Intelligence artificielle symbolique
Une douce introduction à la logique propositionnelle pour l'intelligence artificielle

Pierre Mercuriali

Résumé

Ce document est une douce introduction -non exhaustive !- à certains concepts importants en logique et en intelligence artificielle symbolique, en particulier en logique propositionnelle.

Table des matières

1 Introduction

1.1 Contenu du cours

Les symboles nous permettent de communiquer et de raisonner à propos de choses absentes. Si je veux vous dire que je pense que le comté, qui est un fromage de lait cru de vache à pâte pressée cuite produit dans la région de Franche-Comté, est délicieux, je n'ai pas besoin de vous montrer une vraie tranche de fromage, si tant est que vous sachiez ce que les symboles

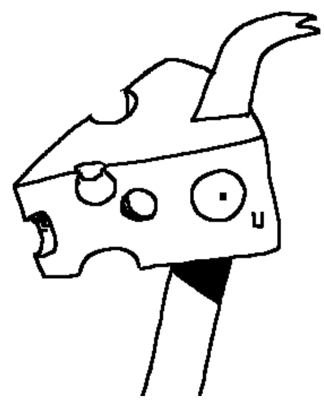
— `le`,
— `comté`
— `est`, et
— `délicieux`

signifient, et que vous sachiez ce qui se passe lorsque je combine ces symboles en la phrase

`le comté est délicieux`.

Il me suffit d'utiliser le symbole `comté`, écrit avec des lettres de l'alphabet ou prononcé avec des phonèmes, pour faire référence au comté réel, ou au concept général du comté. J'aurais également pu utiliser un croquis représentant un comté, ou une photo d'une tranche de comté.

Figure 1 – Un dessin de fromage.



Le symbole revêt donc une fonction *communicative* : je pense à une certaine chose, et je souhaite vous faire penser à cette même chose ; j'utilise donc, pour ce faire, un symbole spécifique, en espérant que nous partagions la même compréhension de ce symbole ou, en d'autres termes, en espérant que nous interprétions le symbole de la même manière.

Je peux également *raisonner* à partir et à propos de la connaissance que j'ai du comté, de moi-même, et, en général, du monde qui m'entoure et me contient. Par exemple, je sais que le comté est une chose délicieuse. Je sais également que j'aime les choses délicieuses. Pouvez-vous alors en déduire que j'aime le comté ?

Une approche possible à l'intelligence artificielle (IA) est de faire en sorte qu'une IA, qu'elle soit un robot, un logiciel de discussion automatique, un ordinateur ou une intelligence acentrique distribuée ayant pour but de détruire l'humanité, peut effectuer les différentes tâches que nous avons suggérées ci-dessus. Plus précisément,

— nous pouvons représenter notre connaissance du monde en utilisant des symboles ;
— nous pouvons utiliser ces symboles pour communiquer, pour partager cette connaissance du monde que nous avons, comme nous le faisons en ce moment ; et
— nous pouvons manipuler ces symboles nous-même, avec notre cerveau, pour déduire et découvrir des connaissances nouvelles à propos du monde.

Dans ce court document sur l'IA symbolique, nous allons découvrir comment les robots et les humains peuvent effectuer ces différentes tâches, à travers des exemples tirés de l'histoire, de la littérature, et de la recherche scientifique.

Premièrement, nous explorerons la notion de *symbole* et la définirons convenablement dans les Sections 2 et 3. Deuxièmement, nous allons découvrir comment manipuler et représenter la connaissance que l'on a du monde en utilisant des symboles ainsi qu'une *logique* ; plus précisément, nous allons utiliser le langage logique connu sous le nom de *logique propositionnelle* (LP) pour exprimer et, plus tard, déduire certaines connaissances sur le monde en utilisant les connaissance que l'on a déjà, dans la Section 4. Troisièmement, nous allons parler de la notion de vérité et d'interprétation de la vérité, et nous allons mentionner quelques pièges amusants et révélateurs dans lesquels nous (humains, et intelligences artificielles) pouvons tomber et que le raisonnement logique cherche à éviter dans la Section 5. Quatrièmement et finalement, nous ouvrons la discussion à d'autres logiques, différentes de la LP, et certaines motivations pour leur utilisation, dans la Section 6.

1.2 Objectifs d'apprentissage

Dans la Table 1, nous avons résumé les objectifs d'apprentissage principaux (OAs) que vous allez atteindre durant ce cours. Vous trouverez plus de détails sur ces OAs et des exemples d'exercices et de problèmes types dans l'Appendice A.

☞ `Pour aller plus loin`

Un rectangle informatif tel que celui-ci accompagnera le texte à la fin des sections pour vous inviter à chercher les notions et les thèmes qui y sont décrits et qui, malgré leur intérêt, ne tombent pas dans le cadre du cours.

Thème	Section	Objectif d'apprentissage	Termes importants
Présentation du sujet	1	OA1 : connaître les objectifs d'apprentissage du cours	
Qu'est-ce qu'un symbole ?	2	OA2 : être capable de donner des exemples de signes présents dans le monde qui vous entoure OA2' : connaître l'étymologie du mot 'symbole' et savoir la relier à son utilisation moderne	symbole, signe, IA explicable
Pourquoi utiliser des symboles et la logique pour représenter la connaissance que l'on a du monde ?	3	OA3 : expliquer certaines ambiguïtés de la langue naturelle et expliquer comment la logique peut aider à les lever OA3' : énoncer au moins deux raisons qui motivent l'utilisation de la logique	ambiguïté, syntaxe, sémantique, fondement des symboles raisonnement formel, raisonnement, inférence, déduction
Comment construire des formules de logique propositionnelle (LP) ?	4	OA4 : savoir reconnaître et construire des expressions bien formées en LP OA4' : construire une formule pour modéliser un problème réel OA4" : savoir lire une formule et la traduire en langue naturelle	définition récursive, syntaxe, symboles et noms des opérateurs usuels modélisation, fondement des variables
Interprétation(s) et vérité(s)	5	OA5 : déterminer si une formule est valide, satisfiable, ou pas, en construisant sa table de vérité	validité, satisfiabilité, table de vérité, interprétation
Limites de la LP. Y a-t-il d'autres logiques ?	6	OA6 : connaître d'autres types de valeurs de vérité OA6' : connaître certaines limites de la LP et quelle logique utiliser pour quel but.	logique multi-valuée, logique floue logique du premier ordre, modale, temporelle, floue, de description

Table 1 – Les différents Objectifs d'Apprentissage (OA) de ce cours.

2 Qu'est-ce qu'un symbole ?

2.1 Signes

Nous ne nous aventurerons pas trop profondément dans les subtilités des théories philosophiques des signes, symboles et références, et l'étude générale des symboles, aussi appelée *sémiotique*, mais si vous êtes curieux, je fournis des références dans le texte pour aller plus loin. Les symboles, aussi appelés signes par certains auteurs (mais bien distingués par d'autres), nous permettent de parler et de raisonner à propos de choses qui sont absentes, de choses abstraites, et de communiquer de l'information, simple ou complexe.

Les signes sont présents à travers le monde, dans les villes et au bord des routes, sous la forme de panneaux. A la vue d'un panneau octogonal rouge, vous savez qu'il vous faut stopper, même si le mot `STOP`, ou encore l'action même de s'arrêter, n'est ni écrite, ni visualisée, ni exprimée oralement. Parfois, les signes sont effectivement plus explicites : à la place d'un cercle vert lumineux et abstrait qui indiquent à un piéton qu'il lui est possible de traverser la route, certains panneaux lumineux représentent une figure en train de marcher.

2.2 Étymologie : symboles

Étymologiquement [1], le mot 'symbole' vient du grec ancien 'σύμβολον' (sumbolon) qui signifie

> 'comparer', 'mettre (des choses) ensemble'.

Le mot lui-même peut être séparé en 'σύν' (sun) qui signifie 'ensemble' et '' (ballo) qui signifie 'je jette', 'je place'. En grèce antique, ce mot avait un sens bien concret. Lors d'un accord, commercial ou autre, les deux partis cassaient une pièce de poterie en deux. Chacun gardait une pièce, qui fonctionnait alors comme une preuve d'identité lorsque les deux pièces étaient placées côte-à-côte pour reconstituer la pièce originale, à la manière d'un cadenas et de sa clé, ou du pied de Cendrillon et de son délicat mais dangereux chausson de verre. En latin, et dans le monde romain, le sens devint plus général :

> 'marque ou signe comme preuve d'identification', voire même 'une preuve d'identité',

de manière similaire aux pièces d'identités qui, dans la société actuelle, font référence à la personne physique et réelle elle-même (ou son équivalent légal – à quelle profondeur de représentations sommes-nous) ? Plus tard encore, le sens évolue en

> 'un fait naturel ou un objet qui évoque par sa forme ou sa nature une association d'idées avec quelque chose d'abstrait ou d'absent'

et, ainsi, nous retrouvons l'idée que j'ai présentée ci-haut : la représentation de l'abstrait, ou de ce qui est absent.

En anglais, le sens spécifique de "quelque chose qui représente quelque chose d'autre" apparaît dans le *Faerie Queene* [2] (1590) d'Edmund Spenser. Dans cette partie de l'histoire, l'un des héros, Sire Guyon, chevalier de la Tempérance (la valeur chrétienne similaire à la valeur moderne du *contrôle de soi* et de la *résistance à la tentation*) rencontre une femme, Amavia, qui est sur le point de se suicider, car son conjoint a été empoisonné par la sorcière Acrasia. Elle enfonce un couteau dans sa poitrine et succombe à ses blessures dans les bras du chevalier. Celui-ci se retrouve seul avec le nouveau-né d'Amavia. Il essaye de laver l'enfant, maculé du sang de sa propre mère, mais sans succès. Le sang, ici, agit comme un *symbole* de l'innocence, de la chasteté, et de la dévotion de sa mère, comme un rappel de son suicide, et comme un appel à la vengeance.

From thence it comes, that this babes bloudy hand	Ainsi les mains sanglantes de cet enfant
May not be clensd with water of this well :	Ne peuvent par cette eau être lavées :
Ne certes Sir striue you it to withstand,	Ainsi, Sire, efforcez-vous d'en supporter l'état,
But let them still be bloudy, as befell,	Mais laissez-les toujours ainsi ensanglantées,
That they his mothers innocence may tell,	Afin qu'elles puissent témoigner de l'innocence de sa mère
As she bequeathd in her last testament ;	Comme elle l'a laissé dans ce testament ;
That as a sacred Symbole it may dwell	Afin que reste un *symbole* sacré
In her sonnes flesh, to minde reuengement,	Dans la chair de cet enfant, pour attiser la vengeance,
And be for all chast Dames an endlesse moniment.	Et être pour toute chaste dame un monument éternel.

Edmund Spenser, *Faerie Queene*, Livre II, Chant II (Traduction et adaptation PM)

Le mot 'symbole' apparaît aussi tôt que dans le *Gargantua* de Rabelais, publié en 1534. Dans le prologue, Rabelais encourage le lecteur à ne pas juger trop hâtivement un livre par son titre et ses thème joyeux et, en

1. https://www.cnrtl.fr/definition/symbole
2. http://www.luminarium.org/renascence-editions/fqintro.html ; voir également https://www.gutenberg.org/cache/epub/6930/pg6930.txt

apparence, folâtres. Il utilise la métaphore d'un chien, 'la bête du monde la plus philosophe' selon Platon, qui ronge son os avec fort travail pour en extraire la moelle.

A l'exemple d'icelluy vous convient estre saiges pour fleurer, sentir, et estimer ces beaulx livres de haulte gresse, legiers au prochaz : et hardiz à la rencontre ? Puis par curieuse leçon, et meditation frequente rompre l'os, et sugcer la sustantifique mouelle. C'est à dire : ce que j'entends par ces symboles Pythagoricques avecques espoir certain d'estre faictz escors et preux à ladicte lecture. Car en icelle bien aultre goust trouverez, et doctrine plus absconce, laquelle vous revelera de tres haultz sacremens et mysteres horrificques, tant en ce que concerne nostre religion, que aussi l'estat politicq et vie oeconomicque.

