

Synbiotic

ANR BLANC SIMI3

2010

Version résumée 1.18

IBISC – Université d’Evry/Genopole

Franck Delaplace, Jean-Louis Giavitto, Hanna Kludel, Sylvain Sené
prenom.nom@ibisc.fr

LACL – Université de Paris-Est

Olivier Michel, Gaétan Hains, Franck Pommereau, Antoine Spicher, Serghei Verlan
prenom.nom@lacl.fr

CREA –Ecole Polytechnique - ISC-PIF

René Doursat, Paul Bourgine
prenom.nom@lpolytechnique.fr

CONTEXTE ET POSITIONNEMENT DU PROJET

La *biologie synthétique* est un domaine scientifique émergent [End05] qui concerne la conception et la fabrication banalisée et standardisée de composants et de systèmes biologiques sans correspondants naturels. Elle est actuellement en quête de principes de conception permettant une réalisation fiable et sécurisée à partir de composants biologiques réutilisables (BioBrick [CLE08, Kn03]).

Dans ce contexte, notre objectif est de concevoir et développer les outils permettant de « compiler » (au sens de la compilation des langages de programmation) le comportement global d'une population, par exemple, de bactéries, en des processus cellulaires locaux à chaque entité. Notre motivation à très long terme est de permettre l'exploitation des propriétés collectives d'une population (bactérienne) pour créer des biosystèmes artificiels répondant à divers besoins dans le domaine de la santé, des nanotechnologies, de l'énergie et de la chimie verte.

L'approche originale que nous proposons se fonde sur *une tour de langages* de programmation, dont le plus abstrait définit un modèle computationnel pour une population cellulaire et le plus primitif correspond à un agencement de séquences d'ADN. Chaque langage compile ses constructions propres vers le langage de la couche inférieure et ce, jusqu'au *bioware* (le « hardware biologique »). Cette approche, similaire à l'approche suivie avec succès dans le domaine de la synthèse d'architecture matérielle (chaîne de compilation vers le silicium), permet de combler le fossé existant entre la description d'un système au niveau d'abstraction pertinent pour l'application et la prise en compte de tous les détails de son implantation par des processus physico-chimiques. Elle permet de modulariser la conception d'un système, divisant les difficultés et isolant des niveaux d'abstraction qui peuvent évoluer indépendamment. Dans cette approche, un programme ne définit pas une fonction qui associe une sortie à une entrée mais spécifie un système dynamique (biologique) distribué qui essaie de maintenir des invariants en dépit de perturbations et des changements de l'environnement.

SYNBIOTIC est un projet de *recherche fondamentale à long terme* qui s'inscrit dans le domaine à la fois des langages de programmation non conventionnels et dans l'analyse/validation de propriété de systèmes dynamiques. Il s'agit d'un projet informatique qui vise à utiliser les nouveaux supports de calcul fournis par les avancées de la biologie synthétique et non à les produire. **SYNBIOTIC** a pour objectif d'étendre les techniques et les outils développés dans la *modélisation et la simulation de processus biologiques complexes* et d'intégrer de nouvelles approches de la programmation (*programmation spatiale, programmation amorphe et programmation autonome*) afin de faire face à de nouvelles classes d'applications caractérisées par l'émergence d'un comportement global dans une grande population d'entités localisées et irrégulièrement interconnectées de manière dynamique.

CONTEXTE ET ENJEUX ECONOMIQUES ET SOCIETAUX

La gamme d'applications de la biologie synthétique est extrêmement vaste et dépasse largement le cadre de cette proposition. Le lecteur intéressé peut se référer à [TES08] pour une présentation des initiatives de recherche européennes et [LEF08] pour tout ce qui concerne le développement économique de ce domaine. Très efficaces du point de vue énergétique, peu coûteux à mettre en œuvre et à diffuser une fois conçus (par simple

autoréplication), les biosystèmes offrent une interface extraordinairement souple vers le monde chimique et une ingénierie à l'échelle moléculaire et cellulaire. Un seul exemple dans le domaine de la santé : on imagine d'utiliser à moyen terme des bactéries¹ pour le diagnostic et à plus long terme, pour la thérapie. L'objectif visé est d'intégrer plusieurs fonctions biochimiques et de les contrôler afin de réaliser les différentes étapes d'un processus complexe de perception de l'environnement et de réaction à travers un réseau de nano-bio-machines distribuées. Les fonctions une fois agrégées et coordonnées, sont capables de tâches complexes : routage vers les cellules potentiellement malades, diagnostic en intégrant différents signaux biochimiques localisés à divers endroits, libération et dosage de composants actifs, vérification et surveillance en thérapie. On envisage ainsi de modifier des bactéries *E. Coli* pour la détection du cancer du colon et la destruction des tumeurs (stratégie qualifiée de « seek-and-destroy »). On peut multiplier les exemples dans le domaine de l'énergie (e.g. production d'hydrogène [BMH2]), de l'environnement (décontamination des sols pollués [dLor08]) ou en « chimie verte » (avec la synthèse de molécules complexes via l'assemblage de circuits de régulation codant pour les enzymes intervenant à chaque étape d'une cascade métabolique synthétique complexe dans un organisme-hôte [Dae06]).

Les promesses de la biologie synthétique sont soutenues par l'augmentation exponentielle des capacités de lecture et synthèse de l'ADN. Cet accroissement exponentiel des capacités technologiques est comparé à la loi de Moore qui a caractérisé la croissance de l'industrie informatique à base de silicium [Car03].

La biologie synthétique nécessite des outils informatiques pour la modélisation des processus biologiques, leur simulation, le stockage et l'analyse des données, etc. En cela, la biologie synthétique est très similaire à la biologie des systèmes (ou biologie intégrative). Mais le besoin de concepts et d'outils informatiques va bien plus loin. La biologie synthétique se caractérise en effet par l'importance accordée à l'aspect ingénierie réutilisable :

- comment spécifier et développer un ensemble de composants standards ayant des caractéristiques et des performances bien définies et qui peuvent être utilisés et réutilisés pour construire de nouveaux systèmes biologiques ?
- comment composer et assembler ces composants de la manière la plus « orthogonale » possible, c'est-à-dire de façon modulaire et en conservant l'indépendance mutuelle des composants ?
- comment développer et intégrer des méthodes de conception et des outils (analyse, simulation, validation) permettant d'automatiser la conception de tels systèmes ?
- comment réaliser l'ingénierie inverse de modules biologiques existants afin de les améliorer et de les adapter à nos besoins ?

Cette attention portée à la notion de librairie de composants biologiques réutilisables, la modularisation et l'abstraction de l'information, distinguent *in fine* la biologie synthétique du génie génétique (on peut prendre exemple sur les apports de la chimie combinatoire au génie chimique). C'est cette approche de construction de systèmes qui nécessite l'utilisation de méthodologies issues de l'informatique [Cous08]. dans la modélisation et la synthèse de séquences ADN adéquates à partir de descriptions fonctionnelles de haut niveau.

¹ La prolifération des bactéries dans un organisme humain pouvant être potentiellement dangereuse, une approche alternative est de concevoir des vésicules lipidiques contenant toute la machine biochimique permettant de réaliser les tâches de senseur/actuateur sans la machinerie de réplication.

Il n'est donc pas étonnant que ce soit le laboratoire informatique du MIT qui ait créé le concours iGEM [iGE], proposé un standard de composition de briques de base (BioBricks) et de procédures de fabrication [Wei03], et qui en maintienne le catalogue [REG]. Les BioBricks permettent de fabriquer rapidement des systèmes complexes en offrant des interfaces standard à tous les segments d'ADN susceptibles d'être combinés en séquences utilisant des enzymes commercialement disponibles [Kn03, CLE08].

Le développement des concepts et des outils informatiques nécessaires à l'automatisation de la synthèse de circuits de régulation biologique et à leur simulation, est un enjeu stratégique comparable à celui de la maîtrise des grands outils logiciels permettant la conception des SOC dans le domaine de l'électronique. Cependant, l'analogie avec l'électronique digitale ne peut pas être poussée trop loin ; un système biologique, comme par exemple une colonie bactérienne, est un système possédant un très grand nombre d'entités autonomes (plus grand que le nombre de processeurs reliés à internet ou de portes sur une puce VLSI), en interaction purement locale, arbitrairement défailants, pouvant se reproduire et évoluer. La conception de tels systèmes requiert de nouvelles approches non-conventionnelles, comme la *programmation amorphe* (amorphous computing) [Abe01], la programmation spatiale (spatial computing) [Dag06, SASO08], l'« ingénierie morphogénétique » [MEW09] ou les *systèmes autonomes* [Hor01].

Pour terminer ce bref panorama des enjeux et impacts potentiels de la biologie synthétique, il faut évoquer aussi les risques spécifiques à ces technologies. Ces biosystèmes, que l'on conçoit et construit en programmant des bactéries, sont des systèmes ouverts, qui interagissent fortement avec leur environnement et sont hautement dynamiques, dans leur comportement et leur structure. Le risque d'erreur de conception est grand ; on parle alors de *bio-erreur* [ETC07]. Par ailleurs, ils deviendront faciles à produire et leur détournement est un risque, la *bio-terreur* [CIA03, Whe04, Ven05], que l'on doit intégrer. La maîtrise de ces risques passe bien évidemment par le développement d'une expertise de ces technologies, indépendamment des enjeux économiques.

Dans ce cadre, le développement d'environnements de conception pour la biologie synthétique doit nécessairement intégrer des protocoles garantissant la sûreté des artefacts réalisés. Ceci constitue un prérequis essentiel, à la fois sur le plan scientifique et éthique, au développement de ces environnements. L'intégration des protocoles de sûreté dans ces environnements permet donc d'assurer un contrôle contribuant ainsi à diminuer une utilisation à risque. De ce fait un développement conjoint de méthodes de spécification et de validation est essentiel. Ici, l'informatique joue un rôle fondamental pour la maîtrise du contrôle de grands systèmes complexes critiques (aéronautique, spatial, Internet). Cependant, la particularité du médium, un organisme vivant, implique de revisiter les méthodes en ces domaines.

POSITIONNEMENT DU PROJET

Une approche informatique en amont de l'ingénierie génétique. Si la plupart des études actuelles cherchent à formaliser, concevoir, caractériser et valider des composants biologiques réutilisables, nous nous positionnons en amont de cette étape. La biologie synthétique regroupe des stratégies scientifiques et des technologies très différentes qui incluent la conception et la construction de génomes, la conception de protéines, la synthèse

de composés biochimiques par de nouvelles voies métaboliques et la construction de circuits de régulation génique dans des cellules et des micro-organismes. Pour ce qui nous concerne, nous faisons l'hypothèse qu'il existe des bibliothèques standardisées de comportements biochimiques élémentaires qui peuvent être composés dans une bactérie, comme par exemple les BioBricks [Kn03, Kni03, REG, CLE08].

