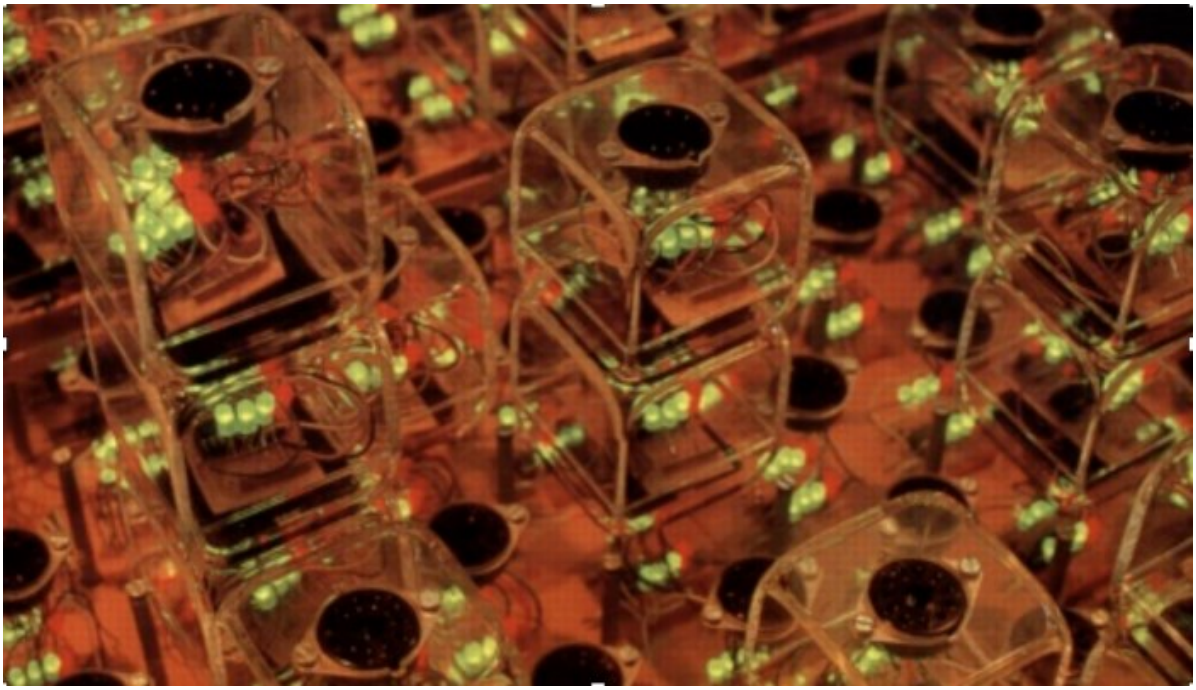
Réinventer l'informatique : Au-delà des architectures Von Neumann

EDF-Lab – Palaiseau

Jeudi 6 décembre 2018



Coordination Scientifique

Christophe Calvin (CEA), David Menga & Christophe Denis (EDF Lab)

Editorial Board

Dr. Christophe Calvin (CEA)

Mr. Laurent Duplouy (BnF)

Mr. Philippe Wlodyka (Polytechnique)

Mr. Pascal Pavel (CEA)

Dr. Thiên-Hiêp Lê (ONERA)

Ms. Katia Castor (Association Aristote)

Réinventer l'informatique : Au-delà des architectures Von Neumann

Séminaire Aristote, 06/12/2018 à EDF-Lab

Coordination scientifique

Christophe Calvin (CEA), David Menga & Christophe Denis (EDF Lab)

Table des matières

Réinventer l'informatique : Au-delà des architectures Von Neumann 3

Compte-rendu des interventions 5

1. Fin de l'ère « logique binaire-architecture Von Neumann-CMOS » : tapage ou véritable « nouveau monde » ? 6

2. Les challenges posés par les systèmes de calcul pour les applications cyber-physiques cognitives 8

3. Utiliser la lumière pour éclairer l'avenir de l'intelligence artificielle 10

4. Vers une IA organique intégrée 12

5. Bayes, de la cellule à la machine 14

6. Introduction au calcul quantique et état de l'art 16

7. Vers un traitement de l'information quantique à grande échelle basée sur les spins électroniques dans les semi-conducteurs 18

8. La programmation quantique 20

9. Investir dans les technologies quantiques : au-delà de la science, des marchés et des opportunités pour les entrepreneurs 22

10. L'optimisation quantique – le problème, le potentiel et l'application 24

11. Le calcul quantique et ses applications en chimie et en physique 26

Compte-rendu des interventions

Introduction

David Menga (EDF), qui a préparé ce séminaire avec Christophe Calvin (CEA) et Christophe Denis (EdF-Lab) qui ne peut être présent, souhaite la bienvenue à ce séminaire spécial qui a l'ambition de révolutionner l'informatique.

Il donne la parole à Laurent Billet, délégué scientifique d'EDF R&D, qui nous accueille dans ses locaux. Celui-ci a relu le rapport de Cédric Villani. Il s'est rendu compte que comme lui il a été influencé par Charles Babbage et Alan Turing. L'architecture de Von Neumann constitue la base de tous les ordinateurs d'aujourd'hui. Alain Aspect avait prouvé que Dieu ne jouait pas au dé, que l'aspect probabiliste était constitutif de notre réalité. Aujourd'hui l'ordinateur quantique sort des prophéties de Richard Feynman pour rejoindre la réalité. EDF, qui utilise beaucoup les simulations numériques, surveille avec attention les percées de l'informatique moderne. Laurent Billet est heureux d'accueillir le séminaire Aristote sur le sujet. EDF s'intéresse plus particulièrement à l'ordinateur quantique pour explorer les arbres de défaillances dans les centrales ou programmer, mais aussi à la cryptographie en vue de la sécurité. Il indique le projet piloté par Cédric Baudry, associé au CEA (Bruyères-le-Châtel), Atos et d'autres industriels (Dassault Aviation, Total, etc.) et des instituts académiques. Car la loi de Moore ne sera bientôt plus valable. Les ruptures qui se profilent peuvent ouvrir aux Européens des opportunités.