Rabelais, *Gargantua*, Prologue de l'auteur, p. 34

A l'exemple de ce chien il vous convient d'avoir la sagesse de flairer, sentir et estimer ces beaux livres de haute graisse, légers à l'approche et hardis à la rencontre ; puis par une lecture minutieuse et une méditation assidue, de rompre l'os et sucer la substantifique moelle - c'est ce que j'entends par ces symboles pythagoriques, avec l'espoir certain de vous rendre sages et valeureux par ladite lecture. Car en elle vous trouverez un bien autre goût, et un savoir plus secret, lequel vous révélera de très sacrées énigmes et des mystères horrifiques, en ce qui concerne tant notre religion, que l'état politique et la vie économique.
(Translation Myriam Marrache-Gouraud)

Que sont ces 'symboles Pythagoricques' ? Marrache-Gouraud nous explique que ce sont des 'signes ou mots énigmatiques, dont le sens ne se révèle pas à la première lecture'.

Comme vous l'avez sans doute déjà remarqué, ce que j'utilise en ce moment pour communiquer peuvent être, et ont été, considéré comme des symboles. Lorsque j'utilise des mots en français ou en anglais, je les utilise comme des référents à des sens abstraits ou non, en supposant que nous sommes assez d'accord sur le sens d'un nombre suffisant de mots pour communiquer de manière efficace.

2.3 Communication

Nous entrerons en détails dans le lien entre les symboles et les sens dans la Section 5. Pour l'instant, je voudrais mettre l'accent sur un bénéfice de l'utilisation de symboles et de logiques en IA. Le théologiste et philosophe Augustin d'Hippo (354 - 430), qui vécut en Algérie actuelle, écrit, à propos des signes :

> *"Il n'y a aucune raison de signer, c'est-à-dire de donner des signes, si ce n'est pour communiquer dans l'esprit d'un autre ce que le donneur de signe a dans son propre esprit."*
> Augustine, De doctr. chr. II 3, 1963, 34 : 17–20

S'il nous est possible de nous mettre d'accord, avec les robots, sur le sens de certains signes, alors nous pouvons communiquer avec eux, leur exprimer nos intentions, mais aussi les comprendre et savoir ce qu'ils font avec ces signes. Par exemple, observer la manière selon laquelle ils manipulent les signes peuvent nous informer sur leur raisonnement. Cette problématique est liée au courant d'IA explicable ('explainable AI' en anglais), qui prône l'importance de savoir comment une intelligence artificielle atteint un certain but, ce qui est particulièrement important lorsque l'on cherche a comprendre quand un robot échoue et comment réparer ou prévenir les erreurs.

☞ Pour aller plus loin

— Sémiotique médiévale, sur https://plato.stanford.edu/entries/semiotics-medieval/. Le SEP (Stanford Encyclopedia of Philosophy - 'encyclopédie de philosophie de Stanford') contient d'excellents articles sur des domaines connexes à la logique, au langage, à l'intelligence, et à l'intelligence artificielle, pour ne citer que quelques thèmes.
— IA explicable
— Fiabilité de l'IA

3 Pourquoi utiliser des symboles et la logique pour représenter la connaissance que l'on a du monde ?

Dans le contexte de ce cours, le terme de 'logique' signifie, sans rentrer dans les détails pour l'instant, un langage, c'est-à-dire un ensemble de phrases, auquel on ajoute une manière de les comprendre, c'est-à-dire une interprétation, un sens, ainsi qu'une manière codifiée de manipuler et de mettre en relation ces phrases. Nous parlerons de deux principales motivations pour l'utilisation de la logique :

1. l'élimination d'ambiguïtés, et
2. l'étude de structures de raisonnement formelles.

Lorsque nous utilisons une langue naturelle telle que le français, l'anglais, le néerlandais, le polonais, le grec... nous tombons parfois sur des ambiguïtés particulières.

Dans cette section, je vais illustrer l'ambiguïté par l'exemple sinistre des têtes de bronze. Dans les mythes de l'époque moderne, c'est-à-dire du XVème au XIXème siècle environ, les têtes de bronze étaient des bustes humanoïdes coulés en bronze auxquels étaient attribués des pouvoirs magiques ou inexpliqués de prédiction et de vérité.

3.1 Ambiguïté

Gerbert d'Aurillac (c.946 - 12 May 1003) était un lettré français qui devint le Pape Sylvestre II. C'était un scientifique accompli, qui fabriqua la première horloge mécanique et un orgue musical hydraulique. Il étudia et publia des ouvrages sur l'arithmétique, la géométrie, l'astronomie, la musique, mais également la grammaire, la logique, et la rhétorique. Parmi les légendes à propos du personnage, on raconte qu'il aurait construit une de ces *têtes de bronze*, une machine astrologique qui pouvait lui apprendre la vérité grâce à la lecture des astres, après avoir volé un grimoire magique à un philosophe sarrasin en espagne et fait un pacte avec le diable pour devenir pape. La machine causera éventuellement sa perte.

> *I have inserted this narrative of the Aquitanian to the intent that what is reported of Gerbert should not seem wonderful to any person ; which is, that he cast, for his own purposes, the head of a statue, by a certain inspection of the stars when all the planets were about to begin their courses, which spake not unless spoken to, but then pronounced the truth, either in the affirmative or negative. For instance, when Gerbert would say, "Shall I be pope ?" the statue would reply "Yes." "Am I to die, ere I sing mass at Jerusalem ?" "No." They relate, that he was so much deceived by this ambiguity, that he thought nothing of repentance : for when he would think of going to Jerusalem, to accelerate his own death ? Nor did he foresee that at Rome there is a church called Jerusalem, that is, "the vision of peace," because whoever flies thither finds safety, whatsoever crime he may be guilty of. We have heard, that this was called an asylum in the very infancy of the city, because Romulus, to increase the number of his subjects, had appointed it to be a refuge for the guilty of every description. The pope sings mass there on three Sundays, which are called "The station at Jerusalem." Wherefore upon one of those days Gerbert, preparing himself for mass, was suddenly struck with sickness ; which increased so that he took to his bed : and consulting his statue, he became convinced of his delusion and his approaching death. Calling, therefore, the cardinals together, he lamented his crimes for a long space of time. They, being struck with sudden fear were unable to make any reply, whereupon he began to rave, and losing his reason through excess of pain, commanded himself to be maimed, and cast forth piecemeal, saying "Let him have the service of my limbs, who before sought their homage ; for my mind never consented to that abominable oath."*
>
> William of Malmesbury, *Chronique des rois d'angleterre.* Traduction J.A. Giles.

Gerbert, après avoir obtenu le grimoire, coule le bronze pour faire sa statue. Celle-ci possède les propriétés suivantes, simples, mais extrêmement puissantes.

1. La statue ne parle pas sans qu'on lui parle (ce n'est pas la propriété la plus puissante, mais peut-être la plus utile si Gerbert ne veut pas être dérangé).
2. La statue prononce la vérité en répondant à une question par `OUI` ou `NON`.

Ainsi, Gerbert demande à la statue :

'Vais-je devenir pape' ?

ce à quoi la statue répond :

'`OUI`.'

Satisfait, Gerbert pose une seconde question :

'Vais-je mourir avant de chanter la messe à Jérusalem ?'

ce à quoi la statue répond :

`'NON.'`

Gerbert décide alors de ne jamais visiter Jérusalem, pensant qu'il s'agit de la ville du Moyen-Orient. Selon lui, s'il ne chante jamais la messe là-bas, il ne mourra pas ! Il devient le pape Sylvestre II, et accomplit ainsi la première prophétie de la tête de bronze. Satisfait de ses fonctions, il chante la messe dans les églises de Rome. Mais, à Rome, il existe une certaine église que l'on appelle Sainte Marie de Jérusalem, et qui est tout simplement appelée 'Jérusalem' par la population locale. Gerbert, sans le savoir, s'y rend et, alors qu'il se prépare à y chanter la messe, il tombe gravement malade. Fou de douleur, il devient tellement sûr de sa mort prochaine qu'il demande à ses cardinaux de le découper en morceaux et de lancer ses membres aux quatre coins de Rome en guise de rédemption.

Il y a clairement un problème. Gerbert ne voulait certainement pas mourir ainsi, et, pourtant, il fut réassuré par les réponses de la tête de bronze. Pour analyser ce problème de communication, nous pouvons réécrire le dialogue de manière un peu plus formelle, dans la Table 2.

Le problème principal est celui du *fondement de la variable* `Gerbert_est_à_Jérusalem`, c'est-à-dire la manière dont le symbole `Gerbert_est_à_Jérusalem` est lié au monde réel. Gerbert essaie de rendre cette variable "fausse", en n'allant *pas* à Jérusalem, auquel cas il ne mourra pas, même s'il chante la messe – ce qu'il est obligé de faire, en tant que pape. Malheureusement, il ne se rend pas compte que le symbole `Jérusalem`, qu'il utilise oralement pour parler avec la tête de bronze, fait aussi référence à cette petite église à Rome dans laquelle il va chanter. La tête de bronze n'a pas menti, et le raisonnement de Gerbert est solide : mais la tête de bronze faisait référence à la petite église romaine appelée Jérusalem, ou, peut-être, à tout lieu appelé Jérusalem. Comment auriez-vous résolu ce problème, si vous aviez été à la place de Gerbert (et plus attentif que lui) ?

3.2 L'étude de structures de raisonnement formelles

L'étude précise de l'histoire de Gerbert nous permet d'exhiber précisément le problème qui mena Gerbert à sa perte et, peut-être, d'en trouver une solution possible. L'étude du raisonnement humain de manière formelle est une autre motivation pour l'étude de la logique formelle en intelligence artificielle. Examinons ces phrases, toutes très similaires.

1. Gerbert meurt s'il chante la messe à Jérusalem. Gerbert chante la messe à Jérusalem. Donc, Gerbert meurt.
2. Socrate est un homme. Tous les hommes sont mortels. Donc, Socrate est mortel (syllogisme très connu !)
3. Platon est un philosophe. Si quelqu'un est un philosophe, alors ils sont forts ! Par conséquent, Platon est fort.
4. Socrate est connu. Une personne connue est aussi une personne célèbre. Donc, Socrate est célèbre.

Toutes ces phrases suivent la même structure, et mettent en jeu le même type de raisonnement, suggéré par l'utilisation des termes 'donc' et 'par conséquent'. La structure est présente quels que soient les autres symboles utilisés : 'homme', 'philosophe', 'Platon', 'fort'... Aristote, le philosophe grec, est un des premiers à concevoir l'idée d'étudier la structure générale de ces raisonnements, qu'il appelait *syllogismes*. C'est là le but de la logique formelle : étudier la *forme*, la structure de ces raisonnements. Par la suite, les philosophes et logiciens tels que Leibniz, Frege, Péano et Russell ont étudié la notion de vérité et la réalité du monde en utilisant des langages artificiels comme la logique pour représenter ces structures de raisonnement.

Dans la section suivante, nous allons définir un tel langage artificiel : la logique propositionnelle.

Formule logique	Réponse
`Gerbert_sera_pape`	`OUI`
SI `Gerbert_chante_la_messe` ET `Gerbert_est_à_Jérusalem`, ALORS `Gerbert_meurt`	`NON`

Table 2 – Le dialogue de Gerbert, un peu plus formel.

☞ Pour aller plus loin

- L'hubris des humains qui se frottent aux absolus - tels que la vérité absolue de la tête de bronze) est un thème récurrent dans la littérature sur l'intelligence artificielle. Lorsque les têtes de bronze ne refusent pas tout simplement de marcher à cause des caractéristiques bien humaines de leur inventeur, tels que Bacon qui s'endort à un point crucial de la fabrication, les têtes de bronze mènent leurs inventeurs à leur perte, à cause d'une erreur de jugement de leur part, comme avec Gerbert, ou même par intention malfaisante. Pour plus d'information, lire *The Persistent Peril of the Artificial Slave* de Kévin LaGrandeur.
- le Loglan et son successeur, le Lojban, sont des conlangues (langues construites) qui ont pour but d'éliminer les ambiguïtés de la langue naturelle.
- Loglan and its successor Lojban are conlangs (constructed languages) that aim at eliminating ambiguities in natural language. `https://omniglot.com/writing/lojban.htm`, `https://mw.lojban.org/index.php?title=Lojban&setlang=en-US`.
- Phrases en Garden path, donkey sentences, paraproskodia, sont des exemples d'effets de langue qui jouent sur les ambiguïtés.
- Les écrits de William de Mamlesbury à propos de Gerbert ont été critiqués comme étant biaisés contre Gerbert ; voir `https://www.icysedgwick.com/brazen-head/` (un excellent article sur les têtes de bronze).
- Pourquoi ai-je mentionné les philosophes forts ? Socrate, dans un des dialogues de Xénophon (*Memorabilia* 3.12), mentionne l'importance de bien s'occuper de son corps, pour la cité – car, d'après lui, il est important d'avoir un corps en bonne santé pour défendre sa propre cité et partir en guerre pour elle.