Notre objectif est d'adresser la conception de grands systèmes biologiques par *une approche langage*, de la même manière que VHDL permet la conception de système de traitement de l'information à partir de portes et de blocs logiques élémentaires. Ce projet informatique repose sur trois hypothèses : l'apport des formalismes discrets, un processus de conception fondé sur la compilation d'une tour de langages et la prise en compte des aspects spatiaux.

Des formalismes discrets. Notre première hypothèse est que des modèles informatiques discrets sont adéquats pour décrire des biosystèmes et parfois plus pertinents que des approches mathématiques traditionnelles comme les équations différentielles [Car09]. Cette hypothèse est corroborée par l'important développement actuel des formalismes informatiques dans le domaine la biologie des systèmes. En particulier, ces formalismes sont plus à même de capturer de manière concise les aspects qualitatifs et quantitatifs des grands réseaux d'interactions biochimiques impliqués dans les processus biologiques.

Ces formalismes permettent de découpler les abstractions utilisés dans le processus de conception (signal, gradient, mémoire, propagation, information de position...) des processus biochimiques utilisés pour leur implémentation, de la même manière qu'un bit abstrait de manière robuste une implantation électrique dans une électronique à base de silicium.

Par ailleurs, ces formalismes permettent d'aborder la question de la validation : que peut-on garantir sur les comportements du système biologique artificiel, quel est le domaine de viabilité du système, quels sont les perturbations de l'environnement qui sont tolérables, quels est la résilience du système, peut-on garantir que certains états son inatteignables, tracer les processus, tester les comportements attendus, etc.

Une approche compilation. Notre seconde hypothèse est qu'à partir de ces formalismes, la compilation est une approche descendante plus souple que l'assemblage direct de composants biologiques prédéfinis. Le processus de compilation permet d'instancier des composants élémentaires génériques dans un organisme particulier, permet de prendre en compte des contraintes d'assemblage (comme l'évitement de cross-talk entre circuits de régulation) ainsi que la simplification et l'optimisation des circuits obtenus par assemblage. Cette approche de haut-niveau correspond à la synthèse d'un système à partir de ses spécifications et repose sur la possibilité de dériver le comportement des parties à partir du comportement d'un tout. Ce problème est notoirement plus simple que celui de l'inférence de propriétés globales à partir de comportements locaux (émergence) et a montré toute son utilité dans le domaine de la synthèse d'architecture, mais aussi dans le cadre de la programmation spatiale et de la programmation amorphe [NaCo98, NaMa03]. Un des enjeux du projet est de montrer que ces techniques peuvent être appliquées avec succès à la synthèse de biosystèmes.

La prise en compte du spatial. Enfin, notre dernière hypothèse est que, même si pour l'instant la biologie synthétique se focalise sur la « programmation d'une seule bactérie », le développement de biosystèmes un tant soit peu complexe reposera sur le fonctionnement

intégré de colonies bactériennes et donc sur la prise en compte d'interactions spatiales au sein d'une population de cellules différenciées. Il est en effet douteux qu'une cellule puisse supporter un nombre arbitraire de comportements artificiellement imposés. Au contraire, l'exemple des processus biologiques naturels montrent toute l'importance de l'organisation spatiale et de la compartimentalisation (membrane, vésicule, cargo, compartiment, cellule, biofilm, tissus, organe, etc.) permettant la spécialisation et le fonctionnement intégré au sein d'un système compris comme une écologie. Par ailleurs, la maîtrise des interactions spatiales ouvre la voie à une ingénierie du développement (« développement » au sens biologique du terme), ce qui permet de rêver à des applications qui vont bien au-delà de la conception de la cellule comme « usine chimique ».

Un volet validation. Le positionnement du projet est résolument informatique et se concentre sur le développement de techniques de compilation et de validation en amont de l'ingénierie génétique. Cependant, afin de valider les outils développés, nous souhaitons les mettre en œuvre concrètement à travers leur utilisation à l'intérieur du projet par une application de morphogenèse et, à l'extérieur, par une équipe iGEM.

Un volet pédagogique. Outre cet aspect validation, l'implication d'étudiants dans ce projet nous semble remplir un volet pédagogique indispensable : nous y voyons l'opportunité d'attirer et d'initier des étudiants à un domaine stratégique. Ce volet pédagogique se complète par l'organisation d'une école permettant de réunir et de faire interagir des communautés scientifiques a priori éloignées (compilation, bioinformatique, synthèse d'architecture, biologie des systèmes, langages non-conventionnels). Cette action pédagogique est indispensable au vu de la complexité du domaine et de la fragmentation des expertises. Enfin des actions de présentation vers le grand public (par exemple à travers un site web comme [Interstice](#) [Int] ou [Vivagora](#) [Viv], et des débats publics comme le « Café du gène » organisé par Genopole Évry) sont nécessaires afin que la société civile puisse s'approprier de manière informée des enjeux et des impacts potentiels de ces nouvelles technologies.

Le positionnement français et européen. Si la France est présente dans le domaine de la biologie des systèmes, elle a tardé à investir le domaine de la biologie synthétique. Par exemple le Kenya, la Turquie, l'Espagne, la Suisse, l'Inde, la Chine... ont proposé des équipes à iGEM avant la France. La mobilisation de l'Europe dans ce domaine est plus importante [TES08] mais se développe principalement à partir de la biologie. C'est principalement aux Etats-Unis, à partir du projet Amorphous Computing au MIT, que se développe l'idée que la « programmation des bactéries » est plus qu'une simple métaphore et doit reposer sur des outils informatiques de conception qui vont au-delà de techniques standardisées de génie génétique.

L'Europe, et dans une moindre mesure la France, ont cependant des atouts pour se faire une place dans ce domaine de recherche. Nous sommes particulièrement présents dans le domaine des formalismes pour la modélisation et la simulation de réseaux de régulation et dans le domaine des modèles de programmation non-conventionnels (par exemple en Angleterre, un « Grand Challenge » est dédié aux modèles de calcul non conventionnels [GC7] et ce thème est soutenu par le programme FET au niveau européen). Ces deux domaines constituent des fondations à partir desquels il est possible de développer des approches de synthèse de haut-niveau. Une des difficultés spécifiquement française est de

reconnaitre le domaine des modèles de calculs non-conventionnel, la communauté étant surtout développée autour du calcul quantique.

Les partenaires de ce projet de recherche sont particulièrement bien placés pour créer un réseau international fort à partir de leurs activités d'animation dans le domaine (voir en particulier la section 5.2 de ce document) et par leurs implications dans des montages de projets de collaborations internationales britanniques (proposition EPSRC de M. Gheorghe au programme « Grand Challenge GC7 » [GC7]) et états-uniens (proposition PIRE de C. Teuscher auprès de la NSF [Pir]). Le projet déposé ici a aussi pour objectif de consolider et pérenniser ces liens.

DESCRIPTION SCIENTIFIQUE ET TECHNIQUE

ÉTAT DE L'ART

Le projet **SYNBIOTIC** s'appuie sur les développements récents de la biologie synthétique, de la modélisation et de la simulation de processus biologiques complexes, ainsi que sur de nouvelles approches non conventionnelles de programmation : la programmation *spatiale*, la programmation *amorphe* et les systèmes *autonomes*. Dans cette section, nous faisons un état des lieux sur ces domaines, uniquement du point de vue « langage » qui caractérise notre approche.

La biologie synthétique. La biologie synthétique combine biologie et ingénierie génétique pour concevoir et construire de manière systématique de nouveaux systèmes biologiques (ou biochimiques) qui exhibent des fonctionnalités nouvelles [BS05, Voi06]. Ce domaine de recherche s'est surtout développé autour de la conception de circuit de régulation génétique [Chi06, Gue02, HP06, Kob04, McDW05] après la construction du premier réseau de régulation artificiel en 2000 [EL00, Gar00]. Plusieurs chercheurs ont défendu la nécessité d'une approche d'ingénierie pour la biologie synthétique et cette voie est largement impulsée par des informaticiens [End05], comme le prouve le développement du « Registry of Standard Biological Parts » [REG] par des chercheurs issus du Computer Science and Artificial Intelligence Laboratory (CSAIL) du MIT. Les entités décrites dans cette base de données correspondent à différents niveaux d'abstraction [End05] : au plus bas on trouve des « parts », séquence d'ADN (promoteur, région codante, etc.), qui peuvent s'assembler en « devices » qui eux-mêmes se composent en systèmes. La biologie synthétique se développe rapidement aux US (Princeton, Harvard, MIT...) et apparaît en Europe (Edinburgh, ETH Zurich, Université de Rome, quelques projets européens comme PACE, principalement orientés les techniques de génie génétique [TES08]). Un certain nombre d'applications ont été développées lors de la compétition internationale iGEM [iGE] organisées aux États-Unis entre de grandes universités : film bactérien photosensible [KKA+04] ou bio-senseur permettant de détecter des explosifs [Gib04].

Le développement d'une ingénierie de composants biologiques réutilisables repose de manière cruciale sur la possibilité de modéliser ces composants et de simuler leur assemblage. La biologie synthétique profite donc du développement très important de la biologie des systèmes [AIRRH03]. Depuis le début des années 2000, de nombreux formalismes utilisés par ailleurs en informatique, ont été appliqués à la *modélisation* et à l'analyse de réseaux biochimiques : le π -calcul [Reg01], Pepa [Cal06], les réseaux de Petri [Red93], les systèmes à membranes [PR02, Paun02], la modélisation logique [TK01], les

systèmes de réécritures [GMM04, EKL02] ; tandis que des formalismes dédiés ont aussi été développés comme le κ -calcul [Dan07], BioNetGen [Fae05], Biocham [Cha04] ou BioAmbient [Reg04]. Et des techniques de *model-checking* [Ric06] ainsi que d'extraction d'invariant ou de test de systèmes [FMP00] ont été utilisées avec succès dans l'analyse des systèmes spécifiés.

Parmi les formalismes utilisés, peu ont été conçus pour la *conception* de réseaux de régulation. On peut citer² BioJade, GeNetDes, GenoCAD, TinkerCell, qui ont tous pour objectif de composer des réseaux d'interaction à partir de bibliothèques de composants prédéfinis (BioBrick ou autre). Outre la simulation du réseau biochimique, ils permettent généralement de produire la suite des oligonucléotides et/ou la suite des opérations d'assemblage des séquences à synthétiser pour construire le circuit. Ces outils ne permettent donc pas de décrire un réseau de manière abstraite puis de l'optimiser et enfin de l'instancier à partir de ressources disponibles dans une ontologie. On est donc plus proche des environnements de simulation que d'un processus de compilation. Seul le formalisme LBS [PP08] est de niveau réseau et explicitement prévu pour être compilé vers une cible : réseaux de Petri ou équations différentielles pour la simulation, ou bien, très récemment, vers un ensemble de BioBricks [CLE08, Kn03]. Enfin, le langage PROTO [Beal08] est à notre connaissance la première tentative de compiler des abstractions spatiales de haut-niveau vers un substrat biologique sous la forme de réseau de régulation.

