

KIMONO: a modelling methodology for hypotheses evaluation and research orientation

Mots clés :

- **Directeur de thèse** : alexis DROGOUL
- **Co-encadrant(s)** :
- **Unité de recherche** : Unité de modélisation mathématique et informatique de systèmes complexes
- **Ecole doctorale** : École Doctorale Informatique, Télécommunications, Électronique de Paris
- **Domaine scientifique principal**: Divers

Résumé du projet de recherche (Langue 1)

Depuis plusieurs années, on peut observer une évolution de la modélisation dans de nombreux domaines (Sociologie, Ecologie, Economie, etc.), des modèles qui sont plus orientés vers l'exploration que simplement vers la prédiction. Se basant sur le paradigme de "système complexe", ces modèles exploratoires utilisent généralement une représentation explicite et détaillée des composants du système étudié. Ils offrent ainsi aux utilisateurs finaux (chercheurs, décideurs, parties prenantes) une grande liberté d'adaptation en termes de paramètres et de structure du modèle. Ces modèles sont conçus comme un outil permettant de réaliser des expériences "as-if" car ils permettent la formulation d'hypothèses détaillées et ce, à différents niveaux de descriptions du système. A partir de ces hypothèses, des scénarios sont construits dont les résultats sont explorés et analysés grâce à des simulations contrôlées et répétées. L'épidémiologie est exemplaire de cette situation. Cette discipline a déjà une longue histoire de modélisation qui peut être caractérisée comme la recherche de l'amélioration de la capacité prédictive des modèles. Cependant, des situations réelles comme l'émergence des foyers épidémiques de grippe aviaire en Asie du Sud Est montre la limite d'une telle approche: sans prise en compte adéquate des interactions entre les dynamiques sociale, écologique et biologique, l'utilisation de modèles à visée prédictive est hors de propos. De plus, quand l'on tente de prendre en compte ces dynamiques, les modèles obtenus sont alors dépendant de données incomplètes et/ou qualitative (le processus du décision des acteurs sociaux ou bien le comportement des oiseaux, par exemple). Par conséquent, il y a eu une remise en question de l'orientation de la communauté vers la conception de modèles exploratoires qui sont adaptés pour la génération et études d'hypothèses variées, et pour mesurer, par rapport à ces hypothèses, l'impact local et global de changement de politiques lors de scénarios complexes. Cependant, concevoir et utiliser de tels modèles met en lumière de sérieux problèmes méthodologiques auxquels ne peuvent réellement répondre les méthodologies de modélisation et de simulation existantes. Quand celle-ci sont adaptées pour prendre ces spécificités en compte, cela résulte en des solutions ad hoc qui ne peuvent être réutiliser dans le domaine en général, sans parler d'autres domaines. L'objectif de cette thèse est donc de proposer une méthodologie que ne soit pas spécifique à un domaine particulier mais une méthodologie générique et qui facilite la conception et l'utilisation de ces modèles exploratoires. En partant d'une série d'exemples de différents domaines (trafic routier, ségrégation sociale, dynamique du sol mais aussi, et de manière bien plus extensive, en épidémiologie), je commence par une caractérisation des besoins de conception (prise en compte d'hypothèses contradictoires et évolutives lors du processus de modélisation, génération de modèles extrêmement modulaires, rendant possible un cycle de modélisation itératif, permettant la collaboration entre différents experts et la combinaisons de différents formalismes, etc.) et ensuite une proposition concrète impliquant des outils informatiques dédiés et un formalisme commun et accessible, orienté aussi bien vers la facilitation de la collaboration, la communication et l'implémentation de "modèles monde" (le nom donné à ces modèles exploratoires. La méthodologie que je propose se concentre sur deux éléments: l'implication des experts et la représentation détaillée du système. Les experts sont au cœur de processus de modélisation, celui-ci commence par une description étendue de leur connaissance, possiblement exprimés dans leurs propres formalismes, par la suite cette description est amendée itérative (que ce soit en la complexifiant ou en la simplifiant). Ce processus itératif s'arrête seulement quand les experts estiment qu'ils ont obtenus suffisamment de précisions sur le système étudié ou quand la poursuite de ces itérations nécessite de nouvelles expériences ou données de terrain à collecter. Concernant les types de représentations adaptés pour un tel processus, je propose une combinaison modulaire et adaptable de deux systèmes implémentés: les modèles à base d'agent et (MBA) et les Systèmes d'Information Géo-référencées (SIG). Je montre que cette combinaison fournit une représentation des composants du système quelque soit le niveau de description et ce que l'information soit qualitative ou quantitative. De plus, cela permet un haut niveau d'évolution des hypothèses lors du processus de modélisation. L'interaction entre les modélisateurs et les thématiciens se base sur deux abstractions de cette implémentation: ODD (Overview, Design concepts, Details) qui est un protocole de communication de modèle et GAML un langage de modélisation pour la programmation collaborative du modèle. La méthodologie proposée a été appliquée et validée dans le contexte d'une étude détaillée située en Asie du Sud Est (essentiellement au Nord Vietnam) par des épidémiologistes et des vétérinaires. Ceux-ci voulaient évaluer l'effet d'hypothèses diverses expliquant la ré-apparition récurrente de foyers épidémiques de grippe aviaire parmi les volailles domestiques. Pendant cette coopération interdisciplinaire qui a duré quatre ans, plusieurs "modèles monde" ont été co-conçu et implémenté sur la plate-forme GAMA et utilisé comme "laboratoires virtuels" par les experts. Cette collaboration, et ces résultats, ont permis de tester un large champ d'hypothèses (en particulier sur les conditions locales de persistances), d'avoir une meilleure compréhension du rôle de l'environnement spatiale, des facteurs écologiques et sociaux dans la survie et la propagation du virus et a aussi permis de réorienter certaines des études de terrains.

