

07

SCIENCES ET TECHNOLOGIES DE L'INFORMATION (INFORMATIQUE, AUTOMATIQUE, SIGNAL ET COMMUNICATION)

Président de la section
François PIERROT

Membres de la section
Philippe BAPTISTE
Nicole BIDOIT-TOLLU
Prosper CHEMOUIL
Annick CHOISIER
Bruno DURAND
Cécile DURIEU
Christophe FONTE
Bruno JOUVENCEL
Rogélio LOZANO*
Isabelle MAGNIN
Éric MOULINES*
Christian MICHEL
Hélène PAUGAM-MOISY
Bernard ORIOLA
Henri PRADE
Luc PRONZATO
Isabelle QUEINNEC
Michel RIVEILL
Dominique ROSSIN
Éric SANLAVILLE

(*) *Remplacés, en avril 2007, par*
Gérard BAILLY
Michel MALABRE

1 – LES GRANDS DOMAINES DE LA SECTION 07

La section 07 regroupe plusieurs domaines scientifiques : l'informatique, l'automatique, le traitement du signal et de l'image, la robotique et l'électronique intégrée, qui ont tous en commun de se rapporter à une entité, l'information, dont le traitement scientifique, devenu capital aujourd'hui, a débuté il y a à peine plus de 70 ans. Chacun de ces domaines est en lui-même très varié, et recouvre de multiples disciplines connexes et complémentaires. Par ailleurs, ces domaines sont en interaction avec d'autres disciplines scientifiques dans des perspectives d'enrichissement mutuel ou d'applications : c'est le cas en particulier avec les mathématiques, les sciences de la vie et les sciences de l'homme et de la société.

Schématiquement, le terme informatique désigne tout ce qui concerne la science du calcul sur machine, de l'architecture des calculateurs aux langages de programmation, les réseaux informatiques, les systèmes d'accès à l'information, la communication homme-

machine, mais aussi des disciplines transversales comme l'algorithmique ou la théorie des graphes, ainsi que des disciplines qui exploitent l'information dans différentes perspectives comme l'intelligence artificielle, la recherche opérationnelle ou la bioinformatique par exemple. Le traitement du signal, de la parole, de l'écrit et de l'image pose des problèmes de codage, de reconnaissance de formes, de synthèse, mais aussi d'analyse et de modélisation des contenus dont l'interprétation diffère selon la modalité physique d'acquisition. L'automatique s'intéresse à la modélisation, à la commande et à la supervision de systèmes dynamiques. La robotique est la science de la perception, du mouvement et de leur intégration en une machine physique. La microélectronique s'intéresse à la conception et au test des circuits intégrés envisagés comme des systèmes à part entière.

2 – CONTEXTE GÉNÉRAL : UNE PRIORITÉ INTERNATIONALE

La création en 2000 du département STIC au CNRS, dont la section 07 était un des deux piliers, avait enfin marqué la reconnaissance de la spécificité de la problématique commune aux grands domaines cités plus haut.

Cette création, et l'élan donné à cette communauté scientifique au sein du CNRS à partir de 2001 ont été certes profitables (*Voir* par exemple le § 3.5), mais ils sont intervenus tardivement en comparaison d'autres évolutions nationales et avec une ampleur restée modeste à l'échelle internationale.

En effet, sur le plan national, on peut se rappeler que « l'essor » de l'INRIA date déjà des années 1980 et que, entre 2000 et 2003, ses effectifs ont crû de plus de 50 % (1). On peut se rappeler également que le programme d'action gouvernemental « *Préparer l'entrée de la France dans la société de l'information* », qui

ne pouvait pourtant pas être considéré comme exagérément précurseur, a été rendu public le 16 janvier 1998. On doit, enfin, garder à l'esprit le formidable essor des STIC dans les Universités et Établissements d'Enseignement Supérieur, et les recrutements massifs d'étudiants et d'enseignants/chercheurs qui y ont contribué : les formations universitaires et supérieures, en Informatique, mais aussi dans le domaine Télécommunications et Réseaux ou en Automatique, étaient déjà largement en place dans les années 1990.

À l'échelle internationale, la place éminente des STIC dans les priorités de recherche et de développement économique est attestée par les choix volontaristes de nombreux pays (2) (Japon, Corée, USA, Suède, etc.).

Ces priorités ne sont pas que des affichages, comme le montrent à la fois le montant et l'évolution des budgets de R&D des grands pays industrialisés (3). Ainsi les dépenses pour la R&D en STIC restent nettement plus faibles en France qu'aux USA et au Japon, qu'elles soient exprimées en \$PPP (4) par habitant ou en % du PIB. En Europe, les efforts sont très hétérogènes avec une variation d'un pays à l'autre dans un rapport de 1 à 13 (Italie : 38 \$PPP ; Finlande : 483 \$PPP).

Tableau 1. Efforts budgétaires de R&D

	Japon	USA	Corée	France	Europe
\$PPP	270	250	221	121	84
% PIB	0,92	0,62	1,15	0,42	0,27

Le domaine des STIC est le seul où on observe un écart majeur entre Europe et USA ; cet écart est de plus de 30 milliards \$PPP en considérant l'ensemble des budgets publics et privés, et il s'est accru entre 1999 et 2005.

L'écart est moins grand en ce qui concerne les seules politiques publiques, mais demeure nettement en faveur des USA : en effet, si l'effort du budget public civil européen est deux fois plus grand qu'aux USA, le budget public militaire y est dix fois plus petit. Bien

entendu, ce budget public militaire a un impact notable dans les centres de recherche académiques, y compris dans ceux dont la vocation n'est pas la recherche militaire, et participe ainsi à l'essor de ces centres. À titre d'exemple anecdotique, le fameux Institut de Robotique de Carnegie Mellon (Pittsburgh, USA) a reçu environ 25 M\$ de crédits d'origine militaire en 2006 (sur un total d'environ 65 M\$), et dans le même temps s'est vu confier un programme de recherche d'un montant de 15 M\$ par la NSF sur le thème des technologies d'assistance aux personnes âgées et dépendantes : de nombreuses technologies issues des recherches en STIC peuvent ainsi être vues comme des « technologies duales » et c'est bien alors l'effort national global qui doit être considéré.

Enfin, les budgets publics évoluent positivement partout, mais avec une ampleur variable : la France est ici légèrement en retrait de la moyenne des grands pays.

Tableau 2. Évolution des dépenses publiques de R&D STIC entre 1999 et 2005 pour dix pays majeurs

Suède	Corée	G.-B.	USA	Espagne
144 %	111 %	83 %	68 %	68 %
Moyenne des 12 plus grands pays : 58 %				
France	Japon	Canada	Finlande	Europe
42 %	38 %	37 %	31 %	27 %

Compte tenu de ces évolutions des dépenses publiques, et des efforts privés correspondants, derrière les *leaders* que sont les USA et le Japon dont l'effort de R&D en STIC représente respectivement 45 % et 22 % de l'effort mondial, on trouve maintenant la Corée avec près de 8 %. Les deux pays suivants dans le classement des pays qui pèsent au plan mondial sont l'Allemagne (moins de 6 %) et la France (moins de 5 %), mais, au contraire de la Corée, ces deux pays n'ont pas vu leur part de l'effort mondial sensiblement évoluer entre 1999 et 2005.

Aujourd'hui, la priorité « STIC » est toujours d'actualité en France, mais le Département STIC du CNRS n'existe plus, et la section 07 fait partie d'un Département qui regroupe également les sections 8, 9, et 10, formant un ensemble dont l'orientation générale est plus large et concerne les « Sciences et Technologies de l'Information et de l'Ingénieur ». La manière d'orchestrer la priorité STIC au sein du CNRS doit donc maintenant être réinventée.

3 – LABORATOIRES ET PERSONNELS DE LA SECTION 07

Ce chapitre brosse un rapide tableau de la section 07, vue sous l'angle de sa structuration en unités de recherche, et donne quelques points de repère sur les personnels, en type, en nombre, et quant à leur recrutement. Bien entendu, l'ensemble de la communauté de recherche qui s'intéresse aux thématiques couvertes par la section 07 est plus large que celle qui travaille dans les laboratoires évalués par cette section, et le tableau 3 donne une estimation des nombres de chercheurs pour les principaux organismes concernés. Il ne s'agit ici que des chercheurs permanents, qui encadrent plusieurs milliers de doctorants et qui collaborent quotidiennement avec de nombreux personnels techniques et administratifs, permanents ou temporaires.

Tableau 3. Chercheurs permanents dans le domaine couvert par la section 07

CEA LIST	CNRS Section 07	INRIA	GET	Univ. CNU 27 et 61	Total
~ 400	~ 500	~ 500	~ 250	~ 2 900 et ~ 1 500	~ 6 000

3.1 CHERCHEURS

Les chercheurs CNRS de la section 07 sont principalement répartis dans 100 laboratoires (5), dont 60 dépendent en priorité de la section 07 (Voir Tableau 4). Ces unités, qui sont quasiment toutes des UMR, accueillent 90 % des chercheurs CNRS et 80 % de l'ensemble du personnel chercheur : il sera très souvent question de ces unités « section 07 principale » ici.

Tableau 4. Répartition des structures relevant de la section 07

	Toutes structures	Section 07 principale
FR/IFR	6	3
FRE	14	9
GDR	20	11
UMI	3	2
UMS/UPS	4	1
UMR	80	47
UPR	3	2
Total	130	75
Dont labos	100	60

Notons que depuis 2001, l'actualité de création, transformation et suppression d'unités a été très chargée : sur cette période, par exemple, plus de 20 FRE ont été créées, puis fermées au titre de la section 07 principale, conduisant à la création d'une dizaine UMR (ceci, en plus des changements de statut, ou de périmètre, des unités existant antérieurement).

