

Kapitel III

Programmanalyse und Typsysteme

Inhalt Kapitel 3

- 1 Einführung
 - While Sprache und Kontrollflussgraphen
- 2 Reichweite von Zuweisungen
- 3 Datenflussgleichungen
 - Lösung der Datenflussgleichungen
- 4 Constraint-basierte Analyse
- 5 Weitere Analysen
 - Verfügbare Ausdrücke
 - Liveness von Variablen
- 6 Formulierung von Analysen als Typsystem
 - Operationelle Semantik
 - Semantische Korrektheit von Typsystemen
- 7 Typ- und Effektsysteme
 - Funktionale Sprache
 - Typinferenz
 - Polymorphie
 - Typsystem für Kontrollflussanalyse

Motivation

- “Exakte Methoden” wie in den ersten beiden Kapiteln sind auf Systeme mit nicht zu großem Zustandsraum beschränkt.
- Durch entsprechend abstrakte Modellierung kann man zwar den Zustandsraum hinreichend verkleinern; der Modellierungsschritt bleibt dann aber unverifiziert.
- Dieses und das nächste Kapitel befassen sich mit Analysemethoden für Software realistischer Größe.
- Die zu beweisenden Eigenschaften sind allerdings entweder schwächer, oder erfordern menschliche Hilfe.
- Beispiele “schwächerer Eigenschaften”: Abwesenheit von Laufzeitfehlern wie Division durch Null, falscher Speicherzugriff, ungefangene Exceptions, Verletzung von Ressourcenschranken und -protokollen (z.B. Dateizugriff).
- Beispiele für “menschliche Hilfe”: Typannotate, Zusicherungen, Invarianten, formale Beweise.

Einfache While-Sprache

- Arithmetische Ausdrücke:

$$a ::= x \mid n \mid a_1 op_a a_2$$

- Boole'sche Ausdrücke:

$$b ::= \text{true} \mid \text{false} \mid \text{not } b \mid b_1 op_b b_2 \mid a_1 op_r a_2$$

- Programmstücke (statements):

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

Hierbei bezeichnen: op_a arithmetische Operatoren wie $+$; op_b Boole'sche Operatoren wie \wedge ; op_r Relationsoperatoren wie \leq .

x rangiert über integer-wertige Programmvariablen, n über Integerkonstanten, ℓ über Labels (konkret: natürliche Zahlen), die elementare Programmteile markieren. Jedes Label soll nur einmal vorkommen (vgl.: Zeilennummern)

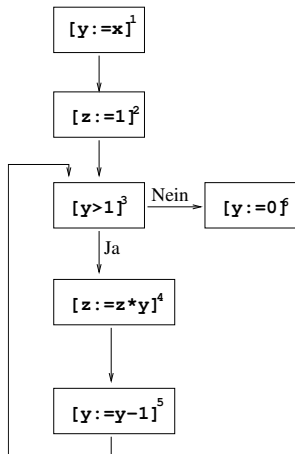
Beispielprogramm

$$[y:=x]^1; [z:=1]^2; \text{while } [y>1]^3 \text{ do } ([z:=z*y]^4; [y:=y-1]^5); [y:=0]^6$$

Alternative Notation:

```
1: y:=x;  
2: z:=1;  
3: while y>1 do (  
4:     z:=z*y;  
5:     y:=y-1);  
6: y:=0
```

Kontrollflussgraph



- Pfade im Kontrollflussgraphen entsprechen möglichen Ausführungssequenzen;
- Erinnerung: Kontrollfluss = “der Computer”, also was und in welcher Reihenfolge es gemacht wird;
- Knoten des Kontrollflussgraphen sind die elementaren Anweisungen, also die, die ein Label tragen;
- Knoten, die unmittelbar aufeinanderfolgen *können*, werden durch Kanten verbunden.
- Bei Fallunterscheidungen können die Kanten mit Antworten beschriftet werden.

Reichweite von Zuweisungen

Definition

Eine *Wertzuweisung* $x := a$ *erreicht* einen bestimmten Programmpunkt, wenn es eine Programmausführung gibt, die diesen Programmpunkt über die Wertzuweisung erreicht und zwischen der Wertzuweisung und dem Programmpunkt die Variable x nicht verändert wird.

Im Beispiel erreicht die Zuweisung $y := x$ bei Label 1 den Eingang zur Instruktion $z := 1$ bei 2.

Für Programmpunkt ℓ bezeichnet $RD_{entry}(\ell)$ die Zuweisungen, die seinen (ℓ 's) Eingang erreichen und $RD_{exit}(\ell)$ die, die seinen Ausgang erreichen.

Detailliertes Beispiel

$$[y:=x]^1; [z:=1]^2; \text{while } [y>1]^3 \text{ do } ([z:=z*y]^4; [y:=y-1]^5); [y:=0]^6$$

Mit $(x, ?)$ bezeichnen wir die Zuweisung eines Startwertes außerhalb des analysierten Programmfragments.

Für Programmpunkte ℓ in denen keine Variable verändert wird, stimmen $RD_{\text{entry}}(\ell)$ und $RD_{\text{exit}}(\ell)$ überein.

ℓ	$RD_{\text{entry}}(\ell)$	$RD_{\text{exit}}(\ell)$
1	$(x, ?), (y, ?), (z, ?)$	$(x, ?), (y, 1), (z, ?)$
2	$(x, ?), (y, 1), (z, ?)$	$(x, ?), (y, 1), (z, 2)$
3	$(x, ?), (y, 1), (y, 5), (z, 2), (z, 4)$	$(x, ?), (y, 1), (y, 5), (z, 2), (z, 4)$
4	$(x, ?), (y, 1), (y, 5), (z, 2), (z, 4)$	$(x, ?), (y, 1), (y, 5), (z, 4)$
5	$(x, ?), (y, 1), (y, 5), (z, 4)$	$(x, ?), (y, 5), (z, 4)$
6	$(x, ?), (y, 1), (y, 5), (z, 2), (z, 4)$	$(x, ?), (y, 6), (z, 2), (z, 4)$

Datenflussgleichungen für RD_{exit}

$$RD_{exit}(1) = (RD_{entry}(1) \setminus \{(y, \ell) \mid \ell = 1..6, ?\}) \cup \{(y, 1)\}$$

$$RD_{exit}(2) = (RD_{entry}(2) \setminus \{(z, \ell) \mid \ell = 1..6, ?\}) \cup \{(z, 2)\}$$

$$RD_{exit}(3) = RD_{entry}(3)$$

$$RD_{exit}(4) = (RD_{entry}(4) \setminus \{(z, \ell) \mid \ell = 1..6, ?\}) \cup \{(z, 4)\}$$

$$RD_{exit}(5) = (RD_{entry}(5) \setminus \{(y, \ell) \mid \ell = 1..6, ?\}) \cup \{(y, 5)\}$$

$$RD_{exit}(6) = (RD_{entry}(6) \setminus \{(y, \ell) \mid \ell = 1..6, ?\}) \cup \{(y, 6)\}$$

Allgemeines Schema:

$$RD_{exit}(i) = (RD_{entry}(i) \setminus \text{“Zuweisungen, die entwertet werden”}) \cup \text{“neu gemachte Zuweisungen”}$$

Datenflussgleichungen für RD_{entry}

$$RD_{entry}(1) = \{(x, ?), (y, ?), (z, ?)\}$$

$$RD_{entry}(2) = RD_{exit}(1)$$

$$RD_{entry}(3) = RD_{exit}(2) \cup RD_{exit}(5)$$

$$RD_{entry}(4) = RD_{exit}(3)$$

$$RD_{entry}(5) = RD_{exit}(4)$$

$$RD_{entry}(6) = RD_{exit}(3)$$

Allgemeines Schema:

$$RD_{entry}(\ell) = \{(x, ?) \mid x \text{ Programmvariable}\} \cup \bigcup_{\ell': \ell' \rightarrow \ell} RD_{exit}(\ell'), \text{ falls } \ell \text{ Einstiegspunkt}$$
$$RD_{entry}(\ell) = \bigcup_{\ell': \ell' \rightarrow \ell} RD_{exit}(\ell'), \text{ sonst}$$

Kleinste Lösung

- Wir interessieren uns für die kleinste Lösung dieser Gleichungen. (“Kleinst” im Sinne von komponentenweiser Inklusion)
- Jede Lösung ist sicher in dem Sinne, dass jede Reichweitenbeziehung auch erfasst wird.
- Es gibt Lösungen, die mehr Reichweitenbeziehungen aufführen, als tatsächlich vorhanden sind, z.B. bleiben die Gleichungen gültig, wenn man zu den RD -Mengen in der Schleife jeweils den Eintrag $(x, 5)$ hinzunimmt.
- Auch die kleinste Lösung kann überflüssige Einträge enthalten. Beispiel:

$$\text{while } [\text{true}]^1 \text{ do } [\text{skip}]^2; [x:=y]^3$$

Hier gilt selbst für die kleinste Lösung $RD_{\text{entry}}(3) = (y, ?)$.

Es gibt kompliziertere Beispiele und im allgemeinen sind die semantisch definierten Reichweiten unentscheidbar.

Berechnung der kleinsten Lösung

Die kleinste Lösung eines Systems von solchen Mengengleichungen kann man berechnen, indem man für jede der gesuchten Mengen RD eine mit \emptyset initialisierte *dynamische Menge* vorsieht und solange Mengen durch die rechten Seiten ihrer definierenden Gleichungen ersetzt, bis nichts mehr passiert.

