

x = 1

let x = 1 in ...

x(1).

!x(1)

x.set(1)

Programming Language Theory

Small-step Operational Semantics (aka Structured Operational Semantics)

Ralf Lämmel

Big-step style

$$[\text{comp}_{\text{ns}}] \quad \frac{\langle S_1, s \rangle \rightarrow s', \langle S_2, s' \rangle \rightarrow s''}{\langle S_1; S_2, s \rangle \rightarrow s''}$$



Easier to understand

Small-step style

$$[\text{comp}_{\text{sos}}^1] \quad \frac{\langle S_1, s \rangle \Rightarrow \langle S'_1, s' \rangle}{\langle S_1; S_2, s \rangle \Rightarrow \langle S'_1; S_2, s' \rangle}$$



“More versatile”

$$[\text{comp}_{\text{sos}}^2] \quad \frac{\langle S_1, s \rangle \Rightarrow s'}{\langle S_1; S_2, s \rangle \Rightarrow \langle S_2, s' \rangle}$$

SOS (statements)

$$[\text{ass}_{\text{SOS}}] \quad \langle x := a, s \rangle \Rightarrow s[x \mapsto \mathcal{A}[[a]]s]$$

$$[\text{skip}_{\text{SOS}}] \quad \langle \text{skip}, s \rangle \Rightarrow s$$

$$[\text{comp}_{\text{SOS}}^1] \quad \frac{\langle S_1, s \rangle \Rightarrow \langle S'_1, s' \rangle}{\langle S_1; S_2, s \rangle \Rightarrow \langle S'_1; S_2, s' \rangle}$$

$$[\text{comp}_{\text{SOS}}^2] \quad \frac{\langle S_1, s \rangle \Rightarrow s'}{\langle S_1; S_2, s \rangle \Rightarrow \langle S_2, s' \rangle}$$

$$[\text{if}_{\text{SOS}}^{\text{tt}}] \quad \langle \text{if } b \text{ then } S_1 \text{ else } S_2, s \rangle \Rightarrow \langle S_1, s \rangle \text{ if } \mathcal{B}[[b]]s = \text{tt}$$

$$[\text{if}_{\text{SOS}}^{\text{ff}}] \quad \langle \text{if } b \text{ then } S_1 \text{ else } S_2, s \rangle \Rightarrow \langle S_2, s \rangle \text{ if } \mathcal{B}[[b]]s = \text{ff}$$

$$[\text{while}_{\text{SOS}}] \quad \langle \text{while } b \text{ do } S, s \rangle \Rightarrow \\ \langle \text{if } b \text{ then } (S; \text{while } b \text{ do } S) \text{ else skip}, s \rangle$$

Prolog as a sandbox for small-step operational semantics

<https://slps.svn.sourceforge.net/svnroot/slps/topics/NielsonN07/Prolog/While/SOS/>

Architecture of the interpreter

- **Makefile**: see “make test”
- **main.pro**: main module to compose all other modules
- **exec.pro**: statement execution

- **eval.pro**: expression evaluation
- **map.pro**: abstract data type for maps (states)
- **test.pro**: framework for unit testing

Same as
NS

Empty statement

step((skip,M),
M).

Sequential composition

$\text{step}(\text{seq}(S1,S2),M1),$
 $\text{seq}(S3,S2),M2)$

:-

$\text{step}((S1,M1),(S3,M2)).$

$\text{step}(\text{seq}(S1,S2),M1),$
 $S2,M2)$

:-

$\text{step}((S1,M1),M2),$
 $\backslash + M2 = (_,_).$

Assignment

step((assign(X,A),M1),
M2)

:-

evala(A,M1,Y),
update(M1,X,Y,M2).

Conditional statement

step((ifthenelse(B,S1,_),M),
 (S1,M))

:-

 evalb(B,M,tt).

step((ifthenelse(B,_,S2),M),
 (S2,M))

:-

 evalb(B,M,ff).

Loop statement

step((while(B,S),M),
 (ifthenelse(B,seq(S,while(B,S)),skip),M)).