Les chercheurs CNRS sont toujours associés à un grand nombre de collègues, qu'ils soient chercheurs « non CNRS » ou enseignants chercheurs. Dans les unités qui relèvent prin-

cipalement de la section 07, la proportion des chercheurs CNRS varie énormément d'un laboratoire à l'autre, de 0 % à 60 %, avec une moyenne de 13 % : le tableau 5 montre que pour deux chercheurs CNRS, on trouve six enseignants-chercheurs et un chercheur d'autres organismes (surtout INRIA, et dans une moindre mesure, GET ou INSERM). Sur les 49 UPR et UMR qui dépendent au premier rang de la section 07, 20 atteignent ou dépassent le nombre moyen de chercheurs CNRS par unité ; de plus, la moitié des chercheurs CNRS de la section 07 est concentrée dans une douzaine d'unités. Il existe *a contrario* 2 UMR sans chercheur CNRS, et 9 UMR avec un seul chercheur CNRS.

Si l'on considère l'ensemble du personnel chercheur (chercheurs CNRS, chercheurs d'autres organismes et enseignants/chercheurs), on trouve 22 unités dont le personnel chercheur est supérieur à la moyenne des UPR et UMR relevant au premier rang de la section 07. On peut noter que les « grosses » UMR au sens du personnel CNRS ne sont pas toujours les mêmes que les « grosses » UMR au sens de l'ensemble des partenaires (Voir Figures 1 et 2).

Tableau 5. Nombre de chercheurs dans les laboratoires de la section 07

	Chercheurs CNRS	Chercheurs non CNRS	E/C	Total
Tous laboratoires	500	200	~ 3 000	~ 3 700
Section 07 principale	420	190	2 400	~ 3 000
Moyenne par unité	8	4	50	61

Sans s'appesantir sur la répartition des laboratoires sur le territoire national qui met en évidence les trois grands centres que sont, pour le personnel CNRS, Toulouse, Paris et Grenoble, on peut toutefois noter deux particularités locales : Grenoble, qui au contraire de

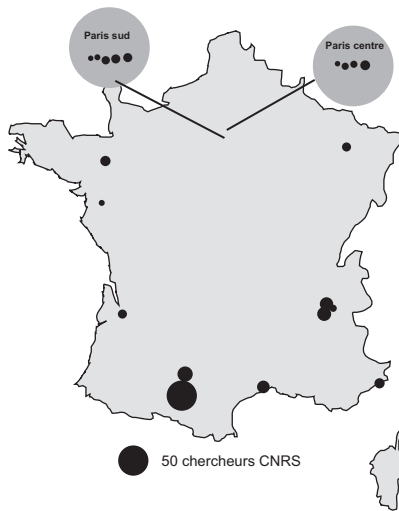


Figure 1 : Répartition géographique des 20 laboratoires atteignant ou dépassant le nombre moyen de chercheurs CNRS – 07.

la plupart des grandes villes de province structurées autour d'une, voire deux, unités, abrite plusieurs laboratoires assez importants au sens du personnel chercheur CNRS ; et Paris, où émergent deux « groupes » de laboratoires, l'un dans une zone « Paris Sud », et l'autre dans « Paris Centre ».

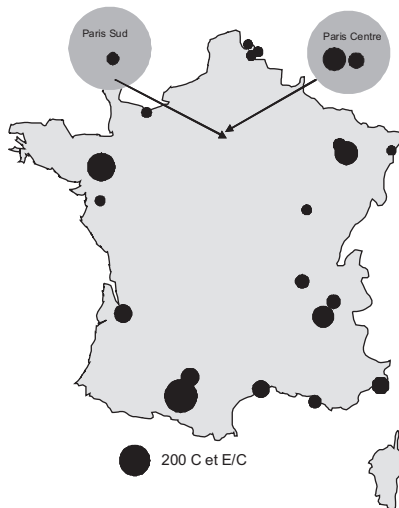


Figure 2 : Répartition géographique des 22 laboratoires regroupant le plus grand nombre d'enseignants/chercheurs et chercheurs relevant de la section 07.

3.2 ITA ET IATOS

La question des personnels (6) ITA et IATOS est plus délicate à décrire car ces personnels ne sont pas « attachés » à une section du Comité National, mais dépendent, pour ce qui concerne les agents CNRS, des Départements Scientifiques et sont affectés à un laboratoire sans qu'il soit facile d'identifier la spécificité scientifique de leur activité, même quand on s'intéresse uniquement aux unités « section 07 principale ». Cette remarque vaut surtout pour les UPR qui ont toutes deux une part très forte de leur activité hors de la section 07 (en section 08 pour le LAAS, et section 10 pour le LIMSI) ; pour établir le présent document, les chiffres bruts d'ITA de ces deux unités ont été divisés par deux pour estimer « la part » section 07. Il est important de noter la part non négligeable d'ITA « non CNRS », appartenant à d'autres organismes, en particulier l'INRIA, et le nombre très modeste de personnels des Universités.

Tableau 6. Répartition des ITA IATOS

	ITA CNRS	ITA non CNRS	IATOS	Total
Part estimée section 07	380	210	250	840
Moyenne par unité	8	4	5	17

Une comparaison des populations de chercheurs-enseignants/chercheurs et ITA-IATOS montre qu'un ITA-IATOS est chargé en moyenne d'assurer gestion ou soutien à la recherche pour environ quatre chercheurs et près du double si on considère les doctorants. Les ITA-IATOS de la section 07 représentent ainsi environ un tiers des ITA-IATOS du Département, alors que les chercheurs représentent près de la moitié du total des chercheurs du Département.

Sur le plan géographique, on retrouve assez logiquement les grands centres que

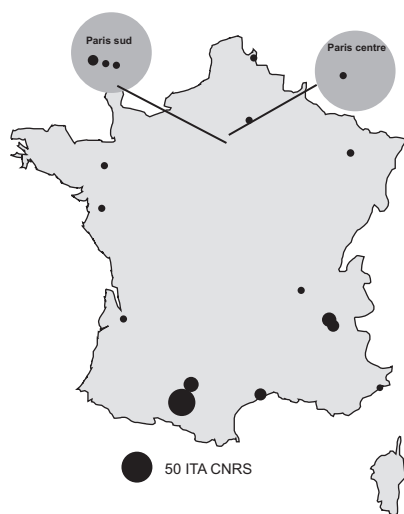


Figure 3 : Répartition géographique des 17 laboratoires atteignant ou dépassant le nombre moyen d'ITA CNRS – 07.

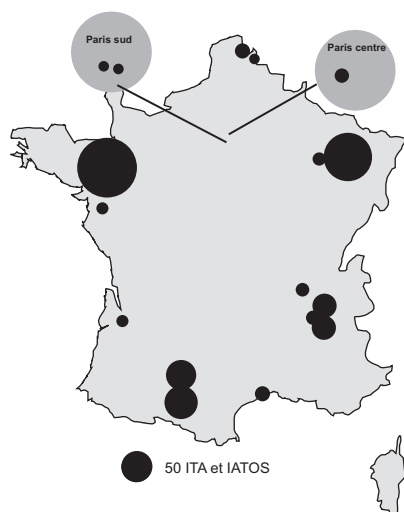


Figure 4 : Répartition géographique des 17 laboratoires regroupant le plus d'ITA et IATOS.

sont Toulouse, Paris et Grenoble pour le personnel CNRS. La situation est modifiée quand on considère l'ensemble des personnels ITA et IATOS, avec la mise en évidence de deux autres sites forts, Nancy et Rennes : l'influence de l'INRIA est ici très sensible.

Enfin, sur le plan des spécialités, on trouve parmi les ITA CNRS une majorité

d'agents relevant de la BAP E (environ 45%), puis des BAP F, G et H (au total 40%) et en fin des BAP A, B et C (15%).

3.3 ENSEMBLE DES PERMANENTS

En moyenne, un laboratoire relevant principalement de la section 07 accueille, pour «sa part section 07», près de quarantevingt agents tous établissements confondus (dont une quinzaine d'agents CNRS; Voir Tableau 7), les extrêmes allant d'une dizaine d'agents jusqu'à près de trois cents. Bien entendu, certains laboratoires sont plus importants si on considère les personnels relevant d'autres sections (c'est le cas en particulier du LAAS).

Tableau 7. Répartition des permanents

	CNRS	Autres organismes	Universités	Total
Ensemble unités Section 07 principale	~ 800	~ 400	~ 2 650	~ 3 800
Moyenne par unité	16	8	54	77

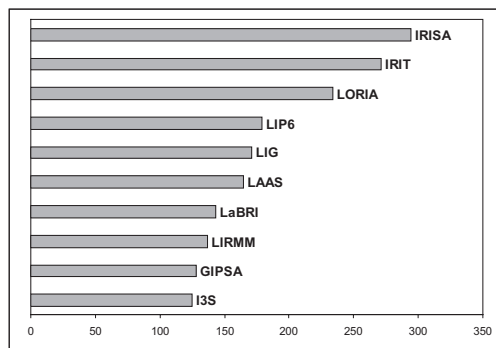


Figure 5 : Les dix plus gros laboratoires de la section 07 (seuls les personnels relevant de la section 07 sont pris en compte).

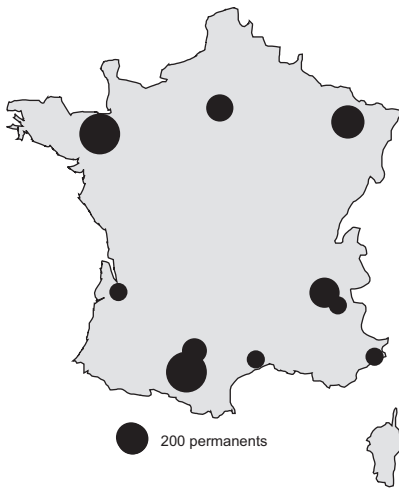


Figure 6 : Les 10 laboratoires regroupant plus de 100 permanents relevant de la section 07.

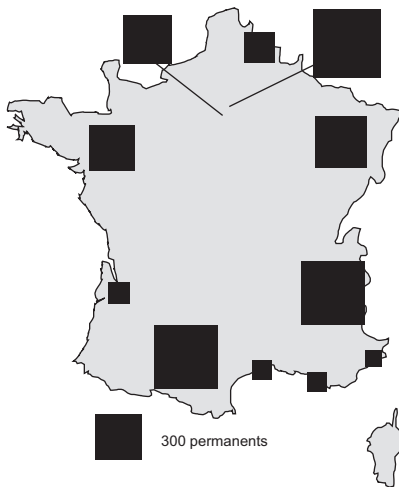


Figure 7 : Les plus grands sites « section 07 » qui regroupent 80 % des permanents (Paris est ici divisé en « Paris Sud » et « Paris Centre »).