ℓ	$RD_{entry}(\ell)$	$RD_{exit}(\ell)$		ℓ	$RD_{entry}(\ell)$	$RD_{exit}(\ell)$	
1	\emptyset	\emptyset		1	$(x, ?), (y, ?), (z, ?)$	$(y, 1)$	
2	\emptyset	\emptyset		2	\emptyset	$(z, 2)$	
3	\emptyset	\emptyset	\rightsquigarrow	3	\emptyset	\emptyset	\rightsquigarrow
4	\emptyset	\emptyset		4	\emptyset	$(z, 4)$	
5	\emptyset	\emptyset		5	\emptyset	$(y, 5)$	
6	\emptyset	\emptyset		6	\emptyset	$(y, 6)$	

```
[y:=x]1; [z:=1]2; while [y>1]3 do ([z:=z*y]4; [y:=y-1]5); [y:=0]6
```

Fortsetzung 1

$[y:=x]^1; [z:=1]^2; \text{while } [y>1]^3 \text{ do } ([z:=z*y]^4; [y:=y-1]^5); [y:=0]^6$

ℓ	$RD_{\text{entry}}(\ell)$	$RD_{\text{exit}}(\ell)$
1	$(x, ?), (y, ?), (z, ?)$	$(y, 1)$
2	\emptyset	$(z, 2)$
\rightsquigarrow 3	\emptyset	\emptyset
4	\emptyset	$(z, 4)$
5	\emptyset	$(y, 5)$
6	\emptyset	$(y, 6)$

ℓ	$RD_{\text{entry}}(\ell)$	$RD_{\text{exit}}(\ell)$
1	$(x, ?), (y, ?), (z, ?)$	$(x, ?), (y, 1), (z, ?)$
2	$(x, ?), (y, 1), (z, ?)$	$(x, ?), (y, 1), (z, 2)$
\rightsquigarrow 3	$(x, ?), (z, 2), (y, 1), (y, 5)$	$(x, ?), (z, 2), (y, 1), (y, 5)$
4	$(x, ?), (z, 2), (y, 1), (y, 5)$	$(x, ?), (z, 4), (y, 1), (y, 5)$
5	$(x, ?), (z, 4), (y, 1), (y, 5)$	$(x, ?), (z, 4), (y, 5)$
6	$(x, ?), (z, 2), (y, 1), (y, 5)$	$(x, ?), (z, 2), (y, 6)$

Fortsetzung 2

$[y:=x]^1; [z:=1]^2; \text{while } [y>1]^3 \text{ do } ([z:=z*y]^4; [y:=y-1]^5); [y:=0]^6$

ℓ	$RD_{\text{entry}}(\ell)$	$RD_{\text{exit}}(\ell)$
1	$(x, ?), (y, ?), (z, ?)$	$(x, ?), (y, 1), (z, ?)$
2	$(x, ?), (y, 1), (z, ?)$	$(x, ?), (y, 1), (z, 2)$
\rightsquigarrow 3	$(x, ?), (z, 2), (y, 1), (y, 5)$	$(x, ?), (z, 2), (y, 1), (y, 5)$ \rightsquigarrow
4	$(x, ?), (z, 2), (y, 1), (y, 5)$	$(x, ?), (z, 4), (y, 1), (y, 5)$
5	$(x, ?), (z, 4), (y, 1), (y, 5)$	$(x, ?), (z, 4), (y, 5)$
6	$(x, ?), (z, 2), (y, 1), (y, 5)$	$(x, ?), (z, 2), (y, 6)$
ℓ	$RD_{\text{entry}}(\ell)$	$RD_{\text{exit}}(\ell)$
1	$(x, ?), (y, ?), (z, ?)$	$(x, ?), (y, 1), (z, ?)$
2	$(x, ?), (y, 1), (z, ?)$	$(x, ?), (y, 1), (z, 2)$
\rightsquigarrow 3	$(x, ?), (z, 2), (z, 4), (y, 1), (y, 5)$	$(x, ?), (z, 2), (z, 4), (y, 1), (y, 5)$
4	$(x, ?), (z, 2), (z, 4), (y, 1), (y, 5)$	$(x, ?), (z, 4), (y, 1), (y, 5)$
5	$(x, ?), (z, 4), (y, 1), (y, 5)$	$(x, ?), (z, 4), (y, 5)$
6	$(x, ?), (z, 2), (z, 4), (y, 1), (y, 5)$	$(x, ?), (z, 2), (z, 4), (y, 6)$

Constraint-basierter Ansatz

Anstatt Gleichungen aufzustellen, kann man auch Bedingungen formulieren, denen die unbekannteten Mengen RD_{entry} und RD_{exit} genügen müssen.

Diese Bedingungen (“*constraints*”) unterscheiden sich von den Gleichungen nur durch die Verwendung von \supseteq anstelle von $=$. Z.B. hat man

$$RD_{exit}(5) \supseteq (RD_{entry}(1) \setminus \{(y, \ell) \mid \ell = 1..6, ?\}) \cup \{(y, 5)\}$$

oder noch “konkreter” aber äquivalent:

$$\begin{aligned} RD_{exit}(5) &\supseteq RD_{entry}(5) \setminus \{(y, \ell) \mid \ell = 1..6, ?\} \\ RD_{exit}(5) &\supseteq \{(y, 5)\} \end{aligned}$$

Die erste Bedingung kommt daher, alle eingangs vorhandenen Reichweiten, die nicht überschrieben werden, fortgelten. Die zweite Bedingung kommt von der explizit hinzugekommenen Zuweisung und deren Reichweite.

Constraint-basierter Ansatz, Forts.

Die Bedeutung des Constraint-basierten Ansatzes liegt darin, dass man sie inkrementell erzeugen kann und nicht die gesamte Gleichung für eine linke Seite (hier RD_{exit}) auf einmal aufstellen muss.

Die "Gleichung" $RD_{exit}(5) = \{(y, 5)\}$ wäre ja nicht richtig.

Man löst ein System von Constraints ganz ähnlich wie ein Gleichungssystem: Für jede linke Seite eine mit \emptyset initialisierte dynamische Menge vorhalten; solange die rechten Seiten auswerten und dazugeben, bis sich nichts mehr ändert.

Anwendung: Constant folding

Mit den folgenden Regeln kann man ein gegebenes Programm u.U. verbessern:

- Werte rechte Seiten ohne Variablen zu Konstante aus. Z.B.:
 $x := 5 * 9 - 1 \rightsquigarrow x := 44.$
- Propagiere Konstantendefinitionen: Falls $[x := n]^\ell$ und (x, ℓ) der einzige Eintrag zu x in $RD_{entry}(\ell')$ ist, so ersetze $[y := a]^\ell$ durch $[y := a[n/x]]^{\ell'}$. Analog ersetze $[b]^\ell$ durch $[b[n/x]]^{\ell'}$
Z.B.: Falls $[x := 44]^5$ und $RD_{entry}(7) = \{(x, 5), (y, ?)\}$, dann
 $[y := 2 * x - 3]^7 \rightsquigarrow [y := 85]^7$

Verfügbare Ausdrücke

Wir interessieren uns jetzt dafür, welche arithmetischen Ausdrücke an einem bestimmten Programmpunkt verfügbar sind, also nicht neu berechnet werden müssen, wenn man bereit ist, berechnete Ausdrücke in einer Tabelle abzuspeichern.

$$[x:=a+b]^1; [y:=a*b]^2; \text{while } [y>a+b]^3 \text{ do } ([a:=a+1]^4; [x:=a+b]^5)$$

Beim Test $[y>a+b]^3$ ist der Ausdruck $a+b$ verfügbar, müsste also dort nicht neu berechnet werden. Man könnte das Programm also durch folgende effizientere Version ersetzen:

$$[\text{mem}:=a+b]^{1a}; [x:=\text{mem}]^1; [y:=a*b]^2; \\ \text{while } [y>\text{mem}]^3 \text{ do } ([a:=a+1]^4; [\text{mem}:=a+b]^{5a}; [x:=\text{mem}]^5)$$

Datenflussgleichungen

Wir bezeichnen mit $AE_{entry}(\ell)$ bzw. $AE_{exit}(\ell)$ die bei Ein- und Ausgang eines Labels verfügbaren Ausdrücke. Es gelten die folgenden Datenflussgleichungen:

$$\begin{aligned}
 AE_{entry}(\ell) &= \bigcap_{\ell': \ell' \rightarrow \ell} AE_{exit}(\ell') \\
 AE_{exit}(\ell) &= (AE_{entry}(\ell) \setminus kill_{AE}(B^\ell)) \cup gen_{AE}(B^\ell)
 \end{aligned}$$

wobei B^ℓ das durch ℓ beschriftete Programmstück (“Block”) ist und

$$\begin{aligned}
 kill_{AE}([x:=a]^\ell) &= \{a' \mid x \in FV(a')\} \\
 kill_{AE}([\text{skip}]^\ell) &= \emptyset \\
 kill_{AE}([b]^\ell) &= \emptyset \\
 gen_{AE}([x:=a]^\ell) &= \{a' \mid a' \text{ Teilausdruck von } a \text{ und } x \notin FV(a)\} \\
 gen_{AE}([\text{skip}]^\ell) &= \emptyset \\
 gen_{AE}([b]^\ell) &= \{a' \mid a' \text{ Teilausdruck von } b\}
 \end{aligned}$$

Datenflussgleichungen im Beispiel

$[x:=a+b]^1; [y:=a*b]^2; \text{while } [y>a+b]^3 \text{ do } ([a:=a+1]^4; [x:=a+b]^5)$

ℓ	$kill_{AE}$	gen_{AE}
1	\emptyset	$\{a+b\}$
2	\emptyset	$\{a*b\}$
3	\emptyset	$\{a+b\}$
4	$\{a+b, a*b, a+1\}$	\emptyset
5	\emptyset	$\{a+b\}$

$$\begin{aligned}
 AE_{entry}(1) &= \emptyset \\
 AE_{entry}(2) &= AE_{exit}(1) \\
 AE_{entry}(3) &= AE_{exit}(2) \cap AE_{exit}(5) \\
 AE_{entry}(4) &= AE_{exit}(3) \\
 AE_{entry}(5) &= AE_{exit}(4)
 \end{aligned}$$

$$\begin{aligned}
 AE_{exit}(1) &= AE_{entry}(1) \cup \{a+b\} \\
 AE_{exit}(2) &= AE_{entry}(2) \cup \{a*b\} \\
 AE_{exit}(3) &= AE_{entry}(3) \cup \{a+b\} \\
 AE_{exit}(4) &= AE_{entry}(4) \setminus \{a+b, a*b, a+1\} \\
 AE_{exit}(5) &= AE_{entry}(5) \cup \{a+b\}
 \end{aligned}$$

Größte Lösung

Auch hier ist jede Lösung der Datenflussgleichung gerechtfertigt, allerdings sind wir an möglichst großen AE -Mengen interessiert. Wir berechnen daher die *größte* Lösung. Man erhält sie wie die kleinste Lösung, initialisiert allerdings die dynamischen Mengen jeweils mit der Menge *aller Ausdrücke* (die im Programm vorkommen) anstatt mit \emptyset .

