

Photo ECAL / Milo Keller, Design : ECAL / Singal Mösch et Francis Chabloz



Comment les robots construisent leur monde

par Frédéric Kaplan et Pierre-Yves Oudeyer

Pourrait-on apprendre à un robot le nom des objets qui l'entourent ? Ne suffirait-il pas de le programmer pour qu'il associe les caractéristiques des images qu'il perçoit avec des séquences sonores particulières ? La forme arrondie d'une balle avec le son du mot « balle ». Une gamme de couleurs particulière associée au mot « rouge ». Peut-être qu'en accumulant les exemples, le robot finirait par prédire les contextes dans lesquelles chacun de ces mots s'applique ? Est-ce aussi simple que cela d'apprendre à parler ?

Telles étaient les questions que nous posions en 1999 avant de débiter une longue série d'expériences destinées à explorer la manière dont une machine pouvait apprendre ses premiers mots. Tous les matins, nous prenions une demi-heure pour montrer au robot des objets de couleur et de formes différentes. Nous changions d'environnement chaque jour (pièces très éclairées, zone sombres) de manière à ce que le robot puisse percevoir chaque objet dans les conditions les plus variées possible.

Nous avons utilisé un système d'apprentissage simple et efficace : l'apprentissage par

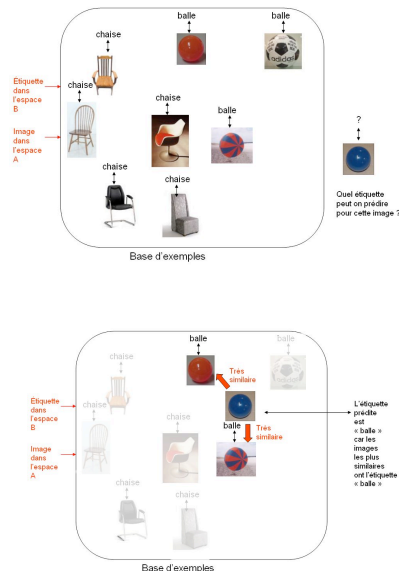
prototypes. Pour prédire le mot associé à une nouvelle situation, le robot en comparait les caractéristiques à celles rencontrées dans

le passé et choisissait le mot qui avait été utilisé dans la situation la plus similaire (voir encadré 1). Le robot a pu par cette méthode maîtriser l'usage d'une dizaine de mots associés à des objets facilement distinguables par leur couleur et leur forme. Les performances obtenues dans ces conditions étaient satisfaisantes (entre 70% et 80% de réussite) mais elles devenaient rapidement médiocres sitôt que nous tentions d'introduire des objets supplémentaires. L'approche n'était pas généralisable.

Une analyse plus fine de la base d'exemples qui contenait toutes les informations (images et sons) que le robot avait accumulées durant ces expériences fut riche d'enseignements. Dans un certain nombre de ces exemples,



1 / Comment un robot peut-il apprendre?



Tout un ensemble de techniques permettent aujourd'hui à une machine d'apprendre à faire des prédictions à partir d'exemples. Par exemple, on peut donner quelques images de balles et de chaises associées aux étiquettes « balle » et « chaise », et la machine peut ensuite deviner, quand on lui donne une nouvelle image qu'elle n'a jamais vue, si c'est une balle ou un chaise. La tâche de prédiction peut être aussi temporelle : par exemple, la machine observe un certain nombre de trajectoires d'une balle qu'on lance de diverses manières, et ensuite est capable de prédire où cette balle va atterrir dès les premières millisecondes d'un nouveau lancé. Il est possible de résumer ces différents contextes de prédiction en disant que la machine essaie d'apprendre à chaque fois à faire correspondre les points d'un espace A (e.g. images ou début de trajectoire) à des points dans un espace B (e.g. étiquettes ou fin de trajectoire), à partir d'un nombre limité d'exemples. Cette formulation correspond à ce qu'on appelle en mathématiques le problème de la régression, c'est-à-dire l'essai de reconstruction d'une fonction dans son ensemble à partir de quelques points. L'idée est souvent de présupposer certaines propriétés de cette fonction à reconstruire, par exemple qu'elle peut être

représentée par un polynôme où une combinaison linéaire de tangentes, et de chercher dans ce sous-ensemble paramétrable de fonctions candidates quelles sont celles qui approximent le mieux les points donnés en exemple. Il existe tout un arsenal d'outils pour représenter ces fonctions candidates : réseaux de neurones, réseaux bayésiens, chaînes de markov,

Il existe également une approche alternative pour résoudre le problème de l'apprentissage, que l'on appelle non-paramétrique, et qui bien que plus simple se montre souvent tout aussi efficace : c'est l'apprentissage par prototypes. Il s'agit dans ce cas pour la machine de garder en mémoire les exemples qu'elle a obtenu, chaque exemple étant une association entre un point de l'espace A et un point de l'espace B. Quand elle a une prédiction à faire, étant donnée par exemple une image a , alors la machine utilise une mesure de similarité pour calculer quels sont les images de sa base d'exemples qui sont les plus proches. La prédiction fournie est donc dans ce cas une étiquette qui est calculée en combinant les étiquettes des exemples les plus proches (par exemple par un vote à la majorité).