Laurent Billet passe la parole à Christophe Calvin, président d'Aristote, qui donne un éclairage sur ce qui a guidé David Menga et lui-même à proposer cette journée. Les nouveaux enjeux sur le big data, l'augmentation de la puissance de traitement et la maitrise de la consommation électrique indique que les architectures Von Neumann arrivent à leur limites. Et il y a les limites de la physique. Comme Laurent Billet, selon lui, la loi de Moore ne sera plus d'actualité. Le moment est-il venu d'inventer l'informatique du XXIe siècle ? David Menga reprend la parole pour parler de la structure de la journée. Il s'agit d'abord de présenter le problème. Est-ce que tout ce qu'on a connu est dépassé ? Et pourquoi changer de paradigme ? Quels sont les nouveaux défis ? Ensuite on entrera dans le détail. Comme l'OPU à la place du GPU. L'après-midi est consacré au quantique avec les diverses architectures, la programmation et les nouveaux usages envisagés comme Volkswagen qui veut trouver la bonne molécule pour les batteries. Mais avant tout, une revue de la question par le professeur Alain Cappy.

1. Fin de l'ère « logique binaire-architecture Von Neumann-CMOS » : tapage ou véritable « nouveau monde » ?

Alain Cappy (Université de Lille)

Alain Cappy, professeur à l'Institut de Recherche sur les Composants logiciels et matériels pour l'Information et la Communication Avancée de l'université de Lille, va balayer très largement la problématique avant que les autres intervenants ne regardent dans le détail.

Il y a 50 ans, l'homme allait sur la Lune. Mais l'ordinateur de bord d'Apollo comportant 2800 circuits intégrés avec de la logique peu performante, de fréquence faible (80 Hz). Aujourd'hui, chaque minute, on envoie 187 millions d'emails, 18 millions de SMS, 400 000 gazouillis. La progression vient du codage binaire de l'information, l'architecture de Von Neumann et le CMOS, une technologie matériel qui possède la propriété de changement d'échelle. Le nombre de transistors par puce a augmenté à quelques milliards. À partir de 2005, le nombre de cœurs s'est multiplié, la fréquence d'horloge a commencé à stagner. Comme la puissance de 100 W/cm^2 est une limite indépassable, les fréquences des processeurs a donc aussi stagné. La limite de 100 W/cm^2 est énorme. On peut refroidir par eau comme IBM ou même par azote liquide. Alain Cappy revient 50 ans en arrière, avec la parution de « *2001, l'Odyssée de l'espace* » avec l'intelligence artificielle Hal qui prend le contrôle du vaisseau. Aujourd'hui, on assiste à la victoire d'AlphaGO sur Lee Sedol. Or AlphaGO c'est 1202 GPU et 300 kW au minimum. Lee Sedol c'est seulement 20 à 30 W. La compétition aurait été équitable à puissance équivalente. Le véhicule autonome, c'est la même chose. Les 50 à 100 processeurs ont une puissance

de 5 kW alors qu'un conducteur humain a une puissance de 30 W. Alain Cappy se demande donc si le véhicule autonome est une si bonne chose.

Puis il se demande quels seraient les besoins actuels. La robotique, l'intelligence d'appareils nomades… Mais à quelle consommation énergétique ? Pourquoi le CMOS dépense-t-il autant d'énergie ? En logique binaire, la tension (ou la charge) donne la valeur. En thermodynamique, si on bouge des charges, on dissipe de l'énergie. Si on change de transistor, le problème sera le même. La majorité des pertes se fait dans l'interconnection. Le transistor est conducteur si on applique une tension de 0,3 volt. Pour le fermer, il faut appliquer une tension de l'ordre du volt. Comment faire pour changer ça ? Evoluer ou révolutionner ? À partir du triplet « logique binaire-von Neumann-CMOS », on peut faire du computing memory ou de la logique adiabatique (avec tension variable plutôt que fixe). Autre solution, travailler à moins de 1 volt avec de nouveaux switchs ou changer la variable d'état. Ensuite, la révolution serait de ne plus se baser sur le triplet. De nouveaux transistors sont proposés. Les variables d'états différentes foisonnent depuis 20 ans, mais rien n'a percé.

La révolution est de ne plus faire de codage binaire comme l'ordinateur quantique ou neuromorphique. Il faut des approches multidisciplinaires et démontrer au niveau système que cela fonctionne. Le quantique repose sur des ions, des photons, le caractère aléatoire des mesures et le principe de superposition d'état. En lançant une pièce 50 fois, il faudrait 100 terabytes de stockage mais seulement 50 qubits. Les annonces d'IBM ou de Google n'ont pas de base scientifique. Il faut s'en méfier. Il y a des problèmes d'algorithmique, de temps de cohérence quantique (100 µs seulement), de température très faible (IBM 100 mK).

