Lecture Notes on
Classical Logic

15-317: Constructive Logic
William Lovas

Lecture 7
September 15, 2009

1   Introduction

In this lecture, we design a judgmental formulation of classical logic. To gain an intuition, we explore various equivalent notions of the essence of classical reasoning including the Law of the Excluded Middle and Double-Negation Elimination. Throughout the discussion a common theme is the indirectness and "dishonesty" of classical proofs, an idea which will later be key to understanding their computational interpretation. Eventually, we arrive at a judgmentally parsimonious system based on the principle of Proof by Contradiction and founded on two new forms of judgment: *A is false* (written *A false*) and *contradiction* (written #).

2   Example

Classical reasoning is pervasive in classical mathematics, so we begin with a typical example of a theorem proven using classical methods.

> **Theorem:** $\exists a, b \in \mathbb{R}.\ irrational(a) \wedge irrational(b) \wedge rational(a^b)$
>
> **Proof:** Consider $\sqrt{2}^{\sqrt{2}}$: this number is either rational or irrational. Suppose it is rational: then $a = \sqrt{2}$, $b = \sqrt{2}$ gives the required result. Suppose it is not: then $a = \sqrt{2}^{\sqrt{2}}$, $b = \sqrt{2}$ gives the required result, as $a^b = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^{\sqrt{2}\cdot\sqrt{2}} = (\sqrt{2})^2 = 2$.

Although this claims to be a proof of an existential theorem, it gives the reader no grasp on what numbers actually witness the result—it offers two possible choices but no further guidance as to which one is correct.

A slightly whimsical characterization of the proof offers a glimpse into the computational content of classical reasoning. Imagine a prominent mathematician delivering the above proof aloud as part of a lecture before a large audience. Initially, he delivers only the first half, saying, "Let $a = \sqrt{2}$ and $b = \sqrt{2}$; then a and b are irrational while a^b is rational—*trust me*." Amidst some mumbling, the audience nods and accepts the proof; after all, he *is* a very prominent mathematician, and he probably knows what he's talking about. But then, halfway through the lecture, a student from the back suddenly leaps up and exclaims, "His proof is no good—*I have a proof that $\sqrt{2}^{\sqrt{2}}$ is irrational!*"[1]. The audience gasps and a murmur runs through the crowd, but before anyone else can speak, the mathematician calmly responds, "May I see your proof?" After checking it over, the mathematician addresses the crowd again: "My apologies—I misspoke earlier. What I meant to say was this: Let $a = \sqrt{2}^{\sqrt{2}}$ and $b = \sqrt{2}$; then a and b are irrational—I have this proof, if you don't believe me!—while $a^b = 2$ and is therefore rational."

The poor student at the back thought she could attain fame and fortune by debunking the prominent mathematician, but in fact, the mathematician stole the glory by leveraging her proof for his own ends. Classical proofs exhibit a similar time-traveling behavior when executed, as we'll see in the next lecture.

3   What is classical logic?

Classical logic can be characterized by a number of equivalent axioms:

- Proof by Contradiction

- Law of the Excluded Middle: $A \vee \neg A$

- Double-Negation Elimination: $\neg\neg A \supset A$

- Peirce's Law: $((A \supset B) \supset A) \supset A$

We might consider making the logic we've seen so far classical by adding one or more rules that correspond to these axioms. For instance, we might

---

[1]It is, in fact, though the proof is non-trivial.

add:

$$\frac{}{A \vee \neg A \; true} \; LEM \quad \text{or} \quad \frac{\neg\neg A \; true}{A \; true} \; DNE.$$

Of course, we need only add one or the other, and not both, since they are interderivable. First, let's show how we can derive DNE from LEM:

$$\frac{\dfrac{}{A \vee \neg A \; true} \; LEM \quad \dfrac{}{A \; true} \; u \quad \dfrac{\dfrac{\neg\neg A \; true \quad \overline{\neg A \; true} \; v}{\perp \; true} \; \supset E}{\dfrac{A \; true}{} \; \perp E}}{A \; true} \; \vee E^{u,v}$$

As you can see, it is precisely the power of having the middle excluded that lets us turn our proof of $\neg\neg A \; true$ into a proof of $\perp \; true$, and thus a proof of $A \; true$ as required.