Quelle que soit la méthode d'apprentissage utilisée, elle impose toujours des contraintes, ou biais d'apprentissage, sur la forme de la fonction reconstruite : une méthode donnée ne peut jamais approximer toutes les fonctions avec précision et efficacité. Autrement dit, pour une méthode donnée, certaines fonctions seront faciles à deviner, et d'autres difficiles voir impossibles. En outre, chaque méthode utilise des biais différents, donc ce qui est facile pour l'une peut être difficile pour l'autre et vice versa.

utilisés dans les situations d'erreurs, l'objet qui était le sujet de l'interaction n'apparaissait tout simplement pas dans le champ de vision de la machine. Il s'agissait par exemple de situations où quand nous disions le mot « balle », le robot regardait en fait ailleurs. Il mémorisait ainsi un grand nombre d'expériences dans lesquelles ce mot semblait être associé à la moquette, au tapis ou à une ombre sur le sol. Difficile d'apprendre dans ces conditions.

Ce qui limitait les performances du robot, ce qui l'empêchait d'apprendre à bien à parler, ce n'était pas tant les techniques d'apprentissage automatique (voir encadré 1). Si on donnait à l'algorithme d'apprentissage des données pré-formatées, c'est-à-dire qu'on lui donnait manuellement les bonnes images associées aux bons sons, il se débrouillait plutôt bien. Le facteur limitant résidait plutôt dans les difficultés à faire que le robot se concentre sur la même chose que l'expéri-

mentateur quand ce dernier prononçait un mot : l'attention de l'homme et de la machine n'était pas partagée ...

L'attention est cette capacité qu'ont la plupart des animaux de se concentrer sélectivement sur certains aspects d'une situation à l'exclusion des autres. Lorsque nous marchons, lorsque nous montons un escalier, lorsque nous escaladons une montagne, nous prêtons attention à certains aspects de notre monde perceptif, mais pas à d'autres. Le comportement attentionnel peut donc s'interpréter dans le contexte d'une action motivée, dirigée vers un but.

Dans notre exemple, rien ne poussait le robot à s'intéresser à la balle que nous lui présentions. Nous aurions pu bien sûr lui programmer un « détecteur de balle » et un comportement de recherche associé aux caractéristiques visuelles correspondantes. Dans ce cas il aurait été facile de lui apprendre un mot pour cet objet. Mais le problème se serait

reposé à l'identique pour un nouvel objet. Comment un robot pourrait-il apprendre à reconnaître les caractéristiques d'un objet qu'il ne connaîtrait pas à l'avance ? Comment ferions-nous pour diriger son attention vers cet objet ?

Imaginons un salon dans lequel se trouve une table, quatre chaises, un canapé, un tabouret, un secrétaire et une bibliothèque murale (voir les illustrations de Jakob von Uexküll, *Mondes animaux et monde humain*, Denoël, 1965). Au plafond il y a un luminaire. Quatre assiettes sont posées sur la table. Un chien dans ce salon voit les choses quelque peu différemment. Pour lui il y a les objets sur lesquels il peut monter pour se reposer : le canapé et les chaises, mais sans doute pas le tabouret. Ce dernier, comme le secrétaire et la bibliothèque, fait partie des obstacles, éléments qui ne sont associés à aucune action particulière si ce n'est celle de gêner les déplacements. Imaginons maintenant le même salon vu par une mouche. Pour cet insecte, seule la lumière et peut-être le reste de nourriture dans les assiettes constituent des éléments significatifs.

Ce que nous reconnaissons dans les objets, ce sont les actions potentielles qui leur sont associées. Pour la même raison développer de nouveaux savoir-faire revient à élargir sa perception. Le monde de certains insectes reste simple parce que leurs différents moyens d'actions n'évoluent pas. Au contraire, le monde du jeune enfant humain ne cesse de se complexifier au fur et à mesure qu'il apprend à interagir avec son environnement. En quelque mois à peine, l'enfant apprend à contrôler son corps, à manipuler des objets, à échanger avec ses proches pour devenir un être autonome capable d'interactions complexes tant physiques que sociales. Durant ces mois d'intenses activités, se mettent en place les capacités qui seront à la base de l'apprentissage des premiers mots au début de la seconde année.

