

Analyse formelle d'exigences en langue naturelle pour la conception de systèmes cyber-physiques

Aurélien lamercherie
aurelien.lamercherie@inria.fr

Introduction

La motivation de cette proposition est d'apporter une aide à la **conception des systèmes cyber-physiques**, en permettant de simuler et vérifier, par des méthodes automatiques ou assistées, des cahiers des charges exprimés en **langue naturelle**.

Un **cahier des charges** peut se ramener à un ensemble de propriétés qui pourrait être traité en deux phases :

1. **Analyse formelle** des exigences pour aboutir à leur représentation sémantique
2. **Composition** de ces exigences pour obtenir une représentation du comportement du système

En partant d'un ensemble de propriétés, les objectifs suivants sont visés :

1. Obtenir une représentation visuelle du comportement du système
2. Mettre en évidence les incohérences de conception

Analyse syntaxico-sémantique

Les représentations sémantiques sont obtenues à partir de **dérivations CCG** (grammaire catégorielle combinatoire). Chaque élément lexical est associé à une catégorie lexicale et une **représentation sémantique** (lambda-calcul). Le traitement est une application récursive de fonctions sur les arguments jusqu'à obtenir la représentation sémantique.

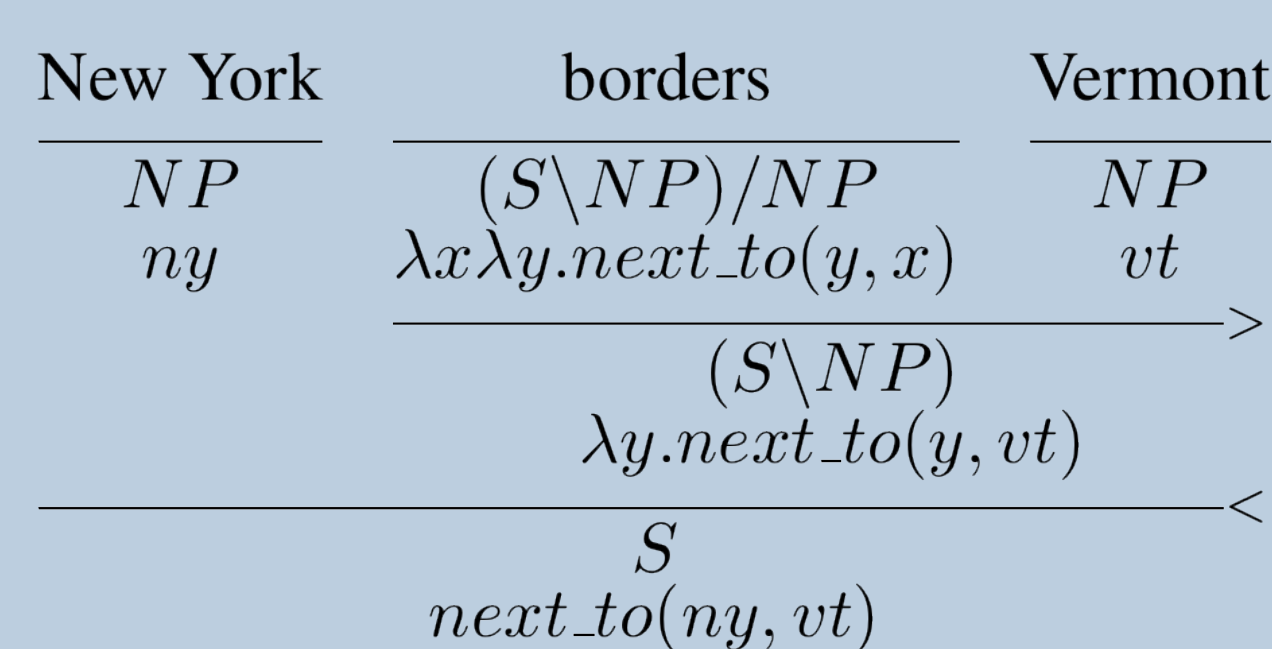


FIGURE 1: Exemple de dérivation syntaxico-sémantique

Résolution des références

Un module spécifique permet de résoudre les **références aux concepts** en s'appuyant sur une ontologie du domaine d'étude. Cette étape intermédiaire permet de relier les formules en lambda-calcul à leur représentation sous forme de spécifications modales.

