

Un robot curieux

Un robot apprend à découvrir son environnement de façon autonome, grâce à un nouveau système de motivation : la curiosité.

Frédéric Kaplan et Pierre-Yves Oudeyer

Aujourd'hui, certains robots reconnaissent plusieurs milliers de mots. Pour autant, ils n'en comprennent pas le sens et sont incapables d'en apprendre de nouveaux. En général, les mots reconnus par le robot à partir de signaux sonores sont associés à des systèmes de commande définis à l'avance par les programmeurs, similaires à ceux qui lui permettent d'effectuer une danse ou d'aller chercher un objet. Depuis plusieurs années, nous tentons d'aller au-delà de cette limitation et nous étudions si un robot peut apprendre à reconnaître et à nommer les objets qui l'entourent sans qu'ils ne soient déjà inscrits dans sa « mémoire ». Peut-on utiliser un système d'apprentissage classique ? Ne suffirait-il pas de programmer le robot pour qu'il associe les caractéristiques des images observées avec les séquences sonores qu'il entend ? Il pourrait alors par exemple associer la forme arrondie d'une balle avec le son du mot *balle* ou une gamme de couleurs particulière avec le mot *rouge*. En accumulant les exemples, le robot réussirait peut-être à prédire les contextes dans lesquels chaque mot s'applique. Est-ce aussi simple d'apprendre à parler ? Nos recherches nous ont conduits à répondre à ces questions par la négative.

En effet, nous nous sommes heurtés à un obstacle : pour apprendre à parler, un robot doit porter son attention sur des objets qu'il ne connaît pas encore. Pourquoi s'intéresserait-il à ces objets ? Pourrait-on le doter d'une forme de curiosité ? Pour mener à bien nos recherches, nous nous sommes inspirés des travaux réalisés en psychologie, en pédagogie et en statistiques. Nous verrons que la curiosité du robot est liée à l'acquisition de nouvelles compétences, et nous préciserons comment nous avons inventé un robot « motivé » pour découvrir son environnement de façon autonome.

1. Le robot AIBO est utilisé depuis de nombreuses années comme un outil de recherche en intelligence artificielle. Celui des auteurs est aujourd'hui doté d'un système de curiosité qui lui permet de découvrir seul son environnement.





Nous avons utilisé le robot *AIBO* développé pendant de nombreuses années par la Société SONY. Ce robot est équipé d'un ordinateur, de nombreux capteurs et de moteurs qui actionnent sa tête, sa bouche, ses pattes, ses oreilles et sa queue (voir la figure 3). Son comportement est contrôlé par l'ordinateur relié à ces capteurs et actionneurs. En général, des algorithmes d'apprentissage permettent au robot d'apprendre à réaliser une tâche particulière définie à l'avance par l'expérimentateur, par exemple tourner sur lui-même. Le robot tente un mouvement, le résultat est comparé avec celui qu'il doit réaliser, et le mouvement est recalculé pour qu'il se rapproche au mieux de l'objectif. L'approche que nous allons décrire est fondamentalement différente : un robot curieux choisit lui-même ce qu'il explore, puis apprendra, et devient capable non plus de maîtriser une tâche, mais une succession de tâches de complexité croissante.

Des mots vides de sens

En 1999, nous avons commencé une série d'expériences pour explorer la façon dont une machine peut apprendre des mots associés aux objets qui l'entourent, en interagissant avec un être humain. Tous les matins, nous montrions au robot, pendant une demi-heure, des objets de couleur et de forme différentes. Nous changions d'environnement chaque jour (pièce éclairée ou sombre, par exemple) afin que le robot voie chaque objet dans différentes conditions. Pour ces expériences, nous utilisons un système d'apprentissage simple et efficace : l'apprentissage par prototypes (voir l'encadré de la page 36). Pour prédire le mot associé à un objet dans une nouvelle situation – une balle présentée dans l'obscurité par exemple –, le robot compare les caractéristiques visuelles de l'objet à celles rencontrées au préalable, et il choisissait le mot retenu dans la situation la plus similaire. Le robot a pu maîtriser une dizaine de mots correspondant à des objets facilement identifiables par leur couleur et leur forme. Les performances d'apprentissage étaient satisfaisantes (entre 70 et 80 pour cent de réussite), mais elles devenaient vite médiocres dès l'introduction d'objets supplémentaires. Ainsi, cette approche n'était pas généralisable. Examinons pourquoi.

