Syntaks og semantik

Lektion 9

11 marts 2008

Semantik

Syntaks: Læren om sprogs *form*

- hvordan *ser* et lovligt program *ud*?
- beskriv byggesten (*alfabet*) og hvordan de kan sættes sammen (*grammatik*, *automat* etc.)

Semantik: Læren om sprogs *betydning*

- hvordan *opfører* et givet program sig?
- beskriv *betydningen* af byggesten og hvordan betydningen af sammensætninger af byggesten fås ud fra de enkelte betydninger

- denotationel semantik
  - beskriv et programs betydning som funktion fra *input* til *output*
  - *Hvad* laver det her program?
- operationel semantik
  - beskriv et programs betydning som *transitionssystem*
  - *Hvordan* udføres det her program?
- aksiomatisk semantik
  - beskriv et program ved *præ-* og *post-betingelser*
  - *Hvilke egenskaber* har det her program?
- (algebraisk semantik: variant af aksiomatisk semantik)

- præcis beskrivelse af programmeringssprog
  - “rettesnor” til implementation
- automatisk generering af compilere og fortolkere
- automatisk verifikation af programmer
  - det kan være *dyrt* at finde fejl i et program ved aftestning

  $\Rightarrow$ heller finde fejl *før*

Husk: Definition:

- Et transitionssystem er et par $(\Gamma, \rightarrow)$, hvor delene er
  1. Γ : en mængde af tilstande (eller konfigurationer)
  2. $\rightarrow \subseteq \Gamma \times \Gamma$: transitions-relationen

  – en *orienteret graf*
- Et afmærket transitionssystem er en tripel $(\Gamma, \Sigma, \rightarrow)$, hvor delene er
  1. Γ : en mængde af tilstande (eller konfigurationer)
  2. Σ : en mængde af mærker
  3. $\rightarrow \subseteq \Gamma \times \Sigma \times \Gamma$: transitions-relationen
- De (afmærkede) transitionssystemer vi er interesserede her, har alle specificeret et antal sluttilstande $T \subseteq \Gamma$.
- Nogle gange er vi også interesserede i (afmærkede) transitionssystemer der har en starttilstand $\gamma_0 \in \Gamma$.
- *Hüttels* definition 3.2 inkluderer sluttilstande.
- Jeg har i lektion 4 givet en definition af transitionssystemer med starttilstand, men uden sluttilstande.

- En NFA er et afmærket transitionssystem med start- og sluttilstande $(\Gamma, \Sigma, \gamma_0, T, \rightarrow)$, hvor både Γ og Σ er endelige.

gammel notation	Q	Σ	q_0	F	δ
ny notation	Γ	Σ	γ_0	T	$\rightarrow$

- En DFA er en NFA der er deterministisk, dvs.
  1. $\forall \gamma \in \Gamma : \forall a \in \Sigma : \exists \gamma' \in \Gamma : \gamma \xrightarrow{a} \gamma'$
  2. $\forall \gamma \in \Gamma : \forall a \in \Sigma : \forall \gamma_1', \gamma_2' \in \Gamma : (\gamma \xrightarrow{a} \gamma_1' \wedge \gamma \xrightarrow{a} \gamma_2') \Rightarrow \gamma_1' = \gamma_2'$
- En PDA er et afmærket transitionssystem med start- og sluttilstande $(\Gamma, \Sigma_1 \times \Sigma_2 \times \Sigma_2, \gamma_0, T, \rightarrow)$, hvor Γ, Σ_1 og Σ_2 er endelige.
  - Σ_1 : inputalfabet, Σ_2 : stackalfabet
  - transitioner $\gamma \xrightarrow[b,c]{a} \gamma'$: læs a, pop b, push c
- dvs. transitionssystemer giver en fælles ramme for syntaktisk beskrivelse af NFAs, DFAs og PDAs, nice!
- men hvad med deres semantik?