4 Comment construire des formules de logique propositionnelle (LP)

En langue naturelle, comme le français ou l'anglais, et, sans se perdre dans les détails complexes de la grammaire, nous construisons des phrases en utilisant des mots. De manière très simplifiée, ces mots peuvent être vus comme des briques de construction, et les règles de grammaire indiquent comment les placer pour construire des morceaux de phrases ou des phrases entières. Par exemple, en face d'un nom tel que 'chat' ou 'robot', il est commun de trouver un article, tel que 'le chat' ou 'un robot'. En logique, nous procédons de la même manière. Vous avez à votre disposition des briques et du mortier, qui sont des symboles préalablement choisis, ainsi que les règles qui vous permettent de les combiner. Dans cette section, je vais vous montrer les briques et les règles qui vous permettrons de construire des formules de logique propositionnelle.

Premièrement, considérons notre ensemble de briques : un ensemble de symboles Var, qui représentent des concepts ou des individus que l'on veut modéliser. Le plus souvent, ce sont des noms tels que '`femme`', '`homme`', '`Platon`', '`fort`', '`philosophe`', '`AdaLovelace`', ou même les symboles apparemment plus abstraits '`X`', '`Y`', '`Z`', '`A`', '`B`'. Nous appelons ces symboles des *variables*, ou encore les *variables propositionnelles*.

Deuxièmement, considérons notre mortier, qui sont un autre ensemble de symboles que l'on appelle *connecteurs logiques* ou *opérateurs logiques* :

$$Op = \{\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow, (,)\}$$

Ceux-ci portent chacun un nom particulier que je donne dans la Table 3.

Personnellement, je confonds toujours la disjonction et la conjonction, comme la gauche et la droite, d'ailleurs. Un moyen de se souvenir de la différence est de remarquer que *con*, en latin et en italien, signifie 'avec', 'ensemble', qui est un cognate [3] de 'et'.

Les parenthèses ont une fonction syntactique : elles permettent d'éliminer l'ambiguïté de lecture dûe à la précédence des opérateurs, comme en mathématique. Sans aucune règle sur la précédence des opérateurs $+$ and $\times$, c'est-à-dire sans savoir quel opérateur doit être lu en premier, l'expression $1+2\times3$ est ambiguë, mais l'expression $(1+2)\times3$ ne l'est pas. Elle peut être évaluée sans question, en commençant par évaluer la partie entre parenthèses. De manière similaire, la formule logique $A \wedge B \vee C$ est ambiguë, mais la formule $(A \wedge B) \vee C$ ne l'est pas. En logique propositionnelle, les opérateurs $\Rightarrow$ et $\Leftrightarrow$ ont précédence sur $\wedge$ et $\vee$.

Finalement, considérons un ensemble de symboles de *constantes* $\{V, F\}$, que l'on lit 'Vrai' et 'Faux', respectivement. Les ensembles Op, Var, et les constantes $\{T, F\}$ sont disjoints deux-à-deux (pour ne pas tout mélanger et confondre une variable avec un connecteur, par exemple).

Nous avons maintenant tous les ingrédients pour construire nos formules ! Étant donnés les ensembles Op, Var, and $\{V, F\}$, l'*ensemble des formules propositionnelles* est défini inductivement comme suit.

— Les constantes V et F sont des formules (atomiques). Elles sont appelées 'atomiques' parce qu'elles forment les plus petits composants des formules et ne peuvent donc pas être divisées.

— Les variables propositionnelles sont aussi des formules atomiques.

— Si f et g sont des formules, atomiques ou pas, alors $\neg f$, $f \wedge g$, $f \vee g$, $f \Rightarrow g$, $f \Leftrightarrow g$, et (f) sont aussi des formules.

Par exemple, l'expression

$$\texttt{Gerbert\_chante} \Rightarrow \texttt{Gerbert\_meurt}$$

3. C'est-à-dire de sens similaire.

Connecteur	Comment le lire ?	Nom	Exemple
$\neg$A	'non A'	négation	Socrate n'est pas jeune.
A $\wedge$ B	'A et B'	conjonction	Carol Karp et Susanne Langer sont des logiciennes.
A $\vee$ B	'A ou B'	disjonction	'Êtes-vous une philosophe ou une logicienne ?' 'Oui.'
A $\Rightarrow$ B	'A entraîne B', ou 'si A, alors B'	implication	Si Gerbert chante, alors il meurt.
A $\Leftrightarrow$ B	'A si et seulement si B', ou 'A est équivalent à B'	équivalence	Une tête de bronze est une tête faite en bronze (quelque chose est une tête de bronze si et seulement si c'est une tête, et elle est faite de bronze).

Table 3 – Connecteurs principaux en logique propositionnelle.

est une formule de LP, où $\{\texttt{Gerbert\_chante}, \texttt{Gerbert\_meurt}\}$ est l'ensemble des variables propositionnelles qui apparaissent dans l'expression. L'expression

$$\texttt{Gerbert\_chante\_la\_messe} \land \texttt{Gerbert\_est\_à\_Jérusalem} \Rightarrow \texttt{Gerbert\_meurt}$$

est aussi une formule de LP. A titre d'exercice, essayez de reformuler ces formules en la langue naturelle de votre choix (français, anglais, italien, etc.). Remarquez qu'il existe plusieurs réponses à cet exercice. Il n'y a pas une correspondance exacte entre la logique et la langue naturelle, ou entre la logique et le monde qui nous entoure. Ce problème est lié à celui de la *modélisation* : comment trouver une approximation du monde qui nous permette d'obtenir des explications ou des prédictions pertinentes ?

Finalement, l'expression

$$(\texttt{Gerbert\_chante\_la\_messe} \land)\texttt{Gerbert\_est\_à\_Jérusalem} \Rightarrow \neg\neg\texttt{Gerbert\_meurt}$$

n'est *pas* une formule de LP. Pouvez-vous expliquer pourquoi ?

☞ Pour aller plus loin

— Les chapitres 7, 8, and 10 de Russell SJ, Norvig P. Artificial Intelligence : A Modern Approach. (Fourth edition, 2021) contiennent une excellente introduction à la logique propositionnelle.

5 Interprétation(s) et vérité(s)

Pour l'instant, nous avons examiné des formules de PL faites de combinaisons de symboles, construites de manière presque mécanique en suivant un petit ensemble de règles de construction. Comment mettre en relation ces formules avec une situation, réelle ou fictive, que l'on voudrait décrire en utilisant la logique ?

5.1 Interprétations

Considérons le cas de Gerbert chantant. A un certain moment, nous considérons (ce qui est déjà une décision de modélisation !) que Gerbert est soit en train de chanter, soit qu'il n'est pas en train de chanter. Nous pouvons capturer cette décision de modélisation en utilisant la simple variable propositionnelle

$$\texttt{Gerbert\_chante}.$$

En elle-même, elle n'en dit pas plus sur l'état du monde : elle est juste un symbol. Mais nous avons vu précédemment que les symboles peuvent transporter du *sens*. Dans ce cas, les deux sens possibles du symbole sont les suivants. Si, effectivement, Gerbert est en train de chanter, nous disons que la variable `Gerbert_chante` is VRAIE ; sinon, si Gerbert n'est pas en train de chanter, nous disons que la variable `Gerbert_chante` is FAUSSE.

En LP, nous pouvons définir la notion de vérité, de sens, ou de contenu sémantique d'une formule, grâve à ce que l'on appelle *l'interprétation* d'une formule. L'interprétation est très proche de ce que nous faisons déjà quand nous communiquons : nous essayons de comprendre ce que notre interlocutrice nous dit, en attribuant un sens aux phonèmes qu'elle prononce, aux mouvements de son corps, aux expressions de visage !...

En LP, il existe deux valeurs de vérité : 'vrai', et 'faux', symbolisés par les symboles V et F, respectivement. Remarquez que j'utilise les mêmes symboles pour les constantes V et F dans nos formules ; c'est un abus de notation, car dans un cas les symboles sont utilisés dans un contexte syntaxique (les formules) et dans l'autre dans un contexte sémantique (l'interprétation) ; mais, puisqu'il n'y a pas d'ambiguïté, cet abus de notation est acceptable...

Une *interprétation* $\mathcal{I}$ est une fonction $\mathcal{I} : Var \rightarrow \{V, F\}$ qui à chaque variable propositionnelle assigne une valeur de vérité : soit V, soit F.

Dans l'exemple précédent, si `pape` est une variable qui fonctionne, dans notre modélisation, comme un symbole qui exprime le fait que Gerbert est pape ou pas, alors nous pouvons indiquer sa valeur de vérité explicitement grâce à une interprétation. Si, dans notre situation, Gerbert n'est pas encore pape, alors nous écrirons

$$\mathcal{I}(\texttt{pape}) = F$$

et lierons ainsi le symbole à la situation correspondante. Notez que chaque situation différente peut correspondre à une interprétation différente. Dans ce cas, $\mathcal{I}$ correspond à une situation dans laquelle Gerbert n'est pas encore pape ; mais nous aurions pu définir une autre interprétation $\mathcal{I}'$ qui correspond à une situation dans laquelle Gerbert est pape :

$$\mathcal{I}(\texttt{pape}) = V.$$

Ce n'est pas fini ! Après les variables, nous pouvons aussi interpréter des *formules plus complexes*, et pas seulement des formules atomiques. Les formules qui ne sont pas atomiques contiennent des connecteurs logiques. Nous devons donc définir la manière selon laquelle ces connecteurs se comportent sémantiquement. Nous avons déjà une compréhension naturelle du comportement du 'et' et du 'ou', par exemple. Grâce à la notion d'interprétation, nous pouvons définir cette compréhension naturelle.

Par exemple, pour calculer la valeur de vérité de la formule

$$\texttt{sera\_pape} \wedge \texttt{mourra},$$

étant donnée une interprétation $\mathcal{I}$, nous avons besoin de connaître les valeurs de vérités $\mathcal{I}(\texttt{sera\_pape})$ et $\mathcal{I}(\texttt{mourra})$ des variables propositionnelles `sera_pape` et `mourra`, et nous avons besoin de connaître le comportement sémantique du connecteur $\wedge$. Ici, $\mathcal{I}(\wedge)$ est une *function* à deux entrées. Supposons en effet que dans l'histoire $\mathcal{I}(\texttt{sera\_pape}) = V$ et que $\mathcal{I}(\texttt{mourra}) = T$. Alors, la valeur de vérité de la formule est

$$\begin{aligned}\mathcal{I}(\texttt{sera\_pape} \wedge \texttt{mourra}) &= \mathcal{I}(\wedge)(\mathcal{I}(\texttt{sera\_pape}), \mathcal{I}(\texttt{mourra})))\\ &= \mathcal{I}(\wedge)(V, V)\\ &= T,\end{aligned}$$

ce que Gerbert n'a pas pu conclure.

La dernière étape du calcul ci-dessus a été possible par définition de la fonction $\mathcal{I}(\wedge)$. Pour définir nos connecteurs $\{\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow\}$, nous pouvons en décrire le comportement de manière exhaustive, c'est-à-dire, donner la valeur qu'ils prennent, une fois interprétés, pour toute valeur d'entrée. On appelle ce procédé la construction d'une '*table de vérité*'. Par exemple, pour définir la négation $\neg$ (étant donnée une interprétation), nous donnons la valeur de l'interprétation de la formule $\neg$`A` pour toute valeur possible de `A` : si `A` est (interprétée comme) vraie, alors $\neg$`A` est (interprétée comme) fausse ; si `A` est fausse, alors $\neg$`A` est vraie.