Programmation spatiale et programmation amorphe. Considérer une colonie de bactéries, chacune dotée de certains comportements, comme un système programmé [Wei01], pose des problèmes nouveaux et nécessite de développer de nouvelles approches de la programmation. Les langages de programmation non-conventionnels (voir les conférences [BFGM05, GC05, Dag06, MBFG06, UC07] pour n'en citer que quelques unes) offrent de nouvelles pistes permettant de faire face à aux défis soulevés par la conception/programmation de systèmes biologiques artificiels. Nous en évoquerons ici deux : la décentralisation massive et l'autonomie. En effet, les bactéries interagissent de manière très locale et de manière autonome. Il n'est pas possible d'intervenir à un niveau fin (au niveau de chaque bactérie) lors de « l'exécution » du programme.

L'objectif du calcul autonome [Hor01] est de concevoir des systèmes matériels et logiciels *auto-gérés* (self-managed) en s'appuyant sur des propriétés d'*auto-organisation* : la configuration autonome des composants du système en une architecture dynamique répondant aux objectifs du système, d'*auto-réparation* : la détection et la correction autonomes des erreurs matérielles ou logicielles, d'*auto-optimisation* : l'auto-surveillance, le contrôle des ressources et la reconfiguration du système par lui-même pour assurer un fonctionnement optimal. Ces propriétés (et quelques autres) sont regroupées sous le nom générique de *self-**. Les concepts développés dans le domaine des systèmes autonomes sont à même de répondre aux problèmes posés par l'autonomie des bactéries. La population de bactérie peut être vue comme un système dynamique distribué qui se stabilise dans un état nominal. Des perturbations écartent ce système dynamique de son point de fonctionnement admissible mais les propriétés *self-** permettent d'y revenir.

Le calcul amorphe [Abe01] s'est explicitement préoccupé de concevoir des langages permettant de programmer des propriétés *self-**. Un système amorphe est un système composé d'un grand nombre (d'un ordre de grandeur supérieur à 10^7) d'unités de calcul

² Cf. http://openwetware.org/wiki/Computational_Tools

autonomes, possiblement défaillantes et interconnectées de façon irrégulière et dynamique. Chaque unité de calcul est capable d'acquérir des informations sur son voisinage et d'interagir uniquement avec des entités spatialement voisines. De tels systèmes naturels ou artificiels constitués d'une myriade d'unités élémentaires et dont la structure est dynamique, sont à l'heure actuelle, disponibles et peuvent servir de support de calcul. Outre les populations bactériennes qui nous intéressent ici, on y trouve aussi les réseaux pair-à-pair, les réseaux de senseurs (ceux du projet *smart dust* par exemple [KKP99]), les *smart matters* ou MEMS [HH98], les nanostructures auto-assemblées [RPW04], etc. Un support de calcul amorphe [Abe01] est donc un système hautement redondant, spatialement organisé et asynchrone ; par l'absence d'hypothèse sur la précision des connexions ou sur l'organisation des parties, ces systèmes constituent l'environnement idéal pour explorer les problématiques de la programmation spatiale : une conception robuste dans un contexte présentant une topologie dynamique, une organisation locale variable et arbitraire, et un couplage local avec le monde physique environnant. Les études sur le calcul amorphe ont particulièrement développé des techniques de compilation permettant de passer d'une spécification globale au comportement local à chaque entité élémentaire [Coo99, Nag01, Coo04, Sto04, Yam07].

Les systèmes autonomes et le calcul amorphe focalisent généralement leur attention sur des entités localisées mais ne permettent pas de manipuler directement des abstractions spatiales. C'est l'objectif du calcul spatial, un domaine de recherche émergent [Dag06] dont l'objectif est le développement de modèles de calcul, de langages ou encore d'architectures tenant compte de l'espace afin de dépasser le modèle de von Neumann. Ce besoin de manipulation d'entités spatiales apparaît par exemple dans la simulation pour la biologie [MSG08], les nouvelles architectures [GLRT04], le domaine de la Claytronics et des MEM [Clay04, Upim03] et récemment dans le domaine du calcul amorphe [Beal08].

Ces approches alternatives de la programmation renouvellent la notion de système dynamique (par exemple avec la notion de systèmes dynamiques à structure dynamique : (DS)² [GM01b, GGMP02a, GM02c, GM03]) et fournissent des techniques qui permettent la spécification de processus *dynamiquement structurés* et *spatialement distribués*.

OBJECTIFS ET CARACTERE AMBITIEUX/NOVATEUR DU PROJET

L'objectif du projet **SYNBIOTIC** est de concevoir et développer une approche langage de haut-niveau en biologie synthétique pour la conception de systèmes biologiques artificiels. Notre motivation finale est de permettre la synthèse de bioware de manière analogue à la synthèse actuelle de VLSI.

Plus précisément, nous voulons prouver la faisabilité d'une chaîne de compilation permettant de compiler la spécification d'un comportement spatial global en des réseaux artificiels de régulation dans une bactérie (régulation génétique, métabolique, signalisation). Ce processus de compilation s'effectue à travers une tour de langages intermédiaires permettant de combler le fossé existant entre la description de haut-niveau pertinente pour l'application et la prise en compte de tous les détails des processus physico-chimiques. Cette chaîne de compilation se complète par des outils de simulation et de validation.

Cet objectif ambitieux va bien au-delà de ce qui est actuellement développé dans le domaine de la biologie synthétique où les outils de conception actuels se limitent à la composition de réseaux de régulation. Il se complète par un volet d'animation scientifique dont l'objectif est

de familiariser une communauté d'informaticiens aux spécificités de ce domaine et de consolider et ancrer un réseau de collaborations scientifiques internationales.

Quatre verrous scientifiques et techniques principaux sont soulevés par ce projet. Le premier concerne la conception et la caractérisation de langages intermédiaires permettant de faciliter le passage de spécifications globales au fonctionnement asynchrone et défaillant d'entités en interaction purement locale. Notre approche est de traiter au niveau d'un premier langage le passage d'une spécification globale au comportement local des entités composant le système et de traiter au niveau d'un second langage, le caractère amorphe (asynchronisme et défaillance) de la cible visée.

Le second verrou porte sur le développement d'algorithmes de simplification et d'optimisation des réseaux de régulation générés. Par comparaison, les phases de génération de code dans un compilateur classique doivent faire face à des ressources de calcul relativement orthogonales. Ce n'est bien évidemment pas le cas quand il faut générer des réseaux de régulation à partir de base de données contenant plusieurs milliers de circuits élémentaires en prenant en compte toutes les contraintes biochimiques.

Le troisième verrou est de nature technique : afin de tester le résultat de chaque étape de compilation, sans attendre d'avoir une chaîne complète, nous développerons un interprète pour chaque langage intermédiaire. Ces interprètes devront s'appuyer sur des techniques de simulation *ad hoc* (approximation, représentation hybride, représentation multi-niveaux, représentation symbolique) permettant la représentation de grandes populations de molécules (gradient chimique) ou de bactéries (colonies bactériennes).

Enfin, concernant le dernier verrou, la validation des programmes se fera par des approches de vérification de type *model-checking* permettant de prouver une propriété pour toutes les exécutions possibles d'un programme. Les techniques de *model-checking* (tant sur des systèmes booléens que sur des systèmes stochastiques ou hybrides) ont été appliquées avec succès à l'analyse de systèmes biologiques. L'enjeu ici est d'étendre ces techniques afin de prendre en compte la localisation spatiale des processus.

Les résultats de ce projet de recherche se concrétiseront à travers : (1) la définition et la caractérisation théorique de langages de programmation non-conventionnels ; (2) le développement de prototypes d'interprètes et de compilateurs pour chacun des langages intermédiaires (visant à faire la preuve de faisabilité sur un sous-ensemble pertinent du langage) ; (3) des outils de traduction de formalismes dédiés vers des formalismes existants permettant la vérification ; (4) le développement d'un exemple paradigmatique dans le domaine de l'ingénierie morphogénétique permettant de valider sur un exemple la chaîne de compilation complète.

Concernant le volet animation de ce projet, un des résultats escomptés est l'émergence d'une équipe iGEM « théorique » permettant de valider sur un cas concret, et par des personnes extérieures au projet, l'approche et les outils développés. Par delà cette action ponctuelle, nous espérons attirer et mobiliser de nouveaux jeunes chercheurs français sur ce domaine. Enfin, la consolidation de liens de recherche au niveau international devrait permettre d'envisager la constitution d'un réseau européen, *en informatique*, dans ce domaine émergent.

Bien que très ambitieux, la prise de risque dans ce projet reste bien maîtrisée. En effet, nous nous appuyons sur l'expérience des participants et sur des résultats déjà acquis dans des domaines connexes. De plus, chaque langage intermédiaire est conçu pour « cloisonner » les

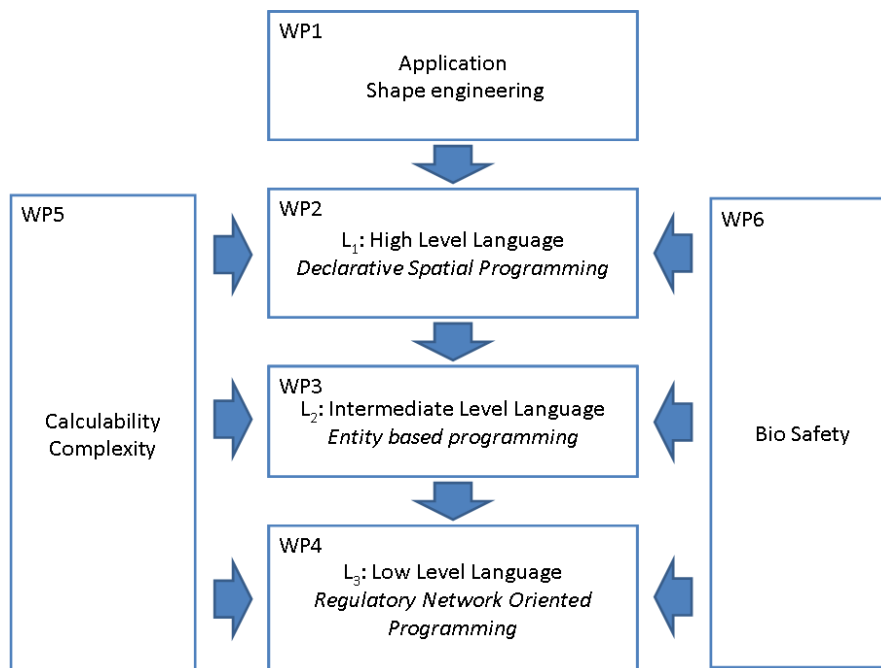
difficultés. La possibilité de disposer d'un interprète pour chaque niveau permet de valider chaque phase indépendamment de la réalisation de la chaîne complète. Enfin, la faisabilité de chaque phase de compilation dépend largement des primitives disponibles à chaque niveau de description, ce qui permet de moduler les résultats attendus.

