Derivability

Professor Cian Dorr

13th October 2022

New York University

Syntax of first-order terms (review)

For a first-order signature Σ, $\mathrm{Terms}(\Sigma)$ is the smallest set of strings meeting the following conditions:

$$\frac{v \in \mathrm{Var}}{v \in \mathrm{Terms}(\Sigma)}$$

$$\frac{f \in F_\Sigma, a_\Sigma(f) = n \qquad t_1 \in \mathrm{Terms}(\Sigma) \qquad \cdots \qquad t_n \in \mathrm{Terms}(\Sigma)}{f(t_1, \ldots, t_n) \in \mathrm{Terms}(\Sigma)}$$

Syntax of first-order formulae (review)

For a first-order signature Σ, $\mathcal{L}(\Sigma)$ is the smallest set of strings meeting the following closure conditions:

$$\frac{F \in R_\Sigma, a_\Sigma(F) = n \qquad t_1 \in \mathrm{Terms}(\Sigma) \qquad \cdots \qquad t_n \in \mathrm{Terms}(\Sigma)}{R(t_1, \ldots, t_n) \in \mathcal{L}(\Sigma)}$$

$$\frac{P \in \mathcal{L}(\Sigma)}{\neg P \in \mathcal{L}(\Sigma)} \qquad \frac{P \in \mathcal{L}(\Sigma) \qquad Q \in \mathcal{L}(\Sigma)}{P \rightarrow Q \in \mathcal{L}(\Sigma)}$$

$$\frac{P \in \mathcal{L}(\Sigma) \qquad Q \in \mathcal{L}(\Sigma)}{P \wedge Q \in \mathcal{L}(\Sigma)} \qquad \frac{P \in \mathcal{L}(\Sigma) \qquad Q \in \mathcal{L}(\Sigma)}{P \vee Q \in \mathcal{L}(\Sigma)}$$

$$\frac{P \in \mathcal{L}(\Sigma) \qquad v \in \mathrm{Var}}{\forall v P \in \mathcal{L}(\Sigma)} \qquad \frac{P \in \mathcal{L}(\Sigma) \qquad v \in \mathrm{Var}}{\exists v P \in \mathcal{L}(\Sigma)}$$

Provability

Provability

We are interested in studying the properties of *classical first order logic*, the system of logic for first-order languages that formalizes the reasoning we have been doing in our informal mathematical proofs.

To this end, we are going to define a relation $\vdash \subseteq \mathcal{P}(\mathcal{L}(\Sigma)) \times \mathcal{L}(\Sigma)$: that is, a relation that can hold between a *set* of formulae Γ and a *single* formula P.

- We write the relation in infix position: $\Gamma \vdash P$ is short for $\langle \Gamma, P \rangle \in \vdash$.
- We pronounce $\Gamma \vdash P$ as `$P$ is [classically] provable/derivable from $\Gamma$' or `Γ proves P'. This is a misnomer given that it'll turn out that every sentence is ``provable from'' some Γ, but we're stuck with it.
- An ordered pair $\langle \Gamma, P \rangle$ of a set of formulae and a formula is often called a *sequent*, and notated as something like $\Gamma \triangleright P$.
- When referring to sets of formulae Γ, P abbreviates $\Gamma \cup \{P\}$, and Γ, Δ abbreviates $\Gamma \cup \Delta$. $P \vdash Q$ abbreviates $\{P\} \vdash Q$, and $\vdash P$ abbreviates $\varnothing \vdash P$.