  Mål: fælles ramme for beskrivelsen af virkemåden for NFA, PDA og en masse andre maskiner

Idé i operationel semantik:

- transitionssystemer (uden mærker) som den mest basale model for beregninger
- "abstrakt maskine"
- modeller (automater, grammatikker, programmeringssprog, ...) gives mening ved at angive hvordan man konverterer dem til transitionssystemer

Eksempel: En operationel semantik for endelige automater:

Givet en NFA $M = (Q, \Sigma, \delta, q_0, F)$:

- konfigurationer: tilstand i Q plus tilbageværende del af inputstrengen
  dvs. $\Gamma = Q \times \Sigma^*$ *(uendeligt mange konfigurationer!)*
- slutkonfigurationer: sluttilstand i F plus tom streng
  dvs. $T = \{(q, \varepsilon) \mid q \in F\}$
- transitioner: at læse et tegn (eller ε) og gå i en anden tilstand
  dvs. $(q, aw) \to (q', w)$ hver gang $q' \in \delta(q, a)$, og for alle $w \in \Sigma^*$

M accepterer en streng w hvis og kun hvis der findes $\gamma \in T$ således at $(q_0, w) \xrightarrow{*} \gamma$.

Eksempel: En operationel semantik for PDAs:

Givet en PDA $M = (Q, \Sigma_1, \Sigma_2, \delta, q_0, F)$ (Σ_2 er stackalfabetet):

- konfigurationer: tilstand i Q plus tilbageværende del af inputstrengen plus stackindhold
  dvs. $\Gamma = Q \times \Sigma_1^* \times \Sigma_2^*$
- slutkonfigurationer: sluttilstand i F plus tom streng plus vilkårlig stackstreng
  dvs. $T = \{(q, \varepsilon, s) \mid q \in F, s \in \Sigma_2^*\}$
- transitioner: at læse et tegn (eller ε) fra input og fra stacken, gå i en anden tilstand og pushe et tegn (eller ε) på stacken
  dvs. $(q, aw, bs) \to (q', w, cs)$ hver gang $(q', c) \in \delta(q, a, b)$, og for alle $w \in \Sigma_1^*$, $s \in \Sigma_2^*$

M accepterer en streng w hvis og kun hvis der findes $\gamma \in T$ således at $(q_0, w, \varepsilon) \xrightarrow{*} \gamma$.

Eksempel: En operationel semantik for *kontekstfrie grammatikker*:

Givet en CFG $G = (V, \Sigma, R, S)$:

- konfigurationer: strenge af variable og terminaler:
  $\Gamma = (V \cup \Sigma)^*$
- slutkonfigurationer: strenge af terminaler:
  $T = \Sigma^*$
- transitioner: derivationsskridt!
  $uAv \Rightarrow uwv$ hvis $A \to w$ er i R

G genererer en streng $w \in T$ hvis og kun hvis $S \overset{*}{\Rightarrow} w$.

Definition 3.11: Lad $(\Gamma, \Longrightarrow, T)$ være et transitionssystem. Transitionsaflukningen i k skridt $\overset{k}{\Longrightarrow}$ er defineret induktivt ved

$$\gamma \overset{0}{\Longrightarrow} \gamma \quad \text{for alle } \gamma$$

$$\gamma \overset{n+1}{\Longrightarrow} \gamma' \quad \text{hvis der findes } \gamma'' \text{ for hvilket } \gamma \Longrightarrow \gamma'' \overset{n}{\Longrightarrow} \gamma'$$

Vi skriver $\gamma \overset{*}{\Longrightarrow} \gamma'$ hvis der findes et k så $\gamma \overset{k}{\Longrightarrow} \gamma'$.

– dvs. $\gamma \overset{k}{\Longrightarrow} \gamma'$ hvis der findes en *transitionsfølge*

$$\gamma \Longrightarrow \gamma_1 \Longrightarrow \gamma_2 \Longrightarrow \ldots \Longrightarrow \gamma_{k-1} \Longrightarrow \gamma'$$

– vi har allerede brugt aflukningen $\overset{*}{\Longrightarrow}$ adskillige gange!

Operationel semantik

$n \in$ **Num** – Numeraler

$x \in$ **Var** – Variable

$a \in$ **Aud** – Aritmetiske udtryk

$b \in$ **Bud** – Boolske udtryk

$S \in$ **Kom** – Kommandoer

$$
\begin{aligned}
S &::= x{:=}a \mid \texttt{skip} \mid S_1 ; S_2 \mid \texttt{if } b \texttt{ then } S_1 \texttt{ else } S_2 \\
&\quad \mid \texttt{while } b \texttt{ do } S \\
b &::= a_1 = a_2 \mid a_1 < a_2 \mid \neg b_1 \mid b_1 \wedge b_2 \mid (\, b_1 \,) \\
a &::= n \mid x \mid a_1 + a_2 \mid a_1 \star a_2 \mid a_1 - a_2 \mid (\, a_1 \,)
\end{aligned}
$$

basiselementer

sammensatte elementer

umiddelbare bestanddele

Aritmetiske udtryk uden variable:

$$\textbf{Aud:} \qquad a ::= n \mid a_1 + a_2 \mid a_1 * a_2 \mid a_1 - a_2 \mid (\ a_1\ )$$

hvor n er et numeral (talord) (en *streng*!), *ikke et tal*

- numeraler skrives $\underline{42}$, tal skrives 42
- *værdien* af $\underline{42}$ er 42
- vi har en *semantisk funktion* $\mathcal{N} : \mathbf{Num} \to \mathbb{Z}$ som giver værdien af en numeral