Man erhält:

ℓ	$AE_{entry}(\ell)$	$AE_{exit}(\ell)$
1	\emptyset	$\{a+b\}$
2	$\{a+b\}$	$\{a+b, a*b\}$
3	$\{a+b\}$	$\{a+b\}$
4	$\{a+b\}$	\emptyset
5	\emptyset	$\{a+b\}$

Größte vs. kleinste Lösung

Im Beispiel stimmen größte und kleinste Lösung zufällig überein.
 Das liegt daran, dass die Schleife alle Ausdrücke "killed".
 Anders bei diesem Beispiel:

```
[x:=a+b]1; while [x>y]2 do [x:=x-1]3;
```

Hier gelten die Gleichungen

$$\begin{aligned} AE_{entry}(1) &= \emptyset \\ AE_{entry}(2) &= AE_{exit}(1) \cap AE_{exit}(4) \\ AE_{entry}(3) &= AE_{exit}(2) \end{aligned}$$

$$\begin{aligned} AE_{exit}(1) &= AE_{entry}(1) \cup \{a+b\} \\ AE_{exit}(2) &= AE_{entry}(2) \\ AE_{exit}(3) &= AE_{entry}(3) \end{aligned}$$

Die größte Lösung liefert $AE_{exit}(3) = \{a+b\}$ während bei der kleinsten Lösung $AE_{exit}(3) = \emptyset$ gilt.

Liveness von Variablen

Eine Programmvariable ist an einem bestimmten Programmpunkt *live*, wenn ihr vor diesem Programmpunkt ein Wert zugewiesen wurde und dieser Wert später noch gebraucht wird.

Anders gesagt, kann man eine Variable, die *nicht* live ist, ungestraft überschreiben.

Anwendungen:

- Zuweisungen zu Variablen, die nach der Zuweisung nicht live sind, sind redundant und können eliminiert werden.
- Zwei Variablen, die nie gleichzeitig live sind, können zu einer einzigen verschmolzen werden.

Beispiel:

```
[x:=2]1; [y:=4]2; [x:=1]3; (if [y>x]4 then [z:=y]5 else [z:=y*y]6); [x:=z]7
```

Live-Variablen Gleichungen

Wir bezeichnen mit $LV_{entry}(\ell)$ bzw. $LV_{exit}(\ell)$ die Mengen der Variablen, die bei Ein- und Ausgang eines Labels live sind. Es gelten die folgenden Datenflussgleichungen:

$$\begin{aligned} LV_{exit}(\ell) &= \bigcup_{\ell': \ell \rightarrow \ell'} LV_{entry}(\ell') \\ LV_{entry}(\ell) &= (LV_{exit}(\ell) \setminus kill_{LV}(B^\ell)) \cup gen_{LV}(B^\ell) \end{aligned}$$

$$\begin{aligned} kill_{LV}([x:=a]^\ell) &= \{x\} \\ kill_{LV}([\text{skip}]^\ell) &= kill_{LV}([b]^\ell) = gen_{LV}([\text{skip}]^\ell) = \emptyset \\ gen_{LV}([x:=a]^\ell) &= FV(a) \\ gen_{LV}([b]^\ell) &= FV(b) \end{aligned}$$

Es handelt sich um eine Rückwärtsanalyse, da Information entgegen der Kanten im Kontrollflussgraphen propagiert wird. Wir gehen davon aus, dass am Programmende keine Variable live ist. Alternativ können wir bestimmte Variablen als "Output" deklarieren und dann zur entsprechenden LV_{exit} hinzugeben.

Datenflussgleichungen im Beispiel

$[x:=2]^1; [y:=4]^2; [x:=1]^3; (\text{if } [y>x]^4 \text{ then } [z:=y]^5 \text{ else } [z:=y*y]^6); [x:=z]^7$

$$\begin{aligned}
 LV_{\text{entry}}(1) &= LV_{\text{exit}}(1) \setminus \{x\} \\
 LV_{\text{entry}}(2) &= LV_{\text{exit}}(2) \setminus \{y\} \\
 LV_{\text{entry}}(3) &= LV_{\text{exit}}(3) \setminus \{x\} \\
 LV_{\text{entry}}(4) &= LV_{\text{exit}}(4) \cup \{x, y\}
 \end{aligned}$$

$$\begin{aligned}
 LV_{\text{entry}}(5) &= (LV_{\text{exit}}(5) \setminus \{z\}) \cup \{y\} \\
 LV_{\text{entry}}(6) &= (LV_{\text{exit}}(6) \setminus \{z\}) \cup \{y\} \\
 LV_{\text{entry}}(7) &= \{z\}
 \end{aligned}$$

$$\begin{aligned}
 LV_{\text{exit}}(1) &= LV_{\text{entry}}(2) \\
 LV_{\text{exit}}(2) &= LV_{\text{entry}}(3) \\
 LV_{\text{exit}}(3) &= LV_{\text{entry}}(4) \\
 LV_{\text{exit}}(4) &= LV_{\text{entry}}(5) \cup \\
 &\quad LV_{\text{entry}}(6) \\
 LV_{\text{exit}}(5) &= LV_{\text{entry}}(7) \\
 LV_{\text{exit}}(6) &= LV_{\text{entry}}(7) \\
 LV_{\text{exit}}(7) &= \emptyset
 \end{aligned}$$

Kleinste Lösung

$[x:=2]^1; [y:=4]^2; [x:=1]^3; (\text{if } [y>x]^4 \text{ then } [z:=y]^5 \text{ else } [z:=y*y]^6); [x:=z]^7$

ℓ	LV_{entry}	LV_{exit}
1	\emptyset	\emptyset
2	\emptyset	$\{y\}$
3	$\{y\}$	$\{x, y\}$
4	$\{x, y\}$	$\{y\}$
5	$\{y\}$	$\{z\}$
6	$\{y\}$	$\{z\}$
7	$\{z\}$	\emptyset

Wir interessieren uns für die kleinste Lösung, da die Gleichungen eine Approximation von oben darstellen.

Typsystem für Reaching Definitions

$$\frac{}{[x:=a]^{\ell'} : D \rightarrow D \setminus \{(x, \ell) \mid \ell \in \mathbf{Lab} \cup \{?\}\} \cup \{(x, \ell')\}} \quad (\text{ASS})$$

$$\frac{}{[\text{skip}]^{\ell} : D \rightarrow D} \quad (\text{SKIP})$$

$$\frac{S_1 : D_1 \rightarrow D_2 \quad S_2 : D_2 \rightarrow D_3}{S_1; S_2 : D_1 \rightarrow D_3} \quad (\text{SEQ})$$

$$\frac{S_1 : D_1 \rightarrow D_2 \quad S_2 : D_1 \rightarrow D_2}{\text{if } [b]^{\ell} \text{ then } S_1 \text{ else } S_2 : D_1 \rightarrow D_2} \quad (\text{IF})$$

$$\frac{S : D \rightarrow D}{\text{while } [b]^{\ell} \text{ do } S : D \rightarrow D} \quad (\text{WHILE})$$

$$\frac{S : D_1 \rightarrow D_2 \quad D'_1 \subseteq D_1 \quad D_2 \subseteq D'_2}{S : D'_1 \rightarrow D'_2} \quad (\text{SUB})$$

Bedeutung des Typsystems

- Mit dem Typsystem lassen sich *Typurteile* der Form $S : D_1 \rightarrow D_2$, wobei D_1, D_2 Mengen von Zuweisungen, also (x, ℓ) und $(x, ?)$ sind.
-
- Ist $RD_{entry/exit}(-)$ Lösung des Constraintsystems, so lässt sich für jedes Programmstück S das Typurteil

$$S : RD_{entry}(init(S)) \rightarrow \bigcup_{\ell \in final(S)} RD_{exit}(\ell)$$

herleiten.

Hier bezeichnet $init(S)$ das Eingangslabel von S und $final(S)$ die Ausgangslabels von S . Wenn etwa S ein if-then-else ist, dann gibt es mehr als ein Ausgangslabel.