En considérant l'ensemble des personnels permanents, les dix plus gros laboratoires regroupent environ la moitié des 3 800 permanents mentionnés dans le tableau 7 : on les retrouve sur la figure 5 et la figure 6.

Si l'on raisonne plus globalement, par site et non plus par laboratoire, la région parisienne représente un quart des personnels, et

on retrouve 80 % des permanents sur une dizaine de sites majeurs qui sont reportés sur la figure 7.

3.4 PLACE DES FEMMES DANS LE PERSONNEL CHERCHEUR

La place des femmes dans l'ensemble du personnel chercheur relevant de la section 07 mérite d'être évoquée (Voir Tableau 8) ; elle est nettement plus faible que pour l'ensemble des chercheurs du CNRS, que ce soit en considérant les enseignants/chercheurs universitaires des laboratoires de la section 07, ou plus encore les chercheurs CNRS qui relèvent directement de cette section. Cette situation de déficit se prolonge lors des concours de recrutement récents.

Tableau 8. La place des femmes dans le personnel chercheur

	CR ou MC	DR ou PR	Tous
Chercheurs CNRS toutes sections	37 %	22 %	31 %
E/C labos section 07	25 %	15 %	22 %
Chercheurs CNRS labos section 07	21 %	15 %	19 %

3.5 RECRUTEMENTS DES CHERCHEURS CNRS

Le nombre des postes de chercheurs CNRS, CR et DR, ouverts aux concours entre 1990 et 2006 a fluctué entre 17 et 62 pour s'établir à une moyenne de 25 sur la période considérée (Voir Figure 8).

Au-delà des fluctuations annuelles, difficilement compréhensibles et sans doute pré-

judiciables à la lisibilité et à la sérénité du recrutement, on peut remarquer l'événement majeur de cette période : le choix d'une priorité nationale « STIC » et la création en 2000 du Département éponyme. En effet, une fois les données lissées par une moyenne glissante sur 4 ans (la durée d'une mandature au Comité National) afin de s'affranchir des aléas annuels, on distingue deux phases (Voir Figure 9); jusqu'en 2000, la section 07 gérait environ 3,1% des postes mis au concours chaque année, ou environ 18 postes; depuis 2001, cette part est montée à 5,1%, ou environ 31 postes.

4 - INFORMATIQUE

L'informatique est aujourd'hui une science majeure, mature, mais toujours extrêmement dynamique et dont la richesse des thèmes rend difficile la description synthétique et complète; les paragraphes qui suivent tendent néanmoins d'éclairer ce vaste champ de recherche.

4.1 ALGORITHMIQUE

L'algorithmique, science des algorithmes, est un outil utilisé bien au-delà de la seule section 07. Les recherches menées sur les algorithmes en lien avec les structures discrètes qui modélisent les systèmes informatiques mobilisent de nombreux chercheurs de notre section. Ces recherches algorithmiques ont un caractère fondamental. Cependant, il est fréquent que des recherches aient des retombées dans des domaines aussi divers que le routage pour les réseaux, les télécommunications, ou la bioinformatique.

L'algorithmique s'articule avec la théorie de la complexité, domaine très actif de nos jours, particulièrement chez nos voisins européens. Les travaux d'approximation, permettent de mettre au point des solutions approchées à des problèmes difficiles (au sens de la complexité). L'utilisation de méthodes aléatoires pour la résolution de ces problèmes a permis d'introduire de nouvelles classes d'algorithmes dont les résultats sont « garantis » par des ratios. L'utilisation de l'aléa permet aussi de résoudre des problèmes plus rapidement que de manière déterministe. Un des nouveaux problèmes de l'informatique est la croissance de la taille des données à traiter. Il s'agit donc de mettre au point de nouveaux modèles de calculs qui permettent de n'examiner qu'une partie des données et qui fournissent quand même de bonnes approximations à des problèmes insolubles exactement.

De nombreux résultats fondamentaux d'algorithmique sont obtenus en étudiant les

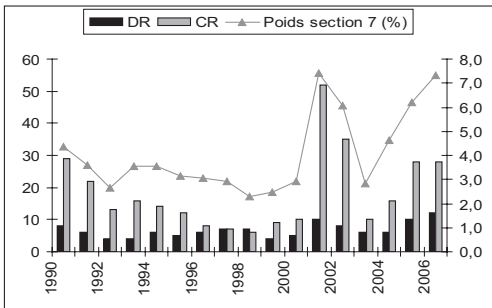


Figure 8 : Évolution des postes en section 07 (histogramme), et proportion par rapport à l'ensemble du concours (courbe).

Il en résulte un âge moyen nettement moins élevé pour les chercheurs de la section 07 que pour l'ensemble du CNRS : environ 43 ans, soit trois années de moins que pour la moyenne d'ensemble; cette différence est essentiellement due à la population des CR.

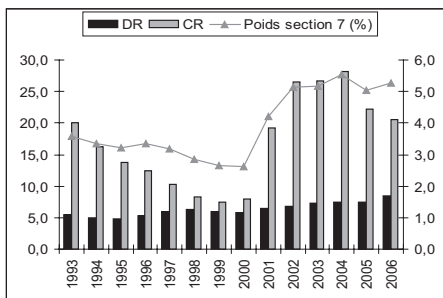


Figure 9 : Évolution des postes en section 07 (histogramme), et proportion par rapport à l'ensemble du concours, en moyenne glissante sur 4 ans.

structures combinatoires d'objets mathématiques. Les applications de cette branche de la combinatoire discrète en algorithmique du texte et du génome sont nombreuses. Les travaux de combinatoire enrichissent aussi l'optimisation combinatoire, discipline qui traite de la résolution exacte (ou approchée) de problèmes difficiles au sens de la complexité. Ces problèmes sont de nature discrète mais, pour les résoudre, il est indispensable d'utiliser des résultats de programmation mathématique et de recherche opérationnelle.

L'étude des graphes et des automates joue un rôle central non seulement en algorithmique mais aussi pour la modélisation des systèmes informatiques. En dehors des résultats obtenus sur des problèmes fondamentaux, la théorie des graphes est aussi au cœur des développements récents de la théorie des jeux et des équilibres de Nash (équilibre dans lequel aucun agent n'a intérêt à changer sa stratégie).

La théorie des jeux s'articule aussi avec la sémantique et le typage, mais il ne faut pas oublier qu'elle joue un rôle fondamental pour la conception et l'analyse de modèles de calculs qui nous permettent de mieux comprendre la nature la notion d'information. La théorie algorithmique de l'information (ou complexité de Kolmogorov) est un des outils les plus utilisés de nos jours dans différentes branches de l'informatique théorique pour mesurer les quantités d'informations. Cette théorie constitue un des développements les plus actifs de la théorie générale des algorithmes (calculabilité).

La géométrie algorithmique recouvre l'ensemble des algorithmes qui manipulent des objets géométriques (surfaces, courbes, etc.). Elle s'articule harmonieusement avec des approches complémentaires plus proches du signal (voir § 6) dans le cadre des divers problèmes d'infographie, de vision et de synthèse d'images, mais elle contribue aussi à la résolution de problèmes classiques d'algorithmique.

La cryptologie dans ses aspects tant théoriques que pratiques est un domaine vertical, à la fois proche de la complexité abstraite, de la conception d'algorithmes, de l'implémentation, etc. jusqu'à la carte à puce. La cryptogra-

phie traditionnelle (cryptage et décryptage des communications entre deux entités) ne peut répondre aux nouveaux paramètres des échanges sur internet et de nouveaux scénarios tels que le vote électronique, les enchères en ligne, etc. En effet, réservée autrefois aux communications militaires et au secteur bancaire, son industrialisation récente liée au développement d'internet a déplacé l'effort des chercheurs vers des problèmes de protocoles où la nature du réseau des échange intervient quelques fois plus dans les failles de sécurité que la solidité de la cryptographie sous-jacente. Les approches heuristiques utilisées dans un premier temps sont remplacées par une « sécurité prouvée » et les recherches en cryptologie s'attachent à valider les algorithmes ou schémas cryptographiques non plus en se basant sur leur résistance décennale aux attaques des cryptanalystes mais en produisant des preuves de sécurité. Enfin, la sécurité de tout protocole cryptographique reposant sur la difficulté d'un problème algorithmique, la recherche de nouveaux problèmes difficiles (dans le sens où ils contiennent non seulement des instances difficiles mais où ces dernières sont faciles à générer) constitue un socle important du domaine.

4.2 GÉNIE LOGICIEL, FIABILITÉ, PROGRAMMATION

Le génie logiciel recherche des solutions pratiques, scientifiquement fondées, afin de produire et maintenir des logiciels avec l'assurance d'en maîtriser les coûts, la qualité et les délais. Pour ce faire, on s'appuie sur la notion de composants logiciels développés indépendamment les uns des autres. L'assemblage, essentiellement manuel, de ces composants permet de construire des familles de produits partageant de nombreux modules, tout en restant ouvertes à de nouvelles évolutions. Il faut donc outiller le processus de conception par aspect et inventer des techniques en permettant la validation. On doit donc représenter et raisonner sur de tels assemblages de composants, ce qui se fait par l'élaboration de

modèles représentant et isolant différents aspects d'une ligne de produits, comme par exemple les variations fonctionnelles, les aspects structurels (paradigme objet), les aspects dynamiques (langages de scénarios), sans négliger les aspects non fonctionnels comme la qualité de service (performance, fiabilité) ou encore les caractéristiques de déploiement, qui deviennent prépondérants pour les systèmes répartis réactifs et temps réel. Cette ingénierie du logiciel fondée sur les modèles pose de nombreuses questions fondamentales sur la composition d'aspects. Pour pallier la vulnérabilité des systèmes informatiques et des réseaux il devient indispensable de savoir définir et garantir leur fiabilité : sûreté de fonctionnement, sécurité, confidentialité et intégrité de l'information.