En analysant les images et les sons des exemples accumulés par le robot durant ces expériences, nous avons remarqué que le robot ne portait pas son attention là où nous le souhaitions. Dans certains exemples, l'objet que nous lui montrions n'apparaissait même pas dans son champ de vision. Quand nous disions le mot *balle* en lui montrant une balle, sa caméra fixait autre chose : le robot mémorisait des expériences où ce mot était associé à la moquette, au tapis ou à une ombre sur le sol. Difficile d'apprendre dans

ces conditions. Or les techniques d'apprentissage automatique ne limitaient pas les performances du robot et ne l'empêchaient pas d'apprendre à parler : si nous utilisons l'algorithme d'apprentissage en lui donnant manuellement les bonnes images associées aux bons sons, le robot apprenait bien. Le problème résidait dans le fait que le robot ne voyait pas la même chose que nous quand nous prononcions un mot : il ne portait pas son attention là où nous souhaitons qu'il la porte. Lorsque nous marchons, montons un escalier ou escaladons une montagne, nous prêtons attention à des caractéristiques de notre environnement, mais pas à d'autres. L'attention est liée à une action dirigée vers un but : nous percevons ce qui est utile pour réaliser une tâche. Dans nos expériences, rien ne prédisposait le robot à s'intéresser à la balle que nous lui présentions. Nous aurions pu programmer un détecteur de balles et un comportement de recherche associé aux caractéristiques visuelles d'une balle : il aurait alors été facile de lui apprendre le mot *balle*. Cependant le même problème se serait posé pour l'apprentissage de chaque nouveau mot. Un robot peut-il apprendre à reconnaître les caractéristiques d'un objet qu'il ne connaît pas à l'avance ? Comment diriger son attention vers cet objet ? La solution consiste à le pousser à développer de nouvelles capacités.

Imaginons un salon où se trouvent une table, quatre chaises, un canapé, un tabouret, un secrétaire et une bibliothèque murale. Au plafond, il y a un luminaire. Quatre assiettes sont sur la table. Un chien dans ce salon voit les choses différemment. Il peut monter pour se reposer sur le canapé et les chaises, mais probablement pas sur le tabouret. Ce dernier, comme le secrétaire et la bibliothèque, sont des obstacles, c'est-à-dire des éléments qui ne correspondent à aucune action particulière si ce n'est celle de gêner les déplacements. Com-

ment une mouche voit-elle ce salon ? Pour cet insecte, seule la lumière et peut-être le reste de nourriture dans les assiettes sont des éléments significatifs. En d'autres termes, nous reconnaissons dans les objets les actions potentielles qui leur sont associées. De même, développer de nouveaux savoir-faire revient à élargir sa perception. Le monde de certains insectes reste simple parce que leurs différents moyens d'action n'évoluent pas. Au contraire, le monde du jeune enfant ne cesse de se complexifier à mesure qu'il apprend à agir dans son environnement. En quelques 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 physiques et sociales. L'enfant développe les capacités qui seront à la base de l'apprentissage des premiers mots au début de sa deuxième année. Attention, perception et savoir-faire étant si intimement liés, un robot portera-t-il son attention sur de nouveaux objets si on le conduit à découvrir de nouvelles compétences ? Pourrait-on doter le robot d'une forme de curiosité qui le pousserait à explorer des situations qu'il ne connaît pas encore ?