La deuxième solution est le neuromorphique ou comment copier le cerveau (2% du poids de l'homme mais 20% de sa consommation d'énergie). Il contient 10^{11} neurones et 10^4 synapses par neurone. Dans le cortex, 100 000 neurones par cm3. Une machine doit faire du calcul et communiquer. Le cerveau fait de l'intégration analogique et numérique pour la partie communication. On peut copier le cerveau comme McCulloch et Pitts en 1943 avec une avancée de F. Rosenblatt en 1958. John Von Neumann s'est aussi intéressé au cerveau (*The commuter and the Brain*, 1958). Ce n'est donc pas nouveau. Le Deep Learning permet d'apprendre en faisant des erreurs. Mais c'est très loin du fonctionnement du cerveau, qui agit par prédiction et par feedback. Ce n'est pas pris en compte dans les travaux. Autre possibilité, faire des composants qui émuleraient des neurones et des synapses, comme Intel et sa puce Loihi ou l'architecture TrueNorth d'IBM, qui permet d'ajouter des puces aux autres. Aujourd'hui, on est dans une période faste sur ce sujet, mais cela a beaucoup varié dans le passé. On ne sait pas traiter des réseaux avec beaucoup d'entrées-sorties (feedback). Le cerveau a une architecture très interconnecté 3D, l'informatique va avoir des problèmes pour le copier vraiment.

En conclusion, Alain Cappy insiste pour que les solutions soient être ultra-économes en énergie.

2. Les challenges posés par les systèmes de calcul pour les applications cyber-physiques cognitives

Marc Duranton (CEA/DRT)

Marc Duranton est chercheur sénior au CEA. Sa présentation est le miroir de celle Alain Cappy avec des cas concrets. Il commence par la citation d'Alan Kay : « la meilleure façon de prédire le future est de l'inventer ». Il y a 70 ans, il y avait un seul d'ordinateur sur terre. Il y a trente ans, on pensait qu'il n'y aurait pas d'ordinateurs à la maison. Et l'Iphone n'a que 11 ans. Alors que la puissance de calcul de Summit est de 200 Pflops ($2x\ 10^{17}$)), elle n'était que de 2 Gflops ($2x\ 10^{9}$) il y a 33 ans avec le Cray 2.

L'amélioration est gigantesque. La loi de Moore qui était une loi économique s'est transformée en une loi physique. Qu'est-ce qu'il y aura après le CMOS, qui arrive à ses limites ? Avant, on diminuait la taille par 2, le courant par 2, mais on ne peut pas diminuer le voltage. Les transistors commencent à fuir par le substrat. On a trouvé des solutions pour éviter ces fuites. La prochaine génération des nanowires sera contrôlé de tous les côtés avec des tailles de 3 à 5 nm, mais les effets quantiques commencent à se faire sentir. Si on les refroidit, on parvient à des qubits. Autre solution, revenir à des électroniques à relais avec des microswitchs ou empiler les transistors plutôt que les mettre en plan. On gagne de la surface, la performance est améliorée et la consommation électrique est plus basse.

Le film *Her* montre une informatique qui est devenue une entité, une collaboration entre l'homme et la machine, permise par le Deep Learning. Cela a effectivement sorti l'ordinateur du cyberespace pour le mettre en lien avec le monde physique. Il ne reconnaît plus seulement des chiffres ou des lettres mais un monde réel. Mais

l'informatique n'a pas été créée pour ce monde réel. Très peu de langages informatiques tiennent compte du temps, or c'est un problème pour les transports où la machine doit répondre dans un temps donné précis et non moyen. C'est un des challenges à relever. En France, on essaie de régler ce type de problème comme l'énergie en diminuant la communication des données, qui sont dupliquées n fois. Autre point, une machine cyberphysique doit interagir avec le monde physique. Il faut repenser l'informatique et les modèles de calculs. Comme tout se fait sur le Cloud, s'il y a une rupture de communication, cela ne fonctionne pas. Une voiture autonome a le coffre plein d'ordinateurs qui consomment et envoient leurs données au Cloud. À quoi cela sert ? , se demande Marc Duranton. C'est pareil pour les assistants personnels. Google s'est rendu compte que si tout le monde utilisait son assistant personnel pendant 3 minutes, ils devraient doubler leur nombre d'ordinateurs. Ils ont donc transformé leur puissance de calcul.

AlphaGo Zéro a appris tout seul alors que AlphaGo a été nourri du big data. Ces centres de data vont poser un problème énergétique. Pour récupérer les données, il faut 100 fois plus d'énergie que pour les calculer. Il faut donc rapprocher la communication des calculs. Le CEA a multiplié par 10 ainsi sa capacité énergétique et par 100 sa puissance de calcul. Autre solution, utiliser l'optique, car un photon, bien que cher à créer, va très loin.

Comme Alain Cappy, Marc Duranton évoque le modèle du cerveau avec la méthode Spike Timing Dependent Plasticity. La fabrication de neurones synthétiques consomme effectivement très peu. Et cela peut permettre de mieux comprendre comment fonctionne le cerveau. À côté du TrueNorth d'IBM et du Loichi d'Intel, DynapSel est une approche européenne. La machine de demain comprendra différents accélérateurs comme des réseaux de neurones, une haute densité 3D, de nouvelles mémoires plus denses utilisant l'optique. Les orateurs suivant vont aller plus loin sur les différentes approches à l'étude. Pour Marc Duranton, « *nous sommes à la croisée de chemins, l'avenir nous appartient* ».

3. Utiliser la lumière pour éclairer l'avenir de l'intelligence artificielle

Igor Carron (LighOn)

Igor Carron, PDG de la société LightOn, une société qui crée une technologie disruptive en utilisant la lumière.

