Pierre Bessière
Eric Dedieu
Olivier Lebeltel
Emmanuel Mazer
Kamel Mekhnacha

Laboratoire LEIBNIZ - CNRS

INTERPRETATION OU DESCRIPTION

Proposition pour une théorie probabiliste des systèmes cognitifs sensori-moteurs

RESUME : Jusqu'à quel point agir et percevoir supposent-ils de « comprendre » ou même plus simplement de se « représenter » le monde ? Telle est l'une des préoccupations fondamentales des recherches sur la sensori-motricité au centre de nombreux débats en sciences cognitives. Cette question une fois formalisée, expurgée, simplifiée et traduite en termes mathématiques, nous amène à nous interroger tout au long de cet article, sur les liens qui peuvent exister entre les inférences formelles mécanisées informatiquement et leurs contreparties dans le monde physique où évolue un robot. Ainsi reformulée, la question centrale débattue devient : comment rendre effectives *les inférences formelles ?*

MOTS-CLES : science cognitive, intelligence artificielle, robotique, logique, inférence probabiliste, inférence bayésienne, entropie maximum, perception, action, décision

ABSTRACT : Is it necessary to « understand » or even to « represent » the world to perceive and act ? This is one of the fundamental issue in sensori-motor research. Numerous debates in cognitive sciences turn on this subject. After simplification, formalisation and translation into mathematical terms, we reformulate this question as : what is the nature of the link between the automatic inferences of a computer and their intended counterparts in the physical world where a robot is to be « embodied » ?

KEY WORDS : cognitive science, artificial intelligence, robotics, logic, probabilistic inference, bayesian inference, maximum entropy, perception, action, decision

Pierre Bessière
Eric Dedieu *Laboratoire LEIBNIZ - CNRS/*
Olivier Lebeltel
Emmanuel Mazerw
Kamel Mekhnacha

INTERPRETATION OU DESCRIPTION

Proposition pour une théorie probabiliste des systèmes cognitifs sensori-moteurs

I. INTRODUCTION

The most incomprehensible thing about the world is that it is comprehensible[1]

Albert Einstein

En deçà du langage et de toute activité intellectuelle verbalisable, existe l'immense domaine cognitif de la sensori-motricité. Cet article propose une réflexion à ce sujet.

La démarche suivie est à la fois modélisatrice et expérimentale. Modélisatrice, car nos sujets de réflexions sont des théories mathématiques. Expérimentale, car nous nous attachons à tester et illustrer la pertinence des concepts proposés et des arguments avancés par des expériences concrètes qui soulèvent en retour de nouvelles interrogations. Ces expériences sont rendues possibles, car les modèles mathématiques étudiés peuvent être utilisés pour animer des dispositifs sensori-moteurs artificiels, mécaniques et électroniques, souvent appelés robots.

Jusqu'à quel point agir et percevoir supposent-ils de « comprendre » ou même plus simplement de se « représenter » le monde ? Telle est l'une des préoccupations fondamentales des recherches sur la sensori-motricité. Question élémentaire dont les échos, pourtant, résonnent depuis bien longtemps entre diverses disciplines des sciences cognitives : philosophie, logique, linguistique, éthologie, neurosciences, psychologie ou intelligence artificielle.

Cette question une fois formalisée, expurgée, simplifiée et traduite en termes mathématiques, nous amène à nous interroger tout au long de cet article, sur les liens qui peuvent exister entre les inférences formelles mécanisées informatiquement et leurs contreparties dans le monde physique où évolue un robot. Ainsi reformulée, la question centrale débattue devient : comment rendre *effectives* les inférences formelles ?

I.1. Ceci est une pipe - Ceci n'est pas une pipe

Le paradigme[2] dominant en programmation des robots, comme en science cognitive, est celui de la modélisation de l'environnement. La modélisation d'un *phénomène* suppose 2 éléments indissociables :

- un *système formel* ou système de manipulation de *signes* qui mécanise le raisonnement

---

[1] La chose la plus incompréhensible a propos du monde est qu'il soit compréhensible

[2] Au sens de Kuhn : noyau dur épistémologique, théorique, formel et méthodologique dont les qualités sont telles qu'il suscite l'adhésion consensuelle et sans restriction d'une large majorité des chercheurs d'un domaine donné à un moment donné.

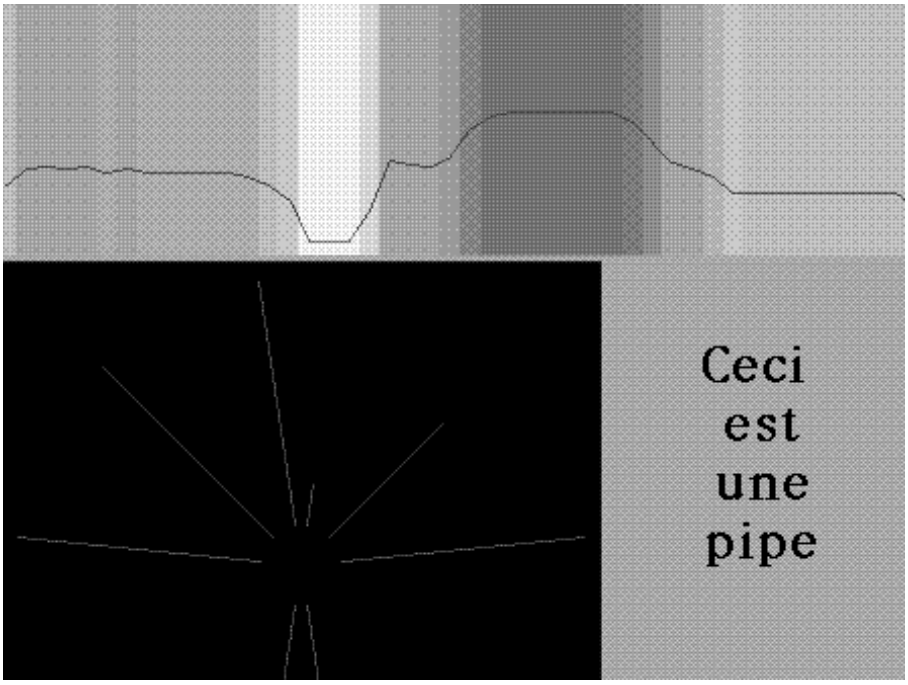


Figure 1 : ceci est une pipe

Figure 2 : ceci n'est pas une pipe

- Une *interprétation* qui définit mathématiquement une analogie entre le système formel et un phénomène modélisé. L'interprétation met en relation les signes utilisés avec les *objets* et *ensembles d'objets* étudiés. L'interprétation est indispensable, car c'est elle qui est censée rendre effectifs les calculs du système formel en permettant de les coupler avec l'environnement.

Dans la suite, nous appellerons ce paradigme « l'approche *F+I* » (système Formel + Interprétation).

Une telle approche suppose donc implicitement d'admettre les 2 postulats suivants :

- *Le robot est capable, tout comme son concepteur, chacun à travers son appareil sensori-moteur, de percevoir une STRUCTURE ENSEMBLISTE (objets et ensemble d'objets) de l'environnement.*
- *Le robot et son concepteur peuvent percevoir la MEME structure ensembliste.*

Dans le paragraphe II nous contestons les fondements de cette approche en mettant en doute ces deux postulats. L'argumentation est double et peut se résumer par la formule, « Ceci est une pipe - Ceci n'est pas une pipe ».

La figure 1 montre comment l'un de nos sujets d'expérience, le robot Khépéra, « voit » une pipe. On conçoit aisément les difficultés que peuvent avoir Khépéra et son concepteur pour se mettre d'accord sur l'ensemble des objets qu'il convient de ranger dans la catégorie « pipe ». Les différences entre les appareils sensori-moteurs du programmeur et du robot rendent en grande partie illusoire qu'ils puissent partager une même structure ensembliste de l'environnement (postulat 2).

La figure 2, célèbre tableau de Magritte, pose la question « qu'est ce qu'une pipe ? ». Est-il possible à un être sensori-moteur de se prononcer définitivement et catégoriquement sur ce qu'est une pipe ? Lui est-il possible de rester constant quant à son propre jugement de ce qu'est une pipe ? « Ceci » n'est certainement pas une pipe si vous voulez la fumer, mais que répondrez-vous si on vous la montre en vous demandant ce que c'est ? Si une pipe est brisée en deux, pourrez vous la fumer ? Certainement pas, serait-elle moins « pipe » pour autant ? En combien de morceaux faudrait-il qu'elle soit brisée pour perdre sa qualité de pipe ? Les magnifiques et brûlants narguilés qu'on vous propose dans les ruelles d'Istanbul, sont-ils des pipes ? Sans doute, si vous ressentez une nécessité de fumer. Certainement pas s'il s'agit d'en ranger un dans votre poche.

L'ensemble des pipes n'est pas clairement défini. « Ceci n'est pas une pipe » va donc nous entraîner très loin. En remettant en cause la validité du premier postulat, nous nous attaquons aux fondements même de la formalisation. En effet, le concept d'appartenance à un ensemble est parfaitement dual de la notion de vérité. Il y a toujours équivalence entre la vérité d'une proposition et la définition d'un certain ensemble. Munir un phénomène d'une structure ensembliste revient à établir la vérité (ou la fausseté) des propositions correspondantes. Finalement, ce que nous remettrons en cause, c'est qu'il soit possible pour des êtres sensori-moteurs de parler et d'accéder à des *vérités objectives* indépendantes de leurs connaissances préalables et de leurs états sensori-moteur et mental.

I.2. Etre pipe ou ne pas être pipe, telle n'est pas la question

F+I est l'aboutissement moderne de la très ancienne association entre raison et discours. Le système formel F y est un « langage » avec lequel il est à la fois possible de « parler » et de « penser » l'environnement.

Une alternative à cette conception langagière de la pensée a toujours existé. Elle associe pensée et image. Dans ce cadre, nous proposons une théorie alternative des systèmes cognitifs sensori-moteurs, baptisée approche *F+D* (système Formel + Description). L'essence de cette approche est de proposer de rendre effectifs les calculs formels par un autre moyen que l'interprétation, appelé description.

L'approche F+D est introduite au paragraphe II. La théorie mathématique qui la sous-tend est présentée dans un autre article de ce même numéro intitulé *Fondements mathématiques de l'approche F+D*.

L'approche F+D suppose 2 éléments indissociables :

- un système formel qui mécanise le raisonnement ;
- une *description* qui définit mathématiquement une relation entre le système formel et le *flux d'information*s sensori-motrices du robot. Une description permet, ainsi, de rendre effectifs les calculs du système formel en les couplant directement avec les capteurs et les actionneurs du robot.

Dans l'approche F+I les signes manipulés par le système formel sont des *symboles* correspondant à des objets et classes d'objets de l'environnement. Dans l'approche F+D les signes utilisés sont les *valeurs* des variables sensorielles, motrices et internes du robot. Lorsque le robot interagit de manière pertinente avec son environnement toutes ces variables ne sont pas indépendantes, certaines varient de manières étroitement corrélées. Par exemple, lorsqu'on fume une pipe, les sensations olfactives, tactiles, kinesthésiques et visuelles ne sont pas quelconques. Odeur du tabac, chaleur, poids de la pipe et fumée forment un tout perceptif cohérent. La perception des unes rendent probables celles des autres. Que l'une manque ou diffère, révèle en général un fait important. Les signes manipulés par le système formel, c'est-à-dire les valeurs de ces différentes variables, ne sont pas des symboles : ils n'ont, pris isolément, aucune contrepartie dans l'environnement, aucune valeur distale. Aucune des informations transmises par le nez, la main, le bras ou la rétine, n'est en soit significative. Pas plus que la couleur d'un pixel sur une rétine n'a de signification et ne permet de se faire une idée quelconque de l'image. C'est l'ensemble des interdépendances entre ces variables qui est caractéristique et que nous avons baptisé « description ». La connaissance des valeurs de certaines d'entre elles permet de prédire les valeurs probables des autres. Lorsqu'on saisit une pipe qui fume, on s'attend à la sensation de chaleur. Ceci s'applique aussi bien aux variables motrices qu'aux variables sensorielles. Si l'on tient une pipe allumée, il est bien probable que les prochains ordres moteurs envoyés au bras soient pour la porter à la bouche. Etre pipe ou ne pas être pipe, telle n'est plus la question. Seule reste la question de savoir quelles actions peuvent être prises étant donné la situation considérée.

Dans l'approche F+I, l'interprétation est donnée *a priori* par le concepteur. Dans l'approche F+D les descriptions sont apprises, certes sur la base de connaissances préalables fournies par le concepteur, mais surtout petit à petit, par la somme des expériences cumulées du robot. Construire une description va consister pour le robot à élaborer des structures internes rendant compte des dépendances observables. Utiliser l'approche F+D lui permettra de percevoir, décider et agir en menant des calculs formels basés sur ces structures internes acquises expérimentalement.

L'approche F+D repose sur un fondement mathématique parfaitement clair : les probabilités. Nous proposons de coder les dépendances entre variables sensorielles, motrices et internes, i.e. ce que nous appelons : descriptions, par des distributions de probabilités. Nous montrons comment les règles de l'inférence probabiliste permettent de faire des raisonnements aussi riches et complexes que ceux faits avec les règles formelles logiques fondant l'approche F+I qui apparait ainsi comme une *approximation* de F+D, historiquement justifiée car souvent suffisante, telles la mécanique newtonienne ou la théorie des gaz parfaits. Nous explicitons comment ces calculs formels probabilistes sont rendus effectifs par la prédiction des valeurs probables des variables motrices. Nous expliquons comment le principe de maximum d'entropie fournit un cadre théorique général et clair pour gérer et comprendre les problèmes d'apprentissage.

II. CRITIQUE DE L'APPROCHE F+I DES SYSTEMES SENSORI-MOTEURS

Mais le rapport du langage à la peinture est un rapport infini. Non pas que la parole soit imparfaite, et en face du visible dans un déficit qu'elle s'efforcerait en vain de rattraper. Ils sont irréductibles l'un à l'autre : on a beau dire ce qu'on voit, ce qu'on voit ne loge jamais dans ce qu'on dit, et on a beau faire voir, par des images, des métaphores, des comparaisons, ce qu'on est en train de dire, le lieu où elles resplendissent n'est pas celui que déploient les yeux, mais celui que définissent les successions de la syntaxe.

Michel Foucault[1]

... les terrifiants pépins de la réalité !

Jacques Prévert

II.1. Principe de l'approche F+I

Concevoir la théorie formelle d'un certain phénomène consiste à produire un système de manipulation de signes obéissant à des règles purement syntaxiques et dont le comportement présente des analogies intéressantes avec le phénomène étudié. Un tel système de manipulation de signes est appelé système formel.

L'analogie entre le système de manipulation de signes et le phénomène est donnée par une interprétation qui met en relation les signes utilisés et les objets et ensemble d'objets étudiés.

La caractéristique la plus fondamentale et la plus intéressante des systèmes formels est qu'ils peuvent exister et fonctionner par eux-mêmes, indépendamment du phénomène qu'ils sont censés représenter. On peut produire de manière systématique, à partir de suites de signes et à partir de règles syntaxiques, de nouvelles suites de signes. Point besoin pour cela d'attacher aucune espèce de signification à aucun des signes. Point besoin de savoir de quoi on parle pour pouvoir en parler avec certitude ! Cette caractéristique permet d'assurer qu'en donnant le même système formel à deux mathématiciens (ou à deux machines), ils (elles) seront capables de produire les mêmes suites de signes. Voilà qui assure ce qu'on appelle communément la rigueur d'une théorie. Voilà mise en pratique et en équation l'ambition aristotélicienne et leibnizienne de voir inéluctablement se mettre d'accord les différents « raisonneurs » partant des mêmes prémisses.

On peut envisager de produire et d'étudier des systèmes purement formels en dehors de toute analogie avec une quelconque autre réalité. Cependant, l'intérêt de ce « jeu » syntaxique n'est pleinement révélé que par la donnée d'une interprétation, quand les objets et ensemble d'objets étudiés peuvent être dénotés par des signes, les règles de manipulation de signes mises en correspondance avec des opérations sur ces objets et ces classes, et les chaînes de signes produites interprétées en terme de prédictions (ou d'explications) sur le comportement du phénomène étudié. Alors, et alors seulement, nous aurons une théorie formelle de notre phénomène.

L'interprétation est une notion mathématique bien définie. Elle peut même être, elle-même, formalisée comme le prouvent les travaux de Tarski et Carnap sur la sémantique extensionnelle et la théorie des modèles. Pour simplifier, disons qu'une interprétation est définie par :

- l'ensemble des objets du phénomène modélisé appelé domaine de discours D ;
- une application (X) de C (l'ensemble des symboles de constante du système formel) dans D, qui associe à chaque symbole de constante un objet de D ;

---

[1] « *Les mots et les choses* » [Foucault66]

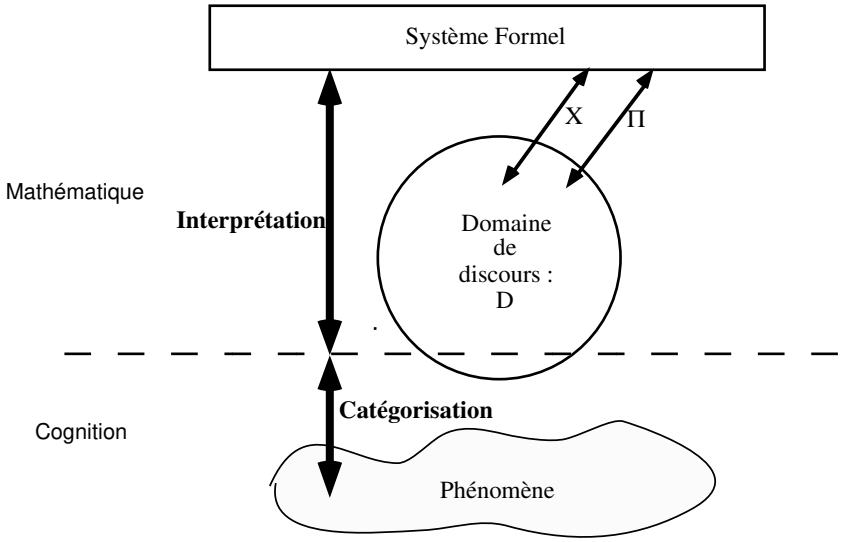


Figure 3 : principe de l'approche F+I

- une application (Π) de P (ensemble des symboles de prédicat du système formel) dans l'ensemble des relations sur D (classes d'objets) qui associe à chaque symbole de prédicat n-aire une relation n-aire sur D.

