Natural Deduction for Classical 1^{st}-Order Logic

1 Background on Logic

Logic was developed as a way to reason about valid forms of argument. Consider the case of the magic rock that keeps tigers away (from the Simpsons, paraphrased):

```
 Lisa: By your logic I could claim that this rock keeps tigers away.
Homer: Oh, how does it work?
 Lisa: It doesn't work...it's just a stupid rock. But I don't see any tigers around, do you?
Homer: Lisa, I want to buy your rock.
```

Clearly there's something wrong with that argument. Logic is the study of what kinds of arguments make sense and what kinds do not—in other words, given certain facts, what conclusions can I reasonably draw? In the West, formal logic started with Aristotle and his syllogistic reasoning. Logic was also developed independently from the Greeks in India and China. Aristotelian logic had a huge impact in the West for a long time, but modern forms of logic started in the 1800s with George Boole (boolean algebra) and Gottlob Frege (predicate logic). When we look at formal logic, it seems like it's just symbol pushing, with no real meaning, e.g., $(A \wedge B \implies C) \implies (\neg C \implies \neg A \vee \neg B)$. And that's exactly what it is! In fact, that's why formal logic is so powerful. It depends on the syntactic form of the argument, *not* the semantic meaning of the argument. This means that logic tells us what forms of argument are acceptable without depending on what the argument is actually about—sports, politics, science, it doesn't matter. Here are some important definitions related to logic:

Interpretation. An *interpretation* maps the logical syntax (i.e., the symbols used to make logical formulae) to a specific domain of discourse; this gives an argument a specific meaning. For example, in the logical formula above we could assign the interpretation A = "you attend class", B = "you attend discussion section", and C = "you get a good grade". Thus, if you don't get a good grade it must be that you didn't attend class or you didn't attend discussion section.

Validity. A logical formula is *valid* if it is guaranteed to be true no matter what interpretation you give it. The logical formula in the previous example is a valid formula.

Satisfiability. A logical formula is *satisfiable* if there is *some* interpretation that can make it true. For example, the formula $A \wedge B$ is satisfied under the interpretation A = "my hair is brown" and B = "I have blue eyes", but there are also interpretations where it is not satisfied, e.g., A = "my hair is pink" and B = "I have purple eyes". In computer science we're often interested in taking a formula and asking if it's satisfiable and, in particular, what interpretation satisfies it.

Unsatisfiability. A logical formula is *unsatisfiable* if there is *no* interpretation that makes it true. For example, $A \wedge \neg A$.

2 Natural Deduction

There are several ways to define first-order logic. In Computer Science most people are introduced to it via the Hilbert-style axiomatic formulation; in philosophy most people learn it via the Natural Deduction formulation. There are also other formulations possible, such as the sequent calculus. These all define the same thing, they just provide different perspectives and different ways of getting to the same end. For reasons that will become clear later in the course, we'll use the natural deduction style. Natural deduction was invented by Gerhard Gentzen in the early 1900s. He wanted to develop a definition of logic that comes as close as possible to the way that people actually *think*, hence the term "natural". Gentzen made huge contributions to the study of logic; he was also a Nazi and member of the SS. We will separate the man from the logic and only look at the logic.

The fundamental notion of natural deduction is a **judgement** on the truth of a **proposition** based on **evidence**. A *proposition* is something that can be either true or false, e.g., "it is raining". A *judgement* says whether a proposition is true based on some evidence (e.g., observation, or a derivation from known facts). For example, we could have the proposition "it is raining" and the judgement "the proposition 'it is raining' is true", based on the evidence that I can see it raining. Propositions are given as formulae in the syntax of first-order logic.

3 Syntax for Logical Propositions

$$x \in \mathit{Variable} \qquad f \in \mathit{Function} \qquad p \in \mathit{Predicate}$$

$$t \in \mathit{Term} ::= x \mid f(\vec{t})$$
$$A, B \in \mathit{Proposition} ::= p(\vec{t}) \mid \top \mid \bot \mid A \wedge B \mid A \vee B \mid A \Rightarrow B \mid \neg A \mid \forall x.A \mid \exists x.A$$

Where $\vec{t}$ is a sequence of terms of length equal to the arity of the given function or predicate symbol. When we need multiple variables we'll use letters u, v, w, x, y, z. When we need multiple function symbols we'll use f, g, h, i, j, k. When we need multiple predicate symbols we'll use m, n, o, p, q, r, s. When we need multiple propositions we'll use A, B, C, D, E. As a last resort, we'll use subscripts to distinguish names, so that, for example, x_1 and x_2 indicate different variables.

