

Chapitre 1

Modéliser le sens commun : Objectifs et méthodes

L'objectif de ce travail est de développer une théorie du mouvement de "sens commun" qui soit exploitable automatiquement dans des tâches qui nécessitent de manipuler des données spatio-temporelles. Nous allons donc présenter dans ce chapitre ce que nous entendons par sens commun, quelle est l'importance de la problématique spatiale qui se pose dans cette perspective et nous allons définir le cadre méthodologique que nous suivons au cours de cette étude.

1.1 La nature du sens commun

1.1.1 Intelligence et sens commun

L'intelligence artificielle (par la suite I.A.) est une notion dont la signification même est controversée, et dont la pertinence est fortement discutée ; en tant que science son objet n'est pas clairement défini dans la mesure où le terme "intelligence" recouvre des réalités diverses et difficilement unifiables. Il est d'ailleurs remarquable que toutes les tâches que l'on a un temps considérées comme relevant de l'intelligence humaine et qui ont pu être simulées avec succès par un ordinateur (jouer aux échecs par exemple) ont pratiquement disparu du champ d'étude de la communauté scientifique se reconnaissant sous l'étiquette "intelligence artificielle". La modélisation de "l'intelligence" a laissé place à la modélisation de capacités plus localisées (traitement de l'information focalisée sur certains aspects : gestion de l'imprécis, de l'incertain ; formes de raisonnement particulières : non monotone, par défaut ; traitement du langage naturel ; vision de haut niveau, etc.). Ces études ont lentement mis en lumière l'importance dans la cognition humaine de capacités, non pas intelligentes dans le sens où on parle de personnes "intelligentes" (plus que d'autres) mais relevant plutôt d'un ensemble de connaissances *partagées* par tous qui sont mises en jeu dans nos moindres activités, intellectuelles ou non. La moindre phrase échangée entre deux humains fait appel à de nombreuses notions qui permettent la compréhension des situations décrites. Prenons par exemple, dans *Alice au pays des merveilles* :

A ce moment son pied glissa et, plouf! elle se trouva plongée jusqu'au cou dans l'eau salée.
Sa première idée fut qu'elle était tombée dans la mer (...).

La compréhension de ce petit texte implique d'avoir des notions de l'espace et du temps, de la physique (des solides et des liquides), de la structure et de la dynamique du corps humain, et même des

connaissances sur les connaissances d’Alice (sur ce qu’est une croyance, et pourquoi elle peut penser qu’elle est dans la mer). Ce petit exemple indique simplement l’ensemble de concepts qui ne semblent avoir *a priori* rien de remarquable, mais qui sont partout. Ce que l’on peut donc désigner comme “sens commun” a alors servi à caractériser la modélisation de ces connaissances naturelles jusque là ignorées par l’I.A. mais indispensables pour pouvoir prétendre rendre compte des capacités humaines dans des domaines variés où l’automatisation est un objectif. On peut faire remonter les premiers travaux soulignant ces aspects à (McCarthy, 1959), qui commençait à soulever l’importance de la modélisation de ces notions. Mais un des principaux problèmes théoriques qui s’est toujours posé à l’I.A. en tant que discipline est l’absence de procédure de validation des travaux comparable à ce qui peut exister dans d’autres disciplines scientifiques. L’absence d’un but commun significatif rend difficile la définition de procédures d’évaluation des résultats admises et respectées¹. Un article phare pour la redéfinition des objectifs de la communauté I.A. dans une perspective mieux cadrée que ce qui existait alors a été la proposition explicite de Hayes dans son manifeste de 1978² de tenter de modéliser cet ensemble de connaissances informelles qui nous permet de communiquer, d’agir et d’interagir avec notre environnement d’une façon satisfaisante. S’opposant aux problèmes “jouets” jusque là considérés par beaucoup de travaux (le monde des cubes en planification par exemple), Hayes définissait un objectif précis beaucoup plus ambitieux : modéliser formellement l’ensemble des connaissances “intuitives” que nous possédons sur le monde physique (la physique “naïve”), sur les objets, sur notre environnement, en assumant pleinement leur complexité, avec quelques indications des méthodes qu’il paraissait raisonnable d’employer dans ce but (un formalisme équivalent à la logique du premier ordre par exemple) mais sans préconiser un cadre rigide.

Seule l’évaluation empirique des résultats formels n’était pas en effet une de ses préoccupations majeures³. Hayes se contentait de considérer le pouvoir expressif d’une théorie comme garant de son intérêt, l’évaluant grossièrement comme le rapport entre la quantité d’information qu’elle représente (“mesurée” par le nombre d’axiomes, et les définitions qu’elle permet) et le nombre de concepts primitifs qu’elle introduit. Bien que nous plaçons nos travaux dans ce contexte général, nous allons essayer de préciser ses motivations, quelles méthodes sont suivies dans sa réalisation, et quelles procédures permettent d’évaluer la pertinence et le succès de cette entreprise, d’une façon peut-être moins discutable.

1.1.2 La représentation des connaissances

Dans les travaux qui se réclament de cette prise de conscience et qui tentent de modéliser les aspects du sens commun, beaucoup se regroupent sous l’appellation “représentation des connaissances” (knowledge representation), ce qui correspond en fait à plusieurs courants : on a d’une part un grand nombre de travaux qui insistent sur les capacités de raisonnement propres aux humains pour expliquer le raisonnement de sens commun; cette approche a donné notamment une grande impulsion au raisonnement non monotone et à beaucoup de logiques non classiques censées être plus proches du raisonnement naturel que la logique mathématique. De nombreux travaux mettent l’accent sur des formalismes de re-

1. Il y a alors très souvent une grande confusion entre la modélisation de certaines capacités attribuées de façon indiscutables à l’intelligence humaine et les réalisations pratiques d’applications ayant recours aux notions recouvertes par ces capacités ; tout ceci rend beaucoup de travaux non cumulatifs et peu exploitables pour des gens ayant des objectifs pratiques différents.

