

# S'orienter dans un monde inconnu

**Simon Lacroix et Raja Chatila**

**LAAS/CNRS**

**7, Ave du Colonel Roche**

**31077 TOULOUSE Cedex**

Les futurs robots planétaires devront accomplir sans intervention d'opérateurs des missions d'exploration de plus en plus longues et complexes. Plus proches de nous, les robots domestiques qui réaliseront des tâches de ménage par exemple, ou les robots de service qui distribueront le courrier dans des bureaux devront le faire sans notre assistance : il nous suffira de leur donner des ordres. Ces différents types de robots mobiles opéreront dans un environnement dont ils ne connaissent pratiquement rien au départ, et se trouvent alors dans une situation analogue à celle d'un explorateur face à une contrée inconnue. Pour réaliser efficacement les missions qui leurs sont allouées, ils doivent décider de stratégies de navigation, planifier des trajectoires et les exécuter. Cela implique qu'ils perçoivent l'environnement qui les entoure, qu'ils le modélisent, en en construisant une représentation intelligible à partir des données fournies par leurs capteurs : c'est la cartographie. À cette fin, mais aussi pour surveiller le contrôle de leur mission et des trajectoires qu'ils décident d'exécuter, il doivent connaître en permanence leur position dans cet environnement : c'est la localisation.

En revanche, si un robot dispose d'une connaissance *a priori* de son environnement, telle qu'un plan des locaux où il doit opérer, sa situation n'est alors plus celle d'un explorateur en terrain inconnu, mais plutôt celle d'un piéton muni d'un plan de la ville dans laquelle il se situe. On pourrait alors penser qu'il n'a pas à construire de modèle de l'environnement, et qu'il lui suffit de se localiser dans la carte initiale. Mais cela implique de mettre en correspondance ce qu'il perçoit avec sa carte, et nécessite la construction d'un modèle dédié à la localisation. Par ailleurs, la carte dont il dispose n'est pas forcément précise, complète et à jour, et il lui faut donc l'amender.

La modélisation de l'environnement et la localisation du robot sont incontournables en robotique mobile : c'est en représentant son environnement et en s'y situant que le robot appréhende l'état du monde, et peut donc y prendre les décisions adéquates pour mener à bien ses missions. Ces deux capacités sont intimement liées, car s'il est nécessaire de représenter des éléments de l'environnement pour s'y localiser, il est aussi indispensable de se localiser pour représenter l'environnement.

C'est en exploitant les données de ses capteurs qu'un robot peut réaliser ces fonctions. On distingue les capteurs *proprioceptifs*, qui le renseignent sur ses mouvements (accéléromètres, compas, gyromètres et odomètres par exemple), et les capteurs *extéroceptifs*, qui le renseignent sur son environnement. Ces derniers se regroupent principalement en deux catégories : les *télémetres*, qui fournissent au robot des informations sur la géométrie de son entourage, et bien sûr les caméras, qui lui permettent d'en appréhender l'aspect visuel. La construction de modèles de l'environnement et la localisation du robot à partir de ces données

met en œuvre des processus de structuration et d'abstraction, qui analysent les données pour en extraire des représentations exploitables. La difficulté est liée à la nature de ces données, qui sont toujours imprécises, ce qui est inhérent à tout processus physique de mesure, et peuvent parfois être erronées, et ne correspondre à aucune réalité. Elles sont par ailleurs incomplètes ou partielles, notamment à cause des occultations générées par les différents éléments présents dans l'environnement.

## Représenter son environnement...

Une représentation de l'environnement est essentielle à un robot autonome pour planifier des trajectoires (de déplacement, de saisie), pour déterminer des stratégies d'action et d'exploration, et pour se localiser. Afin de permettre la réalisation de ces différentes fonctions, les modèles doivent expliciter des informations *géométriques*, qui représentent la forme des lieux et des objets, *topologiques*, qui représentent les relations spatiales entre différents lieux et objets, et *sémantiques*, qui représentent la nature (le sens, la fonction) des lieux et objets. Ainsi planifier une trajectoire locale ne requiert que la détermination de l'espace accessible à partir de la position courante, information exclusivement géométrique, tandis que la détermination de stratégies de déplacement requiert une connaissance plus abstraite de l'environnement, qui explicite sa topologie et peut inclure des informations sémantiques, à partir desquelles le robot peut par exemple planifier une séquence de déplacements décrits à la manière d'un itinéraire (« emprunter le couloir A, pénétrer dans la pièce B par la porte C... »).