Représentation du système

Les **spécifications modales** permettent de représenter des automates dont les transitions sont typées avec les modalités *must* et *may*. Ce formalisme est un bon compromis en termes d'expressivité, de propriétés algébriques et de complexité algorithmique.

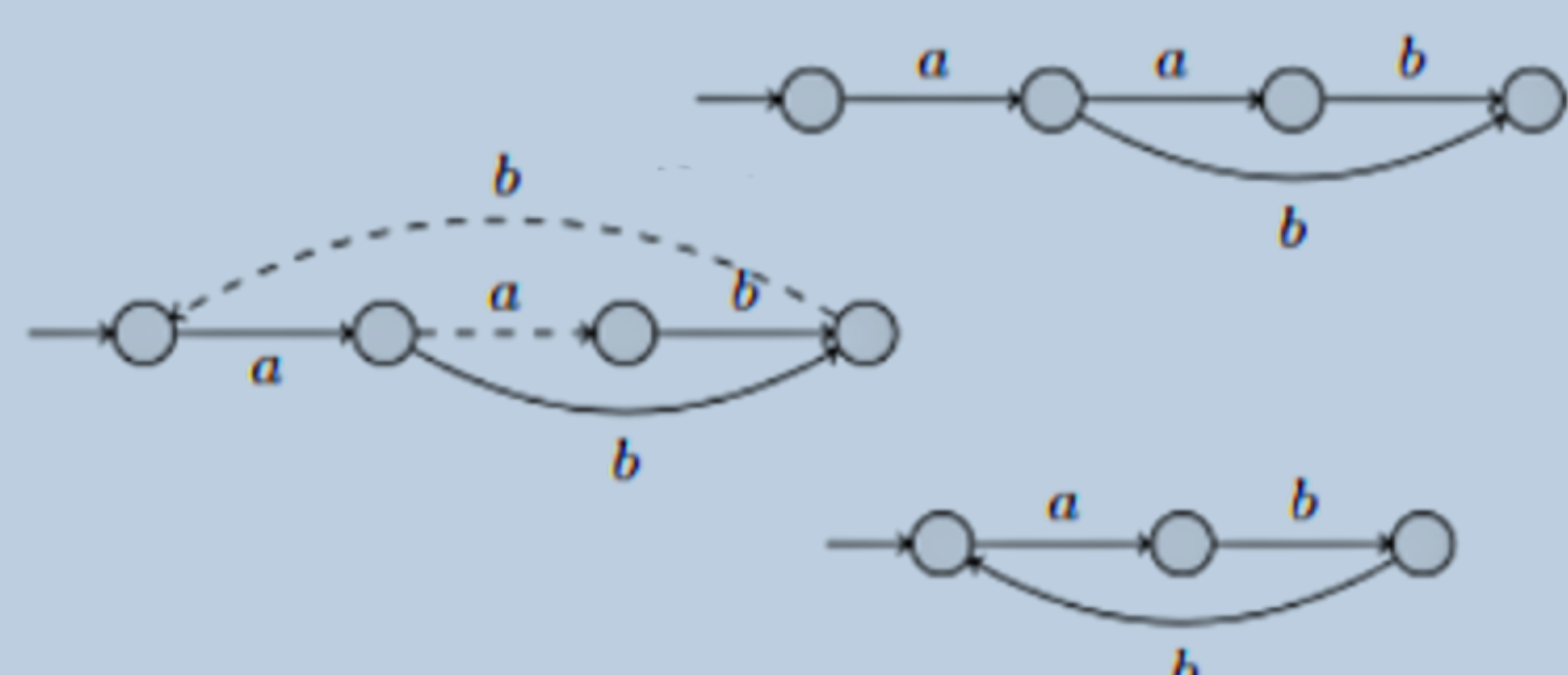
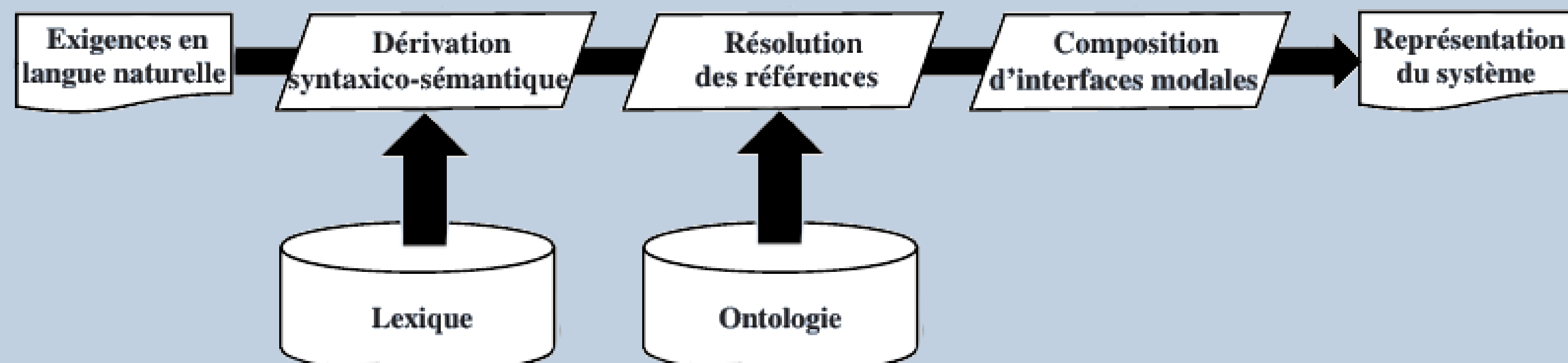