Si attention, perception et savoir-faire sont si intimement liés, la solution pour conduire un robot à porter son attention sur de nouveaux objets ne serait-elle pas de l'amener à découvrir de nouveaux savoir-faire ? Serait-il

possible de doter le robot d'une forme de curiosité qui le pousserait à explorer des situations qu'il ne connaît pas encore ?

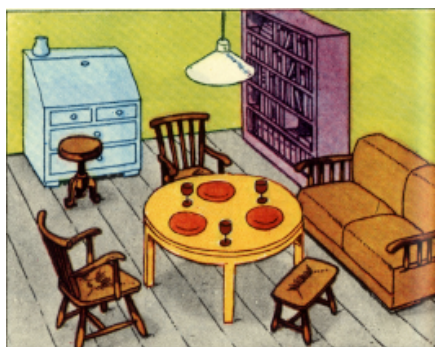
Avant tout chose, il convient de préciser ce que serait un « savoir-faire » pour un robot ? Un robot peut apprendre à anticiper les

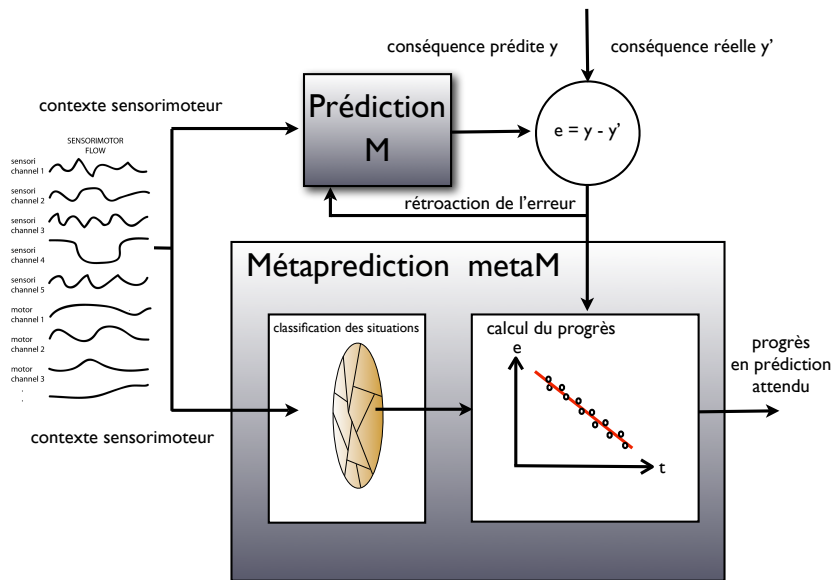
conséquences perceptives de ces actions. Le même type d'algorithme que ceux employés pour associer les caractéristiques d'une image à un son peut alors être utilisé. A partir de la mémorisation des expériences passées, la machine peut prédire les conséquences futures d'une action entreprise. Cette capacité d'anticipation est à la base des savoir faire potentiels qu'il pourra développer (voir encadré 1).

Quelle action le robot choisira-t-il ? La plupart des robots sélectionnent les actions qu'ils vont effectuer à partir d'un système de valeurs défini par ses programmeurs. Ce système de valeurs définit ce que la machine cherche à obtenir et ce qu'elle cherche à éviter. Il peut par exemple s'agir de rechercher des stimulations tactiles de la part de son propriétaire ou tout simplement des signes de sa présence (son de la voix, détection de son visage). A l'inverse des chocs trop

importants peuvent être associés à des valeurs négatives. Ces principes déterminent tout le comportement de la machine, et sont souvent implémentés sous la forme d'un système de récompenses positives ou négatives qu'elle va tenter d'optimiser. Typiquement, celle-ci va au départ essayer des actions ou des combinaisons d'actions au hasard, et progressivement par « essai-erreur » va sélectionner celles qui maximisent les récompenses positives et minimisent les récompenses négatives. Les actions sélectionnées peuvent être alors aléatoirement modifiées ou combinées pour former de nouvelles stratégies qui sont à leur tour évaluées par le robot. Ainsi, au fur et à mesure de ses expériences, chaque robot peut construire des stratégies propres pour satisfaire ces besoins artificiels définis par son créateur.

