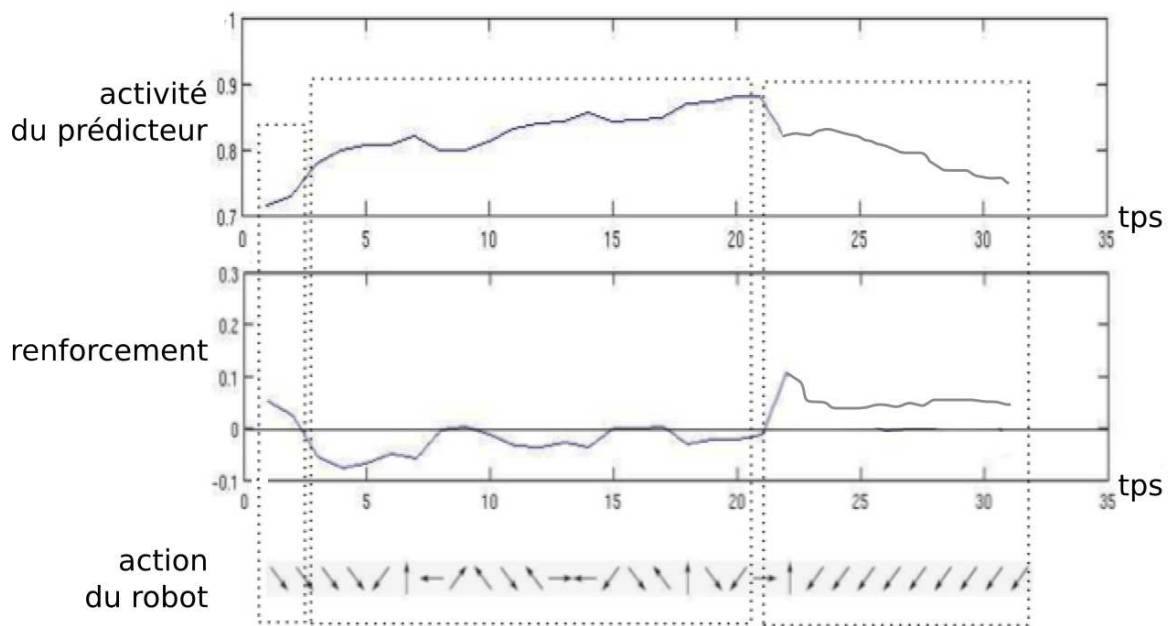


FIGURE 8.16 – Le robot a appris à associer un lieu avec une expression faciale négative. En se rapprochant de ce lieu le robot reçoit des renforcements négatifs et change de direction. Quand une direction gagnante l'éloigne de ce lieu, elle est renforcée positivement et le robot garde sa direction et continue de s'éloigner.

Le couplage environnemental a été utilisé pour apprendre à associer une valeur émotionnelle négative à un lieu dangereux et une valeur émotionnelle positive à un lieu but. Si le robot parvient à utiliser cette source de renforcement conditionnée à sa perception pour éviter un lieu dangereux, il lui est par contre impossible de rejoindre un lieu but. L'évitement repose sur des informations locales. Tant que le robot n'est pas à proximité du lieu dangereux, il n'a pas besoin de choisir une action en particulier. Toutes les directions peuvent être prises. Quand il s'agit par contre de rejoindre un lieu but, le robot doit en tout point de sa trajectoire définir l'orientation à prendre. Dépendant de la dynamique de la réponse de la cellule de lieu conditionnée, le signal de renforcement ne peut pas être utilisé à de grandes distances.

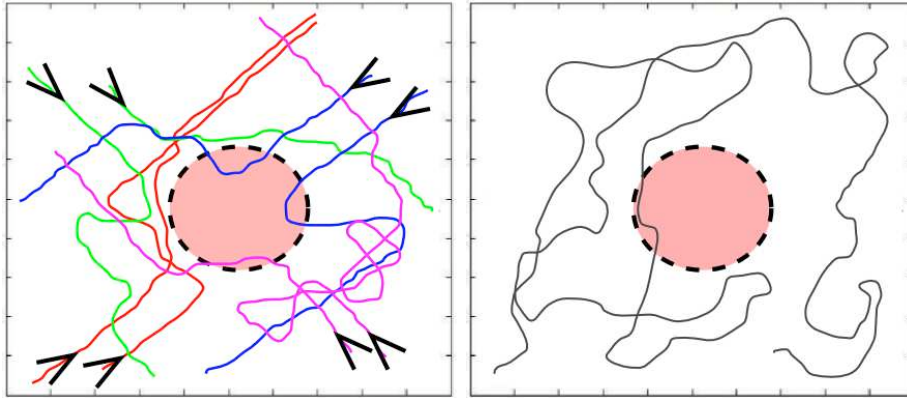


FIGURE 8.17 – A gauche, le robot a appris à associer une valeur négative au lieu correspondant à la zone dangereuse. Le robot évite effectivement ce lieu source de renforcement négatifs. A droite, le robot ne parvient pas à rejoindre le but uniquement en renforçant ses actions grâce au signal positif venant du lieu but.

5 Discussion

Dans les interactions Homme-Machine, l'expressivité émotionnelle a une place capitale. Nous montrons dans ce chapitre que l'expressivité émotionnelle peut servir de base à la construction de signaux d'interaction ayant un sens pour le robot. Certes, l'humain accède à l'état interne du robot si celui-ci est capable de l'exprimer. En soit, cela renseigne déjà sur la façon dont le robot perçoit le monde qui l'entoure. Mais de plus, c'est parce que le robot exprime son état émotionnel interne que l'humain peut lui apprendre le sens, ou plutôt la valence émotionnelle de n'importe quelle stimulation auparavant neutre. Cette approche minimale a pour but de montrer que n'importe quelle stimulation peut servir de base à un conditionnement émotionnel et donc à une communication entre l'humain et le robot. Utiliser une communication minimale n'est pas une fin en soit. Elle sert à illustrer les principes fondamentaux de fonctionnement de la communication émotionnelle de bas niveau. Le robot doit pouvoir exprimer son état interne et il doit pouvoir apprendre des conditionnements. C'est à partir de ces principes minimaux que le robot apprend ensuite à reconnaître des expressions faciales (voir modèle de l'apprentissage d'expression faciale présenté en annexe). La compréhension, ou plutôt la reconnaissance d'une stimulation émotionnelle quelle qu'elle soit repose donc sur le fait que cette stimulation provoque chez le robot un état émotionnel similaire à celui de l'agent (humain ou robot) ayant produit la stimulation. Le fait que les interactions émotionnelles prennent un sens repose donc sur une forme d'empathie. Cela revient pour l'humain à produire une expression faciale afin de reproduire chez le robot l'état qui en est à la cause chez lui. Si cette communication entre l'homme et le robot reste aujourd'hui encore grossière et superficielle, elle permet néanmoins de modéliser les premières étapes de l'apprentissage de la communication expressive (ou de la communication en générale).

Dans le contexte l'homme est considéré comme une sorte de tuteur bienveillant dont le but est d'aider le robot à apprendre à réagir de manière adéquat dans son envi-

ronnement. C'est pourquoi nous avons fait appel à la notion de référencement social. Les signaux expressifs d'interaction peuvent être utilisés par le robot pour apprendre à qualifier son environnement ou son propre comportement. Nous avons présenté deux types de couplage correspondants aux utilisations différentes des signaux d'interaction par l'amygdale et le nucleus accumbens. Le robot peut apprendre à qualifier les différents lieux de son environnement ainsi que les différentes actions qu'il peut y effectuer. Bien sûr, il n'y a pas un bon et un mauvais couplage. Nous avons montré que chacun de ces couplages remplit un rôle spécifique. Cependant, la communication expressive comporte une ambiguïté intrinsèque. Cette ambiguïté peut se formuler ainsi : "A quel couplage l'humain fait-il référence lorsqu'il interagit avec le robot?". Cette question est loin d'être triviale puisqu'elle peut poser de grandes difficultés aux humains aussi bien qu'aux robots. Dans la communication entre humains, l'expressivité émotionnelle n'est qu'exceptionnellement le seul support de sens. Les ambiguïtés peuvent alors être résolues grâce à la redondance sémantique des différents canaux de communication (expressivité émotionnelle, langage, geste déictique, ...).

D'autre part, les signaux des deux couplages pourraient être utilisés pour s'enrichir mutuellement. Par exemple, le robot pourrait utiliser les signaux du couplage environnemental comme signaux d'entrée pour le couplage comportemental. Le robot apprendrait interactivement avec l'humain quels lieux de son environnement sont bons (contenant un but) et quels lieux sont mauvais (contenant un danger). Le monitoring de l'évolution des réponses conditionnées pourrait alors servir de base à la génération autonome des signaux positifs et négatifs permettant au robot d'apprendre à renforcer ou inhiber les différents comportements de navigation. Il s'agirait d'une utilisation des signaux de prédiction des buts et des dangers assez similaire à l'automonitoring présenté dans le chapitre concernant la frustration.

Conclusion

1 Résumé des travaux

L'introduction de ce manuscrit soulignait le double intérêt de la modélisation des émotions pour un robot autonome dans une perspective bio-inspirée. D'un côté les émotions peuvent apporter aux robots des mécanismes de régulation du comportement. Sur le plan de l'autonomie, cette régulation permet de poser les premières bases de l'apprentissage des comportements élémentaires d'attraction et de répulsion ainsi que de détection de la nouveauté. Elle permet aussi au robot d'évaluer l'efficacité de son propre comportement et ainsi d'orienter au second ordre les processus de sélection de l'action. Sur le plan de la communication, les émotions permettent d'échanger des informations entre l'humain et le robot. Elles assurent aussi un cadre plus naturel et plus chargé de sens pour les humains dans les interactions Homme-Robot. D'un autre côté, les robots peuvent servir de plates-formes expérimentales pour évaluer les modèles des émotions proposés. L'inspiration neurobiologique qui caractérise ce travail permet ainsi de proposer des modèles plausibles de l'interaction des principales structures cérébrales impliquées dans le fonctionnement des émotions.

Le deuxième chapitre de ce manuscrit introduit l'approche psychologique des émotions. Un bref historique des grands mouvements de pensée permet d'associer les principales théories sur la nature des émotions à leurs racines philosophiques.

Le troisième chapitre est une introduction à la neurobiologie des émotions. Après un rapide historique de l'étude des bases neurobiologiques des émotions, nous avons décrit les principales structures cérébrales impliquées. Cet inventaire, très loin d'être exhaustif, permet de présenter les structures cérébrales modélisées dans nos architectures de contrôle.

Le quatrième chapitre est une introduction aux principes de la modélisation neuronales et des différentes méthodes d'apprentissage utilisées tout du long de ce travail. Dans ce chapitre, nous avons aussi présenté les deux stratégies de navigation d'inspiration biologique à la disposition du robot : la navigation sensori-motrice visuelle et la navigation proprioceptive (intégration de chemin).

Le cinquième chapitre présente le premier niveau d'apprentissage des conditionnements émotionnels. Dans un premier temps nous décrivons plus en détails le fonctionnement du robot. Suivant une perspective *animat* le robot est conçu comme un animal possédant une physiologie et ayant donc des contraintes de survie. Cette physiologie simulée est à l'origine des signaux de douleur et de plaisir. Un modèle de l'hypothalamus assure le fonctionnement des drives du robot. Les sensations de douleur et de plaisir sont présentées comme les bases essentielles de la notion de valence émotionnelle. Nous présentons la peur comme un mécanisme d'apprentissage de conditionnement à la douleur

(réalisé par l'amygdale). Ce modèle est illustré par une expérience robotique d'apprentissage de l'évitement d'obstacle par des conditionnements visuels à la douleur. Le robot apprend à associer la sensation de douleur (chocs contre les obstacles) avec le lieu dans lequel il se trouve (cellule de lieu la plus active). Si cette méthode n'est sûrement pas la plus efficace pour éviter des obstacles, elle permet néanmoins d'illustrer l'intérêt de ces conditionnements. Nous montrons à l'aide de cette expérience sur robot réel la possibilité de faire émerger les comportements de peur depuis les conditionnements à la douleur. A l'opposé, nous insistons sur l'importance des conditionnements avec le plaisir et les drives pour apprendre les comportements de navigation motivée au niveau du nucleus accumbens. Aussi bien avec la navigation visuelle que proprioceptive, nous illustrons ces conditionnements par des expériences robotiques de navigation multi-buts. Pour la navigation proprioceptive multi-buts, nous avons développé un modèle d'intégration de chemin utilisant une mémoire motrice à court terme pour stocker dynamiquement plusieurs buts. Les comportements rendus possibles grâce à l'apprentissage de conditionnements démontrent une certaine forme d'anticipation des événements à venir. Cette anticipation basée sur les expériences passées est la force d'organisation fondamentale de l'autonomie comportementale.

Le sixième chapitre présente notre modèle d'auto-évaluation comportementale assuré par le cortex préfrontal et insiste sur la réaction de frustration et l'inhibition frontale. Le constat est fait que l'une des causes du manque d'autonomie des robots est leur susceptibilité à entrer dans une dynamique vicieuse de comportements répétitifs et inefficaces : le *deadlock*. C'est pourquoi nous défendons le besoin capital pour le robot d'apprendre à prédire ses buts. Nous décrivons comment le robot peut utiliser ses différentes perceptions pour construire des anticipations ou prédictions des buts. Ces prédictions peuvent alors être utilisées pour évaluer les progrès du robot vers ses buts. De la même manière le robot est alors en mesure de détecter les situations d'échec. L'échec est alors défini comme une absence prolongée de progrès. Cette auto-évaluation permet au robot de générer un potentiel inhibiteur dont le but est de changer le comportement causant le *deadlock*. Ce potentiel inhibiteur est une modélisation du rôle régulateur du cortex préfrontal. Dans ce chapitre, nous explorons les différentes possibilités d'action de cette inhibition frontale. Celle-ci peut intervenir au niveau des drives, des buts ou des stratégies. Ces différentes inhibitions se révèlent efficace lorsqu'elles sont testées expérimentalement dans des contextes où une perturbation empêche le robot de satisfaire son but.