Être curieux

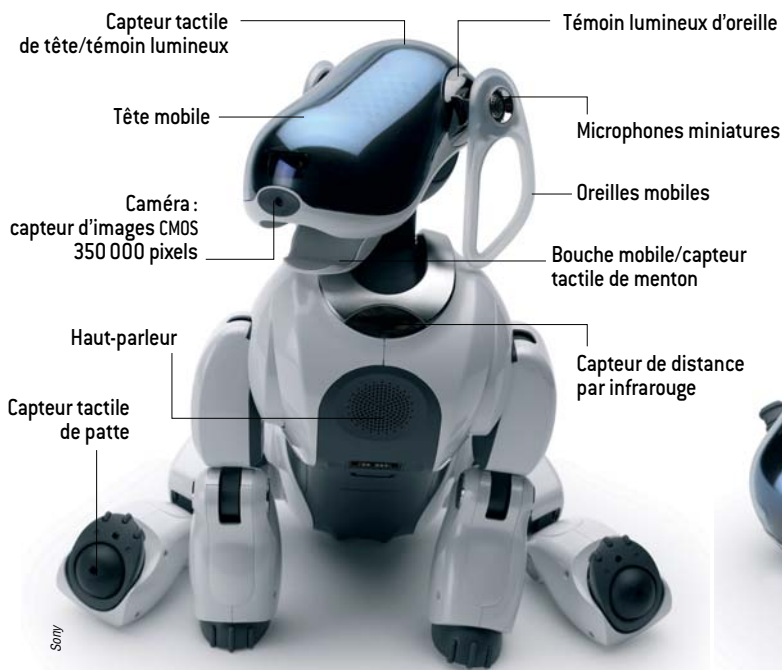
Précisons d'abord ce qu'est un « savoir-faire » pour un robot. Un robot peut apprendre à anticiper les conséquences perceptives (visuelles, sonores, tactiles, proprioceptives, etc.) de ses actions grâce au même type d'algorithmes que ceux employés pour associer les caractéristiques d'une image à un son. Quand la machine a mémorisé des expériences passées, elle prédit les conséquences futures d'un nouvel acte. La plupart des robots sélectionnent leurs actes à partir d'un système de valeurs défini par leurs programmeurs. Ce système de valeurs représente ce que la machine tente d'obtenir ou d'éviter. Par exemple, le robot recherchera des contacts physiques avec son propriétaire ou des signes de sa présence (son de la voix, détection de son visage) si ces stimulus sont associés à des valeurs positives. À l'inverse, des chocs brutaux pourront correspondre à des valeurs négatives. Ces valeurs déterminent le comportement de la machine : elles sont en permanence optimisées par le robot. Ainsi, au départ, le robot agit au hasard, et progressivement, par tâtonnements, il sélectionne les actions qui favorisent les valeurs positives et minimisent les valeurs négatives. Les actions retenues sont ensuite aléatoirement modifiées ou combinées pour élaborer de nouvelles stratégies, à leur tour évaluées par le robot. Au fil de ses expériences, chaque robot construit ses stratégies pour satisfaire les besoins définis par son créateur.

Dans beaucoup d'expériences actuelles de robotique, un robot doit, par exemple, rester près des humains ou ne pas se cogner contre les murs. Une fois ces objectifs atteints, il n'a pas de raison de continuer à évoluer. C'est pour surmonter cette limite que nous avons voulu doter notre robot d'un système de motivation « intrinsèque ». Il s'agit d'un système de valeurs qui n'est pas lié à des tâches spécifiques, définies à l'avance, mais qui pousse le robot vers des « situations d'apprentissage » : une forme de curiosité. Ce système peut conduire le robot à explorer les occasions offertes par son environnement et à découvrir des situations lui permettant de développer de nouvelles compétences.



F. Kaplan et P.-Y. Oudeyer

2. Pour découvrir son environnement et communiquer, le robot est placé sur un tapis de jeu pour bébé. Certains objets peuvent être frappés, d'autres mordus. Un autre robot qui imite les sons émis par le premier est aussi présent sur le tapis, ce qui donnera au premier l'occasion d'apprendre des formes rudimentaires de communication.



3. Le robot AIBO est un « ordinateur sur pattes ». Il contient un processeur RISC de 64 bits à 576 mégahertz, 64 mégaoctets de mémoire vive et il est doté de nombreux capteurs : des microphones, une caméra CMOS de 350 000 pixels, des capteurs tactiles ou électrostatiques sur les pattes, sur le dos, sur la tête et sous le menton, des capteurs de distance à infrarouges, trois capteurs de vitesse, de vibrations, de

température. Il est également pourvu d'un haut-parleur et d'une trentaine de diodes lumineuses. Sa tête, sa bouche, ses pattes, ses oreilles et sa queue sont motorisées. Il peut être relié à Internet ou à un ordinateur grâce à une antenne WIFI. Le comportement du robot est guidé par un système de commande et d'apprentissage, relié à la fois à ses capteurs et à ses actionneurs.

Les motivations intrinsèques et leur rôle dans l'apprentissage sont étudiés depuis longtemps en psychologie et en pédagogie. Mais les modèles élaborés dans ces disciplines ne sont pas décrits avec suffisamment de précision pour être transférables dans un algorithme. Ils ne servent que d'inspiration. En revanche, depuis une dizaine d'années, des travaux en statistiques sur la « création d'expériences optimales » – *Optimal Experiment Design* – et en apprentissage artificiel dans le domaine de l'apprentissage actif – *Active Learning* – se sont penchés sur des questions semblables à celles qui nous intéressent. Plusieurs approches mathématiques ont été proposées pour tenter de définir la meilleure façon d'explorer un espace inconnu afin d'en apprendre le maximum en un minimum de temps. Ces recherches ont amélioré les performances des algorithmes d'apprentissage. Toutefois, la plupart de ces méthodes ne sont pas directement applicables à des robots qui évoluent dans des espaces réels. Qui plus est, nous ne cherchons pas tant que nos robots apprennent le plus vite possible, mais plutôt qu'ils se développent et progressent plus longtemps. Ainsi, avant de décrire le comportement de nos nouveaux robots, examinons comment nous avons programmé une telle architecture.