La notion d'interprétation postule donc implicitement que le phénomène modélisé ait une structure ensembliste et que cette structure soit « accessible » aux « utilisateurs » du système formel. Il faut, en effet, que ces utilisateurs aient la capacité de distinguer parfaitement et sans ambiguïté les uns des autres les objets de D, les ensembles d'objets de D et les ensembles d'ensembles d'objet de D (une relation n-aire est un ensemble d'ensembles à n éléments d'objets de D). De façon duale[1], il faut que ces utilisateurs aient la capacité d'assigner sans ambiguïté la valeur « vrai » ou « faux » à n'importe quelle relation entre éléments de D correspondant à un prédicat défini dans le système formel. S'ils ont cette capacité, alors, les utilisateurs du système formel pourront se servir de sa mécanique syntaxique pour démontrer des « théorèmes » qui seront, sans doute, pour eux interprétables comme des « vérités » dans le domaine de discours et pour le phénomène modélisé.

Nous appellerons *catégorisation* l'opération consistant, pour chaque utilisateur du système formel, à « projeter » sur le phénomène cette structure ensembliste (voir figure 3).

Si l'interprétation est une notion mathématique parfaitement définie, la catégorisation, elle, est une notion cognitive qui prend en compte nécessairement les capacités perceptives, intellectuelles et motrices de l'utilisateur considéré. Adopter cette définition de l'interprétation, et notamment le concept de domaine de discours, c'est admettre que cette catégorisation soit possible. Non seulement qu'elle soit possible, mais qu'elle puisse être mise en œuvre avec toute la rigueur nécessaire, afin de distinguer, sans aucune ambiguïté, les limites des ensembles correspondant aux symboles de constantes et de prédicats du système formel. Adhérer sans réserve à cette thèse revient purement et simplement à écarter tout le problème de la perception et de l'action, toutes les difficultés dues à l'incertitude et l'incomplétude, et toutes les si anciennes, si nombreuses et si subtiles réflexions épistémologiques sur la triade « réalité - observateur - représentation » et son pendant « référé - signifié - signifiant ».

---

[1]Cette dualité est à mettre en parallèle avec celle entre « Théorie des modèles » et « Théorie de la démonstration » [Kleene67].

II.2. Les systèmes formels comme langages des sciences

Au cours du XIX^e et du XX^e siècle, les théories formalisées prennent un essor considérable. Toutes se formalisent, la mathématique bien sûr, la physique, la chimie, l'astronomie, toutes les sciences dites « dures », mais aussi des pans entiers des sciences de la vie ou des sciences de l'homme. La linguistique, par exemple, voit se développer une énorme branche entièrement formalisée, l'économie ou la sociologie aussi. Même si pour de nombreuses sciences une formalisation précise ou exhaustive n'est pas possible, l'influence du paradigme formel est partout et les méthodes de raisonnement correspondantes deviennent un passage quasi obligé à la plupart des démarches scientifiques. Le succès est à la clef et l'extraordinaire progrès de la science moderne est en grande partie dû à la démarche formalisatrice.

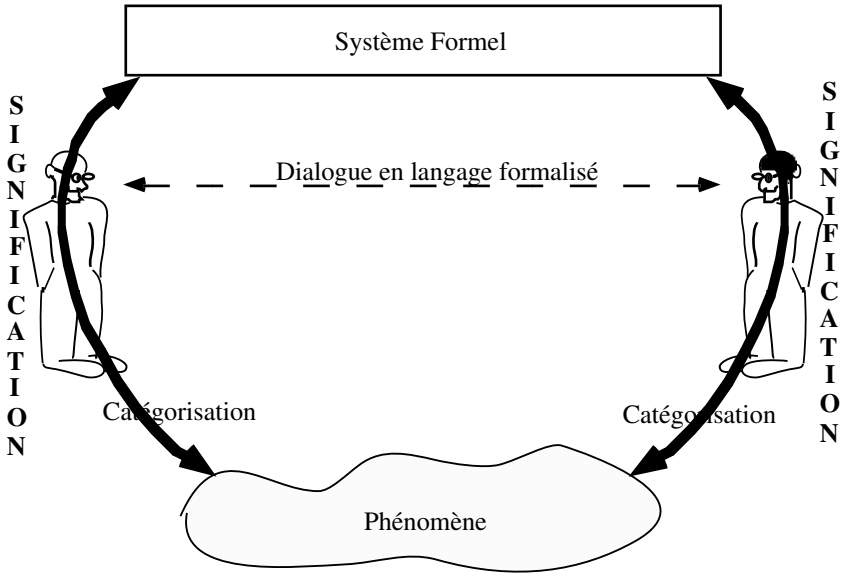


Figure 4 : les systèmes formels comme langage des sciences

Le succès de cette démarche s'explique aisément, ses avantages sont en effet multiples :
- garantir une rigueur « maximum » ;
- permettre une certaine « mécanisation » du raisonnement qui écarte une grande partie des sources d'erreur ;
- inciter à la réduction, l'abstraction et la simplification ;
- offrir un langage commun, supposé non ambigu, pour communiquer et débattre ;
- obliger à un jugement intrinsèque sur la validité de la théorie avancée, par exemple, en examinant des critères tels que consistance et complétude ;
- codifier le jugement extrinsèque sur la qualité des prédictions fournies ;
- enfin, et peut-être surtout, conduire à de formidables synthèses du savoir et d'éblouissants progrès scientifiques lorsque le même système formel s'applique à plusieurs interprétations différentes, c'est-à-dire lorsque deux phénomènes parfois en apparence complètement différents ont en commun certaines structures ou certains principes de fonctionnement.

Cependant, si l'on regarde de près ces diverses tentatives de formalisation, on s'aperçoit que dans la très grande majorité des cas, seule la première moitié du chemin a été parcourue. Un système formel est construit mais son interprétation n'est pas donnée de manière explicite et rigoureuse. La catégorisation est laissée à l'appréciation de chacun. Qu'à cela ne tienne ! Le système formel est dans ce cadre utilisé par les scientifiques pour communiquer et argumenter. Ils se comprennent mutuellement si, et seulement si, ils attachent des significations assez proches les unes des autres à ce système. De ce point de vue, ils sont alors dans une situation de communication tout à fait comparable à celle que l'on rencontre avec l'usage du langage naturel. D'ailleurs, comme pour le langage naturel, les situations où la communication est intéressante sont celles ou les

compréhensions de chacun sont assez proches pour permettre l'échange, tout en étant assez différentes pour susciter le débat. Si l'interprétation était spécifiée rigoureusement, le débat scientifique n'aurait pas de raison d'être. Deux scientifiques ne peuvent qu'être d'accord sur un raisonnement purement formel. Leur seule latitude pour débattre est dans la signification que chacun d'eux attache aux résultats d'un raisonnement formel. La signification de ces résultats et calculs (théorèmes, équations, résultats numériques) est laissée à l'appréciation de chacun, qui dispose de son sens commun, de sa culture du domaine, de son appareillage expérimental et de ses sens pour juger de leur pertinence et de leur portés. C'est là que s'exprime réellement la compétence de ces hommes, par leurs capacités intimes à comprendre, appréhender et traduire les résultats formels en termes de prédiction sur le phénomène modélisé.

II.3. L'informatique : une nouvelle dimension de la formalisation

L'importance de la démarche F+I a acquis une nouvelle dimension au cours de la seconde guerre mondiale avec la naissance de l'ordinateur. D'une part, parce que l'informatique est l'incarnation électronique du concept de système formel. Données et programmes, sont les éléments constitutifs

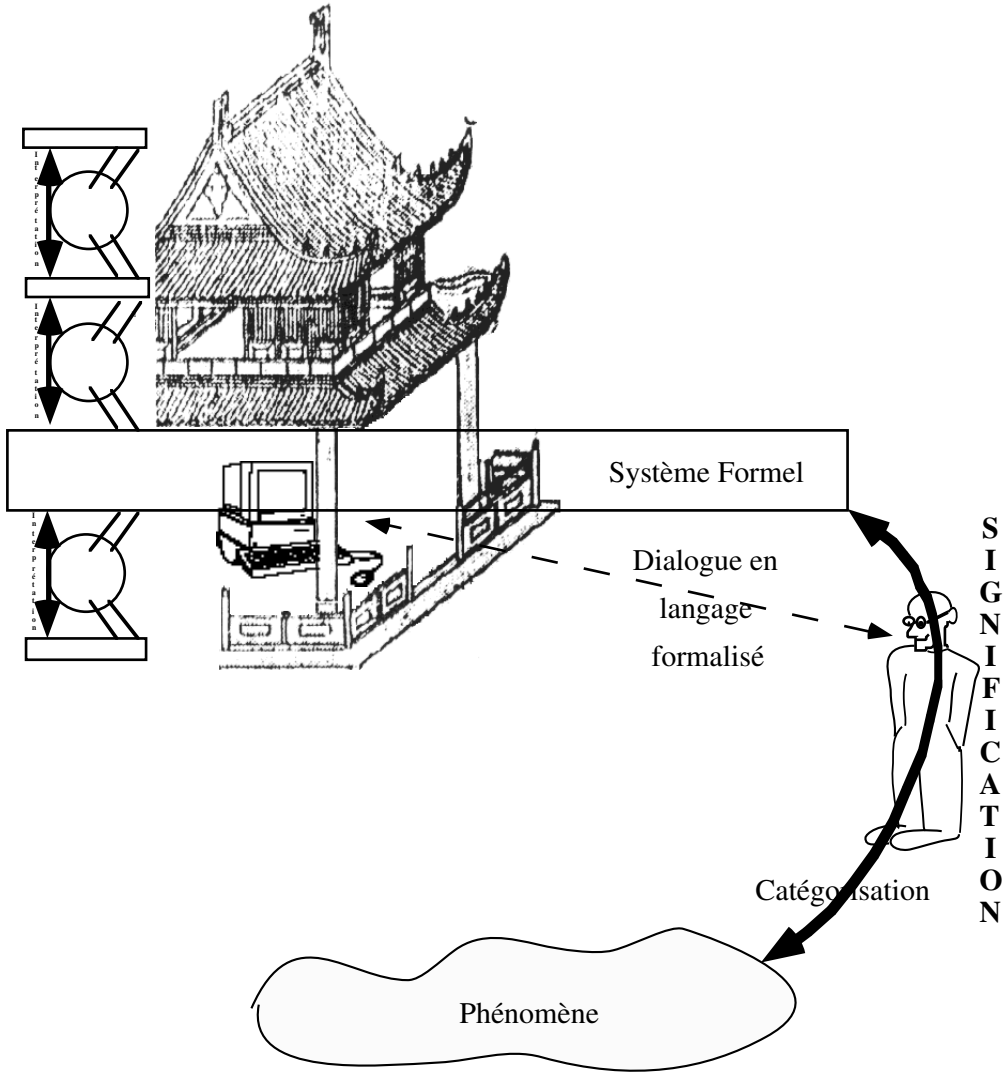


Figure 5 : l'ordinateur dans sa « chambre chinoise »

Il y manipule les « idéogrammes » mais pas les « idées » correspondantes.

d'un système formel, et l'ordinateur, la « mécanique » qui automatise les inférences. D'autre part, parce que l'informatique a ouvert des possibilités entièrement nouvelles à la démarche F+I. Avoir une théorie scientifique formalisée permettant d'élaborer de nouveaux modèles et de concevoir des dispositifs technologiques est une chose, mais disposer de modèles formels, interrogeables en temps réel, pour prendre des décisions ou commander une machine en est une tout autre !

Les programmes d'un ordinateur peuvent être vus comme un ensemble de systèmes formels communiquant les uns avec les autres. Quand il s'agit de faire communiquer entre eux deux programmes, spécifier une interprétation est une tâche relativement facile, étant donné que ces deux programmes sont de nature formelle et donc, par essence, catégorisable.

Reste le problème du lien entre les différents systèmes formels (programmes) présents dans l'ordinateur et le phénomène modélisé. Là encore, on constate que cette interprétation n'est pas donnée de manière explicite et rigoureuse. La machine communique avec son utilisateur en langage formalisé sans avoir aucunement besoin de rattacher ses inférences au phénomène (voir figure 5). Ce dialogue formalisé repose sur le système formel correspondant à l'interface homme-machine. Cette communication est possible, car l'utilisateur et l'ordinateur partagent une même interprétation de ce système formel. Ils sont capables d'identifier les mêmes objets et ensembles d'objets dans l'environnement formel de ce dialogue. Par exemple, l'utilisateur d'un traitement de texte et le programme correspondant savent reconnaître les mêmes caractères, ils sont d'accord sur ce qu'est un mot (une suite de caractère entre deux blancs ou entre un blanc et un signe de ponctuation) ou sur ce qu'est un alinéa. Tout l'intérêt de l'informatique vient de ce que l'utilisateur accorde aux objets et ensemble d'objets manipulés au cours de ce dialogue une autre signification. Il attribue aux manipulations formelles effectuées par la machine, une signification relative au phénomène modélisé. Cette deuxième interprétation échappe totalement aux programmes - alors que c'est celle qui importe à l'utilisateur. C'est le sens du texte et l'esthétique de la mise en page qui compte pour l'utilisateur du traitement de texte, mais il est bien content d'avoir l'ordinateur pour l'aider à en manipuler la forme.

II.4. Difficultés de l'approche F+I des systèmes sensori-moteur

Les deux paragraphes précédents ont montré comment l'approche F+I échappait à la nécessité de spécifier une interprétation lorsqu'on l'applique, soit aux sciences, soit à l'informatique. Quand l'approche F+I est appliquée à un système sensori-moteur, comme nous allons le voir, il n'est plus possible d'échapper à la spécification rigoureuse d'une interprétation. Cette approche se heurte donc dans ce cas, de plein fouet, aux insurmontables difficultés de la catégorisation.

Le phénomène modélisé par un système sensori-moteur est son propre environnement physique. Le système interagit directement avec cet environnement, il fait partie intégrante du phénomène qu'il modélise. Cette interaction ne passe plus par la médiation d'un utilisateur humain, mais par le biais des capteurs et actionneurs du robot. Dès lors, les calculs formels effectués par le robot doivent pouvoir se traduire en termes sensori-moteurs. Pour le robot, chaque signe manipulé doit impérativement référer à un objet ou une classe d'objet de l'environnement. L'interprétation du système formel devient indispensable. Il doit donc exister un programme capable de faire la catégorisation. C'est-à-dire capable de combler l'immense fossé séparant, d'une part, les signaux bruts arrivant des capteurs et envoyés aux actionneurs et, d'autre part, la description abstraite et rigoureuse de l'environnement que le système formel suppose.

Le programmeur a sa propre conception de l'environnement. *Via* ses sens, grâce à ses appareils expérimentaux et vu l'ensemble des expériences et des connaissances qu'il a accumulées au cours de sa vie, il est capable de « projeter » une structure ensembliste sur l'environnement. Il peut nommer les objets et ensemble d'objets qu'il observe. Il peut construire un modèle formel de l'environnement et attacher, à chacun des termes et à chacune des propositions de ce modèle, une signification (*i.e.* les relier à l'environnement *via* ses perceptions et ses actions). Il est capable pour certaines des propositions du système formel de dire (ou de faire calculer par une machine) si elles sont

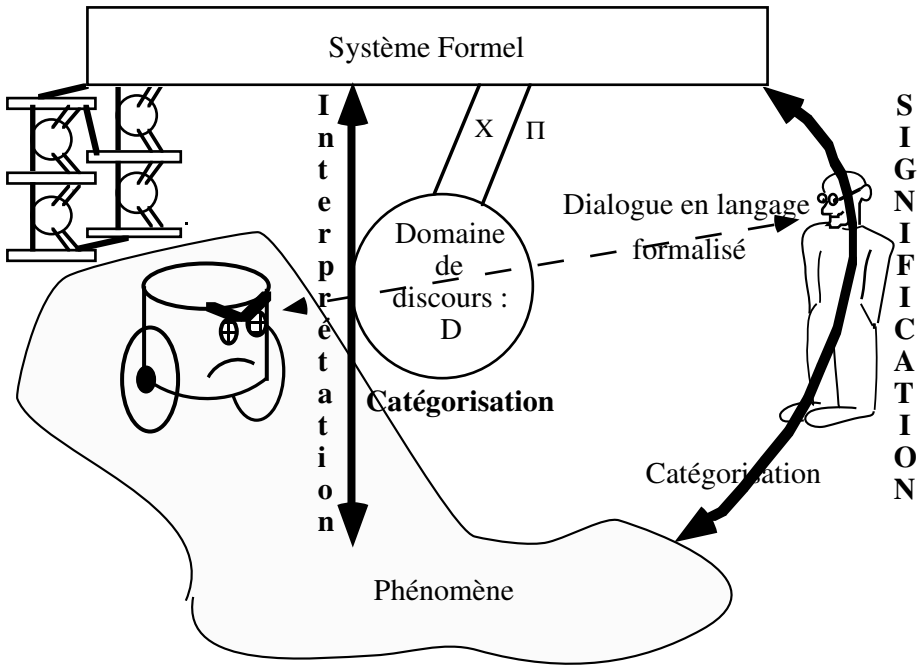


Figure 6 : difficultés de l'approche F+I des systèmes sensori-moteur

dérivables ou réfutables. Il est en mesure de déterminer si la signification de ces formules correspond à une vérité ou une contre-vérité dans l'environnement.