2. Et republié dans (Hayes, 1985b).

3. Ce que résumait une boutade finale de la version remaniée de 1985 (Hayes, 1985b) : “Is this science? Who knows, and who cares?”.

présentation plutôt que sur la représentation effective de connaissances. Une autre tendance se consacre plutôt à la modélisation de concepts importants, comme la prise en compte du temps et de l'espace, la représentation de l'action... Les problèmes sont alors de modéliser de façon adéquate des domaines sur lesquels on possède quelques intuitions, quitte à utiliser des méthodes dédiées au domaine particulier concerné quand il s'agit de raisonner sur ces représentations.

Les approches concernées par la représentation des connaissances proprement dite connaissent un certain essor récent avec le développement de ce que l'on appelle des "ontologies", cf. par exemple CYC (Lenat, 1995), sortes de bases de données censées capturer l'ensemble des concepts humains sous forme d'un ensemble d'objets conceptuels et de relations entre ces objets. Une sémantique appropriée est censée modéliser logiquement ces relations avec des règles, qui ne sont généralement pas rendues publiques. On peut s'interroger sur le succès véritable de ces entreprises dans la mesure où elles ne sont pas vraiment mises à l'épreuve, mais elles résument la démarche de la représentation des connaissances : choix des objets à modéliser, choix des relations pertinentes entre ces objets, et modélisation de ces relations de façon logique. Le terme "ontologie" est un peu ambigu, dans la mesure où il désigne plus généralement des travaux qui se préoccupent des fondements conceptuels de la représentation des connaissances, que ce soit les objets du monde réel, ou des catégories de concepts pertinents en général (on peut voir notamment (Guarino, 1998), pour une vue récente sur l'ontologie en I.A.). La question du choix d'entités primitives et de leur nature pour la représentation est notamment un problème central (en raisonnement temporel, cela concerne par exemple la discussion sur le choix des primitives temporelles à modéliser : instants, intervalles, ou événements)⁴.

Une grande caractéristique de ces approches propres à l'I.A. est donc ce côté modélisation par des relations plutôt que par des grandeurs mesurables, ce que l'on désigne un peu vite par le terme "qualitatif". Nous allons voir maintenant ce que nous entendrons par là, notamment dans le domaine du temps, de l'espace et du mouvement.

1.1.3 Le raisonnement qualitatif

Il est sans doute utile de clarifier ce que l'on entend par théorie qualitative, ce terme étant en effet appliqué dans des acceptions très diverses en I.A., parfois contradictoires. Pour certains auteurs il explique l'essence même de nos capacités de raisonnement sur des connaissances complexes. Si nous n'avons pas besoin de recourir à des équations différentielles pour appréhender le monde physique, cela est probablement dû en effet au caractère *qualitatif* de nos connaissances, ce qui s'oppose aux données numériques quantitatives. Hernández (Hernández, 1994) caractérise les connaissances qualitatives comme le savoir pertinent essentiel par rapport à une tâche donnée, ce qui est généralement traduit par des relations entre entités ou entre grandeurs numériques pas forcément spécifiées. On considère souvent la possession de données "précises" (sous-entendu : numériques) sur le monde comme indispensable à la manipulation d'information, d'autant que les ordinateurs manipulent très bien des nombres.

Hernández recense pourtant plusieurs raisons pour lesquelles les données numériques sont parfois inadaptées à la résolution d'un problème, une analyse partagée par beaucoup d'auteurs (cf. notamment (Cohn, 1996)), et que nous résumons ainsi :

- La complexité : l'utilisation de données numériques peut forcer à faire plus de distinctions qu'il n'est réellement nécessaire à une tâche donnée, et entraîner ainsi un coût plus important en calcul.

4. L'ontologie formelle est quant à elle la branche de la métaphysique qui applique les méthodes de la philosophie analytique.

- Les erreurs de discrétisation : forcer certaines variables à prendre une valeur peut faire perdre certaines informations qualitatives (l'égalité de deux valeurs par exemple).
- En présence d'information partielle, il faut donner un ensemble de valeurs à des variables numériques ce qui rend les calculs plus complexes.
- Enfin, l'inadéquation des données numériques pour communiquer naturellement avec l'utilisateur d'un système informatique oblige à des traductions internes malaisées quant il s'agit de données imprécises ou incomplètes comme le sont généralement les données spatiales, (ce qui se traduit par l'apparence "vague" de la plupart des expressions spatiales en langage naturel).

Par opposition, l'utilisation de données qualitatives présente les avantages suivants :

- Un nombre de distinctions adapté à la tâche visée ; en physique qualitative par exemple, il suffit souvent de connaître le signe d'une grandeur physique pour raisonner sur un système.
- La comparaison entre valeurs potentiellement inconnues (par exemple pour savoir si un véhicule peut en rattraper un autre il suffit de savoir quelle est la relation entre leurs vitesses : $<$, $=$ ou $>$).
- La sous-détermination permet de ne pas avoir à fixer la granularité d'une représentation *a priori* ; ceci n'est pas vrai pour certains travaux qualifiés de qualitatifs dès qu'il opèrent sur une discrétisation sur un ensemble de valeurs réelles, par exemple des ordres de grandeurs (pour raisonner sur des températures par exemple, on va définir des ordres de grandeurs : $0, \leq 10$, compris entre 10 et 50, etc.). L'indépendance à la granularité où bien la possibilité de s'accommoder de raffinements de la représentation est un défi important qui reste souvent un vœu pieux dans certaines théories qualitatives (comme le système de distances de (Hernández, 1994)).