Notre système de motivation intrinsèque comprend deux modules (voir la figure 4). Le premier est un système de prédiction classique (ou « prédicteur ») qui apprend les conséquences perceptives (visuelles, sonores, tactiles, etc.) $S(t+1)$ d'une action entreprise à un instant t dans un contexte sensoriel et moteur donné. Il calcule une erreur $-e_n(t)$ – qui est la différence entre la conséquence réelle de l'action $S(t+1)$ et la conséquence prédite $S'(t+1)$. Par exemple, imaginons que le robot prédise que, sous l'impulsion d'une commande motrice, sa patte gauche bouge de deux centimètres. L'ac-

tion effectuée, ses capteurs enregistrent en fait un mouvement de 2,2 centimètres : l'erreur est la différence entre ce qui était prédit et les conséquences réelles du geste. Le second système ou « métaprédicteur » apprend à prédire les erreurs faites par le prédicteur : il évalue « l'erreur en prédiction ». Dans notre exemple, il prédira que le système de prédiction ne se trompe que de quelques dixièmes de centimètres. Pour ce robot, l'action « bouger la patte » est simple à prédire. En revanche, pour une autre commande motrice, les conséquences pourraient être plus difficiles à prédire. En d'autres termes, ce système associe à chaque situation sensorielle et motrice rencontrée un niveau de difficulté en termes de prédiction.

Optimiser les progrès en apprentissage

Pour 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, d'une part, et celles trop difficiles à prédire, d'autre part : le robot privilégie les situations où le progrès en apprentissage est optimal. Le robot ne choisira pas les situations où l'erreur en prédiction est minimale, ni celles où elle est maximale, mais celles où l'erreur diminue le plus vite. En effet, le métaprédicteur calcule la dérivée locale $D_n(t)$ de la courbe d'erreurs, qui correspond à diverses situations que le robot peut atteindre à partir d'un état sensorimoteur donné. Le robot sélectionne alors l'action qui conduit à la situation dont la dérivée est la plus négative. Pour ce calcul, il compare l'erreur obtenue dans des situations similaires. Comment fait-il ? Le système de métaprédiction utilise un algorithme qui découpe

progressivement l'espace des situations possibles en groupes de situations similaires R_n , que l'on nomme « régions ». Pour chaque groupe R_n , il tient à jour un relevé de l'évolution des erreurs : le système garde en mémoire toutes les erreurs, du début jusqu'à l'instant t , ce qui permet de calculer comment l'erreur diminue en moyenne pour cette région. La valeur attribuée à cette région, encore nommée degré d'intérêt, représente le progrès en apprentissage, qui est égal à l'opposé de la dérivée locale : $L_n(t) = -D_n(t)$. Ainsi, la plupart du temps, dans un contexte sensorimoteur donné, le système d'apprentissage du robot choisit une action qui le mènera dans un contexte appartenant à la région où le degré d'intérêt est maximal (voir l'encadré de la page 38).

En outre, pour ne pas rester bloqué dans une situation particulière alors que d'autres situations intéressantes sont disponibles, le robot effectue régulièrement des actions aléatoires. En général, dans 70 pour cent des cas, il adopte les situations qui lui apportent un progrès en apprentissage maximal, et dans 30 pour cent des cas, il agit aléatoirement. Ainsi, le robot peut découvrir de nouvelles situations correspondant à des diminutions optimales de l'erreur en prédiction : on parle de « 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 caractéristiques de ses mécanismes d'apprentissage, ses expériences passées et son environnement. Après avoir été découverte et exploitée, une niche de progrès disparaît à mesure que la situation à laquelle elle correspond devient plus prédictible. Ainsi, une trajectoire de développement – c'est-à-dire une séquence d'étapes où le robot effectue des activités de complexité croissante – se forme sans qu'elle soit préprogrammée par le concepteur. Voyons à présent comment le robot se comporte si on le dote d'un tel système et, plus particulièrement, comment il peut apprendre à marcher simplement par curiosité.

L'architecture que nous venons de décrire n'est pas spécifique d'une tâche particulière et est applicable à n'importe quel ensemble de capteurs et de moteurs d'un robot. Dans



Sauf mention contraire, les photos du robot sont de ECOL / Milio Keller.