Dans la plupart des expériences de robotique actuelle, ces situations recherchées sont toujours « extérieures » au robot. Dans ces con-



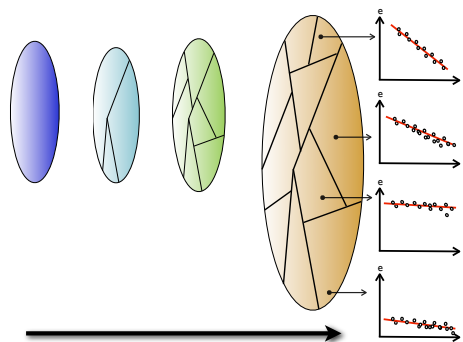


Le système de motivation est composé de deux modules: un prédicteur et un métaprédicteur

ditions, une fois les objectifs atteints (rester près des humains, ne pas de cogner contre les murs) le robot n'a pas de raison de continuer à évoluer. C'est pour tenter de palier à cette limite que nous avons commencé à réfléchir à la manière de doter le robot d'un système de motivation « intrinsèque ». L'idée serait d'équiper le robot d'un système de valeurs qui ne serait pas lié à des tâches spécifiques, définies à l'avance, mais qui le pousserait vers des « situations d'apprentissage » : une forme de curiosité. Ce système le conduirait à explorer les opportunités de son environnement et à découvrir des situations lui permettant de développer de nouvelles compétences.

De façon à doter le robot d'une forme de curiosité, nous avons associé à ces prédicteurs un système de valeurs qui l'incite à éviter les situations trop familières et les situations trop difficiles à prédire, pour privilégier les situations où le progrès en apprentissage est maximal. Il ne s'agit donc pas de choisir les situations pour lesquelles l'erreur en prédiction est minimale, ni celles pour lesquelles elle est maximale, mais celle pour lesquelles l'erreur diminue maximale. Il suffit pour cela de calculer la dérivée locale de la courbe d'erreur prédite par metaM et correspondant à diverses situations atteignables à partir d'un état sensorimoteur donné, et de choisir une action qui va conduire à la situation dont la dérivée correspondante est

Pour décrire le fonctionnement d'un tel système nous pouvons considérer qu'il est constitué de deux modules. Le premier module implémente un système de prédiction M qui apprend les conséquences perçues d'une action entreprise dans un contexte sensoriel et moteur donné. Le second système est un « métaprédicteur » metaM qui apprend à prédire les erreurs faites par le prédicteur M. En d'autres termes, ce second système modélise le premier et associe à chaque situations rencontrées un niveau de difficulté prédictive.

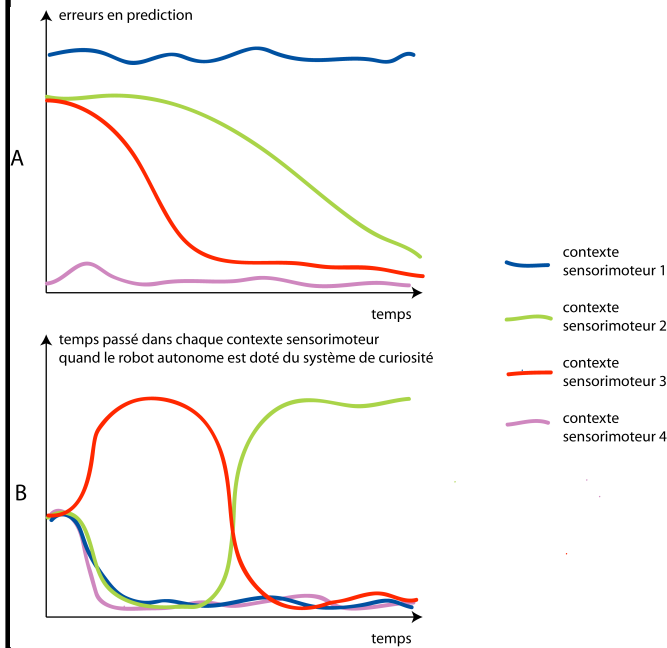


Pour calculer le progrès attendu dans chaque région, un algorithme découpe de manière incrémentale l'espace de situations possibles en groupe de situations similaires

la plus négative. Dans ce calcul, il est important de comparer l'erreur obtenue dans des situations similaires. Pour cela, le système de prédiction de metaM utilise un algorithme qui découpe de manière incrémentale l'espace des situations possibles en groupe de situations similaires. Pour chaque groupe il tient à jour un relevé de l'évolution des erreurs en prédiction, ce qui permet de leur attribuer une valeur.

En outre, de manière à ne pas rester bloquer dans des situations particulières alors que potentiellement d'autres situations intéressantes sont disponibles, des actions aléatoires sont choisies

2 / Comment une séquence de développement s'organise-t-elle ?