Transitive closure

```
execute( (S1,M1),  
        M3 )
```

```
:-
```

```
    step((S1,M1),(S2,M2)),  
    execute((S2,M2),M3).
```

```
execute( (S1,M1),  
        M2 )
```

```
:-
```

```
    step((S1,M1),M2),  
    \+ M2 = (_,_).
```

Transition systems in semantics

Semantics of statements

Syntactic Category

$$\begin{aligned} S ::= & x := a \mid \text{skip} \mid S_1; S_2 \\ & \mid \text{if } b \text{ then } S_1 \text{ else } S_2 \\ & \mid \text{while } b \text{ do } S \end{aligned}$$

Meaning of the syntactic category:

$$\mathcal{S} : \text{Stm} \rightarrow (\text{State} \leftrightarrow \text{State})$$

Two operational semantics

- Natural Semantics
- Structural Operational Semantics

specified by transition systems

Transition systems

$(\Gamma, T, \triangleright)$

- Γ : a set of configurations
- T : a set of terminal configurations
 $T \subseteq \Gamma$
- \triangleright : a transition relation
 $\triangleright \subseteq \Gamma \times \Gamma$

Big step versus small step

- Big-step semantics
 - ◆ aka Natural semantics
 - ◆ **One** (fewer) transition(s)
 - ◆ *Computation steps modeled by derivation tree*
- Small-step semantics
 - ◆ aka Structured Operational Semantics (SOS)
 - ◆ **Many** transitions
 - ◆ *Computation steps modeled by transitions*

Big step operational semantics: describe how the “final” result of the computation is obtained.

Transition system: (Γ, T, \rightarrow)

- $\Gamma = \{(S, s) \mid S \in \text{While}, s \in \text{State}\} \cup \text{State}$
- $T = \text{State}$
- $\rightarrow \subseteq \{(S, s) \mid S \in \text{While}, s \in \text{State}\} \times \text{State}$

Typical transition:

$$(S, s) \rightarrow s'$$

where

S is the program

s is the initial state

s' is the final state

Small step operational semantics: describe how the individual steps of the computation take place.

Transition system: (Γ, T, \Rightarrow)

- $\Gamma = \{(S, s) \mid S \in \text{While}, s \in \text{State}\} \cup \text{State}$
- $T = \text{State}$
- $\Rightarrow \subseteq \{(S, s) \mid S \in \text{While}, s \in \text{State}\} \times \Gamma$

Two typical transitions:

- the computation has not been completed after one step of computation:
 $(S, s) \Rightarrow (S', s')$
- the computation is completed after one step of computation:
 $(S, s) \Rightarrow s'$

Big step versus small step

$$\langle z:=x, s_0 \rangle \rightarrow s_1 \quad \langle x:=y, s_1 \rangle \rightarrow s_2$$