Every function and predicate symbol has an *arity*, which is simply the number of arguments that it accepts. For example, a function $f(u, v)$ has arity 2 and a function $g(u, v, w)$ has arity 3. We'll sometimes indicate the arity using notation such as $f/2$ (for function symbol f with arity 2) and $g/3$ (for function symbol g with arity 3). Functions and predicates can have arity 0, in which case we write them without parentheses. For example, if we have a predicate symbol $p/0$ we'll write p instead of $p()$. Function symbols with 0 arity are called *constants* and predicate symbols with 0 arity are called *propositional variables*.

Function symbols f, g, etc. stand for functions that map objects to other objects. Of course, without an interpretation we don't know what those functions and objects are—an *interpretation* will include a specific domain of objects and map each function symbol to a function are over that domain. For example, one interpretation might specify that the objects are people, and connect the function symbol $\mathit{fatherOf}/1$ with the function that takes a person and returns that person's father, e.g., $\mathit{fatherOf}(\text{Mary}) = \text{John}$. Terms, which as indicated above are made up of variables and functions over terms, are just names for objects in some unknown domain. They don't have any inherent meaning until we give them an interpretation, which will provide a specific domain of objects. Different interpretations may give them different meanings.

Predicate symbols p, q, etc. stand for relations between objects. Recall that a relation is a set of tuples. For example, we might define the relation $\{(\mathit{grass}, \mathit{green}), (\mathit{sky}, \mathit{blue}), (\mathit{apple}, \mathit{red})\}$. Again, without an interpretation we don't know what the relations or objects are; an interpretation will map each predicate symbol to a specific relation over the domain of objects for that interpretation. For example, one interpretation might map the predicate symbol $\mathit{color}/2$ to the relation defined above.

3.1 Example Terms (Specifying Objects in a Domain)

- x
- $f(x, g(a, b))$
- $f(g(h(x, y), i, j(x, z)), h(y, z))$

3.2 Example Propositions (Specifying Relations Between Objects)

- $\neg\forall x.(q(x) \vee \neg p(x))$
- $\exists x.q(x, f(x), g) \wedge s(x) \Rightarrow \forall x.r(h, x)$
- $\forall x \exists y.(r(x, y) \Rightarrow r(y, x))$

4 Making Judgements

Given a proposition, we want to be able to make a judgement about it. There are a number of different kinds of judgements that we could make, but we'll focus on one of the most important: *truth*. We make judgements based on evidence. Some evidence will be given to us as facts (i.e., *axioms*). Other evidence will come from derivations based on *inference rules*. These rules give us guidelines for how to make new judgements based on existing judgements. Judgements will often make use of *hypotheses*, i.e., propositions that we will temporarily assume are true while trying to make the judgement. A hypothesis is just a sequence of propositions; we will symbolize an arbitrary hypothesis using the letter Γ.

A judgement will be of the form '$\Gamma \vdash A$ *true*'; this says that if we assume the propositions contained in Γ are true, then we are justified in saying that proposition A is true. We will take the judgement $\Gamma, A \vdash A$ *true* as an axiom; i.e., we always know that under the assumption that A is true, we can conclude that A is true. The notation 'Γ, A' means that we're appending the proposition A to the list of propositions Γ.

5 Inference Rules

Inference rules are just a compact way of writing `if..then` statements. They consist of a horizontal line with zero or more judgements on top of the line, called *premises*, and exactly one judgement on the bottom of the line, called the *conclusion*. An inference rule is saying that *if* all of the premises can be proven true, *then* the conclusion must also be true.

