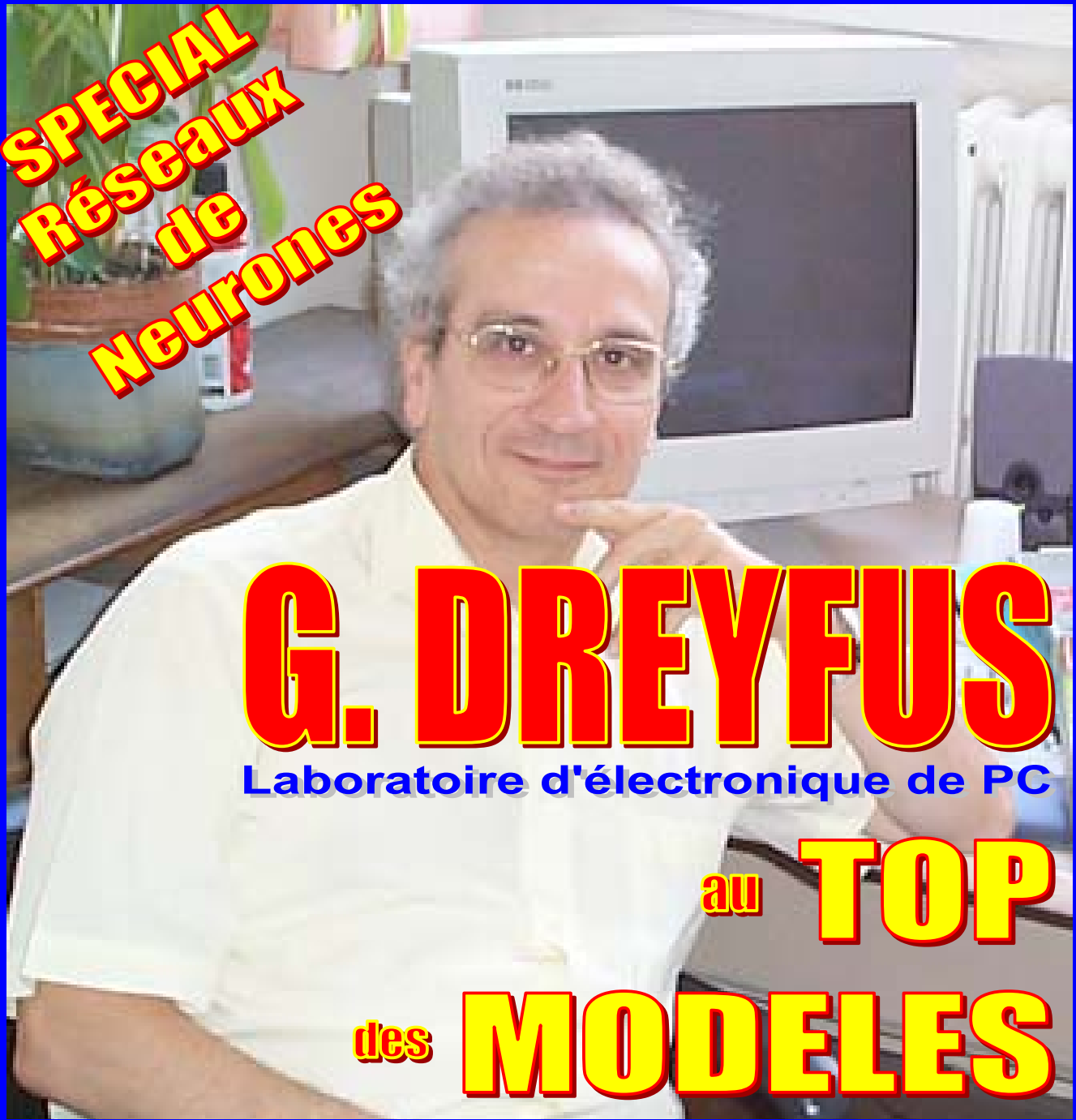


10_{rue} V@uquelin

La feuille de liaison des PC

N°11 Mars 2004



SPECIAL
Réseaux
de
Neurones

G. DREYFUS
Laboratoire d'électronique de PC

au **TOP**
des **MODELES**

Gérard Dreyfus (86) : l'apprentissage, c'est pas sorcier !

Directeur du Laboratoire d'Electronique, Gérard est Professeur à l'Ecole où il enseigne en 1^{ère} année l'électronique des circuits intégrés et la commande de processus. En 3^{ème} année, son cours de physique des composants micro-électroniques familiarise les élèves avec le fonctionnement physique des composants électroniques qui ont été étudiés et mis en oeuvre à l'occasion du cours et des TP d'Électronique de première année. Egalement en 3^{ème} année, son cours sur la modélisation et sur l'apprentissage est destiné à fournir aux physiciens et chimistes les éléments essentiels des méthodes qui permettent de réaliser des modèles prédictifs et/ou explicatifs à partir de données expérimentales. La présente livraison de 10rueV@uquelin est centrée sur ce dernier domaine.

Gérard est également le père de notre camarade Rémi Dreyfus (117).

Enfin sachez qu'il est tombé dans le chant quand il était étudiant et qu'il n'en est jamais sorti (voir le site de l'Association Science et Musique <http://www.scm.espci.fr>).

Qu'appelle-t-on réseau de neurones ?

Les réseaux de neurones sont des outils mathématiques qui permettent de créer des modèles non linéaires à partir de données expérimentales. Leurs applications couvrent tous les domaines où l'on peut effectuer des mesures, et où l'on cherche à prédire des phénomènes non linéaires : physique, chimie, biologie, mécanique, astronomie, logistique, systèmes d'informations et d'aide à la décision, marketing, finance, etc.

L'expression « réseau de neurones » pour désigner des algorithmes, voire des logiciels fait rêver ! Comment est-elle née ?

Historiquement, ces méthodes cherchaient à imiter un des processus les plus utiles et les plus mystérieux du cerveau humain et animal : celui de l'apprentissage. Mais l'on s'est rapidement rendu compte que le paradigme neurobiologique pouvait certes fournir une inspiration initiale, mais qu'il était complètement inutile pour progresser dans la compréhension de l'apprentissage artificiel, et dans la conception d'algorithmes efficaces.

Pour reprendre une image chère à notre camarade Michel Weinfeld (78), Directeur de Recherches au Laboratoire d'Informatique de l'École Polytechnique, si les ingénieurs en aéronautique s'étaient contentés d'imiter la nature, ils chercheraient encore à concevoir des avions aux fuselages couverts de plumes, et qui battent des ailes !

Les vrais progrès dans l'apprentissage artificiel (« machine learning ») sont venus de l'utilisation des concepts issus des sciences « dures », notamment, dans les années 1980, de la physique statistique. Néanmoins, le terme de réseaux de neurones est resté, mais l'ingénieur doit considérer ceux-ci simplement comme des outils très efficaces de modélisation non linéaire. À présent, il y a d'autres outils du même genre, notamment les « Machines à Vecteurs Supports » inventées par Isabelle Guyon (100), et sur lesquelles nous travaillons aussi.

Existe-t-il un lien quelconque avec la physiologie du système nerveux, voire du cerveau ?

Les réseaux de neurones « mathématiques » sont naturellement d'excellents candidats pour réaliser des modèles des systèmes neuronaux vivants. Dans ce cadre, l'objectif de la recherche n'est pas de concevoir des algorithmes efficaces de traitement de l'information, mais d'essayer de répondre à des questions fondamentales des « neurosciences computationnelles », cette branche des neurosciences qui s'intéresse à la manière dont le cerveau calcule, raisonne et prend des décisions. Au laboratoire d'Électronique, nous avons une équipe qui cherche à comprendre comment les informations sont codées dans les systèmes nerveux, notamment dans le système olfactif. C'est aussi un sujet absolument passionnant, mais qui nécessiterait à lui seul de longs développements !

Pour en finir avec les questions de terminologie, que recouvrent les expressions « intelligence artificielle » et « logique floue » ?

Ce genre d'exercice de définition de frontières peut donner lieu à des querelles interminables et à des brouilles définitives, donc prudence ! On pourrait dire que l'Intelligence Artificielle est la branche de l'informatique qui traite des données symboliques, et qui permet notamment de formaliser les raisonnements. Les produits les plus connus de l'Intelligence Artificielle sont les systèmes experts, qui permettent d'implanter dans des ordinateurs les connaissances d'experts sous forme de règles du genre « Si la température dépasse 250°, il faut fermer telle vanne ». En ce sens, l'apprentissage numérique, donc les réseaux de neurones, ne fait pas partie de l'intelligence artificielle.

La logique floue a pour objectif de formaliser les « raisonnements flous », du genre « Si la température augmente trop, il faut tourner un peu telle vanne vers la droite ». Cela n'a rien à voir avec les réseaux de neurones ; néanmoins, dans certains cas, il peut être avantageux de combiner les deux techniques.

Parlons donc d'apprentissage. On voit bien ce qu'est l'apprentissage pour des humains ou même des animaux. Mais pour des ordinateurs ?