Le calcul optique n'est pas récent. Cela a commencé en 1935 pour régler des questions mathématiques sur les nombres premiers. Mais l'essor est venu des années 60 avec les militaires. Puis ce fut l'hiver. Mais le printemps arrive aujourd'hui. La question n'est pas seulement comment changer la loi de Moore, mais comment régler des problèmes de calculs pour tout le monde. Ce qui se fait un peu aujourd'hui dans les data centers.

Certains disent qu'en 2040, 100% de l'énergie sera utilisée dans les calculs. La demande est énorme, si énorme qu'on se demande si la capacité est faisable. Or la majorité de la dissipation d'énergie est utilisée dans les interconnections comme les deux autres intervenants l'ont déjà dit. Il faut interconnecter les composants afin d'être le plus efficace. Avec Laurent Daudet (Paris-Diderot, ESPCI) et les autres collaborateurs de LightOn on s'intéresse au Compress Sensing afin de résoudre des problèmes dits sous-déterminés. De nombreux articles sont publiés sur le sujet depuis 2005. Les projections aléatoires permettent de comprendre le monde au niveau de l'information. Sylvain Gigan, le spécialiste du milieu diffus de LightOn, a montré que les informations concernant les photons dans ces milieux sont perdues. À Light on, on fait des franges d'interférences dans ces milieux, ce qui permet de coder les informations et de faire de la compression universelle. Ces matrices de projections aléatoires dans les milieux diffus sont comme des réseaux de neurones. Ces milieux sont des mémoires. Il est possible d'effectuer 10^{15} opérations par seconde. lightOn a été créée en 2016, de l'argent a été

levé en juin 2017 quand a été créé le premier OPU (puissance 30 W) afin de répondre à la demande de calcul. La communauté de Machine Learning ne s'intéresse pas à la technologie, mais veut que cela marche. Avec lightOn, on multiplie par 8 la vitesse et on réduit la consommation de 90%. Igor Carron donne des exemples effectués par LightOn comme NewMa, la détection de changement dans des grandes images vidéo. ou le crowdsourcing pour les recommandations de DVD pour Netflix. LightOn voit cette informatique comme hybride, car ce n'est pas un ordinateur programmable mais il demande des calculs.

Il propose en conclusion d'inviter la communauté de beta-users à tester la machine. À suivre sur http://Lighton.io

4. Vers une IA organique intégrée

Patrick Pirim (Another Brain)

Patrick Pirim est le directeur technique d'Another Brain. Il est ingénieur électronicien de formation, mais autodidacte pour la création de circuits qu'il a commencé à fabriquer en 86.

Comme les autres orateurs l'ont dit, la question de l'énergie est importante. Il faut copier ce que l'évolution a créé. Le cortex est un ensemble de 360 zones selon un article publié en 2016, avec des gros modules (arrière pour la perception, centrale pour les commandes, frontal pour la cognition). En regardant la colonne corticale, on peut détecter des mouvements ou des taches selon Hubel et Wiesel, les prix Nobel de médecine 1981, alors que Bach-y-Rita a montré avec des aveugles qu'on ne voit pas avec ses yeux mais avec son cerveau. Pour Patrick Pirim, quand on détecte des mouvements, il y a une modalité dynamique, une modalité structurale (détection de formes) et une modalité globale (détection des couleurs). Cela se fait en électronique beaucoup plus vite.

En 1986, il a récupéré son premier circuit bioinspiré, un traducteur universel. Avec des mémoires associatives, cela crée des liens et permet d'anticiper les réactions. Il montre deux vidéos, la première étant le suivi d'une balle qui oscille et la seconde, la conduite automatique en suivant le tracé de lignes blanches. Mais en 2001, les constructeurs pensaient que cela ne servait à rien. Il montre ensuite comment il fait pour récupérer des saillances et montre deux modes de perception visuelle, avec des différences de couleurs ou de spatialité. Pour la nature, la vision est normale jusqu'à 15°. Les connections dans le cerveau sont très rapides (il y a des analogies très rapides entre phonèmes et graphèmes).

Depuis 1986, il a conçu de nombreux circuits dont le dernier, le blue phophor, est prévu pour 2020.

5. Bayes, de la cellule à la machine

Pierre Bessière (UPMC/CNRS)

Pierre Bessière est directeur de recherche au CNRS et co-fondateur de la société ProbaYes.

Il commence par poser plusieurs questions. Les probabilités sont-elles une bonne suite de la logique. Peut-on faire de la programmation avec les probabilités ? Le cerveau est-il bayesien ? et enfin est-il possible de créer du hardware probabiliste ? Pierre Bessière prend comme exemple le jeu d'échec. Qui est le plus intelligent ? On le sait, c'est l'ordinateur. Mais quand il s'agit de jouer et de déplacer les pièces, un enfant le fait facilement, alors qu'un robot ne peut pas le faire pour tous les types d'échiquier. Le jeu est un problème fermé : il suffit de connaître l'échiquier, les pièces et les règles. Le déplace ment des pièces est quant à lui un problème ouvert car il faut savoir comment commander les mouvements, comme se fait la réflexion de la lumière sur les pièces, etc. Faire de la compression d'images est un problème fermé, faire de la classification est un problème ouvert. Mais il y a aussi des problèmes ouverts qui n'ont pas de correspondance en problème fermé comme le diagnostic médical, la détection de fraudes, la prévision du risque financier, etc. Les problèmes fermés sont résolus par les ordinateurs classiques, mais les problèmes ouverts demandent une informatique du futur.