4. Le système de motivation intrinsèque comprend un système de prédiction (ou « prédicteur ») et un « métaprédicteur ». Le premier évalue les conséquences perceptives (visuelles, sonores, tactiles, etc.) d'une action entreprise dans un contexte sensoriel et moteur donné. Par exemple, le robot aperçoit un os. Ce contexte est analysé par le prédicteur

un premier temps, le robot contrôle le mouvement de ses différents moteurs en appliquant des signaux périodiques particuliers dits sinusoïdaux. Pour chaque moteur, il fixe la période, la phase et l'amplitude du signal. Son système de prédiction tente de prédire, en fonction de ces différents paramètres, la façon dont l'image captée par sa caméra est modifiée, ce qui reflète indirectement le mouvement du buste (où il n'y a pas de capteur). À chaque itération de l'algorithme, le système de commande du robot fixe les paramètres à tester pour

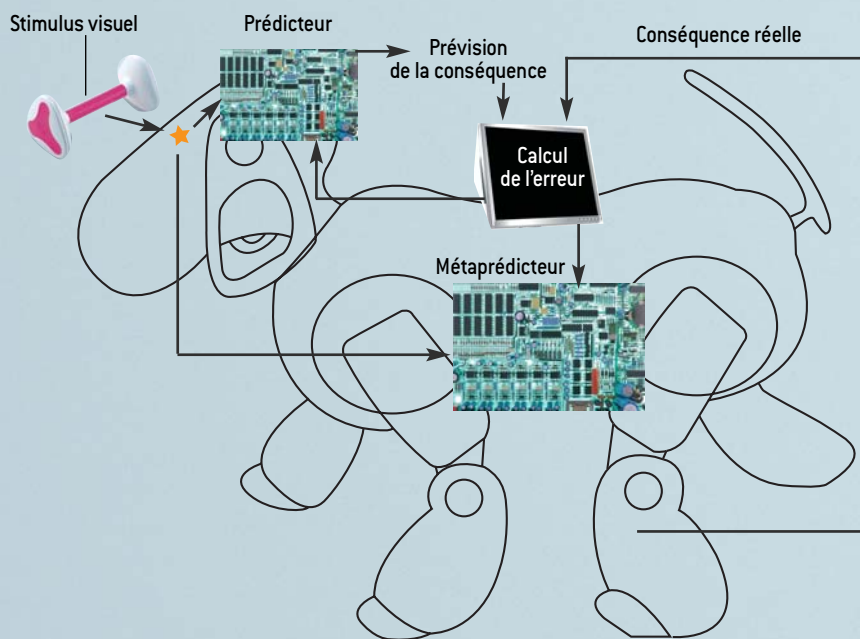
Comment un robot apprend-il ?

Aujourd'hui, une machine peut apprendre à faire des prédictions à partir d'exemples grâce à plusieurs techniques. Par exemple, on associe quelques images de balles et de chaises aux étiquettes « balle » et « chaise », et la machine devine ensuite, quand on lui présente une nouvelle image, si c'est une balle ou une chaise. La tâche de prédiction peut aussi être temporelle : par exemple, la machine observe un certain nombre de trajectoires d'une balle lancée de diverses manières, et prédit ensuite où cette balle va atterrir quelques millisecondes après un nouveau lancer. Ainsi, à partir d'un nombre limité d'exemples, la machine essaie de faire correspondre les points d'un espace A, par exemple des images ou le début d'une trajectoire, à des points d'un espace B, par exemple des étiquettes ou la fin de la trajectoire. En mathématiques, cette formulation correspond à un problème de régression, c'est-à-dire un essai de reconstruction d'une fonction dans son ensemble à partir de quelques points. Souvent, on présuppose certaines propriétés de la fonction à reconstruire (par exemple, elle est représentée par un poly-

nôme) et on cherche, dans un sous-ensemble de fonctions candidates, celles qui coïncident le mieux avec les points donnés.

Une autre approche, dite non paramétrique, est plus simple, mais tout aussi efficace : l'apprentissage par prototypes. Dans ce cas, la machine garde en mémoire les exemples qu'elle a obtenus ; chaque exemple est une association entre un point de l'espace A et un point de l'espace B. Quand elle a une prédiction à faire, par exemple donner le nom d'une image, elle recherche les images les plus proches dans sa base d'exemples. La prédiction fournie est donc une étiquette calculée en combinant les étiquettes des exemples les plus proches déjà mémorisés.