La robustesse des applications informatiques et leur rapidité de développement sont accrues par l'utilisation de langages de programmation expressifs et sûrs. Ces dernières années ont permis de hisser les environnements de programmation typée, et notamment les langages déclaratifs ou fonctionnels, à un niveau de performance et de sûreté remarquable. Le typage statique accroît la sécurité de la programmation, la rapidité du développement d'applications et facilite leur maintenance. Un des objectifs du typage est la détection rapide d'erreurs de programmation. Les systèmes de types figurent parmi les formalismes principaux de recherche de preuves de programmes où les types sont vus comme des spécifications. En particulier, la spécification et la vérification de politiques de sécurité sont des domaines actuellement très actifs à cause du succès du code mobile et des cartes à puce « ouvertes » (cartes Java), car il s'agit de programmes qui s'exécutent dans un contexte très particulier, et pour lesquels il semble possible de définir des politiques de sécurité raisonnables. Dans le même esprit, on assiste à une rencontre entre les compétences algorithmiques traditionnelles de la cryptographie et celles plus logiques ou sémantiques qui abordent la sécurité des protocoles cryptographiques avec l'expérience des assistants de preuve et de la sémantique dénotationnelle.

Un autre domaine très actif est celui de la programmation par contraintes, né du rapprochement de techniques de programmation logique, d'optimisation combinatoire, de recherche opérationnelle, et de propagation de contraintes issues de l'intelligence artificielle. Ceci illustre l'apport de la recherche opérationnelle aux langages déclaratifs pour combiner efficacement des techniques de résolution hétérogènes (numériques, symboliques, deductives, heuristiques) pour la modélisation à la fois du problème à résoudre et des stratégies de résolution, entraînant une diminution importante des coûts de développement et de maintenance du logiciel. Cela permet surtout d'attaquer des problèmes NP-difficiles grâce à des outils à la fois génériques et efficaces. La programmation orientée objets, elle, évolue vers des paradigmes à plus grande échelle à base d'aspects et de composants. On citera enfin les langages de programmation de haut niveau pour le temps réel, pour la manipulation de documents XML (eXtensible Markup Language) et plus généralement de données semi structurées, pour le code mobile.

En amont, la correspondance de Curry-Howard entre programmes et preuves mathématiques, entre spécifications et formules, continue de fournir des clés pour la compréhension, la conception et l'implantation des langages de programmation. Citons par exemple la récente découverte d'une interprétation calculatoire de l'axiome du choix de la théorie des ensembles, qui correspond à un mécanisme d'horloge. Cette correspondance est à la base du système Coq sur les preuves de programmes, assistant de preuve utilisé par de nombreuses équipes, aussi bien pour la formalisation de preuves mathématiques que pour la certification de programmes.

4.3 SYSTÈMES RÉPARTIS ET RÉSEAUX DE COMMUNICATION

Les recherches dans ce domaine s'intéressent à la construction de systèmes informati-

ques mettant en œuvre des ressources de traitement et de stockage de l'information géographiquement dispersées, et à celle des infrastructures matérielles et logicielles qui leur sont liées (incluant par exemple les systèmes d'exploitation, les logiciels intermédiaires, les protocoles de communication). Un fait déterminant pour ces recherches réside dans l'émergence d'une informatique omniprésente, conséquence de deux tendances de fond qui entraînent un accroissement continu de la complexité des systèmes : d'une part la multiplication des capacités de traitement, de stockage et de communication de l'information à des échelles toujours plus vastes (des étiquettes intelligentes aux grilles de calcul, des réseaux spontanés à l'internet planétaire), d'autre part l'imbrication croissante d'éléments informatiques dans les artefacts et les procédés humains. La popularité croissante de systèmes de partage de données à grande échelle (réseaux « pair à pair ») a suscité des avancées en matière d'algorithmique répartie (développement de protocoles probabilistes, et d'algorithmes de routage et de localisation basés sur des tables de hachage réparties). La dispersion et l'irrégularité des systèmes concernés commencent à être prises en compte dans les modèles de programmation avec le développement de nouveaux modèles formels de la concurrence (« join calcul », « calcul des ambients » basés sur des notions explicites de localité), et dans les modèles de réseaux et de systèmes utilisés en algorithmique.

En dépit de ces progrès, la construction d'infrastructures de communication et d'infrastructures logicielles sûres pour une informatique omniprésente continue de poser de nombreux problèmes :

– i) le passage à l'échelle concerne la possibilité de déployer et de faire fonctionner ces infrastructures de manière efficace selon plusieurs dimensions : capacités (des processeurs, des réseaux), nombre (des objets matériels et logiciels impliqués, des utilisateurs), espace (dispersion spatiale), organisation (différentes formes d'organisation, depuis des structures locales spontanées jusqu'à des fédérations multi-entreprises) ;

– ii) la qualité de service concerne la fourchette de garanties de correction et de niveau de service au cours de l'exécution d'un système ou d'une application. Un accord sur un niveau de service peut concerner notamment des contraintes temporelles (temps de réponse), de ressources (bande passante disponible), ou de sûreté de fonctionnement (taux de disponibilité, niveau de sécurité) ;

– iii) l'administrabilité concerne la possibilité de superviser et de contrôler, à des échelles multiples, le comportement et l'évolution des systèmes visés ;

– iv) enfin, la programmabilité concerne la possibilité de construire de nouveaux systèmes et applications par composition et coordination, à plusieurs échelles, tout en offrant aux concepteurs et aux programmeurs des garanties de sûreté (par exemple en termes d'absence d'erreurs de communication, de consommation de ressources, de sécurité) comparables à celles dont ils peuvent disposer dans une programmation centralisée classique.

La mise en œuvre même des fonctions d'administration pose de redoutables problèmes, surtout à grande échelle : par exemple, d'instrumentation dynamique et multi-échelles, à des fins d'observation, de mesure et de comptabilité ; de construction de systèmes de notification et d'orchestration « scalables » ; de systèmes en ligne et fédérés d'analyse, de diagnostic et de planification, à des fins de supervision et de commande.

Enfin, des réflexions actuelles posent la question d'une rupture importante : en effet, la couche protocolaire IP a été conçue pour interconnecter quelques milliers de gros ordinateurs principalement pour des applications de transfert de données dans un univers considéré jusqu'alors comme fiable. Le développement des usages (voix, données, images, télévision), la multiplicité des terminaux et des accès ainsi que le détournement de la sécurité posent la question d'une remise en cause, à terme, de l'Internet tel qu'il existe aujourd'hui. On remarquera d'ailleurs que les programmes de la NSF (GENI) et de la Commission Européenne (dans le cadre du FP7) envisagent de

considérer ces aspects, au moins comme une cible idéale.

4.4 BASES DE DONNÉES, RECHERCHE ET SYSTÈMES D'INFORMATION

La croissance des volumes d'information, les multiples formes d'acquisition, de codage et de restitution de l'information, l'explosion des usages et donc des besoins d'accès et de traitement de l'information à tout moment, de n'importe où, suivant des modalités matérielles et des médias très différents sont à l'origine des enjeux des nouvelles générations de systèmes d'information. Si l'internet (Web et commerce électronique) fut l'un des déclencheurs majeurs de la globalisation de l'information par multiplication de ses sources et de ses usages, les sciences (de la physique aux sciences de la terre en passant par la biologie sans oublier les sciences humaines) se révèlent être des domaines fortement consommateurs et producteurs d'information qui requièrent de nouveaux processus de traitement ayant des capacités d'intégration et d'analyse sophistiquées et puissantes, inexistantes actuellement.

L'information est disponible aujourd'hui à partir de nombreuses sources autonomes, hétérogènes : bases de données, documents multimédia, pages Web, fichiers, etc.

L'hétérogénéité de l'information/des données se déclinent à tous les niveaux d'abstraction et suivant des dimensions diverses dont les combinaisons sont un des facteurs de complexité : hétérogénéité sémantique (terminologie, espace-temps, etc.), hétérogénéité de types (texte, image, son, vidéo, codes, stream, etc.) hétérogénéité de formats (RTF, Html, Xml, MP4, etc.), hétérogénéité des représentations (structurées ou non, descripteurs, etc.). L'information/les données évoluent de manière plus ou moins fréquente, parfois de manière permanente et intensive. La restitution de cette évolution et la qualité de l'information sont

dépendantes entre autre du choix de granularité temporelle (discret, continu) ainsi que de la localisation des sources (centralisation, distribution, répartition, réplication, etc.).

Les systèmes d'information ont pour objet la gestion de l'information tout au long d'une chaîne complexe : capture et saisie des données, constitution, production, structuration de l'information, ainsi qu'extraction, analyse, visualisation de l'information.

Par exemple, l'explosion récente de l'usage de capteurs, mobiles ou non, produisant des données sous forme de signaux (son, image, mesure, etc.) génère de nombreux problèmes difficiles d'acquisition et de traitement. À l'autre extrémité de la chaîne, là où les usages constituent la cible du système d'information, la visualisation de données massives s'avère une des voies incontournables pour présenter l'information à l'utilisateur, pour lui permettre et l'assister dans une analyse exploratoire, pour une simulation. La multiplicité d'usages et d'utilisateurs d'un même système d'information impose que des méthodes d'interaction personnalisées soient développées pour prendre en compte le contexte d'usage et le profil de l'utilisateur. Ces systèmes d'information nécessitent des modèles de représentation, de manipulation et des méthodes d'indexation permettant de modéliser et opérer à tous les niveaux d'abstraction évoqués précédemment. Assurer la qualité des données (cohérence, fraîcheur, pertinence, etc.), la confidentialité et la sécurité des accès sont certainement deux des objectifs rendus particulièrement critiques par la masse des données à disposition et l'accès banalisé aux sources de données. Ces systèmes d'information doivent assembler, faire coopérer, composer de nombreuses méthodes issues des bases de données, de la recherche d'information, de l'intelligence artificielle : recherche exacte, recherche approximative, fouille de données et apprentissage, raisonnement incertain/probabiliste, traitement de la langue naturelle, visualisation, etc. Ils doivent être portables et s'intégrer à des environnements matériels et logiciels différents, pour des configurations allant de la carte à puce aux grilles de données en passant

par les gisements de données. Les fonctionnalités de ces systèmes doivent être développées indépendamment tout en étant facilement composables.