En résumé, une approche qualitative s'attache à modéliser les propriétés pertinentes par rapport à une classe de problèmes, en essayant de s'abstraire de données quantitatives qui peuvent être imprécises ou incomplètes. On peut voir les avantages d'une approche qualitative sur l'exemple suivant : imaginons un ensemble d'informations portant sur des parcelles de terrains, où on a 4 parcelles, x, x_1, x_2, x_3 , telles que : x est l'union de x_1 et x_2 , et x_3 recouvre x . Si ce qui nous intéresse est seulement les intersections ou les contacts de parcelles, un modèle qualitatif topologique utilise directement le type d'information mentionné et peut donner simplement l'information que x_3 doit nécessairement recouvrir x_1 ou x_2 . Avec un modèle numérique il serait nécessaire de traduire les informations précédentes probablement avec un système d'équations après avoir réduit arbitrairement les parcelles à des polygones (ce qui entraîne presque certainement d'avoir à réécrire symboliquement l'égalité $x = x_1$ union x_2 pour ne pas perdre l'information), puis à calculer les intersections entre x_3 et les autres polygones, ce qui en présence de l'information partielle " x_3 recouvre x " est assez coûteux. Il ne s'agit évidemment pas de dire que les représentations numériques sont sans intérêt dans tous les cas ; simplement pour certaines tâches, et en présence d'un certain type d'information, il y a des avantages à considérer une représentation plus sous-déterminée. Dans la pratique (et notamment pour le type d'exemple que nous avons utilisé), la combinaison des deux est inévitable et il est alors important de caractériser les liens qui existent entre ces deux versants d'une même réalité.

1.1.4 L'espace et le sens commun

Le travail présenté dans cette thèse est une étude du mouvement sous le double point de vue du raisonnement qualitatif et de la représentation de la connaissance spatiale humaine, à des fins computation-

nelles. Les connaissances spatiales sont en effet centrales dans de nombreux domaines de l'Intelligence Artificielle, que ce soit en traitement du langage naturel, pour l'interaction homme-machine, pour le raisonnement automatique, en vision artificielle ou pour le traitement de l'information dans les bases de données géographiques. Un effort important a été mené depuis une dizaine d'années pour construire des modèles formels pour la manipulation d'information spatiales diverses dans des contextes plus ou moins incertain et imprécis, et qui seraient plus facilement manipulables que les données quantitatives brutes, largement répandues en robotique par exemple. La masse de données est en effet un obstacle à une bonne exploitation dans d'autres contextes (bases de données, vision), et le côté éloigné de la cognition humaine de ces modèles est une gêne dans les interactions homme-machine, quand elle n'est pas franchement inadéquate (c'est le cas en traitement du langage naturel où la nature de l'espace manipulé est très éloignée de celle de la géométrie euclidienne). Le "raisonnement spatial qualitatif" (RSQ) (Cohn, 1996; Vieu, 1997) est une branche récente en représentation des connaissances qui s'est focalisé sur ces problèmes et dans laquelle nous plaçons donc nos travaux. La place centrale de l'espace dans le raisonnement de sens commun est manifeste dans les travaux qui se réclament du programme de Hayes, comme certains travaux en physique qualitative, et l'échec relatif de celle-ci à modéliser les notions élémentaires de notre conception de l'espace a sans doute été la raison principale de l'émergence du RSQ comme domaine propre. Il faut enfin remarquer que l'étude du mouvement, qui était un des buts de la physique qualitative (Forbus *et al.*, 1987), a été laissée de côté pendant le temps nécessaire à la reconsidération de la nature de l'espace de sens commun, et que peu de travaux ont abordé le problème de sa représentation et de sa caractérisation dans le type d'approche plus cognitif qui est appliqué au traitement de l'information spatiale. Nous nous intéressons ici justement au traitement du mouvement à la suite de ce courant. L'importance du mouvement est indéniable dans le traitement d'informations spatiales dans la mesure où la plupart des systèmes qui manipulent des données de ce type sont des systèmes évolutifs. La prise en compte du temps dans les systèmes d'informations géographiques est par exemple un point crucial, et la plupart des changements que ces systèmes doivent gérer sont liés à des propriétés spatiales (localisation, évolution de régions, etc.), (Frank et Kuhn, 1995; Hirtle et Frank, 1997); on peut mentionner aussi l'importance du mouvement dans les systèmes de vision artificielle, la prise en compte à un haut niveau nécessitant alors d'avoir des représentations du mouvement proches de l'opérateur humain (ecc, 1996). Enfin les bases de données multi-média, essentiellement vidéos, ont besoin de modèles qualitatifs du mouvement, pour l'indexation par exemple, cf. notamment (Li *et al.*, 1997).

Une perspective orientée vers le sens commun place notre démarche sur un terrain qu'on ne peut isoler de certaines traditions de recherche dans des disciplines connexes telles que la linguistique ou la philosophie quand elle s'intéresse à la nature des objets de notre quotidien, car l'espace, le temps et les phénomènes spatio-temporels sont au cœur des réflexions menées sur l'ontologie de notre monde mental. Les travaux récents en RSQ sur l'espace ont en effet donné une nouvelle jeunesse à des travaux qui s'intéressaient à développer des bases pour la géométrie (entre autres) à partir de choix primitifs différents des mathématiques mieux établies. Notamment certains travaux se sont concentrés sur l'utilisation de primitives différentes du classique point de la géométrie (cf (Tarski, 1969; Tarski, 1972)); les structures de connexion de Whitehead (Whitehead, 1929) sont un autre exemple, en topologie cette fois, de travaux basés sur des primitives non standards; les travaux de l'école polonaise d'entre-deux guerres (Lesniewski, 1989) ont par ailleurs fourni le point de départ de ce qui était vu à l'époque comme une alternative à la théorie des ensembles (et encore aujourd'hui par quelques auteurs comme Barry Smith (Smith, 1993)), la méréologie, et qui est maintenant plus modérément considérée comme une

théorie des relations de partie à tout pertinente pour la modélisation de l'inclusion spatiale sur une base plus intuitive qu'une alternative cartésienne à base de points et d'ensembles de points. A partir de ces théories on trouve des travaux de logique philosophique (plus précisément d'ontologie formelle) qui ont exploré les propriétés formelles de ces structures (Leonard et Goodman, 1940; Goodman, 1977; Clarke, 1981; Simons, 1987) pour finalement inspirer des travaux récents en RSQ⁵. Cela montre l'importance d'une réflexion ontologique sur les objets de notre perception courante, dont la spatialité et le mouvement sont des aspects essentiels et occupent ainsi une place centrale pour le sens commun. L'échec avoué de la physique qualitative dans ses premières tentatives à rendre compte de l'espace de ce point de vue est maintenant en partie expliqué par l'ignorance de ces travaux philosophiques, due à un préjugé maintenant caduc sur leur pertinence pour l'I.A. Il semble que le domaine soit maintenant arrivé à un point où les problèmes de représentation (à travers les choix ontologiques et les propriétés structurelles du domaine étudié) ont pris le pas sur les approches qui accordent plus d'importance au formalisme de raisonnement employé pour résoudre les problèmes (par l'utilisation de logiques non classiques particulières), même si on peut encore voir quelques résurgences (cf. surtout (Shanahan, 1995) et le chapitre 2, section 2.1.3).