Lorsque l'on veut enseigner la lecture à un enfant, on lui montre des exemples de lettres ou de chiffres, dans des styles d'écriture ou des fontes diverses ; lorsque l'apprentissage de la lecture est terminé, on attend de l'enfant qu'il sache non seulement lire les lettres de son livre de lecture, mais également n'importe quel mot ou n'importe quel chiffre, écrit dans n'importe quel style. On veut qu'il soit capable de généraliser ce qu'il a appris à partir des exemples, sans qu'il soit nécessaire de lui fournir une description discursive de la topologie des chiffres et des lettres.

De même, lorsque l'on cherche à modéliser un processus que l'on observe, on peut effectuer des mesures des causes et des effets – observer des exemples de comportement – et en inférer des relations mathématiques qui permettent de généraliser à partir de ces exemples, c'est-à-dire de prédire le comportement du processus dans des situations différentes de celles de l'apprentissage.

C'est une alternative à la démarche classique qui consiste à appliquer les lois de la physique (ou de la chimie, de la biologie, etc.) et à résoudre les équations qui en résultent.

Comment fait-on en pratique ?

Supposons que l'on cherche à prédire le taux de pollution de l'air par l'ozone, 24 heures à l'avance.

La mise en équations du phénomène nécessite des connaissances sur la physico-chimie de la production d'ozone en fonction des conditions de température, d'ensoleillement, de concentration en polluants, ainsi que des connaissances sur la circulation des masses gazeuses entre les points de production et les points de mesure, et d'autres encore.

C'est possible ?

Bien sûr, mais c'est très complexe, et la résolution numérique des équations différentielles qui en résultent peut nécessiter un temps de calcul considérable, éventuellement supérieur au délai de prédiction !

Que faire alors ?

L'approche par apprentissage consiste à utiliser les données enregistrées dans le passé pour effectuer l'apprentissage d'un modèle dont les variables sont les données climatiques du ou des jours précédents, les taux d'ozone atmosphérique du ou des jours précédents, des éléments de prévision météorologiques pour le jour suivant, et toute autre grandeur mesurable ou calculable que l'on peut considérer comme pertinente pour la prédiction du taux de pollution.

Le modèle proprement dit est une fonction non-linéaire, paramétrée, des variables choisies, et l'apprentissage est un algorithme qui calcule les paramètres du modèle de manière à ce que ce dernier s'ajuste aux données. Cette fonction paramétrée peut être quelconque.

Les « réseaux de neurones » constituent une famille particulière de telles fonctions paramétrées, qui possède une propriété très intéressante : la *parcimonie*. Ils permettent de construire des modèles à partir d'un nombre de mesures plus restreint que d'autres familles de modèles, avec de meilleures qualités de généralisation.

Est-ce différent des méthodes statistiques de régression du type moindres carrés ?

Techniquement, l'apprentissage numérique est très semblable à la méthode des moindres carrés : on cherche un jeu de paramètres qui minimise la somme des carrés des écarts entre les mesures utilisées pour l'apprentissage et les prédictions du modèle.

Mais l'objectif est complètement différent : lorsque l'on « fait un fit », comme disent les physiciens, on cherche à estimer des paramètres physiques : imaginons que j'aie construit un modèle physico-chimique d'un processus, et que j'aie trouvé qu'il est activé thermiquement, c'est-à-dire qu'il suit une loi du genre $y = \exp(-E/kT)$. Alors, je fais quelques mesures de y en fonction de la température T , je trace $\ln(y)$ en fonction de $1/kT$, j'applique la méthode des moindres carrés pour obtenir la pente de la droite, ce qui me fournit une estimation de l'énergie d'activation.

Dans ce cadre classique, on cherche une estimation aussi précise que possible de cette énergie, et des tests statistiques bien connus permettent d'avoir une idée de la confiance que l'on peut avoir en ces estimations.

Et dans le cas de l'apprentissage ?

Dans le cadre de la modélisation par apprentissage, on ne cherche pas à déterminer de manière précise un ou plusieurs paramètres d'un modèle établi à partir de connaissances physiques, mais on cherche les paramètres qui donnent à un modèle les meilleures propriétés de prédiction, ou de généralisation, compte tenu des seuls résultats expérimentaux dont on dispose ; c'est un problème très différent, et beaucoup plus difficile.

Comment l'apprentissage intègre-t-il l'état des connaissances du moment ?

Si l'on dispose, outre des mesures, de connaissances certaines sur le processus, on peut les utiliser pour structurer le modèle. Par exemple, si la conservation de la masse constitue une des équations du modèle, on ne va pas faire un réseau de neurones pour redécouvrir par apprentissage, à partir des mesures, que la masse est conservée : on utilise directement cette information dans la structure du modèle. C'est la technique de *modélisation semi-physique*, que nous avons

introduite il y a quelques années, et que nous avons développée à l'occasion de plusieurs applications industrielles.

Quels sont les domaines d'applications concernés ?

Tous ! Chaque fois que l'on ne dispose pas d'un modèle suffisamment précis, ou suffisamment simple pour être utilisable en temps réel, on peut faire appel à des techniques d'apprentissage (par exemple des réseaux de neurones, ou des machines à vecteurs supports), si l'on dispose de mesures.

Mais on n'utilise pas les réseaux de neurones uniquement pour faire des modèles prédictifs du comportement de processus. Il y a aussi de très nombreuses applications dans le domaine de la classification automatique, pour la reconnaissance de formes, ou pour le traitement d'informations.

Peux-tu nous donner des exemples concrets ?

Nous avons travaillé avec des entreprises très variées :

Avec Sollac, sur la modélisation du soudage par points, qui est le procédé de soudage le plus utilisé dans l'industrie automobile.

Avec Michelin, pour concevoir des méthodes d'aide à la conception de pneumatiques. Chaque fois que tu achètes un nouveau pneu Michelin, tu achètes des réseaux de neurones conçus à PC !

Pour 3M, nous avons effectué une modélisation semi-physique du séchage du ruban Scotch

Pour Sagem, nous avons mis au point le pilotage d'un véhicule Mercedes 4x4 autonome – c'était l'objet de la thèse d'Isabelle Rivals (106).

Pour Renault, nous travaillons actuellement sur la : commande des moteurs – thèse de Marc Lucéa (116)

Avec Ela Médical, nous avons conçu de nouvelles méthodes d'apprentissage pour la détection précoce des anomalies cardiaques en électrocardiographie

Nous avons également travaillé avec PSA, et bien d'autres...

Vous avez collaboré avec la Caisse des Dépôts et Consignations, qui n'est pas un partenaire habituel de PC ?

En effet, dans une première étude, nous avons abordé le problème de l'évaluation des capacités financières des collectivités locales, dans le cadre de la thèse d'Hervé Stoppiglia (107). Puis la Caisse des Dépôts nous a demandé de concevoir et de réaliser avec eux un système de filtrage automatique des informations : il s'agit de trouver, dans un grand corpus de textes (par exemple, le flot des dépêches de l'Agence France-Presse ou de Reuters), ceux qui traitent un thème donné, défini par quelques mots. C'est un problème difficile, qui va bien au-delà de la simple recherche par mots-clés. Il a fait l'objet de la thèse de Mathieu Stricker (112).

Exemple : intéressons-nous aux « prises de participation entre entreprises ».

Une dépêche qui dit « Alcan rachète Pêchiney » ne contient aucun des mots qui définissent le thème, mais elle est tout à fait pertinente. En revanche, une dépêche qui dit « la participation des communistes au gouvernement inquiète les chefs d'entreprises » contient deux des trois mots qui définissent le thème, mais n'est pas pertinente.

En utilisant des techniques d'apprentissage, nous avons remporté l'épreuve de filtrage d'informations d'une compétition internationale organisée dans le cadre de la conférence TREC-9 (Text **RE**trieval Conference) ; nous avons devancé plusieurs labos français et étrangers, y compris Microsoft !

Ce qui prouve, s'il en était encore besoin, qu'il n'est pas nécessaire d'être dans un énorme labo pour avoir de bonnes idées...

Ce qui est particulièrement intéressant, c'est que les mêmes idées qui nous ont permis de tenir compte du contexte dans l'analyse automatique des textes devraient être précieuses dans l'aide à la découverte de médicaments ; c'est le thème d'une collaboration avec le laboratoire de Chimie Organique de l'École, collaboration dont la cheville ouvrière est Arthur Duprat (97).

As-tu été tenté de lancer ta propre entreprise?

Oui, car il est important que les enseignants de PC, tout en faisant une recherche fondamentale de pointe, restent en prise directe avec le monde industriel auquel se destinent la majorité des étudiants.

Participer à la genèse d'entreprises innovantes est une bonne façon de concilier ces objectifs. À ma connaissance, il y a au moins six professeurs de matières majeures de l'École qui ont participé à la création d'entreprises, ce qui est une proportion énorme !

Qu'entends-tu par participer ?

J'entends par là qu'ils ont fait en sorte, de différentes manières, que les recherches menées dans leur labo débouche sur la création d'une entreprise. Cela peut consister, par exemple, à donner aux créateurs un accès aux moyens du laboratoire, dans des conditions préférentielles, pour faciliter le démarrage.

Et toi ?

J'ai créé une première entreprise en 1973, avec Jacques Lewiner. Elle avait pour objectif de fabriquer des capteurs à électrets utilisant des principes que j'avais développés dans ma thèse. Ce fut un échec, mais nous avons su en tirer les leçons.

