Proof by Contradiction

Scott Martin

March 5, 2012

Not So

- Notice that, so far, our natural deduction (ND) rules have only captured the ‘true’ cases in the truth tables for propositional logic (PL).

Not So

- Notice that, so far, our natural deduction (ND) rules have only captured the 'true' cases in the truth tables for propositional logic (PL).
- This could be seen as a drawback: what about the 'false' cases?

Not So

- Notice that, so far, our natural deduction (ND) rules have only captured the ‘true’ cases in the truth tables for propositional logic (PL).
- This could be seen as a drawback: what about the ‘false’ cases?
- That is, if the only way to say that something’s not provable is just to say that there’s no proof of it, how do we know it’s not the case that we just haven’t yet been clever enough to *find* a valid proof of it?

Not So

- Notice that, so far, our natural deduction (ND) rules have only captured the 'true' cases in the truth tables for propositional logic (PL).
- This could be seen as a drawback: what about the 'false' cases?
- That is, if the only way to say that something's not provable is just to say that there's no proof of it, how do we know it's not the case that we just haven't yet been clever enough to *find* a valid proof of it?
- So it would sometimes be nice to be able to say when an argument is definitely not valid.

Not So

- Notice that, so far, our natural deduction (ND) rules have only captured the ‘true’ cases in the truth tables for propositional logic (PL).
- This could be seen as a drawback: what about the ‘false’ cases?
- That is, if the only way to say that something’s not provable is just to say that there’s no proof of it, how do we know it’s not the case that we just haven’t yet been clever enough to *find* a valid proof of it?
- So it would sometimes be nice to be able to say when an argument is definitely not valid.
- For us, this will mean making one of the premises false. (Remember that an argument can’t be valid if we know for sure that one of the premises *can’t* be true no matter how things are.)

Not So

- Notice that, so far, our natural deduction (ND) rules have only captured the ‘true’ cases in the truth tables for propositional logic (PL).
- This could be seen as a drawback: what about the ‘false’ cases?
- That is, if the only way to say that something’s not provable is just to say that there’s no proof of it, how do we know it’s not the case that we just haven’t yet been clever enough to *find* a valid proof of it?
- So it would sometimes be nice to be able to say when an argument is definitely not valid.
- For us, this will mean making one of the premises false. (Remember that an argument can’t be valid if we know for sure that one of the premises *can’t* be true no matter how things are.)
- This corresponds to what we do in real-world reasoning: it’s referred to as “the process of elimination” or “proof by contradiction.”

Hitting Bottom

- We say an argument is a contradiction when one of the premises (that is, one of the things assumed to be true) can be demonstrated to be false.

Hitting Bottom

- We say an argument is a contradiction when one of the premises (that is, one of the things assumed to be true) can be demonstrated to be false.
- Since this process revolves around contradiction, we'll need a new (nullary) connective to represent this contradiction, symbolized by $\bot$.

Hitting Bottom

- We say an argument is a contradiction when one of the premises (that is, one of the things assumed to be true) can be demonstrated to be false.
- Since this process revolves around contradiction, we'll need a new (nullary) connective to represent this contradiction, symbolized by $\bot$.
- Our universe only contains propositions, so $\bot$ is still a proposition–it's just a very special proposition that is *always* false.

Hitting Bottom

- We say an argument is a contradiction when one of the premises (that is, one of the things assumed to be true) can be demonstrated to be false.
- Since this process revolves around contradiction, we'll need a new (nullary) connective to represent this contradiction, symbolized by $\perp$.
- Our universe only contains propositions, so $\perp$ is still a proposition–it's just a very special proposition that is *always* false.
- To see proof by contradiction in action, consider the argument in (1).

(1)
  a. If Clint doesn't go fishing, he doesn't eat Walleye for dinner.
  b. Clint didn't go fishing.
  c. Clint is eating Walleye for dinner.
  d. Therefore, Clint must have gone fishing.

Negation Rules I

- Notice that the argument in (1) must be invalid. At least one of the premises must be false.
- To reflect this reasoning pattern, we need two new rules.

Inference Rule 12 (Negation Introduction)

In non-sequent style:

$$\begin{array}{c} \dfrac{}{[\varphi]_i}\,(\text{Hyp}) \\ \vdots \\ i\ \dfrac{\bot}{\neg\varphi}\,(\neg\text{I}) \end{array}$$

In sequent style:

$$\frac{\Gamma, \varphi \vdash \bot}{\Gamma \vdash \neg\varphi}\,(\neg\text{I})$$

Negation Rules II

Inference Rule 13 (Negation Elimination)

In non-sequent style:

$$\frac{\varphi \qquad \neg\varphi}{\bot}\,(\neg\mathrm{E})$$

In sequent style:

$$\frac{\Gamma \vdash \varphi \qquad \Delta \vdash \neg\varphi}{\Gamma, \Delta \vdash \bot}\,(\neg\mathrm{E})$$

The Rules Elaborated

- Rule 12 corresponds to what's going on when we use "process of elimination"-type reasoning: from a contradictory conclusion (represented by $\perp$), we know at least one of the premises must have been false.