Considérons comment un robot peut construire une représentation à partir de données acquises d'une seule position. La figure 1 est un exemple de processus de modélisation géométrique dans un cas simple, en deux dimensions, à partir de données télémétriques acquises dans un appartement. La méthode consiste à représenter les données par des *primitives*, formes géométriques analytiques, ici des segments dans un plan, représentées et situées par rapport au robot par les équations qui les définissent. Mais cette technique de modélisation se ramène à un processus de structuration des données, l'abstraction est faible, on passe simplement de données géométriques « brutes » à des données géométriques plus structurées. Le problème est cependant délicat, notamment à cause de la présence des imprécisions dans les données, connues grâce à un modèle de fonctionnement du capteur. Les primitives héritent de ces imprécisions, les paramètres qui les définissent étant eux-mêmes imprécis. La maîtrise de ces imprécisions est nécessaire aux algorithmes qui vont exploiter le modèle construit : ainsi le logiciel qui va planifier une trajectoire ne pourra décider de passer entre deux segments que s'il est certain que la distance qui les sépare permet le passage du robot. Nous allons voir que ces imprécisions sont aussi nécessaires pour analyser le modèle, et pour le compléter lorsque de nouvelles données seront acquises au cours des déplacements. Cette approche de la modélisation d'environnements fortement basée sur la géométrie est très représentative des travaux développés par les chercheurs dès le milieu des années 80, qui les ont conduit à définir des algorithmes de manipulation de géométrie imprécise. Plus récemment, les progrès dans la technologie des capteurs et des calculateurs, mais aussi dans les algorithmes ont permis aux robots de construire des représentations tridimensionnelles de l'environnement (figure 2).

L'extraction d'informations topologiques à partir d'un modèle géométrique est assez aisée, tant que l'on ne cherche pas à identifier les lieux perçus. Une description polygonale peut par exemple être partitionnée en régions connexes, dont les relations de voisinage définissent un graphe à partir duquel le robot peut déterminer des chemins.

Si une description purement géométrique permet de déterminer l'espace accessible au robot et donc de générer des trajectoires et des chemins, elle n'explique pas les notions sémantiques dont il a besoin pour planifier des itinéraires ou interpréter des missions. De telles informations ne sont pas directement perçues par les capteurs, et elles doivent donc être déduites d'une analyse des formes modélisées, en considérant par exemple que deux segments colinéaires distants d'un mètre définissent une porte. Il s'agit cette fois d'un processus d'abstraction, d'interprétation, qui transforme des informations numériques en des informations symboliques, et traduit les imprécisions sur les paramètres des segments en une *incertitude* sur le fait qu'ils représentent bien une porte. Posée ainsi, la construction de représentations sémantiques repose sur le succès de la modélisation géométrique, et l'on voit sur la figure 1 que les informations perçues d'une seule position sont insuffisantes (incomplètes) pour reconnaître la porte. Il faudra donc considérer des données acquises d'autres positions, ce qui implique de déplacer le robot de manière adéquate, afin de confirmer ou d'infirmer la présence d'une porte. Le processus de modélisation est donc un processus *actif*, qui ne se contente pas d'interpréter les données perçues au fur et à mesure de leur arrivée, mais qui décide des données à acquérir, et qui pour cela va générer des déplacements du robot. Curieusement, ceci n'a vraiment été considéré par les chercheurs qu'au début des années 90.

De plus, définir une porte de manière exclusivement géométrique est très pauvre, particulièrement en deux dimensions, où deux meubles parallèles et distants d'un mètre seront interprétés comme une porte, et où les portes fermées ne seront pas détectées. Il faut donc enrichir cette définition, en ajoutant par exemple la forme verticale ou la présence d'une poignée. Ces informations peuvent nécessiter l'intégration dans la représentation d'autres données perceptibles le robot à partir de capteurs plus riches, notamment la vision. C'est la *fusion multisensorielle*, qui va permettre d'affiner l'interprétation, en mixant dans un même modèle des attributs de natures différentes, mais toujours imprécis ou incertains, tels que la géométrie et la couleur.