Lesquelles ?

Il faut conserver uniquement l'activité proprement « créative », et sous-traiter tout ce qui peut l'être. C'est une banalité maintenant, mais ce n'était pas évident à une époque où le terme de « start-up » n'existait même pas. D'autre part, il faut que l'entreprise soit géographiquement aussi proche que possible du laboratoire.

Ta seconde tentative a-t-elle été la bonne ?

Oui. Lorsque, en 1994, mon ancien binôme de TP Jean-Luc Ploix (86) est venu me demander s'il n'y avait pas quelque chose à faire avec les réseaux de neurones, ma réponse a été positive sans hésitation !

En effet, j'avais observé que le laboratoire était à même de faire des démonstrateurs qui répondaient aux préoccupations des industriels, mais que ces derniers hésitaient à faire l'investissement supplémentaire qui permettait de transformer l'essai.

Quel était leur cahier des charges ?

Implanter les techniques d'apprentissage et de modélisation découvertes au labo dans des outils logiciels de qualité industrielle, optimisés, fiables, maintenables, utilisables par des non-spécialistes après une brève formation, et facilement intégrables dans des applications. Bref, faire ce que les chercheurs universitaires ne savent pas faire, car ce n'est pas leur métier. La société de Jean-Luc - Netral S.A. - occupe remarquablement cette niche depuis presque 10 ans.

Quel est ton rôle chez Netral?

C'est un rôle de conseiller scientifique : attirer l'attention de l'entreprise sur les nouveautés qu'il serait utile d'implanter dans les futures versions de leur logiciel, éventuellement les aider à répondre aux questions de leurs clients. Lorsqu'un industriel vient me poser un problème qui ne me paraît pas relever de la recherche, mais de la simple

application des réseaux de neurones en modélisation ou en reconnaissance de formes, je l'envoie vers Netral.

À l'inverse, lorsqu'un prospect de Netral a un problème dont la résolution nécessite un réel effort de recherche, ils leur conseillent de prendre contact avec le labo.

En quoi ta formation PC te préparait-elle à cette évolution ?

En rien... et en tout ! En rien, puisque tous les sujets sur lesquels je travaille, et les techniques que je mets en œuvre, étaient inimaginables lorsque j'étais étudiant.

En tout, parce que la formation pluridisciplinaire de l'École permet d'aborder sans complexes les sujets les plus divers, et parce que la formation par la pratique est une irremplaçable école de rigueur et de bon sens.

Quel avait été ton cursus ?

En 4^{ème} année, j'ai fait le DEA de Physique des Solides, dirigé par un jeune prof à qui l'on prédisait un bel avenir, un certain Pierre-Gilles de Gennes.

Puis j'ai fait une thèse sur les propriétés électriques des films de polymères, sous la direction de Jacques Lewiner.

Après ma thèse, au milieu des années 70, je me suis impliqué dans les développements industriels de capteurs, auxquels il est apparu rapidement indispensable d'associer de l'électronique intégrée, qui en était à ses débuts.

Et il a fallu s'intéresser ensuite au traitement « intelligent » des signaux fournis par cette électronique...

Quelle est la place de PC dans ce domaine ?

PC a été pionnier dans le domaine, et conserve une place de premier rang, en dépit du développement énorme des recherches dans le monde. Si PC a pu prendre une avance appréciable c'est précisément parce que notre formation a permis d'oser se lancer dans un sujet qui paraissait farfelu, en prenant appui sur des bases solides.

La première doctorante du labo, Isabelle Guyon (100), était chimiste ! Elle a joué un rôle capital dans l'impulsion initiale de nos recherches. Elle est devenue un des meilleurs chercheurs de sa génération en « machine learning » : l'École peut être fière d'elle. Et en plus ses enfants sont adorables : je ne manque pas de passer la voir à Berkeley chaque fois que je vais en Californie !

Que pensait Pierre-Gilles de Gennes de tout ça ?

Il a su considérer avec bienveillance le fait que, en 1982, je lance un labo tout nouvellement créé sur un sujet aussi improbable. Dès que nous avons publié nos premiers résultats dans des revues de physique réputées, et que nous avons pu exhiber des applications, il a fermement soutenu cette orientation.

À l'inverse, les chercheurs de ma génération qui avaient un peu les mêmes idées que moi ont connu les plus grandes difficultés pour les faire admettre au niveau de leurs organismes de recherche.

Quels organismes ?

Le CNRS, l'INRIA, certaines universités. Je me rappelle un ayatollah des mathématiques, membre de l'Académie des Sciences, qui écrivait que ce thème de recherche était « du brandouillage de connexions ». Il faut dire, à sa décharge, que l'apprentissage numérique avait été affublé, à ses débuts, du nom malheureux de « connexionnisme ». Je ne sais plus qui disait que le seul mot en « isme » qui soit utile en sciences est le mot « prisme »...

Les premiers résultats en question remontent à quand ? Dans quelles revues ont-ils été publiés ? Quelles étaient ces premières applications ?

Nos premiers résultats théoriques ont été publiés en 1985, donc nous avons dû les obtenir en 1983-84. Nous les

avons publiés sous la forme d'une lettre au Journal de Physique, puis d'articles dans l'austère Physical Review. Nous avons notamment publié un article dans lequel nous décrivions comment les réseaux de neurones pouvaient mémoriser des séquences d'informations, et les retrouver à partir d'un fragment de cette séquence, même partiellement erronée. Cela s'appelle la faculté de « mémoire associative ». Pour illustrer nos algorithmes, nous avons réalisé un réseau qui mémorisait et retrouvait des poèmes. C'est ainsi qu'Isabelle Guyon (100), Léon Personnaz (ESPCI), Jean-Pierre Nadal (physicien de l'École Normale Supérieure) et moi-même avons publié des poèmes de Verlaine dans Physical Review¹ !

Les premières applications significatives étaient relatives à la reconnaissance des codes postaux manuscrits. Un ancien doctorant du laboratoire, Stefan Knerr, a créé la société Vision Objects, qui est l'un des leaders mondiaux de la lecture automatique de documents manuscrits (<http://www.visionobjects.com>).

Comment ton labo fait-il connaître ses travaux au monde extérieur ?

Par nos publications dans les journaux et congrès internationaux, comme il se doit. Mais notre vitrine vers l'industrie et le public est notre site Web (<http://www.neurones.espci.fr>), régulièrement tenu à jour. On peut y trouver des détails sur nos recherches en cours ou récentes, savoir ce que sont devenus nos anciens doctorants, télécharger de nombreux articles (scientifiques et de vulgarisation) ainsi que des thèses, ...

¹ pour les curieux : Storage and Retrieval of Complex Sequences in Neural Networks, Physical. Review A, vol. 38, 6365 (1988)

Le livre dont j'ai dirigé la rédaction (*Réseaux de neurones, méthodologie et applications* (Eyrolles)) est aussi un élément important de notre communication². J'ai fait appel à six des meilleurs spécialistes français des différents aspects du domaine, appartenant à divers organismes (CEA, CNAM, UVSQ, IMAG, Sup Aéro, CENG), car il faudrait être bien présomptueux pour prétendre couvrir, avec les seules forces du laboratoire, l'ensemble d'un sujet aussi riche et foisonnant. J'ai écrit deux des chapitres, mes collègues ont écrit les autres, mais nous avons veillé à ce que le vocabulaire et les notations soient unifiés, de manière à produire un ouvrage réellement pédagogique. De nombreuses applications y sont décrites. Trois chapitres peuvent être téléchargés gratuitement².

Les relations avec les « sociétés savantes » internationales sont également un outil de communication utile. Dans notre domaine, c'est l'IEEE (Institute of Electrical and Electronics Engineers) qui est l'interlocuteur naturel. J'ai été élu membre du Technical Committee « Machine Learning for Signal Processing », et je viens de créer une section française de la « IEEE Neural Networks Society », elle-même de création récente.

Bien entendu, nous nous efforçons d'apparaître dans les journaux français plus ou moins « grand public ». Récemment, une demi-page de *l'Usine Nouvelle* a été consacrée à notre labo, ainsi qu'une pleine page dans *Industrie et Technologies*.

Enfin, nous organisons avec Netral deux sessions de formation permanente par an (cours et Travaux Pratiques) destinées aux ingénieurs.

Merci !

² <http://www.neurones.espci.fr/Francais.Docs/livre.htm>

Netral et Jean-Luc Ploix (86) : bientôt 10 ans de vie commune !



Jean-Luc Ploix crée Netral (www.netral.com) en 1994. Le capital social de Netral est détenu uniquement par des personnes physiques. Pas de banques dans le tour de table. Netral est spécialisée dans le calcul scientifique à l'aide de réseaux de neurones. Le chiffre d'affaires se répartit en trois tiers entre études/conseil, formation et vente de logiciels. Parmi les clients : PSA Peugeot Citroën, Renault, Renault Trucks, Snecma, Dassault Aviation, Arcelor, Institut Français du Pétrole, Total, Rhodia, EDF, SNET Electricité, CEA, Saint-Gobain, CF Gomma, Altadis, Direction des Eaux et de l'Assainissement de la Seine-St-Denis.

Te souviens-tu pourquoi tu as fait PC?

