# CIS 500

# Software Foundations

# Fall 2004

## More on induction

# Reasoning about evaluation

## Induction principles

We've seen three definitions of sets and their associated induction principles:

- Ordinary natural numbers
- Boolean terms
- Arithmetic terms

Given a set defined in BNF notation, it is not too hard to describe the structural induction principle for that set.

For example:

```
t ::= brillig
      tove
      snicker t
      gyre t gimble t
```

What is the structural induction principle for this language?

## More induction principles

However, these are not the only sets that we've defined inductively so far.

We defined the semantics of the boolean and arithmetic languages using inductively defined relations — i.e., inductively defined sets of pairs (of terms).

These sets also have induction principles.

## Induction on evaluation

We can define an induction principle for small-step evaluation. Recall the definition (just for booleans, for now):

$$\texttt{if true then } t_2 \texttt{ else } t_3 \longrightarrow t_2 \qquad \text{E-IfTrue}$$

$$\texttt{if false then } t_2 \texttt{ else } t_3 \longrightarrow t_3 \qquad \text{E-IfFalse}$$

$$\frac{t_1 \longrightarrow t_1'}{\texttt{if } t_1 \texttt{ then } t_2 \texttt{ else } t_3 \longrightarrow \texttt{if } t_1' \texttt{ then } t_2 \texttt{ else } t_3} \qquad \text{E-If}$$

What is the induction principle for this relation?

## Induction on evaluation

Induction principle for the evaluation relation:

Suppose $P$ is a property of pairs of terms.

If we can show

- $P(\texttt{if true then } t_2 \texttt{ else } t_3, t_2)$ for all $t_2$ and $t_3$, and
- $P(\texttt{if false then } t_2 \texttt{ else } t_3, t_3)$ for all $t_2$ and $t_3$, and
- $P(\texttt{if } t_1 \texttt{ then } t_2 \texttt{ else } t_3, \texttt{if } t_1' \texttt{ then } t_2 \texttt{ else } t_3)$ for all $t_1$, $t_2$, and $t_3$ with $P(t_1, t_1')$,

then we may conclude that $P(t, t')$ for all $t$ and $t'$ such that $t \to t'$.

## Derivations

Another way to look at induction on evaluation is in terms of derivations.

A derivation records the "justification" for a particular pair of terms that are in the evaluation relation, in the form of a tree. We've already seen one example on the board last time.

Terminology:

- These trees are called derivation trees (or just derivations)
- The final statement in a derivation tree is its conclusion
- We say that a derivation is proof of its conclusion (or a witness for its conclusion) — it records the reasoning steps that justify the conclusion

Saying that "$t \to t'$" (i.e., "the pair $(t, t')$ is in the relation $\to$") is equivalent to saying "there exists an evaluation derivation $\mathcal{D}$ whose conclusion is $t \to t'$."

## Observation

Lemma: Suppose we are given a derivation $\mathcal{D}$ witnessing the pair $(\mathtt{t}, \mathtt{t}')$ in the $\longrightarrow$ relation. Then exactly one of the following holds:

1. the final rule used in $\mathcal{D}$ is E-IfTrue and $\mathtt{t} = \mathtt{if\ true\ then\ t_2\ else\ t_3}$ and $\mathtt{t}' = \mathtt{t_2}$ for some $\mathtt{t_2}$ and $\mathtt{t_3}$; or
2. the final rule used in $\mathcal{D}$ is E-IfFalse and $\mathtt{t} = \mathtt{if\ false\ then\ t_2\ else\ t_3}$ and $\mathtt{t}' = \mathtt{t_3}$ for some $\mathtt{t_2}$ and $\mathtt{t_3}$; or
3. the final rule used in $\mathcal{D}$ is E-If and $\mathtt{t} = \mathtt{if\ t_1\ then\ t_2\ else\ t_3}$ and $\mathtt{t}' = \mathtt{if\ t_1'\ then\ t_2\ else\ t_3}$, for some $\mathtt{t_1}, \mathtt{t_1'}, \mathtt{t_2}$ and $\mathtt{t_3}$; moreover the immediate subderivation of $\mathcal{D}$ witnesses $\mathtt{t_1} \longrightarrow \mathtt{t_1'}$.

## Induction on Derivations

We can now write proofs about evaluation "by induction on derivation trees."

Given an arbitrary derivation $\mathcal{D}$ with conclusion $t \to t'$, we assume the desired property $P$ for its immediate sub-derivations (if any) and try to show that $P$ holds for $\mathcal{D}$ itself, using a case analysis (applying the previous lemma) of the final evaluation rule used in constructing the derivation tree.

E.g....

## Induction on small-step evaluation

For example, let us show that small-step evaluation is deterministic.