On voit vite les limites d'une approche qui ne définit la sémantique des éléments de l'environnement que par rapport à ce que le robot peut percevoir. Le concepteur est en effet obligé de fournir au robot dans le détail la description de la totalité des objets possibles. Ainsi, décrire au robot une chaise comme « un plateau horizontal posé sur quatre pieds, auquel est attaché un plateau vertical » ne lui permettra d'identifier que les chaises de ce type, et non les milliers de modèles de chaises qu'ont pu définir les designers. Par contre, définir une chaise conceptuellement, par sa fonction, comme « un objet sur lequel un bipède peut s'asseoir en faisant reposer son dos » est bien plus général. Les chercheurs en robotique commencent seulement à exploiter de telles considérations encore difficiles à modéliser, en définissant par exemple une porte comme une « frontière franchissable entre deux lieux », et un lieu comme une « zone dans laquelle le robot peut évoluer aisément ». Ceci illustre le lien très fort qui existe entre la perception du monde et l'action ou la capacité d'action dans ce monde, lien qui a été observé depuis longtemps dans les êtres vivants.

Les difficultés d'une approche de la modélisation d'environnements exclusivement basée sur la géométrie, que l'on peut qualifier de « reconstructionniste », et particulièrement les difficultés posées par la manipulation de données imprécises, ont conduit dès les débuts de la robotique le développement d'autres méthodes, plus adaptées aux possibilités des calculateurs. Ainsi, un moyen devenu classique de représenter l'espace accessible consiste à décrire l'environnement par un ensemble de petites cellules, analogues aux pavés d'un carrelage, pour chacune desquelles on détermine la probabilité qu'elle corresponde à un obstacle (figure 3).

L'intérêt d'une telle représentation est qu'elle explicite directement la notion sémantique « obstacle », avec une incertitude qui n'est pas dérivée d'imprécisions géométriques. Des informations topologiques sont présentes dans les relations de voisinage entre les cellules, et peuvent être compactées en regroupant les cellules similaires selon des procédures classiques de segmentation d'images, qui permettent d'exhiber des régions de pixels (de cellules) homogènes. Ce type de représentation, très aisé à construire et à manipuler, a aussi l'avantage de représenter clairement la quantité d'information dont dispose le robot sur une zone donnée, ce qui permet de déterminer des stratégies d'exploration, en cherchant à percevoir de nouveau les cellules pour lesquelles l'incertitude est grande. Cependant, la sémantique de la représentation est restreinte à la notion d'obstacle : il est difficile de déduire la présence d'une porte à partir d'une telle représentation par exemple. De manière analogue, des méthodes permettent d'extraire des informations sémantiques d'autres natures, à partir des données brutes sans les structurer en primitives, mais en calculant des attributs particuliers qui permettent d'identifier ce qui est perçu en les comparant à un ensemble d'exemples préalablement appris par exemple. Ces dernières techniques permettent d'aborder les problèmes de reconnaissance d'objets ou de lieux par vision notamment, elles commencent à peine à être appliquées à la robotique mais semblent prometteuses.

### **... et s'y localiser.**

Pour disposer d'informations plus riches sur son environnement et en améliorer sa compréhension, nous avons vu que le robot doit agréger au cours de ses déplacements les différentes perceptions locales en une représentation globale. Le principe de cette agrégation peut paraître simple : il suffit de se déplacer, et de concaténer les informations nouvellement perçues aux informations déjà connues. Cela nécessite de considérer les imprécisions sur les données et primitives pour mettre à jour leur description, et surtout de connaître le déplacement effectué entre deux perceptions. À cette fin, le robot peut exploiter les informations de ses capteurs proprioceptifs, ou par exemple analyser le mouvement des pixels tel que perçu par ses caméras lors de ses déplacements (VOIR ARTICLE VIEVILLE-FAUGERAS). Ces différents capteurs lui permettent de naviguer à l'estime, à la manière d'un marin qui mesure ses déplacements en intégrant son cap et sa vitesse dans le temps, sans mémoriser et se référer à la moindre information relative à l'environnement.