Le septième chapitre s'intéresse à la détection de nouveauté et à la réaction de surprise. Inspiré des travaux de Kaplan et Oudeyer sur l'analyse du signal d'un prédicteur, nous avons développé un modèle générique de détection de la nouveauté basé sur la détection d'erreur de prédiction des signaux sensoriels. Cette détection de nouveauté est une modélisation partielle de l'hippocampe. Nous introduisons la notion de cohérence perceptive (prédiction sensorielle intermodale). Le robot apprend une cohérence perceptive en apprenant à prédire chaque signal sensoriel à partir de tous les autres signaux sensoriels. C'est ce qui permet ensuite au robot non seulement de détecter lorsqu'il y a une anomalie dans ses perceptions mais en plus de définir lequel de ses sens est perturbé. Le modèle de la cohérence perceptive permet de modéliser une ré-

action de surprise de bas niveau.

Le huitième chapitre enfin aborde la question des interactions émotionnelles Homme-Robot. Dans un premier temps, ces interactions sont abordées de manière minimale. Le but est de définir les compétences minimales nécessaires à une interaction émotionnelle. Dans cette perspective, nous montrons que dès lors que le robot peut exprimer son état émotionnel propre (de manière compréhensible pour l'humain), il peut apprendre par conditionnement à donner un sens ou une valeur émotionnelle à n'importe quelle stimulation auparavant neutre. Une stimulation conditionnée à un état émotionnel positif ou négatif peut alors servir à donner des signaux de récompense ou de punition. À un niveau plus complexe, c'est essentiellement le même principe qui est utilisé pour apprendre au robot à reconnaître et à donner une valeur émotionnelle aux expressions faciales de l'humain. Que ce soit de manière minimale ou par l'utilisation de l'expressivité faciale, le but est d'utiliser ces signaux d'interaction pour aider le robot à apprendre des tâches de navigation motivée. Nous montrons qu'il existe une ambiguïté fondamentale dans l'utilisation des signaux d'interaction émotionnels et que ceux-ci peuvent être utilisés par l'amygdale ou par le nucleus accumbens. Nous avons réalisé deux couplages entre les signaux d'interaction et les réseaux de la navigation. Le couplage comportemental consiste à attribuer les signaux d'interaction au comportement du robot. Dans ce cas, le robot apprend à conditionner ses actions au niveau de son modèle de nucleus accumbens. Le couplage environnemental consiste quand à lui à associer les signaux d'interaction à la perception de l'environnement du robot. Il s'agit alors de conditionnements amygdaliens. Dans ce cas le robot apprend à conditionner sa localisation (cellule de lieu). Nous montrons que le couplage comportemental permet à l'humain d'apprendre au robot un bassin d'attraction ou de répulsion. Par contre le couplage environnemental tel que nous l'avons utilisé se montre plus limité. Si le robot apprend à qualifier un lieu de son environnement par conditionnement directe de la cellule de lieu correspondante, il peut l'éviter mais pas le rejoindre. Cependant, cette limitation est prévisible étant donné le besoin d'une dynamique attractrice entre plusieurs lieux-action pour surmonter les imprécisions dues au bruit dans la perception. L'interaction émotionnelle avec le robot permet l'échange d'informations pertinentes. Ce référencement social permet une interaction efficace et naturelle entre l'homme et le robot.

2 Discussion et Perspectives

La modélisation des mécanismes émotionnels pour un robot autonome passent nécessairement par la question de la nature des émotions. Après avoir examiné les différentes théories psychologiques, nous avons décrit l'état des connaissances actuelles sur les bases neurobiologiques des émotions. Cette source d'informations nous a servi d'inspiration pour le développement de modèles simplifiés de l'interaction des principales structures cérébrales impliquées dans les émotions. Selon cette perspective, les mécanismes émotionnels n'apparaissent pas structurellement plus apparentés à une vision purement dimensionnelle qu'à une vision purement discrète. Les émotions et proto-émotions à la base des comportements d'attraction et de répulsion sont bipola-

risées par définition. Cette bipolarité trouve son ancrage biologique dans les sensations de plaisir et de douleur. Les conditionnements avec ces deux sensations permettent de produire des comportements simples mais fondamentaux pour la survie : satisfaire ses besoins primaires et éviter les atteintes à l'intégrité corporelle. En d'autres termes ces conditionnements permettent l'émergence des motivations et des peurs. La frustration est plus complexe puisqu'elle résulte du monitoring de ces conditionnements et du rôle inhibiteur du cortex préfrontal. Cependant elle aussi est organisée autour de cette bipolarité. Tandis que la frustration résulte de la détection de l'échec, la satisfaction doit résulter de la détection du succès. Et au lieu de recourir à l'inhibition frontale pour changer de comportement, la satisfaction devrait résulter dans une activation du système de récompense qui permet alors de renforcer le comportement en question. Ce qui distingue alors la frustration de la peur est un aspect peu abordé de la nature des émotions : la complexité de leurs mécanismes sous-jacents (i.e du réseau de structures cérébrales impliquées). Cette notion de gradient de complexité des émotions pose problème à la vision dimensionnelle puisque celle-ci définit toutes les émotions comme les résultantes des mêmes dimensions fondamentales. D'un autre côté, la vision des émotions comme sous-tendues par des mécanismes spécifiques et distincts n'est pas conciliable avec l'idée d'un système cérébral de la valence émotionnelle. Au travers de ces travaux nous souhaitons donc défendre une définition de la nature des émotions qui soit plus orientée vers l'aspect structurel et fonctionnel des mécanismes biologiques la sous-tendant. Nous définissons alors les émotions comme le résultat de la dynamique d'interactions entre deux systèmes permettant l'évaluation des interactions avec l'environnement physique d'une part et l'environnement social d'autre part. Tandis que les interactions avec l'environnement physique provoquent des émotions, celles-ci déclenchent des modifications des comportements vis à vis de l'environnement physique (choix de l'action) aussi bien que social (expressivité). En retour, les émotions reconnues dans l'environnement social affectent les émotions ressenties et affectent le comportement aussi bien dans l'environnement physique (e.g. conditionnement d'une action) que social (e.g. imitation). Il manque cependant une étape essentielle de la caractérisation des émotions : la catégorisation dynamique des états émotionnels (patterning). Cette catégorisation pourrait ainsi permettre de mieux expliquer le lien entre la genèse d'un signal émotionnel et l'expressivité qui y est associée.

Mais comment faire le lien entre la description minimale des émotions issue de ce travail et l'expérience quotidienne riche et complexe que nous faisons des émotions. Pour cela il faut distinguer deux choses. Premièrement, nous n'avons abordé que les émotions les plus simples. Si les modèles proposés permettent par exemple d'expliquer les comportements de peur ou de surprise, ils n'ont pas l'ambition de modéliser des émotions comme la honte. La différence fondamentale entre ces deux émotions est le niveau de complexité des mécanismes cognitifs nécessaire à leur fonctionnement. La honte peut être définie comme la sensation négative qui accompagne sa propre dévaluation aux yeux d'autrui. Il n'est donc pas raisonnable d'espérer la modéliser par des mécanismes neuronaux caricaturalement simples. La cognition des robots que nous développons est centrée sur la satisfaction des besoins primaires. Deuxièmement, il est très difficile de dissocier les mécanismes et les comportements émotionnels de

l'expérience phénoménologique que l'on fait d'un épisode émotionnel. Cependant, la question du ressenti émotionnel n'a pas été abordée dans ce travail. Non pas que la question soit inintéressante, mais elle semble plutôt inappropriée à une perspective développementale constructiviste des mécanismes émotionnels. D'autre part, et plus fondamentalement, la question de la nature de la conscience nous échappe plus encore que la question des mécanismes émotionnels.

Un progrès important en terme d'adaptabilité déjà évoqué dans le chapitre 6 serait l'introduction d'une évaluation du potentiel de satisfaction et de frustration des différents contrôleurs du comportements (drive, but, stratégie). Cette évaluation apprise par conditionnement pourrait ainsi permettre de dépasser un contrôle réactif du comportement pour faire émerger des préférences comportementales liées à la nature de l'environnement (disponibilité des différents buts, conditions perceptives, ...). En allant un peu plus loin dans cette démarche, le robot pourrait apprendre à estimer le potentiel de succès et d'échec de ses drives, buts et stratégies en fonction des conditions telles qu'elles sont définies par la cohérence perceptive (modèle d'interprédictibilité). Le robot pourrait ainsi par exemple apprendre que dans le noir, la navigation visuelle est inefficace et qu'il n'est pas nécessaire d'attendre la perception de l'échec pour utiliser la navigation proprioceptive.

Un aspect très important de l'organisation du comportement a été laissé de côté durant ce travail : la curiosité. Cette caractéristique a notamment été étudiée par Kaplan et Oudeyer et définie comme la maximisation du progrès d'apprentissage. Selon notre approche, la curiosité peut être comprise comme le drive du progrès. Il s'agit d'un drive cognitif et non plus uniquement physiologique. Dans la plupart de nos expérimentations robotiques, lorsque le robot découvre complètement son environnement, il commence toujours par l'explorer en navigant aléatoirement. Cette approche a le mérite d'être simple mais elle n'est pas du tout efficace. La curiosité oriente le robot vers l'inconnu et lui permet au contraire d'optimiser la découverte de nouveaux environnements.

Même si l'on se restreint aux émotions les plus souvent considérées comme basiques, le présent travail ne fait pas un panorama complet. Deux émotions fondamentales manquent à l'appel : la colère et la tristesse. La colère est l'émotion qui se caractérise par un comportement agressif envers l'agent empêchant la satisfaction d'un besoin (ou la préservation de l'intégrité corporelle). D'un côté, son contraire pourrait être compris comme le comportement d'attachement envers l'agent facilitateur de la satisfaction des besoins (amour pour les humains). D'un autre côté, et cela est probablement plus pertinent, son contraire pourrait être compris comme la réaction d'évitement envers l'agent empêchant la satisfaction d'un besoin (i.e. la peur). Cette opposition entre la colère et la peur se comprend bien par la célèbre alternative comportementale *fight or flight* (se battre ou s'enfuir). La colère est un outil de communication puissant pour transmettre ses intentions. Sans aller jusqu'à faire se battre des robots, en exprimant sa colère, un robot utilise un moyen de communication puissant et précis quand à ses intentions. La plupart du temps d'ailleurs l'expression seule de colère suffit à déterminer l'issue d'une confrontation. La modélisation de la colère permettrait ainsi d'étudier bien plus en détails l'émergence de comportements sociaux entre robots. La tristesse

quant à elle se caractérise par une incapacité chronique à satisfaire ses buts. Son intérêt adaptatif doit être compris dans une perspective sociale. L'expression de sa tristesse peut avoir un impact sur le comportement des autres. Elle peut par exemple susciter des comportements d'aide ou de coopération. La tristesse se caractérise aussi par un ralentissement global des fonctions vitales. Ceci pourrait donc aussi correspondre à une adaptation à la rigueur de l'environnement ou à l'inefficacité du comportement. Pour conclure, nous pouvons répondre au moins partiellement à la question de la nature des émotions. La figure 9.1 illustre notre conception de l'émergence des signaux émotionnels depuis les signaux sensoriels fondamentaux : la douleur, le plaisir, les sensations externes (vue, ouïe, odorat, ...) et les sensations centrées sur le corps (toucher et proprioception). Bien que peu utilisée dans ces travaux, nous avons montré comment la surprise peut émerger de la détection de nouveauté qui elle-même peut résulter de la construction d'une cohérence perceptive. Les conditionnements à la douleur permettent de produire des comportements de peur. Les conditionnements de la sensation de plaisir aux différentes sensations externes permettent de construire des prédictions des différents buts. Ces prédictions sont à la base de l'anticipation de la satisfaction. Nous avons montré comment ces anticipations permettent l'auto-évaluation à l'origine de la réaction de frustration. Enfin, nous n'avons pas approfondi nos travaux sur les émotions suivantes, mais nous pouvons les définir dans le cadre de notre conception des émotions : la satisfaction (joie) en tant que pendant de la frustration et la colère et la tristesse en tant que conséquences de la frustration. Selon cette perspective, les émotions "basiques" ne semblent avoir de basique que le nom. Elles ne sont pas liées à des structures cérébrales spécifiques. Il semble plutôt qu'un ensemble de structures fondamentales assure l'apprentissage des conditionnements liant les sensations internes et externes et permet d'orienter le comportement afin d'optimiser la survie.

