Using Natural Deduction to Represent Arguments (Part 1)

Scott Martin

August 10, 2010

Propositions, Arguments, and Logic

Recall that an argument, as we technically defined it, is a collection of propositions some of which are premises and some of which are conclusions. Now consider the following informal argument:

(1) a. If Clarence doesn't catch a Walleye, Myrtle will have to go shop for dinner.
b. If Clarence does catch a Walleye, Myrtle will cook it for dinner.
c. Myrtle didn't go shop for dinner.
d. So, Myrtle is cooking Walleye for dinner.

Notice that this meets our definition of an argument because

1. All the sentences in (1) are declaratives, and

2. There are some premises (namely, (1a-1c)) and a conclusion (1d).

If we want to represent the argument in (1) using PL, the first step is to represent all the declaratives as atomic propositions. Here's one way:

W is the proposition expressed by *Clarence catches a Walleye.*

S is the one denoted by *Myrtle shops for dinner.*

C corresponds to *Myrtle cooks Walleye for dinner.*

With these three atomic propositions, we can represent (1) as follows:

$$\neg W \to S \tag{1a}$$
$$W \to C \tag{1b}$$
$$\neg S \tag{1c}$$
$$C \tag{1d}$$

So we can see how this argument progresses: we're given that $\neg W \to S$. Then we find out that $\neg S$. Remembering that $\neg W \to S$ is equivalent[1] to $\neg S \to \neg\neg W$, we can now conclude that $\neg\neg W$, i.e. W must be true. But that means that C, since we're given that $W \to C$.

This is OK, but ideally we don't want to rely on writing things out in prose all of the time. How can we make this process of deduction more formal and precise?

[1]If you don't remember this, or find this confusing, doing the truth table for both should make it clear.

Natural Deduction for Propositional Logic

- We know how to represent propositions and how to interpret those representations. But now we want to know how propositions go together to form valid or invalid arguments.
- That is, we want a formal system that will model deductively valid reasoning patterns.
- Such a system should let us distinguish arguments with true premises (sound arguments) from ones that don't have true premises, and arguments whose conclusions follow from their premises (valid arguments) from ones whose conclusions don't follow from their premises.
- Natural Deduction (ND) is a system for encoding syntactically what we can find out about validity using truth tables. So it's a kind of shorthand.
- Remembering the truth table for implication ($\rightarrow$), φ and ψ are both true when $\varphi \rightarrow \psi$ is true. Likewise, the truth table for conjunction ($\wedge$) has $\varphi \wedge \psi$ true when φ and ψ are also true.
- ND uses **inference rules** to express validity so we don't have to keep referencing the truth tables for the connectives.
- These deductive rules give us a way to combine the premises of the argument in (1) to see that its conclusion is, in fact, valid.

Inference Rules

In all of our ND rules, φ and ψ are propositions in PL. These rules follow the general format

$$\frac{P_0 \qquad \dots \qquad P_n}{C}\ (R)$$

where P_0 through P_n are premises, C is the conclusion, and R is a label corresponding to the inference rule that was invoked. Notice that this rule format means that an inference rule can involve any number of premises. In fact, the most basic rule has no premises at all:

Inference Rule 1 (Hypothesis)**.**

$$\frac{}{\varphi}\ (\text{Hyp})$$

This rule is so simple it almost seems trivial. But it is in many ways the most powerful rule we have–It's the way all premises are introduced in any argument. It essentially says that you can assume anything you want, which is OK, because remember that in general we're not interested in the truth or falsehood of an argument's premises, just what can be deduced from them. For the informal argument in (1), we would need to use Hypothesis to introduce the complex propositions denoted by (1a) through (1c).

Inference Rule 2 ($\rightarrow$-Elimination)**.**

$$\frac{\varphi \rightarrow \psi \qquad \varphi}{\psi}\ (\rightarrow\text{E})$$

$$\dfrac{\dfrac{}{\neg S \to \neg\neg W}\,(\text{Hyp}) \qquad \dfrac{}{\neg S}\,(\text{Hyp})}{\neg\neg W}\,(\to\text{E})$$

Figure 1: Partial proof of the argument in (1).

Rule 2, sometimes called Modus Ponens, just says that if you've proved (or assumed) an implication and you've also proved the antecedent of that implication, then you've proved the consequent.

Given rule 2, we can already represent part of the process of deduction used in the informal argument in (1). A graph of an argument like the one in Figure 1 is called a **proof**. Notice that the last inference step of the above proof is just an instance of Rule 2 with $\neg S$ inserted for φ and $\neg\neg W$ inserted for ψ. The first inference steps of this proof are both instances of Rule 1.

Sometimes we can rewrite PL sentences as other PL sentences they are equivalent to. We did this in our informal reasoning process to turn $\neg W \to S$ into $\neg S \to \neg\neg W$ so that we could get things to work out right. But notice that one of the steps in our deduction process requires us to treat $\neg\neg W$ as W so that we can get to the conclusion C. We'd ideally like to be able to write proofs so that deduction steps always invoke rules, not equivalences. That is, we'd like to write equivalences only in the assumptions and let the inference rules do all of the deductive "work" for us. So to turn $\neg\neg W$ into W, we'll use a special rule:

Inference Rule 3 (Double Negation Elimination)**.**

$$\dfrac{\neg\neg\varphi}{\varphi}\,(\neg\neg\text{E})$$

The Rule in 3 expresses in syntactic terms what we could easily verify using truth tables–that the truth value of any PL sentence φ is the same as its double negation $\neg\neg\varphi$. With Rules 1-3 in place, we have everything we need in order to give a formal proof representing the deductive process going on in (1). Figure 2 shows this proof.

$$\dfrac{\dfrac{\dfrac{\dfrac{}{\neg S \to \neg\neg W}\,(\text{Hyp}) \qquad \dfrac{}{\neg S}\,(\text{Hyp})}{\neg\neg W}\,(\to\text{E})}{W}\,(\neg\neg\text{E}) \qquad \dfrac{}{W \to C}\,(\text{Hyp})}{C}\,(\to\text{E})$$

Figure 2: Full proof of the argument in (1).

Homework

Problem 1. Consider the following informal argument:

(2) a. If a man owns a Lake Wobegon car dealership, he is the brother of the man who owns the other Lake Wobegon car dealership.
b. Clint owns Lake Wobegon's Ford dealership.
c. Clarence owns Lake Wobegon's Chevy dealership.
d. Clint and Clarence are brothers.

Answer the following questions about the argument in (2):

a. Which propositions expressed by this argument are premises, and which are the conclusion?

b. Write down all the propositions expressed in the argument.

c. If you needed to diagram the pattern of reasoning reflected in this argument, which inference rules from our ND rules would probably be used?

Problem 2. The following informal argument is a slight simplification of the argument in (2):

(3) a. If a man owns a Lake Wobegon car dealership, he is the brother of the man who owns the other Lake Wobegon car dealership.

b. Clint owns a Lake Wobegon car dealership.

c. Clint's brother owns Lake Wobegon's other car dealership.

Translate the informal argument in (3) into a formal proof using the rules introduced in Rules 1-3.