Natural Deduction for IfArith

CIS352 — Fall 2023

Kris Micinski

In this lecture, we'll introduce **natural deduction**

Natural deduction is a mathematical formalism that helps ground the ideas in metacircular interpreters

Natural deduction first used in mathematical logic, to specify **proofs** using inductive data

We will use natural deduction as a framework for specifying semantics of various languages throughout the course

Introduction Rules	Elimination Rules
$$\dfrac{\begin{array}{c}\overline{\vdash^N A}^{\,u} \\ \vdots \\ \vdash^N B\end{array}}{\vdash^N A \supset B}\supset\!\mathrm{I}^u$$	$$\dfrac{\vdash^N A \supset B \qquad \vdash^N A}{\vdash^N B}\supset\!\mathrm{E}$$
$$\dfrac{\begin{array}{c}\overline{\vdash^N A}^{\,u} \\ \vdots \\ \vdash^N p\end{array}}{\vdash^N \neg A}\neg\mathrm{I}^{p,u}$$	$$\dfrac{\vdash^N \neg A \qquad \vdash^N A}{\vdash^N C}\neg\mathrm{E}$$
$$\dfrac{\vdash^N [a/x]A}{\quad}\forall\mathrm{I}^a$$	$$\dfrac{\vdash^N \forall x.\, A}{\quad}\forall\mathrm{E}$$

When we specify the semantics of a language using natural deduction, we give its semantics via a set of **inference rules**

Rules read: if the thing on the **top** is true, then the thing on the **bottom** is also true.

This rule says: "if c is an integer (mathematically: c ∈ ℚ), then c evaluates to c."

$$\textbf{Const}: \frac{c \in \mathbb{Q}}{c \Downarrow c}$$

Note: the notation e ⇓ v is read "e evaluates to v."

Some rules will have more than one **antecedent** (thing on the top).

You read these: "if the first thing, and second thing, and … are **all** true, then the thing on the bottom is true."

$$\textbf{Plus}: \frac{e_0 \Downarrow n_0 \quad e_1 \Downarrow n_1 \quad n' = n_0 + n_1}{(\text{plus } e_0\ e_1) \Downarrow n'}$$

“If $e_0 \Downarrow n_0$, and $e_1 \Downarrow n_1$, and n' = n_0 + n1, **then** I can say (plus e_0 e_1) $\Downarrow$ n′.”

$$\textbf{Plus}: \frac{e_0 \Downarrow n_0 \quad e_1 \Downarrow n_1 \quad n' = n_0 + n_1}{(\mathsf{plus}\ e_0\ e_1) \Downarrow n'}$$

$$\textbf{Const} : \frac{c \in \mathbb{Q}}{c \Downarrow c} \qquad \textbf{Plus} : \frac{e_0 \Downarrow n_0 \quad e_1 \Downarrow n_1 \quad n' = n_0 + n_1}{(\mathsf{plus}\ e_0\ e_1) \Downarrow n'}$$

$$\textbf{Div} : \frac{e_0 \Downarrow n_0 \quad e_1 \Downarrow n_1 \quad n' = n_0 / n_1}{(\mathsf{div}\ e_0\ e_1) \Downarrow n'}$$

The natural deduction rule for **div** is similar

$$\textbf{Const}: \frac{c \in \mathbb{Q}}{c \Downarrow c} \qquad \textbf{Plus}: \frac{e_0 \Downarrow n_0 \quad e_1 \Downarrow n_1 \quad n' = n_0 + n_1}{(\mathsf{plus}\ e_0\ e_1) \Downarrow n'}$$

$$\textbf{Div}: \frac{e_0 \Downarrow n_0 \quad e_1 \Downarrow n_1 \quad n' = n_0 / n_1}{(\mathsf{div}\ e_0\ e_1) \Downarrow n'}$$

$$\textbf{Not}_\mathbf{0}: \frac{e \Downarrow 0}{(\mathsf{not}\ e) \Downarrow 1} \qquad \textbf{Not}_\mathbf{1}: \frac{e \Downarrow n \quad n \neq 0}{(\mathsf{not}\ e) \Downarrow 0}$$