1.1.5 Sens commun, perception et “réalité” physique

En considérant l'espace et le mouvement à travers la conception usuelle qu'en ont les humains on ne peut éviter de se poser la question des rapports entre d'une part la *perception* subjective de l'espace et du mouvement (la physique vraiment “naïve” en quelque sorte), et d'autre part ce qu'on pourrait appeler la réalité physique donnée par les théories scientifiques modernes (la relativité par exemple). En fait nous pensons qu'il s'agit, avec le sens commun, de trois choses différentes et que seul le dernier est pertinent pour la formalisation du raisonnement humain. Nous pensons en effet qu'il ne faut se préoccuper que de ce qui est partagé par les humains, les connaissances qui sont *communicables* et sont donc indépendantes d'un observateur particulier, d'un moment ou un lieu particulier.

Sens commun et perception du temps et de l'espace. Les propriétés du temps, de l'espace et du mouvement qui sont perçus subjectivement et qu'à ce titre il ne nous semble pas souhaitable d'intégrer dans une théorie de sens commun sont les suivantes :

- pour le temps : la perception que le temps s'écoule dans un seul sens. Dans beaucoup de théories formelles du temps, seul l'ordre est formalisé avec une symétrie complète du futur et du passé d'un point de vue structurel. Ces théories constituent un niveau d'abstraction qui voit le temps comme un ordre linéaire. Les structures de temps arborescentes visent souvent à modéliser l'incertitude des connaissances par rapport au futur d'un point donné. Il nous semble qu'il est plus naturel et plus conforme à l'intuition de considérer le temps comme un ordre linéaire et de reporter d'éventuelles incertitudes sur une théorie spécifique de la connaissance.
- pour l'espace : le même genre de phénomène apparaît dans le domaine spatial avec la distinction entre, par exemple, une représentation totale (comme une carte) avec un certain niveau d'abstraction, et une représentation subjective (la vue d'un observateur par exemple) dans laquelle les objets de l'espace peuvent se cacher les uns les autres (le phénomène d'occlusion en vision). Nous pensons qu'une théorie du sens commun doit s'occuper de cette partie des connaissances qui

5. Ces travaux philosophiques continuent, notamment dans les travaux de Varzi (Varzi, 1996).

n'est pas concernée par ces limitations de la perception. Il faut noter que les travaux de Barbara Tversky (Taylor et Tversky, 1992; Tversky, 1993; Tversky, 1996) ont montré que les humains manipulaient différentes sortes de représentations de leur environnement spatial, les deux plus basiques étant le "tour", une vue subjective avec mouvement fictif, et le "survey", sorte d'image mentale reconstruite à partir de la perception et que l'on peut assimiler à une carte mentale. Ces deux types de représentation semblent révéler l'existence d'un modèle mental en 3D dans lequel un humain peut se "promener" et en extraire les deux autres formes de représentation (vue subjective et carte). C'est cette dernière forme de représentation, indépendante du point de vue, sur laquelle nous désirons mettre l'accent.

- pour le mouvement : c'est peut-être là que la différence est la plus marquée. Il y a peut-être eu un malentendu sur la nature du sens commun tel que le comprenait Hayes en parlant de "physique naïve" dans son manifeste. Certaines expériences ont en effet montré que des sujets sans connaissances poussées en physique avaient des théories (fausses) communes à propos du mouvement ; dans (McCloskey, 1983) sont présentées quelques-unes de ces idées erronées ; certains sujets croient par exemple qu'un jet d'eau sortant d'un tuyau en arc de cercle aura une trajectoire circulaire à la sortie du tuyau. Ou encore, beaucoup de sujets n'ont pas assimilé la réalité de l'inertie du mouvement et prédisent incorrectement certaines trajectoires d'objets en chute libre. Il nous semble qu'une théorie du sens commun qui veut formaliser les connaissances humaines de base sur l'espace et le mouvement ne peut se donner pour but d'intégrer des notions qui mènent assez facilement à des incohérences (de plus il est généralement assez simple d'expliquer leur erreur aux sujets concernés), même si cela correspond à des erreurs "naturelles". Ce point de vue n'est pas partagé par certains auteurs, comme Forbus (voir le chapitre suivant), qui propose plusieurs modèles de la dynamique des objets dont un doit correspondre à certaines visions naïves de la physique, quitte à gérer les incohérences à un autre niveau et choisir celui qui s'adapte à une situation donnée. Remarquons quand même que de tels modèles "naïfs" n'ont pas, à notre connaissance, été mis en œuvre automatiquement par aucun des auteurs qui les ont mentionnés.

Le sens commun n'est pas la physique. Le sens commun n'est pas non plus une approximation des théories physiques établies, comme ont tenté de le modéliser les tenants de la physique qualitative. Les buts sont différents et même si les deux sont les reflets d'une certaine réalité, il se préoccupent de choses différentes. Il n'est pas question par exemple de modéliser la structure de l'espace-temps telle que nous l'explique la relativité. Il est difficile de prétendre sans mauvaise foi que les connaissances communes aux êtres humains intègrent la nature relative du temps et la courbure de l'espace-temps. En conséquence, une théorie du sens commun doit s'intéresser à un espace ancré dans notre perception du monde courant, avec un temps non relatif aux observateurs notamment. Par ailleurs les questions sur la nature conventionnelle ou non de la géométrie de l'espace-temps ne nous concernent pas vraiment. Dans tous les cas c'est ce qui est observable empiriquement qui fait notre champ d'investigation.