Le programmeur utilise ce système formel pour programmer le robot, c'est-à-dire pour lui spécifier ce qu'il doit faire et comment il doit le faire. Il l'utilise aussi pour recueillir les informations que le robot lui renvoie. Il utilise donc le système formel comme un langage pour communiquer avec le robot. Ils ne peuvent tous deux se « comprendre » que si la signification que le programmeur attache aux signes et aux propositions du système formel est assez proche de l'interprétation de ces signes qu'il a fournie au robot. Concrètement, cela signifie que le programmeur a dû fournir au robot un programme lui donnant la faculté de catégoriser correctement l'environnement à partir de ses données sensori-motrice. « Catégoriser correctement » signifie que le robot doit être capable, à partir de ses données sensori-motrices propres, d'identifier, sans aucune ambiguïté ni incertitude, les objets et ensembles d'objets du domaine de discours sous-tendant le langage formel utilisé par le programmeur pour dialoguer avec lui (voir figure 6).

Une première difficulté dans l'écriture de ce programme tient aux différences existant entre les appareils sensori-moteurs du robot et du concepteur. Le concepteur voudrait que le robot soit capable de distinguer les mêmes objets et ensembles que lui. Cela peut-il être possible quand l'un est fait de chair et de sang, de neurones et d'hormones, et l'autre de métal et de silicium, de micro-processeurs et d'algorithmes?

La principale des difficultés n'est cependant pas là. Un système formel lorsqu'il sert à décrire un environnement physique ne peut pas être *complet.* Il n'est pas possible de prendre en compte exhaustivement les facteurs influençant le phénomène étudié, pour au moins trois raisons de principe :

- l'idée même d'exhaustivité est en contradiction avec le concept de modèle. Un modèle n'est intéressant que s'il est (beaucoup) plus simple que le phénomène étudié.
- les lois de la physique font que très rapidement il faudrait tenir compte de l'univers entier pour pouvoir décrire un phénomène physique même en apparence simple.

- les systèmes dits « chaotiques » prouvent qu'il est impossible de fournir de modèle « simplifié » de certains phénomènes, même élémentaires, dans la mesure où il faut connaître « exactement » leurs conditions initiales pour être capable de prédire leur évolution.

En pratique, pas besoin d'invoquer ces raisons fondamentales. Un système sensori-moteur, qu'il soit vivant ou artificiel, doit de manière évidente prendre des décisions en n'ayant qu'une connaissance extrêmement partielle de son environnement. Les systèmes vivants même les plus simples nous prouvent que cela est possible.

Dès lors, il est impossible de catégoriser de manière absolue les données sensorielles et motrices. Le modèle utilisé pour interpréter les données n'est jamais suffisamment précis et complet pour pouvoir rendre compte de toutes les variations possibles des mesures. Il existe toujours des variables « cachées » qui « brouillent » les données sensori-motrices. Le diagnostic souvent porté sur ce genre de situation consiste à dire que les données sensori-motrices sont « bruitées », voire même quelquefois « aberrantes ». Etrange retournement de causalité qui semble considérer le modèle comme exact et imputer au monde physique quelque tare inconnue.

Une des premières expériences que l'on a menées permet d'illustrer de manière très simple ce problème fondamental.

Le dispositif expérimental consiste en un « robot » qui se limite à un axe de rotation vertical commandé en position (angle Θ) portant une cellule photoélectrique mesurant une intensité lumineuse (intensité I). L'environnement est un seau de plastique vert dans la paroi duquel on a fixé une lampe éclairant l'intérieur (figure 7).

Nous avons « exploré » l'environnement du seau avec le protocole consistant à tirer au hasard (suivant une loi uniforme) des angles θ, à faire tourner l'axe en fonction de cette commande et à mesurer l'intensité ι lue par la cellule après déplacement. Nous avons ainsi obtenu l'histogramme des couples de valeurs (θ,ι) observés présenté figure 8.

L'interprétation de cet histogramme est très claire. Nous observons de hautes valeurs de l'intensité et une saturation de la cellule photoélectrique pour les angles où la cellule fait face à la lampe. Nous observons un deuxième pic d'intensité à 180° du précédent correspondant à la réflexion principale de la lumière sur les parois du seau.

Il est possible de construire de nombreux modèles formels de cet environnement, depuis de simples modèles géométriques jusqu'à de complexes modèles optiques cherchant à décrire précisément les réflexions à l'aide, par exemple, de lancer de rayons.

Sur le papier, le problème possède une symétrie évidente autour de l'axe principal d'illumination de la lampe. En effet, que le robot tourne d'un angle θ vers la gauche ou vers la droite à partir de cette position de référence devrait conduire à des situations parfaitement identiques en ce qui concerne l'illumination de la cellule. Un modélisateur ne manquera certainement pas de tenir compte d'une telle symétrie. Le modèle résultant prédira donc des valeurs identiques pour ι dans ces 2 situations. Or, nous constatons que cette symétrie n'est pas vérifiée par les données expérimentales.

La première idée qui vient à un ingénieur dans ce genre de situation est de penser à un problème de *calibration*. Un modèle est dit « calibré » quand les valeurs des paramètres du modèle permettent une description « satisfaisante » du phénomène. Ici, par exemple, on peut penser que les paramètres géométriques ne sont pas exactement les bons, que l'axe vertical n'est peut-être pas parfaitement vertical ou peut être pas parfaitement centré. On peut alors, soit aller réaligner l'axe vertical, soit complexifier le modèle pour tenir compte du centrage et de l'inclinaison de l'axe. Nous avons essayé de le faire sans y parvenir. Les raisons de l'absence de symétrie sont ailleurs. Quelles sont-elles exactement ? Nous n'en savons rien. De nombreuses possibilités sont imaginables :

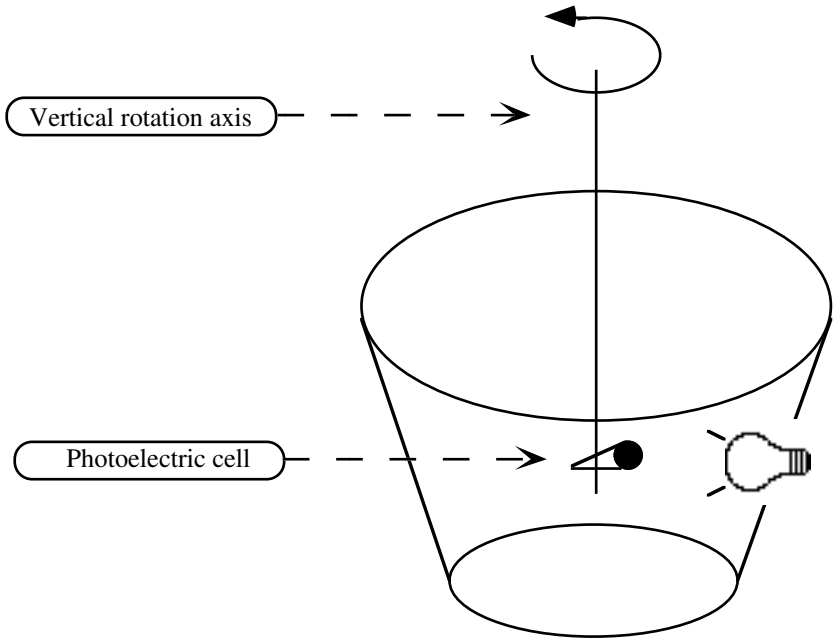


Figure 7 : dispositif expérimental

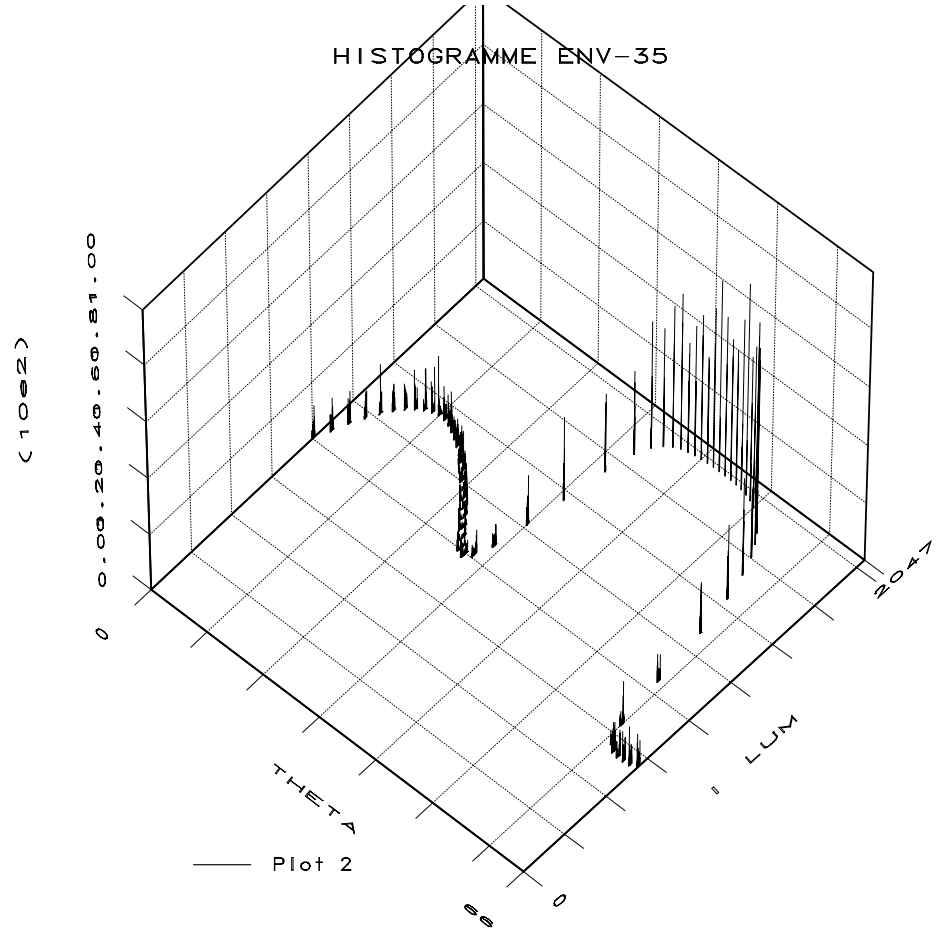


Figure 8 : histogramme des couples de valeurs (θ,ι) observés

- la cellule présente des phénomènes d'hystérésis qui font que les valeurs lues dépendent de l'illumination mais aussi du « passé » des mesures effectuées précédemment (cette explication peut paraître compliquée, elle s'est pourtant avérée partiellement vraie) ;
- un néon était allumé dans la salle au moment de l'expérience, brisant la supposée symétrie de l'illumination ;
- la commande angulaire du robot est plus imprécise pour certains angles que pour d'autres, les angles réellement atteints n'étant plus alors exactement les θ commandés (cette explication semble, elle aussi, bien compliquée, et elle s'est, elle aussi, avérée partiellement vraie, un défaut électronique bloquant à 0 certains bits) ;
- le seau est plus poussiéreux à certains endroits qu'à d'autres ;
- etc..

Complexifier le modèle pour pouvoir tenir compte de ces différentes causes possibles devient, manifestement, beaucoup plus difficile. Il faudrait modéliser l'électronique de la cellule, celle du robot, l'éclairage de la salle et le comportement de la femme de ménage. Finalement, on voit clairement par cette expérience que, même pour un système sensori-moteur et un environnement aussi simple, il est très délicat de prétendre fournir un modèle formel exhaustif et complet de leur interaction.

Cette difficulté à faire le lien entre les signes du système formel et les données sensori-motrices est largement reconnue en intelligence artificielle et en robotique, où elle prend de multiples visages et divers noms dont nous ne retiendrons que celui de « problème de l'ancrage perceptif des symboles »[1].

La parade habituelle pour pallier cette difficulté consiste à contraindre l'environnement dans lequel évolue le système sensori-moteur et les fonctions qu'il doit remplir. Dans l'exemple précédent, comme on l'a déjà dit, le réflexe d'un roboticien serait de rechercher les causes de la dissymétrie pour modifier soit le robot, soit l'environnement, afin de supprimer ce « défaut ». « Défaut » signifiant ici que le monde physique a le mauvais goût de ne pas vouloir correspondre au modèle mathématique supposé le décrire. En robotique industrielle, cette approche est légitime et permet de développer des applications robotiques opérationnelles. L'environnement du robot peut toujours être étroitement contrôlé et les tâches qu'il doit remplir bien spécifiées. C'est cela qui assure la pertinence du modèle employé.

Contrôler et contraindre l'environnement n'est cependant pas toujours possible. Les systèmes sensori-moteurs vivants n'évoluent pas dans des environnements spécifiquement conçus pour eux. Étant donné les buts scientifiques de ce travail, nous nous intéressons à la robotique en *environnement naturel*, c'est-à-dire dans des environnements donnés au robot indépendamment de sa propre conception. L'approche F+I des systèmes sensori-moteurs nous semble dans ce cas dans une impasse, car nous ne voyons pas comment surmonter la difficulté à catégoriser. Seule une approche radicalement différente des systèmes sensori-moteurs peut espérer résoudre un tel problème.

II.5. Notes bibliographiques

Nous n'avons pas pu dans cet article rappeler la définition « formelle » d'une formalisation. Des exemples introductifs de formalisation peuvent être trouvés dans (Hofstadter79, notamment aux chapitres 1, 2, 3, 7 et 8) ainsi que dans (Nilsson80, aux chapitres 1 et 4). Un ouvrage de référence, non mathématique, sur l'approche F+I en science, d'une richesse et d'une profondeur sans pareille, est certainement (Piaget67a). On y trouve notamment une très belle introduction à la formalisation dans le chapitre de Grize (Grize67). Un ouvrage de référence, cette fois-ci mathématique, complet et rigoureux, est sans aucun doute (Kleene67) où on trouve en particulier une intéressante présentation de la théorie du modèle et de la théorie de la démonstration qui clarifie la notion

[1] symbol grounding problem, voir Harnad (Harnad90)

d'interprétation. Une présentation synthétique et claire des systèmes formels du point de vue langage est faite dans (Vauquois76).

Nous n'avons pas pu non plus traiter ici le très vaste et très important sujet des difficultés mathématiques rencontrées par la formalisation et les nombreux théorèmes de limitation qui les explicitent. Une introduction amusante sur ce sujet peut être trouvée dans (Hofstadter79). Dans (Piaget67a), l'article de Ladrière (Ladrière67) en fait une bonne analyse. Les démonstrations de la plupart de ces théorèmes peuvent être trouvées dans (Kleene67). Enfin, une discussion des conséquences philosophiques et surtout cognitives de ces théorèmes peut être trouvée dans (Lucas61).

Les difficultés de l'intelligence artificielle et de la robotique avec la notion d'interprétation ont fait couler beaucoup d'encre. Une analyse de la logique en tant que langage est présentée dans les deux articles de Léo Apostel (Apostel67a) et (Apostel67b) toujours dans (Piaget67a). On pourra se référer notamment à (Harnad89) et (Harnad90) concernant le « symbol grounding problem ».

Diverses critiques de l'intelligence artificielle sont fondées sur des arguments similaires à ceux proposés ici, par exemple celles de Dreyfus (Dreyfuse79), de Searle (Searle81), de Reeke et Edelman (Reeke88), de Malcom et Smithers (Malcom89), de Winograd et Flores (Winograd86), ou enfin, dans l'ouvrage très complet de Bickhard et Terveen (Bickhard95). Pour tenter de surmonter les difficultés que les approches « classiques » (F+I) rencontrent en robotique, Brooks a proposé l'approche réactive (voir (Brooks86a), (Brooks86b), (Brooks89) et (Brooks90)) qui a été à l'origine de nombreux travaux. L'apprentissage et la modélisation, qui au début n'en étaient pas les préoccupations majeures, ont été abordées depuis sous diverses formes (voir, entre autres, (Mataric91), (Koza91), (Brooks91a) et (Horswill93)).

Cette critique de l'approche F+I mériterait d'être précisément située, dans le cadre de l'immense réflexion épistémologique déjà citée, sur la triade « réalité - observateur - représentation » et son pendant « référé - signifié - signifiant ». Il est évident qu'il n'est pas possible dans le cadre de cet article de donner ne serait-ce qu'un aperçu de la bibliographie sur le sujet. Contentons-nous de donner un point de départ intéressant avec le livre de Michel Foucault intitulé « *Les mots et les choses* » (Foucault66) qui a le mérite de présenter toute l'évolution historique des concepts de signe et de langage, et avec le livre de Joëlle Proust « *Question de forme* » (Proust86) plus centré sur les temps modernes et beaucoup plus « technique ».