Theorem: If $t \to t'$ and $t \to t''$ then $t' = t''$.

Proof: By induction on a derivation $\mathcal{D}$ of $t \to t'$. (Check: exactly what is $\mathbf{P}$ here?)

1. Suppose the final rule used in $\mathcal{D}$ is E-IfTrue, with $t = \texttt{if } t_1 \texttt{ then } t_2 \texttt{ else } t_3$ and $t_1 = \texttt{true}$ and $t' = t_2$. Then the last rule of the derivation of $t \to t'$ cannot be E-IfFalse, because $t_1$ is not $\texttt{false}$. Furthermore, the last rule cannot be E-If either, because this rule requires that $t_1 \to t_1'$, and $\texttt{true}$ does not step to anything. So the last rule can only be E-IfTrue, and $t' = t''$.

2. Suppose the final rule used in $\mathcal{D}$ is E-IfFalse, with $t = \texttt{if false then } t_2 \texttt{ else } t_3$ and $t' = t_3$. This case is similar to the previous.

3. Suppose the final rule used in $\mathcal{D}$ is E-IF, with $t = \texttt{if } t_1 \texttt{ then } t_2 \texttt{ else } t_3$ and $t' = \texttt{if } t_1' \texttt{ then } t_2 \texttt{ else } t_3$, where $t_1 \longrightarrow t_1'$ is witnessed by a derivation $\mathcal{D}_1$. The last rule in the derivation of $t \longrightarrow t''$ can only be E-If, so it must be that $t_1 \longrightarrow t_1''$. By the inductive hypothesis, $t_1' = t_1''$, from which we conclude $t' = t''$.

## What principle to use?

We've proven the same theorem using two different induction principles.

Q: Which one is the best one to use in a given case?

A: The one that works in that case!

For these simple languages, anything you can prove by induction on derivations of $\mathtt{t} \longrightarrow \mathtt{t}'$, you can also prove by structural induction on $\mathtt{t}$. But that will not be the case for every language.

# Well-founded induction

## A Sceptic Asks...

**Question:** Why are any of these induction principles true? Why should I believe a proof that employs one?

**Answer:** These are all instances of a general principle called **well-founded induction**.

## Well-founded induction

Let $\prec$ be a well-founded relation on a set $A$ and let $P$ be a property. If

$$\forall a \in A.\quad [\forall b \prec a.\, P(b)] \Rightarrow P(a)$$

then $\forall a \in A.\, P(a)$.

Choosing the set $A$ and relation $\prec$ determines the induction principle.

## Well-founded induction

For example, we let $A = \mathcal{N}$ and $n \prec m \stackrel{\text{def}}{=} m = n + 1$. In this case, we can rewrite previous principle as:

If

$$\forall a \in \mathcal{N}.([\forall b \prec a.P(b)] \Rightarrow P(a)$$

then $\forall a \in \mathcal{N}.P(a)$.

Now, by definition $a$ is either $0$ or $i + 1$ for some $i$:

If

$$[\forall b \prec 0.P(b)] \Rightarrow P(0) \wedge$$
$$\forall i \in \mathcal{N}.[\forall b \prec i + 1.P(b)] \Rightarrow P(i + 1)$$

then $\forall a \in \mathcal{N}.P(a)$.

Or, simplifying:

If $P(0)$ and $\forall i \in \mathcal{N}.P(i) \Rightarrow P(i + 1)$ then $\forall a \in \mathcal{N}.P(a)$.

## Strong induction

If we take $\prec$ to be the "strictly less than" relation $<$ on natural numbers, then the principle we get is strong (or "complete") induction:

If

$$\forall a \in \mathcal{N}.([\forall b < a.P(b)] \Rightarrow P(a)$$

then $\forall a \in \mathcal{N}.P(a)$.

## Well-founded relation

The induction principle holds *only* when the relation $\prec$ is well-founded.

Definition: A well-founded relation is a binary relation $\prec$ on a set $A$ such that there are no infinite descending chains $\cdots \prec a_i \prec \cdots \prec a_1 \prec a_0$.

Are the successor and $>$ relations well-founded?

## Validity of well-founded induction

**Theorem:** Let $\prec$ is a well-founded relation on a set $A$. Let $P$ be a property. Then $\forall a \in A.P(a)$ iff

$$\forall a \in A.([\forall b \prec a.P(b)] \Rightarrow P(a)$$

**Proof:** The ($\Rightarrow$) direction is trivial. We'll show the ($\Leftarrow$) direction.

First, observe that any nonempty subset Q of A has a minimal element, even if Q is infinite.

Now, suppose $\neg P(a)$ for some $a$ in $A$. There must be a minimal element $m$ of the set $\{a \in A | \neg P(a)\}$. But then, $\neg P(m)$ yet $[\forall b \prec m.P(b)]$ which is a contradiction.