Beispiel

$[y:=x]^1; [z:=1]^2; \text{while } [y>1]^3 \text{ do } ([z:=z*y]^4; [y:=y-1]^5); [y:=0]^6$

Es sei $D := \{(x, ?), (y, 1), (y, 5), (z, 2), (z, 4)\}$.

$$\frac{\frac{\frac{[z:=z*y]^4 : D \rightarrow D \setminus \{(z, 2)\}}{[z:=z*y]^4; [y:=y-1]^5 : D \setminus \{(z, 2), (y, 1)\}}}{[z:=z*y]^4; [y:=y-1]^5 : D \rightarrow D}}{\text{while } [y>1]^3 \text{ do } ([z:=z*y]^4; [y:=y-1]^5) : D \rightarrow D}}$$

Operationelle Semantik

Programmzustände (σ) sind Abbildungen von Variablen auf Werte.
Die Urteile

- $\langle S, \sigma \rangle \rightarrow \sigma'$ (Programmstück S liefert bei Startzustand σ den Endzustand σ')
- $\langle S, \sigma \rangle \rightarrow \langle S', \sigma' \rangle$ (Die Auswertung von S im Zustand σ führt zum Zwischenzustand σ' ; es bleibt noch S' abzuarbeiten.)

werden durch folgende Regeln definiert:

Regeln für die operationelle Semantik

$$\frac{}{\langle [x := a]^{\ell}, \sigma \rangle \rightarrow \sigma[x \mapsto \llbracket a \rrbracket \sigma]} \quad (\text{ASS})$$

$$\frac{}{\langle \text{skip}, \sigma \rangle \rightarrow \sigma} \quad (\text{SKIP})$$

$$\frac{\langle S_1, \sigma \rangle \rightarrow \langle \sigma', S'_1 \rangle}{\langle S_1; S_2, \sigma \rangle \rightarrow \langle S'_1; S_2, \sigma' \rangle} \quad (\text{SEQ1})$$

$$\frac{\langle S_1, \sigma \rangle \rightarrow \sigma' \quad \langle S_2, \sigma' \rangle \rightarrow \langle S'_2, \sigma'' \rangle}{\langle S_1; S_2, \sigma \rangle \rightarrow \langle S'_2, \sigma'' \rangle} \quad (\text{SEQ2})$$

$$\frac{\langle S_1, \sigma \rangle \rightarrow \sigma' \quad \langle S_2, \sigma' \rangle \rightarrow \sigma''}{\langle S_1; S_2, \sigma \rangle \rightarrow \sigma''} \quad (\text{SEQ3})$$

Regeln für die operationelle Semantik, Forts.

$$\frac{\llbracket b \rrbracket \sigma = \text{true} \quad \langle S_1, \sigma \rangle \rightarrow X}{\langle \text{if } b \text{ then } S_1 \text{ else } S_2, \sigma \rangle \rightarrow X} \quad (\text{IFT})$$

$$\frac{\llbracket b \rrbracket \sigma = \text{false} \quad \langle S_2, \sigma \rangle \rightarrow X}{\langle \text{if } b \text{ then } S_1 \text{ else } S_2, \sigma \rangle \rightarrow X} \quad (\text{IFD})$$

$$\frac{\llbracket b \rrbracket \sigma = \text{false}}{\langle \text{while } b \text{ do } S, \sigma \rangle \rightarrow \sigma} \quad (\text{WHILEF})$$

$$\frac{\llbracket b \rrbracket \sigma = \text{true}}{\langle \text{while } b \text{ do } S, \sigma \rangle \rightarrow \langle S; \text{while } b \text{ do } S, \sigma \rangle} \quad (\text{WHILET})$$

Instrumentierung der operationellen Semantik

- Für den Korrektheitsbeweis der *RD*-Analyse reichen wir die Zustände um eine Komponente an, die die Menge der aktiven Zuweisungen mitführt.
- Mit $\sigma.D$ bezeichnen wir diese Komponente.
- Der Anfangszustand σ_I setzt alle Variablen auf 0 und es gilt $\sigma_I.D = \{(x, ?) \mid x \text{ Programmvariable}\}$.
- Die Regel [ass] ist wie folgt anzupassen:

$$\frac{\langle [x := a]^\ell, \sigma \rangle \rightarrow \sigma[x \mapsto \llbracket a \rrbracket \sigma][D \mapsto \sigma.D \setminus \{(x, \ell) \mid \ell \text{ Label oder ?}\} \cup \{(x, \ell)\}]}{\text{(ASS)}}$$

Korrektheit des Typsystems für RD

Korrektheit des RD -Typsystems

Es sei $S : D_1 \rightarrow D_2$ und $\sigma.D \subseteq D_1$.

- Wenn $\langle S, \sigma \rangle \rightarrow \sigma'$ dann $\sigma'.D \subseteq D_2$.
- Wenn $\langle S, \sigma \rangle \rightarrow \langle S', \sigma' \rangle$ dann $\sigma'.D \subseteq D_m$ und $S' : D_m \rightarrow D_2$.

Der Beweis erfolgt durch simultane Induktion über die Herleitung von $\langle S, \sigma \rangle \rightarrow X$ (erste Priorität) und über die Herleitung von $S : D_1 \rightarrow D_2$ (zweite Priorität).

Typinferenz

Man kann das Typsystem als Zwischenschritt zum Beweis der Korrektheit der Programmanalyse einsetzen: Jede Lösung der Datenflussgleichungen liefert Typherleitung; jede Typherleitung liefert semantisch korrekte Aussage.

Man kann alternativ auch gleich einen Algorithmus zur Findung einer Typherleitung konstruieren: Hierzu baut man eine "Skelett-Typherleitung", die Mengenvariablen anstelle von konkreten Mengen enthält. Die Nebenbedingungen in den Regeln wie *sub* werden dann zu Constraints deren Lösungen den Typherleitungen entsprechen.

Typsystem für Live-Variablen

Für live Variablen verwenden wir ein ähnliches Format für Typurteile: $S : L_1 \rightarrow L_2$, wobei L_1, L_2 Mengen von Variablen sind. Die Bedeutung kann hier deklarativ gefasst werden und bedarf auch keiner Instrumentierung:

Wir schreiben $\sigma \upharpoonright_L$ für die Einschränkung der Abbildung σ auf L . Also bedeutet $\sigma \upharpoonright_L = \sigma' \upharpoonright_L$, dass σ und σ' auf den Variablen aus L übereinstimmen.

Korrektheitseigenschaft des Live-Variablen Typsystems

Es gelte $S : L_1 \rightarrow L_2$ und $\sigma \upharpoonright_{L_1} = \sigma' \upharpoonright_{L_1}$.

- Falls $\langle S, \sigma \rangle \rightarrow \sigma_1$, so existiert σ'_1 mit $\langle S, \sigma' \rangle \rightarrow \sigma'_1$ und $\sigma_1 \upharpoonright_{L_2} = \sigma'_1 \upharpoonright_{L_2}$.
- Falls $\langle S, \sigma \rangle \rightarrow \langle S', \sigma_1 \rangle$, so existieren L_m und σ'_1 mit $S' : L_m \rightarrow L_2$ und $\langle S, \sigma' \rangle \rightarrow \langle S', \sigma'_1 \rangle$ und $\sigma_1 \upharpoonright_{L_m} = \sigma'_1 \upharpoonright_{L_m}$.

Typsystem für Live Variablen

$$\frac{}{[x:=a]^{\ell'} : L \setminus \{x\} \cup FV(a) \rightarrow L} \quad (\text{ASS})$$

$$\frac{}{[\text{skip}]^{\ell} : L \rightarrow L} \quad (\text{SKIP})$$

$$\frac{S_1 : L_1 \rightarrow L_2 \quad S_2 : L_2 \rightarrow L_3}{S_1; S_2 : L_1 \rightarrow L_3} \quad (\text{SEQ})$$

$$\frac{S_1 : L_1 \rightarrow L_2 \quad S_2 : L_1 \rightarrow L_2}{\text{if } [b]^{\ell} \text{ then } S_1 \text{ else } S_2 : L_1 \rightarrow L_2} \quad (\text{IF})$$

$$\frac{S : L \rightarrow L}{\text{while } [b]^{\ell} \text{ do } S : L \rightarrow L} \quad (\text{WHILE})$$

$$\frac{S : L_1 \rightarrow L_2 \quad L'_1 \subseteq L_1 \quad L_1 \subseteq L'_2}{S : L'_1 \rightarrow L'_2} \quad (\text{SUB})$$

Zusammenfassung Typsysteme

- Herleitbare Typurteile werden durch Typisierungsregeln induktiv definiert.
- Meist gibt es eine Typregel für jedes syntaktische Konstrukt und darüberhinaus Anpassungsregeln, wie *sub*.
- Typurteile können semantisch interpretiert werden. Korrektheit besagt, dass jedes herleitbare Typurteil in diesem Sinne semantisch gültig ist.

Zusammenfassung Typsysteme, Forts.

- Typkorrektheit wird in der Regel durch doppelte Induktion über operationelle Semantik und Typherleitungen bewiesen. Die Möglichkeit über die operationelle Semantik zu induzieren ergibt sich, da Typkorrektheit meist die Form “Wenn $\langle S, \sigma \rangle \rightarrow \dots$, dann ... hat.
- Durch Rückwärtsanwendung der Typregeln (“Skelett-Herleitung”) können Constraints generiert werden und durch deren Lösung Typherleitungen automatisch gefunden werden (Typinferenz).
- Ergebnisse einer Programmanalyse können oft in Typherleitungen übersetzt werden.
- Typsysteme erleichtern die Formulierung von semantischen Korrektheitsaussagen.

Typ- und Effektsysteme

- Analysen für funktionale Sprachen und Sprachen mit Prozeduren und Methoden lassen sich am besten als Typsystem formulieren.
- Diese Typsysteme verallgemeinern das von SML bekannte Typsystem um Effektannotationen (“Typ- und Effektsysteme”)
- Wir betrachten hier Typsysteme für die Erkennung von Speicherzugriffen und Exceptions in (unreinen) funktionalen Sprachen.
- Die Begriffe Polymorphie und polymorphe Rekursion werden dabei erläutert.

Funktionale Sprache

Terme:

$$e ::= c \mid x \mid \text{fn}_{\pi} x \Rightarrow e_0 \mid \text{fun}_{\pi} f \ x \Rightarrow e_0 \mid e_1 \ e_2 \\ \mid \text{if } e_0 \text{ then } e_1 \text{ else } e_2 \mid \text{let } x=e_1 \text{ in } e_2 \mid e_1 \ \text{op} \ e_2$$

π rangiert über "Labels" und dient der Markierung von Programmpunkten.