Au cœur du débat actuel sur le sujet se trouve, d'après nous, la notion d'*objectivité* dans sa double acception : (1) possibilité de juger de la validité d'une théorie indépendamment de toute connaissance sur celui qui l'a formulée, en faisant complètement abstraction de l'observateur, et par conséquent, (2) possibilité qu'à un niveau de description donné on puisse confondre la réalité avec la connaissance qu'on en a. Admettre qu'il puisse exister des modèles indépendants de tout observateur, c'est supposer que le problème de la catégorisation puisse être parfaitement résolu et que l'environnement puisse être muni d'une structure ensembliste indépendamment de toutes considérations sensori-motrices. Toute cette première partie n'a pas eu d'autres buts que de contester cette possibilité.

Le passionnant livre de Roland Omnès (Omnès94) est pratiquement totalement consacré à l'analyse de l'approche formalisatrice en physique et notamment aux difficultés rencontrées. On trouvera aussi de très intéressantes considérations sur ce sujet dans (Poincaré02) même si, évidemment, les progrès de la physique du XX^{e} siècle ne sont pas pris en compte. Ce sujet est aussi amplement abordé dans (Piaget67a) notamment dans les articles d'André Lichnerowicz (Lichnerowicz67), Jean Piaget (Piaget67b), (Piaget67c) & (Piaget67d), Jean Ullmo (Ullmo67) et Louis de Broglie (Broglie67). Le livre d'Ivar Ekeland (Ekeland84) fait un tour d'horizon passionnant et abordable des difficultés et critiques de l'objectivisme (et du déterminisme) et est agréablement complété par son livre de 91 (Ekeland91). Une introduction intéressante aux sciences cognitives, revenant sur le rôle épistémologique de l'objectivisme, est donné dans (Stewart92).

Enfin, une présentation beaucoup plus détaillée de toute cette problématique peut être trouvée dans (Dedieu95) (notamment dans le chapitre 2).

III. INTRODUCTION A L'APPROCHE F+D DES SYSTEMES SENSORI- MOTEURS

Far better an approximate answer to the RIGHT question,which is often vague, than an exact answer to the WRONG question,which can always be made precise.[1]

John W. Tukey[2]

Cette citation de John W. Tukey nous semble bien résumer les difficultés rencontrées par l'approche F+I et indiquer quelle voie suivre pour les surmonter. Comme on l'a vu, l'approche F+I des systèmes sensori-moteurs se pose des questions précises sur les modèles mathématiques décrivant l'environnement des robots et trouve des réponses exactes à ces questions en faisant des calculs formels. Malheureusement, ces modèles mathématiques ont souvent bien peu à voir avec la réalité sensori-motrice d'un robot. Les questions sont rendues précises par la modélisation mais perdent ainsi, trop souvent, toute pertinence pratique. Il nous faut donc reconsidérer la notion d'interprétation en imaginant un autre moyen de rendre effectifs les résultats des calculs. Comme nous allons le voir, le prix à payer pour obtenir ce résultat sera, pour l'approche F+D (Système Formel + Description), de ne pouvoir fournir que des réponses approximatives aux problèmes sensori-moteurs. Cependant, mêmes approximatives, ces réponses seront suffisantes pour permettre à un robot d'interagir avec son environnement.

III.1. Espace sensori-moteur et espace des phases

Pour un robot, nous appelons *espace sensori-moteur* l'ensemble des valeurs possibles pour les signaux venant de ses capteurs et envoyés à ses actionneurs.

Le robot peut, en partie, contrôler son interaction avec l'environnement en choisissant des valeurs pour ses variables motrices. L'ensemble des calculs qui le conduit à faire ces choix se base non seulement sur les variables sensori-motrices, mais aussi sur des variables internes, dites variables d'état. Nous appellerons *espace des phases*, l'ensemble des variables manipulées par le système au cours de ses raisonnements. L'espace sensori-moteur est un sous ensemble de l'espace des phases. L'espace des phases est hautement structuré. Quand les valeurs de certaines des variables sont imposées, les valeurs possibles pour certaines autres sont fortement contraintes.

Prenons comme exemple le petit robot mobile Khépéra construit par l'EPFL[3]. Khépéra est un robot mobile à deux roues, de 57 mm de diamètre et 29 mm de hauteur, pour un poids total de 80g (voir figure 9).

Il est équipé de 8 capteurs de lumières (6 devant et 2 derrière) prenant des valeurs entre 0 et 511 en fonction décroissante de l'intensité lumineuse (stockées dans les variables L_1 ... L_8, voir figure 10). Ces 8 capteurs peuvent aussi être exploités en tant que proximètres infrarouges prenant des valeurs entre 0 et 1023 en fonction décroissante de la distance[4] de l'obstacle (stockées dans les variables V_1 ... V_8, voir figure 10).

---

[1]Il vaut beaucoup mieux une réponse approximative à une question pertinente, souvent vague, qu'une réponse exacte à une question sans intéret, qui peut toujours être rendue précise.

[2]Citation emprunté à [Jaynes95]

[3]Ecole Polytechnique Fédérale de Lausanne

[4] En fait les valeurs de ces capteurs dépendent de plusieurs facteurs: la distance de l'objet, sa couleur, sa matière et la qualité de sa surface.

Khépéra vu de face

Khépéra vu de dessus

Figure 9

Il est commandé par les vitesses de ses roues gauche et droite (stockées dans les variables M_g et M_d).

L'espace sensori-moteur de ce robot est donc un espace de dimension 18 construit sur les variables L_1 ... L_8, V_1 ... V_8, M_g et M_d.

De ces variables sensorielles et motrices de base on tire par prétraitement les 4 variables *Lum*, *Dir*, *Prox* et *Vrot*:

- *Lum* est une variable correspondant à peu près à la direction de plus grande illumination. *Lum* prend une valeur entre 0 et 7 correspondant au numéro (minoré de 1) de la cellule photoélectrique la plus éclairée.
- *Dir* est une variable correspondant très approximativement à la direction de l'obstacle le plus proche (voir figure 10). *Dir* prend des valeurs entières entre -10 (obstacle à gauche du robot) et 10 (obstacle à droite du robot). *Dir* est défini par la formule :

$$Dir = PartieEntiere\left(\frac{-90V_1 - 45V_2 - 5V_3 + 5V_4 + 45V_5 + 90V_6}{9(V_1 + V_2 + V_3 + V_4 + V_5 + V_6) + 1}\right) \qquad [3.1]$$

- *Prox* est une variable correspondant à peu près à la proximité de l'obstacle le plus proche (voir figure 10). *Prox* prend des valeurs entières entre 0 (obstacle très loin du robot) et 15 (obstacle très proche du robot). *Prox* est défini par la formule :

$$Prox = PartieEntiere\left(\max(V_1, V_2, V_3, V_4, V_5, V_6)/16\right) \qquad [3.2]$$

- Le robot est piloté uniquement en rotation à vitesse de translation constante. Les ordres moteurs lui sont envoyés par la variable *Vrot* calculée comme la différence entre la vitesse de la roue gauche et celle de la roue droite. Elle prend des valeurs entières entre -10 (à fond vers la gauche) et +10 (à fond vers la droite).

Ces 4 variables, *Lum*, *Dir*, *Prox* et *Vrot* viennent s'ajouter aux 18 précédentes et on obtient un espace des phases de dimension 22.

Cet espace des phases est effectivement très structuré. Il existe d'innombrables et très fortes dépendances entre ces 22 variables. Par exemple :

- les valeurs des 8 variables Li sont *a priori* corrélées par le niveau moyen d'éclairage ;
- si le proximètre V3 a une valeur très élevée, il est probable que V2 et V4 aient aussi des valeurs élevées ;

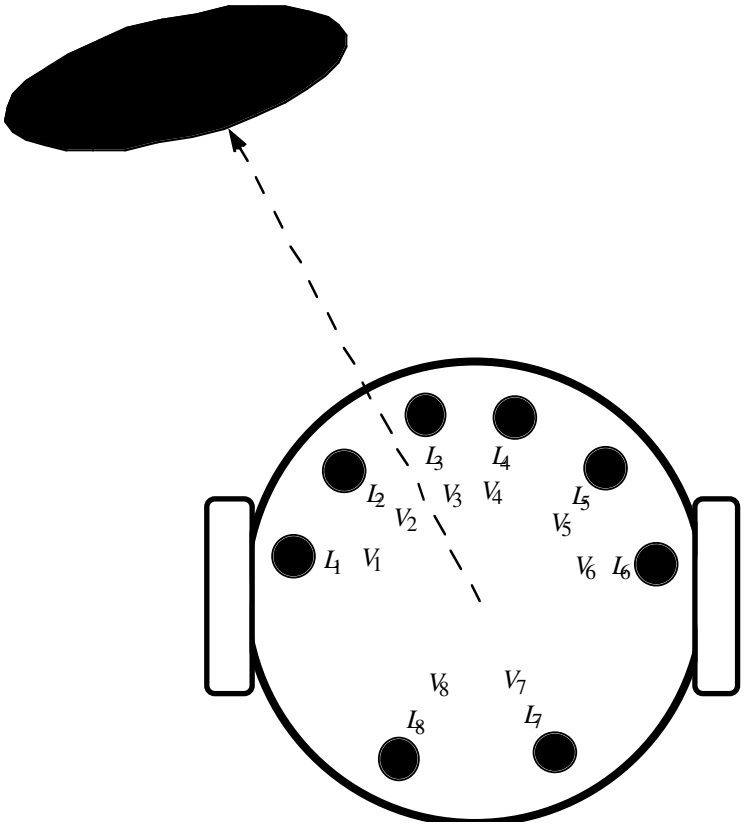


Figure 10 : les variables sensori-motrices de Khépéra

- si Mg et Md sont tous 2 négatifs (le robot recule), il est prévisible que les valeurs de V2, V3, V4 et V5 vont diminuer puisque le robot va s'éloigner des obstacles frontaux ;
- si le robot avance et évite les obstacles alors il est probable qu'une valeur de *Dir* négative (obstacle à gauche) corresponde à une valeur de *Vrot* positive (le robot tourne vers la droite) et inversement ;
- etc..

III.2. Principe de l'approche F+D

Comme l'approche F+I, l'approche F+D repose sur l'emploi d'un système formel manipulant des signes et obéissant à des règles purement syntaxiques. Par contre, les calculs formels ne sont plus rendus effectifs par le biais d'une interprétation reliant signes et objets de l'environnement, mais par le biais de *descriptions* codant les dépendances entre valeurs des variables de l'espace des phases.

Les signes manipulés par le système formel sont les valeurs des variables sensorielles, motrices et internes du robot. Ils n'ont pas pris individuellement de signification particulière. Pas plus que la couleur d'un pixel sur une rétine n'a de signification et ne permet de se faire une idée quelconque de l'image. Leurs seuls référés sont des courants électriques dans des capteurs ou des actionneurs, ou des valeurs de variables informatiques. C'est collectivement, par leurs dépendances mutuelles, que les valeurs des variables de l'espace des phases prennent de l'intérêt. Ce sont ces dépendances que les descriptions vont coder sous forme de distributions de probabilités.

Une description est une notion mathématique tout aussi précisément définie que l'était le concept d'interprétation. Pour simplifier, disons qu'une description est définie comme une distribution de

probabilités sur un sous espace de dimension n de l'espace des phases (défini par le choix d'une base de n variables $X_1 ... X_n$).

Une description est construite par un processus d'apprentissage sur un ensemble de données expérimentales D à partir d'un ensemble de connaissances préalables C données au système par le concepteur. Une description est donc dénotée formellement :

$$P(X_1 ... X_n \mid DC)^1 \qquad [3.3]$$

On peut, si l'on connaît $P(X_1 ... X_n \mid DC)$, calculer n'importe quelle probabilité :

$$P\left(\left[X_i = x_i\right] ... \left[X_j = x_j\right] \mid \left[X_k = x_k\right] ... \left[X_l = x_l\right] DC\right)^2 \qquad [3.4]$$

Ces calculs sont purement formels : comme pour l'approche F+I, à partir de suite de signes et de règles syntaxiques, on produit de nouvelles suites de signes. Cependant, alors que pour l'approche F+I, les règles syntaxiques sont, en général, inspirées par les règles de la logique, de l'algèbre ou de la géométrie, elles sont, pour l'approche F+D dérivées des probabilités.

Dans le cadre sensori-moteur, une description permet de rendre *effectif* ces calculs formels de manière originale. Supposons que S soit l'ensemble de variables sensorielles de la description, M l'ensemble des variables motrices et I l'ensemble des variables internes (ou contextuelles). On peut alors, grâce à la description, traduire les résultats des calculs formels en :

- *Commandes motrices* : P(M | S I D C) permet de choisir les ordres moteurs pour atteindre une situation sensorielle S souhaitée connaissant le contexte I (voir § III.3.1).
- *Prévisions sensorielles* : P(S | M I D C) permet de prédire les conséquences sensorielles d'un jeu d'ordres moteurs dans un certain contexte I.
- *Reconnaissance de contexte* : P(I | S M D C) permet de reconnaître le contexte le plus probable connaissant une situation sensori-motrice (voir § III.3.3).

Ces calculs ont les remarquables propriétés suivantes :

- Ils restent parfaitement valables en situation d'incomplétude. Il est possible de choisir des commandes motrices, de faire des prévisions sensorielles et de reconnaître des situations même si seulement une partie des valeurs des variables de S, M et I sont connues. Par exemple :
  - * P(M | S D C) permet de choisir des commandes motrices connaissant la situation sensorielle mais ignorant le contexte.
  - * P(M | S1 S4 I D C) permet de choisir les commandes motrices connaissant seulement une partie de la situation sensorielle (par exemple parce que certains capteurs sont en panne).
  - * P(I | M1 M3 S1 S4 D C) permet de trouver le contexte le plus probable en ne connaissant qu'une partie de la situation sensori-motrice.
  - * Etc.
- Ils permettent de propager les incertitudes. Il se peut que certaines valeurs des variables de l'espace des phases ne soient pas connues précisément, soit que le capteur ou l'actionneur correspondant fournisse des valeurs entachées d'incertitude, soit que ces valeurs doivent, elles-mêmes, être inférées par un raisonnement probabiliste. A ces variables, correspond alors une distribution de probabilité. Ces distributions peuvent intégralement être prises en compte dans le raisonnement probabiliste (voir § III.3.2).

---

[1] La barre verticale dénote le conditionnement. X1 ... Xn sont connues conditionnellement à D et C.

[2] On note en majuscule les variables et en minuscule les valeurs de ces variables.

III.3. L'approche F+D mise en pratique

Dans ce paragraphe, nous décrivons quelques unes des expériences menées avec le Khépéra afin d'illustrer le fonctionnement et les capacités de l'approche F+D

III.3.1. Utilisation et apprentissage de comportements réactifs

Dans cette première série d'expériences, nous avons cherché à doter Khépéra de comportements purement réactifs vis-à-vis des obstacles. Les commandes motrices sont directement inférées des valeurs sensorielles.

Prenons l'exemple de l'évitement d'obstacle. Le but poursuivi est de donner la capacité à Khépéra d'éviter les obstacles et d'aller en ligne droite lorsqu'il est en espace libre. Nous avons choisi d'utiliser pour cette expérience les 3 variables *Dir*, *Prox* et *Vrot*. La description correspondante va donc être une distribution de la forme :

$$P(DirProxVrot \mid DC) \qquad [3.5]$$

Le protocole expérimental adopté est constitué de 2 phases:
- Une phase d'apprentissage pendant laquelle Khépéra est piloté par un opérateur à l'aide d'un *joystick*. Durant cette phase les données expérimentales D (valeurs des variables sensorielles *Dir* et *Prox* et de la commande motrice *Vrot*) sont enregistrées et les paramètres, permettant de stocker P(*Dir Prox Vrot* | D C) en machine, sont identifiés
- Une phase de restitution où Khépéra évite les obstacles tout seul. Les calculs formels probabilistes sont alors utilisés pour calculer à chaque instant la commande motrice *Vrot* à envoyer aux roues en fonction des données sensorielles *Dir* et *Prox* provenant des capteurs de proximité.

La description P(*Dir Prox Vrot* | D C) obtenue lors de l'apprentissage dépend évidemment des expériences D, mais aussi, et d'une certaine manière : surtout, des connaissances préalables C.

Les connaissances préalables C données par le concepteur peuvent être regroupées en 4 catégories de natures différentes :
- Les *connaissances préalables structurelles* sont définies par l'ensemble des variables X_1 ... X_n sur lesquelles la distribution de probabilité va être construite. Pour notre exemple, comme il a déjà été dit, 3 variables ont été choisies : *Dir*, *Prox* et *Vrot*.
- Les *connaissances préalables de dépendances* sont données par une structure de dépendance[1] permettant de décomposer de manière unique une description comme un produit de distributions plus simples. Ces simplifications essentielles sont fondées sur des considérations d'indépendance conditionnelle entre variables. Pour l'exemple qui nous occupe, les règles des probabilités donnent :

$$P(DirProxVrot \mid DC) = P(Dir \mid DC)P(Prox \mid DirDC)P(Vrot \mid ProxDirDC) \qquad [3.6]$$

Nous pensons *a priori* qu'il n'existe pas de corrélation particulière entre la direction dans laquelle on observe un obstacle et la distance de cet obstacle. Nous traduisons cette connaissance pour notre robot en lui disant que $Prox$ est indépendante de *Dir* conditionnellement à D et C.

$$P(Prox \mid DirDC) = P(Prox \mid DC) \qquad [3.7]$$

---

[1]Une structure de dépendance est une généralisation de la notion de réseaux bayésiens due à J. Pearl [Pearl91].