Les travaux en bases de données et recherche d'information doivent intensifier l'intégration des méthodes et des approches tout autant que l'exploration de nouvelles méthodes pour développer les concepts et technologies permettant de rendre transparents la complexité des procédés d'acquisition, de traitement, d'analyse, d'interrogation, d'extraction, d'accès aux informations, tout autant que le volume, la nature et la localisation des données accédées et ceci pour des usages allant de l'ingénierie industrielle à l'usage personnel.

4.5 INTELLIGENCE ARTIFICIELLE, COMMUNICATION HOMME-MACHINE ET COGNITION

L'objet de l'intelligence artificielle (IA) est de construire des systèmes informatiques capables d'effectuer des tâches réputées nécessiter de l'intelligence. Les recherches en IA tendent à rendre la machine capable d'acquiescer de l'information, de raisonner sur une situation statique ou dynamique, de résoudre de manière générique des problèmes combinatoires, de faire un diagnostic, de proposer une décision, un plan d'action, d'expliquer et de communiquer les conclusions qu'elle obtient, de comprendre un texte simple ou un dialogue en langage naturel et d'en extraire des éléments d'information. L'ingénierie des connaissances s'attache à résoudre les problèmes posés par la gestion et l'accès à de grandes bases de connaissances hétérogènes (développement par exemple d'« ontologies » pour mieux appréhender des domaines particuliers).

Les grandes problématiques de l'IA sont :

- i) le développement de cadres pour la représentation logique et/ou structurée des connaissances et la formalisation de différents types de raisonnement et de décision ;

- ii) l'algorithmique de l'inférence et des systèmes de contraintes ;

- iii) l'apprentissage et la fouille de données ;

- iv) les systèmes multi-agents, et les modèles formels d'interaction ;

- v) la résolution de problèmes et la planification des actions.

Un fait marquant des recherches actuelles en intelligence artificielle est qu'elles sont de plus en plus développées en synergie avec les sciences cognitives, afin de valider l'adéquation des modèles utilisés avec la façon dont les agents appréhendent l'incertitude, ou perçoivent les relations temporelles ou spatiales, ou dont les agents interagissent entre eux, modifient leurs croyances, ou expriment leurs préférences ou leurs intentions. On s'intéresse aussi à savoir dans quelle mesure les mécanismes artificiels peuvent simuler assez fidèlement les processus qui se déroulent au niveau du cerveau humain.

La communication homme-machine, en dehors des questions de reconnaissance et de synthèse de la parole et du geste liées au développement de systèmes toujours plus performants offrant de nouvelles modalités, du traitement des documents écrits, s'intéresse à la conception et au développement d'interfaces et d'outils de visualisation, cognitivement ergonomiques pour les usagers (qui peuvent éventuellement présenter différents types de handicaps). Les systèmes cognitifs artificiels, comme les êtres vivants, doivent faire face à une difficulté fondamentale : comment percevoir, raisonner, décider et agir avec leur environnement ? Quand cet environnement est peuplé d'humains, le système doit intégrer dans son raisonnement des connaissances sur les codes (linguistiques, interactionnels, sociaux, etc.) de la communication humaine, notamment face-à-face lorsque le système est incarné par un agent conversationnel à apparence humaine. Dans un contexte d'intelligence ambiante, la communication homme-machine, longtemps cantonnée à la conception et au développement d'interfaces et d'outils de visualisation, s'ouvre ainsi à un champ plus

large d'étude et de modélisation de boucles d'interaction multimodales prenant en compte les usagers ainsi que l'environnement physique de l'interaction. Notons que l'étude du comportement humain en interaction située s'appuie sur des données expérimentales nécessitant des plate-formes techniques parfois lourdes (réalité virtuelle, capture de mouvement, imagerie médicale, etc.) dont l'interprétation et la modélisation nécessite la constitution d'équipes pluridisciplinaires de chercheurs de STIC, SDV et SHS.

Les recherches sur le raisonnement portent en particulier sur le changement, le temps et l'espace, sur le développement de représentation graphique ou qualitative de l'incertitude (probabilités, possibilités) et des préférences, sur la gestion des incohérences. De grands progrès continuent d'être réalisés en matière d'algorithmes efficaces pour la déduction en logiques classique et non classiques (modales, conditionnelles, pondérées), ou pour la résolution de problèmes pouvant s'exprimer sous forme de satisfaction d'ensembles de contraintes, éventuellement flexibles (comme par exemple en conception assistée par ordinateur).

L'apprentissage artificiel connaît aussi de grandes avancées, avec le développement de l'apprentissage par machines à vecteurs de support, de l'apprentissage bayésien, de la programmation logique inductive, de l'apprentissage par renforcement, des méthodes neuro-floues, mais également à travers une meilleure compréhension de l'apprentissage supervisé, et par une interaction fructueuse avec les sciences de l'homme sur le thème de la cognition. Un des enjeux majeurs est de déterminer comment combiner efficacement les méthodes d'apprentissage avec des informations de nature statistique. Les problématiques de la fouille de données, du résumé d'informations, de la découverte de règles, en liaison avec les méthodes statistiques et symboliques d'analyse de données ont aussi élargi les perspectives d'application de la recherche en apprentissage, comme en génomique par exemple.

Le paradigme de systèmes multi-agents en IA connaît des développements considéra-

bles, au niveau de la simulation de l'activité de société d'agents, ou de la modélisation de processus coopératifs, de négociation, d'argumentation, de décision de groupe équitable, ou de maximisation de gain dans des problèmes d'enchères combinatoires, conduisant à des applications en commerce électronique, ou en recherche d'information sur la toile, en particulier. La problématique ancienne en IA de la résolution de problèmes se retrouve en planification. Il s'agit de raisonner sur des actions et leurs conséquences pour trouver une solution, à partir de croyances et de perceptions. Au plan algorithmique, un planificateur est un démonstrateur (ou résolveur de contraintes) très spécialisé qui raisonne efficacement sur des axiomes qui décrivent les actions. Des représentations par processus markoviens partiellement observés peuvent être aussi utilisées en cas de situations incertaines. La planification et l'ordonnancement trouvent leurs applications en gestion de trafic aérien, en gestion d'entreprise, en robotique par exemple.

Depuis près d'une décennie le GDR I3 Information-Interaction-Intelligence fédère les communautés des anciens GDR Bases de Données, Communication Homme-Machine, et Intelligence Artificielle regroupant une communauté de près de huit cent chercheurs. Le GDR I3 favorise dans la mesure de ses moyens les collaborations entre les chercheurs des différentes communautés qu'il réunit, afin de concevoir et de développer des outils pour une nouvelle génération de systèmes d'information intelligents et interactifs.

4.6 BIOINFORMATIQUE

La bioinformatique est une science jeune, en pleine expansion et pluridisciplinaire avec de nombreux problèmes théoriques ouverts et des applications en médecine, biotechnologie, pharmacie et agronomie. Elle a pour objectif l'analyse et la modélisation des éléments biologiques en se basant sur diverses disciplines : l'informatique (algorithmique, complexité, combinatoire des mots, cryptogra-

phie, simulation, bases de données, analyse de données et méthodes de classification, calcul distribué et parallèle, développement de logiciels de recherche, etc.), les mathématiques appliquées (mathématiques discrètes, géométrie, topologie, théorie des jeux, systèmes dynamiques, équations différentielles, probabilités, statistiques, etc.), la physique, la chimie, la biologie, la pharmacie et la médecine. Les divers programmes « génomes », de l'homme et d'organismes modèles (levures, mouches, vers, plantes, bactéries, virus, etc.), ont permis de déterminer la séquence de plusieurs milliards de nucléotides (lettres) dans diverses espèces et de leur attribuer une fonction biologique : gènes codant les protéines, gènes codant les ARN, gènes non codants (introns, microsatellites, etc.). Ainsi, le nombre de gènes (codant les protéines) est estimé à 30 000 dans les 23 paires de chromosomes de l'homme, à 6 000 pour la levure, à 18 000 pour le ver, etc. De très nombreux organismes sont en cours de séquençage et en voie d'achèvement suite à l'apparition de nouvelles techniques de séquençage à haut débit, rapides et peu coûteuses. Ainsi, plus de 700 projets de séquençage de génomes bactériens sont en cours, plusieurs souches de *E. Coli* sont entièrement séquencées tandis qu'une vingtaine de génomes complets de mammifères sont aujourd'hui déterminés. Cette ère génomique a donc généré une masse de données biologiques volumineuses, complexes et hétérogènes qui sont disponibles dans des centaines de bases de données, les bases de données de gènes principales et historiques étant EMBL et GenBank. L'ère post-génomique s'intéresse à l'analyse et à la modélisation de cette masse d'information.

Les éléments biologiques peuvent être étudiés à plusieurs niveaux. Au niveau moléculaire, les études des gènes et des protéines peuvent porter sur leurs structures primaire, secondaire, tertiaire et quaternaire. La génomique comparative s'intéresse à la fonction globale du génome (diversité, adaptation, pathogénicité) et à sa comparaison avec d'autres espèces. Au niveau cellulaire, les méthodes bioinformatiques ont pour objectif d'intégrer les différentes molécules et leurs interactions

dans des réseaux de régulations (génomique fonctionnelle). La modélisation d'organes, enjeu moteur également du domaine de l'imagerie (voir § 6), conduit à des simulations à destination, par exemple, d'interventions assistées par ordinateur ou robotisées (voir § 7). Au niveau organisme, l'objectif est de trouver une théorie globale et unificatrice des niveaux sous-jacents, permettant par exemple de comprendre une pathologie intéressant plusieurs organes. Enfin au niveau population, les domaines de recherche s'intéressent notamment à la génétique des populations, à la transmission et la diffusion de maladies au sein des populations.