Je suis entré à PC car je souhaitais une formation à la recherche. J'avais rencontré dans mon entourage des chercheurs, travaillant dans des laboratoires, et j'avais compris que l'esprit de recherche était ce qui me convenait. Et je crois que j'ai gardé cet esprit jusqu'à maintenant.

Si c'était à refaire?

Je le referais sans hésitation

Peux-tu nous dire comment tu en es venu à créer ta propre entreprise ?

En 1993, je me suis trouvé en disponibilité. C'est une façon pudique de dire que j'ai fait partie d'une charrette de licenciements économiques. A 45 ans l'avenir s'assombrit sérieusement dans ces cas là. Alors, j'ai fait ce que tout le monde fait, c'est à dire un tour des amis et connaissances pour explorer les possibilités. Ma visite à l'Ecole chez Gérard Dreyfus a été déterminante. Il avait une technologie en voie de mûrissement, il était un des premiers au monde à la posséder, et on commençait à identifier des applications industrielles.

Après un mois de réflexion, ma décision a été prise : je me consacrerai entièrement à la mise en place de Netral, à l'exclusion de toute autre recherche d'emploi.

Quelles sont les principales difficultés que tu as eu à surmonter?

Pour les entreprises de ce type, il y a des lois classiques de comportement. Les deux premières années sont en général faciles. Mais la première crise arrive très vite. On a vite utilisé les fonds rassemblés, et épuisé les diverses aides obtenues à la création. Les premiers contrats arrivent à échéance, et il faut faire la transition d'une phase de lancement vers une phase de fonctionnement continu. La principale difficulté à surmonter alors est de faire le tri parmi la somme des conseils que l'on reçoit. Le jeune chef d'entreprise est en effet la cible privilégiée des donneurs de conseils de tous genres, banquiers, chambre de commerce, amis, etc. Or, si aucun de ces conseillers n'est compétent dans le domaine d'action de l'entreprise, tous ont un avis définitif sur ce qu'il convient de faire. Si j'avais écouté ces conseils, Netral aurait fermé plusieurs fois depuis le début. Heureusement, je ne les ai

pas suivis, et l'ensemble des amis qui ont investi dans Netral m'a conservé sa confiance.

La création d'une entreprise de haute technologie est une véritable gageure. Les banquiers restent méfiants, n'ayant aucune compétence pour juger de la qualité de nos produits et prestations ni la réalité d'un marché qui n'est encore que latent. De plus, la tendance très répandue à demander un retour sur investissement dans le trimestre qui suit n'aide pas à faire comprendre qu'une nouvelle technologie a besoin de dix ans pour commencer à être admise dans les bureaux d'études. Nos décideurs ont en effet du mal à faire confiance à une technologie qu'ils n'ont pas apprise à l'école ou, plus grave encore, qui n'a pas déjà été testée par leurs concurrents.

Les principales satisfactions qu'elle t'a apportées?

Ceci dit, cette aventure m'a apporté de très nombreuses satisfactions. Aujourd'hui, six personnes y travaillent, et nous avons de très belles perspectives devant nous. Et puis nous avons acquis une bonne réputation auprès de nos clients. Et je dois dire qu'elle n'aurait pas été possible sans le fort soutien de ma famille très proche.

Qu'appelle-t-on réseau de neurones? Pourquoi "neurones"?

Pour nous "réseau de neurones" signifie outil statistique non linéaire. Cela n'a plus rien à voir ni avec le neurone biologique, ni avec le cerveau. Il s'agit là de scories de l'histoire. C'est simplement un outil mathématique bien adapté à la modélisation du comportement des processus déterministes. Mais nous conservons la dénomination tout simplement parce qu'elle est connue de tous et qu'il est désormais impossible d'en changer.

A quoi ça sert?

Cet outil sert à modéliser des processus. Et une fois que le modèle est disponible, il peut servir:

- à simuler, pour éviter une expérience coûteuse ou dangereuse,
- à prendre une décision à la vue d'une prédiction précoce,
- à piloter le processus en introduisant le modèle dans une boucle de régulation,
- à inverser le modèle, et retrouver les variables d'entrée qui conduisent à une caractéristique de sortie visée.

As tu des exemples concrets d'applications?

Aujourd'hui, il s'agit surtout de processus industriels. Les exemples de modélisation sont très nombreux :

- les compositions des distillats et résidus d'une colonne à distiller en fonction des commandes appliquées ;
- la température de transition vitreuse d'un verre en fonction des ses constituants ;
- les caractéristiques d'une peinture en fonction de sa formule ;
- la traction nécessaire sur la tôle dans un laminoir en fonction de la nuance de l'acier et de l'épaisseur désirée ;
- la valeur future de la concentration de l'atmosphère en NOx en fonction des conditions météo ;
- La capacité de fermentation d'une récolte de raisin en fonction d'une analyse spectrale du moût à la récolte.

De nombreuses autres applications ne peuvent pas être citées pour des raisons de confidentialité.

Le domaine couvert, on le voit, est très vaste. Il n'est limité que par l'imagination des utilisateurs. Des essais ont été faits dans le domaine financier. La prédiction des cours de bourse en a excité plus d'un. A ma connaissance, il n'y a eu que des échecs. Est-ce par manque de déterminisme ? Dans ce domaine, nos outils sont plus adaptés à la modélisation des densités de probabilités qu'à celle du processus lui-même.

Très récemment, nous avons démarré le développement d'un nouveau générateur de plan d'expériences. Chacun sait qu'un plan d'expériences permet de limiter le nombre

d'expériences nécessaire à la connaissance d'un processus. Mais on oublie très souvent que la théorie des plans d'expériences a une prémisse importante : le modèle de processus est linéaire. Or la nature est très souvent non linéaire. Il nous a semblé judicieux d'appliquer notre connaissance des modèles non linéaires acquise avec les réseaux de neurones aux plans d'expériences, et faire faire un pas à cette technique vers le non linéaire. Plusieurs grands groupes français nous ont suivis dans cette démarche, et participent au financement de ce développement : le CEA, l'IFP, Rhodia et PSA. Nous espérons que d'autres nous rejoindront très bientôt.

Qui sont les clients?

Nos clients sont en majorité des grands groupes industriels. Ces groupes ont eu les moyens de mettre en place des cellules de compétence chargées de la veille technologique. Ils ont donc été les premiers à utiliser ces nouveaux outils. Quelques entreprises PME innovantes ont rejoint les rangs, mais elles restent peu nombreuses.

Quel est ton environnement concurrentiel?

Les produits concurrents sur l'aspect technique sont des produits statistiques généralistes, qui offrent toute une gamme de fonctionnalités. La force de ces produits tient dans l'éventail de leur offre. Mais dans chacune des fonctionnalités, ils ne sont pas toujours au meilleur niveau. Dans le cas des réseaux de neurones, ils sont très en retard puisque nous avons la chance d'être en relation étroite avec le Laboratoire d'Electronique de l'Ecole, et que nous introduisons dans nos produits les derniers résultats de la recherche.

Un autre type de concurrence vient des produits voisins qui proposent d'autres types de traitement de données ou de modélisations, et qui tendent à répondre aux mêmes problèmes des utilisateurs. Ces produits sont nombreux et offrent en général des prestations intéressantes. Face à ces produits, Netral intervient lorsque la réponse obtenue avec ces produits est insuffisante.

Peux-tu nous décrire une journée typique pour toi?

Netral est une petite entreprise. Je reste moi-même le principal développeur de nos produits. Je passe donc le principal de mon temps devant mon ordinateur. Le reste du temps est consacré à l'animation de l'équipe.

Recrutes-tu? Quels profils?

La principale embauche de la Société a été un ingénieur commercial. Expérience faite, les compétences techniques ne vont pas obligatoirement avec des compétences commerciales. Et le commerce est évidemment une condition majeure de notre développement.

Depuis, j'ai embauché un jeune développeur, plutôt orienté sur l'algorithmique, afin de faire évoluer les produits vers les plans d'expériences. Et si j'en ai la possibilité, j'embaucherai dans un proche avenir un nouveau commercial, puis plus tard un informaticien.

Tu es Chef d'entreprise. Comment t'es-tu formé à ce métier auquel PC ne prépare pas?

Pour moi, et pour la petite entreprise que je dirige, les compétences nécessaires au chef d'entreprise sont de plusieurs ordres :

- la compétence technique,
- le leadership de l'équipe,
- le caractère permettant d'assumer la solitude de la décision, éventuellement contre l'avis des conseils, et le calme dans les moments difficiles,
- le bon sens.

Les deux premières se comprennent bien. La troisième ne s'apprend nulle part, et elle est pourtant très importante. Il n'y a plus de "patron" vers lequel se retourner !

Quant au management sous son aspect technique, c'est d'abord du bon sens.

Pécuniairement, y trouves-tu ton compte?

L'aspect pécuniaire est intéressant à discuter. Le lancement, et surtout la continuation d'une entreprise s'est traduite par une réduction importante de mes revenus. Je suis très loin d'être le seul créateur d'entreprises dans ce cas. L'intérêt financier que je vise est surtout dans la valeur vénale d'une société qui fonctionne bien. C'est très spéculatif, mais ça fait partie du jeu.

Comment vois-tu l'avenir de Netral? Et le tien propre?