$$\langle z:=x; x:=y, s_0 \rangle \rightarrow s_2$$

$$\langle y:=z, s_2 \rangle \rightarrow s_3$$

$$\langle z:=x; x:=y; y:=z, s_0 \rangle \rightarrow s_3$$

Derivation
tree

Transition
= big step

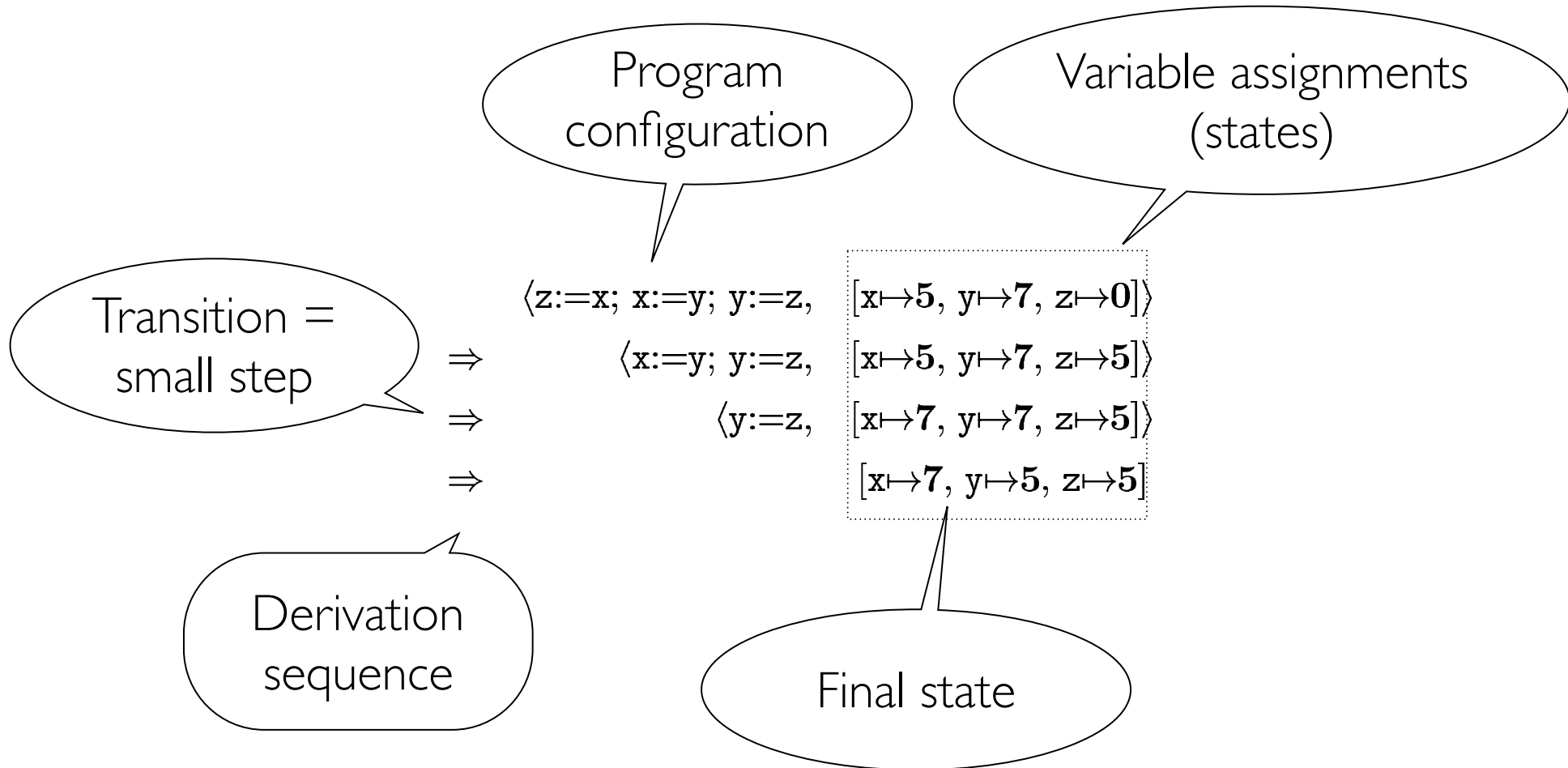
$$s_0 = [x \mapsto 5, y \mapsto 7, z \mapsto 0]$$

$$s_1 = [x \mapsto 5, y \mapsto 7, z \mapsto 5]$$

$$s_2 = [x \mapsto 7, y \mapsto 7, z \mapsto 5]$$

$$s_3 = [x \mapsto 7, y \mapsto 5, z \mapsto 5]$$

Big step versus **small step**



Big step versus **small step**

Derivation sequence
(many transitions)

$$\begin{aligned} &\langle (z := x; x := y); y := z, s_0 \rangle \\ &\Rightarrow \langle x := y; y := z, s_0[z \mapsto 5] \rangle \\ &\Rightarrow \langle y := z, (s_0[z \mapsto 5])[x \mapsto 7] \rangle \\ &\Rightarrow ((s_0[z \mapsto 5])[x \mapsto 7])[y \mapsto 5] \end{aligned}$$

Derivation tree
(for each single step)

$$\frac{\langle z := x, s_0 \rangle \Rightarrow s_0[z \mapsto 5]}{\langle z := x; x := y, s_0 \rangle \Rightarrow \langle x := y, s_0[z \mapsto 5] \rangle}$$
$$\langle (z := x; x := y); y := z, s_0 \rangle \Rightarrow \langle x := y; y := z, s_0[z \mapsto 5] \rangle$$

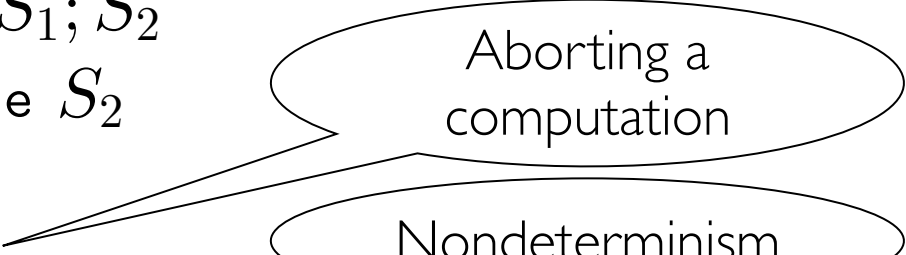
Execution of $\langle S, s \rangle$ terminates successfully if $\langle S, s \rangle \Rightarrow^k s'$ for some k and s' .