## Structural induction

Well-founded induction also generalizes structural induction.

If $\prec$ is the "immediate subterm" relation, then the principle we get is structural induction for terms.

For example, in Arith, the term $t_1$ is an immediate subterm of the term succ $t_1$.

Is the immediate subterm relation well-founded?

Yes, since all terms of Arith are finite.

## Mathematical Digression

If you want to understandn the full story about induction and inductively know defined relations, check out the beginning of Chapter 21 in TAPL.

# Termination of evaluation

## Termination of evaluation

Theorem: For every $t$ there is some normal form $t'$ such that $t \longrightarrow^* t'$.

How can we prove it??

## An Inductive Definition of a Function

We can define the size of a term with the following relation:

$$
\begin{array}{lcl}
\mathsf{size}(\texttt{true}) & = & 1 \\
\mathsf{size}(\texttt{false}) & = & 1 \\
\mathsf{size}(0) & = & 1 \\
\mathsf{size}(\texttt{succ } t_1) & = & \mathsf{size}(t_1) + 1 \\
\mathsf{size}(\texttt{pred } t_1) & = & \mathsf{size}(t_1) + 1 \\
\mathsf{size}(\texttt{iszero } t_1) & = & \mathsf{size}(t_1) + 1 \\
\mathsf{size}(\texttt{if } t_1 \texttt{ then } t_2 \texttt{ else } t_3) & = & \mathsf{size}(t_1) + \mathsf{size}(t_2) + \mathsf{size}(t_3) + 1
\end{array}
$$

Note: this is yet more shorthand. How would we write this definition with inference rules?

## Induction on Derivations — Another Example

**Theorem:** If $t \longrightarrow t'$, then $\mathsf{size}(t) > \mathsf{size}(t')$.

**Proof:** By induction on a derivation $\mathcal{D}$ of $t \longrightarrow t'$.

1. Suppose the final rule used in $\mathcal{D}$ is E-IfTrue, with $t = \texttt{if true then } t_2 \texttt{ else } t_3$ and $t' = t_2$. Then the result is immediate from the definition of size.
2. Suppose the final rule used in $\mathcal{D}$ is E-IfFalse, with $t = \texttt{if false then } t_2 \texttt{ else } t_3$ and $t' = t_3$. Then the result is again immediate from the definition of size.
3. Suppose the final rule used in $\mathcal{D}$ is E-If, with $t = \texttt{if } t_1 \texttt{ then } t_2 \texttt{ else } t_3$ and $t' = \texttt{if } t_1' \texttt{ then } t_2 \texttt{ else } t_3$, where $(t_1, t_1') \in \longrightarrow$ is witnessed by a derivation $\mathcal{D}_1$. By the induction hypothesis, $\mathsf{size}(t_1) > \mathsf{size}(t_1')$. But then, by the definition of size, we have $\mathsf{size}(t) > \mathsf{size}(t')$.

## Termination of evaluation

**Theorem:** For every $t$ there is some normal form $t'$ such that $t \longrightarrow^* t'$.

**Proof:**

- First, recall that single-step evaluation strictly reduces the size of the term:

  if $t \longrightarrow t'$, then $\mathsf{size}(t) > \mathsf{size}(t')$

- Now, assume (for a contradiction) that

  $t_0, t_1, t_2, t_3, t_4, \ldots$

  is an infinite-length sequence such that

  $t_0 \longrightarrow t_1 \longrightarrow t_2 \longrightarrow t_3 \longrightarrow t_4 \longrightarrow \cdots$,

- Then

  $\mathsf{size}(t_0), \mathsf{size}(t_1), \mathsf{size}(t_2), \mathsf{size}(t_3), \mathsf{size}(t_4), \ldots$

  is an infinite, strictly decreasing, sequence of natural numbers.

- But such a sequence cannot exist — contradiction!

## Termination Proofs

Most termination proofs have the same basic form:

> **Theorem:** The relation $R \subseteq X \times X$ is terminating — i.e., there are no infinite sequences $x_0$, $x_1$, $x_2$, etc. such that $(x_i, x_{i+1}) \in R$ for each $i$.
>
> **Proof:**
>
> 1. Choose
>    - a well-founded set $(W, <)$ — i.e., a set $W$ with a partial order $<$ such that there are no infinite descending chains $w_0 > w_1 > w_2 > \ldots$ in $W$
>    - a function $f$ from $X$ to $W$
> 2. Show $f(x) > f(y)$ for all $(x, y) \in R$
> 3. Conclude that there are no infinite sequences $x_0$, $x_1$, $x_2$, etc. such that $(x_i, x_{i+1}) \in R$ for each $i$, since, if there were, we could construct an infinite descending chain in $W$.