Toute méthode d'apprentissage impose des contraintes, ou biais d'apprentissage, sur la fonction reconstruite : elle ne peut jamais s'approcher de toutes les fonctions avec précision et efficacité. Autrement dit, pour une méthode donnée, certaines fonctions seront faciles à deviner et d'autres difficiles, voire impossibles. En outre, chaque méthode a ses propres biais : ce qui est facile pour l'une peut être difficile pour l'autre et *vice versa*.



qui évalue la conséquence d'un mouvement possible à entreprendre : bouger une patte pour se rapprocher de l'os. La conséquence réelle du mouvement est comparée à la conséquence calculée par le prédicteur. Selon l'erreur obtenue, le robot essaie un nouveau mouvement de la patte. Au fil des boucles d'itération, le métaprédicteur évalue à quel point ses

erreurs en prédiction dans ce contexte sensorimoteur diminuent : cela définit le degré d'intérêt du robot associé à cette activité. Si l'erreur diminue beaucoup, alors il s'y intéressera et persévéra (par exemple, il pourra apprendre à manipuler l'os de différentes façons). Si elle stagne ou même augmente, il passera à une autre activité.

optimiser la réduction de l'erreur en prédiction, sauf dans 30 pour cent des cas où les paramètres sont choisis aléatoirement.

Examinons comment le robot qui n'a pas encore « appris » à marcher se comporte quand on le pose sur le sol. Il explore au hasard différents paramètres pendant les premières minutes. Il agite ses pattes de façon désordonnée. L'erreur en prédiction reste minimale : le robot ne bouge presque pas et ses mouvements sont prédictibles. Ces situations ne sont pas intéressantes pour le robot, qui est programmé pour trouver une situation où l'erreur en prédiction est d'abord élevée, puis diminue. Mais, au bout d'une dizaine de minutes, une combinaison aléatoire des paramètres le conduit à effectuer un léger déplacement : cette combinaison le fait reculer légèrement. Cette nouvelle situation se caractérise par une augmentation de l'erreur en prédiction puis, à mesure que le robot a de nouvelles occasions d'effectuer ces mouvements, l'erreur commence à baisser : une niche de progrès a été découverte.

Dans l'heure qui suit, le robot explore les différentes façons de reculer. Mais certaines modifications de paramètres lui font découvrir des mouvements de rotation : c'est un nouvel ensemble de niches de progrès qui seront examinées quand les compétences liées à la marche en arrière seront maîtrisées. En général, plus de trois heures sont nécessaires pour que le robot découvre la marche vers l'arrière, en rotation, vers l'avant, sur le côté... Mais il n'est pas programmé pour apprendre à marcher. Il optimise seulement la réduction de son erreur en prédiction, et il développe pourtant des compétences polyvalentes pour la locomotion. Le caractère non spécifique de son architecture est la cause de cette variété : un robot programmé pour se diriger vers un objet n'aurait pas appris à reculer ou à tourner sur lui-même.

Nous ne pouvons pas prévoir que la marche en arrière serait plus facile à découvrir que les autres marches dans cette expérience. Étant donné la structure physique du robot et le type de sol sur lequel il était placé, les mouvements de recul ont été la première niche explorée. Pour savoir si cette niche est régulièrement découverte par les robots, il faut mettre en place un programme d'expériences systématiques semblable à celui que nous allons examiner.

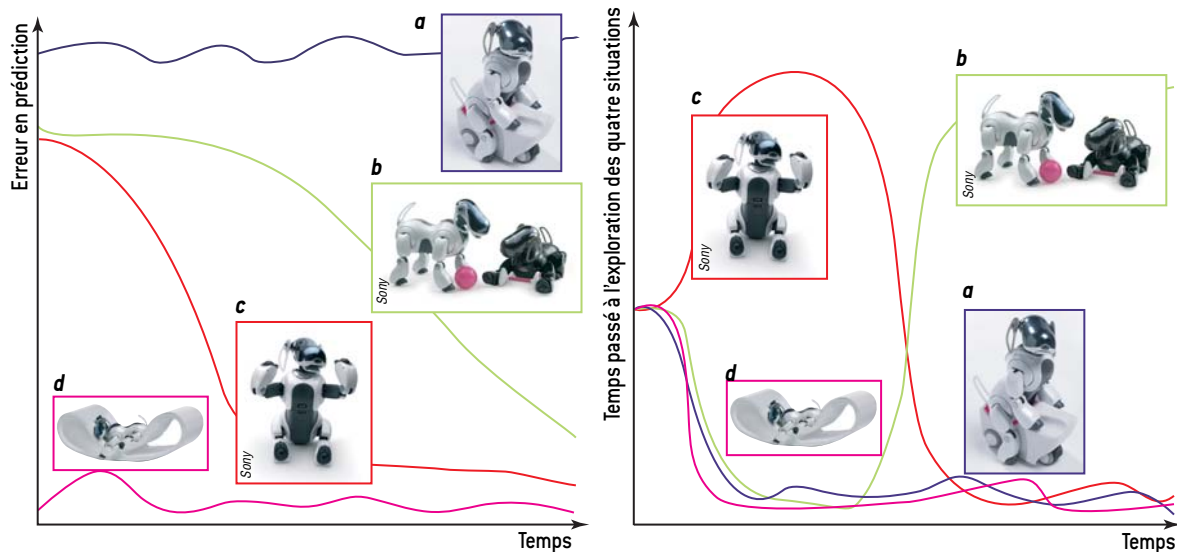
Découvrir les objets... puis communiquer