Beispiele:

- $(\text{fn}_X x \Rightarrow x)(\text{fn}_Y y \Rightarrow y)$.
- $\text{fun}_F f \ x \Rightarrow \text{if } x = 0 \text{ then } 1 \text{ else } x * f(x - 1)$

Einfaches Typsystem

Typen: $\tau ::= \text{int} \mid \text{bool} \mid \tau_1 \rightarrow \tau_2$

Typkontexte Γ : Endliche Abbildungen von Variablen auf Typen.

$$\frac{}{\Gamma \vdash c : \tau^c} \quad (\text{CON})$$

$$\frac{\Gamma(x) = \tau}{\Gamma \vdash x : \tau} \quad (\text{VAR})$$

$$\frac{\Gamma[x \mapsto \tau_1] \vdash e_0 : \tau_2}{\Gamma \vdash \text{fn}_{\tau} x \Rightarrow e_0 : \tau_1 \rightarrow \tau_2} \quad (\text{FN})$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \rightarrow \tau_2 \quad \Gamma \vdash e_2 : \tau_1}{\Gamma \vdash e_1 e_2 : \tau_2} \quad (\text{APP})$$

Beispiel: $\vdash (\text{fn}_X x \Rightarrow x)(\text{fn}_Y y \Rightarrow y) : \tau \rightarrow \tau$ für jedes τ .

Einfaches Typsystem, Forts.

$$\frac{\Gamma[f \mapsto \tau_1 \rightarrow \tau_2][x \mapsto \tau_1] \vdash e_0 : \tau_2}{\Gamma \vdash \text{fun}_\pi f x \Rightarrow e_0 : \tau_1 \rightarrow \tau_2} \quad (\text{FUN})$$

$$\frac{\Gamma \vdash e_0 : \text{bool} \quad \Gamma \vdash e_1 : \tau \quad \Gamma \vdash e_2 : \tau}{\Gamma \vdash \text{if } e_0 \text{ then } e_1 \text{ else } e_2 : \tau} \quad (\text{IF})$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \quad \Gamma[x \mapsto \tau_1] \vdash e_2 : \tau_2}{\Gamma \vdash \text{let } x = e_1 \text{ in } e_2 : \tau_2} \quad (\text{LET})$$

$$\frac{\Gamma \vdash e_1 : \tau_1^{op} \quad \Gamma \vdash e_2 : \tau_2^{op}}{\Gamma \vdash e_1 \text{ op } e_2 : \tau_3^{op}} \quad (\text{OP})$$

Beispiel:

$\text{fun}_F f x \Rightarrow \text{if } x = 0 \text{ then } 1 \text{ else } x * f(x - 1) : \text{int} \rightarrow \text{int}$, falls

$\tau_1^* = \tau_2^* = \tau_3^* = \tau_1^- = \tau_2^- = \tau_3^- = \tau^0 = \tau^1 = \text{int}$.

Typinferenz

Wir erweitern die Typen um Typvariablen:

$$\tau ::= \text{int} \mid \text{bool} \mid \tau_1 \rightarrow \tau_2 \mid \alpha$$

Die Funktion \mathcal{W} nimmt einen erweiterten Typkontext Γ und einen Term e und liefert entweder “fail” zurück, oder ein Paar (θ, τ) bestehend aus einer Einsetzung (Substitution) θ von erweiterten Typen für die Typvariablen in Γ und einem erweiterten Typ τ , sodass gilt:

- Falls $\mathcal{W}(\Gamma, e) = (\theta, \tau)$, so gilt
 - $\Gamma_0 \vdash e : \tau_0$ für alle durch Einsetzung von konkreten Typen für Variablen in $\Gamma[\theta]$ und τ entstehende Γ_0 und τ_0 .
 - Ist $\Gamma_0 \vdash e : \tau$ und entsteht Γ_0 durch Einsetzung aus Γ , so entstehen Γ_0 und τ durch Einsetzung von konkreten Typen für Variablen in $\Gamma[\theta]$ und τ (also wie im ersten Spiegelstrich).
- Falls $\mathcal{W}(\Gamma, e) = \text{“fail”}$, dann gibt es keine Typisierung $\Gamma_0 \vdash e : \tau_0$, sodass Γ_0 durch Einsetzung aus Γ hervorgeht.

Beispiele

- $\mathcal{W}(\emptyset, \text{fn}_X x \Rightarrow x) = (\emptyset, \alpha \rightarrow \alpha)$
- $\mathcal{W}([x \mapsto \text{bool}], x - 1) = \text{"fail"}$
- $\mathcal{W}([x \mapsto \tau], x) = (\emptyset, \tau)$
- $\mathcal{W}([x \mapsto \alpha], x - 1) = ([\alpha \mapsto \text{int}], \text{int})$
- $\mathcal{W}([x \mapsto \alpha, y \mapsto \beta], x y) = ([\alpha \mapsto \beta \rightarrow \gamma], \gamma)$
- $\mathcal{W}([x \mapsto \alpha, y \mapsto \alpha], x y) = \text{"fail"}$

Definition von \mathcal{W}

$$\frac{}{\mathcal{W}(\Gamma, c) = (\emptyset, \tau^c)} \quad (\text{CON})$$

$$\frac{}{\mathcal{W}(\Gamma, x) = (\emptyset, \Gamma(x))} \quad (\text{VAR})$$

$$\frac{\mathcal{W}(\Gamma[x \mapsto \alpha] \vdash_{\mathcal{W}} e_0) = (\theta, \tau) \quad \alpha \text{ frisch}}{\mathcal{W}(\Gamma, \text{fn}_{\pi} x \Rightarrow e_0) = (\theta, \alpha[\theta] \rightarrow \tau)} \quad (\text{FN})$$

$$\frac{\mathcal{W}(\Gamma[f \mapsto \alpha_1 \rightarrow \alpha_2][x \mapsto \alpha_1], e_0) = (\theta, \tau_2) \quad \theta_1 \text{ sodass } \alpha_2[\theta][\theta_1] = \tau_2[\theta_1]}{\mathcal{W}(\Gamma \text{fun } f \ x \Rightarrow e_0) = (\theta; \theta_1, \alpha_1[\theta][\theta_1] \rightarrow \tau_2[\theta_1])} \quad (\text{FUN})$$

In Regel FUN ist θ_1 so allgemein wie möglich zu wählen.

Wir definieren $\mathcal{W}(\Gamma, e) = (\theta, \tau)$, wenn das mit diesen Regeln (und den auf der nächsten Folie) herleitbar ist.

$\mathcal{W}(\Gamma, e) = \text{"fail"}$ wenn keine solche Herleitung existiert

Typinferenz, Forts.

$$\begin{array}{l}
 \mathcal{W}(\Gamma, e_0) = (\theta_0, \tau_0) \\
 \mathcal{W}(\Gamma[\theta_0], e_1) = (\theta_1, \tau_1) \\
 \mathcal{W}(\Gamma[\theta_0][\theta_1], e_2) = (\theta_2, \tau_2) \\
 \hline
 \mathcal{W}(\Gamma, \text{if } e_0 \text{ then } e_1 \text{ else } e_2) = (\theta_0; \theta_1; \theta_2; \theta, \tau)
 \end{array}
 \quad (\text{IF})$$

wobei θ allgemeinstmöglich so zu wählen ist, dass
 $\tau_0[\theta_1][\theta_2]\theta = \text{bool}$ und $\tau_1[\theta_2][\theta] = \tau_2[\theta]$.

$\theta_1; \theta_2$ bezeichnet die Hintereinanderausführung von θ_1 und θ_2 .

Die Regeln für APP, LET und OP verbleiben als Übung.

Allgemeinstmögliche Wahl

Für erweiterte Typen (mit Variablen) τ_1, τ_2 bezeichne $\mathcal{U}(\tau_1, \tau_2)$ die allgemeinstmögliche Einsetzung θ von erweiterten Typen für die Variablen, so dass $\tau_1[\theta] = \tau_2[\theta]$. Wenn es keine solche gibt, dann ist $\mathcal{U}(\tau_1, \tau_2) = \text{"fail"}$.

Beispiele

- $\mathcal{U}(\alpha, \tau) = [\alpha \mapsto \tau]$
- $\mathcal{U}(\text{int}, \text{bool}) = \text{"fail"}$
- $\mathcal{U}(\alpha \rightarrow \beta, (\text{bool} \rightarrow \beta) \rightarrow \text{int} \rightarrow \gamma) =$
 $[\alpha \mapsto \text{bool} \rightarrow \text{int} \rightarrow \gamma, \beta \mapsto \text{int} \rightarrow \gamma]$

$\theta_1; \theta_2$ bezeichnet die Hintereinanderausführung von θ_1 und θ_2 :
 $[\alpha \mapsto \beta \rightarrow \text{bool}]; [\beta \mapsto \text{int}] = [\alpha \mapsto \text{int} \rightarrow \text{bool}, \beta \mapsto \text{int}]$.

Formal kann \mathcal{U} durch Rekursion über die Typsyntax definiert werden. (wird hier nicht behandelt)

Typschemata

Bisher konnten Funktionen nur konkrete Typen haben, wie z.B.
 $\text{fn } x \Rightarrow x : \text{int} \rightarrow \text{int}.$

Wir erweitern jetzt das Typsystem selbst um Typvariablen und Quantifikation, sodass z.B.: $\text{fn } x \Rightarrow x : \forall \alpha. \alpha \rightarrow \alpha.$

SML Notation

In SML *schreibt man* $\alpha \rightarrow \alpha$ oder $'a \rightarrow 'a$, *meint aber* $\forall \alpha. \alpha \rightarrow \alpha.$

Grammatik für Typen (τ) und Typschemata (σ)

$$\tau ::= \text{int} \mid \text{bool} \mid \tau_1 \rightarrow \tau_2 \mid \alpha$$
$$\sigma ::= \forall(\alpha_1, \dots, \alpha_n). \tau$$

Abkürzungen und Beispiele

- Ein Typschema der Form $\forall(\alpha).\tau$ (also $n = 1$) schreiben wir kurz $\forall\alpha.\tau$.
- Ein Typschema der Form $\forall().\tau$ (also $n = 0$) schreiben wir kurz τ . Auf diese Weise werden die Typen zu einer Teilmenge der Typschemata.
- Jeder Typ ist ein Typschema aber nicht umgekehrt.
- Beispiel: $\forall(\alpha, \beta, \gamma).(\alpha \rightarrow \beta \rightarrow \gamma) \rightarrow (\beta \rightarrow \alpha \rightarrow \gamma)$.
- In der Praxis nimmt man auch rekursive Typen, wie z.B. Listen hinzu: $\forall(\alpha, \beta).(\alpha \rightarrow \beta) \rightarrow \text{list}(\alpha) \rightarrow \text{list}(\beta)$.