Defining provability

$\vdash$ is defined to be the smallest relation between sets of formulae and formulae meeting the following conditions.

$$\frac{P \in \mathcal{L}(\Sigma)}{P \vdash P}\ \text{Assumption} \qquad \frac{\Gamma \vdash P \qquad \Delta \subseteq \mathcal{L}(\Sigma)}{\Gamma, \Delta \vdash P}\ \text{Weakening}$$

$$\frac{\Gamma, P \vdash Q}{\Gamma \vdash P \to Q}\ {\to}\text{Intro} \qquad \frac{\Gamma \vdash P \to Q \qquad \Gamma \vdash P}{\Gamma \vdash Q}\ {\to}\text{Elim}$$

$$\frac{\Gamma \vdash P \qquad \Gamma \vdash Q}{\Gamma \vdash P \wedge Q}\ {\wedge}\text{Intro} \qquad \frac{\Gamma \vdash P \wedge Q}{\Gamma \vdash P}\ {\wedge}\text{Elim1} \qquad \frac{\Gamma \vdash P \wedge Q}{\Gamma \vdash Q}\ {\wedge}\text{Elim2}$$

$$\frac{\Gamma \vdash P \qquad Q \in \mathcal{L}(\Sigma)}{\Gamma \vdash P \vee Q}\ {\vee}\text{Intro1} \qquad \frac{\Gamma \vdash Q \qquad P \in \mathcal{L}(\Sigma)}{\Gamma \vdash P \vee Q}\ {\vee}\text{Intro2}$$

$$\frac{\Gamma \vdash P \vee Q \qquad \Gamma, P \vdash R \qquad \Gamma, Q \vdash R}{\Gamma \vdash R}\ {\vee}\text{Elim}$$

Defining provability (contd.)

$$\frac{\Gamma, P \vdash Q \qquad \Gamma, P \vdash \neg Q}{\Gamma \vdash \neg P}\ \neg\text{Intro} \qquad\qquad \frac{\Gamma \vdash \neg\neg P}{\Gamma \vdash P}\ \text{DNE}$$

$$\frac{\Gamma \vdash P[u/v] \qquad u \notin FV(\Gamma, \forall v P)}{\Gamma \vdash \forall v P}\ \forall\text{Intro} \qquad\qquad \frac{\Gamma \vdash \forall v P \qquad t \in \mathrm{Terms}(\Sigma)}{\Gamma \vdash P[t/v]}\ \forall\text{Elim}$$

$$\frac{\Gamma \vdash P[t/v]}{\Gamma \vdash \exists v P}\ \exists\text{Intro} \qquad\qquad \frac{\Gamma \vdash \exists v P \qquad \Gamma, P[u/v] \vdash Q \qquad u \notin FV(\Gamma, Q, \exists v P)}{\Gamma \vdash Q}\ \exists\text{Elim}$$

$$\frac{t \in \mathrm{Terms}(\Sigma)}{\vdash t = t}\ {=}\text{Intro} \qquad\qquad \frac{\Gamma \vdash s = t \qquad \Gamma \vdash P[s/v]}{\Gamma \vdash P[t/v]}\ {=}\text{Elim}$$

Nomenclature

We call a sequent of the form $P \triangleright P$ an *instance* of Assumption. Likewise, we call a sequent of the form $\triangleright t = t$ an *instance* of $=$Elim.

We say that $\Gamma' \triangleright P'$ *follows by Weakening from* $\Gamma \triangleright P$ iff $P' = P$ and $\Gamma \subseteq \Gamma'$.

We say that $\Gamma' \triangleright P'$ *follows by* $\rightarrow$*Intro from* $\Gamma \triangleright P$ iff there exists a formula Q such that $P' = Q \rightarrow P$ and $\Gamma = \Gamma' \cup \{Q\}$.

We say that $\Gamma'' \triangleright P''$ *follows by* $\rightarrow$*Elim from* $\Gamma \triangleright P$ and $\Gamma' \triangleright P'$ iff $\Gamma'' = \Gamma' = \Gamma$ and $P = P' \rightarrow P''$.

Similarly for all the rest. All told we have 2 sets of sequents; 9 binary relations between sequents; 6 ternary relations between sequents; and 1 quaternary relation between sequents. $\vdash$ is being defined as the closure of the union of the two sets under those 16 relations.

Some other provability relations

I mentioned the existence of restricted languages $\mathcal{L}_\Phi(\Sigma)$, where $\Phi \subseteq \{\neg, \to, \wedge, \vee, \forall, \exists\}$. When we are thinking about $\mathcal{L}_\Phi(\Sigma)$, it's natural to consider a provability relation $\vdash_\Phi$ that's defined like $\vdash$ above, but that just drops the rules that mention a logical constant not in Φ.