En conclusion nous pouvons dire que notre conception de ce que doit prendre en compte une théorie de sens commun se situe hors des aspects psychologiques autres que ce qui fait indiscutablement partie de l'appréhension de l'espace commune aux hommes, même s'il est parfois délicat de tracer une frontière bien nette entre ceci et ce que nous apprend la physique, ne serait-ce que parce que ces théories sont plus ou moins connues par beaucoup de monde et influencent notablement ce que l'on

appelle l'intuition. Les conceptions intuitives sur l'espace, le temps et le mouvement d'un homme du XVI^{ème} sont certainement différentes des intuitions modernes. On peut donc reprendre la formule de Lenat, cité par (Davis, 1990) définissant le sens commun comme une "consensus reality" (Lenat *et al.*, 1986), une réalité de consensus, avec ce que cela implique d'évolutif et de partiellement contestable. Par exemple, même s'il n'est pas dans nos objectifs de modéliser la mécanique newtonienne, certains de ses concepts (comme l'attraction des corps par gravité) ont été intégré dans le sens commun, de façon qualitative au moins, et sûrement pas par des équations. Il reste donc la question de savoir où s'arrêter entre les barrières un peu floues que nous avons essayé de distinguer pour une voie moyenne entre l'objectif de la physique et le subjectif lié à la perception.

1.2 Principes méthodologiques

Nous arrivons alors à une question centrale dans toute discipline et qui est une source de controverse interminable en I.A., et qui est celle de la validation des travaux que l'on propose. Quelle est la validité des réductions que l'on fait pour modéliser un phénomène cognitif, et comment la justifier? Cette question est notamment au centre des théories de sens commun de l'action, où les théories se sont longtemps multipliées pour répondre à des problèmes spécifiques posées à des théories précédentes et se contentaient généralement de résoudre un type d'exemple avant qu'un nouveau contre-exemple vienne les balayer à leur tour (généralement en modifiant un peu l'exemple de référence). Erik Sandewall (Sandewall, 1994) a plaidé pour une méthodologie plus précise dans le raisonnement sur l'action pour obliger chaque théorie à définir son domaine de validité, et essayer de mettre fin à des années de "chasse à la dinde" plus ou moins stériles. Plus généralement on peut se demander quels objectifs se fixe une théorie du sens commun et comment on peut vérifier l'adéquation du résultat avec le but initial, notamment quand il s'agit de l'espace. Une fois sorti des "toy problems" tant critiqués par Hayes, qu'est ce qui garantit le succès d'une théorie? Si l'on se contente de vérifier qu'une théorie correspond à "l'intuition" on se heurte au problème des intuitions contradictoires que différentes personnes peuvent avoir sur un même phénomène. De plus, à force de travailler sur un certain sujet, les intuitions d'une même personne évoluent inmanquablement et pervertissent le jugement critique nécessaire à l'évaluation. Dans ce cas le bon usage des intuitions semble être de les laisser aux autres... ou bien de les évaluer sérieusement de façon expérimentale (par des expériences de psychologie). Bien que ce soit l'objet de nombreux travaux en psychologie cognitive, les travaux conjoints entre chercheurs en I.A. et en psychologie restent assez rares; on peut noter quand même les exemples de (Aurnague *et al.*, 1998; Schlieder, 1995; Gryl, 1996).

Parmi les candidats possibles pour remédier à ce problème, certains auteurs pensent que le langage naturel nous fournit une base, "objective" en un sens, de ce qu'est une intuition acceptable par l'humain, à travers l'acceptabilité de phrases décrivant un problème (du moins si on peut exprimer ces problèmes en langage naturel). C'est la méthode empirique adoptée notamment par (Aurnague *et al.*, 1997; Hobbs *et al.*, 1987; Aurnague, 1991).

Il faut évidemment ajouter à ceci un souci de cohérence des théories proposées, ce que la plupart des auteurs vérifient en RSQ avec une rigueur variable. La plupart des travaux auxquels nous nous sommes intéressés sont des théories axiomatiques⁶, et on pourrait classer les méthodes par ordre de rigueur croissant de la façon suivante: d'abord on peut montrer les théorèmes dérivables par la théorie pour

6. Certaines approches du raisonnement spatial sont par contre résolument et consciemment non qualitatives, par exemple (Glasgow, 1993).

caractériser certaines de ses propriétés (c'est l'approche la plus répandue en raisonnement spatial), on peut montrer la consistance logique d'une théorie en exhibant un modèle (ce qui est déjà plus rare, et souvent fait longtemps après la première publication présentant la théorie, comme par exemple pour la théorie du temps de Allen, ou celle de l'espace de Randell *et al.*), on peut caractériser complètement une théorie en donnant la classe de modèles exactement capturée par la théorie (ce qui est encore plus rare en raisonnement spatial qualitatif, on peut noter (Asher et Vieu, 1995; Pratt et Lemon, 1997)). Enfin, il faut également se poser la question de l'application que la théorie peut avoir informatiquement, de l'usage que l'on peut en faire. Ceci est généralement séparé de l'étude théorique, ce qui paraît souhaitable par rapport aux excès des approches *ad hoc*. Il nous paraît cependant utile d'indiquer le genre d'application auquel une théorie, même d'un aspect réduit du sens commun, peut apporter un plus. Cela implique d'une part de considérer la complexité des algorithmes de raisonnement que l'on peut utiliser pour manipuler les données considérées par la théorie, et qui est une source de travaux spécifiques, (en raisonnement temporel de façon très développée, en raisonnement spatial de façon un peu plus balbutiante). D'autre part on peut indiquer dans quelle mesure une théorie est exploitable en considérant un type d'applications précises dans lesquelles les concepts étudiés sont mis en jeu. Dans (Davis, 1990), on trouve une caractérisation schématique d'une méthodologie pour l'automatisation du sens commun (reproduite à la figure 1.1), où l'approche globale est guidée par l'application qui est le but du développement d'une théorie. Il nous semble plus intéressant, pour pouvoir considérer qu'une