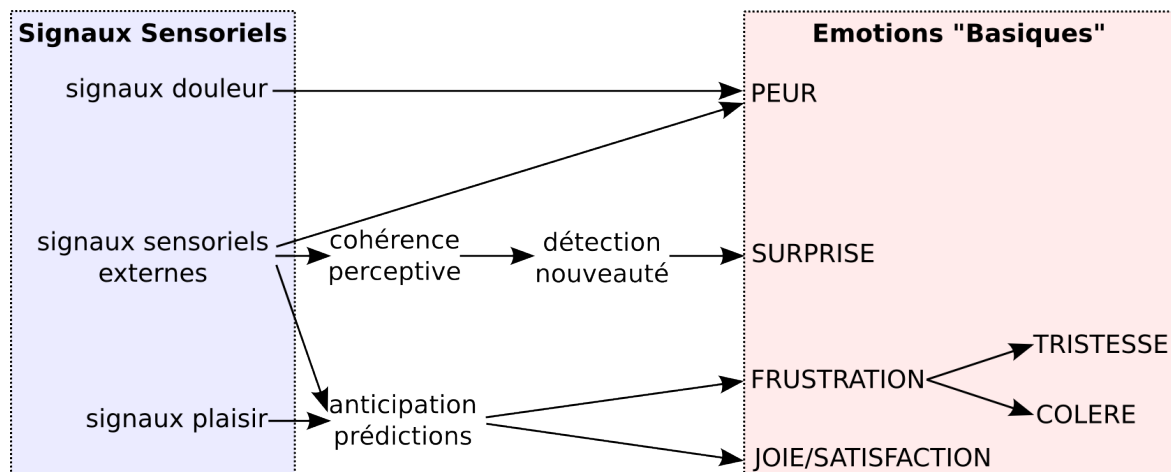


FIGURE 9.1 – Emergence des émotions depuis les sensations fondamentales de douleur et de plaisir en lien avec les sensations externes.

Bibliographie

- [Altier 1998] N. Altier et J. Stewart. *Dopamine receptor antagonists in the nucleus accumbens attenuate analgesia induced by ventral tegmental area substance P or morphine and by nucleus accumbens amphetamine*. *Journal of Pharmacology and Experimental Therapeutics*, vol. 285(1), pages 208–15, 1998. 43
- [Andry 2008] P. Andry, P. Gaussier, M. Lagarde et S. Boucenna. *Proprioception and imitation : on the road to agent individuation*. In *International Conference on Robots and Systems (IROS)*, pages 43–63, 2008. 143
- [Arkin 1998] R. Arkin. *Behavior based robotics*. Cambridge MA : The MIT Press, 1998. 1, 2, 3
- [Arkin 2003] R. Arkin, M. Fujita, T. Takagi et R. Hasegawa. *An ethological and emotional basis for human robot interaction*. *Robotics and Autonomous Systems*, vol. 42(3-4), pages 191–201, 2003. 7, 142
- [Arleo 2001] A. Arleo, F. Smeraldi, S. Hug et W. Gerstner. *Place Cells and Spatial Navigation based on Vision, Path Integration, and Reinforcement Learning*. In *Advances in Neural Information Processing Systems*, pages 89–95, 2001. 91
- [Arnold 1950] M.B. Arnold. *Feelings and emotions : The mooseheart symposium, chapitre An excitatory theory of emotion*. New York : Mc-Graw-Hill, 1950. 23
- [Arnold 1960] M.B. Arnold. *Emotion and personality*. Columbia University Press New York, 1960. 23
- [Aylward 1999] E.H. Aylward, N.J. Minshew, G. Goldstein, N.A. Honeycutt, A.M. Augustine, K.O. Yates, P.E. Barta et G.D. Pearlson. *MRI volumes of amygdala and hippocampus in non mentally retarded autistic adolescents and adults*. *Neurology*, vol. 53, pages 21–45, 1999. 41
- [Balkenius 2009] C. Balkenius, J. Moren et S. Winberg. *Interactions between Motivation, Emotion and Attention : From Biology to Robotics*. In *Proceedings of the Ninth International Conference on Epigenetic Robotics*, 2009. 149
- [Balleine 2003] B.W. Balleine, S.A. Killcross et A. Dickinson. *The effect of lesions of the basolateral amygdala on instrumental conditioning*. *Journal of Neuroscience*, vol. 23(2), pages 666–675, 2003. 40, 149
- [Balleine 2006] B.W. Balleine et S.A. Killcross. *Parallel incentive processing : an integrated view of amygdala function*. *Trends in Neuroscience*, vol. 29(5), pages 272–279, 2006. 40, 149
- [Bals-Kubik 1993] R. Bals-Kubik, A. Ableitner, A. Herz et T.S. Shippenberg. *Neuroanatomical sites mediating the motivational effects of opioids as mapped by the conditioned place preference paradigm in rats*. *Journal of Pharmacology and Experimental Therapeutics*, vol. 264(1), pages 489–495, 1993. 34
- [Bard 1928] P.A. Bard. *A diencephalic mechanism for the expression of rage with special reference to the central nervous system*. *American Journal of Physiology*, vol. 84, pages 490–513, 1928. 23, 30

- [Bard 1937] P.A. Bard et D.M. Rioch. *A study of four cats deprived of neocortex and additional portions of the forebrain*. Johns Hopkins Medical Journal, vol. 60, pages 73–153, 1937. 30
- [Bard 1996] K.A. Bard et K.H. Gardner. Reaching into thought : The minds of the great apes, chapitre Influences on development in infant chimpanzees : Enculturation, temperament and cognition. Cambridge University Press, 1996. 141
- [Bard 1999] K.A. Bard. Intersubjective communication and emotion in early ontogeny, chapitre Social-experiential contributions to imitation and emotion in chimpanzees. Cambridge University Press, 1999. 141
- [Baron-Cohen 2000] S. Baron-Cohen, H.A. Ring, E.T. Bullmore, S. Wheelwright, C. Ashwin et S.C.R. Williams. *The amygdala theory of autism*. Neuroscience and Biobehavioral Reviews, vol. 24(3), pages 355–364, 2000. 41
- [Barto 1981] A. Barto, R. Sutton et P. Brouwer. *Associative search network : A reinforcement learning associative memory*. Biological Cybernetics, vol. 40(3), pages 201–211, 1981. 158
- [Baxter 2002] M.G. Baxter et E.A. Murray. *The amygdala and reward*. Nature Review of Neuroscience, vol. 3, pages 563–573, 2002. 40, 149
- [Bechara 1995] A. Bechara, D. Tranel, H. Damasio, R. Adolphs, C. Rockland et A.R. Damasio. *Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans*. Science, vol. 269(5227), pages 1115–1118, 1995. 36
- [Bechara 2000] A. Bechara, H. Damasio et A.R. Damasio. *Emotion, decision making and the orbitofrontal cortex*. Cerebral Cortex, vol. 10(3), pages 295–307, 2000. 44
- [Berridge 1998] K.C. Berridge et T.E. Robinson. *What is the role of dopamine in reward : hedonic impact, reward learning, or incentive salience ?* Brain Research Reviews, vol. 28, no. 3, pages 309–369, 1998. 48
- [Berridge 2003] K.C. Berridge. *Pleasures of the brain*. Brain and Cognition, vol. 52, no. 1, pages 106–128, 2003. 42
- [Birbaumer 1998] N. Birbaumer, W. Grodd, O. Diedrich, U. Klose, M. Erb, M. Lotze, F. Schneider, U. Weiss et H. Flor. *fMRI reveals amygdala activation to human faces in social phobics*. Brain Imaging, vol. 9(6), pages 1223–1226, 1998. 41
- [Bird 2008] C.M. Bird et N. Burgess. *The hippocampus and memory : insights from spatial processing*. Nature Reviews Neuroscience, vol. 9, pages 182–194, 2008. 37
- [Blair 2005] H.T. Blair, F. Sotres-Bayon, M.A.P. Moita et J.E. Ledoux. *The lateral amygdala processes the value of conditioned and unconditioned aversive stimuli*. Neuroscience, vol. 133(2), pages 561–569, 2005. 40, 149
- [Bogerts 1993] B. Bogerts, J.A. Lieberman, M. Ashtari, R.M. Bilder, G. Degreef, G. Lerner, C. Johns et S. Masiar. *Hippocampus-amygdala volumes and psy-*

- chopathology in chronic schizophrenia*. Biological Psychiatry, vol. 33(4), pages 236–246, 1993. 41
- [Bohbot 1998] V.D. Bohbot, M. Kalina, K. Stepankova, N. Spackova, M. Petrides et L. Nadel. *Spatial memory deficits in patients with lesions to the right hippocampus and to the right parahippocampal cortex*. Neuropsychologia, vol. 36(11), pages 1217–1238, 1998. 37
- [Boucenna 2008] S. Boucenna, P. Gaussier et P. Andry. *What should be taught first : the emotional expression or the face ?* In International conference on Epigenetic Robotics, EPIROB, 2008. 143, 152
- [Breazeal 2000] C. Breazeal et B. Scassellati. *Infant-like Social Interactions between a Robot and a Human Caregiver*. Adaptive Behavior, vol. 8(1), pages 49–74, 2000. 142
- [Breazeal 2001] C. Breazeal. *Affective Interaction between Humans and Robots*. In Advances in Artificial Life, volume 2159, pages 582–591. Springer, 2001. 142
- [Breazeal 2002] C. Breazeal. *Designing sociable robots*. Cambridge MA : The MIT Press, 2002. 7
- [Bressan 2005] R.A. Bressan et J.A. Crippa. *The role of dopamine in reward and pleasure behaviour: review of data from preclinical research*. Acta Psychiatrica Scandinavica, vol. 427, pages 14–21, 2005. 48
- [Brooks 1986] R.A. Brooks. *A robust layered control system for a mobile robot*. Robotics and Automation, vol. 2(1), pages 14–23, 1986. 1, 2
- [Brooks 1991] R.A. Brooks. *Intelligence without representation*. Artificial Intelligence, vol. 47(1-3), pages 139–159, 1991. 2
- [Brown 1888] S. Brown et E. Shafer. *An investigation into the functions of the occipital and temporal lobes of the monkey's brain*. Philosophical Transactions of the Royal Society of London : Biological Sciences, vol. 179, pages 303–327, 1888. 31
- [Bruce 2002] A. Bruce, I. Nourbakhsh et R. Simmons. *The Role of Expressiveness and Attention in Human-Robot Interaction*. In Proceedings of the International Conference on Robotics and Automation IEEE, 2002. 142
- [Bunzeck 2006] N. Bunzeck et E. Düzel. *Absolute Coding of Stimulus Novelty in the Human Substantia Nigra/VTA*. Neuron, vol. 51(3), pages 369–379, 2006. 36
- [Burgdorf 2006] J. Burgdorf et J. Panksepp. *The neurobiology of positive emotions*. Neuroscience and Biobehavioral Reviews, vol. 30, no. 2, pages 173–187, 2006. 42, 73
- [Burke 1992] A. Burke, F. Heuer et D. Reisberg. *Remembering emotional events*. Memory & cognition, vol. 20(3), pages 277–290, 1992. 7
- [Cahilla 1998] L. Cahilla et J.L. McGaugh. *Mechanisms of emotional arousal and lasting declarative memory*. Trends in Neurosciences, vol. 21(7), pages 294–299, 1998. 36

- [Cannon 1920] W.B. Cannon. Bodily changes in pain hunger fear and rage. D.Appleton And Company, 1920. 23
- [Cannon 1927] W.B. Cannon. *The James-Lange theory of emotions : A critical examination and an alternative theory*. American Journal of Psychology, vol. 39, pages 106–124, 1927. 23
- [Cardinal 2002] R.N. Cardinal, J.A. Parkinson, J. Hall et B.J. Everitt. *Emotion and motivation : the role of the amygdala, ventral striatum, and prefrontal cortex*. Neuroscience and Biobehavioral Reviews, vol. 26, pages 321–352, 2002. 40, 149
- [Carpenter 1987] G.A. Carpenter et S. Grossberg. *ART 2 : self-organization of stable category recognition codes for analog input patterns*. Applied Optics, vol. 26(23), pages 4919–4930, 1987. 59
- [Carpenter 1988] G.A. Carpenter et S. Grossberg. *The ART of Adaptive Pattern Recognition by a Self-Organizing Neural Network*. Computer, vol. 21(3), pages 77–88, 1988. 59
- [Carpenter 1990] G.A. Carpenter et S. Grossberg. *ART 3 : Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures*. Neural Networks, vol. 3(2), pages 129–152, 1990. 59
- [Carrive 1993] P. Carrive. *The periaqueductal gray and defensive behavior : functional representation and neuronal organization*. Biobehavioral Brain Research, vol. 58(1-2), pages 27–47, 1993. 34
- [Casey 2006] K.L. Casey et T.D. Tran. *Chapter 12 : Cortical mechanisms mediating acute and chronic pain in humans*. In Fernando Cervero et Troels S. Jensen, editeurs, PAIN, volume 81 of *Handbook of Clinical Neurology*, pages 159–177, I–III. Elsevier, 2006. 46, 47
- [Cañamero 2001a] L.D. Cañamero. *Building emotional artifacts in social worlds : Challenges and perspectives*. Rapport technique FS-01-02, Papers from the AAI Fall Symposium, 2001. 4
- [Cañamero 2001b] L.D. Cañamero et J. Fredslund. *I show you how i like you ? can you read it in my face ?* IEEE Transactions on Systems Man and Cybernetics, vol. 31(5), pages 454–459, 2001. 7
- [Cañamero 2008] L.D. Cañamero. Animating expressive characters for social interaction, chapitre Animating affective robots for social interaction. Amsterdam/Philadelphia : John Benjamins Publishing Company, 2008. 7
- [Chen 2008] B.T. Chen, M.S. Bowers, M. Martin, F.W. Hopf, A.M. Guillory, R.M. Carelli, J.K. Chou et A. Bonci. *Cocaine but Not Natural Reward Self-Administration nor Passive Cocaine Infusion Produces Persistent LTP in the VTA*. Neuron, vol. 59(2), pages 288–297, 2008. 34
- [Chiara 2002] G. Di Chiara. *Nucleus accumbens shell and core dopamine : differential role in behavior and addiction*. Behavioural Brain Research, vol. 137, no. 1-2, pages 75–114, 2002. 42