Unterschied zwischen “let” und Applikation

Man kann schreiben:

$$\text{let } f = \text{fn}_X x \Rightarrow x \text{ in if } f \text{ true then } f \ 0 \text{ else } f \ 1$$

Man kann *nicht* schreiben:

$$(\text{fn}_F f \Rightarrow \text{if } f \text{ true then } f \ 0 \text{ else } f \ 1)(\text{fn}_X x \Rightarrow x)$$

denn Funktionsparameter (hier f) können nicht ein Typschema als “Typ” haben. Let-gebundene Variablen hingegen schon.

System F

Es gibt Typsysteme, die keinen Unterschied zwischen Typen und Typschemata machen und Typen wie $(\forall \alpha. \alpha \rightarrow \alpha) \rightarrow \text{int}$ erlauben. Z.B. System F. Dann wird aber Typinferenz unmöglich (unentscheidbar) und man muss Typannotate zumindest an bestimmten Schlüsselstellen angeben.

Polymorphes Typsystem

Typkontexte binden nunmehr Variablen auf Typschemata ab.

Das Typurteil hat das Format $\Gamma \vdash e : \sigma$.

Die Typregeln CON, VAR, FN, FUN, APP, FUN, OP, IF, APP werden unverändert übernommen mit der gemachten Übereinkunft, dass Typen spezielle Typschemata sind. So ist zum Beispiel die Regel

$$\frac{\Gamma[x \mapsto \tau_1] \vdash e_0 : \tau_2}{\Gamma \vdash \mathbf{fn}_\pi x \Rightarrow e_0 : \tau_1 \rightarrow \tau_2} \quad (\text{FN})$$

nur anwendbar, wenn x in Γ an einen Typen und kein echtes Typschema gebunden ist.

Polymorphes Typsystem, Generalisierungsregel

$$\frac{\Gamma \vdash e : \sigma \quad \alpha_1, \dots, \alpha_n \text{ kommen nicht frei in } \Gamma \text{ vor}}{\Gamma \vdash e : \forall(\alpha_1, \dots, \alpha_n).\sigma} \text{ (GEN)}$$

α kommt frei vor in $\alpha \rightarrow \alpha$ oder in $[x \mapsto \alpha]$, o.ä.

α kommt nicht frei (sondern nur gebunden) vor in $\forall\alpha.\alpha \rightarrow \alpha$ oder $[x \mapsto \forall\alpha.\alpha \rightarrow \alpha]$. Man könnte ja stattdessen auch $\forall\beta.\beta \rightarrow \beta$ oder $[x \mapsto \forall\beta.\beta \rightarrow \beta]$ schreiben.

Beispiel:

$$\frac{\frac{x \mapsto \alpha \vdash x : \alpha}{\emptyset \vdash \text{fn}_X x \Rightarrow x : \alpha \rightarrow \alpha} \text{ FN}}{\emptyset \vdash \text{fn}_X x \Rightarrow x : \forall\alpha.\alpha \rightarrow \alpha} \text{ GEN}$$

Polymorphes Typsystem, Instanzierungsregel

$$\frac{\Gamma \vdash e : \forall(\alpha_1, \dots, \alpha_n).\tau}{\Gamma \vdash e : \tau[\theta]} \quad (\text{INST})$$

Hier ist $\theta = [\alpha_1 \mapsto \tau_1, \dots, \alpha_n \mapsto \tau_n]$ wie vorher eine Einsetzung von Typen für die Variablen $\alpha_1, \dots, \alpha_n$.

Beispiel:

$\frac{}{\Gamma \vdash \forall \alpha. \alpha \rightarrow \alpha \vdash f : \forall \alpha. \alpha \rightarrow \alpha} \text{VAR}$	$\frac{}{\Gamma \vdash \forall \alpha. \alpha \rightarrow \alpha \vdash f : \forall \alpha. \alpha \rightarrow \alpha} \text{VAR}$
$\frac{}{\Gamma \vdash \forall \alpha. \alpha \rightarrow \alpha \vdash f : \text{bool} \rightarrow \text{bool}} \text{INST}$	$\frac{}{\Gamma \vdash \forall \alpha. \alpha \rightarrow \alpha \vdash f : \text{int} \rightarrow \text{int}} \text{INST}$
$\frac{}{\Gamma \vdash \forall \alpha. \alpha \rightarrow \alpha \vdash f \text{ true} : \text{bool}} \text{APP}$	$\frac{}{\Gamma \vdash \forall \alpha. \alpha \rightarrow \alpha \vdash f \ 0 : \text{int}} \text{APP}$
$\frac{}{\Gamma \vdash \forall \alpha. \alpha \rightarrow \alpha \vdash \text{if } f \text{ true then } f \ 0 \text{ else } 1 : \text{int}} \text{IF und CON}$	

Polymorphes Typsystem, Let-regel

$$\frac{\Gamma \vdash e_1 : \sigma_1 \quad \Gamma[x \mapsto \sigma_1] \vdash e_2 : \sigma_2}{\Gamma \vdash \text{let } x = e_1 \text{ in } e_2 : \sigma_2} \quad (\text{LET})$$

Beispiel:

$$\frac{\frac{f \mapsto \forall \alpha. \alpha \rightarrow \alpha \vdash \text{if } f \text{ true then } f \ 0 \text{ else } 1 : \text{int}}{\quad} \quad \frac{}{\emptyset \vdash \text{fn}_X x \Rightarrow x : \forall \alpha. \alpha \rightarrow \alpha}}{\emptyset \vdash \text{let } f = \text{fn}_X x \Rightarrow x \text{ in if } f \text{ true then } f \ 0 \text{ else } 1 : \text{int}}$$

Typinferenz mit Polymorphie

- Die Funktion $\mathcal{W}(\Gamma, e)$ liefert wie vorher einen Typ (kein Typschema) und eine Substitution (für freie Typvariablen in Γ) zurück (oder “fail”!).
- Bei $\mathcal{W}(\Gamma, x)$ muss das Typschema $\Gamma(x)$ mit frischen Variablen instanziiert werden.
- Ist also $\Gamma = [x \mapsto \forall \alpha. \alpha \rightarrow \alpha, y \mapsto \alpha]$, so ist $\mathcal{W}(\Gamma, x) = \beta \rightarrow \beta$.
- Bei $\mathcal{W}(\Gamma, \text{let } x = e_1 \text{ in } e_2)$ muss der rekursiv ermittelte Typ von e_1 *geschlossen* werden.
- Ist also $\mathcal{W}(\Gamma, e_1) = (\theta, \tau)$, so tätigt man den Aufruf $\mathcal{W}(\Gamma[\theta][x \mapsto \forall \vec{\alpha}. \tau], e_2)$, wobei $\vec{\alpha}$ alle Typvariablen umfasst, die frei in τ vorkommen, aber gleichzeitig nicht frei in $\Gamma[\theta]$ vorkommen.

Kontrollflussanalyse

- Wir möchten für jede Funktion wissen, mit welchen `fn`- oder `fun`-Konstrukten sie erzeugt wurde.
- Dazu benutzen wir verfeinerte Funktionstypen der Form $\tau_1 \xrightarrow{\Pi} \tau_2$, wobei Π eine *Menge* von Labels ist.

- Beispiele:

$$\text{fn}_{\chi x \Rightarrow x} : \text{int} \xrightarrow{\{X\}} \text{int}$$

$$\text{if } b \text{ then } \text{fn}_{\chi x \Rightarrow x} \text{ else } \text{fn}_{\gamma y \Rightarrow y + 1} : \text{int} \xrightarrow{\{X, Y\}} \text{int}$$

- (Polymorphie ist hier zunächst wieder “abgeschaltet”).
- Verallgemeinerte Anwendung: Herkunft von Datenstrukturen (hier Funktionen). Z.B.: Strings nur als Ergebnis von Sanitizing Funktionen (Abhilfe gegen XSS Attacken).

Typregeln für Kontrollflussanalyse

Typen: $\tau ::= \text{int} \mid \text{bool} \mid \tau_1 \xrightarrow{\Pi} \tau_2$

$$\frac{\Gamma[x \mapsto \tau_1] \vdash e_0 : \tau_2}{\Gamma \vdash \text{fn}_{\pi} x \Rightarrow e_0 : \tau_1 \xrightarrow{\{\pi\} \cup \Pi} \tau_2} \quad (\text{FN})$$

$$\frac{\Gamma[f \mapsto \tau_1 \xrightarrow{\{\pi\} \cup \Pi} \tau_2][x \mapsto \tau_1] \vdash e_0 : \tau_2}{\Gamma \vdash \text{fun}_{\pi} f x \Rightarrow e_0 : \tau_1 \xrightarrow{\{\pi\} \cup \Pi} \tau_2} \quad (\text{FUN})$$

Die anderen Regeln werden sinngemäß übernommen.

Approximation

Die Kontrollflussanalyse ist approximativ in dem Sinne, dass jede mögliche Funktionsherkunft in den Typen aufgeführt ist, aber nicht jede aufgeführte Herkunft tatsächlich möglich sein muss.

Kontrollflussanalyse: Typinferenz

- Man verwendet *Mengenvariablen* als Annotierungen der Funktionstypen.
- Die Funktionen \mathcal{W} und \mathcal{U} liefern jeweils zusätzlich eine Liste von Constraints über den Mengenvariablen, sodass jede Lösung der Constraints eine Typisierung liefert und umgekehrt.

Beispiel: Typinferenz angewandt auf

$$(\text{fn}_X \ x \Rightarrow x)(\text{fn}_Y \ y \Rightarrow y)$$

liefert den verallgemeinerten Typ $\alpha \xrightarrow{\zeta} \alpha$ und die Constraintmenge $\{Y\} \subseteq \zeta$.

