
Laboratoire 1^{*†}

Raisonnement et systèmes à base de connaissances

Introduction

Un système à **base de connaissances** de type déductif permet de tirer des conclusions à partir de faits véridiques. Tout problème qui s'exprime en ces termes pourrait, en théorie, être résolu grâce à ce type de systèmes. La principale difficulté est que leur construction exige l'écriture d'un ensemble de règles qui guident, d'une certaine manière, la construction des conclusions. L'autre difficulté concerne le langage dans lequel sont exprimés les faits et les règles. Plus le langage est permissif, plus le type de **moteur d'inférence** est complexe. Par exemple, si l'expression des faits utilise un ensemble d'expressions booléennes complexes, le moteur d'inférence devra être un **solveur SAT**. Les **systèmes experts** sont plus contraints et c'est pourquoi ce laboratoire en fera usage.

Vous avez déjà trouvé, en classe, des exemples de systèmes déductifs. Prenons un exemple simple.

1. Zénon est mort entre 1h et 4h.
2. Daniel était au restaurant Le Québec à 8h.
3. Le mort a été trouvé au Blue Mountain.
4. Denis était Blue Mountain entre 18h et 20h.
5. Daniel était au Blue Mountain à 2h.

Ces cinq phrases constituent 5 faits de base. Les règles permettent de faire des **déductions**, donc d'ajouter des faits. Dans le cas précédent, il est clair que nous pouvons conclure un nouveau fait :

6. Daniel était sur les lieux du meurtre.

Le module déductif que vous devez concevoir devra résoudre une énigme de votre choix en respectant un ensemble de contraintes détaillées plus loin. Cette élaboration se fera selon une méthodologie itérative sur tous les plans. Prenons l'exemple précédent. Il est clair qu'une véritable enquête ne pourrait conclure de (1), (3), (5) donc (6) que Daniel est le coupable. Par contre, il devient un suspect et c'est tout ce dont nous avons besoin. Graduellement, vous pouvez ajouter d'autres éléments. Considérez les faits supplémentaires suivants :

7. Zénon a été tué avec un couteau.
8. Le couteau appartient à Daniel.

En supposant (encore une fois, nous devons faire abstraction d'une véritable enquête) qu'il n'y ait qu'un seul couteau dans cette histoire, on pourrait en conclure que Daniel est un suspect encore plus probable.

De manière plus générique, nos histoires mettent minimalement en jeu les **concepts** suivants et pouvant référer à des objets physiques, biologiques ou virtuels.:

LIEU, INSTRUMENT, ÉVÉNEMENT/ÉTAT/PROPRIÉTÉ/RELATION, PERSONNAGE, TEMPS

^{*} Originellement, ce laboratoire utilisait le jeu Clue. Cette version était alors fortement inspirée d'un énoncé proposé par Todd W. Neller, Zdravko Markov, Ingrid Russell et Dave Musicant (<http://cs.gettysburg.edu/~tneller/nsf/clue/index.html>). Les auteurs suggéraient alors d'utiliser un solveur SAT (ex. Zchaff), ce qui impliquait la traduction sous forme normale conjonctive des indices.

[†] Les expressions et les mots soulignés forment un ensemble d'éléments que vous devriez être capables de définir ou de vulgariser.

Objectif

Nous visons ici à construire un système rationnel capable de jouer à l'enquêteur. C'est donc dire que le système doit pouvoir effectuer des déductions à partir d'une base de connaissances. Deux conclusions sont doivent être déduites :

- Le QUI/QUOI : Quel personnage/objet est recherché?
- Le OÙ : Où est le personnage/objet recherché présentement?

Objectifs spécifiques

- Créer un monde imaginaire composé d'objets et de lieux, dans lequel évoluent des personnages dans le temps;
- Créer des banques d'énoncés exprimant des faits passés et présents, chacune permettant d'identifier un personnage et un lieu;
- Exprimer en Jess chacune des banques d'énoncés pour former des bases de connaissances initiales;
- Exprimer en Jess un ensemble de règles de déduction permettant d'identifier, pour chaque base de connaissances initiale, le personnage et le lieu où il se trouve;
- Documenter le système afin de pouvoir créer de nouvelles banques d'énoncés à volonté.