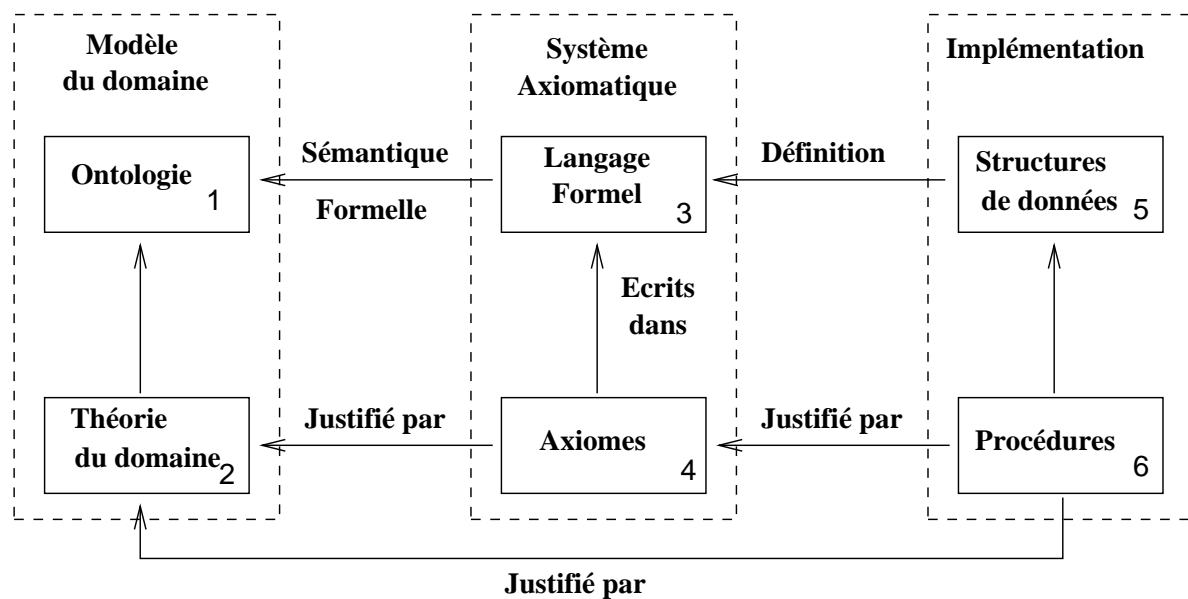


FIG. 1.1 - La méthodologie du sens commun vu par (Davis, 1990).

théorie de sens commun est réutilisable, de la considérer comme un modèle d'une partie de la réalité que l'on veut modéliser, et que l'on peut évaluer suivant les critères esquissés plus haut : d'une part on s'assure de la cohérence de la théorie et on caractérise exactement les structures qu'elle capture, ensuite on vérifie son adéquation (forcément partielle) à la partie du sens commun que l'on vise en testant son pouvoir expressif et les inférences qu'elle permet par rapport à la sémantique du langage naturel, et enfin on indique dans quelle mesure la théorie est opératoire en la mettant en œuvre sur un problème

précis, et en indiquant alors quelle sorte de problèmes on peut résoudre de façon pratique (ce qui est résumé à la figure 1.2). Idéalement, cela correspond à donner des algorithmes de calcul dans certaines situations, ainsi que leur complexité.

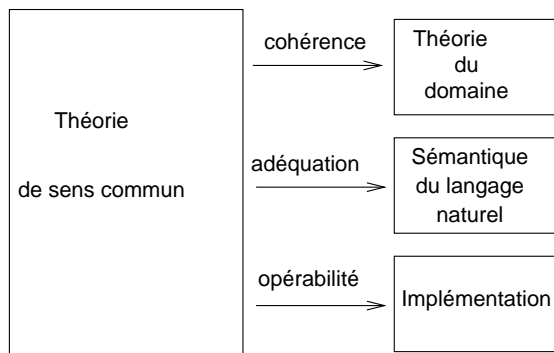


FIG. 1.2 - *Trois aspects de la validation d'une théorie.*

Nous allons maintenant revenir un peu plus en détail sur les caractéristiques principales de notre approche et sur les moyens que l'on veut se donner de l'évaluer.

1.2.1 Une approche axiomatique

Notre approche de la représentation est une approche axiomatique : les concepts étudiés sont modélisés par une théorie logique du premier ordre dont on étudie les propriétés expressives et inférentielles. L'utilisation de la logique étant un sujet de controverse incessant en I.A. de façon générale, en représentation des connaissances (spatiales ou non) et en Traitement Automatique des Langues en particulier, il nous paraît nécessaire de préciser quel rôle nous faisons jouer à la logique dans ce travail. Il y a une confusion tenace dans certains domaines entre d'une part l'utilisation de la logique pour représenter des concepts pertinents pour une tâche donnée et l'utilisation de la logique comme formalisme computationnel (avec un démonstrateur de théorèmes par exemple) ; certains vont même jusqu'à assimiler l'utilisation de la logique du premier ordre et la programmation en Prolog. Il est bien clair dans notre étude que la logique classique est un *moyen* d'étudier un objet de façon claire et rigoureuse, dans la mesure où elle permet l'expression formelle des concepts que l'on tente de représenter, et qu'il est facile alors d'étudier les propriétés des modèles que l'on manipule ainsi. L'étude logico-mathématique des objets auxquels on s'intéresse est juste un moyen de s'assurer que les constructions que l'on va vouloir utiliser ensuite pour des tâches diverses (représentation du langage, manipulation de base de données spatiales par exemple) sont bâties sur des fondations solides et cohérentes. Une autre bonne raison est que la logique est un langage clair et précis, très largement admis comme un "standard", et qui permet de faciliter la communication entre chercheurs (même si ce n'est pas un standard universel)⁷. Notre objectif ici est de montrer le domaine de validité des objets auxquels on s'intéresse en même temps qu'on les introduit, et pour cela la logique est un outil avéré pourvu qu'on respecte certaines conditions. L'utilisation d'un langage formel n'est cependant pas en lui-même le garant de la validité d'une approche, si on s'en sert uniquement pour la représentation. C'est l'étude des modèles de la théorie développée qui permet de dire que l'on a bien caractérisé les notions que l'on étudie. C'est pour ne

7. On peut voir une vision du débat à ce propos dans (Israel, 1985)

pas avoir été au bout de cette démarche que certains travaux dans le domaine de la représentation spatiale ont été l'objet de nombreuses remises en cause ; l'exemple le plus frappant en est la théorie de référence RCC-8 de Randell, Cui et Cohn ; cette théorie pionnière en RSQ et une des plus connues attend pourtant encore une caractérisation exacte de ses modèles (pour l'instant on n'en connaît que quelques-uns), et son extension RCC15 attend toujours un modèle non trivial.