Pour cette seconde expérience, nous plaçons le robot dans un environnement contenant plusieurs objets, tel ceux d'un tapis de jeu pour bébé (voir la figure 2). Un second robot programmé pour imiter les sons émis par le premier est aussi présent. Le robot peut exécuter de nombreux mouvements. Il perçoit et mémorise les conséquences de ses mouvements par le biais de sa caméra, des sons qu'il enregistre et de ses capteurs de distance. Comme pour la découverte de la marche, le robot explore l'espace avec pour unique objectif la maximisation de la réduction de ses erreurs en prédiction initialement élevées. Chaque expérience dure une dizaine d'heures et se déroule ainsi : après une phase d'exploration aléatoire, le robot commence à effectuer de façon systématique des actions répétées (frapper, mordre). Puis il découvre des parties de son environnement (objets à frapper, objets à mordre) qui correspondent à des progrès en apprentissage. Cependant, à ce stade, le robot essaie encore de frapper les objets à mordre et de mordre les objets à frapper. Or quand le robot

Comment un robot peut-il être curieux ?

Le système de curiosité du robot lui permet de trouver de nouvelles « niches de progrès ». Imaginons un environnement où il existe quatre types de contextes sensoriels et moteurs pour le robot : il peut dormir, bouger une patte, taper dans une balle sans bouger ou faire du scooter. Si l'on forçait le robot à se concentrer sur chacune de ces activités séparément, on mesurerait l'évolution de son erreur en prédiction dans chaque contexte (les courbes à gauche). Dans la situation *a* (faire du scooter), l'erreur en prédiction est toujours élevée et ne diminue pas, peut-être parce que cette situation est trop compliquée pour le système d'apprentissage du robot. Dans la situation *d* (dormir), l'erreur est toujours basse et ne change pas (cette situation est facile donc peu intéressante pour le système d'apprentissage), et dans les situations *b* (taper dans une balle) et *c* (bouger une patte), l'erreur en prédiction est importante au départ, mais diminue ensuite plus ou moins rapidement. En pratique, le robot est placé dans cet environnement en ignorant qu'il existe quatre

types de situations, et *a fortiori* les courbes d'apprentissage correspondantes. Au départ, il explore aléatoirement son environnement, découvrant qu'il existe des situations différentes, et évalue l'intérêt de chaque situation en termes de réduction potentielle de ses erreurs en prédiction. Comme on le voit sur les courbes du temps passé à explorer chaque situation (à droite), le robot évite les situations *a* (trop compliquée, courbe bleue) et *d* (trop simple, courbe rose), qui ne permettent pas de progrès en apprentissage. Il les explore cependant de temps en temps et par hasard pour vérifier qu'elles restent peu intéressantes. À l'inverse, il se consacre à la situation *c* (courbe rouge) pour laquelle ses prédictions s'améliorent le plus vite initialement. Après un certain temps, la situation *c* est maîtrisée et par conséquent prédictible : il l'abandonne, car l'erreur en prédiction ne diminue presque plus. Il consacre alors l'essentiel de son temps à la situation *b* (courbe verte) qui, à ce stade de son développement, lui procure le plus de progrès en apprentissage.



frappe l'objet à mordre, qui est un objet lourd et mou posé sur le sol, rien ne se passe. Au contraire, l'objet à frapper peut se déplacer et se balancer de différentes façons, ce qui offre de multiples occasions d'apprentissage au robot. Ainsi, le robot « découvre » qu'il est plus intéressant d'associer la bonne action au bon objet. La maîtrise de ces bonnes associations – nommées « affordances » – se fait en quelques heures.