Using DNE also allows us to derive LEM. Note that since the LEM rule has no premises, this will have to be a completely closed derivation. To help elucidate the thought process of a classical prover, we'll do a step-by-step derivation. We start bottom up from the conclusion:

$$\vdots$$
$$A \vee \neg A \; true$$

If we were to attempt to proceed as we've done previously, we would now have to apply a disjunction-introduction rule, but we have no way of deciding which injection to choose: merely positing all propositions to be true or not does nothing to make apparent which is the case. We know that this classical tautology is unprovable intuitionistically, anyhow, so we have no hope but to begin by employing our classical rule, double-negation elimination.

$$\vdots$$
$$\frac{\neg\neg(A \vee \neg A) \; true}{A \vee \neg A \; true} \; DNE$$

Since $\neg A$ is notationally defined to mean $A \supset \bot$, we can take the usual no-brainer step of applying implication introduction:

$$\cfrac{\cfrac{\cfrac{\overline{\neg(A \vee \neg A) \ true}^{\ u}}{\vdots}}{\cfrac{\bot \ true}{\neg\neg(A \vee \neg A) \ true} \supset I^u}}{A \vee \neg A \ true} DNE$$

Now we are tasked with proving falsehood true: a difficult task, to be sure, but a glimmer of hope shines through in our sole assumption u. Perhaps by applying it to an appropriate argument, we can conclude $\bot$ *true* as required:

$$\cfrac{\cfrac{\cfrac{\overline{\neg(A \vee \neg A) \ true}^{\ u} \quad \cfrac{\overline{\neg(A \vee \neg A) \ true}^{\ u}}{\cfrac{\vdots}{A \vee \neg A \ true}}}{\bot \ true} \supset E}{\cfrac{\neg\neg(A \vee \neg A) \ true}{A \vee \neg A \ true} DNE} \supset I^u}{}$$

At this point, if we were not being careful, we might throw up our hands and quit. We're right back to where we started, trying to prove $A \vee \neg A$ *true*! But now we have an extra hypothesis to help us. Undaunted, we pause to consider: which shall we prove, A or $\neg A$? Since we know nothing of the structure of A, we have no hope of proving it unless an assumption can yield it, but nothing seems appropriate. So instead, we try for the right disjunct, $\neg A$. We can also go ahead and apply implication introduction without thinking twice, this time attaining an assumption v that A *true*.

$$\cfrac{\cfrac{\overline{\neg(A \vee \neg A) \ true}^{\ u} \quad \cfrac{\cfrac{\cfrac{\overline{\neg(A \vee \neg A) \ true}^{\ u}, \ \overline{A \ true}^{\ v}}{\vdots}}{\cfrac{\bot \ true}{\neg A \ true} \supset I^v}}{A \vee \neg A \ true} \vee I_R}{\bot \ true} \supset E}{\cfrac{\neg\neg(A \vee \neg A) \ true}{A \vee \neg A \ true} DNE} \supset I^u$$

This situation is familiar once again: to prove $\perp$ *true*, we must use an assumption, and u still seems to be the only one that can help. So we carry out another implication elimination, feeling only the slightest sense of *deja vu...*

$$
\cfrac{
  \cfrac{\hphantom{\neg(A \vee \neg A)\ true}}{\neg(A \vee \neg A)\ true}\ u
  \quad
  \cfrac{
    \cfrac{
      \cfrac{\hphantom{\neg(A \vee \neg A)\ true}}{\neg(A \vee \neg A)\ true}\ u
      \quad
      \cfrac{
        \cfrac{
          \cfrac{\dfrac{\overline{\neg(A \vee \neg A)\ true}\ u\ , \ \overline{A\ true}\ v}{\vdots}}{A \vee \neg A\ true}
        }{\perp\ true}\ \supset I^v
        \cfrac{}{\neg A\ true}
      }{A \vee \neg A\ true}\ \vee I_R
    }{\perp\ true}\ \supset E
  }{
    \cfrac{\perp\ true}{\cfrac{\neg\neg(A \vee \neg A)\ true}{A \vee \neg A\ true}\ DNE}\ \supset I^u
  }
}{}
$$

Once more we are faced with our great adversary, the *tertium non datur* himself! But now the tables are turned: no longer must we cower in fear behind our security blanket of negation. No, now we can prove *A true* itself, by the very assumption we hypothesized to prove $\neg A$ *true*! The derivation is now complete.

