Logics for Artificial Intelligence- AI Master

Session 1 – Formalization in first-order logic

Logic is the study of arguments. An **argument** is composed by a given set of facts (**premises**) and a potential **conclusion**. An argument is valid if, whenever all the premises hold, the conclusion must also necessarily hold (i.e. it is not possible that all the premises are true and the conclusion false). Any logical formalism provides **a formal language** (to represent the premises and the conclusion in a non-ambiguous, precise way) and **tools to validate arguments** (i.e. to determine if an argument is valid or not).

In the first part of the course we consider **first-order logic**. In order to validate an argument, the first task is to translate the premises and the conclusion from a given natural language (Catalan, Spanish, English, …) to a formal language. This task is called **formalization**.

The basic elements that we want to represent in predicate logic are objects/individuals, properties (sets of objects), relationships between sets of objects (e.g. inclusion, intersection), and relationships between individuals. The objects that we want to talk about in an argument are the **domain** of discourse.

The main elements of the language of predicate logic are **constants** (that represent particular objects), **variables** (generic references to any object), the **parenthesis** '(' and ')', **predicates** (that can express a relationship between *n* individuals), **logical operators** (conjunction (^), disjunction (˅), negation (¬) and conditional/implication (→))[1] and **quantifiers** (universal (∀), existential (∃))[2]. A **term** is a constant or a variable. A binary predicate *P* applied to two terms *a* and *b* may be written P(a,b) or Pab.

Formulas in predicate logic are the following:

- Atomic predicate: *n*-ary predicate applied to *n* terms.
- If A is a formula, (¬A) is a formula.
- If A and B are formulas, (A ^ B), (A ˅ B) and (A → B) are formulas.
- If A is a formula and x is a variable, ∀x(A) and ∃x(A) are formulas, and A is the **scope** of the quantifier.
- There are not any other formulas.

In order to simplify the notation, it is allowed to eliminate some parenthesis, by considering the following priority between operators: the quantifiers have the maximum priority, followed by negation, then conjunction and disjunction have the same priority, and implication has the lowest priority. Therefore, the formula A ^ ¬ B → C has the meaning ((A ^ (¬B)) → C).

---

[1] Other operators, like Equivalence (↔), XOR (Exclusive OR), NOR (not OR) or NAND (not AND) may be considered.

[2] Other elements, such as functions, will be introduced later in the course.

An **atomic predicate** represents an observable fact in a given fixed state of the world, which can only be **true** or **false** in that state. Spatial or temporal aspects are not considered in first-order logic. The intuitive meaning of the basic operators is the following:

- Conjunction: the two conjuncts are true at the same time.
- Disjunction: at least one of the two disjuncts is true (and even both of them may be true).
- Negation: the negated fact does not hold in the world.
- Implication: if the antecedent is true, it can be affirmed that the consequent is also true (it is not possible that the antecedent is true and the consequent is false).

The formalization of implication can be particularly tricky. These are examples of sentences that could be translated into the conditional formula ($P \rightarrow Q$):

- If P then Q
- If P, Q
- Q if P
- Q when P
- P is a sufficient condition for Q
- Whenever P, Q
- It is not possible to have P and not to have Q
- P only if Q
- Q is a necessary condition for P
- Q is needed to have P
- If not Q then not P
- If Q does not hold, it is not possible to have P
- Either P is false or Q is true

An appearance of a variable in a formula is **bound** if it is contained in the scope of a quantifier on that variable; otherwise, the appearance of the variable is **free**. The expression Ax refers to a formula in which *x* appears free.

The intuitive meaning of the quantifiers is the following:

- $\forall x(Ax)$: all the objects in the domain have the property A.
- $\exists x(Ax)$: there is at least one object in the domain that has property A.

Basic material you should study on week 1:

- Book by Brachman and Levesque: chapters 1 and 2 (in especial, section 2.2).

Complementary material:

- Material of First-Order Logic (Catalan): language and formalization in propositional logic (sections 2.1 and 2.2) and predicate logic (sections 3.1 and 3.2). You can test your learning with the formalization and interpretation exercises (4.1, 4.2, 5.1 and 5.2).
- Stanford course *Introduction to Logic*: Propositional Logic and Relational Logic (note that the notation they use has some small differences with the one used in this course).