La production de sons, explorée dans un premier temps comme les mouvements du corps, est rapidement abandonnée au profit de l'interaction avec les objets, qui donne des résultats plus immédiats en termes de réduction d'erreur. En effet, lors des interactions sonores, le son émis par le robot est déformé quand l'autre robot l'imité. Cette déformation est difficile à prédire, elle est trop compliquée... l'erreur en prédiction reste importante. Mais dès que le robot a atteint une certaine maîtrise des objets de son environnement, il réémet des sons et se consacre presque exclusivement à cette compétence : il s'essaie aux échanges vocaux en optimisant la diminution de son erreur en prédiction.

Chaque robot découvre à sa façon son environnement, c'est-à-dire que chaque expérience conduit à une trajectoire de développement unique. Cependant la majorité des

trajectoires sont structurées de la même façon, car chaque robot explore les tâches les plus simples avant de se consacrer aux situations les plus difficiles en termes de prédiction. Les algorithmes qui contrôlent le robot ne définissent pas entièrement son évolution ; son environnement non plus. La trajectoire de développement du robot dépend de son système d'apprentissage, de ses caractéristiques physiques et de l'endroit où il est placé. Nous pouvons donc étudier, en faisant de telles expériences, le rôle respectif de chacun de ces facteurs dans l'organisation du développement.

Grâce à notre système de motivation intrinsèque, un robot apprend à maîtriser son corps, à explorer les objets présents dans son environnement proche et à s'engager vers les premières formes d'interaction. Les compétences que le robot développe sont adaptées à sa morphologie et aux environnements auxquels il est confronté. Le robot « choisit » vers quoi il tourne son attention en fonction de ses expériences passées : il est acteur de son propre développement et construit ainsi son monde. Cependant cette structure développée de façon autonome par le robot n'a pas de raison de correspondre à celle de notre monde. En particulier, il est fortement improbable que le robot ait pu développer seul les concepts

que nous voulions initialement lui apprendre, tel celui d'une balle. Pour le robot, il existe deux organisations indépendantes : celle qu'il a développée seul et celle qui peut lui être proposée de l'extérieur sous la forme des mots de notre langue.

Deux mondes parallèles

Il est probable que les différences physiques 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 le philosophe autrichien Ludwig Wittgenstein (1889-1951). Entre le monde des robots et celui des hommes, il y aura peut-être des passerelles linguistiques, mais leur nombre et leur largeur restent inconnus. Quel est alors le rôle de l'homme dans cette approche de l'apprentissage ? Face à un robot qui se développe seul, il organise seulement les situations d'apprentissage que la machine rencontre. C'est ainsi qu'il peut, dans une certaine mesure, guider le développement de la machine.

L'étude de l'acquisition du langage chez les robots nous a conduits à celle du partage de l'attention, puis à l'acquisition de savoir-faire et au développement de motivations intrinsèques. À mesure que les robots progressent ou se heurtent à des difficultés inattendues, nous comprenons davantage les phénomènes qu'ils tentent d'approcher. Ainsi, nous ne construisons pas seulement des machines « plus intelligentes » : cela nous permet de mieux comprendre ce qu'« intelligent » veut dire. Les hypothèses proposées viennent compléter les travaux entrepris dans d'autres domaines. Les recherches actuelles en neurosciences portent une attention toute particulière sur la façon 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 pluridisciplinaires se multiplient. Dans cette approche, il ne s'agit pas d'imiter l'homme en tout point. C'est au contraire en étudiant le rôle du corps et des motivations de la machine sur son propre développement que l'on peut en retour construire des explications pertinentes sur l'influence de facteurs comparables chez le jeune enfant. La robotique propose ainsi une démarche expérimentale d'un genre nouveau pour mieux comprendre les dynamiques complexes du développement, et éclairer sous un autre jour l'extraordinaire capacité qu'ont les enfants à apprendre...

Frédéric KAPLAN et Pierre-Yves OUDEYER mènent leurs recherches au *Sony Computer Science Laboratory* à Paris.

P.-Y. OUDEYER et F. KAPLAN, *Discovering communication*, in *Connection Science*, vol. 18, pp. 189-206, 2006.

P.-Y. OUDEYER, *Self-organization in the evolution of speech*, in Oxford University Press, 2006.

F. KAPLAN, *Les machines apprivoisées*, Vuibert, 2005.

Sites de Pierre-Yves Oudeyer et Frédéric Kaplan : www.csl.sony.fr/~py, www.csl.sony.fr/~kaplan.

Site du congrès Epigenetic Robotics, 20-22 septembre 2006 à l'Hôpital de la Salpêtrière : www.epigenetic-robotics.org