Il faut une informatique basée sur les probabilités. Beaucoup de gens y travaillent dont l'équipe de Pierre Bessière. Il faut spécifier les variables, la distribution conjointe de ces variables (multiplication des conditions indépendantes). La programmation bayésienne est une façon algébrique de représenter des réseaux bayésiens. C'est un langage déclaratif qui permet de poser des questions. Le bayesian programming se situe entre

les méthodes empiriques et les méthodes logiques. C'est ce qu'a fait ProbaYes, rachetée par La Poste.

Autre question, le cerveau ou la cellule utilisent-ils les probabilités ? Pierre Bessière, quand il était au Collège de France avec Alain Berthoz, a travaillé dessus comme beaucoup d'autres dans le monde. Les réseaux de neurones formels sont une injure au biologiste, car une cellule est très complexe et a des capacités de traitements très étendue comme la capacité à générer des bits stochastiques. Exemple, un canal dans la cellule peut être ouvert ou fermé selon le bruit extérieur. Différents dispositifs physiques peuvent reproduire cela avec le *magnetic tunnel junction* et le bruit thermique. Dans un projet européen terminé il y a un peu plus d'un an, il a été étudié une algue unicellulaire avec des macromolécules qui peut prendre deux états binaires et qui émet des signaux à d'autres macromolécules qui à leur tour prend deux états. Et ainsi de suite. L'algue peut ainsi aller vers la lumière quand elle en a besoin. On a montré qu'un modèle de chaine de Markov pouvait faire la même chose (voir le cours de Stanislas Dehaene au Collège de France sur le sujet). Patrick Bessière s'est aujourd'hui lancé sur la création au 1er janvier 2019 de la société HawAI.tech (harware for IA). Le but est d'approfondir la programmation bayésienne avec des processus de perception et de synthétisation, de calculs et d'apprentissage. Chose étonnante, le CTO d'Intel a la même idée. La première étape est la résolution de problèmes probabilistes avec des simulations de Monte Carlo, puis la fabrication des (B)Asic, des Asics stochastiques de faible consommation ciblé pour l'IoT. Exemple, des constellations de capteurs sur des grands équipements. Contraintes : trouver de l'énergie sur place et transmettre les données de façon hertzienne et épisodiquement afin de consommer moins. (B)asics fait la fusion des capteurs, du filtrage temporel, de la prédiction d'événements et de l'envoi d'alertes.

HawAI.tech veut aussi créer des (B)PU, un ensemble de coprocesseurs dédiés aux simulations Monte Carlo. L'équipe est composée de 5 jeunes doctorants, de 4 chercheurs expérimentés, d'un business développeur. Un industriel y est associé.

6. Introduction au calcul quantique et état de l'art

Denis Vion (CEA/DRF)

Denis Vion est physicien au CEA, spécialiste en circuits quantiques.

Il explique qu'il y a plusieurs types d'ordinateurs quantiques. Le plus commun et le plus avancé est à portes logiques. Il est copié sur l'ordinateur classique. C'est une machine de Turing quantique. Il en existe d'autres comme l'ordinateur quantique adiabatique déjà commercialisé ou le one-way quantic computing. Il commence son exposé par les bases de l'ordinateur quantique. L'ordinateur classique est une machine dans lesquels entrent des données et en sortent d'autres. Tout est codé en information élémentaire, les bits (1 ou 0). En 1937, Turing avait compris qu'on pouvait faire des calculateurs universels sur ce principe. Il peut y avoir des erreurs hardware bien qu'aujourd'hui il n'y en ait plus. La redondance permettait d'éviter ces désagréments. On fait de même pour les ordinateurs quantiques mais dans le monde quantique. Les bits ne sont plus à deux éléments, mais un ensemble d'états superposés. Il donne l'exemple avec des photons passant par deux lames semi-transparentes. En fait, le photon passe à la fois par les deux portes car il a des dons d'ubiquité. C'est l'exemple d'un bit quantique, un qubit. Mais quand il faut mesurer le passage, on retombe dans le monde classique et le résultat est soit 1 soit 0. Ce n'est pas grave car tout dépend de ce qui s'est passé avant. On va alors avoir une probabilité de 0 et de 1 (Qb = □/0> + ß /1>). Alors que la seule opération possible en physique classique est la transformation de 0 en 1 ou inversement, en physique quantique, toutes les rotations sont possibles. Mais ce n'est pas là la puissance de l'ordinateur quantique, car la mesure fixe encore une fois les choses. La puissance est quand on met plusieurs bits quantiques. Avec n qubits, on a 2^n états possibles. La porte à deux qubits est le « Ou » exclusif. Quand deux particules sont intriquées, un bit de contrôle est la superposition des deux états, alors que le deuxième est fixé à 0. On peut

faire ainsi un ensemble d'opérations, mais à la fin on fera la mesure sur une partie de l'échantillon. L'algorithme de Shor (1994) permet de cette façon de calculer la période d'une fonction périodique. La mesure fixe bien un résultat mais c'est le bon car il est représentatif des valeurs possibles. C'est du parallélisme massif et de l'utilisation intelligente de l'intrication. C'est beaucoup plus rapide que la méthode classique. Mais les superpositions quantiques sont sensibles à la décohérence qui survient après 100 µs. C'est la mort subite d'un qubit qui passe de l'état 1 à l'état 0 sur un temps dit de relaxation. Tout devient faux. Autre problème, le pure déphasage où un qubit est sensible au bruit et modifie les probabilités d'états. Comment résoudre ces erreurs ? La redondance n'est pas valable car on ne peut pas recopier un qubit sans le mesurer. Mais la relaxation d'un système à 3 qubits résout la question car un sous-ensemble peut détecter s'il y a eu erreur et le corriger, à condition qu'il y ait un taux d'erreur de seulement 0,01% par porte. L'invention du surface code permet de tolérer des erreurs à 1% mais il faut alors entre 100 et 1000 qubits. Ce n'est pas encore possible. En attendant les physiciens s'arrangent des erreurs.