Execution loops if there is an infinite derivation sequence.

Extensions of *While*

$$S ::= x := a \mid \text{skip} \mid S_1; S_2$$
$$\mid \text{if } b \text{ then } S_1 \text{ else } S_2$$
$$\mid \text{while } b \text{ do } S$$

$$\mid \text{abort}$$
$$\mid S_1 \text{ or } S_2$$
$$\mid S_1 \text{ par } S_2$$



Aborting a computation



Nondeterminism



Parallelism

Adding **abort**

Configurations:

$$\{(S, s) \mid S \in \text{While}^{\text{abort}}, s \in \text{State}\} \\ \cup \text{State}$$

Transition relation for NS:

unchanged

Transition relation for SOS:

unchanged

NS vs. SOS

abort vs. while true do skip ?

- **Natural semantics:** We cannot distinguish between abnormal termination and nontermination. (One could extend the set of final configurations to specifically distinguish “stuck” configurations due to abort.)
- **SOS:** Nontermination is reflected by infinite derivation sequences while abortion is reflected by finite derivation sequences ending in a “stuck” configuration.

Adding nondeterminism

$x := 1 \text{ or } (x := 2; x := x + 2)$ assigns 1 or 4 to x .

$$[\text{or}_{\text{sos}}^1] \quad \langle S_1 \text{ or } S_2, s \rangle \Rightarrow \langle S_1, s \rangle$$

$$[\text{or}_{\text{sos}}^2] \quad \langle S_1 \text{ or } S_2, s \rangle \Rightarrow \langle S_2, s \rangle$$

$$[\text{or}_{\text{ns}}^1] \quad \frac{\langle S_1, s \rangle \rightarrow s'}{\langle S_1 \text{ or } S_2, s \rangle \rightarrow s'}$$

$$[\text{or}_{\text{ns}}^2] \quad \frac{\langle S_2, s \rangle \rightarrow s'}{\langle S_1 \text{ or } S_2, s \rangle \rightarrow s'}$$

NS vs. SOS

Does the following program terminate?

`(while true do skip) or (x := 2; x := x+2)`

- **Natural semantics:** Nondeterminism suppresses looping, if possible. That is, we obtain *one* derivation tree (transition) for the terminating option.
- **SOS:** Nondeterminism does not suppress looping. That is, we obtain *two* transition sequences, and one of them is non-terminating.

Adding parallelism

$x := 1 \text{ par } (x := 2; x := x+2)$ assigns 1, 3, or 4 to x .

Transition relation for SOS:

$$\frac{(S_1, s) \Rightarrow (S'_1, s')}{(S_1 \text{ par } S_2, s) \Rightarrow (S'_1 \text{ par } S_2, s')}$$

$$\frac{(S_1, s) \Rightarrow s'}{(S_1 \text{ par } S_2, s) \Rightarrow (S_2, s')}$$

$$\frac{(S_2, s) \Rightarrow (S'_2, s')}{(S_1 \text{ par } S_2, s) \Rightarrow (S_1 \text{ par } S'_2, s')}$$

$$\frac{(S_2, s) \Rightarrow s'}{(S_1 \text{ par } S_2, s) \Rightarrow (S_1, s')}$$

Transition relation for NS:

$$\frac{(S_1, s) \rightarrow s', (S_2, s') \rightarrow s''}{(S_1 \text{ par } S_2, s) \rightarrow s''}$$

$$\frac{(S_2, s) \rightarrow s', (S_1, s') \rightarrow s''}{(S_1 \text{ par } S_2, s) \rightarrow s''}$$

NS vs. SOS

- **Nat. sem.:** Each constituent of par is executed in one big step. Hence, interleaving of computations is *not* achieved.

$x := 1 \text{ par } (x := 2; x := x+2)$ evaluates to 1, 4.

- **SOS:** The constituents of par are executed in many small steps. Hence interleaving of computations is achieved.

$x := 1 \text{ par } (x := 2; x := x+2)$ evaluates to 1, 3, or 4.

Semantics and proofs

- ◆ Three approaches to semantics:
 - ★ Compositional definitions
 - ★ Natural semantics
 - ★ SOS
- ◆ Three corresponding proof principles:
 - ★ Induction on the syntactic structure
 - ★ Induction on the shape of derivation trees
 - ★ Induction on the length of derivation sequences

Induction on the Length of Derivation Sequences

- 1: Prove that the property holds for all derivation sequences of length 0.
- 2: Prove that the property holds for all other derivation sequences: Assume that the property holds for all derivation sequences of length at most k (this is called the *induction hypothesis*) and show that it holds for derivation sequences of length $k+1$.