- [Chiara 2004] G. Di Chiara, V. Bassareo, S. Fenu, M.A. De Luca, L. Spina, C. Cadoni, E. Acquas, E. Carboni, V. Valentini et D. Lecca. *Dopamine and drug addiction : the nucleus accumbens shell connection*. Neuropharmacology, vol. 47, no. Supplement 1, pages 227–241, 2004. 42
- [Cicchetti 1995] D. Cicchetti, B.P. Ackerman et C.E. Izard. *Emotions and emotion regulation in developmental psychopathology*. Development and Psychopathology, vol. 7(1), pages 1–10, 1995. 22
- [Collett 2006] M. Collett, S.T. Collett et M.V. Srinivasan. *Insect navigation : Measuring travel distance across ground and through air*. Current biology, vol. 16(20), 2006. 68
- [Constantinidis 2002] C. Constantinidis, G.V. Williams et P.S. Goldman-Rakic. *A role for inhibition in shaping the temporal flow of information in prefrontal cortex*. Nature Neuroscience, vol. 5, pages 175–180, 2002. 100
- [Corbit 2001] L.H. Corbit, J.L. Muir et B.W. Balleine. *The Role of the Nucleus Accumbens in Instrumental Conditioning : Evidence of a Functional Dissociation between Accumbens Core and Shell*. The Journal of Neuroscience, vol. 21(9), pages 3251–3260, 2001. 42
- [Craig 2003] A.D. Craig. *A new view of pain as a homeostatic emotion*. Trends in Neurosciences, vol. 26(6), pages 303–307, 2003. 73
- [Crusio 1989] W.E. Crusio, H. Schwegler et J.H. van Abeelen. *Behavioral responses to novelty and structural variation of the hippocampus in mice. II. Multivariate genetic analysis*. Behavioral Brain Research, vol. 32(1), pages 81–88, 1989. 37
- [Damasio 1994] H. Damasio, T. Grabowski, R. Frank, A.M. Galaburda et A.R. Damasio. *The return of Phineas Gage : clues about the brain from the skull of a famous patient*. Science, vol. 264, pages 1102–1105, 1994. 44
- [Damasio 1995] A.R. Damasio. *Descartes' error : Emotion, reason, and the human brain*. Putnam Publishing, 1995. 44
- [Damasio 2003] A. Damasio. *Looking for spinoza : Joy, sorrow and the feeling brain*. Harcourt Inc., Orlando, 2003. 4, 73
- [D'Ardenne 2008] K. D'Ardenne, S.M. McClure, L.E. Nystrom et J.D. Cohen. *BOLD responses reflecting dopaminergic signals in the human ventral tegmental area*. Science, vol. 319(5867), pages 1264–1267, 2008. 36
- [Darwin 1872] C. Darwin. *The expression of the emotions in man and animals*. Londres, John Murray, 1872. 17, 121
- [Dautenhahn 2004] K. Dautenhahn et I. Werry. *Towards interactive robots in autism therapy : Background, motivation and challenges*. Pragmatics and Cognition, vol. 12(1), pages 1–35, 2004. 142
- [Dehaene 1989] S. Dehaene et J.P. Changeux. *A simple model of prefrontal cortex function in delayed-response tasks*. Journal of Cognitive Neuroscience, vol. 1(3), pages 244–261, 1989. 44, 100

- [Dehaene 1991] S. Dehaene et J.P. Changeux. *The Wisconsin Card Sorting Test : theoretical analysis and modeling in a neuronal network*. Cerebral Cortex, vol. 1(1), pages 62–79, 1991. 44, 115
- [Dehaene 1995] S. Dehaene et J.P. Changeux. *Neuronal models of prefrontal cortical functions*. Annals of the NY Academy of Sciences, vol. 769, pages 305–319, 1995. 44, 100, 115
- [Dehaene 2006] S. Dehaene, J.P. Changeux, L. Naccache, J. Sackur et C. Sergent. *Conscious, preconscious, and subliminal processing : a testable taxonomy*. Trends in Cognitive Sciences, vol. 10(5), pages 204–211, 2006. 44
- [Depaulis 1989] A. Depaulis, R. Bandler et M. Vergnes. *Characterization of pre-entorhinal periaqueductal gray matter neurons mediating intraspecific defensive behaviors in the rat by microinjections of kainic acid*. Brain Research, vol. 486, no. 1, pages 121–132, 1989. 34
- [Descartes 1649] R. Descartes. Les passions de l'ame. 1649. 17
- [Dias 2002] R. Dias et R.C. Honey. *Involvement of the rat medial prefrontal cortex in novelty detection*. Behavioral Neurosciences, vol. 116(3), pages 498–503, 2002. 44
- [Diorio 1993] D. Diorio, V. Viau et M.J. Meaney. *Prefrontal Cortical-Amygdalar Metabolism in Major Depression*. Journal of Neuroscience, vol. 13(9), pages 3839–3847, 1993. 44
- [Drevets 1999] W.C. Drevets. *Prefrontal Cortical-Amygdalar Metabolism in Major Depression*. Annals of the New York Academy of Sciences, vol. 877, pages 614–637, 1999. 41
- [Duchenne 1876] G.B.A. Duchenne. Mécanisme de la physionomie humaine ou analyse électro-physiologique de l'expression des passions. Librairie J.-B. Bailliere et Fils, 1876. 17
- [Duffy 1934] E. Duffy. *Emotion : an example of the need for reorientation in psychology*. Psychological Review, vol. 41(2), pages 184–198, 1934. 14
- [Duffy 1941] E. Duffy. *An explanation of "emotional" phenomena without the use of the concept "emotion"*. Journal of General Psychology, vol. 25, pages 283–293, 1941. 14
- [Duffy 2003] B.R. Duffy. *Anthropomorphism and the social robot*. Robotics and Autonomous Systems, vol. 42(3-4), pages 177–190, 2003. 142
- [Duncker 1941] K. Duncker. *On Pleasure, Emotion, and Striving*. Philosophy and Phenomenological Research, vol. 1(4), pages 391–430, 1941. 73
- [Dunsmoor 2009] J. Dunsmoor et N. Schmajuk. *Interpreting patterns of brain activation in human fear conditioning with an attentional associative learning model*. Behavioral Neuroscience, vol. 123(4), pages 851–855, 2009. 40, 149
- [Edelman 2001] G.M. Edelman et G. Tononi. A universe of consciousness how matter becomes imagination. Basic Books, 2001. 44

- [Ehrsson 1988] H.H. Ehrsson, T. Kito, N. Sadato, R.E. Passingham et E. Naito. *Neural substrate of body size : Illusory feeling of shrinking of the waist*. PLoS Biology, vol. 3(12), page e412, 1988. 131
- [Ekman 1969] P. Ekman, E.R. Sorenson et W.V. Friesen. *Pancultural elements in facial displays of emotion*. Science, vol. 164, pages 86–88, 1969. 21, 121
- [Ekman 1971] P. Ekman et W.V. Friesen. *Constants across cultures in the face and emotion*. Journal of Personality and Social Psychology, vol. 17, pages 124–129, 1971. 21, 121
- [Ekman 1972] P. Ekman. *Universals and cultural differences in facial expression of emotion*. Nebraska Symposium on Motivation, pages 207–283, 1972. 21, 121
- [Ekman 1977] P. Ekman et W.V. Friesen. *Manual for the facial action coding system*. Consulting Psychologists Press Palo Alto, 1977. 21
- [Ekman 1978] P. Ekman, W.V. Friesen et J.C. Hager. *Facial action coding system*. Consulting Psychologists Press Palo Alto, 1978. 21
- [Ekman 1983] P. Ekman, R.W. Levenson et W.V. Friesen. *Autonomic nervous system activity distinguishes among emotions*. Science, vol. 221, pages 1208–1210, 1983. 20
- [Ekman 1999] P. Ekman. *Handbook of cognition and emotion*, chapitre Basic Emotions. Sussex, U.K. : John Wiley and Sons, Ltd, 1999. 21
- [Ellsworth 1988] P.C. Ellsworth et C.A. Smith. *From appraisal to emotion : differences among unpleasant feelings*. Motivation and Emotion, vol. 12, pages 271–302, 1988. 24
- [Ellsworth 1991] P.C. Ellsworth. K. strongman (ed.), *international review of studies on emotion*, chapitre Some implications of cognitive appraisal theories of emotion, pages 143–161. New York : Wiley, 1991. 24
- [Ellsworth 1994] P.C. Ellsworth. *William James and emotion : is a century of fame worth a century of misunderstanding?* Psychological Review, vol. 101, pages 222–229, 1994. 24
- [Etienne 1996] A.S. Etienne, R. Maurer et V. Seguinot. *Path integration in mammals and its interaction with visual landmarks*. Journal of Experimental Biology, vol. 199, no. 1, pages 201–209, 1996. 94
- [Etkin 2006] A. Etkin, T. Egner, D.M. Peraza, E.R. Kandel et J. Hirsch. *Resolving emotional conflict : a role for the rostral anterior cingulate cortex in modulating activity in the amygdala*. Neuron, vol. 51(6), pages 871–882, 2006. 44
- [Fanselow 1995] M.S. Fanselow, J.P. Decola, B.M. De Oca et J. Landeira-Fernandez. *Ventral and dorsolateral regions of the midbrain periaqueductal gray (PAG) control different stages of defensive behavior : Dorsolateral PAG lesions enhance the defensive freezing produced by massed and immediate shock*. Aggressive Behavior, vol. 21(1), pages 63–77, 1995. 34
- [Fernández-Carriba 2002] S. Fernández-Carriba, A. Loeches, A. Morcillo et W.D. Hopkins. *Asymmetry in facial expression of emotions by chimpanzees*. Neuropsychologia, vol. 123(4), pages 1523–1533, 2002. 141

- [Ferrari 2003] P.F. Ferrari, A.M.M. Van Erp, W. Tornatzky et K.A. Miczek. *Accumbal dopamine and serotonin in anticipation of the next aggressive episode in rats*. *European Journal of Neuroscience*, vol. 17, pages 371–378, 2003. 42
- [Fields 2007] H.L. Fields, G.O. Hjelmstad, E.B. Margolis et S.M. Nicola. *Ventral Tegmental Area Neurons in Learned Appetitive Behavior and Positive Reinforcement*. *Annual Review of Neuroscience*, vol. 30, pages 289–316, 2007. 36
- [Fitzpatrick 2004] R.C. Fitzpatrick et B.L. Day. *Probing the human vestibular system with galvanic stimulation*. *Journal of Applied Physiology*, vol. 96, pages 2301–2316, 2004. 131
- [Frijda 1986] N.H. Frijda. *The emotions*. New York : Cambridge University Press, 1986. 24
- [Frijda 1987] N.H. Frijda. *Emotion, cognitive structure and action tendency*. *Cognition and Emotion*, vol. 1, pages 115–143, 1987. 24
- [Frijda 1993] N.H. Frijda. *The place of appraisal in emotion*. *Cognition and Emotion*, vol. 7, pages 357–387, 1993. 24
- [Frijda 1995] N.H. Frijda. *Comparative approaches to cognitive science, chapitre Emotions in Robots*. Cambridge MA : The MIT Press, 1995. 6
- [Fudge 2000] J.L. Fudge et S.N. Haber. *The central nucleus of the amygdala projection to dopamine subpopulations in primates*. *Neuroscience*, vol. 97(3), pages 479–494, 2000. 40
- [Fudge 2009] J.L. Fudge et T. Tucker. *Amygdala projections to central amygdaloid nucleus subdivisions and transition zones in the primate*. *Neuroscience*, vol. 159(2), pages 819–841, 2009. 40
- [Fukuda 2004] T. Fukuda, J. Myung-Jin, M. Nakashima, F. Arai et Y. Hasegawa. *Facial expressive robotic head system for human-robot communication and its application in home environment*. In *Proceedings of the IEEE*, volume 92(11), pages 1851–1865, 2004. 7
- [Furmark 2004] T. Furmark, M. Tillfors, H. Garpenstrand, I. Marteinsdottir, B. Långström, L. Oreland et M. Fredrikson. *Serotonin transporter polymorphism related to amygdala excitability and symptom severity in patients with social phobia*. *Neuroscience Letters*, vol. 362(3), pages 189–192, 2004. 41
- [Gadanhó 2003] S.C. Gadanhó et P. Dayan. *Bridging emotion theory and neurobiology through dynamic systems modeling*. *The Journal of Machine Learning Research*, vol. 4, pages 385–412, 2003. 7
- [Gallistel 1996] C.R. Gallistel et A.E. Cramer. *Computations on metric maps in mammals : getting oriented and choosing a multi-destination route*. *Journal of experimental biology*, vol. 199(1), 1996. 67, 68
- [Gallistel 1998] C.R. Gallistel. *An invitation to cognitive science : Methods, models, and conceptual issues, chapitre Symbolic processes in the brain : The case of insect navigation*. The MIT Press, 1998. 68