On a essayé de fabriquer ces machines avec les spins d'atome de molécules organiques, puis avec de la lumière en créant des pistes à photons qui se rejoignent de temps en temps et permettent au photon de passer d'une piste à l'autre. Autre solution, des atomes portant un spin électronique ou nucléaire que Tristan Meunier va détailler dans l'intervention suivante. Le plus précis sont les pièges à ions. Mais les plus prometteurs sont les superconducteurs quantiques comme va le détailler Ivano Tavernelli d'IBM. Jusqu'en 2013, il n'y avait des études que dans le monde académique, mais maintenant les grosses compagnies s'y mettent (IBM avec un computer à 17 qubits, Google et UCSB à 50 qubits, Intel, Microsoft ou Atos), mais aussi des petits compagnies.

Denis Vion a lui-même fabriqué ce genre de machine, mais le temps lui manque pour détailler ces travaux. C'est un monde très concurrentiel où les effets d'annonce sont à prendre avec des pincettes.

7. Vers un traitement de l'information quantique à grande échelle basée sur les spins électroniques dans les semi-conducteurs

Tristan Meunier (CNRS)

Tristan Meunier, physicien quantique de l'Institut Néel, explique qu'à Grenoble, l'écosystème scientifique permet d'avancer sur le sujet de l'ordinateur quantique.

Comme Denis Vion, il explique un peu le monde quantique dans lequel un chat peut être à la fois mort et vivant. En Quantic Computing, on peut faire à la fois des opérations sur 0 et sur 1 contrairement à l'ordinateur classique. Au lieu de doubler le nombre de bits classiques, il suffit d'ajouter un seul qubit. Aujourd'hui la suprématie quantique se situe à 56 qubits sans erreurs. Mais pour trouver des médicaments ou des matériaux innovants, il faudra 200 qubits logiques. Les applications sont exponentielles comme les investissements. L'équipe de Tristan Meunier veut construire un ordinateur quantique et doit résoudre des questions : chaque qubit est-il en connection avec tous les autres ou seulement avec quelques-uns? Il vaut mieux des qubits en connexion avec ses plus proches voisins, car le nombre de contrôles des paramètres est le carré du nombre de qubits. En 1, 2 ou 3D ? En ligne (1D), tous les qubits doivent interagir avec les autres et le nombre d'opérations à effectuer explose. En 2D, cela se passe en deux directions. Tous les qubits ne doivent ainsi pas être connectés ensemble, mais à seulement 4. Tristan Meunier travaille sur les qubits semi-conducteurs, le système le moins avancé. Le qubit est donné par le spin d'un électron dans un semi-conducteur. Il est irradié afin de faire des rotations pour faire des opérations. Alors que les qubits supraconducteurs ont commencé en 1999, on arrive aujourd'hui à faire 1000 opérations sans erreurs. Tristan

Meunier et son équipe ont commencé en 2012, mais a rattrapé son retard en seulement 6 ans grâce à la purification du silicium.

Sur le nombre de qubits, le silicium est en retard (2 qubits contre 20 pour les supraconducteurs ou les ions piégés) mais en taille, ils sont les meilleurs (100 nm^2 contre 100 μm^2 pour les supraconducteurs. Ils ne sont pas les seuls. Il y a HRL, à Malibu, UC Madison ou UC Barkeley, Sandia, UNSW à Sydney. Le challenge est le contrôle à large échelle et la variabilité.

En 2007, on a démontré qu'on pouvait isolé un électron, puis cette année, on a fabriqué un CMOS avec un électron. À l'institut Néel, on arrive à une fidélité de 99%. Les échantillons actuels contiennent 16 grilles. Il va falloir changer d'échelle, mais où mettre les systèmes de mesures dans la grille de qubits ? Pour Tristan Meunier, la solution est dans une couche en dessous. Il termine par deux challenges d'ingénieurs qu'il a repérés : lier verticalement les qubits et les détecteurs ainsi que contrôler l'augmentation d'échelle.

En conclusion, l'information quantique est powerful mais fragile. Le silicium est un grand espoir.

8. La programmation quantique

Cyril Allouche (ATOS)

Cyril Allouche, polytechnicien, est directeur des programmes d'innovation au sein du groupe Atos, en particulier le programme en informatique quantique qu'il a initié en 2016.

Les espoirs de l'ordinateur quantique sont dans le calcul haute performance. Mais il faut savoir programmer avec ces nouveaux systèmes qui sont probabilistes et non déterministes. Le programme n'apporte pas de réponse mais donne l'évolution des résultats. Cela doit être interprété. L'algorithme de Shor donne au mieux 50% de bons résultats. C'est à prendre en compte. Autre différence, on ne peut pas copier. L'information ne peut pas être détruite ni copié sans la perdre. Il faut donc minimiser les lectures. Et le quantique se base sur la superposition comme le parallélisme du classique mais aussi l'intrication, sans comparaison dans le classique. Il n'existe pas d'algorithme classique transposable sur du quantique avec une accélération. Ainsi il ne peut y avoir de convertisseur automatique du classique vers le quantique. Selon Cyril Allouche, le calcul quantique va rester spécialisé, car la gestion de branchements est difficile. On va donc aller vers du calcul hybride : on ne pourra avoir des résultats que classiques car il n'y aura pas superposition. Le nombre d'algorithmes quantiques (à voir sur le site math/nist.gov/quantum/zoo/) est passé de 0 en 1993 à 250 en 2016. Ils ont le mérite d'exister même sans application, à part la réduction de systèmes linéaires, d'équations aux dérivées partielles, en optimisation combinatoire… Mais le reste reste à inventer.