Dans la Table 4, je donne la table de vérité de tous nos connecteurs. Comme vous pouvez le constater, les deux premières colonnes prennent toutes les valeurs de vérité possibles des variables `A` et `B`.

Remarquez que certaines interprétations ne correspondent pas nécessairement à l'interprétation habituelle du terme en langue naturelle. En français, lorsque nous utilisons le mot 'ou', nous l'utilisons typiquement comme un *ou exclusif*, comme dans la phrase

'Veux-tu une glace à la noix de coco `ou` une glace au chocolat ?'

En logique, nous utilisons plutôt le *ou inclusif* ; 'A ou B' est alors interprété comme vrai si au moins une des variables est interprétée comme vraie, pour le plus grand plaisir des étudiants logiciens facétieux qui répondront alors 'oui' à de telles questions.

Notez qu'il est parfaitement possible de définir ses propres connecteurs. Certains domaines de logique appliquée, comme par exemple la théorie des circuits Booléen, utilisent le 'ou exclusif'. Dans ce cours, nous utilisons les connecteurs définis dans la Table 4 car ils sont souvent utilisés dans la littérature logique moderne.

En guise d'exemple final, interprétons la valeur de vérité d'une autre formule en utilisant les définitions de la Table 4. Considérons la formule suivante.

$$p := (\texttt{gerbert\_chante} \wedge \texttt{gerbert\_à\_Jérusalem}) \Rightarrow \neg\texttt{gerbert\_meurt}$$

ainsi que l'interprétation suivante :

$$\mathcal{I} = \{\texttt{gerbert\_chante} : F, \texttt{gerbert\_à\_Jérusalem} : F, \texttt{gerbert\_meurt} : V\}.$$

(Le symbole ' :=' indique que je définis le symbole p comme un symbole qui fait référence à la formule).

Nous calculons que :

$$\begin{aligned}
\mathcal{I}(p) &= \mathcal{I}((\texttt{gerbert_chante} \wedge \texttt{gerbert_à_Jérusalem}) \Rightarrow \neg\texttt{gerbert_meurt}) \\
&= (\mathcal{I}(\texttt{gerbert_chante})\mathcal{I}(\wedge)\mathcal{I}(\texttt{gerbert_à_Jérusalem}))\mathcal{I}(\Rightarrow)\mathcal{I}(\neg)\mathcal{I}(\texttt{gerbert_meurt}) \\
&= (F\mathcal{I}(\wedge)F)\mathcal{I}(\Rightarrow)\mathcal{I}(\neg)T && \text{par définition de } \mathcal{I} \\
&= F\mathcal{I}(\Rightarrow)F && \text{par la Table 4} \\
&= F\mathcal{I}(\Rightarrow)F && \text{par la Table 4} \\
&= V.
\end{aligned}$$

Considérez attentivement ce résultat. Est-il surprenant ? D'où provient la surprise ? Pour plus d'information, voir l'encart à la fin de la section. Donnez une explication de la formule p en langue naturelle, ainsi que de son interprétation.

Remarque à propos de la notation Il existe plusieurs manières d'écrire les fonctions $\mathcal{I}(a)$, étant donné un connecteur logique a. Dans les calculs ci-dessus, j'ai choisis de conserver la notation infixe, comme par exemple $F\mathcal{I}(\wedge)T$, pour garder la même syntaxe que pour les formules de LP. Toutefois, il est parfaitement acceptable d'utiliser une notation fonctionnelle telle que $\mathcal{I}(\wedge)(F, V)$; tout dépend de ce qui est le plus aisé, lisible, et qui empêche le mieux les ambiguïtés étant donnés les autres choix syntactiques que l'on fait. Certains chercheurs utilisent même la notation polonaise, comme $\mathcal{I}(\wedge)FV$, voire même la notation polonaise inversée $FV\mathcal{I}(\wedge)$ (vous souvenez-vous des calculatrices ?) !

5.2 Équivalences dans d'autres domaines

Les opérations logique de disjonction, conjonction, négation... ont des équivalents dans d'autres domaines.

A	B	(A)	¬A	A ∧ B	A ∨ B	A ⇒ B	A ⇔ B
V	V	V	F	V	V	V	V
V	F	V	F	F	V	F	F
F	V	F	V	F	V	V	F
F	F	F	V	F	F	V	V

Table 4 – Définition du comportement sémantique de nos connecteurs.

Théorie des ensembles naïve Cette théorie est appelée 'naïve' non pas parce que ses concepts sont triviaux et simples, mais parce que sa conceptualisation ne nécessite pas d'adopter une position ontologique (c'est-à-dire à propos des choses, notre conception du monde) qui amène aux paradoxes fondamentaux de la théorie des ensembles. Nous pourrions en effet éviter le vocabulaire ensembliste et ne parler que d'objets au pluriel sans avoir de problèmes, mais pour simplifier la présentation du sujet, nous considérerons un ensemble de manière très concrète : comme un sac d'objets.

Tableau récapitulatif : intersection, union, négation, diagrammes de Venn.

Circuits

5.3 Satisfiabilité

Si, pour une formule donnée, il est possible de trouver une interprétation qui la rend vraie, alors on dit qu'elle est *satisfiable*. Si, pour une formule donnée, n'importe quelle interprétation la rend vraie, alors elle est *valide*. On dit aussi que c'est une *tautologie*[4]. Parfois, la formule atomique 'V' est appelée une tautologie pour cette raison !

En général, il est très coûteux en termes de calcul, c'est-à-dire qu'il faut beaucoup de temps et de mémoire de calcul (pour une certaine définition de 'beaucoup' qui est en dehors du cadre de ce cours !) pour vérifier si une formule est satisfiable ou pas.

Les problèmes 'SAT', et le terme de 'SAT solving' font référence au domaine de recherche qui concerne l'étude formelle du problème algorithmique qui consiste à déterminer si une formule est satisfiable ou non. Étonnamment, des problèmes très concrets peuvent être modélisés en termes de problème de satisfiabilité. Par exemple,

- le problème du voyageur de commerce : comment visiter un graphe (des villes, reliées entre elles par des routes) en un temps minimal ?
- vérification de programmes : quand est-ce qu'un programme s'arrête ? Pour quelles raisons ?
- contrôle du trafic aérien ;
- programmation de tournois sportifs...

☞ Pour aller plus loin

- N'est-il pas surprenant que la formule 'Faux implique Faux' soit interprété comme Vrai, et que, plus généralement, 'Faux implique n'importe quoi' soit toujours interprété comme Vrai, quelle que soit la valeur de vérité de ce 'n'importe quoi' ? C'est une propriété de l''implication matérielle', qui est la manière dont nous définissons l'implication logique dans ce cours. Certains logiciens critiquent ce comportement non-intuitif et préconisent l'utilisation d'autres connecteurs. Voir `https://en.wikipedia.org/wiki/Material_conditional` et `https://fr.wikipedia.org/wiki/Implication_(logique)`.
- Satisfiabilité (SAT - not to be confused with the US' 'Scholastic Aptitude Test')
- Problème de satisfaction de contraintes
- Complexité calculatoire
- Very Large Scale Integration (VLSI) pour la synthèse de circuits : comment construire des circuits logiques (tels que des circuits imprimés en informatique, par exemple) qui se comportent comme le veut ?
- Boolean Functions : Theory, Algorithms, and Applications de Peter L. Hammer et Yves Crama est un excellent livre qui couvre la logique propositionnelle en particulier, mais aussi la théorie des circuits et de la résolution de problèmes en général à travers l'étude des fonctions Booléennes - qui sont les fonctions que nous obtenons en interprétant nos connecteurs.

4. Peut-être avez-vous déjà entendu parler de ce terme en linguistique. Le sens est similaire : il signifie une déclaration redondante, le fait de 'dire la même chose deux fois'. Considérez par exemple 'un cadeau gratuit', les acronymes redondants tels que 'code PIN', ou les pléonasmes tels que 'Je l'ai vu, dis-je, vu, de mes propres yeux vu, ce qui s'appelle vu...' (Molière, Le Tartuffe)

6 Limites de la LP. Y a-t-il d'autres logiques ?

6.1 Autres connecteurs

Comme je l'ai déjà suggéré ci-haut, il est important de se souvenir que la plupart des règles et des définitions de nos connecteurs correspondent à *une* définition possible de logique, qui fonctionne relativement bien pour modéliser certains problèmes. La logique propositionnelle correspond à un langage, comme des notes de musique sur une partition ou des numéros sur une tablature. Avec une partition, il est plus facile de partager de la musique que des recettes de cuisines !

Certaines constructions de la logique propositionnelle omettent l'opérateur d'équivalence $\Leftrightarrow$, car il peut être défini en utilisant l'implication (matérielle) $\Rightarrow$. en effet, si vous avez déjà suivi un cours de logique mathématique, vous saurez sans doute que $\mathtt{A} \Rightarrow \mathtt{B}$ et $\mathtt{B} \Rightarrow \mathtt{A}$ est équivalent à $\mathtt{A} \Leftrightarrow \mathtt{B}$ (pour vous en convaincre, construisez et comparez les tables de vérité des formules $\mathtt{A} \Rightarrow \mathtt{B} \wedge \mathtt{B} \Rightarrow \mathtt{A}$ et $\mathtt{A} \Leftrightarrow \mathtt{B}$). En fait, en utilisant un unique opérateur que l'on appelle la 'Sheffer stroke', ainsi que les constantes V et F, il est possible de construire n'importe quelle formule que vous souhaitez ! C'est-à-dire que, pour toute formule que vous pouvez construire en utilisant les connecteurs des Tables 3 and 4, vous pouvez construire une formule qui lui est équivalente et qui ne met en jeu que ce connecteur et les constantes V et F ! Bien sûr, cela rend les formules assez difficiles à lire : j'ai donc choisi de conserver les plus classiques et courants 'et', 'ou', 'implique', 'non'... parce qu'ils sont plus proches de la langue naturelle et plus faciles à comprendre.

6.2 Complexité de la modélisation

La logique propositionnelle devient difficile à gérer dans certains cas pourtant assez simples. Imaginez que Gerbert aie voulu décrire sa position dans la cité de Rome en écrivant le nom de la rue et le numéro du bâtiment dans lesquels il se trouve. En logique propositionnelle, il aurait donc dû utiliser des symboles tels que

`Gerbert_au_numéro_3_de_la_rue_Appia_Antica`,

et ce, pour toute combinaison de numéro et de nom de rue. Pour modéliser de manière précise toutes les positions possibles, même avec seulement 10 numéros de rue et 100 rues différentes (une supposition très raisonnable !) il aurait besoin de 1000 variables différentes ! Pis encore, s'il avait également voulu décrire la position de sa tête de bronze dans Rome, il aurait eu besoin de 1000 variables de plus, telles que

`ma_tête_de_bronze_au_numéro_7_de_la_rue_Via_dei_Fori_Imperiali`!

Très rapidement, ce type de modélisation devient impossible à gérer proprement, et, en fait, la logique propositionnelle n'est pas très adaptée à ce genre de problèmes. Une solution possible est d'utiliser un autre type de logique, appelée *logique des prédicats* ou *logique du premier ordre*, qui permet l'utilisation de symboles spéciaux appelés *symboles de fonctions*. Ces symboles permettent d'accéder aux propriétés des objets que l'on manipule, telle que la position spatiale de l'individu représenté par une variable. Ainsi, en logique du premier ordre, Gerbert peut simplement écrire

$$\mathit{position}(\mathtt{Gerbert}),$$

et

$$\mathit{position}(\mathtt{ma\_tête\_de\_bronze}),$$

ce qui ne nécessite que trois symboles différents, au lieu des 2000 en logique propositionnelle.

Il y a toutefois une contrepartie à l'utilisation de la logique du premier ordre. Donner un tel pouvoir de représentation à notre logique la rend beaucoup, beaucoup plus compliquée pour résoudre des problèmes dans le cas général. Pour éviter cela, les logiciens utilisent plutôt ce qu'on appelle un *fragment* de la logique du premier ordre, légèrement plus puissant que la logique propositionnelle, mais pas tellement puissante qu'elle rend le raisonnement automatique impossible pour des raisons matérielle de temps et de mémoire.