FIGURE 2: Une spécification modale (à gauche) et deux modèles possibles (à droite)

La définition d'un système complexe est réalisable en appliquant un ensemble d'opérations.

Proposition



Cas d'étude : "The Parking Garage"

Un cas d'étude théorique, représentant un garage de voitures, a été notre support pour des expérimentations préliminaires. Le document de spécifications est composé de plusieurs sections : (*gate, payment, supervisor*). Le tableau suivant reprend **quelques exigences** portant sur la section "gate".

Id.	Exigence
1	Entry gate must open once a ticket has been issued.
2	When a ticket has been issued, entry gate must open.
3	When a ticket has been issued, entry gate opening is obligatory.
4	Afer an exit ticket is inserted, exit gate must open.

TABLE 1: Quelques exigences du système

L'outil **Boxer/C&C** a été utilisé pour traiter les exigences et construire des représentations sémantiques. Boxer/C&C est un analyseur sémantique de textes en anglais. Il s'appuie notamment sur la grammaire catégorielle combinatoire et permet de construire des **représentations sémantiques** sous différentes formes. La figure suivante donne une représentation de l'exigence 4 du tableau précédent.

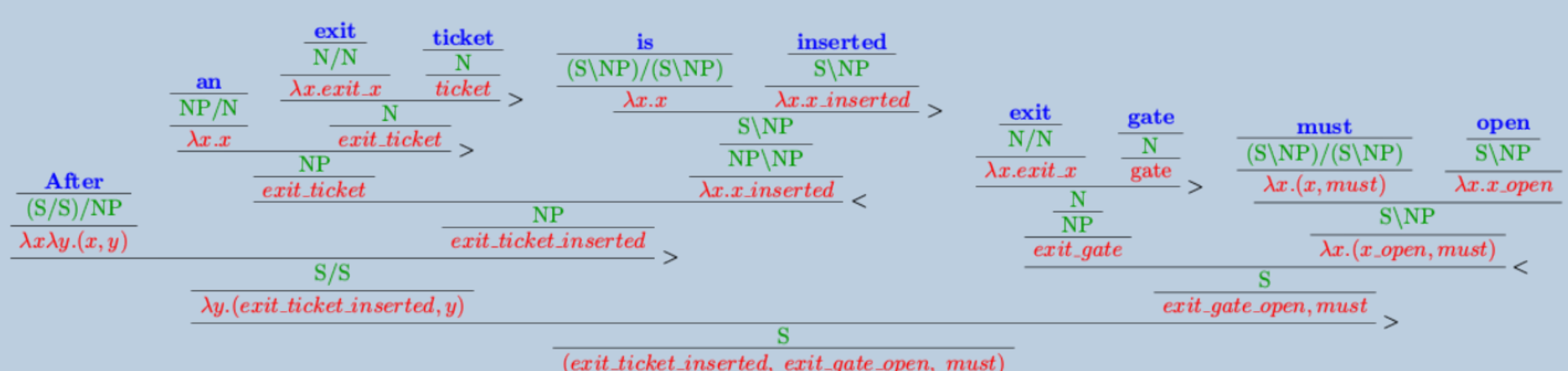


FIGURE 3: Représentation graphique de l'analyse sémantique pour l'exigence 4

