CS 3110: Proof Strategy and Examples

1   Propositional Logic Proof Strategy

The fundamental thing you have to do is figure out where each connective is going to come from. Sometimes the answer is very simple; other times, it may require a substantial detour.

If you are trying to *prove* some formula involving:
$\Rightarrow$: You can usually assume the antecedent, prove the consequent, and use $\Rightarrow I$.
$\wedge$: You can almost always just prove both sides.
$\vee$: If you can prove one of the sides, you can use $\vee I$. If you can't prove either side alone, it becomes significantly harder, and you often need to use something like RAA.
$\neg$: You can often prove that the non-negated form leads to a contradiction, then use $\Rightarrow I$.
Anything: You can try to produce $P \Rightarrow Q$, where P is something you have and Q is either what you want to show or something that will help you.
Anything: You can assume the negation and try to use RAA.

If you are trying to *use* some formula involving:
$\Rightarrow$: You usually want to prove the antecedent and use $\Rightarrow E$ to get the consequent.
$\wedge$: You usually want to break it up into the two subformulas using $\wedge E$.
$\vee$: You usually want to use $\vee E$, which is a complicated rule that requires two other branches.
$\neg$: If you can prove the thing being negated, you can derive a contradiction using $\Rightarrow E$.
Anything: Keep in mind that sometimes you don't need to break down formulas at all. If a whole huge formula appears as the antecedent of a $\Rightarrow$ somewhere, once you get that huge formula you probably don't want to break it up.

When other approaches fail, it is often worthwhile to assume the negation of what you're trying to prove, and try to derive a contradiction and then use RAA.

2   A Proof Walkthrough

Consider trying to prove $\vdash (P \vee Q) \Rightarrow \neg(\neg P \wedge \neg Q)$. We can see that the main connective is a $\Rightarrow$, so our strategy will be to prove $\neg(\neg P \wedge \neg Q)$ by assuming $P \vee Q$. So we have two proof components, one which must be the end of the proof, and another that we can use somewhere in the proof.

$$\cfrac{\overline{[x : P \vee Q]}A \quad \cfrac{\cfrac{\cdots}{\neg(\neg P \wedge \neg Q)}?}{P \vee Q \Rightarrow \neg(\neg P \wedge \neg Q)}}{}\Rightarrow I, x$$

Since $\neg(\neg P \wedge \neg Q)$ has a negation as its primary connection, we try the approach of assuming $\neg P \wedge \neg Q$ and deriving $\bot$. So now we have one primary component of the proof tree that starts with $\bot$ and ends at the conclusion, and two assumptions that we can make

use of to get to $\bot$.

$$\cfrac{}{[x : P \lor Q]}A \qquad \cfrac{}{[y : \neg P \land \neg Q]}A \qquad \cfrac{\cfrac{\cfrac{\cdots ?}{\bot}}{\cfrac{\neg(\neg P \land \neg Q)}{P \lor Q \Rightarrow \neg(\neg P \land \neg Q)} \Rightarrow I, y}}{} \Rightarrow I, x$$

So we have $\neg P \land \neg Q$ and $P \lor Q$, and we want to get $\bot$. We are out of useful things to assume at this point, so we must make use of what we have. Looking at the two formulas we have to work with, splitting $\neg P \land \neg Q$ into $\neg P$ and $\neg Q$ immediately seems to be a dead end - if we do that, there's not really anywhere to go. Using $P \lor Q$, on the other hand, is promising. The $\lor E$ rule can be used to get $\bot$, if we can show show that both P and Q lead to $\bot$. So we can link one of our assumptions into the main proof now, giving us a main proof that has two subproofs missing and one other assumption we can use somewhere.

$$\cfrac{}{[y : \neg P \land \neg Q]}A \qquad \cfrac{\cfrac{\cfrac{}{[x : P \lor Q]}A \quad \cfrac{\overline{P}^A}{\cfrac{\cdots ?}{\bot}}? \quad \cfrac{\overline{Q}^A}{\cfrac{\cdots ?}{\bot}}?}{\cfrac{\bot}{\cfrac{\neg(\neg P \land \neg Q)}{} \Rightarrow I, y}} \lor E}{} \Rightarrow I, x$$