Clairement, **SYNBIOTIC** est un projet fondamentalement ancré en informatique qui relève du domaine du calcul non-conventionnel et de l'analyse des propriétés des systèmes dynamique. Un programme ne définit pas une fonction qui associe une sortie à une entrée mais spécifie un système dynamique (biologique) distribué essayant de maintenir des invariants en dépit de perturbations et des changements de l'environnement. Par son positionnement amont, le projet dépend peu des avancés à court terme des techniques de biologie synthétique.

PROGRAMME SCIENTIFIQUE ET TECHNIQUE, ORGANISATION DU PROJET

PROGRAMME SCIENTIFIQUE ET STRUCTURATION DU PROJET

Le programme de recherche proposé a pour objectif de développer une *tour de langages*, chaque langage compilant ses constructions propres vers le langage de plus bas-niveau. Nous envisageons 3 niveaux inspirés de concepts élaborés dans les domaines de la programmation spatiale, de la programmation amorphe et des systèmes autonomes.



Au niveau L₁ le plus abstrait (tâche **WP2**), le programmeur peut concevoir son système à partir d'entités abstraites globales (ondes, diffusion, synchronisation...) exprimant des corrélations spatiales et temporelles entre un grand nombre d'entités autonomes localisées. (Niveau L₂, tâche **WP3**). Le niveau L₃ (tâche **WP4**) adresse les spécificités de notre « machine cible » : une bactérie qui sera programmée suivant les techniques développées dans le domaine de la biologie synthétique et en s'appuyant sur les BioBricks. La puissance d'expression de chacun de ces niveaux doit être caractérisée formellement, tout comme leurs propriétés (tâche **WP5**).

L'artefact biologique spécifié par cette chaîne de compilation doit être validé avant d'être synthétisé. La validation doit se faire dans un premier temps par simulation, ensuite par des méthodes de vérification de systèmes (tâche **WP6**). Ces dernières peuvent s'appuyer sur les techniques de modélisation informatique des processus biologiques développées dans le domaine de la biologie intégrative (algèbre de processus, réseau d'automates, réseau de Petri, théorie des jeux). Ces approches permettent d'utiliser des méthodes d'extraction d'invariant ou de *model-checking* pour prouver *a priori* certaines propriétés du biosystème.

Cette « approche langage » de la biologie synthétique doit se concrétiser par un environnement permettant de faire la preuve de notre approche. Cet environnement sera validé (tâche **WP1**) par le développement d'applications tests dans le domaine de la morphogenèse (*shape engineering*). Outre son intérêt fondamental dans le domaine de la

biologie, ces applications nécessitent réellement la coordination d'une population cellulaire et permettent d'envisager la « construction biologique » d'objets à l'échelle moléculaire et cellulaire, de manière énergétiquement très efficace.

Ce programme de travail se complète par une tâche de coordination **WP0** visant au pilotage du projet et à la dissémination des résultats. En particulier, le développement et le soutien à une équipe iGEM relève de cette tâche.

COORDINATION DU PROJET

Les partenaires étant tous présent en région parisienne et ayant déjà tous travaillé ensemble, un couplage fort est attendu lors de ce projet. Une réunion bimestrielle est prévue réunissant tous les partenaires.

Coordination. Les décisions concernant le projet seront prises en concertation et le consensus est privilégié. Un comité de pilotage constitué du porteur du projet et d'un représentant de chaque site est chargé de veiller à l'avancement et à la bonne fin du projet.

Bonnes pratiques. Les doctorants et stagiaires seront tous co-encadrés par deux sites. Un comité de thèse, comportant des membres extérieurs au projet, sera mis en place pour chaque doctorant encadré dans le projet. Les rapports de stage des étudiants de M2 travaillant sur le projet donneront lieu à des rapports techniques. Les dépenses de chaque site seront communiquées à l'ensemble des partenaires.

Site collaboratif. Un wiki collaboratif sera mis en place avec une partie publique et une partie privée. Le wiki sera en particulier utilisé pour la mise à disposition des logiciels développés. Il proposera (dans sa partie privée) une bibliothèque d'articles avec une bibliographie raisonnée. En effet, une bibliothèque électronique commentée est indispensable pour adoucir la « courbe d'apprentissage » des chercheurs désireux de s'impliquer dans le domaine (nous pensons en particulier aux étudiants). Le wiki se doublera d'une forge des développements logiciels permettant d'enregistrer les différents codes et leurs versions.

Rapport d'avancement. Chaque coordonnateur de tâche enverra un résumé des travaux effectués dans le semestre à $a.N+6$ (a année du projet, N date de début du projet) au coordonnateur du projet. Aux dates anniversaires du contrat, ils participeront avec le coordonnateur du projet à un rapport reprenant l'état d'avancement de chaque tâche, les différents travaux effectués, les résultats obtenus, les manifestations organisées, les publications effectuées, les réorganisations prévues, etc. qui sera envoyé à l'ANR.

Rapports finaux. Chaque tâche, y compris WP0 pour la partie dissémination, produira un rapport final récapitulant tous les résultats et publications obtenus. Ces rapports de fin de tâche seront joints au rapport final du projet.

DESCRIPTION DES TRAVAUX PAR TACHE

TACHE WP0 : COORDINATION ET DIFFUSION DES RESULTATS

La politique de coordination du projet s'établit selon 3 axes : la coordination des travaux de recherche, la mise en place d'une infrastructure de diffusion des résultats et la participation au développement d'une politique de fédération d'une communauté interdisciplinaire et émergente autour de la biologie synthétique. L'action de coordination s'architecture en deux niveaux : le premier au sein de chaque tâche et le second au niveau de la progression du projet.

[...]

TACHE WP1 : L'INGENIERIE DES FORMES

Participants : Université d'Evry, Université Paris-Est Créteil, CREA-ISC

Objectifs

Le rôle de WP1 sera de concevoir, développer et implémenter divers exemples d'applications de type « calcul spatial » réalisant une ingénierie de motifs et/ou de formes, afin de servir d'environnement de test, validation et révision à l'architecture de langages élaborée dans les autres tâches. Ces applications seront toutes déployées dans des simulations virtuelles et certaines d'entre elles, selon leur niveau de difficulté, aboutiront également à une matérialisation dans un substrat biologique synthétique au niveau L3 (le « bioware » de WP4). Par cet exercice, nous serons conduits à abstraire les principes de base de la formation de motifs, du mouvement collectif et de la morphogenèse biologique, tout d'abord en direction de modèles mathématiques et computationnels multi-agents, puis dans la traduction de ces modèles en primitives spatiales offertes par le langage L1 (brisures de symétrie, champs de forces, gradients morphogénétiques, processus de diffusion, propagation d'ondes, etc. ; cf. WP2). En retour, cela pourra donner lieu à des modifications de l'architecture L1-L3.

[...]

TACHE WP2 : PROGRAMMATION SPATIALE DECLARATIVE

Participants : Université d'Evry, Université Paris-Est Créteil

Contexte de la tâche

Le caractère massivement parallèle des processus réalisés par une population de bactéries repose sur la *spatialisation* de la population qui permet d'effectuer différentes tâches à un même moment mais à différents endroits. Dans ce contexte, il est essentiel de considérer l'espace comme un concept de premier ordre qui doit être décrit et manipulé de façon explicite *au niveau du langage de programmation* permettant de concevoir ces processus.

Objectif de la tâche

Nous proposons d'aborder cette problématique à travers le développement d'un langage *déclaratif* L1 dédié à la programmation spatiale. Ce langage permettra de décrire à un niveau global le comportement de l'ensemble de la population. Ces programmes seront ensuite

compilés à travers les étapes décrites en **WP2** et **WP3** pour finalement être instanciés dans un support bio-synthétique.

Les comportements globaux seront spécifiés à l'aide d'un ensemble de primitives globales (propagation d'ondes, mise en place de gradients, ...) qui permettront la définition de champs de données définis sur la population et qui pourront être composées. Ces primitives seront décrites dans un formalisme fondé sur l'utilisation d'opérateurs différentiels discrets. Ceux-ci permettront de décrire de façon homogène les mouvements de données en tout point de la population. Pour terminer, ce **WP** s'instanciera concrètement dans le développement d'un interpréteur de L1 et dans son utilisation sur divers exemples d'applications.

[...]

TACHE WP3 : PROGRAMMATION ORIENTE ENTITE

Participants : Université d'Evry, Université Paris-Est Créteil

Contexte de la tâche

Les développements de la tâche **WP2** ont pour objectif d'aboutir à une localisation des calculs par leur mise sous-forme d'équations différentielles. Le formalisme proposé reste cependant de trop haut niveau pour être implanté directement sur un support de calcul biologique. Il souffre notamment de deux inconvénients majeurs :

1. Bien que fondamentalement local, le système d'équations différentielles discrètes exprime une loi homogène, globalement vérifiée par l'ensemble de la population et qui régit les communications et les interactions entre les entités qui la composent. A l'inverse, il semble plus naturel d'envisager la programmation des bactéries d'un point de vue centré individu où l'hétérogénéité de la population peut être mise en avant. Le point de vue externe aux bactéries offert par les équations différentielles discrètes est donc éloigné d'une implantation.
2. Afin que les interactions spécifiées par les opérateurs différentiels aboutissent au calcul désiré, les communications doivent être considérées dans un contexte synchrone. Encore une fois, l'homogénéité (cette fois-ci temporelle) requise par le formalisme issue de la tâche **WP2** semble trop contraignante pour une implantation directe dans les bactéries, celles-ci fonctionnant chacune à leur propre rythme.

Il est donc nécessaire avant même d'envisager une programmation des bactéries, d'effectuer le changement de point de vue *centré interaction vers centré individu*, et de *désynchroniser* les communications dans la population.

Objectif de la tâche

L'objectif de la tâche **WP3** est de répondre à ces besoins en définissant un langage L2 de niveau d'expression intermédiaire. Cette tâche se divise en deux parties d'égale importance : l'élaboration du langage L2 et son utilisation pour la résolution des systèmes d'équations différentielles discrètes issues du **WP2**.

Le langage L2. Plus proche du medium biologique final que L1, le langage L2 introduira pour la première fois dans la chaîne de compilation les outils permettant la manipulation

explicite d'une population de bactéries virtuelles. Il sera très proche d'un langage de spécification de systèmes multi-agents réactifs et sera accompagné des outils de simulation nécessaires pour le développement d'exemples d'application. L'accent sera particulièrement mis sur le développement de ces outils ; ils devront en effet rendre compte du caractère amorphe d'une population de bactéries (grand nombre d'entités, organisation spatiale arbitraire et dynamique, comportement erroné des agents, ...). Nous nous appuyons pour cela sur les travaux antérieurs sur le calcul amorphe [Abe01] ainsi que sur des collaborations internationales entreprises sur ce sujet (collaboration avec D. Coore).