Description

Dans le cadre de ce travail, votre monde imaginaire est composé d'un nombre fini d'objets, de lieux, de personnages, d'événements et d'états. Chacun de ces éléments est représenté par un symbole défini par vous. Il n'est donc pas question ici d'intégrer toutes les variantes possibles permettant d'exprimer ces éléments en français!

Reprenons notre histoire de meurtre présentée en introduction. Voici le monde dans lequel les personnages évoluent :

- personnages : zenon, daniel et denis
- états/propriétés : is-dead, have, is-suspect
- relations : at-loc, at-time, from-t, to-t, instr¹
- évènement : meurtre
- objet : couteau

Et voici à quoi pourrait ressembler notre banque d'énoncés initiale en Jess :

(is-dead zenon)

(meurtre from-t 1 to-t 4)

(meurtre at-loc bluemountain)

(at-loc denis bluemountain from-t 18 to-t 20)

(at-loc daniel bluemountain at-time 2)

(meurtre instr couteau)

(have daniel couteau)

¹ Les abréviations présentées ici sont communes pour représenter certaines des relations sémantiques dans une phrase. On les appelle des rôles sémantiques, ou rôles thématiques ou événementiels selon les auteurs.

En utilisant un langage simple, nous pourrions ainsi en arriver à écrire l'ultime règle d'inférence (celle qui conclut au suspect) comme suit² :

```
(defrule VOICI-LE-SUSPECT
  (meurtre at-loc ?loc)
  (meurtre from-t ?t1 to-t ?t2)
  (meurtre instr ?inst)
  (at-loc ?s ?loc at-time ?t3)
  (test (>= ?t3 ?t1))
  (test (<= ?t3 ?t2))
  (have ?s ?inst)
  =>
  (assert (is-suspect ?s)))
```

Après le lancement du moteur d'inférence, un nouveau fait apparaîtra dans la base de connaissances :

```
(is-suspect daniel)
```

Évaluation

Les critères d'évaluation sont évolutifs et se basent d'abord sur ce que votre système peut faire ou non. Obligatoirement, votre système doit être basé sur la technique de construction des systèmes experts. Il n'est pas question ici d'essayer de coder une technique de recherche dans un espace d'états de manière détournée ou encore de dénaturer la notion de règle d'inférence par une utilisation procédurale sournoise. Vous devez appartenir aux « **neats** » et non aux « **scruffies** ».

Votre système (80%)

1. peut uniquement faire des déductions évidentes (du type de celle présentée plus haut)! 0
2. peut résoudre toutes vos énigmes par déduction : max 60/80
3. comporte un ensemble de règles particulièrement complexes³ : 80/80

Vos énigmes (10%)

4. sont présentées et commentées et fournies avec une trace d'exécution;
5. illustrent la complexité du système et permettent d'évaluer sa qualité.

Les critères d'excellence (10%)

6. Vous pouvez répondre aux questions de vos collègues, du chargé de laboratoire ou du professeur concernant votre système
7. Le monde que vous proposez est particulièrement créatif et complexe.

Les critères négatifs sont les suivants :

- Code non commenté : -[0..10]%
- Qualité du français de la documentation : -[0..10]%
- Utilisation procédurale non justifiée : -10%
- Non-respect des critères de remise (au niveau du dépôt et de la date limite) : -[0..5]%

La remise comprend :

- le fichier Jess commenté;
- les cinq (5) banques de faits exprimant vos énigmes;
- un document expliquant votre monde, le langage utilisé et les caractéristiques des cinq énigmes;
- une feuille synthétisant les caractéristiques de votre système sans en dévoiler les secrets! Soyez vendeurs! Ce dernier document servira à présenter, sur le site du cours, la synthèse des systèmes soumis par l'ensemble du groupe.

² Attention : ceci est un exemple simplifié.

³ Si le point 5 est déficient, vous serez appelé à mieux illustrer cette complexité (= défendre votre système) de manière orale.