On obtient donc :

$$P(DirProxVrot \mid DC) = P(Dir \mid DC)P(Prox \mid DC)P(Vrot \mid DirProxDC) \quad [3.8]$$

- Les *connaissances a priori* donnent des valeurs initiales (avant apprentissage) aux distributions ou explicitent une méthode pour les calculer à partir d'autres distributions connues. Une connaissance *a priori* peut aussi, par exemple, exprimer qu'une distribution est supposée ne pas évoluer en fonction des données.
  Nous n'avons aucune raison de penser que le robot sera amené à rencontrer les obstacles suivant une direction privilégiée ou à une distance particulière. Nous adoptons donc comme *a priori* que tous les couples (*Dir, Prox*) sont équiprobables, ce qui se traduit par :

$$P(Dir \mid DC) = P(Dir \mid C) = \frac{1}{21} \quad [3.9]$$

et par :

$$P(Prox \mid DC) = P(Prox \mid C) = \frac{1}{16} \quad [3.10]$$

- Enfin, les *connaissances préalables d'observation*, qui se décomposent en la spécification d'une représentation paramétrique et d'un processus d'apprentissage (identification des paramètres) pour chacune des distributions de la structure de dépendances. Une représentation paramétrique permet d'une part, de représenter ces distributions en machine et d'autre part, d'identifier leurs paramètres par apprentissage.
  Pour cette expérience, on a utilisé une loi de succession de Laplace comme représentation paramétrique de :

$$P([Vrot = v] \mid [Dir = d][Prox = p]DC) = \frac{m_v + 1}{m_{dp} + 21} \quad [3.11]$$

  Où :
  * m_{dp} est le nombre de fois où la situation sensorielle *Dir*=*d* et *Prox*=p a été rencontrée pendant l'apprentissage
  * m_v est le nombre de fois ou *Vrot*=v a été observé pour la situation sensorielle précédente.
  * 21 est le nombre de valeurs possibles pour *Vrot*.

  Pour un nombre important d'expériences, la loi de succession de Laplace se rapproche d'un simple histogramme défini par m_v/m_{dp}. Les termes correctifs 1 et 21 servent à traiter les cas où le nombre d'expériences est faible. En particulier pour $m_v=m_{dp}=0$, on trouve 1/21 soit la distribution uniforme.

Par le calcul formel, on peut alors, en phase de restitution, commander Khépéra en choisissant à chaque instant *Vrot* suivant la distribution P(*Vrot* | *Dir Prox* D C). La commande de Khépéra par cette méthode fonctionne très bien. Après un temps d'apprentissage de quelques dizaines de secondes, on obtient le comportement souhaité (voir figure 11). Le comportement obtenu s'avère être très « robuste » au changement de position, de taille, de forme, de matière, de couleurs[1] et même de vitesse des obstacles, ainsi qu'aux conditions d'éclairage de la scène.

[1]En fait, seul les obstacles noirs et mats, que les proximètres ne « voient » pas du tout, posent un problème.

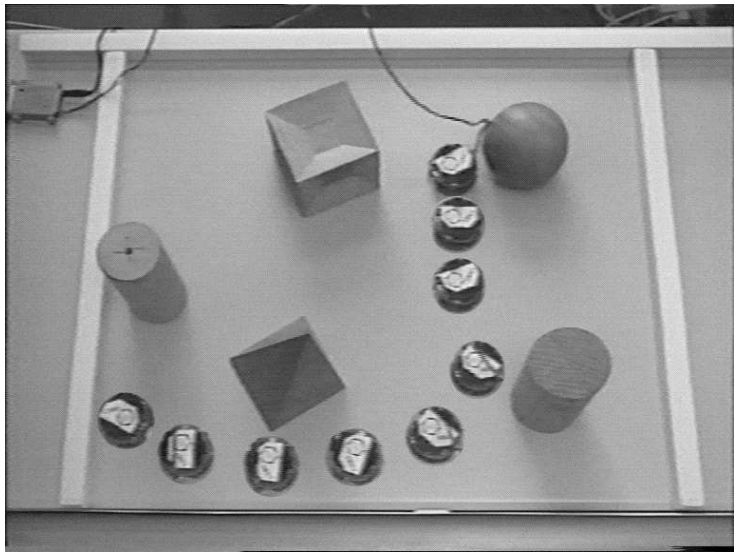


Figure 11 : comportement d'évitement d'obstacles

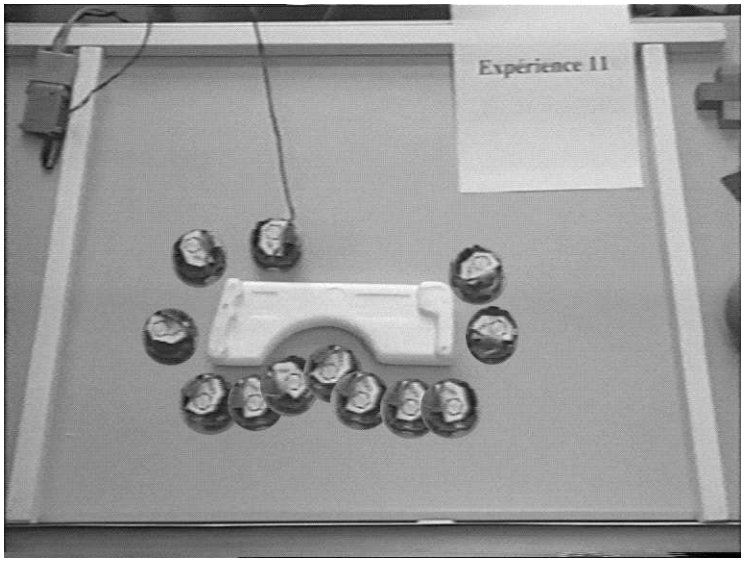


Figure 12 : comportement de suivi de contour

Avec la même méthode et les mêmes connaissances préalables, en changeant l'apprentissage téléopéré au « joystick », on a appris a Khépéra à suivre les murs et contours (voir figure 12), à « déménager les pianos » (i.e. pousser des objets) et à poursuivre les obstacles mobiles. Avec la même méthode, sur la base des variables $Vrot$ et Lum, nous avons appris à Khépéra un comportement photophile représenté par la description P($Vrot\ \ Lum \mid$ D' C').

III.3.2. Composition simple de comportements réactifs

Un but souvent poursuivi et rarement atteint par les approches réactives de la robotique est de pouvoir composer entre eux des comportements appris séparément. Par exemple, ici, nous aimerions obtenir de Khépéra qu'il aille vers la lumière en évitant les obstacles à partir des comportements photophile et d'évitement d'obstacles déjà appris.

L'approche F+D à l'avantage de permettre de poser très clairement ce genre problème:
- Nous cherchons P(*Vrot* | *Lum Dir Prox* D C)
- Nous avons P(*Vrot* | *Dir Prox* D1 C1) et P(*Vrot* | *Lum* D2 C2)
- Nous voulons que (par exemple):
    * en l'absence de données sur la lumière, Khépéra se contente d'éviter les obstacles :

$$P(Vrot \mid DirProxC) = P(Vrot \mid DirProxD_1C_1) \quad [3.12]$$

    * En l'absence de données sur les obstacles, Khépéra se contente d'aller vers la lumière :

$$P(Vrot \mid LumC) = P(Vrot \mid LumD_2C_2) \quad [3.13]$$

    * En l'absence d'obstacle, Khépéra aille vers la lumière :

$$P(Vrot \mid Lum[Prox = 0]C) = P(Vrot \mid LumD_2C_2) \quad [3.14]$$

    * A grande proximité d'un obstacle, Khépéra évite coûte que coûte ceux-ci :

$$P(Vrot \mid LumDir[Prox = 15]C) = P(Vrot \mid DirProxD_1C_1) \quad [3.15]$$

On obtient ainsi une série de contraintes pesant sur la forme paramétrique recherchée de P(*Vrot* | *Lum Dir Prox* D C). Ce genre de problème est en général largement sous-contraint, laissant de nombreuses possibilités ouvertes pour le choix final de P(*Vrot* | *Lum Dir Prox* D C).

Ici, par exemple, les règles des probabilités nous permettent d'introduire une variable *H* pouvant prendre 2 valeurs et d'écrire:

$$\begin{aligned} P(Vrot \mid DirProxLumDC) = {} & P(Vrot \mid DirProxLum[H = 1]DC)P([H = 1] \mid DirProxLumDC) \\ & + P(Vrot \mid DirProxLum[H = 2]DC)P([H = 2] \mid DirProxLumDC) \end{aligned} \quad [3.16]$$

Si l'on considère cette variable comme une manière de choisir entre le comportement photophile et le comportement d'évitement, on peut poser :

$$P(Vrot \mid DirProxLum[H = 1]DC) = P(Vrot \mid DirProxD_1C_1) \quad [3.17]$$

et

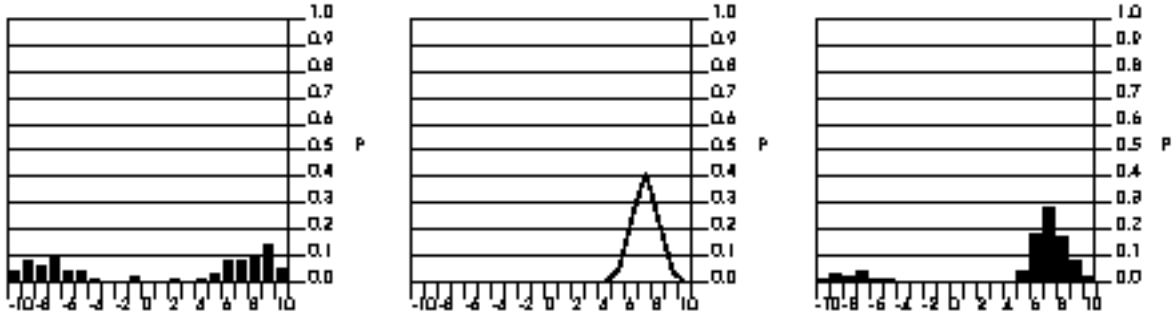
$$P(Vrot \mid DirProxLum[H = 2]DC) = P(Vrot \mid LumD_2C_2) \quad [3.18]$$

Finalement, en posant *a priori* :

$$P([H = 1] \mid DirProxLumDC) = P([H = 1] \mid ProxC) = \frac{Prox}{15} \quad [3.19]$$

On obtient une solution qui satisfait toutes les contraintes ([3.12], [3.13], [3.14] et [3.15]) posées.

En pratique, on obtient ainsi, sans aucun apprentissage supplémentaire, un comportement photophile avec évitement d'obstacle (voir figure 13).



a b c

Figure 13 : composition simple de comportements réactifs
Les ordonées représentent les probabilités P(Vrot | ...)dans diverses sitations :
a : au départ, face à un obstacle à moyenne distance.
Khépéra n'a pas de réelle préférence entre la droite et la gauche
b : comportement photophile, cas où la source lumineuse est à droite.
c : comportement mixte obtenu par la méthode de composition expliquée ci-dessus :
si la source lumineuse est à droite Khépéra préfère éviter l'obstacle par la droite.

III.3.3. Reconnaissance d'objet

Il est bien évident que des comportements un tant soit peu complexe ne peuvent être purement réactifs. En particulier, il est souvent nécessaire de prendre des décisions sur une base de temps plus large que le simple instant présent et sur un domaine spatial plus grand que la pure information sensorielle immédiate.

Dans cette expérience, grâce à l'approche F+D, nous faisons un premier pas dans cette direction. L'expérience va consister pour Khépéra à apprendre à reconnaître un certain nombre d'objets. Cet apprentissage est totalement non supervisé. On présente les objets sans dire au robot à quel objet il a à faire, ni même s'il s'agit d'un nouvel objet ou d'un objet déjà connu. Khépéra doit donc les classer de lui-même, et créer, de lui-même, une nouvelle classe quand un nouvel objet lui est présenté.

Le protocole expérimental est le suivant :

- Khépéra est placé en espace libre face à un objet. Il avance en ligne droite jusqu'à ce que ses proximètres détectent l'objet puis il commence à suivre les bords de cet objet grâce au comportement de « suivi de contours » décrit précédemment (voir figure 12).
- Pendant qu'il effectue ainsi le tour de l'objet, Khépéra enregistre les commandes *Vrot* qu'il envoie à ses moteurs. Il arrête cet enregistrement lorsque la somme algébrique des commandes *Vrot* égale 360° (Ce qui correspond en pratique à une estimation plutôt grossière de la fin du tour). L'enregistrement des valeurs de *Vrot* est lissé par une moyenne mobile sur 50 valeurs successives. Cette série de données lissées est utilisée pour calculer 4 nouvelles variables qui vont servir à discriminer les objets :
  * *Nvd* « Nombre de Virages à Droite » : Un « virage à droite » est une suite d'au moins 30 valeurs négatives de *Vrot* lissées.
  * *Nvg* « Nombre de Virages à Gauche ».
  * *Per* « Périmètre » : Nombre total de valeurs de *Vrot* enregistrées en un « tour ».
  * *Lgc* « Longueur du plus Grand Coté » : Plus grand nombre de valeurs de *Vrot* entre 2 « virages ».

  Par exemple, pour un « tour » effectué autour de l'objet de la figure 12 on obtient : *Nvd* = 1, *Nvg* = 4, *Per* = 710, *Lgc* = 173.
- A partir des valeurs ainsi obtenues pour ces 4 variables, Khépéra, par inférence probabiliste, catégorise l'objet soit dans une classe correspondant au prototype d'un objet déjà rencontré

(qui est alors enrichi par cette nouvelle occurrence) soit dans une nouvelle classe qu'il crée parce qu'il pense être en présence d'un objet nouveau.

Les variables manipulées sont les suivantes : Nvd, Nvg, Per, Lgc et O. O est la variable permettant d'identifier les classes d'objets. Lorsque n classes d'objets ont déjà été identifiées, O peut prendre n+1 valeurs différentes (la valeur O=0 est réservée à une classe de référence « neutre » et est utilisée pour détecter la nouveautée). La description qui nous intéresse est donc :

$$P(NvdNvgPerLgcO \mid DC) \qquad [3.20]$$

On suppose que Nvg, Nvd, Per et Lgc sont des variables indépendantes conditionnellement à O. On obtient donc comme structure de dépendance :

$$P(NvdNvgPerLgcO \mid DC) = P(O \mid DC)P(Nvd \mid ODC)P(Nvg \mid ODC)P(Per \mid ODC)P(Lgc \mid ODC) \qquad [3.21]$$

Comme connaissances préalables d'observation, on choisit des gaussiennes comme représentation paramétrique des 4 distributions à identifier. Pour chaque valeur de O on identifie donc une moyenne et un écart type pour Nvg, Nvd, Per et Lgc.

La valeur O=0 est traitée à part. $P(NvdNvgPerLgc \mid [O=0]DC)$ est fixée *a priori* uniforme. Cette distribution va être utilisée comme référence pour détecter la nouveauté.

Pour répondre à la question qui lui est posée, Khépéra calcule pour chaque valeur possible de O:

$$P(O \mid NvdNvgPerLgcDC) = \frac{P(O \mid DC)P(Nvd \mid ODC)P(Nvg \mid ODC)P(Per \mid ODC)P(Lgc \mid ODC)}{\sum_{O=0}^{O=n+1} P(O \mid DC)P(Nvd \mid ODC)P(Nvg \mid ODC)P(Per \mid ODC)P(Lgc \mid ODC)} \qquad [3.22]$$

Il trouve ainsi la valeur de O la plus probable. Si cette valeur est i≠0, Khépéra considère avoir reconnu l'objet i et modifie les paramètres correspondants pour tenir compte de cette nouvelle instance. Si cette valeur est 0, Khépéra considère être en face d'un nouvel objet. Il crée une nouvelle classe (nouvelle valeur possible pour O) sur la base des données qu'il vient de recueillir.

O est une variable plus abstraite et plus synthétique que les variables sensori-motrices. O a une constante de temps d'acquisition plus longue que les variables sensori-motrices puisqu'elle suppose l'intégration dans l'espace et dans le temps d'un certain nombre de valeurs des ces variables de bases. O peut à son tour être utilisé dans de nouvelles descriptions. On peut espérer ainsi construire de véritables hiérarchies de descriptions, de plus en plus abstraites, correspondant à des comportements de plus en plus complexes, avec des constantes de temps de plus en plus longues. Nous en verons un exemple à l'alinéa suivant.

La figure 15 montre les objets qui ont été appris avec cette méthode. Chacun a été présenté 4 fois, en suivant un ordre aléatoire. Seuls les 2 objets à base carrée, en haut à droite de la photo, ont été confondus. Ils sont en fait parfaitement indiscernables avec les 4 variables choisies. Dans tous les autres cas, les objets ont été reconnus sans erreur. La nouveauté a, elle aussi, été identifiée à chaque fois sans erreur et sans oublis.