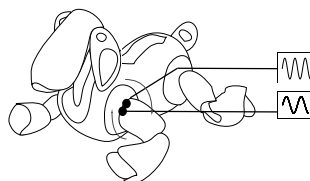
Nous pouvons illustrer le fonctionnement du système de curiosité artificielle sur un exemple abstrait. Imaginons un environnement dans lequel il existe quatre types d'activités sensorimotrices pour le robot, que l'on appelle contextes sensorimoteurs (e.g. shooter dans une balle, foncer et rebondir dans les murs, courir après le chat, dormir). Si l'on forçait le robot à se concentrer sur chacune de ces activités séparément, on pourrait mesurer l'évolution de son erreur en prédiction dans chacun de ces contextes. Les courbes correspondantes sont représentées dans la figure (A). On s'aperçoit qu'il y a une situation (1) dans laquelle l'erreur reste toujours élevée et ne diminue pas, possiblement parce que cette situation est trop compliquée pour son système d'apprentissage, une autre (4) dans laquelle elle est toujours basse et ne change pas, et deux situations (2 et 3) pour lesquelles l'erreur est haute au départ, mais diminue ensuite à des rythmes différents. En pratique, le robot est placé dans un environnement où ces différentes activi-

tés sont possibles, mais il ne connaît rien des courbes d'apprentissage correspondantes ni même le fait qu'il y a quatre types de contextes spécifiques. Dans ce cas, on doit observer d'abord, une phase d'exploration en grande partie aléatoire qui permet au robot d'une part de s'apercevoir qu'il y a des types de situations différentes et d'autre part de se faire une idée initiale de l'intérêt en terme d'apprentissage de chacune de ces situations. On doit alors observer le comportement illustré par le graphique (B). Le robot évite de se trouver dans les situations 1 et 4 car elle ne permettent pas de progrès en apprentissage. Il les explore cependant de temps en temps et par hasard, ce qui lui permet de vérifier qu'elles restent peu intéressantes. A l'inverse, il va se concentrer d'abord sur la situation 3 qui est celle pour laquelle ses prédictions s'améliorent le plus vite initialement. Après une première période, la situation 3 est maîtrisée et prédictible : le robot se met alors spontanément à explorer la situation 2 qui à ce stade de son développement lui procure le plus de progrès en apprentissage.

régulièrement. Typiquement, le robot choisit 70 pourcent du temps les situations qu'il juge lui apporter un progrès en apprentissage maximal et 30 pourcent du temps il effectue une action aléatoire. Cela lui permet de mettre à jour son évaluation des différentes situations, et notamment d'en découvrir de nouvelles correspondant à des diminutions maximales de l'erreur en prédiction. Ces situations sont appelées des « niches de progrès ». Les niches de progrès ne sont pas des propriétés intrinsèques de l'environnement. Elles résultent de la relation entre la structure physique du robot, les biais de ses mécanismes d'apprentissage, ses interactions passées, et l'environnement particulier dans lequel il est placé. Une fois découverte et exploitée, une niche de progrès disparaît au fur et à mesure que la situation à laquelle elle correspond devient plus prédictible. Ainsi, une trajectoire développementale, c'est-à-dire une séquence d'étapes dans lesquelles le robot se focalise

sur des activités de complexité croissante, se forme sans qu'elle ait été pré-programmées par le concepteur. Le fonctionnement d'un tel système est illustré sur un exemple abstrait dans l'encadré 2.

L'architecture que nous venons de décrire est générique dans la mesure où elle peut être appliquée à n'importe quel espace de senseurs et de moteurs qu'un robot peut explorer, et par définition n'est pas spécifique d'une tâche particulière. Considérons un premier exemple dans lequel le robot contrôle le mouvement de ces différents moteurs en appliquant des signaux sinusoïdaux. Pour chaque moteur, il fixe la période, la phase et l'amplitude du signal. Son système de prédiction tente de prédire les effets de ces différents jeux de paramètres sur la manière



Le robot contrôle ses mouvements grâce à des signaux sinusoïdaux.

dont l'image captée par sa caméra est modifiée, ce qui reflète indirectement le mouvement de son buste (pour lequel il n'a pas de capteur). A chaque itération de l'algorithme, le

robot choisit la valeur du prochain jeux de paramètres à essayer de façon à maximiser la réduction de l'erreur en prédiction, sauf dans 30 pourcent des cas où comme on l'a expliqué plus haut il choisit des paramètres aléatoires.