- [Gaussier 1994] P. Gaussier et S. Zrehen. *Navigating with an animal brain : a neural network for landmark identification and navigation*. In Proceedings of Intelligent Vehicles, pages 399–404, 1994. 86
- [Gaussier 1995] P. Gaussier et S. Zrehen. *PerAc : A Neural Architecture to Control Artificial Animals*. Robotics and Autonomous Systems, vol. 16(2-4), 1995. 53, 62
- [Gaussier 1997] P. Gaussier, C. Joulain, S. Zrehen, J.P. Banquet et A. Revel. *Visual navigation in an open environment without map*. In International Conference on Intelligent Robots and Systems - IROS'97, pages 545–550, 1997. 53, 62
- [Gaussier 2000] P. Gaussier, C. Joulain, J.P. Banquet, S. Leprêtre et A. Revel. *The visual homing problem : an example of robotics/biology cross fertilization*. Robotics and Autonomous Systems, vol. 30, 2000. 53, 62
- [Gaussier 2001] P. Gaussier. *Toward a cognitive system algebra : A perception/action perspective*. In European Workshop on Learning Robots (EWLR), pages 88–100, 2001. 62
- [Gaussier 2004] P. Gaussier, K. Prepin et J. Nadel. Embodied artificial intelligence, volume 3139, chapitre *Toward a cognitive system algebra : Application to facial expression learning and imitation*, pages 243–258. 2004. 7, 143, 152
- [Gaussier 2007a] P. Gaussier, J.-P. Banquet, F. Sargolini, C. Giovannangeli, E. Save et B. Pousset. *A model of grid cells involving extra hippocampal path integration, and the hippocampal loop*. Journal of Integrative Neuroscience, vol. 6(3), 2007. 68, 69
- [Gaussier 2007b] P. Gaussier, S. Boucenna et J. Nadel. *Emotional interactions as a way to structure learning*. In International conference on Epigenetic Robotics, EPIROB, pages 193–194, 2007. 7, 143, 152
- [Giovannangeli 2006] C. Giovannangeli, P. Gaussier et G. Desilles. *Robust Mapless Outdoor Vision-Based Navigation*. In Proceeding of the International Conference on Intelligent Robots and Systems IROS, pages 3293–3300, 2006. 62
- [Giovannangeli 2007] C. Giovannangeli et P. Gaussier. *Orientation system in robots : Merging allothetic and idiothetic estimations*. In Proceeding of the 13th International Conference on Advanced Robotics, pages 349–354, 2007. 62
- [Giovannangeli 2008a] C. Giovannangeli. *Navigation autonome bio-inspirée en environnement intérieur et extérieur : apprentissage sensori-moteur et planification dans un cadre interactif*. PhD thesis, Université de Cergy-Pontoise, 2008. 127
- [Giovannangeli 2008b] C. Giovannangeli et P. Gaussier. *Autonomous vision-based navigation : Goal-oriented action planning by transient states prediction, cognitive map building, and sensory-motor learning*. In Proceeding of the International Conference on Intelligent Robots and Systems, pages 676–683, 2008. 62, 116
- [Gockley 2006] R. Gockley, J. Forlizzi et R. Simmons. *Interactions with a moody robot*. In HRI '06 : Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction, pages 186–193, 2006. 142

- [Grossberg 1971] S. Grossberg. *On the dynamics of operant conditioning*. Journal of Theoretical Biology, vol. 33(2), pages 225–255, 1971. 58
- [Grossberg 1974] S. Grossberg. *Classical and instrumental learning by neural networks*. Progress in Theoretical Biology, pages 51–141, 1974. 58
- [Grossberg 1987a] S. Grossberg et D.S. Levine. *Neural dynamics of attentionally modulated Pavlovian conditioning : blocking, interstimulus interval, and secondary reinforcement*. Applied Optics, vol. 26(23), pages 5015–5030, 1987. 58
- [Grossberg 1987b] S. Grossberg et N.A. Schmajuk. *Neural dynamics of attentionally modulated pavlovian conditioning : conditioned reinforcement, inhibition, and opponent processing*. Developmental psychobiology, vol. 15(3), pages 195–240, 1987. 58
- [Grossberg 2008] S. Grossberg, D. Bullock et M. Dranias. *Neural Dynamics Underlying Impaired Autonomic and Conditioned Responses Following Amygdala and Orbitofrontal Lesions*. Behavioral Neuroscience, vol. 122(5), pages 1100–1125, 2008. 40, 149
- [Grunwald 1998] T. Grunwald, K. Lehnertz, H.J. Heinze, C. Helmstaedter et C.E. Elger. *Verbal novelty detection within the human hippocampus proper*. Proceedings of the National Academy of Science USA, vol. 95(6), pages 3193–3197, 1998. 37
- [Haferlach 2007] T. Haferlach, J. Wessnitzer, M. Mangan et B. Webb. *Evolving a Neural Model of Insect Path Integration*. Adaptive Behavior, vol. 15(3), 2007. 68
- [Harlow 1848] J.M. Harlow. *Passage of an iron rod through the head*. Boston Medical and Surgical Journal, vol. 39, pages 389–393, 1848. 44
- [Harnad 1990] S. Harnad. *The Symbol Grounding Problem*. Physica D, vol. 42, pages 335–346, 1990. 3
- [Hasselmo 1999] M.E. Hasselmo et J.L. McClelland. *Neural models of memory*. Current Opinion in Neurobiology, vol. 9(2), pages 184–188, 1999. 59
- [Hasson 2010a] C. Hasson, S. Boucenna, P. Gaussier et L. Hafemeister. *Using Emotional Interactions for Visual Navigation Task Learning*. In Proceedings of the International Conference on Kansei Engineering and Emotion Research, pages 1578–1587, 2010. 152
- [Hasson 2010b] C. Hasson et P. Gaussier. *From Conditioning of a Non Specific Sensor to Emotional Regulation of Behavior*. In Proceedings of the International Conference on Artificial Neural Networks (ICANN 2010), pages 317–326, 2010. 143
- [Hasson 2010c] C. Hasson et P. Gaussier. *Frustration as a Generical Regulatory Mechanism for Motivated Navigation*. In Proceedings of the International Conference on Intelligent Robots and Systems (IROS 2010), pages 4704–4709, 2010. 99

- [Hasson 2010d] C. Hasson et P. Gaussier. *Path Integration Working Memory for Multi Task Dead Reckoning and Visual Navigation*. In Proceedings of Simulation of Adaptive Behavior (SAB'2010), pages 380–389, 2010. 87
- [Hebb 1949] D.O. Hebb. *The organization of behavior*. New York : Wiley, 1949. 56
- [Hecht-Nielsen 1987] R. Hecht-Nielsen. *Counterpropagation networks*. Applied Optics, vol. 26(23), pages 4979–4983, 1987. 62
- [Hecht-Nielsen 1989] R. Hecht-Nielsen. *Theory of the backpropagation neural network*. In International Joint Conference on In Neural Networks, pages 593–605, 1989. 62
- [Heimer 2003] L. Heimer. *A new anatomical framework for neuropsychiatric disorders and drug abuse*. American Journal of Psychiatry, vol. 160, pages 1726–1739, 2003. 37
- [Hornak 2004] J. Hornak, J. O'doherty, J. Bramham, E.T. Rolls, R.G. Morris, P.R. Bullock et C.E. Polkey. *Reward-related Reversal Learning after Surgical Excisions in Orbito-frontal or Dorsolateral Prefrontal Cortex in Humans*. Journal of Cognitive Neuroscience, vol. 16, no. 3, pages 463–478, 2004. 44
- [Howard 2000] M.A Howard, P.E. Boucher, P. Broks, A. Mayes, A. Farrant et N. Roberts. *Convergent neuroanatomical and behavioural evidence of an amygdala hypothesis of autism*. Brain Imaging, vol. 11(13), pages 2931–2935, 2000. 41
- [Hull 2002] A.M. Hull. *Neuroimaging findings in post-traumatic stress disorder*. The British Journal of Psychiatry, vol. 181, pages 102–110, 2002. 41
- [Hulse 2006] L. Hulse et A. Memon. *Fatal Impact ? The Effects of Emotional Arousal and Weapon Presence on Police Officers ? Memories for a Simulated Crime*. Legal & Criminological Psychology, vol. 11, pages 313–325, 2006. 7
- [Hyman 2001] S.E. Hyman et R.C. Malenka. *Addiction and the brain : The neurobiology of compulsion and its persistence*. Nature Reviews Neuroscience, vol. 2, pages 695–703, 2001. 34
- [Ikemoto 2007] S. Ikemoto. *Dopamine reward circuitry : two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex*. Brain Research Review, vol. 56(1), pages 27–78, 2007. 48
- [Izard 1971] C.E. Izard. *The face of emotion*. East Norwalk, CT, US : Appleton-Century-Crofts, 1971. 22
- [Izard 1972] C.E. Izard et E.S. Bartlett. *Patterns of emotions : A new analysis of anxiety and depression*. Oxford, England : Academic Press, 1972. 22
- [Izard 1977] C.E. Izard. *Human emotions*. New York : Plenum Press, 1977. 22
- [Izard 1979] C.E. Izard. *Maximally discriminative affect coding system (max)*. Newark : University of Delaware, Computer Network Service and University Media Service, 1979. 22
- [Izard 1980] C.E. Izard, R.R. Huebner, D. Risser, G. McGinnes et L. Dougherty. *The young infant's ability to produce discrete emotion expressions*. Developmental Psychology, vol. 16, pages 132–140, 1980. 22

- [Izard 1993] C.E. Izard. Handbook of emotions, chapitre Organizational and motivational functions of discrete emotions. New York, NY, US : Guilford, 1993. 22
- [Izard 1994] C.E. Izard. *Innate and universal facial expressions : Evidence from developmental and cross-cultural research*. Psychological Bulletin, vol. 115, pages 288–299, 1994. 22
- [Izard 2004] C.E. Izard et J-A.A. Abe. *Developmental Changes in Facial Expressions of Emotions in the Strange Situation During the Second Year of Life*. Development and Psychopathology, vol. 4(3), pages 251–265, 2004. 22
- [Jakob 1907] C. Jakob. Localizacion del alma y de la inteligencia. El Libro (Buenos Aires), 1907. 30
- [Jalabert 2009] M. Jalabert, G. Aston-Jones, E. Herzog, O. Manzoni et F. Georges. *Role of the bed nucleus of the stria terminalis in the control of ventral tegmental area dopamine neurons*. Progress in neuro-psychopharmacology and biological psychiatry, vol. 33(8), pages 1336–1346, 2009. 34
- [James 1884] W. James. *What is an Emotion ?* Mind, vol. 9, pages 188–205, 1884. 22
- [James 1890] W. James. The principles of psychology. 1890. 22
- [Julius 2001] D. Julius et A.I. Basbaum. *Molecular mechanisms of nociception*. Nature, vol. 413(6852), pages 203–210, 2001. 46
- [Kandel 1961] E.R. Kandel et W.A. Spencer. *Electrophysiological properties of an archicortical neuron. In Problems in Electrobiolgy*. Annals of the New York Academy of Sciences, vol. 94, pages 570–603, 1961. 56
- [Kandel 1968] E.R. Kandel et W.A. Spencer. *Cellular neurophysiological approaches in the study of learning*. Physiological Review, vol. 48, pages 65–134, 1968. 56
- [Kaplan 2003] F. Kaplan et P. Oudeyer. *Motivational principles for visual know-how development*, 2003. 121
- [Kaplan 2004] F. Kaplan et P. Oudeyer. *Maximizing learning progress : an internal reward system for development*. In Embodied Artificial Intelligence, pages 259–270. Springer-Verlag, 2004. 121
- [Keefe 2001] F.J. Keefe, M. Lumley, T. Anderson, T. Lynch, J.L. Studts et K.L. Carson. *Pain and emotion : new research directions*. Journal of Clinical Psychology, vol. 57(4), pages 587–607, 2001. 73
- [Klennert 1983] M. Klennert, J. Campos, J. Sorce, R. Emde et M. Svejda. *The development of the social referencing in infancy*. Emotion in early development, vol. 2, pages 57–86, 1983. 152
- [Klüver 1938] H. Klüver et 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, vol. 5, pages 33–54, 1938. 31
- [Klüver 1939] H. Klüver et P.C. Bucy. *Preliminary analysis of functioning of the temporal lobes in monkeys*. Archives of Neurology and Psychiatry, vol. 42, pages 979–1000, 1939. 31

- [Knight 1995] R.T. Knight, M.F. Grabowecky et D. Scabini. *Role of human prefrontal cortex in attention control*. *Advance in Neurology*, vol. 66, pages 21–34, 1995. 44
- [Knight 1996] R.T. Knight. *Contribution of human hippocampal region to novelty detection*. *Nature*, vol. 383, pages 256–259, 1996. 37
- [Knight 1999] R.T. Knight, W.R. Staines, D. Swick et L.L. Chao. *Prefrontal cortex regulates inhibition and excitation in distributed neural networks*. *Acta Psychologica*, vol. 101, no. 2-3, pages 159 – 178, 1999. 100
- [Knight 2005] D.C. Knight, H.T. Nguyen et P.A. Bandettini. *The role of the human amygdala in the production of conditioned fear responses*. *Neuroimage*, vol. 26(4), pages 1193–1200, 2005. 40, 149
- [Koechlin 2000] E. Koechlin, G. Corrado, P. Pietrini et J. Grafman. *Dissociating the role of the medial and lateral anterior prefrontal cortex in human planning*. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97(13), pages 7651–7656, 2000. 44
- [Kolb 1983] B. Kolb, R.J. Sutherland et I.Q. Wishaw. *A comparison of the contributions of the frontal and parietal cortex to spatial localization in rats*. *Behavioral Neuroscience*, vol. 87, 1983. 64
- [Konishi 1998] S. Konishi, K. Nakajima, I. Uchida, M. Kameyama, K. Nakahara, K. Sekihara et Y Miyashita. *Transient activation of inferior prefrontal cortex during cognitive set shifting*. *Nature Neuroscience*, vol. 1, pages 80–84, 1998. 115
- [Kozima 2005] H. Kozima, C. Nakagawa et Y. Yasuda. *Interactive robots for communication-care : a case-study in autism therapy*. In *IEEE International Workshop on Robot and Human Interactive Communication*, pages 341–346, 2005. 142
- [Kraut 1979] R.E. Kraut et R.E. Johnston. *Social and emotional messages of smiling : An ethological approach*. *Journal of Personality and Social Psychology*, vol. 37(9), pages 1539–1553, 1979. 141
- [Kringelbach 2009] M.L. Kringelbach. *The pleasure center : trust your animal instincts*. Oxford University Press, 2009. 73
- [Kumaran 2007] D. Kumaran et E.A. Maguire. *Which computational mechanisms operate in the hippocampus during novelty detection ?* *Hippocampus*, vol. 17(9), pages 735–748, 2007. 37
- [Kumaran 2009] D. Kumaran et E.A. Maguire. *Novelty signals : a window into hippocampal information processing*. *Trends Cognitive Sciences*, vol. 13(2), pages 47–54, 2009. 37
- [LaBar 2006] K.S. LaBar et R. Cabeza. *Cognitive neuroscience of emotional memory*. *Nature*, vol. 7(1), pages 54–64, 2006. 7, 36
- [Labrosse 2004] F. Labrosse. *Visual compass*. In *Proceedings of Towards Autonomous Robotic Systems*, pages 85–92, 2004. 63