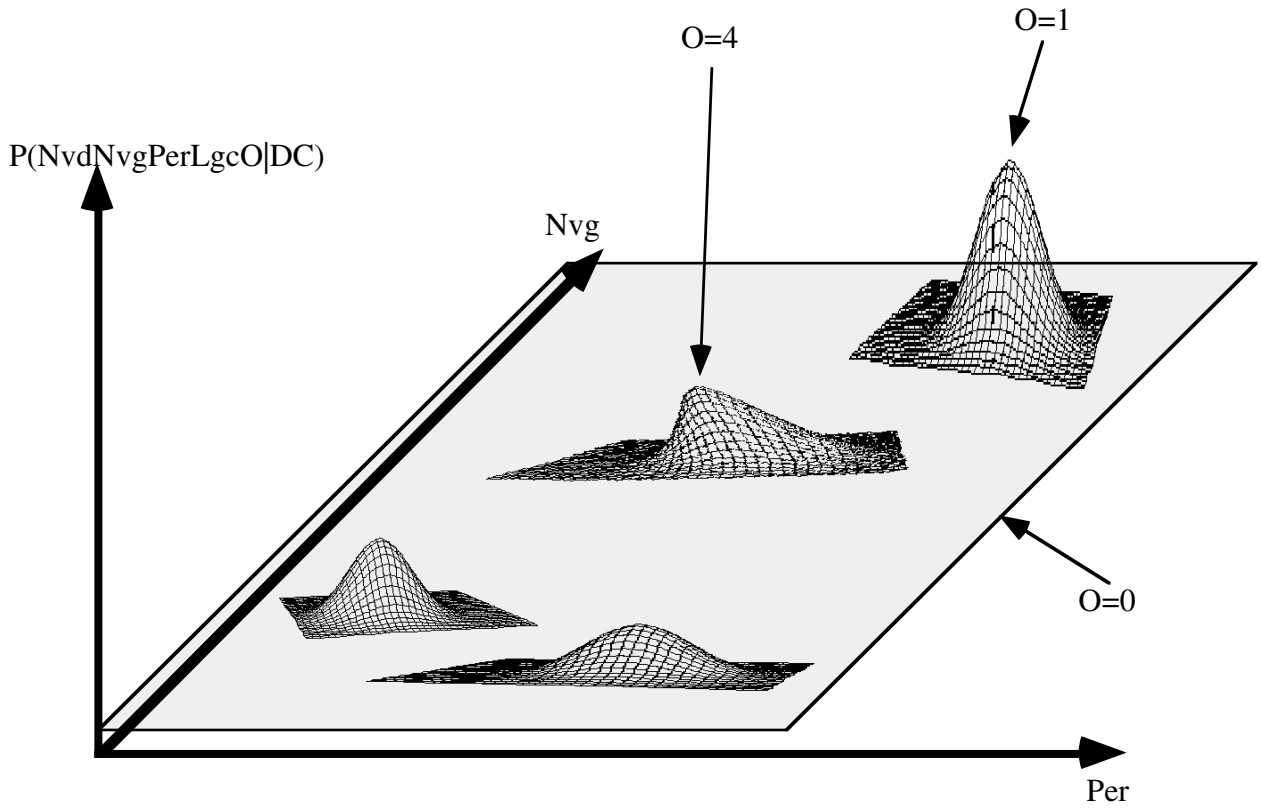


Figure 14 : $P(NvdNvgPerLgcO \mid DC)$

Seules les dimensions Per et Nvg sont représentées. Chaque pic correspond à un objet appris.
La distribution uniforme (plan grisé) sert de référence pour reconnaître la nouveauté.
Si pour un jeu de valeurs de Nvg, Nvd, Per et Lgc, c'est la distribution uniforme qui est la plus probable, Khépéra pense avoir rencontré un nouvel objet.

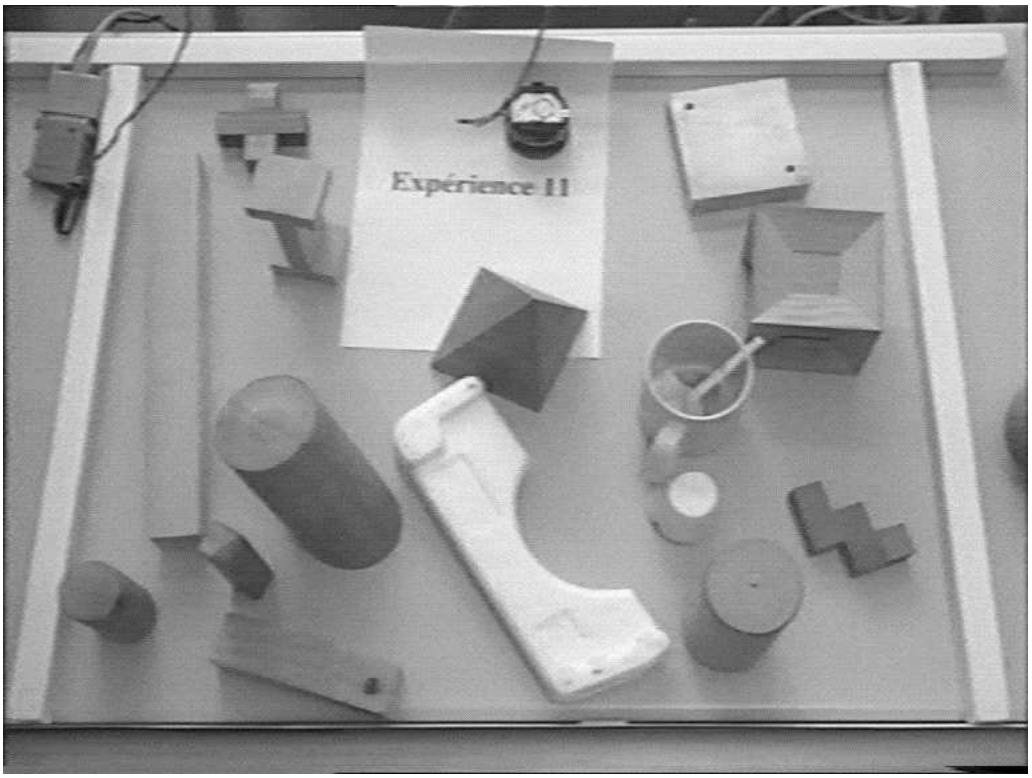


Figure 15 : Les objets présentés à Khépéra

III.3.4. Une tâche de surveillance

Enfin, et surtout, à l'aide de tous les comportements précédents, préalablement acquis, nous avons programmer et « dresser » Khépéra pour une tâche de surveillance beaucoup plus complexe. Dans cette tâche, Khépéra patrouille un environnement inconnu, donne l'alarme s'il détecte des mouvements, éteind (souffle) un incendie (une bougie) s'il en rencontre un (voir figure 16), identifie les objets qu'il cotoie et rentre seul à sa base quand on lui en donne l'ordre ou quand ses réserves d'énergie baissent.

Outre les capteurs et actionneurs précédemment décrits, on utilise pour cette tâche un caméra linéaire à 64 pixels (pour la déctection du mouvement et des bougies), les odomètres (qui permettent de savoir quand le tour d'un objet a été effectué) et une micro-turbine (pour souffler les bougies).

Cette tâche utilise 5 comportement réactifs acquis par apprentissage qui constituent les types de déplacements de bases à sa disposition :
- l'évitement d'obstacle (voir alinéa III.3.1) ;
- le suivi de contour (voir alinéa III.3.1) ;
- l'évitement d'obstacle en allant vers la lumière (voir alinéa III.3.2) ;
- le comportement de pompier ;
- l'immobilité.

Ces cinqs types de déplacement élémentaires sont combinés les uns avec les autres, par inférence probabiliste, pour obtenir 6 comportements plus complexes :
- la patrouille où Khépéra se déplace aléatoirement dans son environnement en s'arrêtant de temps en temps pour vérifier que rien ne bouge en dehors de lui ;
- le traitement de l'incendie où Khépéra se positionne par rapport à la bougie et la souffle ;
- la reconnaissance d'objet (voir alinéa III.3.3) ;
- le retour à la base sur ordre ;
- le retour à la base par manque d'énergie ;
- la recharge de son énergie, immobile, à la base.

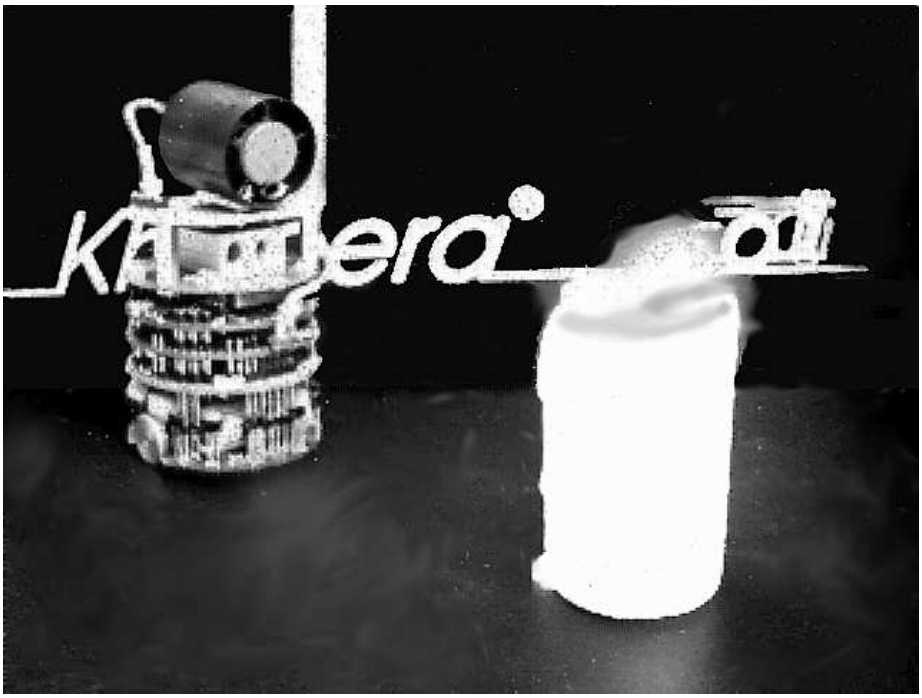


Figure 16 : Khépéra soufflant une bougie

III.4. Les réponses apportées par l'approche F+D aux difficultés de l'approche F+I

L'approche F+D tranche le nœud gordien de la catégorisation.

Elle ne manipule pas de symbole référant aux objets et aux classes d'une hypothétique réalité supposée objective. Elle manipule uniquement des « marques », dont le seul référé est proximal : un courant électrique dans un capteur ou la valeur d'une variable informatique. Aucun lien n'est à fournir entre ces marques et les objets ou classes d'objets de l'environnement. Plus besoin de justifier ces deux postulats si contestables:

- Le robot est capable, tout comme son concepteur, chacun à travers son appareil sensori-moteur, de percevoir une STRUCTURE ENSEMBLISTE (objets et ensemble d'objets) de l'environnement.
- Le robot et son concepteur sont tous deux capables d'identifier les MEMES objets et ensembles d'objets.

L'analogie pertinente pour l'approche F+I est celle de la langue. L'analogie pertinente pour l'approche F+D est celle de l'image. L'environnement n'est pas appréhendé par des mots, forcément réducteurs, sinon caricaturaux, mais par les descriptions, de véritables « images » où les probailités tiennent lieu de « niveaux de gris » (voir figure 14). L'inférence n'est pas une manipulation linguistique contraignante mais peut être vue comme l'application d'un ensemble de filtres et d'opérateurs de composition permettant de combiner directement les unes avec les autres ces différentes « images mentales ».

Reste la raison fondamentale invoquée pour contester ces postulats, à savoir l'inévitable incomplétude d'un système formel pour modéliser un environnement physique. Le système formel probabiliste de l'approche F+D n'échappe pas plus à cette incomplétude que les systèmes formels logiques des approches F+I, mais il le reconnait et sait raisonner au mieux malgré cette incomplètude. Pour cela, l'approche F+D renonce explicitement à construire des modèles complets. Elle postule, au contraire, que l'observateur, avec ses connaissances préalables et ses *a priori*, doit être impérativement pris en compte dans toute tentative d'automatisation de la perception, de la prise de décision ou de l'action. Elle ne cherche donc pas à construire de modèles d'une hypothétique réalité mais toujours à élaborer par l'expérience des descriptions d'une interaction entre un système sensori-moteur et son environnement ou, autrement dit, du point de vue très « subjectif » que ce système peut avoir de son environnement. Adopter ce point de vue subjectif de l'observateur c'est reconnaître la nature par essence incomplète de la connaissance du phénomène et donc accepter qu'il subsiste irrémédiablement des sources d'incertitude. *L'apprentissage et les techniques de maximum d'entropie permettent de transformer l'incomplétude (ce qui n'a pas été explicitement modélisé) en incertitude (les « descriptions »). L'inférence probabiliste prend le relais ensuite et permet de raisonner avec cette incertitude.* On pourra lire dans ce même numero l'article intitulé *Fondements mathématiques de l'approche F+D* pour une présentation approfondie de l'inférence probabiliste et du principe de maximum d'entropie.

Finalement, l'approche F+D répond bien aux objectifs qu'on lui avait fixés en donnant des réponses approximatives aux questions jugées pertinentes, par essence vagues, et en récusant qu'une réponse exacte puisse être apportée à autre chose qu'une question trop formelle pour être sensori-motrice.

III.5. Notes bibliographiques

L'idée de base de l'approche F+D est à relier à de nombreux travaux de sciences cognitives qui, chacun avec leur propre terminologie et avec de nombreuses nuances, nient la possibilité de modéliser l'environnement par une simple opération de type “traitement de l'information” et défendent la thèse que toute activité cognitive sensori-motrice doit être « enracinée » dans une interaction physique effective avec l'environnement.

Ces travaux vont du concept de « corporéité » de Merleau-Ponty (Merleau-Ponty42), (Merleau-Ponty45) à la notion d'« enaction » de Varela (Varela89a), (Varela89b), (Varela91), (Varela93) en passant par la « psychologie écologique » de Gibson et sa notion d'« affordance » (Gibson79), par le concept d'« action en situation » de Suchman (Suchman87) et finalement par les travaux de Braitenberg (Braitenberg83) et Brooks (Brooks91b).

De nombreuses logiques « alternatives » ont été proposées pour essayer de pallier les difficultés de l'approche F+I et pour gérer l'incertitude et l'incomplétude de l'information venant d'un phénomène physique et naturel. Pour aborder le vaste sujet que constituent ces logiques « alternatives », se référer à (Rosser58) et (Turner84) pour les logiques multivaluées, à (Bradley79), (McCarthy80), (McDermott80), (McDermott82), et (Turner84) pour les logiques modales et non monotones, à (Zadeh65), (Goguen67) et (Dubois80) pour la logique floue, à (Dempster68) et (Shafer76) pour la théorie dite de « Dempster-Shafer » et enfin, à (Zadeh78) et (Dubois85) pour la théorie des possibilités.

Des débats passionnés ont eu lieu autour des mérites relatifs des diverses théories du raisonnement plausible. On trouvera des arguments en faveur des probabilités dans (Cox61), (deFinetti72), (Cheeseman85), (Robert91) et (Pearl91).

Chercher à élaborer des structures internes rendant compte de dépendances entre variables observées est un paradigme déjà largement exploré et connu sous le nom générique d'apprentissage. Il en est ainsi, par exemple, de la théorie de l'identification et des très nombreux et très importants travaux qui y sont liés. Elle cherche, étant donné un modèle mathématique d'un phénomène, à identifier la forme et la valeur des paramètres de ce modèle pour décrire « correctement » un ensemble d'observations. On distingue l'identification paramétrique et l'identification non paramétrique. La première impose des contraintes de type analytique sur le modèle (notamment contraintes de « lissage » et de dérivabilité sur les fonctions) et cherche dans une classe de fonctions laquelle peut le mieux rendre compte des données. La seconde, plus restrictive, part d'une forme fonctionnelle donnée *a priori* et cherche uniquement à l'ajuster aux données. On distingue aussi la théorie de l'approximation et la théorie de l'estimation. La première prend en compte explicitement la gestion de l'incertitude, alors que la seconde assume qu'une description analytique sans incertitude est suffisante. Les applications de l'identification sont très nombreuses. L'identification est un sujet trop vaste pour qu'une bibliographie même superficielle en soit proposée. Contentons-nous de citer deux ouvrages introductifs de qualité, (Tarentola87) plutôt théorique, et (Richallet91) couvrant la pratique de ces techniques, ainsi que les thèses de A. Labbi (Labbi93) et R. Balaniuk (Balaniuk96) qui ont l'avantage de faire le lien entre les réseaux neuronaux et l'identification.

L'approche par les réseaux de neurones formels peut être vue comme un cas particulier de l'identification. La forme fonctionnelle donnée *a priori* correspond à l'architecture du réseau et les paramètres à identifier pour ajuster le modèle aux observations sont alors appelés « poids ». Les réseaux de neurones formels sont fréquemment utilisés en robotique pour faire de l'apprentissage. De nombreuses voies sont en cours d'exploration, certaines sont assez proches de celle préconisée dans cette étude. De plus, la tendance actuelle en théorie des réseaux de neurones formels est de montrer que ces réseaux sont des implantations plus ou moins astucieuses et efficaces d'estimateurs probabilistes. La théorie du raisonnement probabiliste et les techniques de maximum d'entropie introduites précédemment et qui sont développées dans l’article *Fondements mathématiques de l’approche F+D* apparaissent donc de plus en plus comme le support théorique des travaux

effectués dans le domaine des réseaux de neurones formels.On peut d'une part espérer que l'existence d'une théorie mathématique claire des réseaux va pouvoir faire progresser le développement de ceux-ci et, d'autre part, que les algorithmes étudiés pour les réseaux vont constituer des implantations intéressantes (notamment parallèles) des méthodes théoriques proposées par le raisonnement probabiliste.Pour les liens théoriques entre réseaux de neurones formels et probabilités on pourra lire, entre autres, (Geman84), (Smolensky86), (Cotrell88), (Robert90b), (Bessière90), (Hervé90), (Robert91) et (Geman92). Pour les liens théoriques entre réseaux de neurones formels, probabilités et physique statistique on pourra lire aussi, par exemple, (Gardner89a), (Gardner89b), (Brunel92), (Grassberger92), (Nadal92), (Nadal93) et (Gutfreund94).