Big-step-semantik: udtryk evalueres i ét hug

- transitioner fra *udtryk* til *værdier*
- f.x. en transition $(\underline{2}+\underline{4}) * (\underline{6}+\underline{1}) \to 42$

$$[\text{plus}_{\text{bss}}] \quad \frac{a_1 \to v_1 \quad a_2 \to v_2}{a_1 + a_2 \to v} \quad \text{hvor } v = v_1 + v_2$$

$$[\text{minus}_{\text{bss}}] \quad \frac{a_1 \to v_1 \quad a_2 \to v_2}{a_1 - a_2 \to v} \quad \text{hvor } v = v_1 - v_2$$

$$[\text{mult}_{\text{bss}}] \quad \frac{a_1 \to v_1 \quad a_2 \to v_2}{a_1 \star a_2 \to v} \quad \text{hvor } v = v_1 \cdot v_2$$

$$[\text{parent}_{\text{bss}}] \quad \frac{a_1 \to v_1}{(\, a_1 \,) \to v_1}$$

$$[\text{num}_{\text{bss}}] \quad n \to v \quad \text{hvis} \quad \mathcal{N}[\![n]\!] = v$$

præmis

$$[\text{minus}_{\text{bss}}] \quad \frac{a_1 \to v_1 \quad a_2 \to v_2}{a_1 - a_2 \to v} \qquad \text{hvor } v = v_1 - v_2$$

transitionsregel

konklusion

sidebetingelse

aksiom *(transitionsregel uden præmis)*

$$[\text{num}_{\text{bss}}] \qquad n \to v \quad \text{hvis} \quad \mathcal{N}[\![n]\!] = v$$

$$[\text{plus}_{\text{bss}}] \quad \frac{a_1 \to v_1 \quad a_2 \to v_2}{a_1 + a_2 \to v} \quad \text{hvor } v = v_1 + v_2$$

$$[\text{minus}_{\text{bss}}] \quad \frac{a_1 \to v_1 \quad a_2 \to v_2}{a_1 - a_2 \to v} \quad \text{hvor } v = v_1 - v_2$$

$$[\text{mult}_{\text{bss}}] \quad \frac{a_1 \to v_1 \quad a_2 \to v_2}{a_1 \star a_2 \to v} \quad \text{hvor } v = v_1 \cdot v_2$$

$$[\text{parent}_{\text{bss}}] \quad \frac{a_1 \to v_1}{(\, a_1\,) \to v_1}$$

$$[\text{num}_{\text{bss}}] \quad n \to v \quad \text{hvis} \quad \mathcal{N}[\![n]\!] = v$$

Transitionssystemet $(\Gamma, \to, T)$:

- $\Gamma = \textbf{Aud} \cup \mathbb{Z}$, $T = \mathbb{Z}$
- $\to$ består af præcis de transitioner som kan udledes af aksiomerne ved brug af et endeligt antal transitionsregler

At konstruere et derivationstræ for udtrykket $(\underline{2}+\underline{4}) * (\underline{6}+\underline{1})$:

$$(\underline{2}+\underline{4}) * (\underline{6}+\underline{1}) \longrightarrow ?$$

At konstruere et derivationstræ for udtrykket $(\underline{2}+\underline{4}) * (\underline{6}+\underline{1})$:

$$\frac{(\underline{2}+\underline{4}) \longrightarrow ? \qquad\qquad (\underline{6}+\underline{1}) \longrightarrow ?}{(\underline{2}+\underline{4}) * (\underline{6}+\underline{1}) \longrightarrow ?}$$