Equivalence of semantics

$$\mathcal{S}_{\text{ns}}: \text{Stm} \rightarrow (\text{State} \leftrightarrow \text{State})$$

$$\mathcal{S}_{\text{sos}}: \text{Stm} \rightarrow (\text{State} \leftrightarrow \text{State})$$

$$\mathcal{S}_{\text{ns}}[S]s = \begin{cases} s' & \text{if } \langle S, s \rangle \rightarrow s' \\ \underline{\text{undef}} & \text{otherwise} \end{cases}$$

$$\mathcal{S}_{\text{sos}}[S]s = \begin{cases} s' & \text{if } \langle S, s \rangle \Rightarrow^* s' \\ \underline{\text{undef}} & \text{otherwise} \end{cases}$$

Only for basic
While!

Theorem 2.26 For every statement S of **While** we have $\mathcal{S}_{\text{ns}}[S] = \mathcal{S}_{\text{sos}}[S]$.

Theorem 2.26 For every statement S of **While** we have $\mathcal{S}_{\text{ns}}[S] = \mathcal{S}_{\text{sos}}[S]$.

Proof Summary for While:

Equivalence of two Operational Semantics

- 1: Prove by *induction on the shape of derivation trees* that for each derivation tree in the natural semantics there is a corresponding finite derivation sequence in the structural operational semantics.
- 2: Prove by *induction on the length of derivation sequences* that for each finite derivation sequence in the structural operational semantics there is a corresponding derivation tree in the natural semantics.

Theorem 2.26 For every statement S of **While** we have $\mathcal{S}_{\text{ns}}[[S]] = \mathcal{S}_{\text{sos}}[[S]]$.

Lemma 2.27 For every statement S of **While** and states s and s' we have

$$\langle S, s \rangle \rightarrow s' \text{ implies } \langle S, s \rangle \Rightarrow^* s'.$$

So if the execution of S from s terminates in the natural semantics then it will terminate in the same state in the structural operational semantics.

Lemma 2.28 For every statement S of **While**, states s and s' and natural number k we have that

$$\langle S, s \rangle \Rightarrow^k s' \text{ implies } \langle S, s \rangle \rightarrow s'.$$

So if the execution of S from s terminates in the structural operational semantics then it will terminate in the same state in the natural semantics.



Let's focus on this lemma for the sake of exercising induction on length of derivation sequences.

$$\langle S, s \rangle \Rightarrow^k s' \text{ implies } \langle S, s \rangle \rightarrow s'.$$

Proof: The proof proceeds by induction on the length of the derivation sequence $\langle S, s \rangle \Rightarrow^k s'$, that is by induction on k .

If $k=0$ then the result holds vacuously.

To prove the induction step we assume that the lemma holds for $k \leq k_0$ and we shall then prove that it holds for k_0+1 . We proceed by cases on how the first step of $\langle S, s \rangle \Rightarrow^{k_0+1} s'$ is obtained, that is by inspecting the derivation tree for the first step of computation in the structural operational semantics.

The case $[\text{ass}_{\text{sos}}]$: Straightforward (and $k_0 = 0$).

The case $[\text{skip}_{\text{sos}}]$: Straightforward (and $k_0 = 0$).

Clearly, the cases for compound statement forms somehow have to take apart the k transitions to account for the transitions needed for the constituents.

Taking apart statement composition

Lemma [2.19] If $(S_1; S_2, s) \Rightarrow^k s''$ then
there exists s' , k_1 and k_2 such that

$$(S_1, s) \Rightarrow^{k_1} s',$$
$$(S_2, s') \Rightarrow^{k_2} s'' \text{ and}$$
$$k = k_1 + k_2$$

Proof We proceed by induction on the number k .

Proof by induction on the length of
derivation sequences

Lemma 2.19 If $\langle S_1; S_2, s \rangle \Rightarrow^k s''$ then there exists a state s' and natural numbers k_1 and k_2 such that $\langle S_1, s \rangle \Rightarrow^{k_1} s'$ and $\langle S_2, s' \rangle \Rightarrow^{k_2} s''$ where $k = k_1 + k_2$.

Proof: The proof is by induction on the number k , that is by induction on the length of the derivation sequence $\langle S_1; S_2, s \rangle \Rightarrow^k s''$.