- [Lackner 1988] J.R. Lackner. *Some proprioceptive influences on the perceptual representation of body shape and orientation*. *Brain*, vol. 111, pages 281–297, 1988. 131
- [Lang 1990] P.J. Lang, M.M. Bradley et B.N. Cuthbert. *Emotion, attention, and the startle reflex*. *Psychological Review*, vol. 97(3), pages 377–395, 1990. 16
- [Lang 1993] P.J. Lang, M.K. Greenwald, M.M. Bradley et A.O. Hamm. *Looking at pictures : Affective, facial, visceral, and behavioral reactions*. *Psychophysiology*, vol. 30(3), pages 261–273, 1993. 16
- [Lang 1994] P.J. Lang et M.M. Bradley. *Measuring emotion : The self-assessment manikin and the semantic differential*. *Journal of Behavior Therapy and Experimental Psychiatry*, vol. 25(1), pages 49–59, 1994. 16
- [Lange 1885] C. Lange. *The mechanisms of the emotions*. 1885. 22
- [Langford 2010] D.J. Langford et al. *Coding of facial expressions of pain in the laboratory mouse*. *Nature Methods*, vol. 7, pages 447–449, 2010. 141
- [Lawrie 2006] S.M. Lawrie, H.C. Whalley, D.E. Job et E.C. Johnstone. *Structural and Functional Abnormalities of the Amygdala in Schizophrenia*. *Annals of the New York Academy of Sciences*, vol. 985, pages 445–460, 2006. 41
- [Lazarus 1966] R.S. Lazarus. *Psychological stress and the coping process*. New York : McGraw-Hill, 1966. 24
- [Lazarus 1968] R.S. Lazarus. *Nebraska symposium on motivation, chapitre Emotion and adaptation : Conceptual and empirical relations*, pages 175–266. W.J. Arnold, 1968. 24
- [Lazarus 1982] R.S. Lazarus. *Thoughts on the relations between emotion and cognition*. *American Psychologist*, vol. 37(9), pages 1019–1024, 1982. 24
- [Lazarus 1984] R.S. Lazarus. *On the primacy of cognition*. *American Psychologist*, vol. 39, pages 124–129, 1984. 24
- [Lazarus 1991] R.S. Lazarus. *Emotion and adaptation*. New York : Oxford University Press, 1991. 24
- [LeDoux 1990] J.E. LeDoux, P. Cicchetti, A. Xagoraris et L.M. Romanski. *The lateral amygdaloid nucleus : sensory interface of the amygdala in fear conditioning*. *Journal of Neuroscience*, vol. 10, pages 1062–1069, 1990. 40
- [LeDoux 1993] J.E. LeDoux. *Emotional memory systems in the brain*. *Behavioral Brain Research*, vol. 58(1-2), pages 69–79, 1993. 36
- [LeDoux 2003a] J.E. LeDoux. *The Emotional Brain, Fear, and the Amygdala*. *Cellular and Molecular Neurobiology*, vol. 23(4-5), pages 727–773, 2003. 40
- [LeDoux 2003b] J.E. LeDoux. *Synaptic self*. New York : Penguin Books, 2003. 33
- [LeDoux 2007] J.E. LeDoux. *The amygdala*. *Current Biology*, vol. 17(20), pages 868–874, 2007. 40
- [Lee 1983] S.C Lee, T. Yamamoto et S. Ueki. *Characteristics of aggressive behavior induced by nucleus accumbens septi lesions in rats*. *Behavioral and Neural Biology*, vol. 37, no. 2, pages 237–245, 1983. 42

- [Leprêtre 2000] S. Leprêtre, P. Gaussier et J. Cocquerez. *From navigation to active object recognition*. In *Simulation for Adaptive Behavior (SAB)*, pages 266–275, 2000. 63
- [Levita 2002] L. Levita, J.W. Dalley et T.W. Robbins. *Nucleus accumbens dopamine and learned fear revisited : a review and some new findings*. *Behavioural Brain Research*, vol. 137, no. 1-2, pages 115–127, 2002. 42
- [Lewis 2005] M.D. Lewis. *Learning behavior-selection by emotions and cognition in a multi-goal robot task*. *Behavioral and Brain Science*, vol. 28, pages 169–194, 2005. 4
- [Lindsley 1951] D.B. Lindsley. *Handbook of experimental psychology (stevens s.s. ed.)*, chapitre Emotion. New York : Wiley, 1951. 15
- [Lisman 2001] J.E. Lisman et N.A. Otmakhova. *Storage, recall, and novelty detection of sequences by the hippocampus : elaborating on the SOCRATIC model to account for normal and aberrant effects of dopamine*. *Hippocampus*, vol. 11(5), pages 551–568, 2001. 36, 37
- [Lisman 2005] J.E. Lisman et A.A. Grace. *The Hippocampal-VTA Loop : Controlling the Entry of Information into Long-Term Memory*. *Neuron*, vol. 46(5), pages 703–713, 2005. 36
- [Lotstra 2002] F. Lotstra. *Le cerveau émotionnel ou la neuroanatomie des émotions*. *Cahiers critiques de thérapie familiale et de pratiques de réseaux*, vol. 29, pages 73–86, 2002. 44
- [MacLean 1949] P.D. MacLean. *Psychosomatic disease and the "visceral brain" : Recent developments bearing on the Papez theory of emotion*. *Psychosomatic Medicine*, vol. 11, pages 338–353, 1949. 31
- [MacLean 1953] P.D. MacLean et J.M.R. Delgado. *Electrical and chemical stimulation of frontotemporal portion of limbic system in the waking animal*. *Electroencephalography and Clinical Neurophysiology*, vol. 5(1), pages 91–100, 1953. 32
- [Mandler 1985] G. Mandler. *Mind and body*. New York NY : W.W. Norton, 1985. 3
- [Mannella 2008] F. Mannella, S. Zappacosta, M. Mirolli et G. Baldassarre. *A computational model of the amygdala nuclei's role in second order conditioning*. In *From animals to animats 10*, 2008. 149
- [Maren 2005] S. Maren. *Building and burying fear memories in the brain*. *Neuroscientist*, vol. 11(1), pages 89–99, 2005. 40, 149
- [McCulloch 1943] W.S. McCulloch et W. Pitts. *A logical calculus of the ideas immanent in nervous activity*. *Bulletin of Mathematical Biophysics*, vol. 5, pages 115–133, 1943. 54
- [McDonald 1998] A.J. McDonald. *Cortical pathways to the mammalian amygdala*. *Progress in Neurobiology*, vol. 55(3), pages 257–332, 1998. 40, 149
- [Mervaala 2000] E. Mervaala, J. Föhr, M. Könönen, M. Valkonen-Korhonen, P. Vainio, K. Partanen, J. Partanen, J. Tiihonen, H. Viinamäki, A.K. Karjalainen et J. Lehtonen. *Quantitative MRI of the hippocampus and amygdala in severe depression*. *Psychological Medicine*, vol. 30(1), pages 117–125, 2000. 41

- [Mittelstadt 1980] M.L. Mittelstadt et H. Mittelstadt. *Homing by path integration in a mammal*. *Naturwissenschaften*, vol. 67(11), 1980. 68
- [Moe 2006] R.O. Moe, M. Bakken, S. Kittilsen, H. Kingsley-Smith et B.M. Spruijt. *A note on reward-related behaviour and emotional expressions in farmed silver foxes (*Vulpes vulpes*) ? Basis for a novel tool to study animal welfare*. *Applied Animal Behaviour Science*, vol. 101, pages 362–368, 2006. 141
- [Montiel 2006] J.M.M. Montiel et A.J. Davison. *A visual compass based on SLAM*. In *Proceedings of the International Conference on Robotics and Automation (ICRA)*, pages 1917–1922, 2006. 63
- [Mor 2010] D. Mor, A.L. Bembrick, P.J. Austin, P.M. Wyllie, N.J. Creber, G.S. Denyer et K.A. Keay. *Anatomically specific patterns of glial activation in the periaqueductal gray of the sub-population of rats showing pain and disability following chronic constriction injury of the sciatic nerve*. *Neuroscience*, vol. 166, no. 4, pages 1167–1184, 2010. 34
- [Moren 2000] J. Moren et C. Balkenius. *A Computational Model of Emotional Learning in the Amygdala*. In *Proceeding of the 6th International Conference on the Simulation of Adaptive Behavior*, 2000. 149
- [Morris 1982] R.G.M. Morris, P. Garrud, J.N.P. Rawlins et J. O’Keefe. *Place navigation impaired in rats with hippocampal lesions*. *Nature*, vol. 297, pages 681–683, 1982. 37
- [Mudra 2003] R. Mudra et R.J. Douglas. *Self-correction mechanism for path integration in a modular navigation system on the basis of an egocentric spatial map*. *Neural networks*, vol. 16(9), 2003. 68
- [Nadel 2004] J. Nadel, A. Revel, P. Andry et P. Gaussier. *Toward communication : First imitations in infants, low-functioning children with autism and robots*. *Interaction Studies*, vol. 5(1), pages 45–74, 2004. 142
- [Nadel 2006] J. Nadel, M. Simon, P. Canet, R. Soussignan, P. Blancard, L.D. Canamero et P. Gaussier. *Human responses to an expressive robot*. In *Proceedings of the Sixth International Workshop on Epigenetic Robotics*, pages 79–86, 2006. 7, 142
- [Nathanson 1992] D.L. Nathanson. *Shame and pride : Affect, sex, and the birth of the self*. New York : W.W. Norton, 1992. 19
- [Nestler 2005] E.J. Nestler. *Is there a common molecular pathway for addiction ?* *Nature Neuroscience*, vol. 8, pages 1445–1449, 2005. 34
- [Nicola 2005] S.M. Nicola, S.A. Taha, S.W. Kim et H.L. Fields. *Nucleus accumbens dopamine release is necessary and sufficient to promote the behavioral response to reward-predictive cues*. *Neuroscience*, vol. 135, no. 4, pages 1025–1033, 2005. 36
- [Nieuwenhuys 2008] R. Nieuwenhuys, J. Voogd et C. van Huijzen. *The human central nervous system*. Springer-Verlag Berlin and Heidelberg GmbH and Co. K, 2008. 43

- [O'Doherty 2001] J. O'Doherty, M.L. Kringelbach, E.T. Rolls, J. Hornak et C. Andrews. *Abstract reward and punishment representations in the human orbitofrontal cortex*. *Nature Neuroscience*, vol. 4, pages 95–102, 2001. 44
- [O'Keefe 1971] J. O'Keefe et J. Dostrovsky. *The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat*. *Brain Research*, vol. 34(1), pages 171–175, 1971. 37, 62
- [Okun 2007] M.S. Okun, G. Mann, K.D Foote, N.A. Shapira, D. Bowers, U. Springer, W. Knight, P. Martin et W.K. Goodman. *Deep brain stimulation in the internal capsule and nucleus accumbens region : responses observed during active and sham programming*. *Journal of Neurology, Neurosurgery and Psychiatry*, vol. 78, pages 310–314, 2007. 42
- [Olds 1954] J. Olds et P. Milner. *Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain*. *Journal of Comp. Physiol. Psychol.*, vol. 47, pages 419–427, 1954. 48
- [Ortony 1988] A. Ortony, G.L. Clore et A. Collins. *The cognitive structure of emotions*. New York : Cambridge University Press, 1988. 24, 121
- [Ortony 1990] A. Ortony et T. Turner. *What's basic about basic emotions ?* *Psychological Review*, vol. 97, pages 315–331, 1990. 24
- [Osaka 2005] N. Osaka et M. Osaka. *Striatal reward areas activated by implicit laughter induced by mimic words in humans : a functional magnetic resonance imaging study*. *Neuroreport.*, vol. 16(15), pages 1621–1624, 2005. 42
- [Osgood 1952] C. Osgood. *The nature and measurement of meaning*. *Psychological Bulletin*, vol. 49, pages 172–237, 1952. 15
- [Osgood 1957] C. Osgood, G. Suci et P. Tannenbaum. *The measurement of meaning*. Urbana, IL : University of Illinois, 1957. 15
- [Oudeyer 2004] P. Oudeyer et F. Kaplan. *Intelligent Adaptive Curiosity : a source of Self-Development*. In *Lund University Cognitive Studies*, pages 127–130, 2004. 121
- [Oudeyer 2005] P. Oudeyer, F. Kaplan, V.V. Hafner et A. Whyte. *The playground experiment : Task-independent development of a curious robot*. In *Proceedings of the AAAI Spring Symposium on Developmental Robotics*, pages 42–47, 2005. 121
- [Oudeyer 2007] P. Oudeyer, F. Kaplan et V.V. Hafner. *Intrinsic motivation systems for autonomous mental development*. *IEEE Transactions on Evolutionary Computation*, vol. 11, pages 265–286, 2007. 121
- [Papez 1995] J. Papez. *A proposed mechanism of emotion (1937)*. *Journal of Neuro-psychiatry and Clinical Neurosciences*, vol. 7, pages 103–112, 1995. 30
- [Parr 2003] L.A. Parr. *The discrimination of faces and their emotional content by chimpanzees (Pan troglodytes)*. *Annals of the New York Academy of Sciences*, vol. 1000, pages 56–78, 2003. 141