$$
\cfrac{
  \cfrac{\hphantom{\neg(A \vee \neg A)\ true}}{\neg(A \vee \neg A)\ true}\ u
  \quad
  \cfrac{
    \cfrac{
      \cfrac{\hphantom{\neg(A \vee \neg A)\ true}}{\neg(A \vee \neg A)\ true}\ u
      \quad
      \cfrac{\cfrac{\overline{A\ true}\ v}{A \vee \neg A\ true}\ \vee I_L}{}
    }{\perp\ true}\ \supset E
    \ \supset I^v
  }{}
}{}
$$

An interesting point to note about this proof is that, save for the last line, it is a perfectly valid intuitionistic proof. Only at the very end—or beginning, as we've told the story—did we need to appeal to classical reasoning in the form of *DNE*. In fact, this observation hints at a more general one: there is a *double-negation translation*, discussed in the next lecture, that translates any classical theorem into one that is intuitionistic valid.

From the intuitionistic portion of the deduction, we can read off a proof term, the analysis of which will begin to reveal the nature of the computational interpretation of classical proofs.

$$\textbf{fn}\ (u : \neg(A \vee \neg A)) \Rightarrow u\ (\textbf{inr}\ (\textbf{fn}\ (v : A) \Rightarrow u\ (\textbf{inl}\ v)))$$

It begins by supposing a proof u of $\neg(A \vee \neg A)$—i.e., a *refutation* of $A \vee \neg A$—and then proceeds to debunk that refutation, showing that it must have been mistaken by driving it to a contradiction. There are two choices of how to do so: prove A or prove $\neg A$. First, it chooses to show that, in fact, $\neg A$ is the case. Its proof of $\neg A$ proceeds as usual, supposing a proof of A and deriving contradiction. But that contradiction is produced precisely by changing its mind, saying that the refutation u is mistaken because, in fact, by the new assumption, A is the case! This time-travelling mind-changing behavior is essential to classical reasoning, and we'll see in the next lecture how this corresponds to programming with continuations.

Although it should be clear that adding either of *LEM* or *DNE* would suffice to make our logic classical, such rules violate the aesthetic principles we've adhered to thus far: both rules contain connectives, but they are neither introducing nor eliminating any single connective, and both rules contain multiple instances of a connective, suggesting a certain non-orthogonality of principles. Must our logic include negation and disjunction in order to become classical? Perhaps just negation? Or, since negation as we've seen it has been defined as implying falsehood, perhaps implication and falsehood are the important characteristics.

In fact, we can characterize what it means for a logic to be classical without appealing to any connectives at all, using purely judgmental notions. This is what we shall now endeavor to do.

4   Towards a better proof theory

Our proof theory for classical logic will be based on the idea of proof by contradiction. Proof by contradiction is closely related to double-negation elimination: If we take the rule *DNE* and require that its premise be proven by implication-introduction (as we know we may), we find that we could

replace it by a rule that looks like this:

$$\cfrac{\cfrac{\overline{\neg A \; true}^{\;u}}{\vdots}{\cfrac{\bot}{A \; true}}}{} \; ?PBC^u$$

If an assumption that $\neg A$ is true can lead to a proof of falsehood, we may conclude (classically, anyhow) that A must be true.

But we should keep in mind the design principles that keep our logics clear and easy to reason about: at most one connective should occur in a given rule, with all other notions being at the level of judgments. The proposed rule above still contains two connectives, $\neg$ and $\bot$, and does not read like an introduction or elimination rule, so we would like to replace it with a rule that appeals only to judgmental notions. We achieve this by inventing two new judgment forms, *A false* and # (contradiction):

$$\cfrac{\cfrac{\overline{A \; false}^{\;k}}{\vdots}{\cfrac{\#}{A \; true}}}{} \; PBC^k$$

By convention we use letters like k and ℓ to denote hypotheses of falsity.

Now, of course, we must explain the meaning of the new judgment forms *A false* and #. We understand these judgments through certain principles analogous to the Substitution Principle we posited for hypotheses of the form *A true* before. First, we require that contradiction yield anything:

$$\mathcal{D} \qquad \mathcal{D}$$
Contradiction Principle: If #, then J.

In this principle, J stands for any judgment.

The *false* judgment is treated somewhat specially: we derive its meaning from the meaning of *true*. We take *A false* as a conclusion to be a judgment-level notational definition:

$$A \; false := \cfrac{\cfrac{A \; true}{\vdots}}{\#}$$

To prove a proposition false, we assume it true and derive a contradiction.