- These restrictions of $\vdash$ also make *sense* for $\mathcal{L}$, but aren't very interesting there, since e.g., $P \nvdash_{\neg,\wedge,\forall} P \vee Q$ and $P \to Q, P \nvdash_{\neg,\wedge,\forall} Q$ for all P and Q.

Two other famous restrictions of $\vdash$ that *are* interesting on $\mathcal{L}(\Sigma)$

- $\vdash_M$ (provability in *minimal logic*) is defined like $\vdash$ but dropping the DNE rule.
- $\vdash_I$ (provability in *intuitionistic logic*) is defined like $\vdash$ but replacing DNE with

$$\frac{\Gamma \vdash_I P \qquad \Gamma \vdash_I \neg P}{\Gamma \vdash_I Q}\ \text{Explosion}$$

Explosion is a *derived rule* for $\vdash$, i.e. we have $\Gamma \vdash Q$ whenever $\Gamma \vdash P$ and $\Gamma \vdash \neg P$.

A succinct way to show this is to display the following diagram:

$$\dfrac{\dfrac{\dfrac{\Gamma \vdash P}{\Gamma, \neg Q \vdash P}\,\text{Weakening} \qquad \dfrac{\Gamma \vdash \neg P}{\Gamma, \neg Q \vdash \neg P}\,\text{Weakening}}{\Gamma \vdash \neg\neg Q}\,\neg\text{Intro}}{\Gamma \vdash Q}\,\text{DNE}$$

Proving by induction that all provable sequents have some property

Given how $\vdash$ is defined as a closure, if we want to prove something of the form 'For all Γ and P such that $\Gamma \vdash P$, $\phi(\Gamma, P)$', we can do so by a proof by induction. In principle there'll be 18 steps: two base clauses (corresponding to the zero-premise rules Assumption and $=$Intro) and 16 induction steps (corresponding to the 10 one-premise, 5 two-premise, and 1 three-premise rules).

Often many steps can be bundled together. (Don't worry, I'm not going to give you assignments where you have to do anything like 18 separate bits.)

A proof by induction: provability is compact

Compactness of provability

$\Gamma \vdash P$ iff there is a finite $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \vdash P$.

The right-to-left direction is just a matter of applying Weakening.

The left-to-right direction needs an induction. Let's say that a sequent $\Gamma \triangleright P$ is *compactable* iff there is a finite $\Gamma_0 \subseteq \Gamma$ such that $\Gamma_0 \vdash P$; we are trying to show that every provable sequent is compactable.

Assumption: if $\Gamma \rhd P$ is an instance of Assumption, $\Gamma = \{P\}$ which is finite.

Weakening: if $\Gamma \rhd P$ follows by Weakening from some provable compactable sequent, that sequent must be $\Delta \rhd P$ for some $\Delta \subseteq \Gamma$. By the IH, there's a finite $\Delta_0 \subseteq \Delta$ such that $\Delta_0 \vdash P$; since $\Delta_0 \subseteq \Gamma$, this means $\Gamma \rhd P$ is also compactable.

$\rightarrow$Intro: suppose $\Gamma \rhd P$ follows by $\rightarrow$Intro from some provable compactable sequent. Then there must be some Q and R such that $P = Q \rightarrow R$ and that sequent is $\Gamma, Q \rhd R$. By the IH, there's a finite subset Δ of $\Gamma \cup \{Q\}$ such that $\Delta \rhd R$. Let $\Gamma_0 = \Delta \setminus \{Q\}$; note that Γ_0 is finite since Δ is. Then $\Gamma_0 \subseteq \Gamma$ and $\Delta \subseteq \Gamma_0 \cup \{Q\}$, so by Weakening $\Gamma_0, Q \vdash R$, so by $\rightarrow$Intro, $\Gamma_0 \vdash Q \rightarrow R$.