- [Parr 2006] L.A. Parr et B.M. Waller. *Understanding chimpanzee facial expression : insights into the evolution of communication*. Social Cognitive and Affective Neuroscience, vol. 1(3), pages 221–228, 2006. 141
- [Paton 2006] J.J. Paton, M.A. Belova, S.E. Morrison et C.D. Salzman. *The primate amygdala represents the positive and the negative value of visual stimuli during learning*. Nature, vol. 439, pages 865–870, 2006. 40, 149
- [Pavlov 1927] I.P. Pavlov. *Conditional reflexes*. Oxford University Press, 1927. 58
- [Persson 2002] J. Persson, R. Habib et L. Nyberg. *Decreased activity in inferotemporal cortex during explicit memory : dissociating priming, novelty detection, and recognition*. Neuroreport, vol. 13(17), pages 2181–2185, 2002. 44
- [Pezawas 2005] L. Pezawas, A. Meyer-Lindenberg, E.M. Drabant, B.A. Verchinski, K.E. Munoz, B.S. Kolachana, M.F. Egan, V.S. Mattay, A.R. Hariri et D.R. Weinberger. *5-HTTLPR polymorphism impacts human cingulate-amygdala interactions : a genetic susceptibility mechanism for depression*. Nature Neuroscience, vol. 8, pages 828–834, 2005. 41
- [Pfeifer 1991] R. Pfeifer. *A dynamic view of emotion with an application to the classification of emotional disorders*. Vrije Universiteit Brussel AI Memo, pages 91–8, 1991. 4
- [Pfeifer 1999] R. Pfeifer et C. Scheier. *Understanding intelligence*. Cambridge MA : The MIT Press, 1999. 3
- [Phan 2006] K.L. Phan, D.A. Fitzgerald, P.J. Nathan et M.E. Tancer. *Association between Amygdala Hyperactivity to Harsh Faces and Severity of Social Anxiety in Generalized Social Phobia*. Biological Psychiatry, vol. 59(5), pages 424–429, 2006. 41
- [Phelps 2004] E.A. Phelps. *Human emotion and memory : interactions of the amygdala and hippocampal complex*. Current Opinion in Neurobiology, vol. 14, no. 2, pages 198–202, 2004. 36
- [Picard 1997] R. Picard. *Affective computing*. Cambridge MA : The MIT Press, 1997. 1
- [Picard 2002] R.W. Picard, A. Wexelblat et C.I. Nass. *Future interfaces : social and emotional*. In Conference on Human Factors in Computing Systems CHI '02, pages 698–699, 2002. 142
- [Pitkanen 2000] A. Pitkanen, E. Jolkkonen et S. Kempainen. *Anatomic heterogeneity of the rat amygdaloid complex*. Folia Morphologica, vol. 59(1), pages 1–23, 2000. 40, 149
- [Plutchik 1962] R. Plutchik. *The emotions : Facts, theories and a new model*. New York : Crown Publishing Group/Random House, 1962. 19
- [Plutchik 1980] R. Plutchik. *Emotion : Theory, research, and experience : Vol. 1. theories of emotion, chapitre A general psychoevolutionary theory of emotion*, pages 3–33. New York : Academic Press, 1980. 19

- [Plutchik 2002] R. Plutchik. *Emotions and life : Perspectives from psychology, biology, and evolution*. Washington, DC : American Psychological Association, 2002. 19
- [Pobbe 2010] R.L.H. Pobbe, H. Zangrossi Jr., D.C. Blanchard et R.J. Blanchard. *Involvement of dorsal raphe nucleus and dorsal periaqueductal gray 5-HT receptors in the modulation of mouse defensive behaviors*. *European Neuropsychopharmacology*, vol. In Press, Corrected Proof, pages –, 2010. 34
- [Prepin 2008] K. Prepin. *Développement et modélisation de capacités d'interaction homme-robots : L'imitation comme modèle de communication*. PhD thesis, Université de Cergy-Pontoise, 2008. 143
- [Purves 2008] D. Purves, J.A. George, D. Fitzpatrick, W.C. Hall, A.S LaMantia, J.O. McNamara et L.E. White. *Neuroscience (fourth edition)*. Sinauer Associates, Inc., 2008. 46
- [Rani 2006] P. Rani, C. Liu et N. Sarkar. *Affective feedback in closed loop human-robot interaction*. In *Proceedings of the 1st Conference on Human-Robot Interaction*, 2006. 7
- [Rauch 2000] S.L. Rauch, P.J. Whalen, L.M. Shin, S.C. McInerney, M.L. Macklin, N.B. Lasko, S.P. Orr et R.K. Pitman. *Exaggerated amygdala response to masked facial stimuli in posttraumatic stress disorder : a functional MRI study*. *Biological Psychiatry*, vol. 47(9), pages 769–776, 2000. 41
- [Rea 2009] K. Rea, Y. Lang et D.P. Finn. *Alterations in extracellular levels of gamma-aminobutyric acid in the rat basolateral amygdala and periaqueductal gray during conditioned fear, persistent pain and fear-conditioned analgesia*. *The Journal of Pain*, vol. 10(10), pages 1088–1098, 2009. 34
- [Reeves 1996] B. Reeves et C. Nass. *The media equation. how people treat computers television and new media like real people and places*. New York NY : Cambridge University Press, 1996. 2
- [Rescorla 1967a] R.A. Rescorla. *Pavlovian conditioning and its proper control procedures*. *Psychological Review*, vol. 74(1), pages 71–80, 1967. 58
- [Rescorla 1967b] R.A. Rescorla et R.L. Solomon. *Two-process learning theory : Relationships between Pavlovian conditioning and instrumental learning*. *Psychological Review*, vol. 74(3), pages 151–182, 1967. 58
- [Rescorla 1972] R.A. Rescorla et A.R. Wagner. *Classical conditioning ii. current research and theory*, chapitre A *Theory of Pavlovian Conditioning : Variations in the Effectiveness of Reinforcement and Non-reinforcement*. New York : Appleton Century Crofts, 1972. 58
- [Rescorla 1988] R.A. Rescorla. *Behavioral Studies of Pavlovian Conditioning*. *Annual Review of Neuroscience*, vol. 11, pages 329–352, 1988. 58
- [Richardson 2004] M.P. Richardson, B.A. Strange et R.J. Dolan. *Encoding of emotional memories depends on amygdala and hippocampus and their interactions*. *Nature Neuroscience*, vol. 7, pages 278–285, 2004. 36

- [Richter-Levin 2004a] G. Richter-Levin. *The amygdala, the hippocampus, and emotional modulation of memory*. *The Neuroscientist*, vol. 10(1), pages 31–39, 2004. 7
- [Richter-Levin 2004b] G. Richter-Levin. *The amygdala, the hippocampus, and emotional modulation of memory*. *Neuroscientist*, vol. 10(1), pages 31–39, 2004. 36
- [Robins 2005] B. Robins, K. Dautenhahn, R. Boekhorst et A. Billard. *Robotic assistants in therapy and education of children with autism : can a small humanoid robot help encourage social interaction skills ?* *Universal Access in the Information Society*, vol. 4(2), pages 105–120, 2005. 142
- [Rolls 2004] E.T. Rolls. *The functions of the orbitofrontal cortex*. *Brain and Cognition*, vol. 55(1), pages 11–29, 2004. 44
- [Roseman 1984] I.J. Roseman. *Cognitive determinants of emotion : a structural theory*. P. Shaver (Ed.), *Review of Personality and Social Psychology*, vol. 5, pages 11–36, 1984. 24
- [Roseman 1990] I.J. Roseman, M.S. Spindel et P.E. Jose. *Appraisal of emotion-eliciting events : testing a theory of discrete emotions*. *Journal of Personality and Social Psychology*, vol. 59, pages 899–915, 1990. 24
- [Roseman 1991] I.J. Roseman. *Appraisal determinants of discrete emotions*. *Cognition and Emotion*, vol. 5, pages 161–200, 1991. 24
- [Roseman 1996] I.J. Roseman. *Appraisal determinants of emotions : constructing a more accurate and comprehensive theory*. *Cognition and Emotion*, vol. 10, pages 241–277, 1996. 24
- [Rosenblatt 1958] F. Rosenblatt. *The Perceptron : probabilistic model for information storage and organization in the brain*. *Psychological Review*, vol. 65, pages 386–408, 1958. 54
- [Rubia 2003] K. Rubia, A.B. Smith, M.J. Brammer et E. Taylor. *Right inferior prefrontal cortex mediates response inhibition while mesial prefrontal cortex is responsible for error detection*. *NeuroImage*, vol. 20, no. 1, pages 351–358, 2003. 100
- [Rumelhart 1985a] D.E. Rumelhart, G.E. Hinton et R.J. Williams. *Learning Internal Representations by Error Propagation*. Rapport technique ADA164453, California University of San Diego La Jolla, Institute for Cognitive Science, 1985. 63
- [Rumelhart 1985b] D.E. Rumelhart et D. Zipser. *Feature discovery by competitive learning*. *Cognitive Science*, vol. 9, 1985. 59
- [Russell 1977] J.A. Russell et A. Mehrabian. *Evidence for a three-factor theory of emotions*. *Journal of Research in Personality*, vol. 11(3), pages 273 – 294, 1977. 15
- [Russell 1980] J.A. Russell. *A circumplex model of affect*. *Journal of Personality and Social Psychology*, vol. 39, pages 1161–1178, 1980. 15

- [Russell 1994] J.A. Russell. *Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies*. Psychological Bulletin, vol. 115(1), pages 102–141, 1994. 15
- [Russell 1997] C. Russell, K. Bard et L. Adamson. *Social referencing by young chimpanzees (pan troglodytes)*. Journal of Comparative Psychology, vol. 111(2), pages 185–193, 1997. 152
- [Russell 2003] N.A. Russell, A. Horii, P.F. Smith, C.L. Darlington et D.K. Bilkey. *Long-term effects of permanent vestibular lesions on hippocampal spatial firing*. The Journal of Neuroscience, vol. 23(16), pages 6490–6498, 2003. 37
- [Saha 2005] S. Saha, M.J. Drinkhill, J.P. Moore et T.F. Batten. *Central nucleus of amygdala projections to rostral ventrolateral medulla neurones activated by decreased blood pressure*. European Journal of Neuroscience, vol. 21(7), pages 1921–1930, 2005. 40
- [Salamone 1994] J.D. Salamone, M.S. Cousins et S. Bucher. *Anhedonia or anergia? Effects of haloperidol and nucleus accumbens dopamine depletion on instrumental response selection in a T-maze cost/benefit procedure*. Behavioural Brain Research, vol. 65, no. 2, pages 221–229, 1994. 42
- [Schachter 1962] S. Schachter et J. Singer. *Cognitive, social, and physiological determinants of emotional state*. Psychological Review, vol. 69(5), pages 379–399, 1962. 24
- [Schachter 1975] S. Schachter. Handbook of psychobiology, chapitre Cognition and peripheralist-centralist controversies in motivation and emotion, pages 529–564. New York : Academic Press, 1975. 24
- [Scherer 1982] K.R. Scherer. *Emotion as a process : function, origine and regulation*. Social Sciences Information, vol. 21, pages 555–570, 1982. 24
- [Scherer 1984a] K.R. Scherer. Approches to emotions, chapitre On the nature and function of emotion : a component process approach, pages 293–318. Hillsdale, NJ : Erlbaum, 1984. 24
- [Scherer 1984b] K.R. Scherer. P. shaver (ed.), review of personality and social psychology, volume 5, chapitre Emotion as a multicomponent process : a model and some cross-cultural data, pages 37–63. Beverly Hills, CA : Sage, 1984. 24
- [Scherer 1986] K.R. Scherer. *Vocal affect expression : a review and a model for futur research*. Psychological Bulletin, vol. 99, pages 165–143, 1986. 24
- [Scherer 1988] K.R. Scherer. Cognitive perspectives on emotions and motivation, chapitre Criteria for emotion-antecedent appraisal : a review., pages 89–126. Dordrecht : Nijhoff, 1988. 24
- [Scherer 1993] K.R. Scherer. *Studying the emotion-antecedent appraisal process : an expert system approach*. Cognition and Emotion, vol. 7, pages 325–355, 1993. 24
- [Scherer 2001] K.R. Scherer. Appraisal processes in emotion : Theory, methods, research, chapitre Appraisal considered as a process of multi-level sequential checking, pages 92–120. Oxford University Press, 2001. 25