The Rules Elaborated

- Rule 12 corresponds to what's going on when we use "process of elimination"-type reasoning: from a contradictory conclusion (represented by $\perp$), we know at least one of the premises must have been false.
- Rule 13 is a ND representation of what we earlier referred to as the "Law of Excluded Middle": every proposition is either true or false, but not both.

The Rules Elaborated

- Rule 12 corresponds to what's going on when we use "process of elimination"-type reasoning: from a contradictory conclusion (represented by $\perp$), we know at least one of the premises must have been false.
- Rule 13 is a ND representation of what we earlier referred to as the "Law of Excluded Middle": every proposition is either true or false, but not both.
- If it seems strange to call rules 12 and 13 Negation Introduction and Elimination, try thinking of $\neg\varphi$ as the implication $\varphi \to \perp$.

The Rules Elaborated

- Rule 12 corresponds to what's going on when we use "process of elimination"-type reasoning: from a contradictory conclusion (represented by $\perp$), we know at least one of the premises must have been false.
- Rule 13 is a ND representation of what we earlier referred to as the "Law of Excluded Middle": every proposition is either true or false, but not both.
- If it seems strange to call rules 12 and 13 Negation Introduction and Elimination, try thinking of $\neg\varphi$ as the implication $\varphi \rightarrow \perp$.
- That is, if we take $\neg\varphi$ to be a synonym for $\varphi \rightarrow \perp$, then Rules 12 and 13 are just special cases of Implication Introduction and Elimination, respectively (and we didn't really even have to write them!).

Example Proof by Contradiction I

- As an example, let $\neg F$ be the proposition expressed by (1b), and W be the one denoted by (1c). Then Figure 1 shows a proof (in both non-sequent and sequent styles) of the argument in (1).

$$
\begin{prooftree}
\AxiomC{}
\RightLabel{(Hyp)}
\UnaryInfC{$\neg F \to \neg W$}
\AxiomC{}
\RightLabel{(Hyp)}
\UnaryInfC{$[\neg F]_1$}
\RightLabel{($\to$E)}
\BinaryInfC{$\neg W$}
\AxiomC{}
\RightLabel{(Hyp)}
\UnaryInfC{W}
\RightLabel{($\neg$E)}
\BinaryInfC{$\bot$}
\LeftLabel{1}
\RightLabel{($\neg$I)}
\UnaryInfC{$\neg\neg F$}
\RightLabel{($\neg\neg$E)}
\UnaryInfC{F}
\end{prooftree}
$$

Example Proof by Contradiction II

$$\cfrac{\cfrac{\cfrac{\cfrac{\neg F \to \neg W \vdash \neg F \to \neg W \qquad \neg F \vdash \neg F}{\neg F \to \neg W, \neg F \vdash \neg W}\,(\to\mathrm{E}) \qquad W \vdash W}{\neg F \to \neg W, \neg F, W \vdash \bot}\,(\neg\mathrm{E})}{\neg F \to \neg W, W \vdash \neg\neg F}\,(\neg\mathrm{I})}{\neg F \to \neg W, W \vdash F}\,(\neg\neg\mathrm{E})$$

Figure 1: Proof of the argument in (1).

Notably False

Some important things about this example of proof by contradiction:

- In the proof in Figure 1, notice that we could negate *any* of the premises.

Notably False

Some important things about this example of proof by contradiction:

- In the proof in Figure 1, notice that we could negate *any* of the premises.
- That is, we could have said that W was actually false, or that the implication $\neg F \rightarrow \neg W$ was actually false instead of saying that $\neg F$ was false.

Notably False

Some important things about this example of proof by contradiction:

- In the proof in Figure 1, notice that we could negate *any* of the premises.
- That is, we could have said that W was actually false, or that the implication $\neg F \to \neg W$ was actually false instead of saying that $\neg F$ was false.
- This corresponds to the "process of elimination" because we have a choice of which of the premises to say was the one giving rise to the inconsistency.

Notably False

Some important things about this example of proof by contradiction:

- In the proof in Figure 1, notice that we could negate *any* of the premises.
- That is, we could have said that W was actually false, or that the implication $\neg F \to \neg W$ was actually false instead of saying that $\neg F$ was false.
- This corresponds to the "process of elimination" because we have a choice of which of the premises to say was the one giving rise to the inconsistency.
- With Rules 12 and 13, we have everything we need to prove that (for any PL propositions A and B) $A \to B$ is equivalent to both $\neg(A \wedge \neg B)$ and $\neg B \to \neg A$. Partial proof is available in Figures 2, 3, and 4 in both non-sequent (first) and sequent (second) styles.

