

Einführung in die funktionale Programmierung

Prof. Dr. Manfred Schmidt-Schauß

Künstliche Intelligenz und Softwaretechnologie

23. Oktober 2007

Haskell, eine nicht-strikte funktionale Programmiersprache mit polymorphem Typsystem.

Ziel dieses Kapitels ist:

- Darstellung der syntaktischen Grundlagen,
- Unterscheidung zwischen Kernbestandteilen der Programmiersprache und zusätzlichen Features,
- Programmierverständnis in Haskell,
- operationale Semantik: Auswertungen

Die Kernsprache folgt der Lambda-Schreibweise im Lambda-Kalkül, der von Alonzo Church entwickelt wurde.

Der Lambda-Kalkül ist Teil der betrachteten Kernsprache und auch von Haskell.

Einführung der Syntax und Reduktionsregeln

Es gibt *Konstantensymbole*, die jeweils eine feste Stelligkeit haben.
Diese nennen wir auch *Konstruktoren*.

Konstruktoren mit Stelligkeit sind anzugeben
(wir erfinden keine KFPT-Syntax dafür)

Es gibt *Typen*

Jeder Konstruktor gehört zu genau einem Typ.

z.B. seien die Typen `Bool` und `List`; zu `Bool` gehören `True` und `False`, zu `List` gehören `Cons` und `Nil`.

Es gibt pro Typ ein case-Konstrukt: $\text{case}_{\text{Typname}}$.

Kontextfreie Grammatik für KFPT

$Exp ::= V$ (Variable)
| $(\backslash V \rightarrow Exp)$ wobei V eine Variable ist.
| $(Exp_1 Exp_2)$
| $(c Exp_1 \dots Exp_n)$ wobei c Konstruktor und $n = ar(c)$
| $(case_{Typ} Exp \text{ of } \{Pat_1 \rightarrow Exp_1; \dots; Pat_n \rightarrow Exp_n\})$

Hierbei ist Pat_i Pattern zum Konstruktor i ,
Es kommen genau die Konstrukteure zu Typname vor
 $Pat_i \rightarrow Exp_i$ heißt auch case-Alternative.

$Pat ::= (c V_1 \dots V_{ar(c)})$
Die Variablen V_i müssen alle verschieden sein.

$(c \ V_1 \ \dots \ V_{ar(c)})$ sind einfache *Pattern* bzw. *Muster*.

$\lambda x.s$ im Lambda-Ausdrücke entspricht $\backslash x \rightarrow s$ in KFPT

Zum Beispiel:

$\lambda x.1$ ist die Funktion, die konstant 1 liefert.

Die Funktion, die leere Listen erkennt:

```
\xs -> (case_List xs of {Nil -> True; Cons y ys -> False})
```

In Haskell: kein Typ-Index am case,
sondern automatische Typerkennung

- $\lambda x \rightarrow s$ sind *Abstraktionen* oder Funktionsausdrücke.
- $(s\ t)$ sind *Anwendungen (Applikationen)*.
- $(c\ t_1 \dots t_{ar(c)})$ sind *Konstruktor-Anwendungen*.
- $(\text{case}_T\ e\ \text{of}\ \text{Alts})$ sind *case-Ausdrücke (Fallunterscheidungen)*.

$\lambda x \rightarrow x$ ist die Identitätsfunktion.

if A then B else C

ist darstellbar durch:

(case_{Bool} A of {True \rightarrow B; False \rightarrow C})

Sei Paar zweistelliger Konstruktor.

(Paar True 1) ist ein Objekt zu Paar.

Die Selektoren zu Paar sind definierbar:

```
\x -> (case_Paar x of (Paar y1 y2) -> y1)
```

```
\x -> (case_Paar x of (Paar y1 y2) -> y2)
```

Tupel sind vordefiniert in Haskell: Schreibweise ist (x, y, z) .