6.3 Autres types d'interprétations

Une autre manière de considérer d'autres logiques et de mieux modéliser certains problèmes est de changer la nature de la fonction d'interprétation. Jusqu'à présent, nous avons vu que $\mathcal{I}$ assigne, à toute variable, une valeur binaire : soit vrai, soit faux. Cette décision est assez naturelle lorsque l'on veut décrire des propriétés ou des choses qui peuvent être soit vraies, soit fausses, telles que

`gerbert_chante_la_messe`

ou

`gerbert_meurt`.

Toutefois, certaines propriétés sont difficiles à capturer de cette manière. Considérez les propriétés dont le sens peut être modifié par l'adverbe 'très', tel que 'très gentil', 'très riche', 'très vieux' ou 'très gros'. Comparez ces expressions avec 'très chantant la messe' ou 'très mort'[5] !

Si nous augmentons le nombre de valeurs de vérités que l'interprétation peut prendre, nous pouvons capturer cet effet. En *logique multi-valuée*, nous pouvons par exemple choisir d'interpréter nos variables sur une échelle allant de 0 à 5. La valeur la plus grande, 5, peut correspondre à quelqu'un qui est très gentil, et 0 peut correspondre à quelqu'un qui n'est pas très gentil (ou très pas gentil). Une valeur de 3 peut correspondre à quelqu'un qui est plus gentil que quelqu'un qui a un score de 0, mais qui est moins gentil que quelqu'un qui a un score de 5...

La *logique floue* va encore plus loin dans ce sens en assignant des valeurs réelles aux variables, souvent entre 0 et 1 (inclus). Quelqu'un qui est, sans conteste, gentil, aura la valeur 1 une fois interprété, et quelqu'un qui n'est *pas*, sans conteste, gentil, aura la valeur 0. Jusqu'ici, rien de bien différent qu'en logique propositionnelle. Mais si quelqu'un est à la fois gentil et méchant (ou non gentil) à des degrés divers, il aura un score entre 0 et 1 (exclus). Une des motivations pour ce type de modélisation est qu'elle permet de faire des comparaisons subtiles entre variables. De plus, dans les domaines de modélisation et de contrôle de systèmes, la logique floue permet, par exemple, de modéliser des probabilités telles que la probabilité de mauvais fonctionnement d'un composant ou d'un élément d'un système, et ainsi de permettre à une intelligence artificielle de choisir le meilleur comportement, par exemple le comportement ayant le plus de chance de réussir, étant donnée une certaine situation. Certains connecteurs de logique floue correspondent directement à certaines opérations que l'on peut faire sur des variables aléatoires, tel que le calcul de la probabilité que deux évènements aient lieu ou la probabilité de l'évènement contraire.

☞ Pour aller plus loin

— Vous pourrez trouver discussion fondamentale sur les propriétés des choses et la nature des propriétés des choses dans l'*Organon* d'Aristote.
— Sheffer stroke, Peirce's arrow ('flèche de Peirce').
— Un autre excellent article du SEP : https://plato.stanford.edu/entries/logic-firstorder-emergence/, sur l'histoire de la logique du premier ordre.
— D'autres types de logique : logique floue, multi-valuée, temporelle, de description.
— Les logiques de descriptions sont une certaine classe de logiques très utilisées en intelligence artificielle symbolique, et dans le domaine des 'systèmes experts' tels que les systèmes d'aide au diagnostic médical. Les logiques de descriptions apportent une solution aux problèmes du manque de représentativité pour modéliser des situations réelles (comme en LP) et celui d'un langage si puissant que certains calculs deviennent impossibles (comme en LPO). Voir *A Description Logic Primer*, pour une introduction complète et agréable à lire des logiques de description : https://arxiv.org/abs/1201.4089.
— (Pour les logiciens avancés) Pour avoir une idée des choses étranges qui arrivent lorsque l'on utilise des logiques qui sont 'trop puissantes' : https://fr.wikipedia.org/wiki/Th%C3%A9or%C3%A8mes_d%27incompl%C3%A9tude_de_G%C3%B6del et https://blog.plover.com/math/Gdl-Smullyan.html.

7 Conclusion générale

Récapitulons ce que nous avons appris durant ce cours.

Nous avons appris ce que sont les *symboles*. Nous avons appris *à quoi peuvent servir* les symboles : communiquer et éliminer les ambiguïtés de la langue naturelle. Nous avons appris ce que sont les *logiques* ; nous avons appris à quoi elles peuvent servir, comme par exemple l'étude formelle du raisonnement humain ou artificiel. Nous avons étudié une logique bien particulière : la logique propositionnelle (LP). Nous avons appris comment *construire des formules* en LP, et comment les *lire*. Jusqu'à présent, nous n'avions parlé que de syntaxe. Mais nous avons maintenant considéré la notion formelle de vérité et d'*interprétation* d'une formule, pour relier les formules au monde réel. Finalement, nous avons mentionné d'*autres logiques* et moyens de représenter la 'vérité' plus ou moins adaptés suivant le problème à traiter.

> J'espère que ce document vous aura donné un avant-goût agréable de la logique symbolique en intelligence artificielle, ainsi que de bonnes pistes pour continuer à apprendre, si tel est votre souhait. Je vous souhaite ainsi bonne chance dans vos aventures logiques futures !

5. Ces expressions peuvent être modélisées linguistiquement ; mais ce sujet est en dehors du cadre de ce cours.

☞ **Et maintenant ?** ☕

Voici certaines questions intéressantes mais qui n'entrent pas dans le cadre de ce cours. Vous pourrez y trouver les réponses dans les livres cités à travers le texte, en particulier celui à la fin de la Section 4 et, par exemple, https://iep.utm.edu/prop-log/#H5.

— La déduction naturelle. Comment faire pour que les robots et les intelligences artificielles raisonnent et découvrent de nouvelles connaissances, de manière automatique, à partir de ce qu'elles savent déjà ?

— Y a-t-il des règles pour manipuler les formules ? Peut-on simplifier les formules, c'est-à-dire remplacer une formule par une autre, plus petite, mais qui modèle le monde de la même manière (les deux formules sont *logiquement équivalentes)* ? Peut-on combiner des formules pour en produire d'autres ?

— Peut-on prouver des théorèmes mathématiques de manière automatique, avec un ordinateur ?

A Présentation détaillée de certains objectifs d'apprentissage

Dans cette section, je donne des exemples des compétences que vous pourrez être amenées à mettre en jeu après avoir accompli les objectifs d'apprentissage. Je donne, par exemple, des exercices que vous devriez être capables de comprendre et, par la suite, de résoudre ou d'essayer de résoudre.

OA1 : connaître les objectifs d'apprentissage du cours Connaître et réfléchir sur ce que l'on apprend peut être important. Cela vous permet de penser la manière dont vous apprenez, vérifier vos connaissances, et orienter votre étude dans une direction consciente. Pour plus d'information, voir https://educationendowmentfoundation.org.uk/education-evidence/teaching-learning-toolkit/metacognition-and-self-regulation.

OA2 : être capable de donner des exemples de signes présents dans le monde qui vous entoure Un exemple simple de signe ou de symbole dans le monde qui nous entoure sont les panneaux de signalisation près de la rue. Chaque panneau possède un certain sens qui vous indique ce qu'il faut faire dans une certaine situation, ce à quoi il faut faire attention, etc.

Un autre exemple un peu moins visible directement est le langage que nous utilisons : les mots, les phonèmes, les gestes et les expressions dans le cadre de la langue des signes, sont autant de symboles qui portent les sens que nous partageons et qui nous permettent de communiquer.

OA2' : connaître l'étymologie du mot 'symbole' et savoir la relier à son utilisation moderne Voir la Section 2 pour l'étymologie grecque et l'utilisation moderne du mot. Voir aussi l'excellent site du CNRTL https://www.cnrtl.fr/etymologie/symbole.

OA3 : expliquer certaines ambiguïtés de la langue naturelle et expliquer comment la logique peut aider à les lever En langue naturelle, l'ambiguïté syntaxique peut générer pour une même phrase des interprétations différentes :

'Le robot voit la chercheuse avec sa caméra.'

Où est la caméra dans cette situation ? Est-ce que le robot est en train d'utiliser sa propre caméra, ou est-ce la caméra appartenant à la chercheuse, laquelle est en train de la porter ? Ou encore est-ce la caméra du robot ?

D'autres ambiguïtés proviennent du problème dit du *fondement des variables/symboles* ('variable/symbol grounding'), c'est-à-dire le lien entre les variables et le monde. Gerbert questionne sa tête de bronze à propos de 'Jérusalem'. Pour lui, le symbole 'Jérusalem' fait référence exclusivement, dans le monde réel, à la ville du moyen-orient. Toutefois, pour la tête de bronze, le symbole 'Jérusalem' fait aussi référence à l'église de Rome que les voisins appellent 'Jérusalem', d'où le qui-pro-quo, le problème de communication qui finalement fait que Gerbert tombe malade - même si son raisonnement était logiquement correct.

OA3' : énoncer au moins deux raisons qui motivent l'utilisation de la logique Communiquer, éliminer les ambiguïtés de la langue naturelle, étudier le raisonnement formel, automatiser tout cela. Voici une excellente citation de Leibniz, pionnier de la manipulation mathématique de symboles, qui résume tout cela :

> *'La seule manière de rectifier nos raisonnements est de les rendre aussi tangibles de ceux des Mathématiciens, afin que nous puissions découvrir notre erreur d'un regard, et, quand il y a des disputes entre personnes, il nous suffit de dire : "calculons, sans tarder, qui a tord et qui raison."'*
> Leibniz, L'art de la découverte, 1685 (traduction PM)

OA4 : savoir reconnaître et construire des expressions bien formées en LP Une 'expression bien formée' de LP est une expression qui suit les règles de construction que j'ai données dans la Section 4. Remarquez que j'utilise le terme 'expression' avant le terme de 'formule', pour bien faire la distinction entre une expression non nécessairement bien formée et une formule de LP.

Un résultat non trivial mais en dehors du cadre de ce cours est que si une expression ne suit *pas* ces règles, alors elle n'est pas une expression de LP. En d'autres termes, les règles de construction vous permettent de construire *exactement* toutes les formules de LP (rien de plus, rien de moins).

OA4' : construire une formule pour modéliser un problème réel C'est un problème compliqué, qui a inspiré de nombreux ouvrages. Le modèle, résultat de la modélisation, doit être assez proche du monde réel pour pouvoir en expliquer et en prédire le comportement, par exemple en simulant le monde réel grâce au modèle, qui agit comme un proxy ou comme un symbole du monde. Le modèle doit être assez simple pour que les résultats aient du sens pour nous humains, et que l'on puisse obtenir des résultats que l'on puisse utiliser. Par exemple, si un robot conducteur de voiture a une représentation du monde, un modèle, qui lui permet de reconnaître des formes de manière très précise - telles qu'une voiture qui arrive en sens inverse, par exemple - mais que la simulation et la reconnaissance de formes prend 20 secondes pour chaque image que le robot obtient du monde réel, ce robot ne pourra jamais être utilisé en situation réelle !

En général, en logique propositionnelle, vous pouvez utiliser des variables pour modéliser l'état des composants de la situation que vous voulez modéliser. Par exemple, `peut_laver` indique que quelque chose peut être lavé, ou pas, telles que les mains du fils d'Amavia dans le *Faerie Queene*. Vous pouvez utiliser des formules contenant une implication logique $\Rightarrow$ pour exprimer une connaissance générale que vous avez sur une situation. Par exemple, Gerbert est à la fois philosophe et pape : $\texttt{Gerbert} \Rightarrow \texttt{philosopher} \wedge \texttt{Pope}$.

Ce ne sont pas les seules options pour modéliser des problèmes, bien entendu ! Dans l'appendice B, je donne des exemples de modélisation en LP correspondant à différentes situations.