Ces méthodes de localisation à l'estime sont cependant trop imprécises à long terme, car elles estiment la position du robot par intégration de déplacements élémentaires, ce qui a pour effet d'accroître l'erreur sur l'estimation (on dit qu'elles dérivent). Il faut donc leur adjoindre des algorithmes qui utilisent la perception d'éléments de l'environnement, appelés *amers* par analogie à la navigation maritime, pour corriger la dérive (on parle de recalage). Ces algorithmes, dits de cartographie et localisation simultanées, raisonnent explicitement sur les imprécisions des capteurs, et font appel à l'utilisation de théories de l'estimation, qui permettent de déterminer la valeur la plus probable d'une grandeur physique à partir d'un ensemble de mesures bruitées (voir figure 4). Ils nécessitent la construction et la mise à jour d'une description des amers de l'environnement, qui constituent donc une représentation particulière dédiée à la localisation. La mise en correspondance entre des amers mémorisés et des amers perçus, indispensable à ces algorithmes exige un raisonnement probabiliste qui n'est pas toujours simple. En effet, si la seule connaissance de leur position permet d'associer les amers perçus depuis deux positions proches du robot, à l'issue d'un long déplacement l'imprécision sur la position du robot peut devenir grande au point que les positions des amers ne permettent plus de les associer. Il est alors nécessaire d'identifier et de reconnaître les amers par d'autres moyens que leur position. Leur représentation doit alors être suffisamment

riche, les amers n'étant plus seulement définis par leur position, mais aussi par des attributs qui permettent de les identifier à l'aide de techniques de reconnaissance d'objets, ou encore par la configuration topologie qui les relie entre eux.

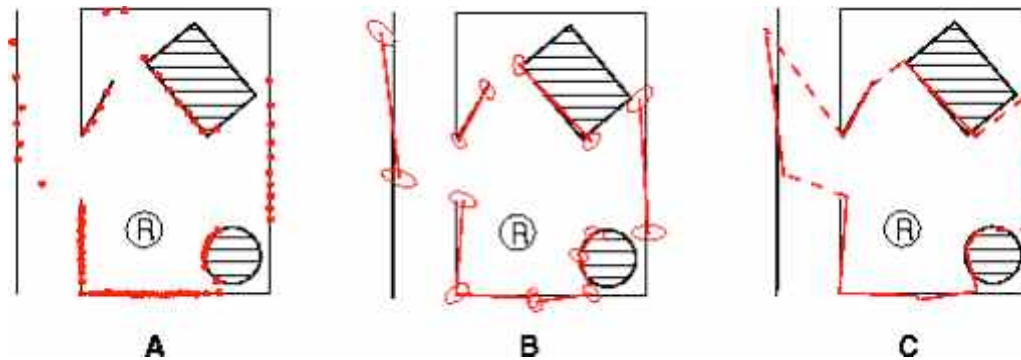
À long terme, l'intégration et la gestion de très nombreux amers nécessite des calculs lourds, et ne garantit pas toujours que le modèle construit soit spatialement cohérent. Pour pallier ces difficultés, des méthodes de localisation plus qualitatives, moins basées sur la géométrie mais plus robustes à long terme, ont été récemment proposées. L'utilisation de techniques de reconnaissance d'images panoramiques, qui permet d'identifier des lieux préalablement traversés, en est un exemple. D'autres approches exploitent aussi la représentation topologique de l'environnement pour se localiser à long terme : on ne cherche plus à estimer précisément la position d'amers, mais on identifie un lieu de manière qualitative, en vérifiant que les relations spatiales entre les objets qui y sont présents correspondent à un endroit déjà modélisé.