Ni la théorie de l'identification (approximation + estimation) ni celle des réseaux ne répondent cependant aux exigences que nous nous sommes fixées car, si elles permettent de construire des structures de données rendant compte des dépendances entre variables observées, elles ne permettent plus, comme c'était le cas dans l'approche F+I, de combiner entre elles ces structures pour dérouler un calcul formel. Tel est l'essence de la critique adressée par Fodor et Pylyshyn au « connexionisme » quand ils lui reprochent de ne plus pouvoir satisfaire les contraintes de *productivity*, *systematicity* et *compositionality* qu'ils considèrent comme indispensable pour tout système cognitif. Concernant ces objections, l'article de référence est (Fodor88). On consultera les deux livres de Fodor (Fodor76) et (Fodor84) pour mieux comprendre cet article en le replaçant dans l'ensemble de la théorie cognitive qu'il préconise.

IV. UNE PERSPECTIVE EN GUISE DE CONCLUSION : L'EVOLUTION CONTINGENTE DES CONNAISSANCES PREALABLES

La logique et l'approche F+I sont la théorie fondatrice et la mécanisation du courant « symbolique » ou « cognitiviste » qui suppose que la « pensée » a comme prérequis le « discours ». L'apport fondamental des probabilités et de l'approche F+D est de proposer une théorie mathématique fondatrice et un moyen de mécaniser l'alternative « sub-symbolique » ou « antéprédicative » qui suggère que la « pensée » puisse préexister au « discours ». Ces deux options épistémologiques fondamentales, qui se sont opposées sous diverses formes dans toute l'histoire de la philosophie et de la science, sont ainsi remises sur un pied d'égalité, quant à la possibilité de les analyser mathématiquement et de les expérimenter informatiquement, possibilité réservée jusque-là au seul paradigme cognitiviste.

C'est dire l'ambition de l'inférence probabiliste et le nombre de perspectives ouvertes pour poursuivre le travail décrit ici, qui peut être vu comme l'amorce d'un très vaste programme de recherche. Plutôt que d'énumérer trop rapidement une trop longue liste de travaux en cours et de perspectives possibles, nous préférons en détailler une qui nous paraît à la fois fondamentale et abordable : l'évolution contingente des connaissances préalables.

L'approche F+D est un moyen de surmonter nombre des difficultés que l'approche F+I rencontre dans le cadre sensori-moteur. Cependant, il incombe au concepteur du système sensori-moteur une tâche absolument essentielle, le choix des connaissances préalables, qui conditionne entièrement le succès de l'entreprise. Sur le chemin vers « l'autonomie[1] » des systèmes sensori-moteurs, l 'étape suivante est donc de se demander s'il est possible d'automatiser au moins partiellement cette tâche de sélection des connaissances préalables. Nous pensons que la formalisation probabiliste rigoureuse et unifiée de l'inférence et de l'apprentissage que propose l'approche F+D permet de poser clairement cette question et d'envisager les voies à suivre pour la résoudre. L'objectif de cette conclusion est de les présenter.

---

[1]Du grec *autos*, soi-même et *nomos*, loi : qui se donne ses propres lois.

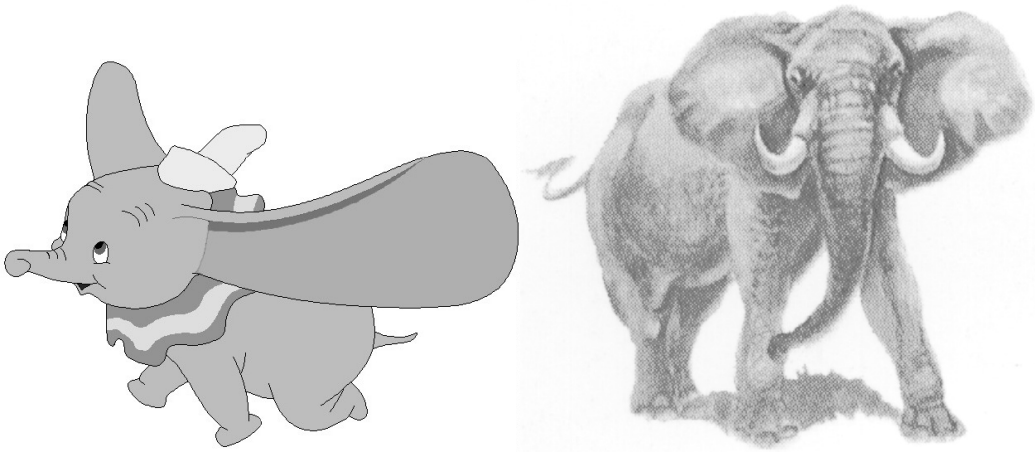


Figure 17 : les ailes n'ont pas été « conçues » pour que les animaux puissent voler : Disney n'a rien inventé

IV.1. Le rôle du concepteur dans l'approche F+D

Dans l'approche F+D des systèmes sensori-moteurs, le rôle du concepteur est clairement identifié, il doit :

1 - concevoir la structure mécanique et électronique du robot ;

2 - spécifier les connaissances préalables (structurelles, de dépendance, *a priori* et d'observation) ;

3 - servir de « moniteur » au robot en le plaçant dans des situations qu'il juge pertinentes et éventuellement « superviser » son apprentissage, soit en lui donnant des « exemples » de comportements (par exemple, pour Khépéra en le pilotant avec le *joystick*), soit en lui fournissant une mesure de « qualité » sur ses comportements (apprentissage par renforcement).

Concernant le premier point, pour des raisons pratiques, il nous semble difficile pour l'instant de modifier automatiquement la structure physique et électronique du robot.

Le troisième point est très intéressant mais n'est pas le sujet de cette conclusion où nous nous concentrerons sur le deuxième problème moins classique : comment rendre (partiellement) automatique la recherche des connaissances préalables.

IV.2. La notion de contingence

Les ailes n'ont pas été « conçues » pour que les animaux puissent voler. Elles furent sans doute, à l'origine, des organes de régulation de la température. Il s'avéra alors que la structure de ces organes donnait des capacités pour planer. Cet « effet de bord » apportant un avantage sélectif, il fut préservé et continua à évoluer pour devenir la fonction principale et conduire aux ailes que nous observons sur les insectes actuels.

De nombreuses structures, de la biologie à la technologie (par exemple, la disposition des touches sur les claviers d'ordinateurs), ne peuvent s'expliquer de manière satisfaisante que si on les replace dans une telle perspective historique et évolutionniste. Elles sont incompréhensibles si l'on suppose qu'elles ont été « conçues » à partir d'une *tabula rasa* en vue de remplir leurs fonctions actuelles.

Cette notion appelée *contingence* est considérée par certains (voir en particulier l'œuvre de Stephen Gould (Gould85), (Gould89) et (Gould91)) comme un aspect essentiel de l'évolution des

espèces animales. Elle n'a jamais, à notre connaissance, été appliquée à la construction des robots. Cela se comprend aisément, dans la mesure où la plupart des robots existant ont été « conçus ». Dans la perspective d'autonomie où nous nous plaçons dans ce paragraphe, il ne s'agit plus de concevoir les robots mais de leur donner les moyens de se « développer ». C'est alors que nous pensons que la contingence devient un facteur capital.

Comme nous l'avons vu tout au long de ce papier, l'incomplétude des modèles est la difficulté majeure que rencontre la robotique. Du « bruit » aux « aberrations » en passant par les « hasards » et les « accidents », tout ce qui est *imprévu*, c'est-à-dire tout ce qui sort des limites étriquées de la modélisation, est considéré comme parasite et nuisible.

L'approche F+D, comme nous l'avons vu, apporte un premier type de réponse à l'incomplétude : les « variables cachées », ce qui n'a pas été modélisé, engendrent de l'incertitude qui est prise en compte dans les descriptions et par le raisonnement probabiliste.

La notion de développement contingent va plus loin, car elle suppose que les effets imprévus peuvent dans certains cas être utilisés comme moteurs du développement. Une anecdote concernant le Khépéra mérite d'être racontée pour illustrer ces propos. Khépéra était utilisé avec un programme photophile réflexe tout à fait élémentaire lui permettant de suivre une lampe torche. Un jour, alors que le porteur de la torche l'éteignit, le robot fit quelque chose de complètement inattendu qui marqua beaucoup tous les spectateurs. Le robot fit demi-tour sur place, se dirigea en *évitant tous les obstacles* vers une pièce éclairée, *franchit la porte* à moitié fermée de ce bureau en *négociant parfaitement ce passage étroit* pour se mettre à *danser sous la lampe* dont la lumière provenait, apparemment content d'avoir atteint son but ! Ce comportement étonnant est en fait facile à comprendre. La source de lumière principale étant le bureau éclairé, il est normal qu'elle ait attiré Khépéra. La combinaison du comportement photophile et des ombres projetées par chaque obstacle explique que le robot en allant vers la lumière évite les obstacles et négocie facilement les étroitures, le gradient de lumière lui indiquant toujours le chemin à suivre. Enfin, la « danse de victoire » traduit seulement l'indécision du robot une fois arrivé dans le « marais » lumineux se trouvant sous la lampe. Cette expérience, déjà ancienne, fut à l'origine d'une bonne partie des travaux décrits ici, la question étant de savoir comment exploiter l'apparition de tels comportements *émergents.*

IV.3. La remise en cause des connaissances préalables structurelles

Reprenant le schéma de l'anecdote qui vient d'être racontée, nous avons tout d'abord appris à Khépéra le comportement photophile (en le pilotant avec le *joystick*). Nous obtenons ainsi une description P(VrotLum|D1C1) encodant les dépendances entre *Lum* et *Vrot* (voir figure 18.a).

Dans une deuxième phase, nous utilisons cette description photophile pour commander Khépéra (P(Vrot|LumD1C1)). Comme prévu, on constate que, placé dans un environnement avec une seule source de lumière assez rasante pour que les obstacles projettent des ombres importantes, Khépéra se dirige vers la lumière en évitant les obstacles. Alors que Khépéra exécute ce comportement, nous apprenons la description P(VrotDirProx|D2C2) encodant les dépendances entre *Vrot*, *Dir* et *Prox* (voir figure 18.b). Cette nouvelle description est utilisée pour commander Khépéra (P(Vrot|DirProxD2C2)). On vérifie qu'elle permet bien d'éviter les obstacles, et ce, même en l'absence d'ombre. On vérifie aussi que les valeurs de Vrot inférées à partir de la description photophile P(VrotLum|D1C1) et à partir de la description d'évitement d'obstacle P(VrotDirProx|D2C2) sont les mêmes aux incertitudes près.

Enfin, dans une troisième phase (voir figure 18.c), on change la forme et la taille des ombres de certains obstacles en changeant leur hauteur. Par exemple, on prend des obstacles plus hauts qui projettent une ombre plus grande. Khépéra est commandé en mode photophile (P(Vrot|LumD1C1)). Les obstacles plus hauts ayant des ombres plus grandes, il les évite de plus loin que dans les deux

premières expériences. Les valeurs de Vrot inférées à partir de la description photophile P(VrotLum|D1C1) et à partir de la description d'évitement d'obstacle P(VrotDirProx|D2C2) ne sont donc plus du tout les mêmes. L'écart entre ces valeurs est représentatif de la hauteur des obstacles et peut éventuellement être utilisé pour construire une description encodant les dépendances entre ces deux valeurs de *Vrot* et la hauteur des obstacles.

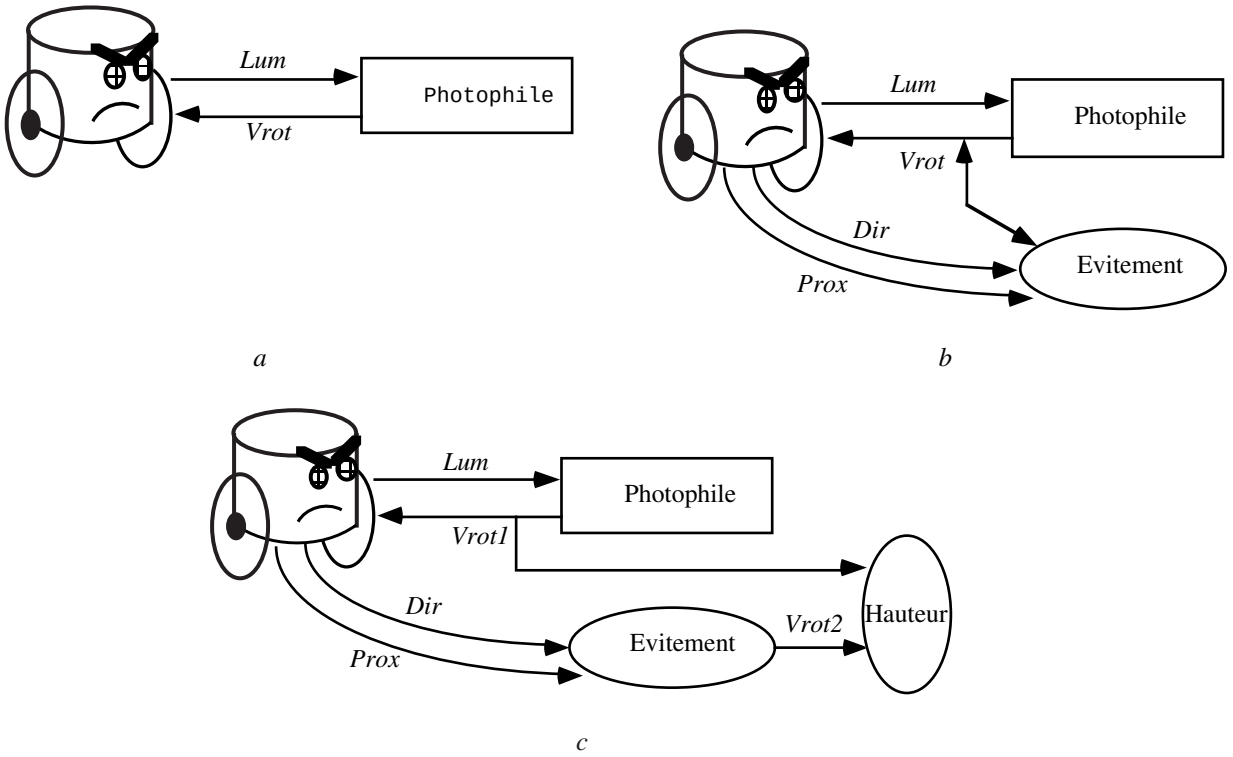


Figure 18 : mesurer la hauteur des obstacles avec des proximètres et des capteurs de lumières.

Les proximètres et les capteurs de lumières n'ont pas été « conçus » pour que Khépéra puisse mesurer la hauteur des obstacles. Ils ont été conçus pour des tâches tout autres. Pourtant cette expérience montre que, dans certaines conditions, ils permettent d'arriver à ce résultat.

L'ensemble des variables données au départ par le concepteur (connaissances préalables structurelles) peut s'avérer très rapidement insuffisant. Nous voyons dans l'expérience qui vient d'être décrite l'indication de la voie à suivre pour créer de nouvelles variables. Elle repose sur l'idée qu'une description une fois acquise peut être utile même dans des contextes différents de celui de l'apprentissage initial, en particulier, pour créer de nouvelles variables identifiant des « différences » entre la description de référence et les données en cours de traitement. Plus généralement, il nous semble ainsi possible d'envisager la création systématique et automatique de nouvelles variables potentiellement intéressantes.

IV.4. Vers une approche évolutionniste des connaissances préalables?

Un processus évolutionniste suppose 4 composantes fondamentales:
- Un code
- Un processus d'ontogenèse qui à partir du code soit capable par l'interaction avec l'environnement de produire des objets.
- Un processus d'échange d'information mutuelle qui à partir de 2 codes (ou plus) soit en mesure d'engendrer de nouveaux codes, assez ressemblant à leurs « ancêtres » pour en garder les acquis, mais assez différents d'eux pour explorer de nouvelles solutions.

- Un processus de sélection qui « évalue » ces objets et sélectionne les meilleurs en favorisant les codes correspondants dans le processus de génération de nouveaux codes.

Nous sommes avec l'approche F+D dans une situation qui rend apparemment possible la mise en place d'un processus évolutionniste pour le développement des facultés de perception, de décision et d'action d'un système sensori-moteur de type robotique.

Les connaissances préalables telles qu'elles ont été définies sont des objets mathématiques bien structurés. En tant que tels, il est certainement possible de les coder.

Par exemple, les connaissances préalables structurelles correspondent au choix d'un ensemble de variables parmi toutes les variables potentiellement utilisables par le système sensori-moteur. Il existe de multiples manières connues de coder simplement le choix de n objets parmi m.

Les connaissances préalables de dépendances peuvent se traduire par le choix parmi les n variables précédemment sélectionnées d'un certain nombre de sous ensembles, eux-mêmes divisés en deux, entre les variables conditionnées et les variables conditionnantes. De multiples codes semblent, aussi ici, possibles.

Pour les connaissances *a priori* et les connaissances préalables d'observation, les formes de code possibles sont moins évidentes à imaginer, mais étant donné qu'il s'agit de connaissances formelles de type mathématique, trouver de tels codes est certainement possible.

L' inférence probabiliste et le principe de maximum offrent le processus de développement du code des connaissances préalables par interaction avec les données expérimentales pour produire les objets appelés descriptions.