$\rightarrow$Elim: suppose $\Gamma \rhd P$ follows by $\rightarrow$Elim from two provable compactable sequents. Then there must be some Q such that one of those sequents is $\Gamma \rhd Q \rightarrow P$ and the other is $\Gamma \rhd Q$. By the IH, there are finite subsets Γ_1, Γ_2 of Γ such that $\Gamma_1 \vdash Q \rightarrow P$ and $\Gamma_2 \vdash Q$. Let $\Gamma_0 = \Gamma_1 \cup \Gamma_2$. By Weakening, $\Gamma_0 \vdash Q \rightarrow P$ and $\Gamma_0 \vdash Q$, so by $\rightarrow$Elim, $\Gamma_0 \vdash P$. Γ_0 is finite since it's the union of two finite sets.

Other rules similar.

Finite-sequent provability

We can shed more light on the last result by introducing a new relation $\vdash_{\mathit{fin}}$ between *finite* sets of formulae and formulae (a subset of $\mathcal{P}_{\mathit{fin}}(\mathcal{L}(\Sigma)) \times \mathcal{L}(\Sigma)$). It is defined just like $\vdash$, but the Weakening rule is changed to:

$$\frac{\Gamma \vdash_{\mathit{fin}} P \qquad Q \in \mathcal{L}(\Sigma)}{\Gamma, Q \vdash_{\mathit{fin}} P}\ \text{One-formula Weakening}$$

Theorem

$\Gamma \vdash P$ iff for some $\Gamma_0 \subseteq \Gamma$, $\Gamma_0 \vdash_{\mathit{fin}} P$

For the right-to-left direction, we first show by a trivial induction that $\Gamma_0 \vdash P$ whenever $\Gamma_0 \vdash_{\mathit{fin}} P$, and then appeal to Weakening to get that when $\Gamma_0 \vdash P$ and $\Gamma_0 \subseteq \Gamma$, $\Gamma \vdash P$.

For the left-to right direction, we first prove that $\vdash_{\mathit{fin}}$ is closed under the version of Weakening restricted to *finite* Δ (by induction on the size of Δ). Then the proof proceeds just like the one on the previous slide.

Provability and proofs

What about proofs?

When you're taught to *use* a formal system of logic, you're taught rules for writing down things called *proofs*. So far, we haven't even mentioned them!

But there's a sense in which a certain very abstract notion of "proof" is in play whenever one defines a set (or relation) as a closure.

Closures and derivations

Supose we have a family (R_i) of relations on a set A (which may be of different arities), such that C is the smallest subset of A closed under all the R_i. (Note that some of the R_i may be 1-ary, i.e. subsets of A, so C need not be $\varnothing$!)

Let a *derivation history* for (R_i) be a list $s \in A^*$ such that for each element y of s, there is some R_i (of arity n) and some $x_1, \ldots, x_{n-1}$ occurring earlier than y in s such that $R_i x_1 \ldots x_{n-1} y$. More carefully:

Definition

The set of derivation histories for (R_i) is the smallest set which contains $[]$ and is such that if it contains s, and $R_i x_1 \ldots x_{n-1} y$ for some $x_1 \ldots x_{n-1} \in \mathrm{elements}\, s$ (where n is the arity of R) then it contains $(y : s)$.

It is easy to show that s and t are derivation histories, $s \oplus t$ is. (Use induction on s.)

The existence of derivations

A *derivation of y* is a derivation history whose last element is y, i.e. which is $(y : s)$ for some $s \in A^*$.

Then we can show that for any $y \in A$, $y \in C$ iff there is a derivation of y.

Proof, left-to-right: When $y \in R_i$ for a singulary R_i, $[y]$ is a derivation of y. Otherwise, suppose $R_i x_1 \ldots x_{n-1} y$ for an n-ary R, where $x_1 \ldots x_{n-1} \in C$, and there is a derivation s_i of each x_i. Then $t := [y] \oplus s_1 \oplus \cdots \oplus s_{n-1}$ is a derivation of y.