We still require hypotheses of the form *A false*, since they appear in our classical rule PBC^k. To explain the meaning of hypotheses of *A false*, we derive a substitution principle from the definition above.

$$
\textbf{Falsity Substitution Principle: If} \quad
\dfrac{\begin{array}{c}\overline{\rule{2em}{0.4pt}}\;u \\[-0.3em] A\;\mathit{true} \\ \mathcal{D} \\ \# \end{array}}{} \quad \textbf{and} \quad
\dfrac{\begin{array}{c}\overline{\rule{2em}{0.4pt}}\;k \\[-0.3em] A\;\mathit{false} \\ \mathcal{E} \\ J \end{array}}{} \quad \textbf{then} \quad
\dfrac{\begin{array}{c}\dfrac{\mathcal{D}}{A\;\mathit{false}}\;k \\ \mathcal{E} \\ J \end{array}}{}
$$

The final deduction in the above no longer depends on the hypothesis k nor on the hypothesis u.

There's an important difference between this substitution principle and the one for true hypotheses: we took that substitution principle as a given, since it arose directly from the meaning of the hypothetical judgment. Here, though, since we did not directly define *A false* as a conclusion, we are not substituting proofs that conclude *A false* for hypotheses of *A false*. Therefore, this substitution principle is one we must prove of our inference rules, and we use the double-ruled notation to denote the operation embodied by this proof. (An alternative notation is $[\![u.\,\mathcal{D}/k]\!]\,\mathcal{E}$.) We defer any further discussion until later when we learn rule induction.

We have one more thing to define before the system is complete, and that is how to derive contradiction! With the rules given above, there are currently no ways to do so. Contradiction means that some proposition is both true and false, so we take the following rule:

$$
\dfrac{A\;\mathit{false} \quad A\;\mathit{true}}{J}\;\mathit{contra}
$$

Note that we do not suppose the rule to conclude # directly, since we wish to ensure that the Contradiction Principle above remains true. Instead, we allow the rule to conclude any judgment J, including #.

This concludes the discussion of the judgments *A false* and #. There is one remaining important technical matter to deal with, though: since we have introduced one new judgment that may appear as a conclusion, #, we must revisit any rules we previously defined that were supposed to conclude an arbitrary conclusion. In particular, the rules $\vee E^{u,v}$ and $\bot E$ must be generalized to allow a conclusion of any judgment, and not merely

one of the form C *true*[2]:

$$\cfrac{A \lor B \; true \qquad \cfrac{\overline{A \; true}^{\;u}}{\vdots} \qquad \cfrac{\overline{B \; true}^{\;v}}{\vdots}}{J} \lor E^{u,v} \qquad\qquad \cfrac{\bot \; true}{J} \bot E$$

To show an example of these rules in action, we derive Peirce's Law, one of the characterizations of classical logic mentioned above. We show only the final deduction, but it has a flavor very similar to the derivation of $A \lor \neg A$ from DNE.

$$\cfrac{\cfrac{\overline{A \; false}^{\;k}}{\cfrac{\cfrac{\overline{(A \supset B) \supset A \; true}^{\;u} \qquad \cfrac{\cfrac{\cfrac{\overline{A \; false}^{\;k} \qquad \overline{A \; true}^{\;v}}{B \; true} \; contra}{A \supset B \; true} \supset I^v}{A \; true} \supset E}{\cfrac{\#}{A \; true} PBC^k} \; contra}}{((A \supset B) \supset A) \supset A} \supset I^u$$

Note the two uses of the hypotheses k that A *false*, which is similar to the two uses of the refutation of $A \lor \neg A$ in the derivation of *LEM* above. Notice also the two uses of the rule *contra* to conclude two different judgments.

Negation revisited. Armed with our new judgmental notions, we can now give a direct definition of negation rather than one that simply defines it in terms of implication and falsehood. The new rules are as follows:

$$\cfrac{\cfrac{\overline{A \; true}^{\;u}}{\vdots}}{\cfrac{\#}{\neg A \; true}} \neg I^u \qquad\qquad \cfrac{\neg A \; true \qquad \cfrac{\overline{A \; false}^{\;k}}{\vdots} \quad J}{J} \neg E^k$$

The introduction rule essentially amounts to saying, "$\neg A$ *true* if A *false*," but replacing A *false* with its judgmental definition. The elimination rule

---

[2]We would also have to similarly generalize the existential elimination rule, $\exists E^a$, but in the present lecture we treat only the propositional fragment.

lets you extract *A false* from a proof of ¬*A true*, but since we only allow *A false* as a hypothesis, the rule introduces a new scope with an extra assumption rather than simply concluding *A false* directly.