Enfin, un des moyens de garantir la cohérence spatiale des représentations construites et la qualité de la position estimée du robot est d'exploiter une carte initialement fournie au robot. Une telle carte est le plus souvent une représentation exogène au robot, elle a été déterminée par d'autres moyens que ceux dont il dispose pour percevoir le monde (images satellites ou plan d'architecte d'un bâtiment par exemple). La mise en correspondance ce qu'il perçoit effectivement avec les éléments contenus dans cette carte existante nécessite donc la construction nouvelle carte, représentation spécifique qui intègre des informations géométriques, topologiques et sémantiques d'un niveau compatible avec celles qui sont présentes dans la carte initiale.

Les méthodes brièvement présentées ici reflètent la diversité des approches qu'ont développées les chercheurs pour résoudre les problèmes de la modélisation d'environnement et de la localisation. Il est intéressant de constater que pour ces différents développements, les roboticiens ne se sont pas contentés d'appliquer différents outils théoriques connus. En effet, de même que pour la planification de mouvements et de tâches ou pour l'automatique, la considération des problématiques de la cartographie et de la localisation dans le contexte de la robotique suscite des extensions aux formalismes existants. Les travaux dans ces domaines sont actuellement très nombreux et fertiles ; ils sont notamment intimement liés à l'évolution des technologies des capteurs et des calculateurs, évolutions qu'il faut en permanence anticiper.

Aucune des approches présentées ici ne s'impose comme la solution idéale, toutes ayant des limites et des imperfections. En fait, après près de deux décennies d'efforts, il paraît évident que c'est la *multiplicité des approches* qui permet de garantir au robot un fonctionnement cohérent et robuste. Se pose alors le problème de *l'intégration* de techniques reposant sur des formalismes variés, problème qui est au centre même de la robotique. Enfin, il ne faut pas considérer la cartographie et la localisation d'un robot comme deux fonctionnalités passives, dont le résultat est simplement exploité par un niveau décisionnel supérieur qui décide des actions à exécuter. Elles constituent des processus actifs, qu'il faut contrôler, et dont le fonctionnement va influencer l'ensemble des décisions prises par le robot. Trouver la réponse à la question « où dois-je aller ? » pour réaliser une mission donnée n'implique pas seulement de répondre de la question « qu'ai-je vu ? », mais aussi aux questions « que dois-je voir ? », « comment me déplacer sans me perdre ? ». Les fonctionnalités de perception sont des actions qui permettent la prise de décision, mais qui sont aussi le résultat de prises de décision. Ceci

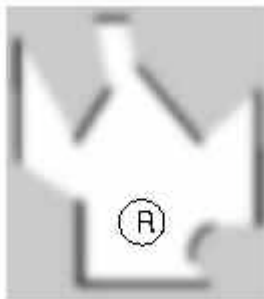
illustre à nouveau combien la discipline de la robotique repose sur une intégration étroite entre des fonctions de natures très variées.



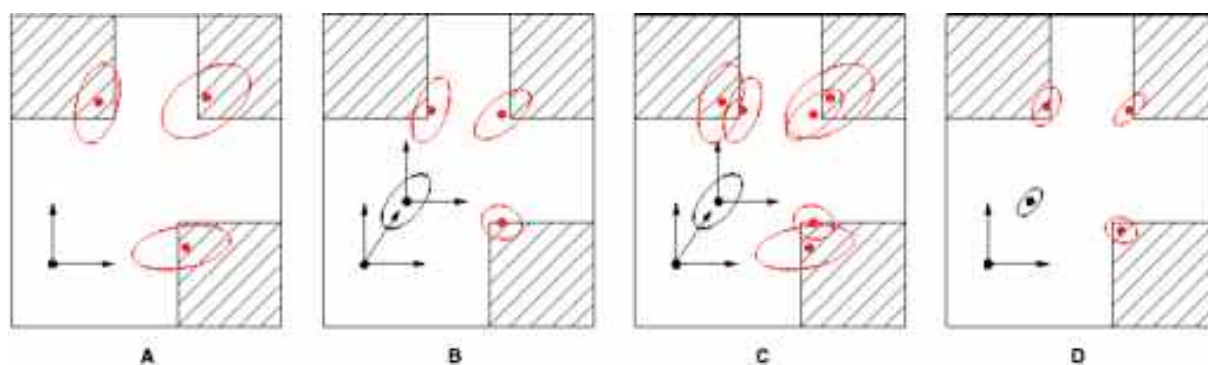
**Figure 1** : Illustration du processus de modélisation géométrique en deux dimensions. Le robot (R) est ici dans une pièce représentée par un plan d'architecte. En (A), les distances aux obstacles qu'il perçoit à l'aide d'un télémètre laser monté sur un système rotatif sont représentées par des points : toutes les données sont imprécises, certaines sont erronées. En (B), un processus de structuration des données permet de les représenter sous une forme plus compacte par des segments de droite. Notons que les imprécisions des données ont conduit à définir des « segments imprécis », d'où les ellipses aux extrémités des segments qui représentent la zone où elles ont une probabilité non nulle de se trouver. Notons aussi qu'à cause des imprécisions des mesures, certains éléments linéaires uniques de l'environnement sont représentés par plusieurs segments, et que bien sûr les segments ne sont pas des primitives adaptées à la représentation du cylindre. En (C), la représentation polygonale déduite des segments permet d'explicitier les zones accessibles par le robot. Les lignes pointillées qui ne correspondent pas à des segments réels, délimitent la zone perçue des zones occultées : elles peuvent par exemple être utilisées pour déterminer où le robot peut se placer pour acquérir de nouvelles informations. Des modèles géométriques tels que celui-ci sont depuis près de dix ans couramment utilisés par les robots qui se déplacent dans les laboratoires, mais ils restent pauvres, dans le sens où le robot peut difficilement les exploiter pour en extraire des informations sémantiques.