Right-to-left: we show by induction that for every s, if s is a derivation history, then $\mathrm{elements}\, s \subseteq C$. Base case trivial since $\mathrm{elements}[] = \varnothing \subseteq C$. Induction step: suppose s is such that if it's a derivation history, then $\mathrm{elements}\, s \subseteq C$. Suppose $(y : s)$ is a derivation history. Then s is a derivation history and there is an n-ary R_i such that for some $x_1 \ldots x_{n-1} \in \mathrm{elements}\, s$, $R_i x_1 \ldots x_{n-1} y$. But by the induction hypothesis, $\mathrm{elements}\, s \subseteq C$. Since C is closed under R_i, it follows that $y \in C$.

The use of derivation histories

A common situation: we are interested in some finite number of (R_i) such that for each one, it is a mechanical matter to check whether $R_i x_1 \ldots x_n$ (when $x_1 \ldots x_n$ are "given" to us in some appropriate way, e.g. they are strings we have written down).

Figuring out whether a given y belongs to the smallest set closed under (R_i) may still be very hard!

But if we are presented (in the same canonical way) with a list of elements of A, it will be a mechanical matter to check whether it is a derivation of y. We just go through it step by step and check, for each step, whether it bears each R_i to an appropriate tuple of previous steps.

- It's easier if we have an ``annotated'' derivation history where for each element of the list we are told which R_i applies and where in the earlier list we are to find the relevant $x_1, \ldots, x_{n-1}$; but even without this information, there are only finitely many possibilities to search through.

Proofs

A derivation history for $\vdash$ is a list of sequents, where each one is either an instance of Assumption, an instance of $=$Intro, follows from some earlier sequent by Weakening, $\to$Intro, ...; follows from two earlier sequents by $\to$Elim, $\wedge$Intro,...; or follows from three earlier sequents by $\vee$Elim.

Such derivation histories may involve sequents $\Gamma \triangleright P$ where Γ is infinite. There is no clear sense in which one could *write down* such a derivation history.

By contrast, derivation histories for $\vdash_{fin}$ are lists of *finite* objects. These are more like what we'd expect a `proof' to be.

Proofs as strings

If we're treating formulas as strings, it's natural to think that proofs should be strings too. Derivation histories in $\vdash_{fin}$ aren't strings; they are lists of ordered pairs of a finite set of formulae and a formula.

But we can represent any such list unambiguously as a string.

1. First, convert each finite set of formulae into a string by listing them in alphabetical order and joining them with a character that never appears in formulae: `;`, say.
2. Second, convert each ordered pair of such a string and a formula into a single string by joining the two with another character that never appears in formulae: `>`, say.
3. Finally, convert the resulting list of strings into a single string by joining them all with some third character that never appears in formulae: newline, say.

Tree-style proofs

In constructing a proof in this sense of some given sequent, one typically ends up making a lot of arbitrary choices about how to order the lines. *Proof theorists* are interested in properties of proofs that don't depend on these arbitrary line-numbering choices: for these purposes, it's more useful to think of formal proofs as *trees* of formulae rather than lists of formulae. E.g. the following (cf. our earlier discussion of Explosion) would be an unambiguous visual representation of a unique proof:

$$\dfrac{\dfrac{\dfrac{\Gamma \triangleright P}{\Gamma, \neg Q \triangleright P}\text{Weakening} \qquad \dfrac{\Gamma \triangleright \neg P}{\Gamma, \neg Q \triangleright \neg P}\text{Weakening}}{\Gamma \triangleright \neg\neg Q}\neg\text{Intro}}{\Gamma \triangleright Q}\text{DNE}$$

The theory of trees can be developed along similar lines to the theory of lists; but we haven't done this, so we'll stick with our less elegant linear conception of proof.

Fitch-style proof formatting

In your introductory Logic class, you probably learnt a way of writing down proofs that look something like this:

1 -	A	
2 -	B	Assumption
3 -	A	Reiteration, 1
4 -	$B \to A$	$\to$-Intro 2--3
5 -	$A \to (B \to A)$	$\to$-Intro 1--4

Any such proof can be understood as a proof in our sense (in a system with some more rules, which are derived rules for $\vdash$) where the vertical lines on the left correspond to the formulae on the left hand side of $\triangleright$.