La traduction de L1 vers L2. Le second objectif est de clairement répondre aux questions sur le changement de points de vue et la désynchronisation des interactions. Ces questions sont délicates et ne peuvent pas être appréhendées dans un contexte général. Cependant, les motivations de ce projet sont assez précises pour restreindre le champ d'investigation à des cas particuliers (propagation d'onde, gradient) pour ensuite élargir le périmètre. Pour pallier tout risque d'échec, nous envisageons également d'autres voies d'études.

[...]

TACHE WP4 : PROGRAMMATION ORIENTE RESEAUX DE REGULATIONS

Participants : Université d'Evry, Université Paris-Est Créteil

Contexte de la tâche

A l'interface entre la description et l'implémentation, la couche basse du langage adresse deux problématiques : La première concerne la définition d'opérateurs permettant d'implanter la spécification de haut niveau ; La seconde concerne l'étude de l'implantation sur un médium biologique.

Objectif de la tâche

L'enjeu de cette partie est d'établir un pont entre spécification et implantation en bioware qui permette : d'une part de conserver l'expressivité et l'abstraction des langages des couches supérieures et d'autre part de se doter des outils d'analyse pour permettre une implantation fiable et sécurisée des fonctions biologiques synthétiques. Ce double enjeu s'incarne par deux tâches identifiées d'un côté par la définition d'un langage de bas niveau fondée une interprétation des processus biologiques comme des opérateurs de calcul et d'un autre côté par le développement d'outils d'analyse afin de modéliser le comportement d'un système biologique, sa dynamique, afin d'en accroître sa maîtrise. Cette approche sera développée en veillant à permettre l'installation de protocoles de sûreté et de sécurité.

[...]

TACHE WP5 : CALCULABILITE ET COMPLEXITE

Participants : Université d'Evry, Université Paris-Est Créteil

Contexte de la tâche

Les modèles de calcul permettent de capturer la notion de « calcul » en précisant des règles strictes qui le définissent. Cela permet de comprendre les limites théoriques du calcul et donne aussi une première idée des choix d'implantation possibles. Les fondements des ordinateurs actuels ont été instaurés par les modèles de calcul classiques, en particulier par la machine de Turing qui inspira l'architecture de von Neumann. Une machine de Turing représente le modèle d'un calcul séquentiel où une seule instruction fournie par une unité de contrôle globale est exécuté à un moment du temps et qui peut accéder à une seule mémoire centrale. Cette particularité, héritée par les ordinateurs actuels, permet de définir d'une manière assez simple comment un comportement complexe peut être obtenu à partir des comportements simples et quelques règles de composition.

Cependant, si l'approche ci-dessus fonctionne bien pour des calculs séquentiels, ou on peut représenter le résultat comme une suite d'opérations sur l'entrée, il n'est pas aisé de décrire des calculs parallèles ou le résultat est une émergence de différents calculs locaux, le problème principal venant du fait que le modèle de calcul inhérent est séquentiel. La situation devient encore plus difficile si on souhaite d'exprimer le comportement des processus qui sont massivement parallèles, comme il est le cas dans la plupart des applications biologiques où le résultat final émerge à partir d'un nombre élevé (de l'ordre de 10^{12}) de « calculs » locaux.

Tout cela conduit à une nécessité de redéfinir les règles de calcul classiques en proposant de nouveaux modèles de calcul plus adaptés aux nouvelles stratégies de l'obtention du résultat. Le point de départ du calcul naturel (natural computing) [KR08] consiste dans la formalisation des processus qui se passent dans la nature, en particulier, au niveau biochimique. Parmi les branches les plus connues de ce domaine on mentionne le calcul moléculaire (DNA computing) [PRS98] et le calcul à membranes (membrane computing) [Paun02, PPage]. Les modèles de calcul issus de ces domaines présentent des propriétés intéressantes comme l'absence de contrôle centralisé, des interactions purement locales, l'organisation hiérarchisée du système et des transformations plus complexes que la réécriture de mots (qui est l'opération de base de la théorie classique). Par exemple, dans les systèmes à membranes on considère un graphe dont les nœuds, les membranes, contiennent des multi-ensembles d'objets et on fait évoluer le système par des opérations qui sont inspirées de la réalité biologique des échanges moléculaires effectués par les membranes cellulaires et qui sont purement locales.

Objectifs de la tâche

On se propose de développer des modèles de calcul correspondant aux langages L1, L2 et L3 et d'étudier leur puissance d'expression et leurs relations avec d'autres modèles du calcul naturel, en particulier, avec les modèles du calcul à membranes. Comme le nombre d'ingrédients des langages considérés sera assez grand, on effectuera des recherches sur le nombre minimal d'ingrédients permettant d'effectuer des calculs intéressants. Cela nous permettra d'avoir une étude approfondie de la puissance d'expression et des classes de

complexité des différents sous-ensembles des langages L1, L2 et L3. Cette étude sera ensuite raffinée par celle des sous-classes de complexité correspondant aux calculs réalisables efficacement en parallèle. Comme la classe séquentielle P contient celle, NC, des calculs parallèles efficaces et réalistes [Gibbons88], on pourra définir et étudier des sous-classes parallélisables des nouvelles classes de complexité obtenues pour le calcul membranaire.

Puis, une étude de l'efficacité des différents ingrédients sera faite et, éventuellement, des consignes d'implantation pourront être proposées. L'étude des classes de complexité « parallèles membranaires » sera aussi spécialisée à des modèles encore plus concrets pour la prévision des performances et accélérations du calcul en fonction du nombre de molécules et de la vitesse des échanges. Ce genre de modèle s'inspirera du Bulk-synchronous computing ou BSP [Valiant89, Loulergue00] avec lequel on peut effectuer des prévisions quantitatives et réalistes de calculs parallèles avant même leur implantation. Ce genre de modélisation nous servira de base pour évaluer et comparer les futures implantations de nos langages et algorithmes de calcul membranaire. Les connaissances sur la parallélisation des opérations de démonstration automatique [Hains90] seront directement applicables à ces problèmes.

[...]

TACHE WP6 : BIO SURETE ET BIO SECURITE

Participants : Université d'Evry, Université Paris-Est Créteil

Contexte de la tâche

Certaines fonctions critiques touchant à la sécurité et sûreté de fonctionnement de l'artefact biologique doivent être vérifiées formellement en complément des simulations. En effet, si la simulation constitue un bon outil pour explorer les comportements d'un système et corriger les erreurs qu'on peut anticiper, la vérification par *model-checking* permet une exploration exhaustive des comportements et la détection d'erreurs inattendues.

Nous différencions les propriétés de sûreté qui sont essentiellement internes (est-ce que le système se comporte comme attendu) des propriétés de sécurité qui sont plutôt externes (est-ce que le système continue de fonctionner dans un environnement hostile). Ce deuxième aspect est important pour le projet car nous devons trouver un compromis entre deux aspects contradictoires liés à l'environnement :

- Robustesse des systèmes synthétisés qui doivent être fonctionnels dans des environnements variés. On pourrait parler de portabilité de ces systèmes entre différentes plateformes biologiques.
- Assurance que ces systèmes ne seront pas fonctionnels hors des environnements auxquels ils sont destinés. Il s'agit d'éviter le risque de dissémination incontrôlée.

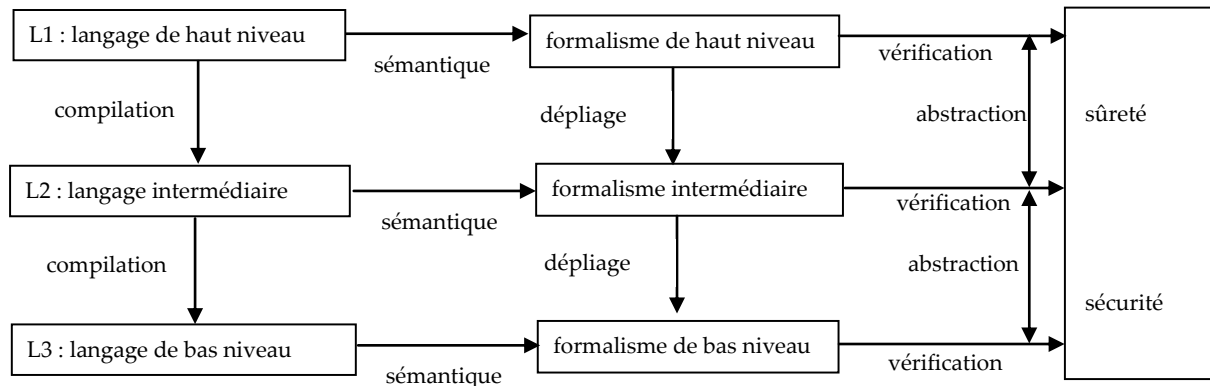
Les systèmes étudiés peuvent être décrits avec des outils de la biologie tels que les graphes d'interaction. Ces derniers peuvent s'exprimer avec des outils de la modélisation informatique tels que les algèbres de processus, les automates, les réseaux de Petri, etc. Cela offre la possibilité d'étudier leurs comportements dynamiques, notamment :

- Au niveau biologique : présence d'états stables, d'attracteurs ou de comportements cycliques.

- Au niveau informatique : accessibilité de configurations particulières, vivacité de sous-systèmes, présence de blocages (*livelocks* ou *deadlocks*), vérification de formules de logique temporelle, présence d'actions irréversibles.

Objectifs de la tâche

Les objectifs se déclinent selon deux axes, verticaux et horizontaux, conformément au schéma suivant :



Sur l'axe horizontal, nous souhaitons proposer pour chaque niveau de langage une sémantique formelle adaptée. Cette sémantique permettra de vérifier des propriétés de sûreté et de sécurité des systèmes. Sur l'axe vertical, deux aspects seront considérés : la certification partielle de la chaîne de compilation $L1 \rightarrow L2 \rightarrow L3$ et le développement de techniques permettant de pallier l'explosion combinatoire des techniques de vérifications classiques.

ORGANISATION DU PARTENARIAT

DESCRIPTION, ADEQUATION ET COMPLEMENTARITE DES PARTENAIRES

PARTENAIRE 1 : UNIVERSITE D'ÉVRY – IBISC

Le laboratoire IBISC est le laboratoire d'informatique de l'université d'Evry. Il a développé depuis 2000 de nombreuses collaborations dans le domaine de la modélisation en biologie et dans la biologie des systèmes dans le cadre du campus Genopole et au-delà en France et en Europe. Une des missions de Genopole est de favoriser le développement des interactions pluridisciplinaires autour de la génomique et Genopole vient d'inscrire la biologie synthétique comme axe prioritaire dans son programme 2010-2020. L'université d'Evry n'est pas en reste : la post-génomique est la priorité du quadriennal 2010-2013 et un bâtiment qui va recevoir les équipes du site travaillant dans le domaine de la bioinformatique, de la modélisation pour la biologie et en biologie synthétique est en cours de construction. Le projet **SYNBIOTIC** bénéficiera de cet environnement très riche et de toutes les interactions possibles avec les chercheurs des nombreux laboratoires impliqués dans des projets de biologies des systèmes sur le site.