Netral peut se développer de façon importante en s'appuyant sur sa technique, sa réputation, et sa gamme de produits élargis. En France, vers les PME qui vont devenir plus nombreuses à ressentir le besoin, soit de modélisation, soit de plan d'expériences. A l'étranger, tout reste à faire.

Pour ce qui me concerne, dans la mesure du possible, je compte rester à Netral jusqu'à mon départ à la retraite. Après, je pense qu'il peut être intéressant de mettre cette société en état d'être vendue.

Quels conseils donnerais-tu à un jeune PC qui envisagerait de lancer sa propre entreprise?

A un jeune, je poserais deux questions :

1. Possèdes-tu le caractère nécessaire à la direction d'une entreprise ?
2. As-tu les moyens de tes ambitions ?

Il te faudra assumer tes décisions dans une certaine solitude, et tenir le coup dans la durée. Par ailleurs, si tu veux proposer de la haute technologie sur le marché, souviens-toi que celui-ci est particulièrement lent à l'accepter. Si la réponse à ces deux questions est oui, alors vas-y, et crois en toi.

A un plus ancien qui souhaiterait se sortir d'un mauvaise passe en créant une entreprise, je dirais ceci :

Contrairement aux apparences, la création d'une entreprise n'est pas la solution pour sortir d'un passage professionnel difficile, c'est plutôt l'inverse : la création d'une entreprise doit être le fruit d'un désir profond et ancien qui a éventuellement été rendu possible par la disponibilité qui fait suite à un passage professionnel difficile.

Merci !

Arthur Duprat (97), chisicien ou phymiste?



Allez visiter le site internet du labo de chimie organique de l'école. Vous y trouverez le CV d'Arthur. Plusieurs points vont probablement susciter votre curiosité. Ses sujets de recherche d'abord : modélisation et structure, oxydations biomimétiques avec des complexes métalliques de calixarènes, utilisation des réseaux de neurones pour la prédiction de propriétés physico-chimiques à partir de la seule structure moléculaire, chimie combinatoire virtuelle (modélisation et réseau de neurones). Ou bien alors son année en tant que Visiting Scholar à l'Université de Californie de San Diego. Ou enfin sa qualité d'Examineur en Chimie du concours X-PC. Comme en plus il a l'air très abordable (il l'est !), peut-être vous enhardirez vous comme moi à lui poser quelques questions.

Tu enseignes la Chimie Organique et tu effectues ta recherche dans le labo d'électronique de PC. Ce n'est pas banal. Comment en es-tu arrivé là ?

Bon, il faut retourner loin dans ma jeunesse pour retrouver la trace de mon attirance pour la chimie. Au cours de mes années de collège, je bricolais déjà dans mon grenier un four à arc électrique en vue de préparer du carbure de calcium pour une antique lampe à acétylène. J'ai bien dû faire disjoncter le compteur plusieurs fois avant d'y parvenir. J'avais déjà mon repaire de chimiste. En classe préparatoire, j'avais la ferme intention d'intégrer une école de chimie, et par la suite PC (concours chimie-centre à l'époque) a été la seule école dont j'ai préparé très sérieusement le concours ! Après une scolarité "sans histoire", option chimie bien sûr, je découvre les systèmes modèles et la "modélisation" au cours de mon stage de DEA au laboratoire de Chimie Organique de PC sous la direction de J.M. Aubry (91). Naturellement, je candidate pour le poste d'assistant qui vient de se libérer dans ce même laboratoire, afin d'y préparer mon doctorat d'état, dans un premier temps sous la direction de J. Rigaudy (58). Mon sujet de thèse, puis mon stage post-doctoral seront une suite logique de cette attirance pour la modélisation. De retour des Etats-Unis, je comprends combien il est important d'introduire la modélisation moléculaire dans la pratique quotidienne au laboratoire de chimie organique et dans l'enseignement. La direction des études de l'époque n'y étant pas favorable, un rapprochement avec le laboratoire d'Electronique me permettait de mettre en œuvre mes compétences en modélisation moléculaire et en chimie organique pour explorer un sujet très en vogue, la prédiction de propriétés physicochimiques avec de nouveaux outils développés au laboratoire de Gérard

Dreyfus : les réseaux de neurones. Petit à petit, mes activités de recherches se sont concentrées au laboratoire d'électronique où je suis actuellement la plupart du temps.

Quels ont été tes principaux thèmes de recherche?

Après avoir synthétisé des copolymères destinés à servir de capteurs de dioxygène singulet pendant mon DEA, j'ai abordé au cours de ma thèse l'étude d'un système oxydant à base de fer et de dioxygène susceptible de modéliser l'action d'enzymes responsables de l'hydroxylation aromatique. Cette prise de contact avec la chimie biomimétique s'est poursuivie au cours de mon stage post-doctoral ; j'ai cherché à comprendre les mécanismes et la cinétique de fixation d'une petite molécule, le monoxyde d'azote NO, tout d'abord sur des systèmes modèles de l'hème (le site actif de la myoglobine), puis sur la protéine elle-même. De retour en France, je me suis assez rapidement consacré à la modélisation moléculaire, en particulier avec le démarrage du projet impliquant l'utilisation des réseaux de neurones. Je poursuis par ailleurs une collaboration active avec O. Re naud (99) qui travaille sur la chimie de molécules amusantes, les calixarènes. Ces supramolécules peuvent adopter des conformations variées suivant le métal qu'elles complexent et la nature de l'invité qu'elles hébergent. La modélisation moléculaire est alors un outil de choix pour comprendre les modifications conformationnelles qui surviennent lors de l'échange d'un invité. Les modifications observées au niveau de tels systèmes modèles censés mimer l'action de certaines enzymes conduisent à proposer des réarrangements de structures de protéines permettant d'expliquer la régulation enzymatique.

Quelles ont été les principales avancées auxquelles tu as contribué?

Et bien tout ce qui touche à mon travail de modélisation dont je parle ci-dessus, à la fois en enseignement et recherche. La recherche est en marche, et nous sommes actuellement en train d'explorer une voie prometteuse d'apprentissage de la structure des molécules organiques à partir de leur graphe, en utilisant des réseaux de neurones, bien sûr.

Qu'appelle-t-on réseaux de neurones? Pourquoi "réseaux"? Pourquoi "neurones"?

Réseau de neurones : voilà une expression qui m'a fasciné. Réseau, pour moi c'est avant tout une idée de connexion, d'échange et de coopération. Quant au neurone, ce sont sa complexité et son mystère qui m'attiraient. Bon, je reconnais qu'au moment où j'ai franchi la porte du bâtiment P d'électronique de l'ESPCI, je ne connaissais rien aux réseaux de neurones. Il m'a fallu un peu de temps pour réaliser qu'un neurone formel n'était rien d'autre qu'une entité mathématique, et que j'étais en train de faire de la modélisation par apprentissage et pas de la biologie. Toutefois, comme le répète Gérard à l'envi, les réseaux de neurones formels sont des candidats hors pair pour réaliser des modèles d'apprentissage, et c'est là que l'on retrouve une similitude fonctionnelle avec les neurones biologiques, enfin quoi, les neurones. Ainsi, un réseau de neurones formels apprend à partir d'informations qui lui sont fournies (qui constituent les entrées) en modifiant la valeur des paramètres de son architecture (les synapses) pour faire coïncider autant que possible la valeur de la propriété calculée par le réseau (la sortie) avec une valeur connue, souvent mesurée expérimentalement, et stockée en mémoire. Passé cette étape d'apprentissage, de nouvelles informations peuvent être traitées par le réseau, c'est à dire classifiées, comparées, et bien sûr prédites... Comme la tâche à réaliser est souvent complexe, un seul neurone ne peut faire le travail, et il faut en utiliser plusieurs qui sont organisés en réseau, donc connectés.

Peut-on réellement prévoir les propriétés physicochimiques à partir de la seule structure moléculaire?

Joker... Je plaisante ; oui, certainement, si l'on dispose de suffisamment de données expérimentales pour la propriété en question, et ceci pour des molécules de structures pas trop différentes de la molécule cible. Exemple : je veux prédire le point de fusion, disons, du bupropion (zyban ou wellbutrine). Je cherche sa structure et comme je dispose d'une base de points de fusion de molécules organiques assez étoffée, je suis en mesure d'avoir rapidement une prédiction relativement précise du point de fusion du bupropion, une substance classée comme antidépresseur, et qui semble relativement efficace dans le sevrage tabagique. Si je souhaite maintenant effectuer des prédictions sur son activité antitabagique, il va me falloir constituer une base de données de molécules ayant cette propriété, et surtout quantifier leur efficacité pharmacologique, ce qui risque d'être plus délicat.

Qu'est-ce que la Chimie Combinatoire virtuelle?