This elimination rule is locally sound, as witnessed by the following local reduction, which makes use of the Falsity Substitution Principle:

$$
\cfrac{\cfrac{\cfrac{\overline{A\ true}^{\ u}}{\begin{array}{c}\mathcal{D}\\ \#\end{array}}}{\neg A\ true}\neg I^u \qquad \cfrac{\overline{A\ false}^{\ k}}{\begin{array}{c}\mathcal{E}\\ J\end{array}}}{J}\neg E^k \quad\Longrightarrow_R \quad \cfrac{\cfrac{\mathcal{D}}{A\ false}^{\ k}}{\begin{array}{c}\mathcal{E}\\ J\end{array}}
$$

It is also locally complete; we can expand an arbitrary deduction of ¬*A true* in two ways:

$$
\cfrac{\mathcal{D}}{\neg A\ true} \quad\Longrightarrow_E\quad \cfrac{\cfrac{\mathcal{D}}{\neg A\ true} \qquad \cfrac{\cfrac{\overline{A\ false}^{\ k}\qquad \overline{A\ true}^{\ u}}{\#}contra}{\neg A\ true}\neg I^u}{\neg A\ true}\neg E^k
$$

$$
\cfrac{\mathcal{D}}{\neg A\ true} \quad\Longrightarrow_E\quad \cfrac{\cfrac{\cfrac{\mathcal{D}}{\neg A\ true}\qquad \cfrac{\overline{A\ false}^{\ k}\qquad \overline{A\ true}^{\ u}}{\#}contra}{\#}\neg E^k}{\neg A\ true}\neg I^u
$$

Note how the elimination rule can be used to conclude either # or ¬*A true*.

To illustrate the use of negation, we give one final example deduction, a proof of the inverse contrapositive:

$$
\cfrac{\cfrac{\cfrac{\overline{\neg B \supset \neg A\ true}^{\ u}\qquad \cfrac{\cfrac{\overline{B\ false}^{\ k}\qquad \overline{B\ true}^{\ w}}{\#}contra}{\neg B\ true}\neg I^w}{\neg A\ true}\supset E\qquad \cfrac{\overline{A\ false}^{\ \ell}\qquad \overline{A\ true}^{\ v}}{\#}contra}{\cfrac{\cfrac{\#}{B\ true}PBC^k}{\cfrac{A\supset B\ true}{(\neg B\supset\neg A)\supset A\supset B\ true}\supset I^u}\supset I^v}}{}\neg E^\ell
$$

The rules for classical logic are summarized in Figure 1. It is worth noting that the only rule which makes this logic actually classical is the rule of proof by contradiction, PBC^k. All of the other rules, including the new, direct explanation of negation, are perfectly valid from an intuitionistic perspective: intuitionists still deal with falsity and contradiction, just in a more controlled fashion.

The rule PBC^k however is problematic from a verificationist perspective, because it changes the meaning of A *true*: a proposition's meaning is no longer derived from its verifications, since we may now choose to prove *any* proposition by contradiction. It is this retroactive change to the meanings of our connectives that now allows us to prove things like the law of the excluded middle or double-negation elimination. Our carefully considered approach has been irreparably compromised.

Judgmental Definitions

$$A \; false := \begin{array}{c} A \; true \\ \vdots \\ \# \end{array} \qquad \dfrac{A \; false \quad A \; true}{J} \; contra$$

Classical Rule

$$\cfrac{\cfrac{\quad}{A \; false} \; k}{\vdots} \atop \cfrac{\#}{A \; true} \; PBC^k$$

Changed Rules

$$\dfrac{A \lor B \; true \qquad \begin{array}{c} \overline{A \; true} \; u \\ \vdots \\ J \end{array} \qquad \begin{array}{c} \overline{B \; true} \; v \\ \vdots \\ J \end{array}}{J} \; \lor E^{u,v} \qquad\qquad \dfrac{\bot \; true}{J} \; \bot E$$

Rules for Negation

$$\cfrac{\begin{array}{c} \overline{A \; true} \; u \\ \vdots \\ \# \end{array}}{\neg A \; true} \; \neg I^u \qquad\qquad \dfrac{\neg A \; true \qquad \begin{array}{c} \overline{A \; false} \; k \\ \vdots \\ J \end{array}}{J} \; \neg E^k$$

Figure 1: Rules for classical natural deduction