We have **two** rules for not

Natural Deduction Rules for IfArith

$$\textbf{Const}: \frac{\overline{c \in \mathbb{Q}}}{c \Downarrow c} \qquad \textbf{Plus}: \frac{e_0 \Downarrow n_0 \quad e_1 \Downarrow n_1 \quad n' = n_0 + n_1}{(\textsf{plus}\ e_0\ e_1) \Downarrow n'}$$

$$\textbf{Div}: \frac{e_0 \Downarrow n_0 \quad e_1 \Downarrow n_1 \quad n' = n_0 / n_1}{(\textsf{div}\ e_0\ e_1) \Downarrow n'}$$

$$\textbf{Not}_\mathbf{0}: \frac{e \Downarrow 0}{(\textsf{not}\ e) \Downarrow 1} \qquad \textbf{Not}_\mathbf{1}: \frac{e \Downarrow n \quad n \neq 0}{(\textsf{not}\ e) \Downarrow 0}$$

$$\textbf{If}_\mathbf{T}: \frac{e_0 \Downarrow 0 \quad e_1 \Downarrow n'}{(\textsf{if}\ e_0\ e_1\ e_2) \Downarrow n'} \qquad \textbf{If}_\mathbf{F}: \frac{e_0 \Downarrow n \quad n = 0 \quad e_2 \Downarrow n'}{(\textsf{if}\ e_0\ e_1\ e_2) \Downarrow n'}$$

Question: Now that we have the rules, what can we do with them?

Answer: Use them to **formally prove** that some program calculates some result

Let’s say I want to prove that the following program evaluates to 4:

```
(if (plus 1 -1) 3 4)
```

What rule could go here..?

$$\frac{???}{\left(\text{if } (\text{plus } 1\ \ -1)\ 3\ 4\right) \Downarrow 4}$$

$$\textbf{If}_\textbf{T} : \frac{e_0 \Downarrow n \quad n \neq 0 \quad e_1 \Downarrow n'}{(\text{if } e_0\ e_1\ e_2) \Downarrow n'} \quad \textbf{If}_\textbf{F} : \frac{e_0 \Downarrow 0 \quad e_2 \Downarrow n'}{(\text{if } e_0\ e_1\ e_2) \Downarrow n'}$$

$$\frac{???}{\left(\text{if (plus } 1\ \ -1)\ 3\ 4\right) \Downarrow 4}$$

$$\textbf{If}_\textbf{T} : \frac{e_0 \Downarrow n \quad n \neq 0 \quad e_1 \Downarrow n'}{(\text{if } e_0\ e_1\ e_2) \Downarrow n'} \qquad \textbf{If}_\textbf{F} : \frac{e_0 \Downarrow 0 \quad e_2 \Downarrow n'}{(\text{if } e_0\ e_1\ e_2) \Downarrow n'}$$

$$\frac{???}{\left(\text{if } (\text{plus } 1 \ \ -1)\ 3\ 4\right) \Downarrow 4}$$

To apply a natural-deduction rule,
we must perform **unification**

There can be no variables in the resulting unification!

$$\mathbf{If_F} : \frac{e_0 \Downarrow 0 \quad e_2 \Downarrow n'}{(\text{if } e_0\ e_1\ e_2) \Downarrow n'}$$

$$\frac{(\text{plus } 1\ \ -1) \Downarrow 0 \qquad\qquad 4 \Downarrow 4}{\left(\text{if } (\text{plus } 1\ \ -1)\ 3\ 4\right) \Downarrow 4}$$

We perform unification:

e_0: (plus 1 -1), e_1: 3

e_2: 4, n’: 4

Not done yet, now we have to prove
these things

$$\frac{(\text{plus } 1 \ \ -1) \Downarrow 0 \qquad 4 \Downarrow 4}{\left(\text{if } (\text{plus } 1 \ \ -1) \ 3 \ 4\right) \Downarrow 4}$$

Why can we say $4 \Downarrow 4$? Because of the **Const** rule

$$\frac{(\text{plus } 1 \ -1) \Downarrow 0 \qquad \dfrac{4 \in \mathbb{Q}}{4 \Downarrow 4}}{\left(\text{if } (\text{plus } 1 \ -1) \ 3 \ 4\right) \Downarrow 4}$$