Faisons simple : nous savons tous que trouver une molécule active pour une pathologie donnée représente une étape longue, difficile et coûteuse. L'idée est donc de générer une base de données virtuelle de composés, puis de prévoir l'activité désirée pour les éléments de cette base en utilisant des réseaux de neurones. L'étape ultime consiste en la synthèse proprement dite des candidats les plus prometteurs, suivi de tests réels de leur activité. Cela suppose bien sûr de disposer déjà de molécules actives et d'avoir pu préalablement quantifier leur activité. Ce qui est intéressant dans cette démarche, c'est de trouver les descripteurs caractéristiques (l'association de plusieurs atomes ou groupements chimiques, leur proximité par exemple) qui vont engendrer une activité donnée. Le bupropion (zyban) et la fluoxétine (prozac) sont tous deux classés comme antidépresseurs mais ont toutefois des structures chimiques

légèrement différentes ; ce que nous envisageons, c'est de construire par combinaison une structure « hybride », qui nous l'espérons, aurait aussi une activité intéressante. Cette démarche s'apparente à la recherche de mots pertinents dans un contexte dont parle Gérard.

Quelle est ton activité d'enseignement?

La chimie organique reste mon activité principale dans ce domaine, sous forme de travaux dirigés, de préceptorats et de travaux pratiques. J'ai pu finalement y apporter une note personnelle avec l'introduction de la recherche bibliographique informatisée dans des bases de données de chimie, et la modélisation de spectres RMN, de spectres IR ou UV-Vis. J'encadre actuellement des stages de projet de recherches de troisième année sur le thème qui me tient à cœur, la prédiction de propriétés physico-chimiques en utilisant des réseaux de neurones.

Comment équilibres-tu tes activités de Recherche et celles d'enseignement?

C'est assez naturel, puisque les semaines sont rythmées par les plages d'enseignement des différentes promotions présentes à l'école ; d'ailleurs cela correspond à toutes si l'on tient compte des mastères et des DEA (diplôme d'études approfondies) qui se font en dernière année. Bien sûr, il y a des moments de petite frustration lorsque les périodes d'enseignement sont très chargées et que le temps disponible pour s'occuper de sa recherche manque.

Où trouves-tu tes principales satisfactions? En a-t-il toujours été ainsi?

A une époque, j'étais à la paillassse du laboratoire de chimie organique lançant plusieurs manipulations dans la même journée. Aujourd'hui la plupart de mon travail est informatique. Les calculs se font sur ordinateur, les rapports, présentations et publications se préparent sur ordinateur, et même la lecture (avec un grand écran...) des revues scientifiques sur ordinateur se démocratise. Cela me plaît, mais j'apprécie toujours autant le contact avec les élèves, et les discussions avec les autres collègues qui souvent concernent... des problèmes d'ordinateurs.

Quel est le côté ingrat?

Le voilà justement le côté ingrat : ce sont les ordinateurs qui ont envahi notre vie quotidienne, et la logistique de la ville de Paris qui n'a pas suivi. Un seul informaticien pour une école comptant plus de 1000 machines diverses pour 1300 utilisateurs : le compte n'y est pas. Alors les laboratoires s'organisent et les responsables informatiques locaux (RIL) qui y sont désignés ont du grain à moudre. Vous l'avez deviné, je fais partie de cette nouvelle race. Lutter contre les virus, recharger les systèmes d'exploitation, créer les comptes, s'occuper du site web du laboratoire, appliquer les patches (rustines), les mises à jour, ou encore comprendre pourquoi une machine démarre en faisant un bruit de casserole, voilà le lot quotidien d'un RIL.

Peux-tu nous décrire une journée typique pour toi?

Bon, comme Jean-Luc ou Gérard, je passe souvent beaucoup de temps devant mon principal outil de travail, l'ordinateur. Cela commence en dehors des périodes d'enseignement par la lecture des courriels du jour, puis par des travaux de rédaction ou de préparation d'exercices pour nos chers élèves ou encore de corrections diverses, pour n'en citer que quelques-uns. S'y glissent les discussions sur la recherche, les conférences, les réunions diverses propres à l'ESPCI. Finalement la journée est assez variée, et les journées se suivent sans se ressembler.

Tu as travaillé en 1992 à l'Université de Californie à San Diégo. Comment t'y es-tu pris pour obtenir ce poste?

Ce qui est remarquable, c'est qu'à PC un jeune enseignant peut se permettre une telle expérience, et cela est

formidable ! Pour trouver mon stage post-doctoral, j'ai tout d'abord recherché dans quel pays je voulais vraiment aller avec ma famille. J'ai hésité entre le Japon et les Etats-Unis un certain temps. Par peur d'un trop grand décalage culturel, j'ai décidé de compléter mon cursus dans un pays avec lequel j'avais pris par ailleurs contact à plusieurs reprises. Ma passion pour la bibliographie informatisée a alors repris le dessus, et j'ai cherché un laboratoire d'accueil en tapant sur mon clavier quelques mots-clés bien choisis. L'examen attentif de la carrière de mon futur patron, suivi d'échanges de courriels (c'était le début, et c'était compliqué...), voilà ce qui m'a finalement décidé. Encore de la modélisation... Inutile de dire que lorsqu'on arrive avec un salaire, c'est beaucoup plus facile de trouver un coin de paillasse outre-atlantique ; et une fois que l'on a fait ses preuves, rien n'empêche de demander un petit complément !

Quel souvenir en gardes-tu?

Exceptionnel, et je tiens à remercier mes collègues enseignants qui m'ont permis de mener à terme une expérience d'une année très riche sur les plans scientifiques et humains. La concentration d'universités, d'instituts scientifiques et de bibliothèques permettait en fait de passer ses journées à s'instruire soit en assistant à des conférences, soit en fouinant dans les kilomètres de travées de livres divers ! Ce que j'ai sans doute le plus apprécié étaient la liberté d'entreprendre et surtout la facilité de collaborer. Bien entendu, mon niveau en anglais s'est élevé à des sommets jamais atteints !

Depuis 1999, tu es examinateur en Chimie du concours X-PC. Pourquoi?

Jacques Duran (81), notre ancien directeur des études a fait un appel au peuple lorsque notre concours d'entrée a changé. PC avait un savoir-faire très solide dans le domaine de la chimie au concours, et le Directeur du concours de Polytechnique nous faisait confiance. Personnellement, j'étais déjà examinateur en travaux pratiques de chimie du concours Chimie-Centre, et j'avais

bien envie de tenter la "riche" aventure, qui je le savais, représentait un gros investissement. La chimie continue à me passionner, et j'ai naïvement pensé qu'une remise à niveau en potassant les nouveaux programmes de classes préparatoires me permettrait de mieux assurer le suivi des élèves après leur intégration à PC. Cela n'était pas faux, et j'ai effectivement réalisé la mise à jour de mes connaissances de taupin. Quelques spécialités de la chimie étaient nouvelles, mais la plupart du temps, il m'a suffi d'apprendre une nouvelle terminologie et d'examiner des concepts similaires sous un angle différent.

Tu es rentré à PC en 1978. À 25 ans de distance, comment se comparent les candidats?

Personnellement, je suis content d'avoir intégré PC en 1978 : je ne sais pas si j'y arriverais aussi bien aujourd'hui ! Les programmes préparatoires sont très chargés et le concours d'entrée est d'un niveau plutôt relevé. La plupart des candidats sont très motivés, et portent à mon avis un intérêt plus précoce à la recherche que par le passé. Nous étions en 1978 une école majoritairement tournée vers l'industrie et l'enseignement nous préparait à la recherche surtout dans la dernière année de la scolarité. Aujourd'hui dès leur arrivée, les élèves sont baignés dans une ambiance "recherche" et commencent leur scolarité par un court séjour dans les laboratoires, en contact direct avec un chercheur.

Si c'était à refaire, que referais-tu différemment?

Et bien, voilà une question embarrassante, et il m'est dur d'y répondre. Il suffit de peu de choses pour changer le cours d'une carrière : un échec à un concours, un stage postdoctoral différent, une opportunité qui n'a pas été saisie... J'ai envie de dire que je ferais la même chose, mais cela n'est pas réaliste ! Pour faire plaisir à Gérard, je peux dire regretter de ne pas avoir utilisé les réseaux de neurones quelques années plus tôt.

Merci !

Isabelle Guyon (100) : sa conquête de l'Ouest !



La Californie ! Berkeley ! Beaucoup d'entre nous en ont en rêvé. Isabelle l'a fait. Elle s'est construit là bas par ses seuls mérites une réputation sur un créneau high tech hyper pointu.

Aujourd'hui entrepreneur indépendante, elle télétravaille avec succès à partir de son domicile, entourée de la famille franco-helvético-américaine qu'elle a fondée là bas avec son époux Bernhard Boser, natif de Zurich et professeur à Berkeley.

Au-delà de sa « success story » personnelle, l'entretien qu'elle a accordé à 10rueV@uquelin stimule la réflexion.

Notamment à propos du « brain drain ». Combien de temps notre chère vieille Europe pourra-t-elle encore financer par son travail à la fois 1) l'intégration de ceux des immigrants peu ou pas formés, voire peu ou pas alphabétisés, attirés par sa protection sociale enviable et 2) la formation suivie d'émigration de certains de ses plus brillants jeunes chercheurs-entrepreneurs ?

Professionnellement, de quoi es-tu la plus fière?

D'avoir réussi à continuer ma carrière tout en élevant trois enfants ! Je suis toujours très fière que mes collègues continuent à me prendre au sérieux malgré le ralentissement de mes activités. Il m'est arrivé de recevoir des clients avec deux petits mômes dans les bras (j'ai eu des jumeaux), ça ne les a pas refroidis.