Equivalence Proofs based on Contradiciton I

$$
2\;\dfrac{1\;\dfrac{\dfrac{\dfrac{\dfrac{}{A \to B}\,(\text{Hyp}) \qquad \dfrac{}{[A]_1}\,(\text{Hyp})}{B}\,(\to\text{E}) \qquad \dfrac{}{[\neg B]_2}\,(\text{Hyp})}{\bot}\,(\neg\text{E})}{\neg A}\,(\neg\text{I})}{\neg B \to \neg A}\,(\to\text{I})
$$

$$
\dfrac{\dfrac{\dfrac{\dfrac{A \to B \vdash A \to B \qquad A \vdash A}{A \to B, A \vdash B}\,(\to\text{E}) \qquad \neg B \vdash \neg B}{A \to B, A, \neg B \vdash \bot}\,(\neg\text{E})}{A \to B, \neg B \vdash \neg A}\,(\neg\text{I})}{A \to B \vdash \neg B \to \neg A}\,(\to\text{I})
$$

Figure 2: Proof of $(\neg B \to \neg A)$ from $A \to B$.

Equivalence Proofs based on Contradiciton II

$$
{\scriptstyle 2}\,\dfrac{\dfrac{{\scriptstyle 1}\,\dfrac{\dfrac{\dfrac{\dfrac{}{\neg B \to \neg A}\,(\mathrm{Hyp}) \qquad \dfrac{}{[\neg B]_1}\,(\mathrm{Hyp})}{\neg A}\,(\to\mathrm{E}) \qquad \dfrac{}{[A]_2}\,(\mathrm{Hyp})}{\bot}\,(\neg\mathrm{E})}{\neg\neg B}\,(\neg\mathrm{I})}{B}\,(\neg\neg E)}{A \to B}\,(\to\mathrm{I})
$$

$$
\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\neg B \to \neg A \vdash \neg B \to \neg A \qquad \neg B \vdash \neg B}{\neg B \to \neg A, \neg B \vdash \neg A}\,(\to\mathrm{E}) \qquad A \vdash A}{\neg B \to \neg A, \neg B, A \vdash \bot}\,(\neg\mathrm{E})}{\neg B \to \neg A, A \vdash \neg\neg B}\,(\neg\mathrm{I})}{\neg B \to \neg A, A \vdash B}\,(\neg\neg E)}{\neg B \to \neg A \vdash A \to B}\,(\to\mathrm{I})
$$

Figure 3: Proof of $A \to B$ from $\neg B \to \neg A$.

Equivalence Proofs based on Contradiciton III

$$
\begin{prooftree}
\AxiomC{}
\RightLabel{(Hyp)}
\UnaryInfC{$A \to B$}
\AxiomC{}
\RightLabel{(Hyp)}
\UnaryInfC{$[A \wedge \neg B]_1$}
\RightLabel{($\wedge$E$_1$)}
\UnaryInfC{A}
\RightLabel{($\to$E)}
\BinaryInfC{B}
\AxiomC{}
\RightLabel{(Hyp)}
\UnaryInfC{$[A \wedge \neg B]_1$}
\RightLabel{($\wedge$E$_2$)}
\UnaryInfC{$\neg B$}
\RightLabel{($\neg$E)}
\BinaryInfC{$\bot$}
\LeftLabel{1}
\RightLabel{($\neg$I)}
\UnaryInfC{$\neg(A \wedge \neg B)$}
\end{prooftree}
$$

$$
\begin{prooftree}
\AxiomC{$A \to B \vdash A \to B$}
\AxiomC{$A \wedge \neg B \vdash A \wedge \neg B$}
\RightLabel{($\wedge$E$_1$)}
\UnaryInfC{$A \wedge \neg B \vdash A$}
\RightLabel{($\to$E)}
\BinaryInfC{$A \to B, A \wedge \neg B \vdash B$}
\AxiomC{$A \wedge \neg B \vdash A \wedge \neg B$}
\RightLabel{($\wedge$}
\UnaryInfC{$A \wedge \neg B \vdash \neg B$}
\RightLabel{($\neg$E)}
\BinaryInfC{$A \to B, A \wedge \neg B \vdash \bot$}
\RightLabel{($\neg$I)}
\UnaryInfC{$A \to B \vdash \neg(A \wedge \neg B)$}
\end{prooftree}
$$

Figure 4: Proof of $\neg(A \wedge \neg B)$ from $A \to B$.

Exercises

These problems are additional problems for Problem Set 4. Any work toward completing these problems will count as bonus points on Problem Set 4 (you can turn in either or both of these as part of it).

Problem 10 (Bonus)

Finish the proof that $A \to B$ is equivalent to $\neg(A \wedge \neg B)$ by proving that assuming $\neg(A \wedge \neg B)$ by itself leads to $A \to B$.

Problem 11 (Bonus)

Complete the actual proofs of equivalence. That is, give a proof that $\vdash (A \to B) \leftrightarrow (\neg B \to \neg A)$ and a proof that $\vdash (A \to B) \leftrightarrow \neg(A \wedge \neg B)$.