Dans la mesure où un code a été trouvé pour représenter un ensemble de connaissances préalables, il est facile d'imaginer divers opérateurs d'échanges d'information entre codes afin de produire de nouveaux codes correspondant à de nouveaux ensembles de connaissances préalables. Les travaux actuels sur les algorithmes génétiques proposent de nombreuses solutions qui devront être étudiées.

La partie la plus délicate à imaginer est sans aucun doute le processus de sélection des meilleurs descriptions. Les critères de cette sélection ne sont pas encore forcément clairs et devront faire l'étude d'une étude approfondie. Cependant, plusieurs possibilités paraissent d'ores et déjà intéressantes à approfondir:

- Les descriptions « plates », les distributions de probabilités proches de la distribution uniforme, celles traduisant une très grande incertitude sont de manière évidente moins intéressantes que celles très « pointues » qui présentent des pics de probabilités. Les premières sont moins informationnelles que les secondes et leur exploitation plus délicate et plus difficile. L'entropie semble, une fois de plus, une intéressante métrique pour quantifier cette notion, les distributions « plates » ayant une entropie beaucoup plus grande que les « pointues ». Les distributions très « plates » révèlent en général que les connaissances préalables ont été mal choisies (présence de variable cachée, mauvais choix de la structure de dépendance, etc...).
- Pour un jeu de données expérimentales, la proximité entre la description et les données, mesurée par ce qui est souvent appelé biais, semble souhaitable. Une méthode précise et rigoureuse pour quantifier ce deuxième critère est proposé dans l'article.
- Inversement, avoir des descriptions complètement « irrégulières », sans aucune « continuité », ayant une très grande variance semble non souhaitable. Ces distributions sont en général le reflet exact des données d'apprentissage mais offrent de très mauvaises capacités de « généralisation ».
- Enfin, pour pouvoir être exploitable en pratique les descriptions doivent être parcimonieuses dans l'emploi de la mémoire (représentation paramétrique économique) et permettre des

calculs probabilistes faciles et surtout rapides. On retrouve ici la vieille notion de « rasoir d'Occam » qui propose de favoriser les représentations simples aux détriments des plus complexes.

Tous ces critères sont évidemment en contradiction les uns avec les autres, notamment le dilemme biais *versus* variance est bien connu et largement étudié (Geman92). Une large voie de recherche s'ouvre donc devant nous afin d'imaginer, étudier et implanter des processus de sélection intéressants. Un premier pas dans cette voie a été fait avec le DEA de Ludovique Mauconduit (Mauconduit97).

V. BIBLIOGRAPHIE

Apostel Léo; *Syntaxe, sémantique et pragmatique* dans *Logique et connaissance scientifique* ; Encyclopédie de la Pléiade ; Editions Gallimard ; Paris, France ; 1967a

Apostel Léo; *Logique et dialectique* dans *Logique et connaissance scientifique* ; Encyclopédie de la Pléiade ; Editions Gallimard ; Paris, France ; 1967b

Balaniuk Remis; *Identification structurelle* ; Thèse de doctorat de l'INPG (Institut Natiaonal Polytechnique de Grenoble), Grenoble, France ; septembre 1996

Bessière Pierre; *Toward a synthetic cognitive paradigm: Probabilistic Inference ;* COGNITIVA 90, Madrid, 1990 ; edité par T. Kohonen & F. Fogelman-Soulie ; Elsevier science publishers, 1991

Bickhard Mark H. & Terveen Loren; *Foundational issues in artificial intelligence and Cognitive Science ; Impasse and solution* ; North-Holland, Elsevier ; 1995

Bradley B. & Swartz M. ; *Possible worlds ;* Oxford University Press, Great Britain, 1979

Braitenberg V.; *Vehicles: Experiments in synthetic psychology* ; MIT Press, Cambridge, USA, 1986

de Broglie Louis ; *Les représentations concrètes en micro-physique* ; dans *Logique et connaissance scientifique* ; Encyclopédie de la Pléiade ; Editions Gallimard ; Paris, France ; 1967

Brooks, R.A. ; *Achieving artificial intelligence through building robots* ; AI MEMO 899, MIT, 1986a

Brooks, R.A. ; *A robust layered control system for a mobile robot*, IEEE journal of robotics and automation, RA-2, Avril 1986b

Brooks, R.A. ; *A robot that walks: emergent behaviors from a carefully evolved network* ; Neural Computations, 1:2, 1989

Brooks, R.A. ; *Elephants don't play chess*, in P. Maes (ed), *Designing autonomous agents: theory and practice from biology to engineeing and back,* MIT Press ; 1990

Brooks Rodney A.; *Artificial life and real robots* ; in ECAL 91 ;*Toward a practice of autonomous systems:* proc. of the 1st european conference on artificial life (Paris 1991), F. Varela & P. Bourgine (eds), MIT Press/Bradford books, 1991a

Brooks Rodney A.; *Intelligence without representation* ; Artificial Intelligence, n°47, 1991b

Brunel N., Nadal J.-P., and Toulouse G. ; *Information capacity of a perceptron* ; J. Phys. A: Math. and Gen., 25:5017--5037, 1992.

Campbell Jerremy ; *Gramatical Man* ; Simon & Schuster, 1982

Cheeseman P. ; *In defense of Probability* ; Proceedings of AAAI85, 1985

Cotrell Marie ; *Modèlisation des réseaux de neurones par des chaînes de Markov et autres application* ; Thèse d'état, Université Paris Sud (Orsay) ; 1988

Cox R. T. ; *The algebra of probable inference* ; The John Hopkins Press, Baltimore, 1961

Dedieu Eric ; *La représentation contingente : Vers une reconciliation des approches fonctionnelles et structurelles de la robotique autonome*. Thèse de troisième cycle INPG ; septembre 1995, Grenoble, France, 1995

de Finetti B. ; *Probability, induction and statistics* ; John Willey & sons, 1972

Dempster A.P. ; *A generalization of Bayesian inference* ; J. of the Royal Statistical Society, B 30 ; 1968

Dreyfus H. ; *What computers can't do : The limits of artificial intelligence* ; Harper and Row, New York, USA, 1979

Dubois D. & Prade H. ; *Fuzzy sets and systems: theory and applications* ; Academic Press, New York, 1980

Dubois D. et Prade H. ; *Théorie des possibilités* ; Masson ; Paris, France ; 1985

Ekeland Ivar ; *Le Calcul, l'Imprévu ; Les figures du temps de Kepler à Thom* ; Editions du Seuil, Paris, France ; 1984

Ekeland Ivar ; *Au hasard ; La chance, la science et le monde* ; Editions du Seuil, Paris, France ; 1991

Fodor Jeremy A. ; *The language of thought* ; Harvester Press ; Sussex ; 1976

Fodor Jerry A. ; *The Modularity of Mind - An essay on faculty psychology* ; A Bradford Book ; The MIT Press, Cambridge, Massachusetts, USA ; 1984

Fodor Jerry A. & Pylyshin Zenon W. ; *Connectionism and cognitive architecture : A critical analysis* in *Connections and Symbols* edited by S. Pinker & J. Mehler ; A Bradford Book, The MIT Press, Cambridge, Massachusetts, USA ; 1988

Foucault Michel ; *Les mots et les choses* ; Gallimard, Paris France ; 1966

Gardner E. ; *Optimal basins of attraction in randmly sparse neural network models ;* Journal of physics A, volume 22, number 12, 1989a

Gardner E. ; *Three unfinished works on the optimal storage capacity of networks ;* Journal of physics A, volume 22, number 12, 1989b

Geman S. & Geman D. ; *Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images ;* IEEE transactions on pattern analysis and machine intelligence, 6, 1984

Geman Stuart, Bienenstock Elie & Doursat René ; *NeuralNetworks and the Bias/Variance Dilemma* ; Neural Computation 4, 1-58 ; MIT Press ; 1992

Gibson J.J. ; *The ecological approach to visual perception* ; Houghton-Misslin company, Boston, USA, 1979

Goguen J.A. ; *L-fuzzy sets* ; Journal of mathematical analysis and applications, N°18, 1967

Gould Stephen J. ; *The flamingo's smile* ; edited by W.W. Norton, 1985 ; En français : *Le sourire du flamant rose* ; Editions du seuil, Paris ; 1988

Gould Stephen J. ; *Wonderful life* ; edited by W.W. Norton ; 1989

Gould Stephen J. ; *Bully for Brontosorus* ; Penguin editors ; 1991

Grassberger Peter & Nadal Jean-Pierre (editors) ; *From statistical physics to statistical inference and back* ; Kluwer Academic Publishers ; Dordrecht, The netherlands ; 1992

Grize Jean-Blaize ; *Historique. Logique des classes et des propositions. Logique des prédicats. Logiques modales* ; dans *Logique et connaissance scientifique* sous la direction de Jean Piaget ; Encyclopédie de la Pléiade ; Editions Gallimard ; Paris, France ; 1967

Gutfreund H. & Toulouse G. (editors) ; *Biology and computation: a physicist's choice* ; Advances series in neuroscience, Vol. 3, World Scientific ; 1994

Harnad S. ; *Mind, Machines and Searle* ; Journal of theoretical and experimental artificial intelligence ; 1: 5-25 ; 1989

Harnad S. ; *The symbol grounding problem* ; Physica D, 42 : 335-346 ; 1990

Hervé T., Dolmazon J.M. & Demongeot J. ; *Random field and neural information ;* Proc. Natl. Acad. of Sciences, U.S.A., Vol 87, 1990

Hofstadter Douglas ; *Gödel, Escher, Bach : an Eternal Golden Braid* ; Basic Books Inc ; New York, U.S.A. ; 1979 ou *Gödel, Escher, Bach : les Brins d'une Guirlande Eternelle* ; InterEditions ; Paris, France ; 1985

Horswill Ian D. ; *Specialization of perceptual processes* ; MIT 1993

Jaynes E.T. ; *Probability theory - The logic of science* ; à paraître. Version provisoire disponible à http://bayes.wustl.edu. ; 1995

Kleene Stephen C. ; *Mathematical Logic* ; John Wiley & sons ; New York, USA ; 1967 ou *Logique mathématique* ; Armand Colin ; Paris, France ; 1971

Koza John ; *Evolving emergent wall following robotic behavior using the genetic programming paradigm* ; ECAL 91 ; Toward a practice of autonomous systems: proc. of the 1st european conference on artificial life (Paris 1991), F. Varela & P. Bourgine (eds), MIT Press/Bradford books, 1992

Labbi Abderrahim ; *Sur l'approximation et les systèmes dynamiques dans les réseaux neuronaux* ; Thèsede lInstitut National Polytechnique de Grenoble, Grenoble, France, 1993

Ladrière Jean ; *Les limites de la formalisation* ; dans *Logique et connaissance scientifique* sous la direction de Jean Piaget ; Encyclopédie de la Pléiade ; Editions Gallimard ; Paris, France ; 1967

Lichnerowicz André ; *Remarques sur les mathématiques et la réalité* dans *Logique et connaissance scientifique* ; Encyclopédie de la Pléiade ; Editions Gallimard ; Paris, France ; 1967

Lucas J.R. ; *Minds, Machines, and Godël* ; Philosophy 36 ; 1961

MacKay David J.C. ; *A practical bayesian framework for backprop networks* ; Neural computation ; 1992

Malcom C. & Smithers T.; *Symbol grounding via a hybrid architecture in an autonomous assembly system* ; DAI-RP 420, University of Edinburgh ; 1989

Mataric Maja J. ; *Navigating with a rat brain : a neurobiologically inspired model for robot spatial representation* ; SAB 90 From animals to animats: proc. of the 1st international conference on the simulation of adaptive behaviour (Paris 1990), J-A. Meyer & S. Wilson (eds), MIT Press/Bradford books, 1991

Mauconduit Ludovique; *Evolution de connaissances a priori : un pas vers l'autonomie* ; DEA INPG ; Grenoble ; 1997

McCarthy J. ; *Circumpscription: a form of non-monotonic reasoning* ; Artificial Intelligence n°13, 1980

McDermott D. & Doyle J. ; *Non-monotonic logic I* ; Artificial Intelligence, n°13, 1980

McDermott D. ; *Non-monotonic logic II* ; JACM n°29, 1982

Merleau-Ponty, M. ; *La structure du comportement. ;* 8ème ed., PUF, 1977

Merleau-Ponty M. ; *Phénoménologie de la perception* ; Editions Gallimard, Paris, France, 1945
Nadal Jean-Pierre ; *Réseaux de neurones : de la physique à la psychologie* ; Armand Colin, Paris, France ; 1992
Nadal J.-P. and Parga N. ; *Duality between learning machines: a bridge between supervised and unsupervised learning* ; LPSENS preprint, to appear in Neural Computation, 1993.
Nilsson Nils J. ; *Principles of Artificial Intelligence* ; Morgan Kaufmann Publishers ; U.S.A. ; 1980 ou *Principes d'intelligence artificielle* ; CEPADUES-EDITIONS ; Toulouse, France ; 1988
Omnès Roland ; *Philosphie de la science contemporaine* ; Gallimard, Paris, France ; 1994
Pearl Judea ; *Probabilistic reasoning in intelligent systems : Networks of plausible inference* ; Morgan Kaufmann Publishers, San Mateo, California, USA ; 1991
Piaget Jean (sous la direction de) ; *Logique et connaissance scientifique* ; Encyclopédie de la Pléiade ; Editions Gallimard ; Paris, France ; 1967a
Piaget Jea ; *Les problèmes principaux de l'épistémologie des mathématiques* dans *Logique et connaissance scientifique* ; Encyclopédie de la Pléiade ; Editions Gallimard ; Paris, France ; 1967b
Piaget Jean ; *Les données génétiques de l'épistémologie physique* ; dans *Logique et connaissance scientifique* ; Encyclopédie de la Pléiade ; Editions Gallimard ; Paris, France ; 1967c
Piaget Jean ; *Les relations entre le sujet et l'objet dans la connaissance physique* ; dans *Logique et connaissance scientifique* ; Encyclopédie de la Pléiade ; Editions Gallimard ; Paris, France ; 1967d
Poincaré Henri ; *La science et l'hypothèse* ; Publié originalement en 1902 ce livre est maintenant réédité aux éditions de la Bohème, Rueil-Malmaison, France ; 1992
Proust Joëlle ; *Questions de forme : Logique et proposition analytique de Kant à Carnap* ; Fayard, 1986
Reeke G. N. & Edelman G. M. ; *Real brains and artificial intelligence* in *The AI debate : false starts, real fundations edited by* S.R. Graudebard ; The MIT Press, Cambridge, USA ; 1988
Richalet, J. 1991 *Pratique de l'identification.* ADERSA, coll. Traité des nouvelles technologies, Hermès.
Robert Claudine ; *Observable networks, Applications to medical decision making* in Proceedings of IEEE PAMI, 1990a
Robert Claudine ; *An entropy concentration theorem: applications in artificial intelligence and descriptive statistics* ; Journal of Applied Probabilities ; september 1990b
Robert Claudine ; *Modèles statistiques pour l'intelligence artificielle : l'exemple du diagnostic médical* ; Masson, Paris, France ; 1991
Rosser J.B. & Turquette A.R. ; *Many-valued logics* ; North-Holland, Amsterdam, Nederlands, 1958
Searle J. R. ; *Mids, Brains, and Programs* in *Mind design* edited by J. Haugeland ; MIT Press ; 1981
Shafer G. ; *A mathematical theory of evidence* ; Princeton University Press, Princeton, USA, 1976
Smolensky P. ; *Information processing in dynamical systems: foundations of Harmony theory* in *Parallel Distributed Processing* by D.E. Rumelhart, J.L. McClelland & the P.D.P. Research Group ; The M.I.T. Press, 1986
Stewart J. ; *Epistemologie des sciences de la cognition et du langage* ; Annales Carnac, 5, 115-134 ; Lausanne, Suisse, 1992
Suchman Lucy A. ; *Plans and situated actions ; The problem of human/machine communication* ; Cambridge University Press, Great Britain, 1987
Tarentola Albert ; *Inverse Problem Theory ; Methods for data fitting and model parameters estimation* ; Elsevier ; New York, USA ; 1987
Turner Raymond ; *Logics for artificial intelligence* ; Ellis Horwood limited, New York, USA, 1984
Ulmo Jean ; *Les concepts physiques* ; dans *Logique et connaissance scientifique* ; Encyclopédie de la Pléiade ; Editions Gallimard ; Paris, France ; 1967
Varela Francisco ; *Connaître les sciences cognitives ; tendances et perspectives* ; Collection « Science ouverte », Editions du Seuil, Paris, France, 1989a
Varela Francisco ; *Autonomie et Connaissance ; Essai sur le vivant* ; Editions du Seuil, Paris, France, 1989b
Varela Francisco ; ECAL 91 *Toward a practice of autonomous systems: proc. of the 1st european conference on artificial life (Paris 1991),* F. Varela & P. Bourgine (eds), MIT Press/Bradford books, 1991
Varela Francisco, Thompson Evan & Rosch Eleanor ; *L'inscription corporelle de l'esprit ; science cognitives et expérience humaine* ; Editions du Seuil, Paris, France, 1993
Vauquois Bernard ; *Calculabilité des langages* ; Université de Grenoble ; Grenoble, France ; 1976
Winograd T. & Flores F. ; *Understanding computers and cognition : A new foundation for design* ; Ablex Press, USA, 1986
Zadeh L. A. ; *Fuzzy sets* ; Information and Control, N°8, 1965
Zadeh L. A. ; *Fuzzy sets as a basis for a theory of possibility* ; Fuzzy Sets and Systems, N°1, 1978