Lorsque nous débutons une expérience de ce genre, le robot explore au hasard différents jeux de paramètres pendant les premières minutes. Il agite ses bras et ses jambes de manière désordonnée. La majeure partie de ces mouvements a un effet très facilement prédictible : le robot ne bouge quasiment pas. Le robot malgré son agitation reste immobile. L'erreur en prédiction reste minimale : ces situations ne sont pas intéressantes pour le robot. Par hasard, au bout d'une dizaine de minutes un mouvement amène en général le robot à effectuer un léger déplacement. Il découvre par exemple qu'une certaine combinaison de paramètres résulte en un léger recul. Cette situation nouvelle résulte d'abord en une augmentation de l'erreur en prédiction puis, au fur et à mesure que le robot a de nouvelles occasions d'effectuer des mouvements similaires, cette erreur commence à baisser : le robot a découvert une « niche de progrès ».

Dans l'heure qui suit, le robot va explorer les différentes manières de reculer. Au cours de cette exploration, il est probable qu'il découvre que certaines légères modifications de paramètres conduisent à effectuer ce qu'un observateur extérieur appellerait des mouvements de rotation : un nouvel ensemble de « niches de progrès » que le robot pourra exploiter quand les compétences liées à la

marche en arrière auront été pour l'essentiel maîtrisées.

Il faut en général plus de trois heures pour que le robot découvre plusieurs ensembles de paramètres lui permettant de marcher en avant, en arrière, latéralement ou de tourner sur lui-même. A aucun moment, il n'a comme objectif d'apprendre à marcher. Guidé par la maximisation de la réduction de l'erreur en prédiction, il développe pourtant des compétences versatiles pour la locomotion. C'est d'ailleurs le caractère non spécifique de l'architecture qui permet cette versatilité. Un robot motivé pour se rapprocher d'un objet n'aurait par exemple sans doute pas appris à reculer ou à tourner sur lui-même.

Le fait que la marche en arrière se soit révélée dans cette trajectoire plus facile à découvrir que les autres n'était pas facile à prévoir. Etant donné la structure physique de ce robot et le type de sol sur lequel il était placé, les mouvements de recul ont été la première niche découverte. Pour savoir si cette niche constitue un « attracteur » récurrent pour ce type de trajectoire il faut mettre en place un programme d'expériences systématique semblable à celui que nous allons présenter dans le cadre de l'expérience qui suit.

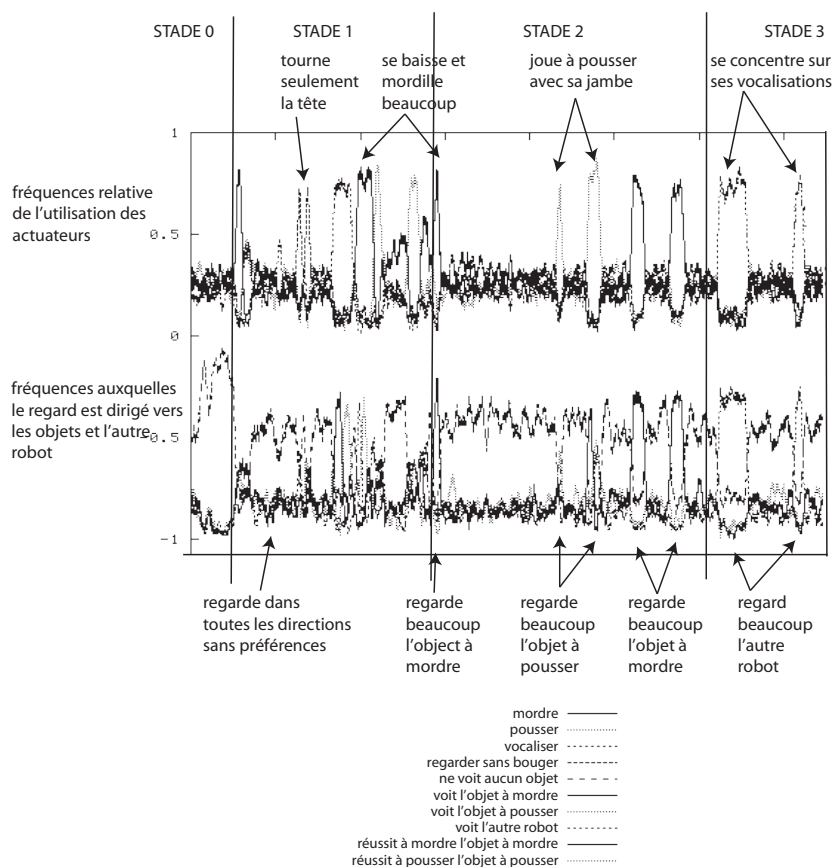
Dans une seconde série d'expériences, le robot est placé dans un environnement où plusieurs types d'objets sont présents. Un objet peut être mordu. Un objet peut-être poussé. Il y a également un autre robot qui imite les sons qu'émet le premier robot quand celui-ci les produit dans sa direction. Le robot



Le robot est placé sur un tapis d'éveil qui propose diverse opportunités d'apprentissage

objet. La découverte de ces associations a lieu quelques heures plus tard.