So now we need to fill in the two subproofs, one that starts with P and ends with $\bot$, and the other that starts with Q and ends with $\bot$. But since we can assume $\neg P \land \neg Q$ wherever we want, this is easy - we can assume it in both subproofs. In the P subproof, we can use the $\neg P$ part to get from P to $\bot$, and similarly for Q. So we can now assemble all the pieces into a complete proof.

$$\cfrac{\cfrac{}{[x : P \lor Q]}A \quad \cfrac{\cfrac{\cfrac{}{[y_1 : \neg P \land \neg Q]}A}{P \Rightarrow \bot} \land E \quad \cfrac{}{[z_1 : P]}A}{\bot} \Rightarrow E \quad \cfrac{\cfrac{\cfrac{}{[y_2 : \neg P \land \neg Q]}A}{Q \Rightarrow \bot} \land E \quad \cfrac{}{[z_2 : Q]}A}{\bot} \Rightarrow E}{\cfrac{\cfrac{\bot}{\cfrac{\neg(\neg P \land \neg Q)}{P \lor Q \Rightarrow \neg(\neg P \land \neg Q)} \Rightarrow I, x}}{} \Rightarrow I, y_1, y_2} \lor E, z_1, z_2}$$

Note that we actually to have a $\neg P \land \neg Q$ assumption twice. This is necessary for the proof, but works out due to the fact that using $\Rightarrow I$ to go from ψ to $\phi \Rightarrow \psi$ can discharge any number of assumptions of the formula ϕ.

3 Sample Propositional Logic Proofs

3.1 Double Negation Elimination (DNE)

Goal: $\vdash \neg\neg P \Rightarrow P$. Recall that $\neg P$ is simply syntactic sugar for $P \Rightarrow \bot$.

$$
\cfrac{
  \cfrac{
    \cfrac{
      \cfrac{\overline{[x : (P \Rightarrow \bot) \Rightarrow \bot]}^{A} \quad \overline{[y : P \Rightarrow \bot]}^{A}}{\bot} \Rightarrow E
    }{P} RAA, y
  }{\neg\neg P \Rightarrow P} \Rightarrow I, x
}{}
$$

3.2 Rule of the Excluded Middle (REM)

Goal: $\vdash P \vee \neg P$.

$$
\cfrac{
  \cfrac{
    \cfrac{
      \overline{[y : \neg(P \vee \neg P)]}^{A} \quad \cfrac{\overline{[x : P]}^{A}}{P \vee \neg P} \vee I
    }{\bot} \Rightarrow E
  }{\neg P} \Rightarrow I, x
  \quad \vee I \quad \overline{[z : \neg(P \vee \neg P)]}^{A}
}{P \vee \neg P} RAA, y, z
$$

with intermediate line $\cfrac{}{P \vee \neg P} \vee I$

3.3 The Contrapositive

Goal: $\vdash (P \Rightarrow Q) \Rightarrow (\neg Q \Rightarrow \neg P)$.

$$
\cfrac{
  \cfrac{
    \cfrac{
      \overline{[z : \neg Q]}^{A} \quad \cfrac{\overline{[x : P \Rightarrow Q]}^{A} \quad \overline{[y : P]}^{A}}{Q} \Rightarrow E
    }{\bot} \Rightarrow E
  }{\neg P} \Rightarrow I, y
  }{\cfrac{\neg Q \Rightarrow \neg P}{(P \Rightarrow Q) \Rightarrow (\neg Q \Rightarrow \neg P)} \Rightarrow I, x} \Rightarrow I, z
$$