For each logical connective $\top, \bot, \wedge, \vee, \Rightarrow, \neg, \forall, \exists$ there are rules that tell us how we can use them to make judgements. Each connective (except for $\bot$) has an *introduction rule* that shows how we can judge that a proposition using that connective is true. Each connective (except for $\top$) has an *elimination rule* that shows how we can judge that a proposition is true based on knowing some other proposition using that connective is true. Note that the rules below usually don't specify the contents of Γ; it is there only to make clear that the rules are valid no matter what assumptions you make, as long as the assumptions are consistent.

Logical Connective	Introduction Rule	Elimination Rule
True	$\dfrac{}{\Gamma \vdash \top\ \mathit{true}}\ \top\mathrm{I}$	
False		$\dfrac{\Gamma \vdash \bot\ \mathit{true}}{\Gamma \vdash A\ \mathit{true}}\ \bot\mathrm{E}$
Conjunction	$\dfrac{\Gamma \vdash A\ \mathit{true} \qquad \Gamma \vdash B\ \mathit{true}}{\Gamma \vdash A \wedge B\ \mathit{true}}\ \wedge\mathrm{I}$	$\dfrac{\Gamma \vdash A \wedge B\ \mathit{true}}{\Gamma \vdash A\ \mathit{true}}\ \wedge\mathrm{E}_L$ $\dfrac{\Gamma \vdash A \wedge B\ \mathit{true}}{\Gamma \vdash B\ \mathit{true}}\ \wedge\mathrm{E}_R$
Disjunction	$\dfrac{\Gamma \vdash A\ \mathit{true}}{\Gamma \vdash A \vee B\ \mathit{true}}\ \vee\mathrm{I}_L$ $\dfrac{\Gamma \vdash B\ \mathit{true}}{\Gamma \vdash A \vee B\ \mathit{true}}\ \vee\mathrm{I}_R$	$\dfrac{\Gamma \vdash A \vee B\ \mathit{true} \quad \Gamma, A \vdash C\ \mathit{true} \quad \Gamma, B \vdash C\ \mathit{true}}{\Gamma \vdash C\ \mathit{true}}\ \vee\mathrm{E}$
Implication	$\dfrac{\Gamma, A \vdash B\ \mathit{true}}{\Gamma \vdash A \Rightarrow B\ \mathit{true}}\ \Rightarrow\mathrm{I}$	$\dfrac{\Gamma \vdash A \Rightarrow B\ \mathit{true} \qquad \Gamma \vdash A\ \mathit{true}}{\Gamma \vdash B\ \mathit{true}}\ \Rightarrow\mathrm{E}$
Negation	$\dfrac{\Gamma, A \vdash \bot\ \mathit{true}}{\Gamma \vdash \neg A\ \mathit{true}}\ \neg\mathrm{I}$	$\dfrac{\Gamma \vdash A\ \mathit{true} \qquad \Gamma \vdash \neg A\ \mathit{true}}{\Gamma \vdash \bot\ \mathit{true}}\ \neg\mathrm{E}$
Universal Quantification	$\dfrac{\Gamma \vdash A[x \mapsto a]\ \mathit{true}}{\Gamma \vdash \forall x.A\ \mathit{true}}\ \forall\mathrm{I}$	$\dfrac{\Gamma \vdash \forall x.A\ \mathit{true}}{\Gamma \vdash A[x \mapsto t]\ \mathit{true}}\ \forall\mathrm{E}$
Existential Quantification	$\dfrac{\Gamma \vdash A[x \mapsto t]\ \mathit{true}}{\Gamma \vdash \exists x.A\ \mathit{true}}\ \exists\mathrm{I}$	$\dfrac{\Gamma \vdash \exists x.A\ \mathit{true} \qquad \Gamma, A[x \mapsto a] \vdash B\ \mathit{true}}{\Gamma \vdash B\ \mathit{true}}\ \exists\mathrm{E}$

In the universal and existential quantifier rules, it is very important that the unknown a is *fresh*, i.e., it cannot occur in any hypothesis or in $\forall x.A$ itself. We will see an example of why this is important in Section 6.4.

6 Proof Examples

Here we give several examples of how to prove a judgement about a given proposition. Note in the following proofs that the desired conclusion is at the bottom of the proof. We can think of these proofs as *derivation trees* rooted in the conclusion, growing upwards according to the appropriate introduction and elimination rules. The leaves of the tree are judgements that are trivially true; in this case, they are all of the form $\Gamma, A \vdash A\ true$.