La production sonore, explorée dans un premier temps au même titre que les mouvements du corps, est rapidement abandonnée au profit de l'interaction avec les objets qui donne des résultats plus immédiats en terme de réduction d'erreur. En effet, lors des interactions sonores, le son émis par le robot est déformé lors de l'imitation par l'autre robot. Cet effet est déterministe mais plus difficile à prédire, du moins dans un premier temps. Ce n'est qu'une fois l'interaction avec les objets maîtrisées que le robot commence à rémettre des sons. Il se consacre alors presque exclusivement à cette modalité.



Le développement du robot s'auto-organise en plusieurs phases

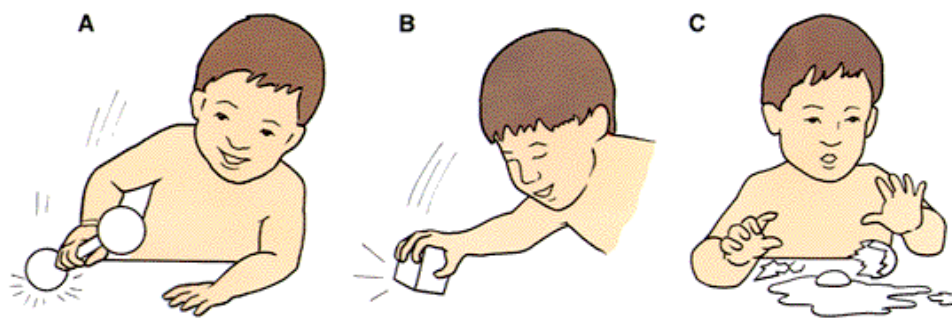
contrôle les paramètres d'un certain nombre de primitives motrices qui lui permettent un large éventail de mouvements. Il perçoit les conséquences de ces mouvements visuellement, de manière sonore et en utilisant ses capteurs de distance.

Comme pour la découverte de la marche, le robot explore cet espace avec comme unique objectif de maximiser la réduction de l'erreur en prédiction. Chaque expérience dure ici une dizaine d'heures et se caractérise par une structure remarquable. Après une première phase d'exploration aléatoire, le robot commence à effectuer de manière systématique des séries d'actions répétées (frapper, morde), puis dans un second temps il découvre les zones de l'environnement (objets frappables, objets mordable) qui correspondent à des progrès en apprentissage. A ce stade cependant, le robot ne fait pas d'association systématique entre une zone contenant un objet et l'action qui lui correspond. Il frappe l'objet à mordre, mord l'objet à frapper. Du point de vue de l'apprentissage il est plus intéressant d'associer la bonne action au bon

Chaque expérience donne lieu à une trajectoire unique. Mais un même type de structuration se retrouve dans la majorité d'entre elles : le robot explore d'abord les tâches les plus simples pour se concentrer progressivement sur les situations les plus difficiles en termes de prédiction. Une trajectoire particulière n'est pas entièrement déterminée par les algorithmes qui contrôlent le robot. Elle n'est pas non plus la conséquence directe des opportunités présentes dans l'environnement. Elle résulte de l'interaction entre un système d'apprentissage générique, un corps robotique particulier et un environnement structuré. Grâce à cette approche, nous avons la possibilité d'étudier le rôle respectif de chacun de ces facteurs structurant dans l'organisation d'une séquence développementale.

Ces deux expériences montrent comment des dynamiques génériques d'apprentissage peuvent conduire un robot à découvrir la maîtrise de son corps, les objets présents dans son environnement proche et à s'engager vers de premières formes d'interactions réciproques. Les compétences pour

3 / L'action structurante



L'organisation des séquences développementales que nous avons décrites présente une similarité intéressante avec la formation des compétences sensorimotrices chez le jeune enfant. Prenons l'exemple canonique de l'enfant qui fait la découverte du « schéma » sensorimoteur correspondant à l'action de frapper un objet sur une table. La plupart des objets solides font un bruit lorsqu'ils touchent la table. L'enfant peut explorer la similarité et la variabilité des réponses de ces objets « frappables ». Il découvre dans cette exploration qu'un certain nombre d'objets (par exemple un œuf) se cassent lorsqu'on les frappe de la même manière. L'action de frapper est ainsi à la base de la distinction entre les objets « frappables » et les objets « cassables ». De façon similaire, le robot de notre expérience catégorise ces expériences en fonction des réponses sensorielles provoquées par certaines actions dans certains contextes. (illustration www.psych.ubc.ca)

percevoir et agir que le robot développe de cette façon sont adaptées à sa morphologie et aux environnements auxquels il a été confronté. C'est le robot qui choisit vers quoi il tourne son attention en fonction de ses expériences passées. Il est acteur de son propre développement et construit ainsi une structuration de son environnement qui lui est propre.