Tu as inventé les « Machines à Vecteurs Supports ». Qu'est ce que c'est?

Il ne s'agit pas de "machines" à proprement parler. C'est une technique statistique semblable aux réseaux de neurones qui permet de construire des modèles à partir d'exemples. L'originalité de la méthode est sa capacité de se focaliser sur les exemples les plus informatifs, les fameux "vecteurs supports".

Lors d'un apprentissage par l'exemple, il n'est pas nécessaire de mémoriser les exemples faciles à prédire. Les machines à vecteurs supports (les SVMs – Support Vector Machines - en anglais) ne retiennent que les exemples ambigus, les cas limites, marginaux. Dans un problème de reconnaissance des formes, il s'agira des exemples de classes différentes se ressemblant beaucoup, tel qu'un "C" presque fermé ressemblant à un "O".

L'ancêtre des SVMs est le « classifieur linéaire de marge optimale », inventé par le célèbre mathématicien russe Vladimir Vapnik dans les années soixante. Les vecteurs supports étaient les exemples marginaux, c'est-à-dire situés sur la marge de chaque côté de la frontière entre les classes. L'algorithme d'apprentissage était un processus d'optimisation efficace conduisant à une solution unique maximisant la marge, c'est-à-dire qui établit une frontière aussi éloignée que possible des exemples appartenant à des classes différentes. Limité à

l'apprentissage de modèles linéaires, l'algorithme n'avait pas attiré beaucoup d'attention. En 1991, mon futur époux Bernhard Boser et moi-même avons généralisé l'algorithme à des modèles non linéaires.

Rétrospectivement, quel a été l'élément déterminant à l'origine de cette invention?

C'est une histoire très amusante. Tous les ingrédients de l'invention existaient depuis trente ans. De plus, ils avaient été inventés au même endroit, à Moscou, dans les années soixante. D'une part, Aizerman, Braverman et Rozonoer avaient mis au point la méthode de noyaux dite des "fonctions de potentiel" et d'autre part, Vapnik avait inventé le classifieur linéaire de marge optimale. Ces deux méthodes de reconnaissance des formes, ancêtres des réseaux de neurones, étaient en compétition. Leurs inventeurs n'eurent jamais l'idée de les combiner.

Pour ma thèse, sous la direction de Gérard Dreyfus et de Léon Personnaz, j'avais utilisé une astuce algorithmique connue qui a pris le nom de "kernel trick". Cela permet de transposer des algorithmes conçus pour des modèles linéaires à des méthodes de noyaux non linéaires. Nous avons appliqué ce "kernel trick" aux réseaux de neurones de Hopfield. En face, à l'École Normale, Marc Mézard et Werner Krauth inventaient un réseau de Hopfield de stabilité optimale, similaire à l'algorithme de Vapnik. Dans la hâte de finir ma thèse, je négligeai d'essayer l'algorithme de Mézard et Krauth et de lui appliquer le "kernel trick". C'était en 1989.

Je suis partie aux Etats-Unis faire un post-doc aux laboratoires Bell. Là, mes collègues firent pression sur moi pour m'associer à l'effort commun de développement des réseaux de neurones. J'abandonnai temporairement mes recherches sur les méthodes de noyaux. Cependant, en 1991, Vladimir Vapnik rejoignit notre groupe aux Laboratoires Bell. Bernhard Boser et moi nous sommes intéressés à son algorithme de marge optimale, qui me rappelait l'algorithme de Mézard et Krauth. Bernhard s'est lancé dans l'implémentation de l'algorithme de Vapnik. Je lui ai suggéré d'appliquer le "kernel trick". Les SVMs étaient nées!

10 ans plus tard, quel développement a-t-elle connu?

De façon intéressante, l'application qui a fait connaître les SVMs, la classification de documents, utilisait des SVMs linéaires. Donc, notre papier a popularisé l'algorithme original de Vapnik qui s'est révélé très efficace dans de nombreuses applications. C'est le cas surtout de problèmes possédant des dizaines de milliers, voire des centaines de milliers de variables, mais pour lesquels nous avons au plus une centaine d'exemples. Les SVMs se sont ensuite illustrés dans les applications de diagnostic biomédical fondé sur des données de génomique ou de protéomique.

L'algorithme original était conçu pour résoudre des problèmes de classification, comme les problèmes de diagnostic médical : le patient est-il malade ou sain? Il a été généralisé depuis par d'autres chercheurs à la régression, à l'estimation de distributions de probabilité, et à la résolution de systèmes d'équations d'opérateurs linéaires. L'intérêt porté aux SVMs a attiré l'attention sur les méthodes de noyaux en général. Beaucoup de nouveaux algorithmes ont vu le jour.

De nombreux résultats théoriques sont venus étayer les résultats empiriques, prouvant que les méthodes de "marge large" ont de bonnes performances de prédiction sur des exemples n'appartenant pas à l'ensemble d'apprentissage. Des connexions ont été faites avec les méthodes de régularisation permettant de résoudre des problèmes dits "mal posés". Les SVMs sont maintenant un membre d'une riche famille de méthodes d'optimisation.

Donc, il m'est de plus en plus difficile de suivre tous les développements. De très éminents mathématiciens et statisticiens travaillent sur le sujet. Je me sens toute petite, je n'ai produit que l'étincelle qui a provoqué l'incendie!

Tu as travaillé 6 ans au sein des prestigieux AT&T Bell Laboratories. Quel souvenir en gardes-tu?

C'était une expérience exceptionnelle. J'ai eu la chance de visiter les Laboratoires Bell lorsque j'étais étudiante, grâce à Gérard Dreyfus, mon directeur de thèse. J'ai souhaité y faire un stage, puis y retourner pour un post-doc. J'étais fascinée par mes collègues, si brillants, et si "grands enfants". Ils étaient capables d'envoyer de vieilles tranches de pizza par la poste ou de faire un lâcher de sacs-poubelles gonflés à l'hélium depuis le toit du bâtiment. Mais nous travaillions énormément, il fallait toujours avoir de nouveaux résultats "excitants" et l'ambiance était très compétitive.

Quand je suis arrivée, nous avions encore beaucoup de liberté pour choisir nos sujets de recherche, dès lors qu'on pouvait prédire un impact important en cas de succès. Si j'avais envie de me lancer sur un nouveau sujet, il y avait toujours un collègue expert dans ce domaine prêt à répondre à mes questions ou à collaborer. Malheureusement, progressivement, des pressions se sont fait sentir. Je suis partie au moment de la scission d'AT&T, à la faveur d'une offre de départ volontaire. La plupart de mes collègues sont restés avec AT&T Labs, mais depuis, l'équipe s'est dispersée. Nous restons cependant en contact et nous nous retrouvons avec plaisir lors de conférences.

Quel a été ton rôle dans l'apparition des premiers "pen computers" (Palm Pilot)?

J'ai un peu l'âme d'une pionnière. Ma thèse était l'une des premières sur les réseaux de neurones, dans les années 80. De même, j'ai commencé à travailler sur l'écriture en tant qu'interface d'ordinateur au tout début des "pen computers", avant même que les premiers ne fassent leur apparition sur le marché. Le premier « touch pad » d'AT&T était équipé de mon réseau de neurones (ce qui m'a valu mon emploi permanent à AT&T). Sous la pression des équipes marketing d'AT&T, nous avons créé un système ambitieux de reconnaissance de l'écriture sans contrainte de style ou de scripteur. Cependant, gardant un œil critique sur notre propre recherche, nous avons organisé une étude d'utilisabilité, qui a révélé que, pour obtenir une bonne interface d'ordinateur, le taux d'erreur toléré doit être très bas et la reconnaissance immédiate. Cela rendait notre système (et tous les autres alors sur le marché) essentiellement inutilisables. À la même époque, j'ai visité Goldberg, un collègue de Xerox, qui m'a présenté ses travaux sur l'alphabet simplifié « unistroke » : chaque lettre était formée d'un seul trait. J'ai immédiatement compris que c'était une meilleure voie. AT&T était bien loin de se lancer dans cette direction. Avec un petit groupe d'amis, nous avons commencé à étudier la possibilité de démarrer une startup. Un mois plus tard, le premier Palm Pilot sortait. Nous nous étions fait prendre de vitesse. Nous avons abandonné nos projets. Donc, j'ai la fierté d'avoir deviné avant beaucoup d'autres que le Palm Pilot serait un succès, mais la frustration de ne pas y avoir directement contribué.

Qu'est ce que la "computational linguistic"?

C'est la modélisation du langage avec des méthodes statistiques. Par exemple, il est possible de créer une machine à parler l'anglais qui engendre des phrases aléatoires bien formées et qui ont plus ou moins un sens, en utilisant seulement des statistiques de la langue. Un des modèles les plus simples, dit de "trigramme", calcule la probabilité qu'un certain mot apparaisse dans une phrase, étant donnés les deux mots précédents. De tels modèles sont utiles en reconnaissance de la parole ou de l'écriture. Ils permettent de guider la reconnaissance en limitant le champ des possibilités.

Tu as aussi appliqué les « Machines à Vecteurs Supports » à certains domaines de la biologie. Que retiens-tu de cette activité?