1.2.2 Le rôle du langage

Le langage joue un rôle double dans le genre d'étude que nous proposons : d'une part il fournit certaines manifestations des conceptions humaines des données que l'on étudie : les expressions spatiales comme les prépositions ou les verbes de mouvement nous renseignent sur les catégories conceptuelles et les configurations spatiales qui semblent pertinentes. Le langage est alors une source de données, un moyen d'accès pour appréhender ce que sont le temps, l'espace et le mouvement pour la cognition humaine. D'autre part, le langage fournit un objectif facile à définir pour une théorie de sens commun : si la théorie permet de représenter simplement et de façon adéquate la sémantique des marqueurs liés au domaine que l'on cherche à représenter (par exemple les prépositions spatiales et les verbes de mouvement dans le cas de l'espace), on a une première validation de la pertinence de cette théorie. Cela peut se vérifier avec les inférences que l'on peut exprimer dans un cadre formel et qui correspondent à celles qu'un locuteur en possession des mêmes informations peut faire. Par exemple, la préposition "dans" a parfois été interprétée comme traduisant simplement l'inclusion spatiale ; celle-ci étant transitive on peut alors faire l'inférence suivante⁸ :

La balle est dans la main.
 La main est dans le gant.
 ⇒ La balle est dans le gant.

Ce qui ne paraît pas acceptable.

Il faut cependant considérer les intuitions que peut fournir l'usage de la langue avec circonspection. Le sens des mots est très dépendant du contexte dans lequel ils sont énoncés, et l'ontologie présupposée par le langage peut avoir quelques problèmes quand on considère certains raisonnements (il suffit de penser aux nombreux paradoxes que l'on peut exprimer en langage naturel et qui proviennent tous plus ou moins d'ambiguïtés du langage lui-même, au caractère vague de certaines expressions, etc). Certains philosophes y ont même vu une raison centrale d'abandonner l'utilisation du langage pour se tourner vers les langages formels comme la logique afin de présenter leurs idées. Le problème de la modélisation du sens commun exige que l'on garde à l'esprit cette ontologie que nous fournit le langage car elle fait partie de notre appareil conceptuel et il est difficile de construire une théorie sans au moins indiquer ce à quoi elle réfère dans nos conceptions. Il faut garder à l'esprit la phrase du philosophe J.L. Austin : "Le langage ordinaire n'est pas le dernier mot , (mais) il faut nous souvenir que c'est le premier mot." D'autant que l'interprétation que l'on peut faire des données linguistiques dans une optique ontologique est sujette à caution : il existe un débat vivace entre philosophes pour décider quelles sont les entités qui constituent la base de notre conception intuitive du monde : points, régions de l'espace, où trajectoires dans l'espace-temps, et chaque option a des défenseurs qui utilisent le langage comme caution ("le langage pense comme moi") ou bien comme repoussoir ("le langage permet d'exprimer des paradoxes avec cette vision du monde, donc il faut proposer autre chose"). Une des critiques les plus radicales (que l'on trouve notamment chez Mark Heller, cf chapitres 4 et 6) de l'ontologie liée au langage naturel affirme qu'aucune des entités introduites dans celui-ci (les *tables*, *chaises*, *chiens*,

8. Cet exemple est dû à Laure Vieu.

chats) n'a d'existence réelle, mais qu'ils correspondent plutôt à des descriptions vagues de certaines catégories d'objets que nous percevons (des régions de l'espace-temps). Nous nous situons dans une optique intermédiaire : les données linguistiques sont prises en compte, mais on tente de s'en abstraire pour proposer un modèle cohérent d'objets du monde, qui peut éventuellement servir à définir la sémantique supposée des mots de la langue qui réfèrent aux mêmes phénomènes. Par exemple nous faisons le choix de considérer comme primitives des entités de l'espace-temps, notamment sur des bases linguistiques, mais ce n'est qu'après avoir élaboré une théorie satisfaisante de ces entités que nous tentons de l'appliquer à la représentation de phénomènes langagiers (à travers une sémantique formelle des expressions spatiales et de mouvement).

1.2.3 Validation opératoire

La dernière forme de validation que l'on doit respecter dans une modélisation du sens commun est de répondre à la question : la théorie a-t-elle une utilité pratique ? L'utilisation de la logique ne doit pas nous faire oublier qu'un but de la théorisation du sens commun est finalement la mise en œuvre informatique, et que toute théorie doit alors se préoccuper des conditions dans lesquelles on peut l'utiliser opératoirement. Notre position est la suivante : il est essentiel de développer de façon cohérente (d'où la démarche formelle) une théorie des objets que l'on étudie afin de fournir des catégories utilisables dans une tâche donnée. L'utilisation pratique d'une théorie peut ensuite s'autoriser une certaine marge dans l'exploitation des propriétés de cette théorie, mais il est nécessaire d'avoir un certain niveau d'abstraction pour prétendre à une quelconque généralisation de n'importe quelle étude. La théorie de Allen est au départ une théorie logique du temps, fondée sur des axiomes précis qui lui donnent certaines propriétés formelles. Le succès de cette théorie qui a largement débordé le cadre de l'I.A. proprement dite est cependant plus liée au vocabulaire précis qu'elle donne pour parler du temps et d'événements qui ont lieu dans le temps qu'à la forme particulière de tel ou tel de ses axiomes. C'est ce vocabulaire qui est connu et utilisé pratiquement et pas la théorie elle-même, qui joue cependant un rôle clarificateur évident. Notre travail peut se comprendre de la même façon : développer une théorie formelle cohérente qui fournit un certain vocabulaire et certaines catégories pour parler du mouvement et raisonner sur le mouvement, puis l'utiliser dans un cadre plus contraint, plus spécifique, qui fait appel à beaucoup d'autres notions qui ne sont pas aussi facilement généralisables, mais qui présente un intérêt pratique.