Mais cette structuration que le robot développe de manière autonome n'a que peu de chance de correspondre à ce que nous considérons d'ordinaire comme étant les objets de notre monde. En particulier, il est fortement improbable que le robot ait pu développer seul les concepts que nous voulions initialement lui apprendre, comme celui d'une balle, et qu'il suffise simplement maintenant de les « étiqueter » avec des mots de notre langue. Pour le robot, il y a deux structurations indépendantes : celle qu'il a développée seul et celle qui peut lui être proposée de l'extérieur sous la forme de mots de la langue. Des passerelles sont peut-être possibles entre ces deux mondes. Au fil des interactions, le mot « balle » pourrait ainsi être associé à plusieurs structures perceptives différentes. Le robot pourra alors chercher ce que ces structures ont en commun. Mais il est également envisageable que les différentes morphologies entre les robots et les hommes limitent de toute façon la possibilité que des machines développent certains concepts importants de nos langues. « Un lion pourrait parler, nous ne pourrions le comprendre » écrivait Wittgenstein. Entre le monde des robots et celui le monde des hommes, il y aura peut-être des passerelles

linguistiques. Mais leur nombre et leur largeur restent encore inconnus.

Ce qu'il nous faut retenir c'est le rôle de l'intervenant humain dans ces processus. Face à un robot qui est acteur de son propre développement, il ne peut qu'organiser les opportunités d'apprentissage que la machine est amenée à rencontrer. Il s'agit pour lui de créer des « niches de progrès » pour la machine. C'est ainsi qu'il peut dans une certaine mesure guider le développement de la machine. C'est en ce sens qu'il est un médiateur, un passeur du monde robotique vers celui des hommes.

La question du langage nous a conduit à la question de l'attention puis à la question du savoir-faire et de l'anticipation, puis à la question des motivations intrinsèques. Au fur et à mesure que les robots progressent ou se heurtent à des difficultés inattendues, notre compréhension des phénomènes qu'ils tentent d'approcher évolue. Ainsi il ne s'agit pas simplement de construire des machines « plus intelligentes », il s'agit aussi de mieux comprendre ce qu'« intelligent » veut dire en construisant des machines.

La démarche technologique qui nous a guidée suggère en effet des hypothèses potentiellement nouvelles sur la manière dont les enfants apprennent et en particulier sur le rôle du corps, des biais d'apprentissage et des motivations intrinsèques dans ce processus. Les recherches actuelles en neurosciences portent une attention toute particulière sur la manière dont le cerveau traite la nouveauté et les signaux d'erreur en prédiction. Les recherches en psychologie comparée

insistent sur l'importance possible de la motivation pour expliquer les différences d'apprentissage entre les grands singes et les hommes. La robotique peut jouer un rôle pertinent dans ce débat et depuis quelques années les initiatives multidisciplinaires se multiplient, à l'image de la série de conférence Epigenetic Robotics.

Dans cette approche, il ne s'agit pas d'imiter l'homme en tout point. C'est au contraire en étudiant le rôle structurant des contraintes physiques et algorithmiques de la machine sur ses propres trajectoires de développement, que l'on peut en retour construire des explications pertinentes sur l'influence de facteurs comparables dans le développement du jeune enfant. En permettant une démarche expérimentale d'un genre nouveau, la robotique ouvre la voie à une meilleure compréhension de l'articulation entre contraintes spécifiques et mécanismes généraux d'apprentissage. Elle invite ainsi à repenser différemment les paradigmes de l'inné et de l'acquis et à construire de nouvelles notions, un nouveau vocabulaire, pour éclairer l'extraordinaire capacité qu'ont les enfants à apprendre.

Pour en savoir plus

Oudeyer, P-Y and Kaplan, F. (2006) Discovering communication, **Connection Science**, 18 (2) : 189 - 206

Kaplan, F. (2005) Les machines apprivoisées, Vuibert, Paris, France

Oudeyer, P-Y (2006) Self-organization in the evolution of speech, Oxford University Press, Oxford, UK.

Pierre-Yves Oudeyer : www.csl.sony.fr/~py

Frederic Kaplan : www.fkaplan.com

Une version modifiée de cet article est apparue dans le numéro 348 du magazine **Pour La Science** (Octobre 2006)