6.1 Example 1

We will prove the following judgement: $\vdash p \wedge q \Rightarrow q \wedge p\ true$.

$$\dfrac{\dfrac{\dfrac{p \wedge q \vdash p \wedge q\ true}{p \wedge q \vdash q\ true}\,{\scriptstyle \wedge E_L} \quad \dfrac{p \wedge q \vdash p \wedge q\ true}{p \wedge q \vdash p\ true}\,{\scriptstyle \wedge E_R}}{p \wedge q \vdash q \wedge p\ true}\,{\scriptstyle \wedge I}}{\vdash p \wedge q \Rightarrow q \wedge p\ true}\,{\scriptstyle \Rightarrow I}$$

6.2 Example 2

We will prove the following judgement: $\vdash p \Rightarrow (q \Rightarrow (p \wedge q))\ true$.

$$\dfrac{\dfrac{\dfrac{p, q \vdash p\ true \quad p, q \vdash q\ true}{p, q \vdash p \wedge q\ true}\,{\scriptstyle \wedge I}}{p \vdash q \Rightarrow (p \wedge q)\ true}\,{\scriptstyle \Rightarrow I}}{\vdash p \Rightarrow (q \Rightarrow (p \wedge q))\ true}\,{\scriptstyle \Rightarrow I}$$

6.3 Example 3

We will prove the following judgement: $\vdash (p \Rightarrow q) \wedge (p \Rightarrow r) \Rightarrow (p \Rightarrow (q \wedge r))\ true$. In the following proof, in order to fit the proof on the page we will abbreviate the assumption $(p \Rightarrow q) \wedge (p \Rightarrow r), p$ as Γ.

$$\dfrac{\dfrac{\dfrac{\dfrac{\Gamma \vdash p\ true \quad \dfrac{\Gamma \vdash (p \Rightarrow q) \wedge (p \Rightarrow r)\ true}{\Gamma \vdash p \Rightarrow q\ true}\,{\scriptstyle \wedge E_R}}{\Gamma \vdash q\ true}\,{\scriptstyle \Rightarrow E} \quad \dfrac{\Gamma \vdash p\ true \quad \dfrac{\Gamma \vdash (p \Rightarrow q) \wedge (p \Rightarrow r)\ true}{\Gamma \vdash p \Rightarrow r\ true}\,{\scriptstyle \wedge E_L}}{\Gamma \vdash r\ true}}{(p \Rightarrow q) \wedge (p \Rightarrow r), p \vdash q \wedge r\ true}\,{\scriptstyle \wedge I}}{(p \Rightarrow q) \wedge (p \Rightarrow r) \vdash p \Rightarrow (q \wedge r)\ true}\,{\scriptstyle \Rightarrow I}}{\vdash (p \Rightarrow q) \wedge (p \Rightarrow r) \Rightarrow (p \Rightarrow (q \wedge r))\ true}\,{\scriptstyle \Rightarrow I}$$

6.4 Example 4

In this example we will illustrate the importance of using a fresh a in the universal quantification introduction rule by giving an **incorrect** proof of the following judgement: $\vdash \forall x \forall y. p(x) \Rightarrow p(y)\ true$. This judgement is obviously wrong; the proof will be incorrect because it fails to properly use a fresh a.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{p(a) \vdash p(a)\ true}{p(a) \vdash \forall x. p(x)\ true}\,{\scriptstyle \forall I}}{p(a) \vdash p(b)\ true}\,{\scriptstyle \forall E}}{\vdash p(a) \Rightarrow p(b)\ true}\,{\scriptstyle \Rightarrow I}}{\vdash \forall y. p(a) \Rightarrow p(y)\ true}\,{\scriptstyle \forall I}}{\vdash \forall x \forall y. p(x) \Rightarrow p(y)\ true}\,{\scriptstyle \forall I}$$

The problem is at the top of the derivation tree, where we used the $\forall I$ rule to turn $p(a)$ into $\forall x.p(x)$. We can't do that because a isn't fresh—it's used in the hypothesis of the top judgement.