3.4 Peirce's Law

Goal: $\vdash ((P \Rightarrow Q) \Rightarrow P) \Rightarrow P$

$$
\cfrac{
  \cfrac{
    \cfrac{
      \overline{[y : (P \Rightarrow Q) \Rightarrow P]}^{A} \quad \cfrac{\cfrac{\overline{[x : \neg P]}}{\neg Q \Rightarrow \neg P}}{P \Rightarrow Q} Contrapositive
    }{P} \Rightarrow E \quad \overline{[z : \neg P]}^{A}
  }{\cfrac{\bot}{P} RAA, x, z} \Rightarrow E
}{((P \Rightarrow Q) \Rightarrow P) \Rightarrow P} \Rightarrow I, y
$$

4   Predicate Logic Proof Strategy

Due to the quantifiers, predicate logic proofs are often harder to write than propositional logic proofs. The rules for quantifiers are less intuitive than the rules for propostional connections. Additionally, the side conditions on the quantifier rules can be tricky.

In general, getting rid of $\forall$s is easy, and so is getting $\exists$s is easy - the hard part is getting $\forall$s and getting rid of $\exists$s. To get a $\forall$, you generally need to start with a $\forall$, apply $\forall E$, manipulate the inner formula in some fashion, and eventually apply $\forall I$ to put the $\forall$ back on. To get rid of a $\exists x$, you need to use the formula inside the $\exists$ to prove something in which x is not free (so it must either not involve x, or x must be bound by some other quantifier).

5   Side Conditions and Invalid Proofs in Predicate Logic

For proofs involving quantifiers, great care must be taken to not violate the side conditions on the $\forall I$ and $\exists E$ rules.

When we use $\forall I$ to introduce $\forall x$, x must not be free in any of the undischarged assumptions. When we use $\exists E$ to get rid of $\exists x$ and conclude ϕ, x must not be free in any of the undischarged assumptions *or* the conclusion ϕ.

If we aren't careful, it is easy to produce invalid proofs. For example, consider the following.

$$
\cfrac{
\cfrac{
\cfrac{
\cfrac{
\cfrac{
\cfrac{[x : \exists x \forall y P(x,y)]^A \quad [y : \forall y P(c,y)]^A}{\forall y P(c,y)} \exists E, y
}{P(c,d)} \forall E
}{\forall x P(x,d)} \forall I
}{\forall y \forall x P(x,y)} \forall I
}{\exists x \forall y P(x,y) \Rightarrow \forall y \forall x P(x,y)} \Rightarrow I, x
}
$$

This proof is clearly false, but what's wrong with it? Most of the steps look reasonable at first glance. You might suspect that one of the $\forall I$ steps is invalid, but in fact, all the $\forall I$ steps are valid; they only introduce variables that are not free in the undischarged assumptions. Even though they are the steps that seem to be doing something unreasonable, the root of the problem lies somewhere else.

As it turns out, the problem is right at the start of the proof, in the use of $\exists E$. One of the side conditions on the use of the $\exists E$ rule is that the specific object we use in place of the variable is not free in the conclusion. However, we are using the specific object c in our assumption $y : \forall y P(c,y)$, and c is free in the conclusion $\forall y P(c,y)$, so this use of $\exists E$ is invalid.

It is easy to fall into traps like this when using predicate logic: one small oversight at the beginning of a proof can result in bizarre conclusions being reached later on. This can make it hard to track down the source of the error, which is why it is important to verify the side conditions every time you use $\exists E$ or $\forall I$.

6 The Rule of Induction

Having a formal system of logic allows us to now formalize the concept of induction by introducing it as another inference rule. Letting the domain of our quantifiers be the natural numbers $\{0, 1, 2, ...\}$, we can write the rule of induction as:

$$\frac{\forall n P(n) \Rightarrow P(n+1) \quad P(0)}{\forall n P(n)}$$

7 Sample Predicate Logic Proofs

If you really want to see if you understand predicate logic, go through every use of $\forall I$ and $\exists E$ and figure out precisely why the side conditions are satisfied!