Cyril Allouche insiste sur trois points : le bruit inhérent à la physique quantique, la réduction d'erreurs et l'obstacle qu'est l'intrication au changement d'échelle. Reste aussi le bruit des portes, celui due à la fabrication et les conditions d'utilisation. Un concept

très récent est celui de NISQ (Noisy Intermédiate Scale Qauntum) avant d'avoir des qubits en nombre. Les limites sont le nombre de qubits (50 à 1000) et la connectivité. On commence à voir des résultats justes alors que le hardware est entaché d'erreurs. Il faut en fait modéliser le bruit. Depuis 3 ans, Atos travaille sur l'Atos Quantum Learning Machine pour des physiciens qui souhaitent voir l'impact de leur design. C'est une machine à mémoire dense qui simule la capacité de 41 qubits intriqués. Elle peut générer du code pour un hardware donné et simule le bruit. L'algorithme QAOA est itératif et converge même avec du bruit. Il résout de façon heuristique les problèmes NP. Il est hybride et dépend de coefficients qui sont optimisés. Le quantique est utilisé pour trouver de façon itérative la bonne fonction.

Cyril Allouche termine est donnant le message que le calcul quantique ne va pas remplacer l'ordinateur classique, mais l'accélérer et la simulation numérique est pour cela un outil indispensable pour prendre en compte les contraintes logiciels.

9. Investir dans les technologies quantiques : au-delà de la science, des marchés et des opportunités pour les entrepreneurs

Christophe Jurczak (Quantonation)

Christophe Jurczak, lui aussi polytechnicien, est directeur général de Quantonation qu'il a cofondé cette année.

C'est un fond d'investissement dédié aux technologies de rupture issues de la physique. Ces dernières semaines, beaucoup de conférences et d'articles de presse sont produites sur ces sujets. Christophe Jurzak explique comment fonctionne un fonds d'investissement tel que Quantonation. Les investisseurs sont tièdes car ils ne comprennent pas la physique quantique. Cela n'est pas son cas car il a fait une thèse dans ce domaine à l'Institut d'optique d'Orsay. En 1995, les expériences ressemblaient à un ensemble de lasers et d'optiques. Il y a 3 ans, à Palo Alto, il a vu que cela bougeait en Europe avec Muquans. L'industrialisation n'était pas loin. Il fallait chercher des fonds pour soutenir ce domaine émergeant. Quantonation, basé à Paris, se concentre sur les technologies quantiques disruptives et les startups de physique dure. Il s'agit de soutenir la création de sociétés et bien se positionner, en sachant que certaines ne vont pas aller au bout de leur projet. À Bristol, ils ont investi dans une startup de cybersécurité. L'horizon d'application de toutes ces sociétés est de 7 à 10 ans. Les investisseurs renâclent un peu car ils espèrent un retour sur 3 à 4 ans. L'ordinateur quantique est de ces technologies intéressantes. Autre point, la communauté quantique se crée et le contact avec le fonds permet de structurer cette communauté. L'intérêt pour le Quantum Computing (QC) est l'usage en chimie, en biologie, en combinatoire, en physique (simulation de matériaux) ou en cryptographie. Les sociétés qui veulent adopter le QC sont surtout intéressées par le Machine Learning (73%) ou le cloud

computing (72%), suivi de la vente digitale, du codage, des nouveaux matériaux et du transport autonome (40%). Le QC suscite beaucoup d'attente. Dans deux ans, on verra que le hardware n'avance pas aussi vite que le software. En 2025, il devrait y avoir des machines commerciales avec des NISQ (100 à 500 qubits non corrigés) qui seront le moteur de la révolution quantique. Selon le cabinet de conseil BCG, 110 milliards de dollars (G$) ont été investis entre 2012 et 2018 sur l'IA et seulement 0,7 G$ sur le QC. Un gros sujet aujourd'hui est de savoir où investir dans la chaine de valeur. Sur la recherche de fixateur d'azote, par exemple, de nombreuses startups agissent à tous les stades (algorithme, langage, microondes, fabrication du dispositif). Un problème sérieux est celui de la mémoire quantique qui n'existe pas, comme l'internet quantique et le blind computing (faire tourner des programmes sans que quelqu'un ne sache ce dont il est question). Quantonation n'a pas de religion sur les sociétés dans lesquelles elle investit. Les business plans présentés n'ont souvent pas sens. Tout dépend du dialogue, de la physique qu'il y a derrière, de la qualité et la cohésion de l'équipe. L'association avec une équipe académique est à conserver longtemps et le potentiel d'upside pour un éventuel changement d'orientation. Et comme les startups de QC vont forcément travailler avec des intégrateurs et non vendre directement ses produits à des clients finaux, il faut s'intéresser aux intégrateurs. Quantonation pousse pour que les startups comprennent que le QC ne va pas tout faire. Il y a un gros développement de QC en Amérique du Nord (États-Unis et Canada). Même si l'Europe a la meilleure académie du monde, les entreprises sont à la traine. En France, on a des composants, du software, mais pas de quantum hardware. Or il y a un vrai enjeu de souveraineté nationale.