Operationelle Semantik

Wir definieren ein Auswerturteil $e \longrightarrow v$, gelesen e wertet zu v aus, wobei e ein geschlossener Ausdruck ist und v ein *Wert* ist, also eine Konstante, oder ein fn-Ausdruck.

$$\frac{}{c \longrightarrow c} \quad (\text{CON})$$

$$\frac{}{\text{fn}_\pi x \Rightarrow e \longrightarrow \text{fn}_\pi x \Rightarrow e} \quad (\text{FN})$$

$$\frac{}{\text{fun}_\pi f x \Rightarrow e \longrightarrow \text{fn}_\pi x \Rightarrow e[f \mapsto \text{fun}_\pi f x \Rightarrow e]} \quad (\text{FUN})$$

$$\frac{e_1 \longrightarrow \text{fn}_\pi x \Rightarrow e'_1 \quad e_2 \longrightarrow v_2 \quad e'_1[x \mapsto v_2] \longrightarrow v}{e_1 e_2 \longrightarrow v} \quad (\text{APP})$$

$$\frac{e_1 \longrightarrow \text{true} \quad e_2 \longrightarrow v}{\text{if } e_1 \text{ then } e_2 \text{ else } e_3 \longrightarrow v} \quad (\text{IF1})$$

Operationelle Semantik, Forts.

$$\frac{e_1 \longrightarrow \text{false} \quad e_3 \longrightarrow v}{\text{if } e_1 \text{ then } e_2 \text{ else } e_3 \longrightarrow v} \quad (\text{IF2})$$

$$\frac{e_1 \longrightarrow v_1 \quad e_2[x \mapsto v] \longrightarrow v}{\text{let } x=e_1 \text{ in } e_2 \longrightarrow v} \quad (\text{LET})$$

$$\frac{e_1 \longrightarrow v_1 \quad e_2 \longrightarrow v_2 \quad v = v_1 \text{ op } v_2}{e_1 \text{ op } e_2 \longrightarrow v} \quad (\text{OP})$$

- $e[x \mapsto v]$ bezeichnet die Einsetzung (Substitution) von v für x in e .
- $v_1 \text{ op } v_2$ ist das entsprechende Ergebnis des Operators op .
Z.B. $3+4 = 7$. Wir verlangen, dass es den Typ τ_3^{op} besitzt.

Semantische Typkorrektheit

Korrektheit des einfachen Typsystems

Wenn $\emptyset \vdash e : \tau$ und $e \longrightarrow v$, dann auch $\emptyset \vdash v : \tau$. Außerdem bleibt die Auswertung von e nicht stecken, in dem Sinne, dass keine Auswertungsregel anwendbar wäre. Z.B. bleibt $0\ 1$ (0 angewendet auf 1) stecken.

Man kann das “Steckenbleiben” sauber definieren, indem man einen Sonderwert *wrong* hinzugibt und Auswerteregeln wie etwa $e_1\ e_2 \longrightarrow \text{wrong}$, falls $e \longrightarrow v$ und v kein fn -Ausdruck ist.

Korrektheit der Kontrollflussanalyse

Wenn $\emptyset \vdash e : \tau$ und $e \longrightarrow v$, dann auch $\emptyset \vdash v : \tau$. Insbesondere gilt:

Wenn $\emptyset \vdash e : \tau_1 \xrightarrow{\Pi} \tau_2$ und $e \longrightarrow \text{fn}_{\pi} x \Rightarrow e_1$, dann $\pi \in \Pi$.

Man beweist die Korrektheit jeweils durch Induktion über $e \longrightarrow v$.

Erweiterung der Sprache um Referenzen

$$e ::= \dots \mid \text{ref}_\pi e \mid !e \mid x=e$$

Abkürzung:

$e_1; e_2$ steht für $\text{let } _ = e_1 \text{ in } e_2$ wobei $_$ eine frische Variable ist.

Programm für Fibonaccizahlen:

```
fnF m=>let o = refO 0 in let n = refN 1 in
  (funG f m=>if m ≤ 1 then 44 else
    f(m - 1); let x=!o in o:=!n;n:=!n + x) m;
  if m ≤ 1 then m else !o
```

Seiteneffektanalyse

Effektannotationen

Eine *Effektannotation* (kurz *Effekt*) ist eine Menge von *elementaren Effektannotationen*; diese sind:

- $!\pi$ (Lesen einer bei π erzeugten Referenz)
- $\pi :=$ (Zuweisen zu einer bei π erzeugten Referenz)
- $\text{ref } \pi$ (Erzeugen einer Referenz bei π)

Zum Beispiel ist $\{O :=, !O, !N\}$ ein Effekt.

Verfeinerte Typen:

$$\tau ::= \text{int} \mid \text{bool} \mid \tau_1 \xrightarrow{\varpi} \tau_2 \mid \text{ref } \Pi \tau$$

Hier bezeichnet ϖ eine Effektannotation und Π wie bisher eine Menge von Labels.

Seiteneffektanalyse als Typsystem

Das Typurteil hat nunmehr die Form $\Gamma \vdash e : \tau \ \& \ \varpi$.

$$\frac{}{\Gamma \vdash c : \tau^c \ \& \ \emptyset} \quad (\text{CON})$$

$$\frac{\Gamma(x) = \tau}{\Gamma \vdash x : \tau \ \& \ \emptyset} \quad (\text{VAR})$$

$$\frac{\Gamma[x \mapsto \tau_1] \vdash e_0 : \tau_2 \ \& \ \varpi}{\Gamma \vdash \text{fn}_\pi x \Rightarrow e_0 : \tau_1 \xrightarrow{\varpi} \tau_2 \ \& \ \emptyset} \quad (\text{FN})$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \xrightarrow{\varpi_1} \tau_2 \ \& \ \varpi_2 \quad \Gamma \vdash e_2 : \tau_1 \ \& \ \varpi_3}{\Gamma \vdash e_1 \ e_2 : \tau_2 \ \& \ \varpi_1 \cup \varpi_2 \cup \varpi_3} \quad (\text{APP})$$

Fortsetzung

$$\frac{\Gamma[f \mapsto \tau_1 \xrightarrow{\varpi} \tau_2][x \mapsto \tau_1] \vdash e_0 : \tau_2 \ \& \ \varpi}{\Gamma \vdash \text{fun}_{\pi} f \ x \Rightarrow e_0 : \tau_1 \xrightarrow{\varpi} \tau_2 \ \& \ \varpi} \quad (\text{FUN})$$

$$\frac{\Gamma \vdash e_0 : \text{bool} \ \& \ \varpi \quad \Gamma \vdash e_1 : \tau \ \& \ \varpi \quad \Gamma \vdash e_2 : \tau \ \& \ \varpi}{\Gamma \vdash \text{if } e_0 \text{ then } e_1 \text{ else } e_2 : \tau \ \& \ \varpi} \quad (\text{IF})$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \ \& \ \varpi \quad \Gamma[x \mapsto \tau_1] \vdash e_2 : \tau_2 \ \& \ \varpi}{\Gamma \vdash \text{let } x = e_1 \text{ in } e_2 : \tau_2 \ \& \ \varpi_1 \cup \varpi} \quad (\text{LET})$$

$$\frac{\Gamma \vdash e_1 : \tau_1^{op} \ \& \ \varpi \quad \Gamma \vdash e_2 : \tau_2^{op} \ \& \ \varpi}{\Gamma \vdash e_1 \text{ op } e_2 : \tau_3^{op} \ \& \ \varpi} \quad (\text{OP})$$

Fortsetzung

$$\frac{\Gamma \vdash e : \text{ref}_{\Pi} \tau \ \& \ \varpi}{\Gamma \vdash !e : \tau \ \& \ \varpi \cup !\Pi} \quad (\text{DEREF})$$

$$\frac{\Gamma \vdash e_1 : \text{ref}_{\Pi} \tau \ \& \ \varpi \quad \Gamma \vdash e_2 : \tau \ \& \ \varpi}{\Gamma \vdash e_1 := e_2 : \tau \ \& \ \varpi \cup \Pi :=} \quad (\text{ASS})$$

$$\frac{\Gamma \vdash e : \tau \ \& \ \varpi}{\Gamma \vdash \text{ref}_{\pi} e : \text{ref}_{\{\pi\}} \tau \ \& \ \varpi \cup \{\text{ref } \pi\}} \quad (\text{REF})$$

$$\frac{\Gamma \vdash e : \tau \ \& \ \varpi \quad \tau \leq \tau' \quad \varpi \subseteq \varpi'}{\Gamma \vdash e : \tau' \ \& \ \varpi'} \quad (\text{SUB})$$

Hier bedeuten $!\Pi = \{!\pi \mid \pi \in \Pi\}$ und $\Pi := = \{\pi := \mid \pi \in \Pi\}$.

Subtyping

- Die Regel SUB erlaubt die Anpassung von Typen und Effekten, sodass z.B. beide Zweige einer Fallunterscheidung denselben Typ und Effekt haben.
- In der Kontrollflussanalyse wurde diese Anpassung in die Regel VAR eingearbeitet. Das würde hier zu aufwendig.
- Das Subtyping-Urteil $\tau_1 \leq \tau_2$ ist definiert durch:

$$\frac{\Pi \subseteq \Pi'}{\text{ref}_{\Pi} \leq \text{ref}_{\Pi'}} \quad ; \quad \frac{\tau_2 \leq \tau_1 \quad \tau'_1 \leq \tau'_2 \quad \varpi \subseteq \varpi'}{\tau_1 \xrightarrow{\varpi} \tau'_1 \leq \tau_2 \xrightarrow{\varpi'} \tau'_2}$$

Beispiele

```
let x = refA 1 in
  (fn f => f (fn y => !x) + f (fn z => (x:=z; z)))
  (fn g => g 1)
```

Die beiden Argumente für f können wie folgt analysiert werden:

$$\text{int} \xrightarrow{\{!A\}} \text{int} \ \& \ \emptyset$$

$$\text{int} \xrightarrow{\{A:=\}} \text{int} \ \& \ \emptyset$$

Daher gibt man f den Typ

$$(\text{int} \xrightarrow{\{!A, A:=\}} \text{int} \xrightarrow{\{!A, A:=\}} \text{int})$$

Die Typen der Argumente werden dann durch Subtyping angepasst.