7.1 $\neg \exists x P(x) \Rightarrow \forall x \neg P(x)$

$$\cfrac{\cfrac{\cfrac{[x : \neg \exists x P(x)]^A \quad \cfrac{\cfrac{[y : P(x)]^A}{\exists x P(x)}\exists I}{\bot}\Rightarrow E}{\neg P(x)}\Rightarrow I, y}{\forall x \neg P(x)}\forall I}{\neg \exists x P(x) \Rightarrow \forall x \neg P(x)}\Rightarrow I, x$$

7.2 $\exists x \forall y P(x, y) \Rightarrow \forall y \forall x P(x, y)$

Goal: $\vdash \exists x \forall y P(x, y) \Rightarrow \forall y \forall x P(x, y)$.

$$\cfrac{\cfrac{[x : \exists x \forall y P(x,y)]^A \quad \cfrac{\cfrac{\cfrac{\cfrac{[y : \forall y P(x, y)]^A}{P(x, y)}\forall E}{\exists x P(x, y)}\exists I}{\forall y \exists x P(x, y)}\forall I}{\forall y \exists x P(x, y)}\exists E}{\exists x \forall y P(x, y) \Rightarrow \forall y \exists x P(x, y)}\Rightarrow I, x$$

7.3 $\forall n \; 0 + 1 + ... + n = \frac{(n+1)(n)}{2}$

Let $P(n)$ be the predicate $0 + 1 + ... + n = \frac{(n+1)(n)}{2}$. Note that $P(0)$ is equivalent to $0 = \frac{(0+1)(0)}{2}$

$$\cfrac{\cfrac{0=0}{0=\frac{(0+1)(0)}{2}}\text{reflexivity}}{}\text{algebra} \qquad \cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\overline{[x:0+...+n=\frac{(n+1)(n)}{2}]}A}{0+...+n+n+1=\frac{(n+1)(n)+2n+2}{2}}\text{algebra}}{0+...+n+n+1=\frac{n^2+3n+2}{2}}\text{algebra}}{0+...+n+n+1=\frac{(n+2)(n+1)}{2}}\text{algebra}}{P(n)\Rightarrow P(n+1)}\Rightarrow I,x}{\forall nP(n)\Rightarrow P(n+1)}\forall I$$

$$\cfrac{}{\forall nP(n)}\text{induction}$$

8 The Mechanics of Writing Proofs

Make sure that all steps in your proof tree have what rule they use written to the right of the step.

If a proof is especially long or uses the same set of rules in the same sequence multiple times, you may be better off breaking the proof into subproofs.

When writing out proofs by hand, it is best to first figure out the proof on scrap paper, then make a clean copy. Proofs quickly become very hard to read when there are mistakes and crossed-out lines.

Typing out proofs in most word processors becomes unmanageable as the proof becomes large. The best option for typing your proofs and having them stay manageable is to use LaTeX, a popular typesetting system. One good implementation of LaTeX is MiKTeX, available from `http://miktex.org/`. A good IDE for working with LaTeX is TeXnicCenter, available from `http://www.toolscenter.org/`.

9 Further Exercises

If you want to practice logic, try the following proofs. Recall that $\phi \Leftrightarrow \psi$ is simply syntactic sugar for $(\phi \Rightarrow \psi) \wedge (\psi \Rightarrow \phi)$. In most cases involving $\Leftrightarrow$, you are best off breaking the problem into two subproofs, one for the $\Rightarrow$ direction and one for the $\Leftarrow$ direction.

$$\vdash ((P \wedge Q) \Rightarrow R) \Leftrightarrow (P \Rightarrow (Q \Rightarrow R))$$
$$\vdash P \Rightarrow Q \Rightarrow P$$
$$\vdash A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$$
$$\vdash A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$$
$$\vdash A \Rightarrow ((A \wedge B) \vee (A \wedge \neg B))$$
$$\vdash (A \Rightarrow \exists x P(x)) \Leftrightarrow \exists x (A \Rightarrow P(x))$$
$$\vdash ((\forall x P(x)) \Rightarrow Q) \Leftrightarrow \exists x (P(x) \Rightarrow Q)$$
$$\vdash \exists x \forall y (P(x) \Rightarrow P(y))$$