L'outil **Mica** permet de définir des systèmes complexes en utilisant une syntaxe du langage O'CamL. Le passage des exigences formelles (en lambda-calcul) vers cet outil pourrait être réalisé avec un module d'interface s'appuyant sur une ontologie. Il a été fait manuellement pour obtenir une représentation du résultat sous la forme d'une **spécification modale**, en prenant en compte les exigences de la section "Gate". Cette représentation intègre l'analyse de l'exigence 4 (états 0 et 1).

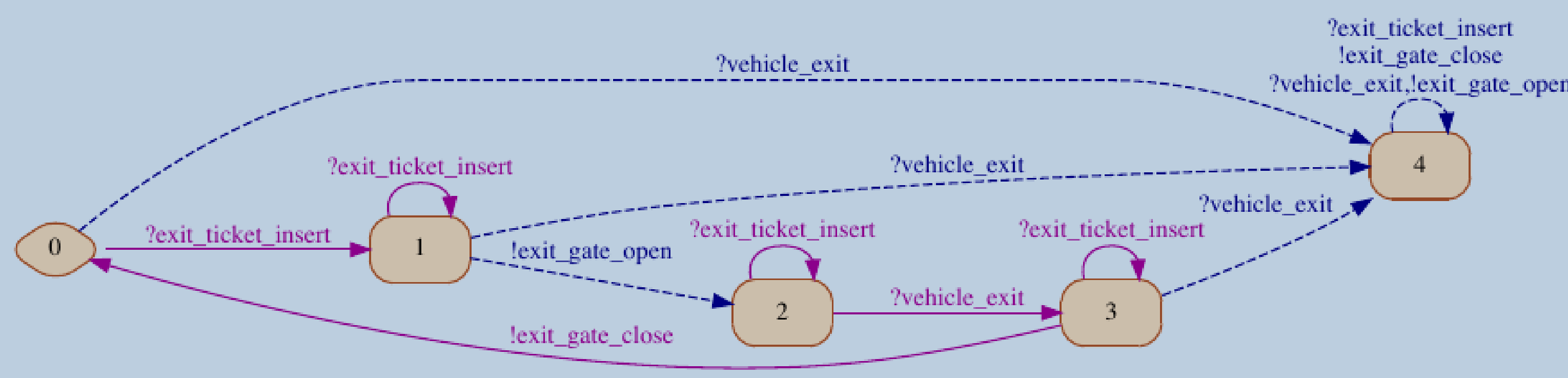


FIGURE 4: Automate modale pour la section "Gate"

Perspectives

1. Validation de la méthodologie par une **mise en œuvre complète**
2. Amélioration de la couverture des phénomènes linguistiques (ambiguïté, anaphore...)
3. Développement de l'**interaction** avec l'utilisateur
4. Apprentissage automatique des concepts par le système, et mise à jour interactive de l'ontologie de référence
5. Extension des interfaces modales pour prendre en compte les **propriétés numériques**

Quelques références

- [1] A. Benveniste, B. Caillaud, D. Nickovic, R. Passerone, J.-B. Raclet, P. Reinkemeier, A. Sangiovanni-Vincentelli, W. Damm, T. A. Henzinger, and K. G. Larsen. Contracts for system design. *FnT Electronic Design Automation*, 12, 2018.
- [2] M. Steedman. *The syntactic process*. MIT Press, 2000.
- [3] T. Kwiatkowski, L. Zettlemoyer, S. Goldwater, and M. Steedman. Inducing probabilistic ccg grammars from logical form with higher-order unification. In *Proc. of the CEMNLP. ACL*, 2010.
- [4] J. Bos. Open-domain semantic parsing with boxer. *Proc. of the NCCL*, 2015.
- [5] B. Caillaud. Mica: A Modal Interface Compositional Analysis Library, 2011.