Cependant, le projet est ancré en informatique qui veut développer une nouvelle méthodologie (informatique) de conception d'organismes biologiques artificiels. Il s'appuie pour cela sur les compétences des équipes LIS (langage, interaction et simulation) et DYNAMIC (dynamique cellulaire). Plus précisément les membres de ces équipes ont

développé des compétences en compilation parallèle (JL Giavitto, F. Delaplace) ; dans les modèles de programmation non-conventionnels (S. Sené, JL Giavitto) ; et en méthodes formels (H. Kludel) en particulier sur les outils de modélisation et de vérification symbolique de systèmes complexes. Cette expertise est complémentaire de celle apporté par les autres partenaires dans le domaine du parallélisme (G. Hains, O. Michel), des modèles de calcul (S. Verlan, O. Michel, A. Spicher, R. Doursat) et des méthodes formels pour la validation (F. Pommereau, G. Hains). Les approches discrètes pour la spécification et l'analyse des systèmes dynamiques sont au cœur de ce projet et correspondent aux compétences de S. Sené (automate cellulaire), JL Giavitto (systèmes de réécriture pour la simulation), H. Kludel (réseau de Petri, algèbre de processus), F. Delaplace (automate booléen, théorie des jeux).

SYNBIOTIC bénéficiera aussi plus largement de toute l'expertise du laboratoire IBISC autour de la modélisation des réseaux moléculaires dans les cellules vivantes et leurs simulations.

Les membres du partenaire 1 sont présent dans la plupart des tâches vu le positionnement des équipes et du laboratoire. Les liens avec les autres partenaires sont forts et existent depuis longtemps, ce qui permettra une mise en route rapide du projet.

PARTENAIRE 2 : UNIVERSITE DE PARIS-EST CRETEIL – LAACL

De par la variété et la complémentarité de leurs compétences, les membres du Partenaire 2 se retrouvent dans tous les **WP**. En effet, ils ont une expertise dans la définition d'outils théoriques et pratiques pour la définition de langages de haut-niveaux appliqués à la modélisation de systèmes biologiques complexes (O. Michel et A. Spicher), les modèles de calculs bio-inspirés (S. Verlan), l'algorithmique parallèle (G. Hains) et enfin la sécurité et la sûreté des grands systèmes complexes (F. Pommereau). Cette variété de champs thématiques sur un même site fait que les collaborations croisées pourront se mettre rapidement en place et bénéficier aux différents **WP** où ils participent. Par exemple, F. Pommereau, acteur majeur du **WP6** transversal (tout comme S. Verlan et G. Hains pour le **WP5**) pourra immédiatement collaborer avec O. Michel et A. Spicher sur la définition des langages L2 et L3.

De plus, O. Michel (ainsi que J.-L. Giavitto, Partenaire 1) ont été les artisans du projet 8 ½ (qui a développé une approche déclarative de la programmation spatiale), du projet MGS (qui a proposé une notion de forme différentielle discrète). Enfin, O. Michel a amorcé la modélisation multi-échelle hybride (continu/discret) de processus de diffusion. La poursuite de ce développement sera intégrée dans le projet **SYNBIOTIC** pour permettre une intégration des trois niveaux L1, L2 et L3.

PARTENAIRE 3 : CREA-ISC

R. Doursat est le Directeur de l'Institut des Systèmes Complexes, Paris Ile-de-France (ISC-PIF), CNRS, et Membre plein du CREA (cf. Paul Bourguine ci-dessous). Informaticien, il s'intéresse à la modélisation computationnelle de systèmes complexes autonomes et décentralisés à toutes les échelles, ainsi qu'à leur capacité d'inspirer une nouvelle forme de conception de systèmes artificiels. Il dispose d'importantes compétences en vie artificielle et ingénierie morphogénétique et co-organisera la première conférence sur le sujet [MEW09]. Ses travaux l'ont naturellement rapproché de la biologie synthétique et il participe, avec la quasi-totalité des membres de ce projet, au groupe de travail sur la biologie synthétique

GREBS. Son implication dans le projet SYNBIOTIC est essentielle car il est maître d'œuvre du **WP1** et intervient aussi dans les autres **WPs**. Ses interactions avec les membres des partenaires 1 et 2 vont nous permettre de concrétiser le projet sur des exemples biologiques.

P. Bourguin est le Directeur du CREA (Centre de Recherche en Epistémologie Appliquée), laboratoire de l'École polytechnique et UMR CNRS. L'originalité du CREA depuis sa fondation en 1982 est l'interaction entre épistémologie et modélisation. Le CREA s'intéresse depuis toujours aux questions fondamentales des systèmes complexes naturels et artificiels comme la reconstruction de leurs dynamiques multi-échelles, l'étude de leur stabilité, ainsi que leur conception et contrôle. Pour l'ensemble de ces questions, le CREA travaille en synergie avec l'ISC, dont il assure la responsabilité financière au travers du CNRS. Le CREA et l'ISC ont coordonné les projets européens *Embryomics* et *BioEmergences* dont le but était la reconstruction des dynamiques multi-échelles de l'embryogenèse, et qui pourront servir de base dans l'élaboration des modèles bio-inspirés et bio-implantés de **WP1**.

QUALIFICATION, ROLE ET IMPLICATION DES PARTICIPANTS

PARTENAIRE 1 – UNIVERSITE D'ÉVRY

Partenaire 1 : Université d'Evry – CNRS	Nom	Prénom	Emploi actuel
Coordinateur/responsable	Delaplace	Franck	PR
Autres membres	Giavitto	Jean- Louis	DR CNRS
	Klaudiel	Hanna	PR
	Sené	Sylvain	MCU

PARTENAIRE 2 - UNIVERSITE DE PARIS-EST CRETEIL – LACL

Partenaire 2 Université de Paris- Est Créteil – LACL	Nom	Prénom	Emploi actuel
Coordinateur/responsable	Michel	Olivier	PR
Autres membres	Hains	Gaétan	PR
	Pommereau	Franck	MC
	Spicher	Antoine	MC
	Verlan	Serghei	MC

PARTENAIRE 3 – CREA-ISC

Partenaire 3 CREA- ISC	Nom	Prénom	Emploi actuel
Coordinateur/responsable	Doursat	René	Chercheur contractuel CNRS
Autre membre	Bourguine	Paul	Directeur de recherche

ANNEXES

REFERENCES BIBLIOGRAPHIQUES

- [Abe01] Harold Abelson, Don Allen, Daniel Coore, Chris Hanson, George Homsy, Thomas F. Knight, Radhika Nagpal, Erik Rauch, Gerald Jay Sussman, and Ron Weiss. [Amorphous Computing](#), Abelson et al, [Communications of the ACM](#), Volume 43, Number 5, May 2001.
- [AIRRH03] Ch. Auffray, S. Imbeaud, M. Roux-Rouquié, and L. Hood. From functional genomics to systems biology: concepts and practices. *CR biologies*, 326(10–11):879–892, 2003.
- [Basu05] Basu, S.; Gerchman, Y.; Collins, C.; Arnold, F. and Weiss, R. A synthetic multicellular system for programmed pattern formation. *Nature* (2005), 434, 1130-1134
- [BDKPP08] Modeling and analysis of security protocols using role based specifications and Petri nets. R. Bouroulet, R. Devillers, H. Klaudel, E. Pelz and F. Pommereau. Proc. of ICATPN'08, LNCS 5062. Springer, 2008.
- [Beal08] J. Beal and J. Bachrach. Cells Are Plausible Targets for High-Level Spatial Languages, Spatial Computing Workshop, October 2008
- [BFGM05] J.-P. Banâtre, P. Fradet, J.-L. Giavitto and O. Michel, editors. Unconventional Programming Paradigms (UPP'04), LNCS volume 3566, Le Mont Saint-Michel, France, September 2005. ERCIM – NFS, Springer Verlag. Revised, selected and invited papers. 367 p. <http://www.springeronline.com/3-540-27884-2>
- [BGK05] F. Bernardini, M. Gheorghe, N. Krasnogor, and J.-L. Giavitto. On self-assembly in population p systems. In Proc. of UC'05 - Fourth International Conference on Unconventional Computation, volume 3699, pages 46 - 57. LNCS, Springer-Verlag, October 3-7, 2005, Sevilla.
- [Bir87] R. S. Bird. An introduction to the theory of lists. In M. Broy, editor, *Logic of Programming and Calculi of Discrete Design*, NATO ASI Series, vol. F36, pages 217-245, 1987.
- [BKP04] C. Bui Thanh, H. Klaudel and F. Pommereau. Box Calculus with High-Level Buffers. Proc. of DADS/ASTC'04, SCS, 2004.
- [BMH2] BioModularH2 : Engineered Modular Bacterial Photoproduction of Hydrogen, FP6-NEST Pathfinder Synthetic Biology project num. 043340 on using Synthetic Biology with cyanobacteria to produce hydrogen, <http://biomodularh2.org/>
- [Bro86] Brooks, R. A. (1986). A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, 2(1):14-23.
- [BS05] Jacob Beal and Gerald Sussman, “Biologically-inspired robust spatial programming,” Tech. Rep. AI Memo 2005-001, MIT, January 2005.
- [BS05] Benner, S.A. and Sismour, A.M. (2005) Synthetic biology. *Nat. Rev. Genet.*, 6, 533–543.
- [Cal06] M. Calder et al. Modelling the influence of RKIP on the ERK signalling pathway using the stochastic process algebra PEPA. *Trans. on Comput. Syst. Biol* VII, 4230:1–23, 2006.
- [Cal09] Travaux en cours (non publiés) au sein du projet Calamar ANR-08-SYSC-003
- [Car03] R. Carlson. The Pace and Proliferation of Biological Technologies. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science*, Volume 1, Number 3, 2003
- [Car04] L. Cardelli. Brane calculi. In Vincent Danos and Vincent Schächter, editors, *Computational Methods in Systems Biology*, volume 3082 of *Lecture Notes in Computer Science*, pages 257–278, Berlin, 2004. Springer.
- [Car09] L. Cardelli. Artificial biochemistry. In *Algorithmic Bioprocesses*. To appear in LNCS. <http://lucacardelli.name/Bibliography.htm#ArtificialBiochemistry>