Finalement, voici une recommandation concernant le choix des noms de variables : essayez, autant que possible, d'utiliser des noms explicites, tels que `Gerbert`, `Amavia`, `mains_propres` instead of `G`, `A`, `m_p`.

OA4" : savoir lire une formule et la traduire en langue naturelle Comme nous l'avons vu, une formule peut-être interprétée de différentes manières. En général, expliquer le sens d'une formule est suffisant pour en lever les ambiguïtés. Un exercice de ce type a plutôt pour but de tester votre connaissance des connecteurs logiques et de leurs combinaisons possibles. Par exemple, je pourrais vous demander une explication des formules

$$\texttt{Amavia} \wedge \texttt{vertueuse} \wedge \texttt{morte},$$

(la situation dans laquelle le sujet est Amavia, et est à la fois vertueuse et morte) et

$$\texttt{robot} \wedge \neg\texttt{can\_move},$$

(la situation dans laquelle le sujet, qui est un robot, ne peut pas bouger) qui nécessitent toutes deux de comprendre le sens des connecteurs $\wedge$ et $\neg$; voir la Table 3.

A5 : déterminer si une formule est valide, satisfiable, ou pas, en construisant sa table de vérité Souvenez-vous qu'une formule est satisfiable s'il existe une interprétation, ou, en d'autres termes, un assignement de variables, qui la rende vraie. Pour construire la table de vérité d'une formule telle que

$$f := (\neg\texttt{peut\_nettoyer} \wedge (\texttt{peut\_nettoyer} \Leftrightarrow \texttt{mains\_propres})) \Rightarrow \neg\texttt{mains\_propres}.$$

vous pouvez suivre la recette que je décris ci-dessous. Cette formule décrit les tentatives de Guyon pour nettoyer les mains du fils d'Amavie : il n'est pas possible de nettoyer les mains (exprimé par $\neg\texttt{peut\_nettoyer}$) et les mains sont propres si et seulement si Guyon peut les nettoyer (exprimé par $(\texttt{peut\_nettoyer} \Leftrightarrow \texttt{mains\_propres})$). Nous voulons maintenant savoir si la conjonction de ces deux connaissances implique que les mains ne sont jamais propres (exprimé par la partie droite de l'implication).

1. Déterminez les variables propositionnelles de votre frormule. Dans ce cas, les variables propositionnelles de f sont `peut_nettoyer` et `mains_propres`.
2. Construisez une table de vérité qui contient toutes les combinaisons de valeurs possibles que vos variables peuvent prendre. Avec n variables, vous aurez besoin de 2^n lignes (savez-vous pourquoi ?)
3. Petit à petit, en commençant là où vous les pouvez, appliquez les règles de la Table 4 pour calculer les valeurs de chaque ligne de la table de vérité. Dans ce cas, commencez par déterminer la table de vérité de chaque variables propositionnelle dans la formule, puis de $(\texttt{peut\_nettoyer} \Leftrightarrow \texttt{mains\_propres})$, puis de $\neg\texttt{peut\_nettoyer}$, puis de $\neg\texttt{mains\_propres}$, puis de $\neg\texttt{peut\_nettoyer} \wedge (\texttt{peut\_nettoyer} \Leftrightarrow \texttt{mains\_propres})$, et enfin de f.

Je résume ce procédé dans la Figure 2, qui est une manière possible de présenter les choses. Vous pouvez aussi simplement tout présenter dans une table comme dans la Table 4.

Finalement, la table de vérité de f est $VVVV$: quelle que soit l'assignation de variable, f est toujours vraie. C'est dont une formule *satisfiable* et même *valide*, aussi appelée une *tautologie*. Que pouvez-vous en conclure sur la situation que la formule modèle ?

(¬	peut_nettoyer	∧	(peut_nettoyer	⇔	mains_propres))	⇒	¬	mains_propres
	V		V		V			V
	V		V		F			F
	F		F		V			V
	F		F		F			F

(¬	peut_nettoyer	∧	(peut_nettoyer	⇔	mains_propres))	⇒	¬	mains_propres
F	V		V	V	V		F	V
F	V		V	F	F		V	F
V	F		F	F	V		F	V
V	F		F	V	F		V	F

(¬	peut_nettoyer	∧	(peut_nettoyer	⇔	mains_propres))	⇒	¬	mains_propres
F	V	F	V	V	V		F	V
F	V	F	V	F	F		V	F
V	F	F	F	F	V		F	V
V	F	V	F	V	F		V	F

(¬	peut_nettoyer	∧	(peut_nettoyer	⇔	mains_propres))	⇒	¬	mains_propres
F	V	F	V	V	V	V	F	V
F	V	F	V	F	F	V	V	F
V	F	F	F	F	V	V	F	V
V	F	V	F	V	F	V	V	F

Le résultat !

FIGURE 2 – Calcul d'une table de vérité pas-à-pas.

OA6 : connaître d'autres types de valeurs de vérité Dans le monde binaire, digital, il existe deux valeurs de vérité : vrai ou faux, 1 ou 0 en logique Booléenne et en logique des circuits, $\top$ ('top') et $\perp$ ('bottom') en algèbre, max et min en analyse, etc. En *logique propositionnelle* et en *logique du premier ordre*, la fonction d'interprétation est une fonction $\mathcal{I} : Var \to \{V, F\}$ qui assigne une valeur de vérité, V ou F, à toute variable propositionnelle de Var. En *logique multi-valuée*, il est possible de définir une interprétation à valeurs dans d'autres ensembles ou même dans des intervalles. En *logique ternaire* par exemple, $\mathcal{I} : Var \to \{V, U, F\}$, U étant un symbole qui peut être compris comme signifiant 'inconnu' ou 'peu importe'. En *logique floue* et en *logique probabiliste*, vous pouvez même être confronté à une interprétation $\mathcal{I} : Var \to [0, 1]$ à valeurs dans l'intervalle continu $[0, 1]$!

Bien sûr, changer la fonction d'interprétation nécessite de donner des définitions alternatives aux connecteurs, dont les tables de vérité changent en conséquence. Par exemple, en logique de Kleene et en logique de Priest, la négation, conjonction, et disjonctions sont définies par les tables de vérité données dans la Table 5.

OA6' : connaître certaines limites de la LP et quelle logique utiliser pour quel but. Modéliser certaines situations du monde réel avec la LP nécessite parfois une grande quantité de variables, ce qui rend les calculs compliqués même avec un ordinateur. La *logique du premier ordre* est une autre logique qui résoud certains de ces problèmes, au prix toutefois d'être si puissante que le raisonnement est parfois impossible. Un compromis possible entre les deux est la famille de logiques dites *de description*.

B Quelques exemples de situations modélisées en logique propositionnelle

Dans cette section, je donne quelques exemples de formules de LP qui modélisent certaines situations variées inspirées de la littérature.

Notez les différences conceptuelles entre les éléments que je décris. Parfois, j'exprime une connaissance d'éléments *spécifiques* ou d'acteurs dans la situation, comme je l'ai fait ci-haut dans le cas de Gerbert, Guyon, ou Amavia ; en revanche, j'exprime parfois une connaissance *générale* à propos du monde, ou des règles telles que le fait que Jérusalem est une ville du moyen orient (n'est-ce pas ?), la plupart du temps à l'aide d'une implication telle que

$$\texttt{Jérusalem} \Rightarrow \texttt{ville} \land \texttt{au\_moyen\_orient}.$$

B.1 Les trois lois de la robotique d'Asimov

Isaac Asimov (1920 – 1992) était un auteur de science fiction et professeur de biochimie américain, né en russie. Il écrivit une grande quantité d'histoires de science fiction [6] et de livres de vulgarisation scientifique. Il eut une influence certaine sur la robotique (il a lui-même inventé le terme 'robotics') et sur l'intelligence artificielle en général [7].

En particulier, il explora les interactions entre les robots et les humains de manière non-conflictuelle, avec des robots considérés non pas comme des ennemis directs de l'humanité, mais plutôt comme des êtres dont les échecs découlent d'erreurs logiques plutôt que d'une intention malveillante, comme c'était souvent le cas dans la science-fiction de l'époque. Dans la nouvelle *Runaround* (1942), il présente un ensemble de règles, connues sous le nom des Trois Lois de la Robotique, qui ont inspiré le reste de ses travaux littéraires, ceux d'autres auteurs, et même la recherche éthique

6. Voir la liste de https://en.wikipedia.org/wiki/Isaac_Asimov_short_stories_bibliography !

7. Voir, par exemple, https://www.latimes.com/archives/la-xpm-1992-04-08-vw-636-story.html ; Marvin Minsky, le spécialiste en intelligence artificielle, se considérait parfois comme un 'psychologue pour robots', en référence aux oeuvres d'Asimov et au personnage de Susan Calvin en particulier.

A	B	¬A	A ∧ B	A ∨ B
V	V	F	V	V
U	V	U	U	V
F	V	V	F	V
V	U	F	U	V
U	U	U	U	U
F	U	V	F	U
V	F	F	F	V
U	F	U	F	U
F	F	V	F	F

TABLE 5 – Définitions de quelques connecteurs en logique de Kleen et de Priest.

en intelligence artificielle. Je les reproduis ici [8] :

— First Law : A robot may not injure a human being or, through inaction, allow a human being to come to harm.
— Second Law : A robot must obey orders given it by human beings except where such orders would conflict with the First Law.
— Third Law : A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

— Première Loi : un robot ne doit pas blesser un être humain ou, par son inaction, permettre à un humain d'être blessé.
— Deuxième Loi : un robot doit obéir aux ordres donnés par les êtres humains, sauf si de tels ordres entrent en conflit avec la Première Loi.
— Troisième Loi : un robot doit assurer sa propre existence, si tant est que cela n'entre pas en conflit avec la Première ou la Deuxième Loi.

Isaac Asimov, *Runaround* (Traduction PM)

La plupart des robots doivent suivre ces règles scrupuleusement. Les conflits apparaissent lorsqu'ils ne peuvent pas, ou suite à certaines interprétations ambiguës des règles.

Comme je le mentionne dans l'introduction de cette section, et parce que ce sont des lois très générales qui peuvent être appliquées à n'importe quel robot et dans n'importe quelle situation, et non pas à un robot *spécifique*, je vais exprimer ces lois à l'aide d'implications logiques.

Première Loi Notez l'importance des parenthèses pour indiquer que la négation s'applique à toute la sous-formule qui exprime l'idée de 'quelque chose qui blesse un humain ou qui permet à un humain dêtre blessé'. Notez également que j'ai choisi de ne pas exprimer la notion 'd'inaction'. Pensez-vous que ce soit important ? Comment auriez-vous exprimé cette notion ?

$$\texttt{première\_loi} := \texttt{robot} \Rightarrow \neg(\texttt{blesse\_humain} \vee \texttt{permettre\_à\_humain\_d'être\_blessé})$$

Deuxième Loi Je peux réutiliser le symbole `première_loi` car je l'ai déjà défini ci-dessus.

$$\texttt{deuxième\_loi} := \texttt{robot} \Rightarrow \texttt{obéir\_ordres\_humains} \wedge \texttt{première\_loi}$$

Troisième Loi

$$\texttt{troisième\_loi} := \texttt{robot} \Rightarrow \texttt{protéger\_propre\_existence} \wedge \texttt{première\_loi} \wedge \texttt{deuxième\_loi}$$

B.2 Talos, l'homme de bronze

Dans son poème épique, l'*Argonautica*, Apollonius de Rhodes (première moitié du troisième siècle avant J.C.), raconte l'histoire du héros Jason et des Argonautes, qui partirent à la recherche de la mythique toison d'or dans la Colchide lointaine (en Géorgie actuelle). Après avoir volé la toison d'or, il fuient les colchidiens à bord de leur vaisseau, l'Argos, et se dirigent sur l'île de Crête. L'île est protégée par un géant de bronze, Talos [9], qu'ils arrivent à vaincre avec l'aide de Médée.

8. https://webhome.auburn.edu/~vestmon/robotics.html
9. Voir également https://www.theoi.com/Gigante/GiganteTalos.html.