At konstruere et derivationstræ for udtrykket $(\underline{2}+\underline{4}) \star (\underline{6}+\underline{1})$:

$$\dfrac{\dfrac{\underline{2}+\underline{4} \rightarrow ?}{(\underline{2}+\underline{4}) \rightarrow ?} \qquad \dfrac{\underline{6}+\underline{1} \rightarrow ?}{(\underline{6}+\underline{1}) \rightarrow ?}}{(\underline{2}+\underline{4}) \star (\underline{6}+\underline{1}) \rightarrow ?}$$

At konstruere et derivationstræ for udtrykket $(\underline{2}+\underline{4}) \star (\underline{6}+\underline{1})$:

$$
\frac{
  \dfrac{\dfrac{\underline{2} \rightarrow 2 \qquad \underline{4} \rightarrow 4}{\underline{2}+\underline{4} \rightarrow 6}}{(\underline{2}+\underline{4}) \rightarrow 6}
  \qquad
  \dfrac{\dfrac{\underline{6} \rightarrow 6 \qquad \underline{1} \rightarrow 1}{\underline{6}+\underline{1} \rightarrow 7}}{(\underline{6}+\underline{1}) \rightarrow 7}
}{
  (\underline{2}+\underline{4}) \star (\underline{6}+\underline{1}) \rightarrow 42
}
$$

At konstruere et derivationstræ for udtrykket $(\underline{2}+\underline{4}) \star (\underline{6}+\underline{1})$:

$$\dfrac{\dfrac{\dfrac{\underline{2} \to 2 \qquad \underline{4} \to 4}{\underline{2}+\underline{4} \to 6}}{(\underline{2}+\underline{4}) \to 6} \qquad \dfrac{\dfrac{\underline{6} \to 6 \qquad \underline{1} \to 1}{\underline{6}+\underline{1} \to 7}}{(\underline{6}+\underline{1}) \to 7}}{(\underline{2}+\underline{4}) \star (\underline{6}+\underline{1}) \to 42}$$

derivationstræer:

- aksiomer i bladene
- knude k har sønner $p_1, p_2, \ldots, p_n$ hvis og kun hvis der er en transitionsregel
$$\frac{p_1, p_2, \ldots, p_n}{k}$$

Small-step-semantik: udtryk evalueres et skridt ad gangen

- transitioner fra *udtryk* til *udtryk* og fra *udtryk* til *værdier*
- f.x.

$$\begin{aligned}(\underline{2}+\underline{4})\star(\underline{6}+\underline{1}) &\Rightarrow (2+\underline{4})\star(\underline{6}+\underline{1})\\ &\Rightarrow (2+4)\star(\underline{6}+\underline{1})\\ &\Rightarrow (6)\star(\underline{6}+\underline{1})\end{aligned}$$

- transitionssystem $(\Gamma, \Rightarrow, T)$:
  - $\Gamma = \mathbf{Aud}' \cup \mathbb{Z}$, $T = \mathbb{Z}$
  - $\Rightarrow$ defineret ved transitionsregler *(coming up!)*

Aritmetiske udtryk uden variable, men *med værdier*:

$$\mathbf{Aud}': \qquad a ::= n \mid v \mid a_1 + a_2 \mid a_1 \star a_2 \mid a_1 - a_2 \mid (\ a_1\ )$$

hvor $n \in \mathbf{Num}$ er et numeral og $v \in \mathbb{Z}$ en værdi [plus-1$_{\text{sss}}$] $$\frac{a_1 \Rightarrow a_1'}{a_1 + a_2 \Rightarrow a_1' + a_2}$$

[plus-2$_{\text{sss}}$] $$\frac{a_2 \Rightarrow a_2'}{a_1 + a_2 \Rightarrow a_1 + a_2'}$$

[plus-3$_{\text{sss}}$] $v_1 + v_2 \Rightarrow v$ hvor $v = v_1 + v_2$

[mult-1$_{\text{sss}}$] $$\frac{a_1 \Rightarrow a_1'}{a_1 \star a_2 \Rightarrow a_1' \star a_2}$$

[mult-2$_{\text{sss}}$] $$\frac{a_2 \Rightarrow a_2'}{a_1 \star a_2 \Rightarrow a_1 \star a_2'}$$

[mult-3$_{\text{sss}}$] $v_1 \star v_2 \Rightarrow v$ hvor $v = v_1 \cdot v_2$

[sub-1$_{\text{sss}}$] $$\frac{a_1 \Rightarrow a_1'}{a_1 - a_2 \Rightarrow a_1' - a_2}$$

[sub-2$_{\text{sss}}$] $$\frac{a_2 \Rightarrow a_2'}{a_1 - a_2 \Rightarrow a_1 - a_2'}$$