- [Scheutz 2006] M. Scheutz, P. Schermerhorn et J. Kramer. *The utility of affect expression in natural language interactions in joint human-robot tasks*. In Proceedings of the 1st Conference on Human-Robot Interaction, 2006. 7
- [Schlosberg 1952] H. Schlosberg. *The description of facial expressions in terms of two dimensions*. Journal of Experimental Psychology, vol. 44(4), pages 229–237, 1952. 15
- [Schlosberg 1954] H. Schlosberg. *Three dimensions of emotion*. Psychological Review, vol. 61(2), pages 81–88, 1954. 15
- [Schmajuk 1986] N.A. Schmajuk et J.W. Moore. *Real-time attentional models for classical conditioning and the hippocampus*. Journal of comparative and physiological psychology, vol. 13(4), pages 278–290, 1986. 59
- [Schmajuk 1992] N.A. Schmajuk et J.J. DiCarlo. *Stimulus configuration, classical conditioning, and hippocampal function*. Psychological Review, vol. 99(2), pages 268–305, 1992. 59
- [Schneider 1998] F. Schneider, U. Weiss, C. Kessler, J.B. Salloum, S. Posse, W. Grodd et H.W. Müller-Gärtner. *Differential amygdala activation in schizophrenia during sadness*. Research Schizophrenia, vol. 34(3), pages 133–142, 1998. 41
- [Schultz 1997] W. Schultz, P. Dayan et P.R. Montague. *A Neural Substrate of Prediction and Reward*. Science, vol. 275(5306), pages 1593–1599, 1997. 34, 36
- [Schultz 1998] W. Schultz. *Predictive Reward Signal of Dopamine Neurons*. The Journal of Neurophysiology, vol. 80(1), pages 1–27, 1998. 48
- [Schultz 2005] R.T. Schultz. *Developmental deficits in social perception in autism : the role of the amygdala and fusiform face area*. International Journal of Developmental Neuroscience, vol. 23, no. 2-3, pages 125–141, 2005. 41
- [Schultz 2007] W. Schultz. *Multiple Dopamine Functions at Different Time Courses*. Annual Review of Neuroscience, vol. 30, pages 259–288, 2007. 38
- [Schwartz 1980] L. Schwartz. *Computational anatomy and functional architecture of striate cortex : a spatial mapping approach to perceptual coding*. Vision Research, vol. 20, 1980. 63
- [Schwienbacher 2004] I. Schwienbacher, M. Fendt, R. Richardson et H.U. Schnitzler. *Temporary inactivation of the nucleus accumbens disrupts acquisition and expression of fear-potentiated startle in rats*. Brain Research, vol. 1027, no. 1-2, pages 87–93, 2004. 42
- [Shaomin 2003] L. Shaomin, K.W. Cullen, R. Anwyl et M.J. Rowan. *Dopamine-dependent facilitation of LTP induction in hippocampal CA1 by exposure to spatial novelty*. Nature Neuroscience, vol. 6, pages 526–531, 2003. 37
- [Shekhar 2006] A. Shekhar, T.J. Sajdyk, D.R. Gehlert et D.G. Rainnie. *The Amygdala, Panic Disorder, and Cardiovascular Responses*. Annals of the New York Academy of Sciences, vol. 985, pages 308–325, 2006. 41
- [Sheline 1998] Y.I. Sheline, M.H. Gado et J.L. Price. *Amygdala core nuclei volumes are decreased in recurrent major depression*. Brain Imaging, vol. 9(9), pages 2023–2028, 1998. 41

- [Sherrington 1906] C.S. Sherrington. *The integrative action of the nervous system*. Scribner, New York, 1906. 46
- [Shimamura 2000] A.P. Shimamura. *The role of the prefrontal cortex in dynamic filtering*. *Psychobiology*, vol. 28, pages 207–218, 2000. 44, 100
- [Sigurdsson 2007] T. Sigurdsson, V. Doyère, C.K. Cain et J.E. LeDoux. *Long-term potentiation in the amygdala : A cellular mechanism of fear learning and memory*. *Neuropharmacology*, vol. 52, no. 1, pages 215–227, 2007. 40
- [Simon 1967] H.A. Simon. *Motivational and emotional controls of cognition*. *Psychological Review*, vol. 74(1), pages 29–39, 1967. 6
- [Skinner 1938] B.F. Skinner. *The behavior of organisms*. Appleton-Century, NY., 1938. 58
- [Skinner 1953] B.F. Skinner. *Science and human behavior*. Macmillan : Oxford, England, 1953. 156
- [Smith 1985] C.A. Smith et P.C. Ellsworth. *Pattern of cognitive appraisal in emotion*. *Journal of Personality and Social Psychology*, vol. 48, pages 813–838, 1985. 24
- [Smith 1987] C.A. Smith et P.C. Ellsworth. *Pattern of appraisal and emotion related to taking an exam*. *Journal of Personality and Social Psychology*, vol. 52, pages 475–488, 1987. 24
- [Smith 1989] C.A. Smith. *Dimensions of appraisal and physiological response in emotion*. *Journal of Personality and Social Psychology*, vol. 56, pages 339–353, 1989. 24
- [Smith 1993] C.A. Smith et R.S. Lazarus. *Appraisal components, core relational themes and the emotions*. *Cognition and Emotion*, vol. 7, pages 233–269, 1993. 24
- [Smith 1996] C.A. Smith, L.A. Griner, L.D. Kirby et H.S. Scott. *Toward a process model of appraisal in emotion*. In *Proceedings of the IXth Conference of the International Society for Research on Emotions, ISRE'96*, pages 101–105, 1996. 24
- [Smith 2006] A.P.R. Smith, E.K. Stephan, M.D. Rugg et R.J. Dolan. *Task and Content Modulate Amygdala-Hippocampal Connectivity in Emotional Retrieval*. *Neuron*, vol. 49(4), pages 631–638, 2006. 37
- [Spencer 1890] H. Spencer. *The principles of psychology (vol.1)*. New York : Appleton, 1890. 13
- [Spinoza 1677] B. Spinoza. *Ethique, partie iii : Concernant la nature et l'origine des émotions*. 1677. 13
- [Stackman 2002] R.W. Stackman, A.S. Clark et J.S. Taube. *Hippocampal spatial representations require vestibular input*. *Hippocampus*, vol. 12(3), pages 291–303, 2002. 37
- [Stein 2002] M.B. Stein, P.R. Goldin, J. Sareen, L.T. Eyler Zorrilla et G.G. Brown. *Increased Amygdala Activation to Angry and Contemptuous Faces in Generalized Social Phobia*. *Archive of General Psychiatry*, vol. 59, pages 1027–1034, 2002. 41

- [Sturm 2009] J. Sturm et A. Visser. *An appearance-based visual compass for mobile robots*. Robotics and Autonomous Systems archive, vol. 57(5), pages 536–545, 2009. 63
- [Sutton 1990] R.S. Sutton et A.G. Barto. Learning and computational neuroscience : Foundations of adaptive networks, chapitre Time-derivative models of Pavlovian reinforcement. Cambridge, MA, US : The MIT Press, 1990. 99
- [Swanson 1998] L.W. Swanson et G.D. Petrovich. *What is the amygdala?* Trends in Neuroscience, vol. 21, pages 323–331, 1998. 37
- [Takanori 1999] S. Takanori, T. Toshihiro et T. Kazuo. *Subjective interpretation of emotional behavior through physical interaction between human and robot*. In Systems, Man, and Cybernetics, 1999. IEEE SMC '99, pages 1024–1029, 1999. 142
- [Terrazas 2005] A. Terrazas, M. Krause, P. Lipa, K.M. Gothard, C.A. Barnes et B.L. McNaughton. *Self-Motion and the Hippocampal Spatial Metric*. The Journal of Neuroscience, vol. 25(35), pages 8085–8096, 2005. 37
- [Thorndike 1901] E.L. Thorndike. *Animal intelligence : An experimental study of the associative processes in animals*. Psychological Review Monograph Supplement, vol. 2, pages 1–109, 1901. 156
- [Tolman 1948] E.C. Tolman. *Cognitive maps in rats and men*. Psychological Review, vol. 55(4), pages 189–208, 1948. 116
- [Tomkins 1962] S.S. Tomkins. Affect imagery consciousness : The positive affects (vol. 1). New York : Springer, 1962. 19
- [Tomkins 1963] S.S. Tomkins. Affect imagery consciousness : The negative affects (vol. 2). New York : Springer, 1963. 19
- [Tomkins 1980] S.S. Tomkins. Emotion, theory, research, and experience : Theories of emotions, chapitre Affect as amplification : Some modifications in theory, pages 188–205. New York : Academic Press, 1980. 19
- [Tomkins 1991] S.S. Tomkins. Affect imagery consciousness : Anger and fear (vol. 3). New York : Springer, 1991. 19
- [Tononi 1998] G. Tononi et G.M. Edelman. *Consciousness and Complexity*. Science, vol. 282(5395), pages 1846–1851, 1998. 44
- [Tracey 2002] I. Tracey, A. Ploghaus, J.S. Gati, S. Clare, S. Smith, R.S. Menon et P.M. Matthews. *Imaging attentional modulation of pain in the periaqueductal gray in humans*. The Journal of Neuroscience, vol. 22(7), pages 2748–2752, 2002. 34
- [Tulving 1994] E. Tulving, H.J. Markowitsch, S. Kapur, R. Habib et S. Houle. *Novelty encoding networks in the human brain : positron emission tomography data*. Neuroreport, vol. 5(18), pages 2525–2528, 1994. 44
- [Vertes 2006] R.P. Vertes. *Interactions among the medial prefrontal cortex, hippocampus and midline thalamus in emotional and cognitive processing in the rat*. Neuroscience, vol. 142(1), pages 1–20, 2006. 44

- [Vickerstaff 2005] R. J. Vickerstaff et E. A. Di Paolo. *Evolving neural models of path integration*. Journal of Experimental Biology, vol. 208, 2005. 68
- [Vogt 2005] B.A. Vogt. *Pain and emotion interactions in subregions of the cingulate gyrus*. Nature Reviews Neuroscience, vol. 6, pages 533–544, 2005. 73
- [Walaas 1980] I. Walaas et F. Fonnum. *Biochemical evidence for [gamma]-aminobutyrate containing fibres from the nucleus accumbens to the substantia nigra and ventral tegmental area in the rat*. Neuroscience, vol. 5, no. 1, pages 63–72, 1980. 42
- [Walden 1988] T. Walden et T. Ogan. *The development of social referencing*. Child development, vol. 59, pages 1230–1240, 1988. 152
- [Walker 2010] J. Walker, A. Dale, N. Waran, N. Clarke, M. Farnworth et F. Wemelsfelder. *The assessment of emotional expression in dogs using a Free Choice Profiling methodology*. Animal Welfare, vol. 19(1), pages 75–84, 2010. 141
- [Watkins 1992] C. Watkins et P. Dayan. *Q-Learning*. Machine Learning, vol. 8(3-4), pages 279–292, 1992. 99
- [Wehner 2003] R. Wehner. *Insect navigation : Measuring travel distance across ground and through air*. Journal of Comparative Physiology A : Neuroethology, Sensory, Neural, and Behavioral Physiology, vol. 189(8), 2003. 87
- [Wehrle 2001] T. Wehrle. *The grounding problem of modeling emotions in adaptive systems*. Cybernetics and Systems, vol. 32(5), pages 561–580, 2001. 6
- [Weiner 1979] B. Weiner, J. Amirkhan, V.S. Folkes et J.A. Verette. *The cognition-emotion process in achievement-related contexts*. Journal of Personality and Social Psychology, vol. 37, pages 1211–1220, 1979. 24
- [Weiner 1982] B. Weiner. Affect and cognition, chapitre The emotional consequences of causal attributions, pages 185–209. Hillsdale, NJ : Erlbaum, 1982. 24
- [Weiner 1986] B. Weiner. An attributional theory of motivation and emotion. New York : Springer, 1986. 24
- [Weiner 1987] B. Weiner, J. Amirkhan, V.S. Folkes et J.A. Verette. *An attributional analysis of excuse giving : study of a naive theory of emotions*. Journal of Personality and Social Psychology, vol. 52, pages 316–324, 1987. 24
- [Widrow 1960] B. Widrow et M.E. Hoff. *Adaptive Switching Circuits*. IRE WESCON Convention Record, vol. 4, pages 96–104, 1960. 58
- [Williams 2004] L.M. Williams et et al. *Dysregulation of Arousal and Amygdala-Prefrontal Systems in Paranoid Schizophrenia*. American Journal of Psychiatry, vol. 161, pages 480–489, 2004. 41
- [Wilson 1991] S.W. Wilson. *The animat path to AI*. In From Animals to Animats (SAB), pages 15–21, 1991. 1, 4, 74
- [Wise 1980] R.A. Wise. *The dopamine synapse and the notion of [‘]pleasure centers’ in the brain*. Trends in Neurosciences, vol. 3, no. 4, pages 91–95, 1980. 48
- [Wise 2002] R.A. Wise. *Brain Reward Circuitry*. Neuron, vol. 36(2), pages 229–240, 2002. 34

- [Wundt 896] W. Wundt. *Outlines of psychology (grundriss der psychologie)*. Wilhelm Engelmann, 1897 (1896). 13
- [Yehuda 2002] R. Yehuda. *Post-Traumatic Stress Disorder*. The New England Journal of Medicine, vol. 346(2), pages 108–114, 2002. 41
- [Zajonc 1980] R.B. Zajonc. *Feeling and thinking : preferences need no inferences*. American Psychologist, vol. 2, pages 151–176, 1980. 24
- [Zajonc 1984a] R.B. Zajonc. *Approches to emotions*, chapitre On the primacy of affect, pages 117–123. Hillsdale, NJ : Erlbaum, 1984. 24
- [Zajonc 1984b] R.B. Zajonc et H. Markus. *Emotions, cognition and behavior*, chapitre Affect and cognition : the hard interface, pages 73–102. Cambridge : Cambridge University Press, 1984. 24
- [Zola-Morgan 1991] S. Zola-Morgan, L.R. Squire, P. Alvarez-Royo et R.P. Clower. *Independence of memory functions and emotional behavior : separate contributions of the hippocampal formation and the amygdala*. Hippocampus, vol. 1(2), pages 207–220, 1991. 36