C'est en fait à l'heure actuelle mon activité principale. En effet, les analyses médicales sont en passe de changer radicalement. Auparavant, des analyses très coûteuses ne

portaient que sur une seule ou un faible nombre de protéines du sérum sanguin. Avec l'avènement d'instruments biomédicaux comme les puces à ADN ou les spectromètres de masse, il est possible de mesurer l'abondance de dizaines de milliers de protéines (ou de leurs précurseurs) simultanément. Ces données se prêtent bien à l'utilisation des SVMs parce qu'on ne dispose généralement que de fort peu d'échantillons de sérum de patients malades et de patients sains. Donc, il est particulièrement important d'extraire les exemples les plus informatifs. Un autre exemple industriel important est l'identification de molécules pharmacologiquement actives. La chimie combinatoire permet d'engendrer une grande variété de molécules et de vérifier *in vitro* leur activité. Ce processus peut être optimisé en sélectionnant à l'avance les molécules les plus prometteuses. Une SVM peut "apprendre" à classifier les molécules actives ou inactives à partir d'exemples. Ces exemples sont des molécules d'activité connue, représentées par des dizaines ou des centaines de milliers de descripteurs physico-chimiques.

Depuis 1996 tu es consultante indépendante. Dans quels domaines?

Mes premiers clients ont exploité mes compétences en matière de reconnaissance de l'écriture. J'ai travaillé sur des essais de stylos équipés d'accéléromètres pour Ricoh, et j'ai développé un système d'enseignement de l'écriture cursive pour une startup. La reconnaissance des formes a continué à me fournir des clients, mais j'ai eu la chance d'être attirée dans une startup de bioinformatique grâce à Vladimir Vapnik. L'idée était d'appliquer les SVMs aux données des puces à ADN. L'idée était bonne, mais personne ne savait comment faire. J'ai proposé un algorithme qui a eu beaucoup de succès et a contribué à rassembler 6 millions de dollars de capitaux. À la suite de ça, le fondateur de la boîte m'a demandé de recruter une équipe des meilleurs chercheurs en SVMs. Malheureusement, la mauvaise gestion et l'éclatement de la bulle spéculative ont été fatals à l'entreprise. Mais je suis restée dans le domaine bioinformatique et je continue de consulter pour des startups qui mettent au point de nouveaux instruments biomédicaux. Je fais autant de caractérisations d'instruments que d'applications des SVMs.

Comment ça marche?

C'est pour moi l'activité idéale, qui me permet de travailler à la maison avec des horaires flexibles, et donc de me consacrer aussi à mes enfants. Je visite mes clients régulièrement, mais je travaille essentiellement seule, avec très peu de directives, sur des données que l'on me fournit. J'ai d'habitude plus de travail que je ne peux en faire. De façon à ne pas être trop sous pression, je travaille à l'heure.

J'aime aussi garder un peu de temps pour rester en contact avec la recherche. J'organise chaque année un workshop dans le cadre d'une grande conférence de réseaux de neurones (Neural Information Processing Systems), sur des thèmes variés.

Comment développes-tu ta clientèle?

J'obtiens très rarement des contrats grâce à mon site Web (NDLR : il vaut pourtant le détour ! <http://www.clopinet.com/>). Au début j'envoyais des brochures, ça ne m'a jamais rien rapporté. Mes clients me trouvent par recommandation d'anciens collègues ou d'anciens clients. Grâce aux SVMs, j'ai maintenant une certaine notoriété qui m'est aussi utile.

Quel futur envisages-tu pour ton activité?

Je ne souhaite pas embaucher, je préfère rester consultante indépendante. Il est certain que je n'en deviendrai pas riche (à moins que les actions d'une des entreprises pour lesquelles je suis conseillère scientifique prennent beaucoup de valeur). Pour l'instant, j'aime beaucoup mettre la main à la pâte et programmer moi-même les solutions que je préconise. Mais je ne vends aucun produit tangible, logiciel ou autre, je ne vends que le produit de ma pensée, d'où le nom de ma compagnie : Clopinet. Certains me provoquent en insinuant que je travaille pour des clopinettes, mais ce n'est pas vrai!

Je crains que bientôt mes neurones qui se rouillent ne me permettent plus d'être aussi efficace que de jeunes programmeuses ou programmeurs. Je vais progressivement agir davantage comme ingénieur conseil.

Comment gères-tu ton temps entre le job et le reste?

Je suis assez stressée pour tout dire, en dépit de l'optimisation de mon emploi du temps. Mon mari est professeur d'électronique à l'université de Berkeley. La plupart des épouses de professeurs sont femmes au foyer, c'est en fait ça qui est le plus difficile à gérer. Mais, malgré tout, je suis assez fière parce que j'ai évité de donner raison à un enseignant mauvaise langue qui, jadis, avait dit à un petit groupe de PCennes: "Vous les filles, vous prenez la place des garçons. Qu'est-ce que vous ferez de votre diplôme plus tard ? Vous l'accrocherez dans votre cuisine ?".

Tu pratiques le télétravail de façon extensive. Peux-tu nous en parler?

Oui, cela fait bien longtemps. Quand mon mari a obtenu son poste à Berkeley, en Californie, j'ai gardé le mien aux laboratoires Bell, sur la côte Est, tout en l'accompagnant sur la côte Ouest! Je travaillais donc à distance et je rendais visite à mon groupe environ une semaine sur cinq. Cela a duré trois ans. Ensuite, les allers-retours devenant incompatibles avec notre désir d'avoir des enfants, j'ai créé ma petite compagnie de consulting. Avec l'expérience du télétravail, il m'a été possible d'obtenir des clients aux quatre coins du continent et quelquefois outre-Atlantique. Il y a beaucoup de gens avec qui j'ai travaillé longtemps sans les avoir rencontrés. Un jour, après avoir travaillé deux ans avec un collègue par courrier électronique, nous nous sommes enfin rencontrés. J'avais imaginé un personnage plein d'assurance, un peu plus âgé que moi, la barbe peut-être. J'ai eu la surprise de voir venir un très jeune homme timide avec de longs cheveux.

Tu es née à Paris. Ton époux est suisse. Vous vivez aux USA depuis 15 ans. Avez-vous plutôt piloté ou plutôt subi cet itinéraire?

Je suis partie pour un an de post-doc. Mon parcours idéal était alors d'entrer au CNRS et de faire de la recherche fondamentale. Mais j'ai rencontré mon mari, un collègue de travail. Au début, nous parlions toujours de rentrer. Au bout d'un certain temps, nous avons eu l'impression désagréable de n'être plus de nulle part, de ne plus savoir où était notre chez nous. Puis, finalement, nous avons élu domicile en Californie, et c'est maintenant notre chez nous. Nos enfants ont trois nationalités et parlent trois langues.

On comprend bien l'enchaînement des circonstances qui ont présidé à ton choix californien. A l'inverse, que faire pour qu'un jour une partie de l'élite intellectuelle formée dans le pays le plus riche du monde ait des raisons de venir s'établir en Europe ? Existe-t-il la moindre chance pour rééquilibrer les flux de part et d'autre de l'Atlantique ?

Personnellement, je pense qu'il ne me serait pas possible de trouver aussi facilement du travail de consulting en Europe. Je travaille surtout avec des startups, ici c'est un véritable vivier de startups. Avec la faillite mondiale des labos de recherches dans le secteur public et les grandes entreprises (les Etats-Unis n'y ont pas échappé), la recherche se fait maintenant, curieusement, dans les startups. Je crois que le dynamisme des petites entreprises est un facteur clé, le système Américain favorise l'esprit d'entreprise. Evidemment, c'est de la recherche appliquée. Comment faire de la recherche fondamentale dans ce contexte défavorable? Actuellement, surtout en milieu universitaire. Mais, les crédits de recherches sont de plus en plus difficiles à obtenir, même aux Etats-Unis. J'ai peur que la recherche ne devienne un sacerdoce ou un "hobby".

Tu es résidente permanente aux USA. Tu as la double nationalité suisse et française. Le regard que tes relations américaines portent sur toi a-t-il changé depuis l'Irak?

J'ai la chance d'habiter à Berkeley. Ici la majorité des gens sont démocrates et opposés à la guerre contre l'Irak. Donc nous n'avons pas ressenti d'hostilité, au contraire. Berkeley est un microcosme. Un grand nombre de gens vivent encore dans le style des années 70. On croise souvent des hommes en costard avec une queue de cheval. Je me sens très à l'aise ici.

Ton site fait référence au flamenco. Quelle est sa place dans ta vie ?

De plus en plus restreinte. Mes enfants et mon mari n'aiment pas ça, je ne suis plus très motivée. J'ai commencé

parce qu'un collègue qui joue de la guitare flamenco m'avait demandé de mettre au point en un temps record un spectacle avec mes collègues des Bell Labs pour un "talent show" à une conférence sur les réseaux de neurones. Les cinq chercheuses des Bell Labs se sont donc produites... avec un certain succès! Je cherche maintenant des activités que je peux faire avec mes enfants. Je me suis remise au judo que je pratiquais quand j'étais adolescente. Les petits aiment bien ça, ils montent en grade, et moi, j'ai regagné mon grade de ceinture marron. J'espère bien être ceinture noire d'ici un an ou deux. Je vais aussi essayer de me remettre à la peinture avec ma fille qui est très douée pour les arts plastiques.

Merci !