- [Cha04] N. Chabrier-Rivier et al. The biochemical abstract machine BIOCHAM. In V. Danos and V. Sch chter, editors, Proc. CMSB, volume 3082 of LNCS, pages 172–191. Springer, 2004.
- [CHG07] Cai Y, Hartnett B, Gustafsson C, Peccoud J (2007) *A syntactic model to design and verify synthetic genetic constructs derived from standard biological parts*. *Bioinformatics* 23(20): 2760-2767.
- [Chi06] Chin, J.W. (2006) Programming and engineering biological networks. *Curr. Opin. Struct. Biol.*, 16, 551–556.
- [CIA03] Central Intelligence Agency, Directorate of Intelligence, “The Darker Bioweapons Future,” November 3, 2003; <http://www.fas.org/irp/cia/product/bw1103.pdf>.
- [Clay04] S. Copen Goldstein and T. C. Mowry. Claytronics: A scalable basis for future robots. In *RoboSphere 2004*, November, 2004.
- [CLE08] B. Canton, A. Labno, and D. Endy, “Refinement and standardization of synthetic biological parts and devices,” *Nature Biotechnology*, vol. 26, pp. 787–93, July 2008.
- [Coo04] D. Coore, “Towards a universal language for amorphous computing,” in Fifth International Conference on Complex Systems, 2004.
- [Coo99] D. Coore. *Botanical Computing: A Developmental Approach to Generating Inter connect Topologies on an Amorphous Computer*. PhD thesis, Massachusetts Institute of Technology, 1999.
- [Cous08] G. Cousineau. Groupe de réflexion « sciences et technologies du logiciel ». http://www.snics.fr/IMG/pdf/Rapport_sur_le_logiciel.pdf
<http://www.enseignementsup-recherche.gouv.fr/cid22477/debut-de-la-mise-en-oeuvre-operationnelle-du-plan-strategique-horizon-2020-du-cnrs.html>
- [CRLM04] E. Coen, A.-G. Rolland-Lagan, M. Matthews, J. A. Bangham, and P. Prusinkiewicz, [The genetics of geometry](#), *PNAS April 6, 2004 vol. 101 no. 14 4728-4735*
- [Dae06] Dae-Kyun Ro, E. M. Paradise, M. Ouellet, K. J. Fisher, K. L. Newman, J. D. Keasling, et al., “[Production of the antimalarial drug precursor artemisinic acid in engineered yeast](#),” *Nature*, vol. 440, no. 7086 (April 13, 2006), pp. 940-943.
- [Dag06] A. DeHon, J.-L. Giavitto, and F. Gruau, editors. [Computing Media and Languages for Space-Oriented Computation](#), 3-8 september 2006. Dagstuhl Seminar Proceedings N 06361.
- [Dan07] V. Danos et al. Rule-based modelling of cellular signalling. In L. Caires and V. T. Vasconcelos, editors, *CONCUR*, volume 4703 of LNCS, pages 17–41. Springer, 2007. Tutorial paper.
- [Des6] M. Desbrun, E. Kanso, and Y. Tong. Discrete differential forms for computational modeling. In *Discrete differential geometry: an applied introduction*, pages 39-54. 2006.
- [DKKP03] H. Kludel, M. Koutny and F. Pommereau. Asynchronous Box Calculus. Raymond Devillers, *Fundamenta Informaticae*, 54(1), IOS Press, 2003.
- [dLor08] de Lorenzo, V. (2008) Systems Biology approaches to Bioremediation. *Curr. Op. Biotech.* 19: 579-89
- [DMS08] J. Demongeot, M. Morvan and S. Sen . Robustness of Dynamical Systems Attraction Basins Against State Perturbations: Theoretical Protocol and Application in Systems Biology. *CISIS 2008 Proceedings*, 675-681, IEEE Press, 2008.
- [DNS10] J. Demongeot, M. Noual and S. Sen . On the Number of Attractors of Boolean Automata Circuits. *BLSMC 2010*, accepted.
- [Dou06] Doursat, R. (2006b) The growing canvas of biological development: Multiscale pattern generation on an expanding lattice of gene regulatory networks. *InterJournal: Complex Systems* 1809.

- [Dou08b] Doursat, R. (2008b) Organically grown architectures: Creating decentralized, autonomous systems by embryomorphic engineering. In *Organic Computing*, R. P. Würtz, ed., pp. 167-200, Springer-Verlag.
- [Dou08d] Doursat, R. (2008d) Programmable architectures that are complex and self-organized: From morphogenesis to engineering. *11th International Conference on the Simulation and Synthesis of Living Systems (ALIFE XI)*, August 5-8, 2008, University of Southampton, Winchester, UK. In *Artificial Life XI*, S. Bullock, J. Noble, R. Watson & M. A. Bedau, eds., pp. 181-188, MIT Press.
- [EKL02] S. Eker, M. Knapp, K. Laderoute, P. Lincoln, and C. L. Talcott. Pathway logic: Executable models of biological networks. *Electr. Notes Theor. Comput. Sci*, 71, 2002.
- [EL00] Elowitz, M.B. and Leibler, S. (2000) A synthetic oscillatory network of transcriptional regulators. *Nature*, 403, 335-338.
- [End05] Endy, D. (2005) Foundations for engineering biology. *Nature*, 438, 449-453.
- [ETC07] Etc Group. Extreme Genetic Engineering: An Introduction to Synthetic Biology. 2007 <http://www.etcgroup.org/en/materials/publications.html?id=602>
- [Fae05] J. R. Faeder et al. Graphical rule-based representation of signal-transduction networks. In L. M. Liebrock, editor, *Proc. 2005 ACM Symp. Appl. Computing*, pages 133-140. ACM Press, 2005.
- [FMP00] M. Fisher, G. Malcolm, and R. Paton. Spatio-logical processes in intracellular signalling. *BioSystems*, 55:83-92, 2000.
- [Gar00] Gardner, T.S. et al. (2000) Construction of a genetic toggle switch in *Escherichia coli*. *Nature*, 403, 339-342.
- [GC05] The Grand Challenge in Non-Classical Computation, The King's Manor, York, United Kingdom, 18-19 April 2005. University of York and Microsoft Research.
- [GC7] R. Milner, S. Stepney *et al.* Grand Challenge 7: "Journeys in Non-Classical Computation", A UK Grand Challenge in Computing Research. Presentation in [Int. J. Parallel, Emergent and Distributed Systems 20\(1\):5-19, 2005](#) and [Int. J. Parallel, Emergent and Distributed Systems 21\(2\):97-125, 2006](#). Cf. also the web site: http://www.ukrc.org.uk/grand_challenges/ and <http://www.cs.york.ac.uk/nature/gc7/>
- [GGMP02a] J.-L. Giavitto, Ch. Godin, O. Michel, and P. Prusinkiewicz. Computational models for integrative and developmental biology. Technical Report 72-2002, LaMI – Université d'Évry Val d'Essonne, March 2002.
- [Gib04] W. Wayt Gibbs. Synthetic life. *Scientific American*, 2004.
- [Gibbons88] Alan Gibbons and Wojciech Rytter, *Efficient Parallel Algorithms*, Cambridge University Press, 1988.
- [GLRT04] F. Gruau, Y. Lhuillier, Ph. Reitz, and O. Temam. Blob Computing. In *Computing Frontiers 2004 ACM SIGMicro.*, 2004
- [GM01b] J.-L. Giavitto and O. Michel. MGS: a programming language for the transformations of topological collections. Technical Report 61-2001, LaMI – Université d'Évry Val d'Essonne, May 2001.
- [GM02c] J.-L. Giavitto and O. Michel. The topological structures of membrane computing. *Fundamenta Informaticae*, 49:107-129, 2002.
- [GM03] J.-L. Giavitto and O. Michel. Modeling the topological organization of cellular processes. *BioSystems*, (70):149-163, 2003.
- [GM72] Gierer, A., and Meinhardt, H. (1972) A theory of biological pattern formation, *Kybernetik* 12: 30-39.

- [GMM04] J.-L. Giavitto, G. Malcolm, and O. Michel. Rewriting systems and the modelling of biological systems. *Comparative and Functional Genomics*, 5:95–99, February 2004.
- [GS08] J.-L. Giavitto, A. Spicher. *Topological Rewriting and the Geometrization of Programming*. *Physica D*, vol. 237, Issue 9, p. 1302–1314. July 2008.
- [Gue02] Guet, C. C. et al. (2002) Combinatorial synthesis of genetic networks. *Science*, 296, 1466–1470.
- [Hains90] G. Hains. La complexité du filtrage associatif-commutatif. *Comptes Rendus de l'Académie des Sciences, Paris*, t.311, Série I(11):741–744, Novembre 1990.
- [HDH09] J. R. Heath, M. E. Davis and L. Hood. Nanomedicine--Revolutionizing the Fight against Cancer. *Scientific American Magazine*. January 2009.
<http://www.sciam.com/article.cfm?id=nanomedicine-targets-cancer>
- [HH98] T. Hogg and B. A. Huberman. Controlling smart matter. *Smart Materials and Structures*, 7:R1, 1998.
- [Hoo06] S. Hoops et al. COPASI – a COMplex PATHway Simulator. *Bioinformatics*, 22(24):3067–74, 2006.
- [Hor01] P. Horn. [Autonomic computing: IBM's perspective on the state of information technology](#). Technical report, IBM Research, October 2001.
- [HP06] Heinemann, M. and Panke, S. (2006) Synthetic biology – putting engineering into biology. *Bioinformatics*, 22, 2790–2799.
- [HTK08] A. Hamez, Y. Thierry-Mieg, F. Kordon. Hierarchical Set Decision Diagrams and Automatic Saturation, ICATPN 2008, LNCS 5062, pp. 211–230, Springer.
- [iGE] International Genetically Engineered Machine competition (iGEM) <http://2009.igem.org> est une compétition internationale dans le domaine de la biologie synthétique. Organisée la première fois en 2004, cette compétition s'adresse aux étudiants « undergraduate ». Les équipes reçoivent un kit de BioBrick issu du catalogue de composants biologiques standards [REG] au début de l'été. Elles présentent en Novembre au MIT les composants qu'ils ont été amenés à créer et les circuits de régulation qu'ils ont synthétisés dans des cellules. La première équipe française, iGEM Paris, a participé à iGEM'07. Les partenaires d'Evry et de Créteil faisaient parti des encadrants.
- [Int] <http://interstices.info>
- [KK01] V. Khomenko and M. Koutny. Towards An Efficient Algorithm for Unfolding Petri Nets, ICATPN 2001, LNCS 2154, Springer.
- [KKA+04] H. Kobayashi, M. Kærn, M. Araki, K. Chung, T. S. Gardner, Ch. R. Cantor, and J. J. Collins. Programmable cells: Interfacing natural and engineered gene networks. *PNAS*, 101(22):8114–8419, June 2004.
- [KKPP08] H. Kludel, M. Koutny, E. Pelz and F. Pommereau. Towards Efficient Verification of Systems with Dynamic Process Creation. *Proc. of ICTAC'08*, LNCS 5160. Springer, 2008.
- [Kn03] T.F. Knight, "Idempotent vector design for standard assembly of biobricks," Tech. Rep., MIT Synthetic Biology Working Group Technical Reports, 2003.
- [Kni03] T.F. Knight, Idempotent vector design for standard assembly of biobricks, Tech. Rep., MIT Synthetic Biology Working Group Technical Reports, 2003
- [Kob04] Kobayashi, H. et al. (2004) Programmable cells: interfacing natural and engineered gene networks. *Proc. Natl Acad. Sci. USA*, 101, 8414–8419.
- [KP08] H. Kludel and F. Pommereau. M-nets, a survey. *Acta Informatica* 45. Springer, 2008.
- [KR08] L. Kari, G. Rozenberg, The many facets of natural computing. *Commun. ACM* 51, 10 (Oct. 2008), 72–83.
- [Law92] P. Lawrence, editor. *The Making of a Fly*. Blackwell Scientific, 1992.