If $k = 0$ then the result holds vacuously.

For the induction step we assume that the lemma holds for $k \leq k_0$ and we shall prove it for $k_0 + 1$. So assume that

$$\langle S_1; S_2, s \rangle \Rightarrow^{k_0+1} s''$$

This means that the derivation sequence can be written as

$$\langle S_1; S_2, s \rangle \Rightarrow \gamma \Rightarrow^{k_0} s''$$

for some configuration γ . Now one of two cases applies depending on which of the two rules $[\text{comp}_{\text{SOS}}^1]$ and $[\text{comp}_{\text{SOS}}^2]$ was used to obtain $\langle S_1; S_2, s \rangle \Rightarrow \gamma$.

In the first case where $[\text{comp}_{\text{sos}}^1]$ is used we have

$$\langle S_1; S_2, s \rangle \Rightarrow \langle S'_1; S_2, s' \rangle$$

because

$$\langle S_1, s \rangle \Rightarrow \langle S'_1, s' \rangle$$

We therefore have

$$\langle S'_1; S_2, s' \rangle \Rightarrow^{k_0} s''$$

and the induction hypothesis can be applied to this derivation sequence because it is shorter than the one we started with. This means that there is a state s_0 and natural numbers k_1 and k_2 such that

$$\langle S'_1, s' \rangle \Rightarrow^{k_1} s_0 \text{ and } \langle S_2, s_0 \rangle \Rightarrow^{k_2} s''$$

where $k_1 + k_2 = k_0$. Using that $\langle S_1, s \rangle \Rightarrow \langle S'_1, s' \rangle$ and $\langle S'_1, s' \rangle \Rightarrow^{k_1} s_0$ we get

$$\langle S_1, s \rangle \Rightarrow^{k_1+1} s_0$$

We have already seen that $\langle S_2, s_0 \rangle \Rightarrow^{k_2} s''$ and since $(k_1+1) + k_2 = k_0 + 1$ we have proved the required result.

$$\frac{\langle S_1, s \rangle \Rightarrow \langle S'_1, s' \rangle}{\langle S_1; S_2, s \rangle \Rightarrow \langle S'_1; S_2, s' \rangle}$$

$$\frac{\langle S_1, s \rangle \Rightarrow s'}{\langle S_1; S_2, s \rangle \Rightarrow \langle S_2, s' \rangle}$$

The second possibility is that $[\text{comp}_{\text{SOS}}^2]$ has been used to obtain the derivation $\langle S_1; S_2, s \rangle \Rightarrow \gamma$. Then we have

$$\langle S_1, s \rangle \Rightarrow s'$$

and γ is $\langle S_2, s' \rangle$ so that

$$\langle S_2, s' \rangle \Rightarrow^{k_0} s''$$

The result now follows by choosing $k_1=1$ and $k_2=k_0$. □

$\langle S, s \rangle \Rightarrow^k s'$ implies $\langle S, s \rangle \rightarrow s'$. cont'd

The cases $[\text{comp}_{\text{SOS}}^1]$ and $[\text{comp}_{\text{SOS}}^2]$: In both cases we assume that

$$\langle S_1; S_2, s \rangle \Rightarrow^{k_0+1} s''$$

We can now apply Lemma 2.19 and get that there exists a state s' and natural numbers k_1 and k_2 such that

$$\langle S_1, s \rangle \Rightarrow^{k_1} s' \text{ and } \langle S_2, s' \rangle \Rightarrow^{k_2} s''$$

where $k_1+k_2=k_0+1$. The induction hypothesis can now be applied to each of these derivation sequences because $k_1 \leq k_0$ and $k_2 \leq k_0$. So we get

$$\langle S_1, s \rangle \rightarrow s' \text{ and } \langle S_2, s' \rangle \rightarrow s''$$

Using $[\text{comp}_{\text{ns}}]$ we now get the required $\langle S_1; S_2, s \rangle \rightarrow s''$.

Further composites omitted.



- **Summary:** *Small-step operational semantics*
 - ♦ *Transitions are steps of computation.*
 - ♦ *Computations are derivation sequences.*
 - ♦ *Some extensions are more convenient with SOS.*
- **Prepping:** *“Semantics with applications”*
 - ♦ *Chapter 2.2 - Chapter 2.5*
- **Lab:** *Operational Semantics in Prolog*
- **Outlook:**
 - ♦ *Type systems*
 - ♦ *The lambda calculus*