Exceptions

Wir erweitern die funktionale Sprache um Exceptions (die Seiteneffekte lassen wir wieder weg).

$$e ::= \dots \mid \text{raise } s \mid \text{handle } s \text{ as } e_1 \text{ in } e_2$$

Hier rangiert s über eine festgewählte Menge von Exceptions.

Beispiel:

```
let comb = fun f x => fn y =>
  if x < 0 then raise y-out-of-range
  elseif y < 0 or y > x then raise y-out-of-range
  else if y = 0 or y = x then 1
       else f (x - 1) y + f (x - 1) (y - 1)
in handle x-out-of-range as 0 in comb x y
```

Operationelle Semantik

Die bisherigen Auswerteregeln bleiben in Kraft, gelten aber nicht für den Fall, dass ein Teilausdruck zu `raise s` für ein `s` auswertet. Hinzu kommen:

$$\frac{e_1 \longrightarrow \text{raise } s}{e_1 \ e_2 \longrightarrow \text{raise } s}$$

$$\frac{e_1 \longrightarrow \text{fn } x \Rightarrow e'_1 \quad e_2 \longrightarrow \text{raise } s}{e_1 \ e_2 \longrightarrow \text{raise } s}$$

$$\frac{e_1 \longrightarrow \text{fn } x \Rightarrow e'_1 \quad e_2 \longrightarrow v_2 \quad e'_1[x \mapsto v_2] \longrightarrow \text{raise } s}{e_1 \ e_2 \longrightarrow \text{raise } s}$$

Analoge Regeln für `fun`, `let`, `op`.

Operationelle Semantik, Forts.

$$\frac{}{\text{raise } s \longrightarrow \text{raise } s}$$

$$\frac{e_2 \longrightarrow v_2 \quad v_2 \neq \text{raise } s}{\text{handle } s \text{ as } e_1 \text{ in } e_2 \longrightarrow v_2}$$

$$\frac{e_2 \longrightarrow \text{raise } s \quad e_1 \longrightarrow v_1}{\text{handle } s \text{ as } e_1 \text{ in } e_2 \longrightarrow v_1}$$

Annotierte Typen mit Effektpolymorphie

Effektausdrücke:

$$\varphi ::= \{s\} \mid \varphi_1 \cup \varphi_2 \mid \epsilon$$

Hierbei bezeichnet s den Effekt des möglichen Auslösens der Exception s und ϵ rangiert über *Effektvariablen*.

Annotierte Typen:

$$\tau ::= \text{int} \mid \text{bool} \mid \tau_1 \xrightarrow{\varphi} \tau_2 \mid \alpha$$

α rangiert wie bisher über Typvariablen.

Typschemata:

$$\sigma ::= \forall(\zeta_1, \dots, \zeta_n). \tau$$

Die ζ_i bezeichnen Typ- oder Effektvariablen.

Typregeln

Das Typurteil hat nunmehr die Form $\Gamma \vdash e : \sigma \ \& \ \varphi$, wobei Γ Variablen an *Typschemata* bindet.

$$\frac{}{\Gamma \vdash c : \tau^c \ \& \ \emptyset} \quad (\text{CON})$$

$$\frac{\Gamma(x) = \tau}{\Gamma \vdash x : \tau \ \& \ \emptyset} \quad (\text{VAR})$$

$$\frac{\Gamma[x \mapsto \tau_1] \vdash e_0 : \tau_2 \ \& \ \varphi}{\Gamma \vdash \text{fn } x \Rightarrow e_0 : \tau_1 \xrightarrow{\varphi} \tau_2 \ \& \ \emptyset} \quad (\text{FN})$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \xrightarrow{\varphi_1} \tau_2 \ \& \ \varphi_2 \quad \Gamma \vdash e_2 : \tau_1 \ \& \ \varphi_3}{\Gamma \vdash e_1 \ e_2 : \tau_2 \ \& \ \varphi_1 \cup \varphi_2 \cup \varphi_3} \quad (\text{APP})$$

Fortsetzung

$$\frac{\Gamma[f \mapsto_{\tau_1} \overset{\varphi}{\rightarrow} \tau_2][x \mapsto_{\tau_1}] \vdash e_0 : \tau_2 \ \& \ \varphi}{\Gamma \vdash \text{fun } f \ x \Rightarrow e_0 : \tau_1 \ \overset{\varphi}{\rightarrow} \ \tau_2 \ \& \ \varphi} \quad (\text{FUN})$$

$$\frac{\Gamma \vdash e_0 : \text{bool} \ \& \ \varphi \quad \Gamma \vdash e_1 : \tau \ \& \ \varphi \quad \Gamma \vdash e_2 : \tau \ \& \ \varphi}{\Gamma \vdash \text{if } e_0 \text{ then } e_1 \text{ else } e_2 : \tau \ \& \ \varphi} \quad (\text{IF})$$

$$\frac{\Gamma \vdash e_1 : \sigma_1 \ \& \ \varphi_1 \quad \Gamma[x \mapsto \sigma_1] \vdash e_2 : \tau_2 \ \& \ \varphi_2}{\Gamma \vdash \text{let } x = e_1 \text{ in } e_2 : \tau_2 \ \& \ \varphi_1 \cup \varphi_2} \quad (\text{LET})$$

$$\frac{\Gamma \vdash e_1 : \tau_1^{op} \ \& \ \varphi \quad \Gamma \vdash e_2 : \tau_2^{op} \ \& \ \varphi}{\Gamma \vdash e_1 \ op \ e_2 : \tau_3^{op} \ \& \ \varphi} \quad (\text{OP})$$

Fortsetzung

$$\frac{\Gamma \vdash e : \sigma \ \& \ \varphi \quad \zeta_1, \dots, \zeta_n \text{ kommen nicht frei in } \Gamma \text{ und } \varphi \text{ vor}}{\Gamma \vdash e : \forall(\zeta_1, \dots, \zeta_n). \tau \ \& \ \varphi} \quad (\text{GEN})$$

$$\frac{\Gamma \vdash e : \forall(\zeta_1, \dots, \zeta_n). \tau \ \& \ \varphi}{\Gamma \vdash e : \tau[\theta] \ \& \ \varphi} \quad (\text{INST})$$

Hier ist θ eine Einsetzung von Typen für Typvariablen und Effektausdrücken für Effektvariablen.

Fortsetzung

$$\frac{}{\Gamma \vdash \text{raise } s : \forall \alpha. \alpha \ \& \ \{s\}} \quad (\text{RAISE})$$

$$\frac{\Gamma \vdash e_1 : \tau \ \& \ \varphi \quad \Gamma \vdash e_2 : \tau \ \& \ \varphi \cup \{s\}}{\Gamma \vdash \text{handle } s \text{ as } e_1 \text{ in } e_2 : \tau \ \& \ \varphi} \quad (\text{HANDLE})$$

$$\frac{\Gamma \vdash e : \tau \ \& \ \varphi \quad \tau \leq \tau' \quad \varphi \leq \varphi'}{\Gamma \vdash e : \tau' \ \& \ \varphi'} \quad (\text{SUB})$$

Das *Subeffecting*-Urteil $\varphi_1 \leq \varphi_2$ versteht sich als punktweise Teilmengenbeziehung.

Subtyping ist wie bisher (Folie 6) definiert.

Beispiel

```
let f = fn g => fn x => g x
in f (fn y => if y < 0 then raise neg else y) 1
  +f (fn z => if z > 0 then raise pos else 0 - z) 1
```

Dieses Programm wertet zu 2 aus.

Der Funktion f kann das folgende Typschema zugewiesen werden:

$$\forall \alpha, \beta, \epsilon. (\alpha \xrightarrow{\epsilon} \beta) \xrightarrow{\emptyset} (\alpha \xrightarrow{\epsilon} \beta)$$

Die Argumentfunktionen für f erhalten die Typen $(\text{int} \xrightarrow{\{\text{neg}\}} \text{int})$
und $(\text{int} \xrightarrow{\{\text{pos}\}} \text{int})$.

Typinferenz

Um Typen im Exception-Typsystem automatisch zu inferieren geht man zunächst so vor wie bei der polymorphen Typinferenz für normale Typen.

Die \mathcal{U} -Funktion erzeugt dann ein System von Constraints, welche auch universell quantifizierte Effektvariablen enthalten. Diese Constraints lassen sich mit verschiedenen Methoden, wie Iteration oder auch SAT-Solver lösen. Die Details werden hier nicht präsentiert.

Polymorphe Rekursion

Erlaubt man Typschemata (statt nur einfacher Typen) für die mit `fun` gebundene Funktion so spricht man von *polymorpher Rekursion*. Für Typschemata mit Typvariablen ist dann die Typinferenz unentscheidbar. Für Typschemata, die nur Effektvariablen quantifizieren lässt sich die Inferenz aber durchführen.

Das ist sinnvoll, wenn zur Rechtfertigung einer Instanz des Typschemas eine andere Instanz im rekursiven Aufruf benötigt wird.

Weiterhin kann man auch über Constraints quantifizieren, also z.B. $\forall \epsilon \geq \{s\} \dots$ und außerdem Kontrollflussinformation zur Findung passender Instanzierungen bei `INST` heranziehen.

Zusammenfassung Typsysteme für funktionale Programme

- Typ- und Effektsysteme verallgemeinern das SML Typsystem durch Effektannotationen für Ausdrücke und Funktionen.
- Sie erlauben die strukturierte Analyse von Programmen mit Funktions- (Prozedur-, Methoden-)Variablen.
- Zur Lösung von Constraintsystemen, die bei der Typinferenz anfallen, werden verschiedene Lösungsalgorithmen herangezogen.
- Man unterscheidet einfache und polymorphe Typsysteme, sowie Typsysteme mit polymorpher Rekursion.