Les applications de la bioinformatique ont indiscutablement des impacts majeurs directs en santé, alimentation et nutrition, et donc des implications économiques et industrielles considérables. Cependant, une recherche fondamentale et théorique est indispensable pour développer une science bioinformatique quantitative et prédictive en identifiant des lois et des propriétés structurales, fonctionnelles et évolutives aux niveaux moléculaire, notamment les gènes et les protéines, cellulaire et organisme. Les approches sont nombreuses et portent en particulier sur :

- le développement de méthodes probabilistes et statistiques pour l'étude des données génomiques et protéiques ;
- les méthodes mathématiques pour la topologie de l'ADN et des protéines ;
- les méthodes d'alignement de séquences et leur significativité, et de façon plus générale l'algorithmique du texte ;
- les méthodes mathématiques numériques et distribuées pour paralléliser le traitement des données biologiques ;
- les modèles d'évolution moléculaire stochastiques qui diffèrent et complètent ceux de la phylogénie ;
- les études des codes dans les gènes et du code génétique dont l'origine reste un problème théorique majeur et ouvert depuis 50 ans malgré des milliers de publications à son sujet.

5 – AUTOMATIQUE

L'automatique est la discipline de la modélisation, de l'analyse, de la commande et de la supervision des systèmes et processus continus ou à événements discrets. Elle s'appuie sur des outils mathématiques, informatiques (calcul scientifique, réseaux), de recherche opérationnelle (optimisation, théorie des jeux), et sur ses propres outils en vue de l'analyse et de la maîtrise des systèmes naturels et artificiels. C'est essentiellement une science de l'action, développée à l'origine sur la notion de rétroaction (ou de boucle fermée), visant à modifier le comportement naturel d'un système, ce qui a imposé des études de stabilité approfondies. Elle a introduit le concept de perturbation, non pas comme une gêne que l'on espère passagère ou de faible importance, mais comme une information à traiter, à mesurer ou à prédire, si l'on veut arriver à un comportement souhaité pour le système, voire à un comportement optimal. Elle est essentiellement une science des systèmes dynamiques, préoccupée par les transitoires. Pour anticiper sur les réactions du système aux perturbations, pour garantir la stabilité de la commande, on a besoin d'un modèle mathématique du système à commander. L'ambition est d'aller de la vision locale, correspondant à un simple procédé bouclé avec un simple contrôleur, vers une vision globale dans laquelle des systèmes physiques et d'informations hétérogènes interagissent et communiquent. Dans ce cadre, une notion de base est celle de système, décrit par des entrées et des sorties, fonctions du temps et reliées par des équations mathématiques aux propriétés très diverses (déterministes, stochastiques, différentielles ou encore aux dérivées partielles). Pour les systèmes hybrides, généralement de grande dimension, qui combinent des opérations logiques avec des dynamiques continues, le besoin de méthodes constructives s'appuyant sur des théories rigoureuses est réel. Les outils théoriques développés en automatique sont implémentés grâce à des outils informatiques adaptés (robustesse des calculs numériques, architectures spécialisées), voire grâce à une implémen-

tation « on chip » de certaines parties de la commande. Les applications concernent de nombreux domaines (aéronautique, chimie, énergie, biologie, environnement, électronique, matériaux, robotique, médecine, réseaux, santé, transports, communication, économie, gestion, etc.).

Pour agir sur un système, il faut avoir une bonne information sur son évolution. L'automatique est donc liée au traitement du signal par son besoin de mise en forme des signaux issus de capteurs très divers, de filtrage, de fusion et/ou de classification de données, de transmission sans perte d'information. Les images constituent des « mesures » aux propriétés particulières, que l'on peut intégrer dans une boucle de régulation, comme par exemple en robotique chirurgicale. Toutefois, toute grandeur n'est pas mesurable, que ce soit pour des raisons physiques ou économiques. Il faut alors reconstruire l'information manquante à partir de celle qui est mesurable. La théorie sous-jacente, celle des « observateurs », a beaucoup évolué depuis le premier filtre de Kalman, mais est toujours en évolution vers des observateurs toujours mieux spécialisés (non linéaires, à grand gain, à horizon glissant, ensemblistes, etc.). Quelques points restent difficiles à régler, comme l'identification en présence de certaines non linéarités dures, l'identification en boucle fermée, en interaction avec la commande. Quelques points de vue novateurs commencent à percer dans le domaine des applications, comme l'estimation ensembliste. Un autre aspect essentiel pour l'avenir est d'être capable de construire des systèmes fiables à partir d'information pas nécessairement sûre. L'imprécision des mesures, la présence de retards, les limitations des actionneurs et capteurs, l'existence de perturbations externes et l'incertitude sur les modèles rendent particulièrement importantes les exigences de robustesse et d'adaptabilité des commandes. Les défauts ou les pannes de certains composants dans les systèmes embarqués, les problèmes de communications dans les systèmes interconnectés font apparaître de nouveaux besoins de reconfigurabilité des systèmes et donc de développement de théories et d'outils spécifiques permettant de configurer

ou reconfigurer en ligne les contrôleurs ou le système physique. Évidemment, avant de reconfigurer, il faut avoir diagnostiqué la source du dysfonctionnement observé et il faut pouvoir, en ligne, valider la nouvelle architecture commande/système. Ceci amène à définir de nouvelles propriétés des systèmes, aux noms un peu barbares, comme la diagnosticabilité, encore très peu étudiées et qui viennent compléter celles de commandabilité et observabilité. Sur les aspects vérification et validation du système contrôlé, les points d'ancrage avec les études des informaticiens sur les logiciels sont clairs.

L'automatique s'intéresse aussi à la productique, c'est-à-dire aux procédés manufacturiers, depuis leur conception ou la conception des produits jusqu'à la phase de destruction ou de recyclage du produit, en passant par le suivi de l'exploitation. L'automatique fournit des modèles divers, proches de ceux de l'informatique (automates, réseaux de Petri, chaînes de Markov, etc.). On retrouve les problèmes de base de l'automatique, modélisation, commande, supervision, diagnostic, maintenance, mais les outils mathématiques sont très différents, et une question ouverte est de savoir si certains concepts peuvent migrer du monde continu vers le monde discret (ou vice-versa). Les nouvelles organisations des entreprises modifient les manières de travailler, la capitalisation ou le partage de connaissances et de responsabilités, et l'automatique doit se tourner vers les sciences humaines (gestion, économie, sociologie) pour avoir une vision d'ensemble. Ceci conduit à de nouveaux modèles de représentation et de spécification (des organisations, des fonctions, des systèmes d'information et de décision, des systèmes de contrôle-commande et de supervision) et d'évaluation de leurs performances. La productique concerne la production de biens, mais aussi de services, ce qui pose de nouveaux problèmes. Enfin, que l'homme soit perçu comme une entrée de commande ou comme une perturbation du système, nous ne pouvons pas ne pas en tenir compte dans la conception d'un système de contrôle-commande. Aussi, les automaticiens sont-ils amenés à collaborer avec des spécialistes de l'interaction homme-machine. Il ne s'agit

pas ici seulement d'ergonomie, mais de travail coopératif, de partage de décision entre l'homme et la machine. On peut penser que les progrès réalisés par ailleurs en réalité augmentés auront un impact assez fort sur ce que les automaticiens proposent à l'heure actuelle comme outils d'assistance à l'opérateur.

Les laboratoires reconnus par le CNRS et ayant une activité significative en automatique sont très généralement rattachés principalement à la section 07, mais ne sont pas monothématiques, et associent aux équipes d'automatique, des équipes d'informatique, de traitement du signal, ou relevant des sections 08, 09, 27, 30 ou 34. Cet ensemble se retrouve depuis de nombreuses années au sein du GDR MACS qui assure efficacement un rôle d'animation de cette communauté et favorise sa visibilité.

Par ailleurs, la tendance actuelle est au rapprochement avec les mathématiques appliquées, tant par le biais de projets communs qu'au niveau des recrutements de doctorants et de chercheurs. Cette typologie de l'organisation de l'automatique en France diffère de ce qui se retrouve à l'étranger, et en particulier aux USA, où les automaticiens sont généralement intégrés dans des départements de génie (électrique, mécanique, des procédés, aéronautique, etc.).

On retrouve ainsi dans ces différentes organisations, sans que l'une soit a priori supérieure aux autres, l'essence même de l'automatique, qui réside dans la dualité entre les systèmes physiques auxquels elle s'adresse et le fait que les avancées majeures ne peuvent venir que du développement de méthodologies génériques, indépendamment de l'application pressenties. La formulation du « modèle standard » est, de ce point de vue, une étape fondamentale dans la formalisation génériques de problèmes de l'automatique moderne, dont la vitrine est certainement son application par le biais de la commande H_∞ sur les lanceurs Ariane nouvelle génération. Ainsi il y aurait tout autant de risque à vouloir déclinier l'automatique sous forme d'objets de recherche que de renoncer à l'interprétation « système » inhérente aux problèmes que l'automatique soulève.

6 – TRAITEMENT DU SIGNAL ET DE L'IMAGE

L'objet du traitement du signal est de développer des méthodes mathématiques (statistiques, algébriques, géométriques, analytiques, variationnelles), des modes de représentation adaptés (temps-fréquence, temps-échelle, multi-résolution, multi-fractales, systèmes fractionnaires, etc.) et de les mettre en œuvre à travers des techniques algorithmiques et informatiques dédiées, dans le but d'extraire et modéliser l'information contenue dans des signaux issus de capteurs.

Historiquement, le traitement d'image est issu du traitement du signal par extension des modes de représentation et d'analyse à des signaux de dimensions égales ou supérieures à 2 (variables d'espace, de temps ou autre paramètre). Aujourd'hui, le traitement d'image est un domaine en soi, qui a su développer ces 20 dernières années des méthodes spécifiques comme la morphologie mathématique ou le tatouage d'images par exemple.