[sub-3$_{\text{sss}}$] $$v_1 - v_2 \Rightarrow v \quad \text{hvor} \quad v = v_1 - v_2$$

[parent-1$_{\text{sss}}$] $$\frac{a_1 \Rightarrow a_1'}{(a_1) \Rightarrow (a_1')}$$

[parent-2$_{\text{sss}}$] $$(v) \Rightarrow v$$

[num$_{\text{sss}}$] $$n \Rightarrow v \quad \text{hvis} \quad \mathcal{N}[\![n]\!] = v$$

Sætning: Vores big-step- og small-step-semantikker for **Aud** er ækvivalente: Givet $a \in$ **Aud** og $v \in \mathbb{Z}$, da har vi $a \to v$ hvis og kun hvis $a \overset{*}{\Rightarrow} v$. *(Bevis næste gang)*

Definition: En operationel semantik givet ved et transitionssystem $(\Gamma, \to, T)$ kaldes deterministisk hvis $\gamma \to \gamma_1$ og $\gamma \to \gamma_2$ medfører $\gamma_1 = \gamma_2$ for alle $\gamma \in \Gamma$ og $\gamma_1, \gamma_2 \in T$ (!).
Semantikken kaldes deterministisk på lang sigt hvis $\gamma \xrightarrow{*} \gamma_1$ og $\gamma \xrightarrow{*} \gamma_2$ medfører $\gamma_1 = \gamma_2$ for alle $\gamma \in \Gamma$ og $\gamma_1, \gamma_2 \in T$.

Sætning 3.13 / 3.15 : Vores big-step-semantik for **Aud** er *deterministisk*. Vores small-step-semantik for **Aud** er *deterministisk på lang sigt*. *(Bevises senere)*

Opgave π: Vores small-step-semantik for **Aud** er *ikke deterministisk*. Lav den om så den er!

Boolske udtryk:

$$\textbf{Bud:} \qquad b ::= a_1 = a_2 \mid a_1 < a_2 \mid \neg b_1 \mid b_1 \wedge b_2 \mid (\ b_1\ )$$

- transitionssystem $(\textbf{Bud} \cup \{tt, ff\}, \rightarrow_b, \{tt, ff\})$
- tt = sandt, ff = falsk
- $\rightarrow_a$ er transitioner fra **Aud**-transitionssystemet

---

$$[\text{ligmed-1}_{\text{bss}}] \quad \frac{a_1 \rightarrow_a v_1 \quad a_2 \rightarrow_a v_2}{a_1 = a_2 \rightarrow_b tt} \quad \text{hvis } v_1 = v_2$$

$$[\text{ligmed-2}_{\text{bss}}] \quad \frac{a_1 \rightarrow_a v_1 \quad a_2 \rightarrow_a v_2}{a_1 = a_2 \rightarrow_b ff} \quad \text{hvis } v_1 \neq v_2$$

$$[\text{størreend-1}_{\text{bss}}] \quad \frac{a_1 \rightarrow_a v_1 \quad a_2 \rightarrow_a v_2}{a_1 < a_2 \rightarrow_b tt} \quad \text{hvis } v_1 < v_2$$

$$[\text{størreend-2}_{\text{bss}}] \quad \frac{a_1 \rightarrow_a v_1 \quad a_2 \rightarrow_a v_2}{a_1 < a_2 \rightarrow_b ff} \quad \text{hvis } v_1 \not< v_2$$

[ikke-1$_{\text{bss}}$] $$\frac{b \rightarrow_b tt}{\neg b \rightarrow_b ff}$$

[ikke-2$_{\text{bss}}$] $$\frac{b \rightarrow_b ff}{\neg b \rightarrow_b tt}$$

[parent-b$_{\text{bss}}$] $$\frac{b_1 \rightarrow_b v}{(b_1) \rightarrow_b v}$$

[og-1$_{\text{bss}}$] $$\frac{b_1 \rightarrow_b tt \quad b_2 \rightarrow_b tt}{b_1 \wedge b_2 \rightarrow_b tt}$$

[og-2$_{\text{bss}}$] $$\frac{b_1 \rightarrow_b ff}{b_1 \wedge b_2 \rightarrow_b ff}$$

[og-3$_{\text{bss}}$] $$\frac{b_2 \rightarrow_b ff}{b_1 \wedge b_2 \rightarrow_b ff}$$