From that point they [the Argonauts] were to cross to Krete (Crete), the greatest island in the sea. But when they sought shelter in the haven of Dikte (Dicte) they were prevented from making fast to the shore by Talos, a bronze giant, who broke off lumps of rock from the cliff to hurl at them. A descendant of the brazen race that sprang from ash-trees, he had survived into the days of the demigods, and Zeus had given him to Europa to keep watch over Krete by running round the island on his bronze feet three times a day. His body and his limbs were brazen and invulnerable, except at one point : under a sinew by his ankle there was a blood-red vein protected only by a thin skin which to him meant life or death. He terrified the Argonauts, and exhausted though they were they hastily backed water. Indeed, what with thirst and other pains, they would have been driven away from Krete in a sorry frame of mind, but for Medeia (Medea), who stopped them as they turned the ship about. 'Listen to me,' she said. 'I think that I and I alone can get the better of that man, whoever he may be, unless there is immortal life in that bronze body. All I ask of you is to stay here keeping the ship out of range of his rocks till I have brought him down.'

Dès lors ils [les Argonautes] se préparaient à dépasser Crête, la plus grande île sur les mers. Mais alors qu'ils cherchaient refuge dans la crique de Dicte, ils en furent empêchés par Talos, un géant de bronze, qui arracha des morceaux de rochers de la falaise et leur projeta. Descendant de la race de bronze, née des arbres de cendres, il avait survécu à l'époque des demi-dieux, et Zeus l'avait confié à Europe pour qu'il garde Crête en parcourant l'île sur ses pieds de bronze trois fois par jour. Son tronc et ses membres étaient faits de bronze et invulnérables, sauf à un seul endroit : sous un nerf, près de sa cheville, se trouvé une veine rouge sang protégée seulement par une fine couche de peau, et qui pour lui signifiait la vie ou la mort. Il terrifia les Argonautes, qui, malgré leur fatigue, se réfugièrent bien vite sur les eaux. Et, en effet, assoiffés et bien en peine, ils auraient été expulsés de Crête si Médée ne les avait pas arrêtés alors qu'ils faisaient demi-tour. 'Écoutez-moi', dit-elle. 'Je pense que moi, et moi seule, peut triompher de cet homme, quel qu'il soit, à moins qu'il n'y ait une vie immortelle dans ce corps de bronze. Tout ce que je vous demande, c'est de rester hors de portée de ses rochers jusqu'à ce que je le mette à terre.'

Apollonius Rhodius, *Argonautica* 4. 1638 ff (translation Rieu) (traduction PM)

Je vais décrire ci-dessous les propriétés spécifiques d'un personnage spécifique, Talos. J'écrirai d'abord les phrases en langue naturelle, puis j'en proposerai une traduction en logique propositionnelle. J'utiliserai tous les connecteurs $\{\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow, (,)\}$ afin d'en montrer l'utilisation.

— Talos est un homme de bronze.

$$\texttt{Talos} \Rightarrow \texttt{homme\_de\_bronze}.$$

Que ce serait-il passé si j'avais utilisé le connecteur $\Leftrightarrow$ au lieu de $\Rightarrow$? Est-ce que cela aurait été un bon model de la situation ? Pourquoi ?

— Un 'homme de bronze' est un géant fait de bronze.

$$\texttt{homme\_de\_bronze} \Leftrightarrow \texttt{géant} \wedge \texttt{fait\_de\_bronze}.$$

Notez qu'il s'agit d'une règle générale à propos de la race des hommes de bronze, et pas seulement à propos de Talos.

— Soit Talos fait le tour de l'île en courant, soit il jette des rochers sur les intrus.

$$\texttt{Talos} \Rightarrow \texttt{court} \vee \texttt{jette\_des\_rochers\_sur\_les\_intrus}.$$

— Talos ne peut pas être tué.

$$\texttt{Talos} \Rightarrow \neg\texttt{mortel}.$$

C Lexique bilingue des termes importants

Vous trouverez ici une liste de vocabulaire important dont vous devriez avoir au moins une petite idée du sens. Selon moi, lorsque vous lirez un texte sur la logique, vous devriez au moins reconnaître le jargon utilisé dans le texte et, même sans connaître le sens exact d'un terme, vous devriez être à même de le deviner, ou de savoir où en chercher le sens, étant donné le contexte dans lequel le terme apparaît. Ce contexte est important : en effet, le terme *formule* n'a pas exactement le même sens suivant qu'il est utilisé en mathématiques ou en logique propositionnelle.

Notez qu'il n'est pas nécessaire de connaître toutes ces définitions et explications pas coeur. J'ai donné les termes que je considère les plus importants dans le cadre de ce cours dans la Table 1, à droite des Objectifs d'Apprentissage.

Glossaire

(IA explicable) explainable AI Aussi appelée 'XAI', cette branche de l'IA considère que la compréhension des mécanismes de décision de l'IA est particulièrement importante pour les humains. La manipulation symbolique offre une solution potentielle, en rendant les opérations logiques que l'IA suit explicites, et en utilisant des concepts qui sont partagés par les humains.

(fragment) fragment Un fragment d'une logique est une partie de celle-ci, obtenue par une restriction quelconque. Un fragment est habituellement moins descriptif mais plus adapté pour certains problèmes. Par exemple, la logique du premier ordre restreinte à des prédicats binaires est un fragment de la logique du premier ordre.

(tiers-exclus) excluded middle La 'loi du tiers-exclus', ou *tertium non datur* en latin ('il n'existe pas de troisième proposition'), énonce que pour toute proposition $\mathtt{A}$, soit $\mathtt{A}$ est vraie, soit $\neg\mathtt{A}$ est vraie. Les deux ne peuvent pas être vraies au même moment. Tout le monde ne suit pas cette loi, et certaines logique multi-valuée telles que certaines logiques indiennes n'y souscrivent pas. Le paradoxe du menteur est un exemple de phrase que le tiers-exclus ne parvient pas à capturer : que pouvez-vous dire de la phrase 'je suis en train de mentir' ? Est-elle vraie ? .

(valide) valid Une formule est valide si c'est une tautologie.

(équivalence) equivalence Deux propositions $\mathtt{A}$ et $\mathtt{B}$ sont équivalentes; ce que l'on note $\mathtt{A} \Leftrightarrow \mathtt{B}$, si et seulement si $\mathtt{A} \Rightarrow \mathtt{B}$ *et* $\mathtt{A} \Rightarrow \mathtt{B}$. Remarquez ici que 'si et seulement si', dans la phrase précédente, est également une équivalence, mais pour le *méta-langage* que j'utilise pour exprimer des propriétés et des définitions d'un autre langage – ici, pour définir l'équivalence logique $\Leftrightarrow$ de la logique propositionnelle..

agent (agent) En modélisation, de manière très générale, les agents désignent des unités individuelles qui ont la capacité d'*agir* (du latin 'agere') et de *percevoir* leur environnement. Gerbert peut être modélisé comme un agent, ainsi que sa tête de bronze, le géant Talos, un robot, un arbre.

ambiguïté (ambiguity) Apparaît lorsque quelque chose peut être interprété de différentes manières. En langage naturel, la même phrase peut avoir différents sens. Considérez la phrase 'elle voit l'homme sur la colline avec son télescope'. Est-ce l'homme qui porte le télescope, ou est-ce ce qu'elle a utilisé pour voir l'homme, tant il était loin ? La logique peut permettre de caractériser et de réduire ces ambiguïtés.

bien formé (well-formed) Une expression qui est bien formée, étant donné un certain ensemble de règle de construction de formules, a été construite à partir de ces règles. Par exemple, dans le contexte de la logique propositionnelle, '$\mathtt{A} \vee \mathtt{B}$' est une expression bien formée de la logique propositionnelle (étant donnés les symboles $\mathtt{A}$, $\mathtt{B}$ comme variables propositionnelles, et le symbole $\vee$ qui indique la disjonction) mais '$\mathtt{A} \vee \vee \mathtt{B}$' n'est pas une expression bien formée.

binaire (binary) En référence à 'deux'. Deux valeurs de vérité en *logique binaire* (0 et 1, Faux et Vrai) ; un connecteur, une fonction, ou une porte logique a deux entrées.

conjonction (conjunction) En logique propositionnelle, elle est parfois écrite '$\mathtt{ET}$', '$\wedge$', ou '.'. Typiquement, la conjonction des variables propositionnelles $\mathtt{A}$ et $\mathtt{B}$ est interprétée comme vraie si et seulement si $\mathtt{A}$ et $\mathtt{B}$ sont *toutes les deux* interprétées comme vraies. Moyen mnémotechnique pour se souvenir du nom : 'con' est un cognate latin de 'avec', et donc de 'et' !.

disjonction (disjunction) En logique propositionnelle, elle est parfois écrite '$\mathtt{OU}$', '$\vee$', ou '$+$'. Typiquement, la disjonction des variables propositionnelles $\mathtt{A}$ et $\mathtt{B}$ est interprétée comme vraie si et seulement si *l'une des variables* $\mathtt{A}$ et $\mathtt{B}$ est interprétée comme vraie. Remarquez que cette définition inclus la possibilité que les deux variables soient interprétées comme vraies en même temps. En langue naturelle, nous interprétons habituellement la disjonction 'ou' comme une disjonction exclusive, comme dans la phrase 'voulez-vous de la glace au chocolat ou à la noix de coco' (l'une ou l'autre, mais pas les deux) ! Certaines formulations forcent l'exclusivité et rendent impossible la réponse 'oui' à une telle question : 'préférez-vous la glace au chocolat ou la glace à la noix de coco' ?.

défini inductivement (recursively defined) Dans le contexte des formules et des langages formels, cette expression fait référence à la construction d'un ensemble de formules en partant de blocs de construction atomiques, c'est-à-dire indivisibles. Typiquement, une définition inductive (ou 'récursive') de formules indiquera d'abord quels sont les blocs les plus petits, indivisibles, qui sont déjà des formules. Ensuite, la définition va indiquer comment combiner ces blocs de construction atomiques pour construire des formules plus complexes.

faux (false) Voir vrai, vérité. .

formule (formula) En logique propositionnelle, une formule est une expression qui est construite en suivant les règles de construction du langage que j'ai données dans la Section 4. Notez qu'il existe d'autres ensembles (équivalents) de règles de construction en fonction des connecteurs que nous autorisons dans le langage.

IA (AI) Voir 'intelligence artificielle'.

implication (implication) En logique propositionelle, elle est parfois notée '$\Rightarrow$' ou '$\rightarrow$'. Voir raisonnement.

inférence (inference) Voir raisonnement.

insatisfiable (unsatisfiable) Une formule est insatisfiable si elle est logiquement équivalente à F. En d'autres termes, toute interprétation la rend fausse ; la table de vérité de la formule n'est constituée que de F. En logique propositionnelle, c'est le cas de la formule $\mathtt{A} \wedge \neg\mathtt{A}$.

intelligence artificielle (artificial intelligence) Ou IA. Un domaine large, préoccupé par l'étude de l'intelligence dans les machines artificielles. Dans le livre *Artificial Intelligence: A Modern Approach*, l'IA est, plus spécifiquement, l'étude des agents qui reçoivent des perceptions de leur environnement et qui agissent dessus. Par exemple, un chien robot reçoit une information visuelle sous la forme d'infrarouge qui lui indique qu'il y a un mur sur son chemin (perception), après quoi il tourne en réponse (action).

interprétation (interpretation) En logique propositionnelle, une interprétation est une fonction qui assigne à chaque variable propositionnelle une valeur de vérité dans $\{V, F\}$ et à chaque connecteur une fonction de $\{V, F\}^n$ dans $\{V, F\}$, où n est le nombre d'entrée du connecteur. Par exemple, le connecteur $\wedge$ est interprété comme une fonction binaire de $\{V, F\}^2$ vers $\{V, F\}$. Dans ce cours, nous interprétons tous les connecteurs usuels comme dans la Table 4.

langage (language) En théorie des calculs, en logique, en mathématiques, un langage est, de manière très générale, un ensemble de phrases. Alternativement, un langage peut être défini comme un alphabet, c'est-à-dire un ensemble de symboles, auquel on associe un ensemble de règles qui indiquent comment combiner ces symboles pour produire des phrases 'bien formées'. C'est le cas en logique propositionnelle tel qu'elle est présentée dans ce cours. L'ensemble des formules propositionnelles forme un langage, dont l'alphabet est constitué des symboles d'opérateurs, des constantes, et des variables, et dont les règles de construction sont détaillées dans la Section 4.