Sur le domaine « signal », après les progrès accomplis en un demi-siècle depuis les travaux pionniers de Shannon, l'effort a porté plus spécifiquement sur des approches visant à élargir les hypothèses de travail conventionnelles vers des approches non conventionnelles (non-linéaire, non-gaussien, non-stationnaire, non-régulier, lois d'échelle, longue dépendance, représentation en $1/f$, etc.).

Le signal et l'image partagent des outils communs de représentation et de codage (ondelettes, analyse multi-résolution, approches pyramidales) et de modèles stochastiques (chaînes de Markov cachées, champs de Markov, méthodes de simulation). Aux méthodes mathématiques issues principalement de l'analyse se sont aussi ajoutés des outils statistiques nouveaux particulièrement efficaces en estimation, filtrage ou débruitage. En image, des travaux récents ont mis en valeur l'intérêt des modèles de graphes pour notamment enrichir les méthodes à échantillonnage régulier avec des méthodes irrégulières où les éléments

de représentation sont gérés par un graphe d'adjacence.

L'image en s'ouvrant sur les dimensions supérieures, donne accès aux images 3D avec la représentation des objets de la scène (modèles surfaciques, maillages, modèles volumiques, voxels). Imager et représenter des formes molles, comme en imagerie médicale, conduit à développer des échantillonnages irréguliers souples capables d'accompagner mouvement et déformations et localement adaptatifs. Ces modèles également développés en synthèse d'images, bénéficient d'une forte activité de recherche en complexité algorithmique et combinatoire avec la géométrie algorithmique.

La nature physique de l'image intervient de plus en plus dès lors que l'on s'oriente vers une imagerie quantitative des formes et des contenus. Les propriétés statistiques des images sont si différentes selon que l'image est fournie par un système d'acquisition RX, IRM, US, radar ou TEP, que des techniques générales de détection, segmentation, analyse ne sont plus suffisantes. En vision par imagerie optique le calibrage des prises de vue multiples et la correction de l'éclairage font l'objet de recherches spécifiques. En échographie, le traitement du signal radiofréquence ouvre des voies extrêmement prometteuses vers l'analyse des milieux traversés en particulier en biologie et médecine ; parmi les nouvelles modalités d'image, l'élastographie permettra bientôt l'accès *in vivo* aux propriétés d'élasticité des tissus.

L'accès aux séquences spatio-temporelles est source de développements de nouveaux modèles de représentation et de suivi des phénomènes dynamiques. D'une manière générale, un effort important est en cours pour « doper » les méthodes d'imagerie en introduisant, outre des *a priori* statistiques (physique de l'image), des *a priori* de formes, de mouvement ou de déformation. En analyse ou synthèse d'image, l'apport des méthodes variationnelles permet de s'appuyer sur les modes de représentation énergétiques par équations aux dérivées partielles, offrant ainsi des méthodes évolutives et convergentes pour la segmentation ou la reconstruction de surfaces.

Dans le domaine de l'interprétation, outre l'effort mené en diagnostic sur la base des nouveaux modes de représentation précitées (temps fréquence, temps-échelle, analyse spectrale, en lien explicite avec des procédures d'apprentissage et de décision), un important effort a porté sur la fusion, enrichissant les méthodes classiques probabilistes avec le flou, les possibilités, les croyances, etc.

On note aussi la place importante prise par des modèles, très souples, mais aussi peu contraints, comme les processus à noyaux ou les décompositions en ondelettes, permettant d'aboutir à des techniques d'apprentissage très puissantes : machines à vecteurs de support et calcul particulière. S'appuyant sur la simulation, sur les approches géométriques, sur la théorie de l'information, ces méthodes permettent d'aborder de front les signaux chaotiques et leurs représentations multi-échelles et fractales. Des emprunts novateurs à la mécanique quantique permettent de revisiter le codage, la transmission et l'analyse spectrale, tandis que des démarches inspirées des systèmes biologiques auto-organisés rendent plus efficaces les techniques d'apprentissage, de reconnaissance, de poursuite ou d'analyse.

Parmi les nombreuses applications du traitement du signal, citons, à titre d'exemples, les systèmes de reconnaissance de locuteurs ou de langues, de transcription et de synthèse de la parole, de traduction multilingues, de traitement de documents écrits, les télécommunications (antennes, codage de sources, codage de canal, séparation de sources, égalisation); et pour l'image, citons l'imagerie satellitaire, l'observation en sismique, l'imagerie biologique, l'imagerie médicale, le contrôle non destructif, la télésurveillance avec les aspects de biométrie, le multimédia avec les problèmes de représentation de scènes à réalité augmentée ou de vidéos.

Parmi les sujets ou thèmes de recherche émergents, on peut citer entre autres les aspects « multi » des informations signal ou images, l'intégration de niveaux d'abstraction permettant de passer du bas niveau au haut niveau, le développement de méthodes d'appréhension du contenu, la synthèse d'image,

les traitements adaptatifs et coopératifs prenant en compte la place de l'opérateur humain dans la boucle de traitement, la vision dynamique et active, la vision pour la réalité virtuelle et la réalité augmentée associant analyse et synthèse dans un mode immersif (avec des applications au *design* et à l'architecture), la modélisation des objets à géométrie et topologie complexe (goudrons, os, milieux poreux) et la modélisation très délicate des milieux vivants.

La quantité de données à traiter pour un passage à l'échelle en imagerie, et la délocalisation des sources sont si grandes que ce domaine est une application de choix pour l'implantation de techniques d'algorithmie parallèle, de calcul distribués sur grilles.

En France, la communauté s'est structurée depuis plus de 15 ans grâce au GDR ISIS. Il y a 3 ans, le GDR d'interface STIC-Santé CNRS-INSERM a été créé, signe de la dynamique forte des recherches émergentes à l'interface des sciences de l'information et du vivant. Enfin, les travaux en signal et image ont joué un rôle essentiel dans les recherches portant sur l'adéquation algorithme-architecture, assurant ainsi un lien important avec la communauté de la microélectronique.

7 – ROBOTIQUE

La robotique fait partie des sciences des systèmes artificiels et peut être vue comme la science de la perception, du mouvement et de leur intégration en une machine physique, mécanique et informatique.

Un robot est un système matériel possédant des capacités de perception, d'action, de décision et de communication, et capable d'agir dans un environnement dynamique, de façon autonome ou en relation avec un humain. Aujourd'hui, si les recherches visant à améliorer les performances « brutes » sont toujours actives, de nombreux efforts tendent

vers une ambition plus forte : le robot autonome et communicant. La robotique peut ainsi être décrite par quatre grandes thématiques :

- conception des systèmes robotiques ; on s'intéresse, à un premier niveau, aussi bien aux approches de conception (bio-inspiration, par exemple) qu'aux méthodes d'intégration, le tout dans des contextes contraints par les domaines d'applications (taille, masse, autonomie énergétique, bio-compatibilité, etc.). À un autre niveau, on considère aussi bien la structure des robots, que l'architecture des systèmes robotiques. Enfin, le niveau du composant (capteur, actionneur, matériau) reste un domaine essentiel et vivifié par de nombreuses technologies émergentes ;

- contrôles sensori-moteurs, perception, action et mouvement ; on fait ici appel en premier lieu à la perception pour les tâches de commande, de localisation et de modélisation du monde physique dans divers modes de représentation. La diversité sensorielle place la fusion de données comme un enjeu majeur. L'étude de l'action par le mouvement, quant à elle, s'appuie en premier lieu sur la modélisation et la commande de systèmes mécaniques d'une grande variété. Le couplage perception/action, l'algorithmique géométrique, et la modélisation pour la simulation physique sont au cœur des problèmes ;

- cognition, décision, autonomie, apprentissage ; à moins que les tâches soient entièrement préprogrammées et l'environnement invariant et parfaitement connu, le robot doit décider de ses actions sur la base de son interprétation de l'état de l'environnement et de son propre état, anticiper l'état futur, raisonner sur les conséquences de ses actions, effectuer des choix. Et, les connaissances opératoires de planification ou de supervision ne pouvant pas toutes être prévues a priori, des mécanismes d'apprentissage sont nécessaires pour permettre l'acquisition de compétences supplémentaires, et l'affinement des compétences existantes ;

- interaction et coopération ; le robot agit dans un environnement peuplé d'êtres

humains et/ou d'autres robots. Il doit donc posséder des capacités conjointes de communication et d'action. Qu'il s'agisse de coordination ou de coopération, l'interaction est cognitive et/ou physique. Il est donc nécessaire d'utiliser des modèles des capacités perceptives, d'action et des contraintes mécaniques de l'homme, ainsi qu'une perception multimodale de son visage, de ses gestes, de ses postures et de ses attitudes pour les reconnaître et les interpréter.

Ces activités sont réalisées dans des laboratoires qui relèvent de nombreux opérateurs de recherche (CNRS, INRIA, CEA, ONERA, INSERM, CEMAGREF, INRETS) et qui « mailent » bien le territoire national. Il n'existe toutefois que très peu de laboratoires dont la seule thématique soit la robotique, et les équipes de roboticiens voisinent ainsi souvent des automaticiens, des informaticiens ou des mécaniciens.

La robotique peut également être abordée sous l'angle des applications, qui dépassent depuis longtemps le cadre des industries manufacturières (encore aujourd'hui trop peu considérées en Europe), pour investir les champs de l'exploration planétaire ou marine, de la défense-sécurité, du transport, de l'aide au geste médical ou aux personnes dépendantes, des activités ludiques et du service à domicile, par exemple. Dans les trois dernières années, le premier aspirateur « robotique » grand public s'est vendu à plus de deux millions d'exemplaires : ses capacités robotiques sont celles que l'on trouvaient dans les laboratoires il y a plus de vingt ans.

Parmi les avancées les plus visibles car matérialisées par des réalisations abouties, on peut remarquer les premiers robots guides de musée, assistants de chirurgiens, animaux ludiques, patrouilleurs aériens ou sous-marins... Ces objets technologiques reposent tous sur des avancées scientifiques majeures en conception, en commande référencée vision, en planification, en interaction homme-robot, etc. Deux réussites peuvent être mises particulièrement en exergue : les premiers robots humanoïdes réalistes sont enfin disponibles, bien qu'encore très loin d'être autonomes, et des

robots mobiles autonomes ont réalisés avec succès, et à grande vitesse, des parcours en milieux naturels de plus de 200 km.