**Figure 2 :** Environnement d'intérieur ou d'extérieur ? Si dans un environnement structuré des primitives géométriques peuvent représenter la réalité, elles ne sont pas adaptées à un environnement naturel, où l'on ne rencontre qu'exceptionnellement des lignes droites ou des volumes représentables par des quadriques. Ici, on voit comment des primitives cylindriques ont pu être extraites d'un maillage en facettes triangulaires d'une scène d'un environnement industriel, lui-même déduit d'un ensemble de points 3D. Dans un environnement extérieur, la géométrie est plutôt représentée par un modèle numérique de terrain, qui décrit l'environnement par un ensemble d'altitudes calculées sur les points d'une grille horizontale régulière.



**Figure 3 :** Représentation de l'environnement par "grille d'occupation". À partir des mêmes données que dans la figure 1, des cellules élémentaires de l'environnement (les pixels de la figure) sont étiquetées par une probabilité de correspondre à un obstacle, ici représentée en niveaux de gris, le plus foncé représentant les probabilités les plus grandes. L'aspect "flou" de la représentation est dû aux imprécisions sur les données d'entrée.



**Figure 4 :** Principe de cartographie et localisation simultanée. À sa position initiale, le robot perçoit des amers dans l'environnement : il s'agit ici de coins, perçus avec une imprécision représentée en (A) par des ellipses rouges. Après un déplacement, le robot perçoit à nouveau ces amers : leur incertitude dans le nouveau repère lié au robot est représentée en (B) par des ellipses rouges, qui sont plus petites qu'en (A), le robot s'étant ici rapproché des amers. Mais la position du nouveau repère lié au robot est elle-même connue avec une incertitude, représentée par une ellipse noire. Si le robot est capable d'associer entre eux les amers perçus depuis les deux positions (ce qui peut nécessiter de les reconnaître), il peut alors agréger les deux perceptions (C), et affiner l'estimée de la position des amers et de sa position (D) grâce à l'utilisation d'un outil d'une théorie d'estimation. Le résultat de cette méthode à partir de données réelles acquises dans un bâtiment est illustré par les trois figures du bas. Le robot a effectué la trajectoire représentée en rouge sur le plan d'architecte de la figure de gauche. Au centre, sont représentées l'ensemble des données perçues, simplement agrégées en considérant la position du robot obtenue par odométrie : cette position étant connue de moins en moins précisément au fur et à mesure des déplacements, l'agrégation de toutes les données perçues est incompréhensible. À droite, à partir des mêmes données mais en recalant le robot au fur et à mesure de ses déplacements, le modèle du monde n'est pratiquement plus distordu, et on peut en extraire une topologie fidèle à la réalité.