- [LEF08] F. Le Fèvre. Biologie Synthétique, quel business modèle ? Mémoire de master d'Administration des Entreprises, Institut d'Administration des Entreprises, Université de Paris 1 Panthéon Sorbonne, 2008
- [Loulergue00] F. Loulergue, G. Hains et C. Foisy. A calculus of functional BSP programs. *Science of Computer Programming*, 37(1-3):253-277, May 2000.
- [MBFG06] O. Michel, J.-P. Banâtre, P. Fradet and J.-L. Giavitto. Challenging questions for the rationale of non-classical programming languages. *International Journal of Unconventional Computing (IJUC)*, 2(4):337-347, 2006.
- [McDW05] McDaniel, R. and Weiss, R. (2005) Advances in synthetic biology: on the path from prototypes to applications. *Curr. Opin. Biotechnol.*, 16, 476-483.
- [Mei82] H. Meinhardt. Models of biological pattern formation. Academic Press, New York, 1982.
- [MEW09] [First International Workshop on Morphogenetic Engineering](#), organized by R. Doursat, H. Sayama, O. Michel. Institut des Systèmes Complexes, Paris Ile-de-France, June 19, 2009.
- [MSG08] O. Michel, A. Spicher, and J.-L. Giavitto. Rule-Based Programming for Integrative Biological Modeling. *Natural Computing*, 2008
- [NaCo98] [An Algorithm for Group Formation in an Amorphous Computer](#), Nagpal and Coore, Intl Conf on Parallel and Distributed Computing Systems (PDCS'98), 1998
- [Nag01] R. Nagpal, Programmable Self-Assembly: Constructing Global Shape using Biologically-inspired Local Interactions and Origami Mathematics, Ph.D. thesis, MIT, 2001.
- [NaMa03] Nagpal, Mamei, Engineering Amorphous Computing Systems, invited chapter in *Methodologies and Software Engineering for Agent Systems*, editors Bergenti, Gleizes, Zambonelli, Kluwer Academic Publishing 2003
- [Paun02] Gh. Paun, Membrane Computing: An Introduction, Springer-Verlag, 2002.
- [PRS98] Gh. Paun, G. Rozenberg, A. Salomaa, DNA Computing: New Computing Paradigms, Springer-Verlag, 1998.
- [PDK07] F. Pommereau, R. Devillers and H. Kludel. Efficient reachability graph representation of Petri nets with unbounded counters. Proc. of Infinity'07, ENTCS. Elsevier, 2007.
- [Pir] NSF Partnerships for International Research and Education (PIRE) projects: http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=12819
- [PLH+90] P. Prusinkiewicz, A. Lindenmayer, J. Hanan, et al. The Algorithmic Beauty of Plants. Springer-Verlag, 1990.
- [SMCG08] A. Spicher, O. Michel, M. Cieslak, J.-L. Giavitto, P. Prusinkiewicz. Stochastic P systems and the simulation of biochemical processes with dynamic compartments. *BioSystems*, 91(3):458-472, March 2008
- [SG08] A. Spicher and J.-L. Giavitto, Systems Self-Assembly: multidisciplinary snapshots chapter: Simulation of self-assembly processes using abstract reduction systems edited by Natalio Krasnogor, Steve Gustafson, David Pelta and Jose L. Verdegay, Elsevier 2008
- [PP08] M. Pedersen, G. Plotkin. A Language for Biochemical Systems, Computational Methods in Systems Biology (CMSB08), Rostock, Germany, October 12-15, 2008. Proceedings. Lecture Notes in Computer Science 5307, p. 63-82, Springer 2008.
- [PA09] M. Pedersen, A. Phillips. Towards programming languages for genetic engineering of living cells, *J. R. Soc. Interface*, vol. 6 no. Suppl 4 S437-S450, August 2009.
- [PP10] M. Pedersen, G. Plotkin. A Language for Biochemical Systems: Design and Formal Specification, to appear in *Transactions on Computational Systems Biology*.
- [PPage] Membrane Computing WEB page: <http://ppage.psystems.eu/>

- [PR02] G. Paun and G. Rozenberg. A guide to membrane computing. *Theor. Comput. Sci.*, 287(1):73–100, 2002.
- [Pur09] Purnick PE, Weiss R. The second wave of synthetic biology: from modules to systems. *Nature reviews. Molecular cell biology*. 2009;10(6):410-22. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19461664>.
- [Rao02] Rao CV, Wolf DM and Arkin AP, Control, exploitation and tolerance of intracellular noise, *Nature* (2002) 420, 231-237
- [Red93] V. N. Reddy et al. Petri net representation in metabolic pathways. In *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, pages 328–336, 1993.
- [REG] The Registry of Standard Biological Parts. <http://partsregistry.org>. Ce site collaboratif catalogue les séquences élémentaires utilisé dans la synthèse de circuit de régulation (principalement chez E. Coli). La plupart de ces séquences ont été développées lors des compétitions iGEM. Elles satisfont au standard des BioBricks, permettant leur composition en aisé en circuits de régulation par clonage. Cette approche est actuellement plus économique que la synthèse directe de la séquence totale représentant le circuit.
- [Reg01] A. Regev et al. Representation and simulation of biochemical processes using the pi-calculus process algebra. In *Pacific Symposium on Biocomputing*, pages 459–470, 2001.
- [Reg04] A. Regev et al. BioAmbients: an abstraction for biological compartments. *Theor. Comput. Sci.*, 325(1):141–167, 2004.
- [Rey87] Reynolds, C. W. (1987) Flocks, herds and schools: a distributed behavioral model. *Computer Graphics*, 21(4): 25-34.
- [Ric06] A. Richard. Modèle formel pour les réseaux de régulation génétique & Influence des circuits de rétroaction. Thèse de doctorat, Université d'Évry, 2006.
- [RPW04] P. W. K. Rothmund, N. Papadakis, and E. Winfree. Algorithmic self-assembly of dna sierpinski triangles. *PLoS Biol*, 2(12):e424, 2004. www.plosbiology.org.
- [SASO08] [Spatial Computing Workshop](#). Satellite workshop of Self-Adaptive and Self-Organizing Systems Conference. October 2008.
- [SFD07] P. Speroni di Fenizio, and P. Dittrich, Chemical Organizations at Different Spatial Scales, *Advances in Artificial Life, LNCS, Volume 4648/2007*, 2007.
- [Spi09] A. Spicher, N. Fatès, O. Simonin. From Reactive Multi-Agents Models to Cellular Automata. *Proceedings of ICAART'09 . INSTICC/Springer/LNCS*, 2009
- [SMCG08] A. Spicher, O. Michel, M. Cieslak, J.-L. Giavitto, Przemyslaw Prusinkiewicz. Stochastic P systems and the simulation of biochemical processes with dynamic compartments. *BioSystems*, 91(3):458-472, March 2008
- [Sto04] K. Stoy, “Controlling self-reconfiguration using cellular automata and gradients,” in 8th int. conf. on intelligent autonomous systems (IAS-8), 2004.
- [TES08] TESSY (Towards a European Strategy for Synthetic biology). Future Perspectives in Synthetic Biology. TESSY Final Report (Deliverable D5.3) December 2008. <http://www.tessy-europe.eu>
- [TK01] R. Thomas and M. Kaufman. Multistationarity, the basis of cell differentiation and memory. i. structural conditions of multistationarity and other nontrivial behavior. *Chaos*, 11(1):170–179, 2001.
- [Tur52] Turing, A. M. (1952) The chemical basis of morphogenesis. *Phil. Trans. R. Soc. London B* 237: 37-72.
- [UC07] Unconventional computation: Quo Vadis, Santa-Fe, 21-23 march 2007. Center for Nonlinear Studies (CNLS). <http://cnls.lanl.gov/uc07/>

- [Upim03] B. Mac Lennan. Combinatory Logic for Autonomous Molecular Computation. *Information Sciences*, 2003
- [Valiant89] L. Valiant, General Purpose Parallel Architectures in Handbook of Theoretical Computer Science, Van Leeuwen editor, North-Holland, 1989.
- [Ven05] J. Craig Venter, prepared statement, [Hearing Examining Twenty-First Century Biological Threats, Focusing on Dual-Purpose Preparedness Against Natural and Deliberate Microbial Threats](#), Subcommittee on Bioterrorism and Public Health Preparedness, Committee on Health, Education, Labor, and Pensions, United States Senate, 109th Congress, 1st sess. (May 11, 2005), p. 40.
- [Vi95] Vicsek, T., Czirók, A., Ben-Jacob, E., Cohen, I. and Shochet, O. (1995). Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 75: 1226–1229.
- [Viv] <http://www.vivagora.org>
- [Voi06] Voigt, C.A. (2006) Genetic parts to program bacteria. *Curr. Opin. Biotechnol.*, 17, 548–557.
- [Wei01] R. Weiss, Cellular Computation and Communications using Engineered Genetic Regulatory Networks, Ph.D. thesis, MIT, 2001.
- [Wei03] Weiss, Basu, Hooshangi, Kalmbach, Karig, Mehreja, Netravali, Genetic circuit building blocks for cellular computation, communications, and signal processing, *Natural Computing* (2003) 2:47-84
- [Wei05] R. McDaniel and R. Weiss, Advances in synthetic biology: on the path from prototypes to applications, *Current Opinion in Biotechnology* (2005) 16:476–483
- [Whe04] M. Wheelis, “Will the New Biology Lead to New Weapons?”, *Arms Control Today* (July/August 2004), http://www.armscontrol.org/act/2004_07-08/Wheelis.asp
- [Wol02] L. Wolpert. Principles of Development. Oxford University Press, 2nd edition, 2002.
- [Yam07] D. Yamins, A Theory of Local-to-Global Algorithms for One-Dimensional Spatial Multi-Agent Systems, Ph.D. thesis, Harvard, December 2007.
- [YKDK07] A. Yartseva, H. Kludel, R. Devillers and F. Képès. Incremental and unifying modelling formalism for biological interaction networks. *BMC-Bioinformatics* 8:433, 2007, pages 1-69, <http://www.biomedcentral.com/1471-2105/8/433>.

—oOo—