langue naturelle (natural language) La définition d'une langue naturelle est un peu différente de la définition très 'calculatoire' de 'langage', même si la différence exacte est sujette à discussions philosophiques. Une langue naturelle a évolué de manière naturelle à travers son utilisation par les humains, sans manipulation consciente, par opposition aux langues artificielles ou 'conlangs' qui ont été construites par une ou plusieurs personnes.

logique (logic) Dans le contexte de ce cours, une logique est un langage, auquel on associe un moyen de donner du sens aux phrases du langage (la fonction d'interprétation) et un système de déduction (voir raisonnement).

logique de description (description logic) Le terme désigne une famille de fragments de logique du premier ordre, particulièrement adaptées pour la représentation des connaissances.

logique du premier ordre (first-order logic) Elle est également appelée *logique des prédicats* (predicate logic). On y autorise plus de constructions qu'en logique propositionnelle, telles que des variables sans nom quantifiées et des prédicats qui sont interprétés comme des fonctions logiques, ainsi que des symboles de fonctions qui sont interprétés comme des symboles à valeurs dans des ensembles d'objets. Par exemple, un symbole de fonction tel que *position* peut indiquer la position de quelque chose qui sera interprété comme un élément d'un ensemble de positions prédéfini.

logique floue (fuzzy logic) Cette logique généralise la logique classique à deux valeurs telle que la logique propositionnelle en considérant des valeurs continues plutôt que discrètes, afin de pouvoir raisonner à partir de connaissances incertaines. Plus généralement, le terme désigne les domaines de recherche qui mettent en jeu les ensembles flous, par opposition aux ensembles classiques. Les ensembles classiques sont, de manière 'naïve', des ensembles d'objets dont l'appartenance est claire. Par exemple, vous pouvez considérer l'ensemble des statues de bronze. Toute statue (tout objet, en fait) sera clairement soit dans cet ensemble, soit n'y sera pas. Par contre, si vous considérez l'ensemble des *statues anciennes*, l'appartenance d'une statue sera plus fidèlement modélisée grâce à une fonction d'appartenance, qui, intuitivement, indique la probabilité que la statue appartienne à l'ensemble, ou, d'une certaine manière, qui en mesure l'ancienneté.

logique modale (modal logic) Le terme désigne une famille de logiques dont le but est d'exprimer la notion philosophique de *modalité*, c'est-à-dire la *nécessité* ainsi que la *possibilité*, qui sont habituellement dénotés par les symboles $\Box$ et $\Diamond$, respectivement. En langue naturelle, vous pouvez rencontrer des phrases telles que

'Par ce temps, j'ai besoin d'un parapluie si je veux sortir !'

ou

'Il risque de pleuvoir...'

Ces phrases expriment la nécessité et la possibilité, respectivement. En logique et en IA symbolique, la logique modale est un domaine très actif, car il est intéressant de construire des robots qui peuvent, par exemple, exprimer leurs désirs et raisonner avec les possibilités relatives à l'état du monde. La logique modale introduit des symboles et des opérateurs spéciaux pour exprimer ces deux notions ainsi que les règles pour les manipuler et les interpréter. La logique K (en l'honneur du logicien contemporain Kripke) est une des plus simples logiques modales : elle est construite à partir de la logique propositionnelle à laquelle on ajoute un seul opérateur pour exprimer la nécessité ainsi que deux règles additionnelles. Remarquez là l'importance de connaître la logique propositionnelle !.

logique multi-valuée (multi-valued logic) Bien que la logique propositionnelle et la logique du premier ordre soient, techniquement, multi-valuées car elles supportent deux valeurs de vérité (vrai et faux), le terme 'logique multi-valuée' fait référence aux logiques qui supportent plus de deux valeurs de vérité, telle que la logique de Kleen.

logique propositionnelle (propositional logic) L'une des logiques les plus simples, mais qui est déjà assez puissante pour beaucoup d'applications de manipulation de connaissances. Elle constitue un point de départ pour des logiques qui sont plus puissantes en terme d'expression conceptuelle : la logique du premier ordre, les logiques de description... La LP met en jeu la manipulation de formules et de variables propositionnelles qui peuvent prendre deux valeurs de vérité : vrai, ou faux. Les propositions sont séparées par des opérateurs logiques.

modèle (model) Ce terme a deux sens. Dans le cadre de la logique, un modèle pour une formule donnée est une interprétation qui la rend vraie. Pour le second sens, voir 'modéliser'.

modéliser (modelling) Modéliser, en général, fait référence à l'activité qui consiste à créer un système (un modèle) pour expliquer un phénomène (au sens large) par analogie. Ce modèle peut, par la suite, être utilisé pour prévoir le comportement du phénomène réel dans le futur ou sous d'autres conditions.

négation (negation) En logique propositionnelle, elle est notée '$\neg$', '$\sim$', ou $\overline{\mathtt{x}}$, étant donnée une variable propositionnelle `x`.

opérateur (operator) Je l'appelle aussi 'connecteur' dans ce cours. Un opérateur logique est une symbole dont on s'accorde sur le comportement et l'utilisation, et qui est utilisé pour construire des formules logiques. On l'appelle parfois un 'symbole logique' ou une 'porte logique' en théorie des circuits. Gardez à l'esprit la différence entre l'opérateur (au niveau syntaxique) et la fonction qu'il représente une fois interprété (au niveau sémantique) !.

paradoxe (paradox) Les paradoxes sont liés aux impossibilités et aux attentes bouleversées. En logique, ils peuvent apparaître comme des phrases qui se contredisent elles-mêmes, telle que

'Cette phrase est fausse'.

Parmi les problèmes liés, nous trouvons les problèmes de l'auto-référence, comme dans la phrase ci-haut. Je présente très formellement la classification des paradoxes de Quine. selon Quine, les paradoxes sont d'une part les choses qui apparaissent absurdes mais que l'on démontre par la suite, tels que certains 'paradoxes' mathématiques qui mettent en jeu l'infini, les choses qui apparaissent absurdes et que l'on infirme par la suite, telles que le paradoxe de la flèche de Zénon ou les preuves mathématiques que $1 = 2$, et finalement les choses qui n'appartiennent ni à l'une ni à l'autre de ces catégories et qui sont liées au paradoxes auto-référenciels tels que le paradoxe de Russell :

'Considère la liste des choses qui ne se contiennent pas. Est-ce que cette liste se contient elle-même ?'

Une quatrième catégorie inclut les choses qui sont à la fois vraies et fausses, ce que l'on peut relier à la loi du tiers-exclus.

problème du fondement des variables (variable grounding problem) Que l'on appelle aussi le *problème du fondement*. Quelle est la relation entre un modèle et le monde réel ? Quelle est la relation entre une variable et ce qu'elle modélise ? C'est un problème compliqué à garder à l'esprit lorsque l'on modélise le monde réel et que l'on veut éviter les paradoxes ou les déductions tragiques comme celle de Gerbert. Le problème de la référence, du lien entre un nom, un symbole, et ce que ceux-ci représentent, est important. Dans le cas de Gerbert, le symbole `Jérusalem` était 'fondé' de manière différente par Gerbert lui-même et par sa tête de bronze. Gerbert pensait que ce symbole faisait référence à la cité du moyen-orient, alors que la tête de bronze faisait référence, au moins, à l'église de Rome, comme le faisaient les gens qui vivaient à proximité.

raisonnement (reasoning) Maintenant que nous avons accès à des formules logiques qui nous permettent de modéliser certains aspects du monde, nous pouvons nous poser la question de l'obtention de nouvelles connaissances à propos du monde (sans le mettre en jeu directement). Les procédés qui nous permettent le faire sont *juste* en dehors du cadre de ce cours, mais pour vous donner un avant-goût, les formules logiques peuvent être modifiées et combinées de différentes manières pour en produire d'autres. Les modifications et les combinaisons doivent habituellement vérifier certaines propriétés pour être sûr que le système ne produira pas de formules absurdes. Par exemple, supposons que je sache seulement que 'j'aime les choses délicieuses' et que 'le comté est une chose délicieuse'. Je peux *inférer*, *déduire* que j'aime le comté en utilisant la règle générale d'*inférence*

$$(\texttt{A} \Rightarrow \texttt{B}) \wedge (\texttt{B} \Rightarrow \texttt{C}) \Rightarrow (\texttt{A} \Rightarrow \texttt{C}),$$

aussi connue sous le nom de *transitivité de l'implication logique*. Vous connaissez peut-être la règle du *Modus Ponens*, qui est une autre règle d'inférence selon laquelle

$$\texttt{A} \wedge (\texttt{A} \Rightarrow \texttt{B}) \Rightarrow \texttt{B}.$$

Il existe même des règles de simplification qui vous permettent de *simplifier* des formules. Par exemple,

$$\texttt{gerbert\_chante} \wedge \texttt{gerbert\_chante}$$

peut immédiatement être simplifier en `gerbert_chante`. Les robots et les intelligences artificielles peuvent être dotées d'un système de raisonnement qui leur permet d'augmenter la connaissance qu'ils ont du monde de manière automatique. Même certains théorèmes mathématiques peuvent être démontrés en utilisant un système de raisonnement ; bien sûr, il est important de prouver que le système est incapable de démontrer des paradoxes ou des formules fausses.

raisonnement formel (formal reasoning) Voir raisonnement.

SAT (SAT) SAT, pour 'satisfiabilité' ou 'satisfiable', fait référence au problème qui consiste à savoir si une formule est satisfiable ou pas, ou au problème qui consiste à trouver une interprétation qui la rende vraie.

satisfiable (satisfiable) Une formue est satisfiable s'il existe une interprétation qui rende la formule vraie. En d'autres termes, la table de vérité de la formule contient au moins un V.

signe (sign) Voir 'symbole'. Certains auteurs font la différence entre les deux, mais ceci est hors du cadre de ce cours.

symbol Du grec ancien 'σύμβολον' (sumbolon) qui signifie 'comparer', 'mettre (des choses) ensemble'. Concrètement, placer des choses côte-à-côte permet de les comparer .

sémantique (semantics) L'étude des sens, références, varités. Du grec 'σημαντικός' ('semantikos'), 'signifiant'.

sémiotique (semiotics) L'étude des signes, des symboles, et de leurs relations au sens et à la sémantique.

table de vérité (truth table) Étant donnée une formule en logique propositionnelle, sa table de vérité est une table qui décrit toutes les valeurs de la formule pour toute valeur que peuvent prendre les variables propositionnelles de cette formule.

tautologie (tautology) Une formule est une tautologie si elle est logiquement équivalente à V. En d'autres termes, toute interprétation la rend vraie ; la table de vérité de la formule n'est constituée que de V. En logique propositionnelle, la formule 'V' est une tautologie. La formule $\texttt{A} \vee \neg\texttt{A}$ est aussi une tautologie (construisez-en la table de vérité pour vous en convaincre). Voir 'tiers-exclus'. En logique mathématique et en théorie de la preuve, les 'théorèmes' sont parfois définis comme étant des tautologies.

tête de bronze (brazen head) Une tête faite de bronze. Certaines légendes du moyen-âge racontent l'histoire de savants qui auraient construit de telles têtes, tels que Gerbert d'Aurillac ou Roger Bacon. On raconte qu'elles avaient des pouvoirs spéciaux de divination. Pour plus d'information, voir *The Persistent Peril of the Artificial Slave* par Kevin LaGrandeur.

variable (variable) Comme en mathématiques, une variable est un symbole qui peut être interprété. En mathématiques, une variable x peut être un symbole qui représente une valeur réelle ou un entier naturel, par exemple. En logique propositionnelle, une variable telle que `gerbert_chante` est un symbole qui peut être interprété comme vrai, ou faux.

vrai (true) Voir faux, vérité.

vérité (truth) La notion de vérité est complexe. En logique Booléenne, nous décidons qu'il existe deux valeurs, vrai et faux, et nous calculons la valeur de vérité de formules en utilisant des règles algébriques suivant les travaux, entre autres, de Boole et de Leibniz.