1.3 Résumé de l'étude

Nous allons maintenant reprendre la structure du travail présenté ici, en insistant sur l'articulation des chapitres par rapport au squelette méthodologique dont nous avons discuté ci-dessus. Dans un premier temps nous verrons comment l'I.A. a abordé le problème de la représentation spatiale et de la modélisation du mouvement dans le domaine du raisonnement qualitatif et les limites avouées de ce genre d'approche. Un survol de travaux en linguistique cognitive donnera alors un panorama des aspects à prendre en compte pour se rapprocher des conceptions humaines des notions qui nous intéressent, tout en soulignant leurs insuffisances quant il s'agit de la représentation effective des concepts. Nous verrons alors certains travaux qui commencent à remettre en cause certains présupposés des travaux précités, en considérant l'espace et le temps conjointement. Il apparaît à ce stade (chapitre 2) que l'inadéquation des choix ontologiques (essentiellement ceux la géométrie classique) est un obstacle à une théorie appropriée à la modélisation du sens commun, et est largement remise en cause par de nombreux travaux qui refondent les théories de l'espace en partant sur des primitives différentes. Nous

verrons alors au chapitre 3 comment les approches qualitatives du temps et de l'espace ont rebâti des théories mieux adaptées à la modélisation de l'espace et du temps, essentiellement en considérant des primitives étendues (intervalles pour le temps, régions pour l'espace) en lieu et place des instant et des points. Ces structures conservent certaines propriétés des modèles classiques (et certains résultats d'équivalence sont alors présentés) et il faut insister sur le fait que c'est essentiellement le côté pratique et plus naturel de ces théories qui en fait leur intérêt. Nous verrons alors au chapitre 4 comment la représentation du mouvement et de l'espace-temps a été considérée dans des cadres similaires, et à partir de leurs insuffisances, nous présentons alors une théorie qualitative de l'espace-temps prenant comme primitives des régions de l'espace-temps, dont nous montrons les propriétés et la définition sémantique. Ensuite, le chapitre 5 introduira une notion de continuité qualitative, qui dans le cadre proposé au chapitre précédent, permet de caractériser une propriété essentielle du mouvement de sens commun et de formaliser certaines intuitions déjà présentes dans certains travaux sur l'espace qualitatif. Le chapitre 7 permet alors de définir un calcul sur des relations de mouvement dans un cadre topologique qualitatif et de voir quelles sortes d'inférences sont réalisables dans un tel cadre. A ce stade, nous aurons une théorie de l'espace-temps dont la cohérence et les modèles auront été caractérisés, et nous aurons vu l'utilisation que l'on peut faire de la théorie pour le raisonnement. Si nous reprenons les types de validation dont nous avons parlé, nous aurons à ce stade considéré uniquement la validation d'un point de vue formel. Le chapitre 6 présentera alors une étude sémantique d'une certaine classe de verbes de mouvement du français qui nous permet d'étudier l'adéquation du cadre formel à la représentation des concepts de sens commun tels qu'ils apparaissent dans la langue et de valider l'approche de façon plus cognitive. Ce travail est poursuivi chapitre 8 avec un modèle de représentation des itinéraires, un domaine de recherche dont nous préciserons l'importance, et nous montrons comment une théorie qualitative du mouvement peut enrichir les approches traditionnelles du problème. Enfin nous verrons au chapitre 9 comment cette représentation d'itinéraires et la sémantique des verbes alliées à un modèle de raisonnement sur le mouvement tel qu'il a été présenté chapitre 7 peuvent être mises en œuvre informatiquement comme un module d'un système de description de haut niveau de scènes numériques, dans le cadre d'un projet plus global (le projet Wire (Borillo et Pensec, 1995; Pensec, 1996)). Ceci constituera un exemple de validation opératoire de l'approche et fournit un test empirique du bien fondé de certains choix effectués au cours de ce travail. Nous aurons alors abordé les trois sortes de mises à l'épreuve dont on peut disposer quand on cherche à valider une théorie de sens commun : cohérence formelle, adéquation cognitive et opérabilité.

1.4 Notations

Afin d'harmoniser l'étude des différents travaux qui mènent à notre propre travail nous avons effectué une réécriture partielle de certaines propriétés exprimées axiomatiquement. Nous utiliserons ainsi les notations suivantes pour tout ce qui est écrit en logique du premier ordre :

- \triangleq introduit une définition logique.
- $=$ est l'égalité logique, axiomatisée classiquement (elle est réflexive, symétrique, transitive et permet la substitution).
- \wedge , \vee , \neg , \rightarrow et \leftrightarrow désignent respectivement la conjonction, la disjonction, la négation, l'implication matérielle et l'équivalence logique.
- \oplus désigne le "ou exclusif" (qui est une abréviation : $p \oplus q \triangleq ((p \vee q) \wedge \neg(p \wedge q))$).

- x, y, z, t, u, v, w , indicés ou non, dénotent des variables.
- \vdash désigne une dérivation syntaxique ; ainsi $P_1, P_2 \vdash P_3$ indique que l'on peut montrer P_3 à partir de P_1 et P_2 .
- Pour des raisons de lisibilité, et suivant en cela de nombreux auteurs, le parenthésage des prédicats et de certaines fonctions sera supprimé lorsqu'il est superflu. Idem pour les quantificateurs universels dont la portée est la formule globale (les variables libres seront donc implicitement quantifiées universellement). Par exemple, $Pxy \rightarrow P(ix)(iy)$ abrège $\forall x \forall y (P(x, y) \rightarrow P(i(x), i(y)))$

Les notations supplémentaires seront introduites au fur et à mesure des besoins.