We're not done yet, because **plus** requires an antecedent:

$$\textbf{Plus}: \frac{e_0 \Downarrow n_0 \quad e_1 \Downarrow n_1 \quad n' = n_0 + n_1}{(\text{plus } e_0\ e_1) \Downarrow n'}$$

$$\frac{(\text{plus } 1\ -1) \Downarrow 0 \qquad \dfrac{4 \in \mathbb{Q}}{4 \Downarrow 4}}{\left(\text{if } (\text{plus } 1\ -1)\ 3\ 4\right) \Downarrow 4}$$

But we’re **still** not done, because we need to finish these three

$$\dfrac{\dfrac{1 \Downarrow 1 \quad -1 \Downarrow -1 \quad 1 + -1 = 0}{(\text{plus } 1\ -1) \Downarrow 0} \qquad \dfrac{4 \in \mathbb{Q}}{4 \Downarrow 4}}{\left(\text{if } (\text{plus } 1\ -1)\ 3\ 4\right) \Downarrow 4}$$

Things that are simply true from algebra require no antecedents, we take them as "axioms."

$$\dfrac{\dfrac{\dfrac{1 \in \mathbb{Q}}{1 \Downarrow 1} \quad \dfrac{-1 \in \mathbb{Q}}{-1 \Downarrow -1} \quad \dfrac{}{1 + -1 = 0}}{(\text{plus } 1\ -1) \Downarrow 0} \qquad \dfrac{4 \in \mathbb{Q}}{4 \Downarrow 4}}{\left(\text{if } (\text{plus } 1\ -1)\ 3\ 4\right) \Downarrow 4}$$

This is a complete proof that the program computes 4

$$\dfrac{\dfrac{\dfrac{1 \in \mathbb{Q}}{1 \Downarrow 1} \quad \dfrac{-1 \in \mathbb{Q}}{-1 \Downarrow -1} \quad \dfrac{}{1 + -1 = 0}}{(\text{plus } 1\ -1) \Downarrow 0} \qquad \dfrac{4 \in \mathbb{Q}}{4 \Downarrow 4}}{\left(\text{if } (\text{plus } 1\ -1)\ 3\ 4\right) \Downarrow 4}$$

Question: could you write this proof..? What would happen if you tried…?

$$\frac{???}{\left(\text{if (plus } 1\ -1)\ 3\ 4 \Downarrow 3\right)}$$

$$\mathbf{If_T} : \frac{e_0 \Downarrow n \quad n \neq 0 \quad e_1 \Downarrow n'}{(\text{if } e_0 \; e_1 \; e_2) \Downarrow n'} \quad \mathbf{If_F} : \frac{e_0 \Downarrow 0 \quad e_2 \Downarrow n'}{(\text{if } e_0 \; e_1 \; e_2) \Downarrow n'}$$

$$\frac{:(}{\left(\text{if (plus } 1 \; -1) \; 3 \; 4\right) \Downarrow 3}$$

Answer: you **can't** write this proof, because IfT will only let you evaluate e1 when e0 is non-0!

$$\frac{???}{\big(\text{plus (plus } 0\ 1)\ 2\big) \Downarrow 3} \qquad \frac{???}{(\text{if } 1\ (\text{div } 1\ 1)\ 2) \Downarrow 1}$$

$$\textbf{Const} : \frac{c \in \mathbb{Q}}{c \Downarrow c} \qquad \textbf{Plus} : \frac{e_0 \Downarrow n_0 \quad e_1 \Downarrow n_1 \quad n' = n_0 + n_1}{(\text{plus } e_0\ e_1) \Downarrow n'}$$

$$\textbf{Div} : \frac{e_0 \Downarrow n_0 \quad e_1 \Downarrow n_1 \quad n' = n_0 / n_1}{(\text{div } e_0\ e_1) \Downarrow n'}$$

$$\textbf{Not}_\mathbf{0} : \frac{e \Downarrow 0}{(\text{not } e) \Downarrow 1} \qquad \textbf{Not}_\mathbf{1} : \frac{e \Downarrow n \quad n \neq 0}{(\text{not } e) \Downarrow 0}$$

$$\textbf{If}_\mathbf{T} : \frac{e_0 \Downarrow n \quad n \neq 0 \quad e_1 \Downarrow n'}{(\text{if } e_0\ e_1\ e_2) \Downarrow n'} \quad \textbf{If}_\mathbf{F} : \frac{e_0 \Downarrow n \quad n = 0 \quad e_2 \Downarrow n'}{(\text{if } e_0\ e_1\ e_2) \Downarrow n'}$$