The Fitch-style proof on the previous line corresponds to the following list of sequents (which is in fact a proof in our sense):

$$
\begin{aligned}
A &\triangleright A \\
A, B &\triangleright B \\
A, B &\triangleright A \\
A &\triangleright B \to A \\
&\triangleright A \to (B \to A)
\end{aligned}
$$

There's a simple algorithm for converting from Fitch notation to ours; the Fitch proof can be regarded as a nicer looking visual representation of the corresponding list-of-sequents.

Fitch proofs enforce certain choices about the ordering of lines, so not every proof in our sense corresponds to a Fitch proof. It turns out that every provable sequent *does* have a Fitch proof, but we won't show this since it would require a rigorous definition of ``Fitch proof'' (which is fiddly!).

Derivable rules

There are many other good-looking inference rules which we might have been tempted to add to the definition of $\vdash$. For example

$$\frac{\Gamma \vdash P \vee Q \qquad \Gamma \vdash \neg P}{\Gamma \vdash Q}\ \text{Disjunctive Syllogism}$$

But this turns out to be already derivable for $\vdash$ as defined earlier, so we would get exactly the same relation if we added this to the definition.

$$\dfrac{\dfrac{\Gamma \vdash P \vee Q \qquad \dfrac{\dfrac{\Gamma \vdash \neg P}{\Gamma, P, \neg Q \vdash \neg P}\,\text{W} \quad \dfrac{\dfrac{}{P \vdash P}\,\text{A}}{\Gamma, P, \neg Q \vdash P}\,\text{W}}{\Gamma, P \vdash \neg\neg Q} \qquad \dfrac{\dfrac{\dfrac{}{Q \vdash Q}\,\text{A}}{\Gamma, Q, \neg Q \vdash Q}\,\text{W} \quad \dfrac{\dfrac{}{\neg Q \vdash \neg Q}\,\text{A}}{\Gamma, Q, \neg Q \vdash \neg Q}\,\text{W}}{\Gamma, Q \vdash \neg\neg Q}\,\neg\text{I}}{\Gamma \vdash \neg\neg Q}}{\Gamma \vdash Q}$$

Provable equivalence

Definition

$P \dashv\vdash Q$ (P and Q are provably equivalent) iff $P \vdash Q$ and $Q \vdash P$.

Fact

$P \dashv\vdash Q$ iff for all Γ, $\Gamma \vdash P$ iff $\Gamma \vdash Q$.

Proof, right to left: by Assumption, $P \vdash P$ and $Q \vdash Q$. Left-to-right: suppose $\Gamma \vdash P$. Since $P \vdash Q$, $\Gamma, P \vdash Q$ by Weakening, so $\Gamma \vdash P \to Q$ by $\to$Intro; hence $\Gamma \vdash Q$ by $\to$Elim.

Provable equivalence

Provable equivalence is a congruence

If $P \dashv\vdash Q$, then $\neg P \dashv\vdash \neg Q$, and for any variable v, $\forall v P \dashv\vdash \forall v Q$ and $\exists v P \dashv\vdash \exists v Q$.

If furthermore $P' \dashv\vdash Q'$, then also $P \wedge P' \dashv\vdash Q \wedge Q'$, $P \vee P' \dashv\vdash Q \vee Q'$, and $P \rightarrow P' \dashv\vdash Q \rightarrow Q'$

Proof for $\wedge$: by $\wedge$Elim, $P \wedge P' \vdash P$ and $P \wedge P' \vdash P'$. But then by the hypothesis, $P \wedge P' \vdash Q$ and $P \wedge P' \vdash Q'$. So by $\wedge$Intro, $P \wedge P' \vdash Q \wedge Q'$.

Proof for $\forall$: $\forall v P \vdash P$ by $\forall$Elim (since $P = P[v/v]$. So by the hypothesis, $\forall v P \vdash Q$. Since v isn't free in $\forall v P$ or $\forall v Q$, we can apply $\forall$Intro to conclude that $\forall v P \vdash \forall v Q$.

Other cases similar.