Pour finir, Christophe Jurzak annonce qu'il travaille sur le refroidissement d'atomes par laser dans lequel Quantonation investit pour faire de la simulation digitale, du NISQ et du QEC pour la simulation de matériaux. Pour faire émerger la communauté, Quantonation écrit (interview, Meetup, conférences, audition de l'OPESCT) et organise des Hackathon de programmation quantique comme celui qui se tiendra le 27 janvier à Paris, à l'école 42, avec l'aide de deux startups anglaises. C'est maintenant qu'il faut bouger.

10. L'optimisation quantique – le problème, le potentiel et l'application

Witold Kowalczyk (Bohr Technology)

Witold Kowalczyk est CEO de Bohr Technology Inc., entreprise fondée en Pologne (optimisation routière pour la poste polonaise), puis parti au Canada dans l'incubateur de l'université de Toronto.

Ils ont comme partenaires Wanadu, Rigetti, DWAve et Microsoft et sont financés par des groupes d'investisseurs pour l'optimisation et les réseaux neuronaux. Comme ils considèrent que pas assez de gens travaillent sur le quantique, ils ont créé une communauté internationale de projets avec laquelle ils travaillent sur le dataprocessing ou l'algorithmique. La question qu'ils se posent tous les jours est à quoi sert l'ordinateur quantique et ce qu'ils peuvent apporter à leurs clients. La réponse est le Machine Learning, la chimie, la combinatoire, la cybersécurité comme cela a déjà été dit aujourd'hui. Le problème que Google avait résolu en 2015 en disant que cela avait été 100 millions de fois plus vite qu'avec un ordinateur classique a au moins eu le mérite de montrer l'intérêt. Bohr Technology cherche des applications où le QC apporte des choses maintenant comme la simulation de molécules en chimie. Mais s'il n'y a pas d'avantage à avoir des ordinateurs quantiques, cela peut apporter quand même beaucoup sur une nouvelle approche d'algorithmes inspirés du quantique. Cela a déjà de la valeur. Le NISQ, l'ordinateur qui nous attend pour les prochaines années, est intéressant et a, selon BCG, beaucoup de potentiel. Les NISQ vont servir pour résoudre des problèmes où les données dispersées sont disponibles, sur les prédictions de précision sur des listes de données pour résoudre des problèmes de graphes comme en logistique avec 400 adresses ou plus. Un autre domaine, prédominant actuellement, est l'amélioration des

temps de calcul de l'IA et des réseaux neuronaux. D'autres applications sont encore potentiels (énergie, transport).

Les challenges techniques sont les problèmes d'échelle qui varie comme le carré des qubits, d'intégration et de cloud et les questions de connectivité.

11. Le calcul quantique et ses applications en chimie et en physique

Ivano Tavernelli (IBM Research)

Le Suisse Ivano Tavernelli travaille à Zurich au centre de recherche d'IBM où il est responsable de l'équipe de calcul quantique théorique.

Il s'est penché avec son équipe sur l'utilisation des supraconducteurs pour le calcul quantique en chimie. Le QC sert selon lui surtout à l'optimisation où se posent de nombreux problèmes. Nous en sommes actuellement au NISQ pour résoudre des questions de découverte de matériaux ou de chimie, ou de Machine Learning. Cette approche hybride permet de créer de 50 à 100 bons qubits. La logique quantique est la superposition (à la fois 1 et 0) et l'intrication (2^n états au lieu de 2). Mais la mesure fait effondrer l'état quantique qui est donc fragile. IBM fait des circuits supraconducteurs à 20 mK avec des durése de vie de 70µs pour des vitesses d'horloge à 50 MHz. Les résonateurs microondes servent à lire les états des qubits, à transporter les particules et à filtrer le bruit. Dans la IBM Q Experience, ils sont passés de 5 qubits en 2016, à 16 en 2017. Le problème est la multiplication des qubits si on veut minimiser les erreurs. La IBM Q Experience est l'ordinateur quantique le plus avancé au monde. En 2016, IBM a proposé une plateforme participative sur son cloud pour permettre aux étudiants et aux scientifiques l'accès aux algorithmes et aux expériences afin qu'ils fassent eux-mêmes des expériences. Il y a maintenant 100 000 utilisateurs d'IBM Q Experience sur les 7 continents qui ont fait 5,2 millions d'expériences. Avec QISkit on peut construire des circuits quantiques mais aussi bien plus. Ivano Tavernelli montre une simulation de caractéristiques physiques du LiH. Pour IBM, la chimie, l'IA et les services financiers seront les grands bénéficiaires du QC.

Il donne plusieurs exemples de l'avantage du QC par rapport à l'ordinateur classique en chimie, comme les questions fermioniques par des simulations Monte Carlo, le calcul d'hamiltoniens, avec ou sans correction d'erreur. Il y a quelques mois, son équipe a développé des méthodes pour des états excités, où la matière interagit avec la lumière, utile pour la photocatalyse, ou la photosynthèse artificielle et va s'attaquer au modèle de Hubbard qui date de 1963.

Il pense, comme les autres conférenciers que nous vivons actuellement dans le NISQ (50 à 100 qubits) mais bientôt les ordinateurs quantiques à 100 à 1000 qubits auront des applications utiles en liaison avec des ordinateurs classiques. Mais il faudra attendra encore pour avoir des QC de 1 à 10 millions de qubits.

David Menga conclut la journée en remerciant l'équipe d'EDF qui a aidé l'association Aristote à monter ce séminaire, les intervenants et les auditeurs et en prévenant que tous les transparents sont sur le site d'Aristote.