Sans atteindre la «force de frappe» de pays comme le Japon où la robotique fait partie des sciences reines (aux plans académique et industriel), ou comme les USA où la robotique est clairement financée comme une technologie duale civile-militaire, la recherche en robotique française peut toutefois afficher des équipes de pointes dans chacun des quatre thèmes cités plus haut. En revanche, elle ne peut pas mettre en avant des réalisations technologiques intégrées aussi spectaculaires que celles que présentent le Japon (humanoïdes) ou les USA (robots mobiles d'extérieur autonomes sur très grande distance). Cette faiblesse dans le volet «intégration» s'explique par des failles qui sont pour certaines internes à la communauté (manque d'une structure pérenne d'animation et de représentation nationale), et pour d'autres, liées au contexte national général (absence d'une industrie support en France; liens distendus avec la recherche de défense).

Toutefois, la tendance générale semble très favorable à la recherche en robotique au niveau international (Japon, USA et Corée mettent en place des budgets très conséquents), au niveau européen (le réseau EURON et la plateforme EUROP sont très dynamiques) et enfin, récemment en France, avec le remarquable P.I.R. ROBEA et le tout nouveau programme PSIROB de l'ANR, même si le niveau d'intervention en France reste encore modeste en comparaison. L'activité et la visibilité d'ensemble de la communauté française de recherche en robotique doivent pouvoir en profiter. Le CNRS, seul opérateur de recherche à couvrir l'intégralité du domaine robotique «de la conception à la cognition» et à offrir une ouverture naturelle vers les SDV et les SHS, a une responsabilité majeure en la matière.

8 – ÉLECTRONIQUE INTÉGRÉE ET ARCHITECTURE DE SYSTÈMES

Ce domaine embrasse la conception et le test de systèmes intégrés complexes, composés de nombre toujours croissants de transistors, et/ou plus récemment de MEMS; il s'agit d'un des domaines qui assurent la continuité thématique entre les sections 07 et 08 du CNRS.

L'évolution rapide des technologies de fabrication de circuits intégrés sur silicium permet déjà de réaliser des systèmes numériques complets intégrés sur une même puce (SoC pour *System on Chip*) et plus récemment de systèmes assemblés (SiP pour *System in Package*). Les techniques de conception de systèmes électroniques évoluent vers l'intégration de systèmes de plus en plus complexes. Des blocs fonctionnels ou «composants virtuels» sont de plus en plus utilisés. Les classiques bibliothèques utilisées en conception de circuits intégrés pour applications spécifiques seront complétées, voire remplacées, par des bases de données de composants virtuels dont la fonctionnalité pourra correspondre à un processeur ou même à un ordinateur complet avec sa mémoire et ses entrées/sorties. Les outils de conception assistée par ordinateur pour la microélectronique doivent ainsi prendre de plus en plus en compte les aspects système. Des alternatives d'architectures (avec des propriétés de reconfiguration) doivent permettre de trouver un équilibre entre contraintes d'efficacité et de flexibilité. Le logiciel embarqué ou enfoui permet de mieux adapter le circuit à une application, le logiciel et le matériel cohabitent dans un système intégré sur un mode nouveau.

Lorsque la fonctionnalité complète du système électronique se trouve intégrée sur un substrat unique de silicium, on assiste à un accroissement de la complexité du problème de conception. Plus que d'un simple changement d'échelle, il s'agit en fait d'une modifica-

tion totale de la méthodologie de conception et de validation pour gérer entre autres les nouveaux degrés de liberté : gestion des ressources matérielles, des interactions logiciel-matériel, des contraintes environnementales, des interactions hétérogènes. Ces systèmes intégrés complexes sont déjà présents dans les objets communicants actuels et seront au cœur de l'évolution des objets et des usages des technologies de l'information.

Les défis qui se posent à la communauté de chercheurs qui se concentrent sur l'électronique intégrée comprennent des volets qui concernent prioritairement la section 08 (par exemple : consommation électrique, circuits analogiques), mais des pans entiers sont bien au cœur de la problématique de la section 07 : méthodologies structurées de conception en vue du test de SoC (problèmes d'accès et de volume de données) ou de SiP (test avant/après assemblage), glissement de la problématique « tolérance aux fautes » vers une approche « robustesse au variation de paramètres » inhérente aux technologies poussées aujourd'hui à leurs limites, prise en compte des architectures reconfigurables dynamiquement, des architectures d'interconnexion ou des architectures multiprocesseurs hétérogènes, recherche de l'adéquation logiciel/matériel en particulier dans un contexte « temps réel ».

9 – CONCLUSION

Depuis ces dix dernières années, toutes les actions stratégiques au niveau international, européen et national ont participé à la révolution des composants logiciels et matériels autant qu'à celle de l'internet. Les avancées scientifiques et technologiques effectuées durant cette décennie, se trouvent à l'origine de l'émergence de nouvelles capacités qui sont loin d'être toutes maîtrisées et qui ont généré

soit de nouveaux besoins, soit des besoins accrus. L'un des moteurs les plus puissants de cette révolution en marche est l'omniprésence des STIC dans toutes les activités humaines : services, santé, transports, environnement, loisirs, etc.

Cette diffusion des STICS modifie jusqu'à la manière de concevoir, réaliser et créer en même temps qu'elle engendre de nouvelles activités et interactivités dans tous les secteurs d'activités : industriel, commercial, culturel et artistique mais bien sûr aussi scientifique.

L'autre aspect saillant des STIC aujourd'hui est le caractère enfoui des nouveaux matériels, processus de calcul et sources de données qui participe à l'intelligence ambiante et ubiquitaire. Les ordinateurs ne constituent désormais qu'un petit pourcentage des ressources de calcul et des sources de données. Plus de 80% des processeurs sont intégrés à des objets divers des plus simples au plus sophistiqués. Le nombre de capteurs intégrés dans des composants avec des unités de traitement et de transmission de signaux connaît une croissance vertigineuse. Chaque capteur est un nœud actif d'un système dont les capacités locales de traitement permettent d'agréger, trier, filtrer des données ou effectuer des traitements plus sophistiqués.

L'informatique ubiquitaire et ambiante en rendant l'information tout autant que les ressources de calcul plus facilement accessibles induisent de nouveaux comportements et aussi de nouveaux risques. Ces risques sont également amplifiés par la puissance croissante des calculateurs qui permettent le croisement et la fouille de masses de données.

On le voit, les enjeux scientifiques portés par les STIC croisent des besoins, des espoirs parfois, des risques aussi, partagés par la société toute entière : nos efforts, nos moyens, notre enthousiasme doivent être à hauteur de ces attentes.

Notes

(1) <http://www.inria.fr/inria/historique.fr.html>

(2) Voir le site du Ministère des Affaires Étrangères <http://www.diplomatie.gouv.fr>

(3) Source : rapport de synthèse « Recherche et développement en STIC dans les grands pays industriels » rédigé par le Groupement Français de l'Industrie de l'Information à la demande du Ministère délégué à l'enseignement supérieur et à la recherche, avril 2006.

(4) Dollar at purchase power parity : les montants bruts sont corrigés pour tenir compte du pouvoir d'achat réel évalué dans chaque pays. Sans être parfaite, cette unité cherche donc à donner une image « juste » permettant une réelle comparaison internationale. Cette correspondance est établie par l'OCDE.

(5) Source : labintel, printemps 2007.

(6) Source : enquête STIC 2004, et Labintel.

ANNEXE

LISTE DES ACRONYMES

ANR	Agence Nationale de la Recherche	CR	Chargé de recherche
BAP	Branche d'Activité Professionnelle	DR	Directeur de Recherche
BAP A	BAP Sciences du vivant	EURON	EUropean RObotics research Network
BAP B	BAP Sciences chimiques et des matériaux	EUROP	European Robotics Platform
BAP C	BAP Sciences de l'ingénieur	FRE	Formation de Recherche en Évolution
BAP E	BAP Informatique	GDR	Groupement de Recherche
BAP F	BAP Documentation et communication	GIPSA	Laboratoire Grenoblois de l'Image, de la Parole, du Signal et de l'Automatique
BAP G	BAP Patrimoine, logistique	GET	Groupe des Écoles des Télécommunications
BAP H	BAP Gestion	IATOS	Personnels Ingénieurs, Administratifs, Techniciens et Ouvriers de service
CEA	Commissariat à l'Énergie Atomique	INRETS	Institut National de Recherche sur les Transports et leur Sécurité
CEMAGREF	Institut de recherche pour l'ingénierie de l'agriculture et de l'environnement	INRIA	Institut National de Recherche en Informatique et en Automatique
CNU	Conseil National des Universités	INSERM	Institut National de la Santé et de la Recherche Médicale
CNU 27	CNU section « informatique »	ITA	Personnels Ingénieurs, Techniciens et Administratifs
CNU 61	CNU section « Génie informatique, automatique et traitement du signal »	LAAS	Laboratoire d'Analyse et d'Architecture des Systèmes

LAGEP	Laboratoire d'Automatique et de Génie des Procédés	P.I.R.	Programme Interdisciplinaire de Recherche
LIG	Laboratoire d'Informatique de Grenoble	PR	Professeur
LIMSI	Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur	PSIROB	Systèmes interactifs et robotiques
LIST	Laboratoire d'Intégration des Systèmes et des Technologies	ROBEA	Robotique et Entités Artificielles
MC	Maître de Conférence	SDV	Sciences du Vivant
MEMS	Micro Electro-Mechanical Systems	SHS	Sciences Humaines et Sociales
NSF	National Science Foundation	STIC	Sciences et Technologies de l'Information et de la Communication
ONERA	Office National d'Études et de Recherches Aérospatiales	UMI	Unité Mixte Internationale
		UMR	Unité Mixte de Recherche
		UMS	Unité Mixte de Service
		UPR	Unité Propre de Recherche