Hat das Paar p als erstes Element ein `False`?

wird implementiert durch:

```
(case_Paar p of {Paar x y ->  
                (case_Bool x of {True -> False; False -> True})})
```

Konstruktoren: `Cons` mit Stelligkeit 2
`Nil` mit Stelligkeit 0

Die Liste `[True, False]` ist darstellbar in KFPT als

`Cons True (Cons False Nil)`

Id für die identische Abbildung

K für den konstanten Kombinator:

$$\begin{aligned} Id & ::= \lambda x \rightarrow x && \text{(Lambda-Notation: } \lambda x.x \text{)} \\ K \ x \ y & ::= \lambda x \rightarrow (\lambda y \rightarrow x) && \text{(Lambda-Notation: } \lambda x.(\lambda y.x) \text{)} \end{aligned}$$

$$bot := (\lambda x.(x\ x)) (\lambda y.(y\ y))$$

wir werden später sehen, dass die Auswertung dieses Ausdrucks nicht terminiert.

Die Bindungsregeln in KFPT sind:

- In $\lambda x . e$ ist x eine gebundene Variable, der Gültigkeitsbereich (Skopus) ist e .
- In der case-Alternative $c x_1 \dots x_n \rightarrow e$ bindet das Pattern alle Variablen x_i .
Deren Gültigkeitsbereich (Skopus) ist e .

Vorkommen von Variablen, die in keinem Gültigkeitsbereich einer Bindung stehen, sind *frei*.

Bei Konflikten gilt die Regel: der innerste Bindungsbereich zählt

Beispiel $(\lambda x . x (\lambda x . (y x)))$

verschiedene Vorkommen können zu verschiedenen Bindungen gehören:

Beispiel

$$\lambda x . (x (\lambda x . x))$$

Sichtbar durch Indizes bzw. Umbenennung $\lambda x_1 . (x_1 (\lambda x_2 . x_2))$

Ausdrücke die bis auf Umbenennung gleich sind,
nennt man auch **α -äquivalent**

Genauer: durch ein Folge von Umbenennungen
ineinander überführt werden können,

Ein Ausdruck ohne freie Variablen heißt *geschlossener Ausdruck*

sonst *offener Ausdruck*.

Ein *KFPT-Programm* ist definiert als ein geschlossener KFPT-Ausdruck

```
\x -> (case_List x of {Cons y ys -> ys; Nil -> Nil})
```

ist ein geschlossener Ausdruck.

Die Variable x ist frei im Unterausdruck

```
case_List x of {Cons y ys -> ys; Nil -> Nil}
```

operationale Semantik:

- Transformationssystem auf den Ausdrücken
- deterministisch
- Bei erfolgreichem Terminieren: Der letzte Ausdruck ist der Wert

Ein *Wert* in KFPT

bzw. eine *WHNF* (weak head normal form, schwache Kopfnormalform)

ist ein Ausdruck der Form

- $(c\ t_1 \dots t_n)$, wobei $n = \text{arity}(c)$ und c ein Konstruktor ist, oder
- eine Abstraktion: $\lambda x . e$.

Eine WHNF kann sein:

- *FWHNF*, wenn die WHNF eine Abstraktion ist, oder
- *CWHNF*, wenn die WHNF eine Konstruktoranwendung der Form $(c\ s_1 \dots s_n)$ ist.

$$s[t/x]$$

Im Ausdruck s werden die freien Vorkommen der Variablen x durch t ersetzt.

Zu Beachten: kein Einfangen von freien Variablen !

Ausreichende Bedingung: In s gebundene Variablen sind von den freien Variablen in t verschieden
Wenn nicht, dann gebundene Variablen in s umbenennen.

- Abstraktion :

$\lambda x.e: \rightarrow \lambda z.e[z/x]:$

z sei bisher nicht verwendete Variable.

In e ersetze alle freien Vorkommen von x durch z .

- case-Alternative:

$c\ x_1 \dots x_n \rightarrow e \quad \rightarrow \quad c\ z_1 \dots z_n \rightarrow e[z_1/x_1, \dots, z_n/x_n]$

ersetze x_i durch neue Variablen z_i , $i = 1, \dots, n$

In e ersetze alle freien Vorkommen von x_i durch z_i :

Vorsicht: Auch beim Umbenennen muss man das Einfangen von Variablen vermeiden!

Was ist die richtige Ersetzung in $(\lambda x.\lambda y.(x\ z))[y/z]$?

falsch: $(\lambda x.\lambda y.(x\ y)).$

richtig: $(\lambda x.\lambda u.(x\ y)).$

mit Umbenennung:

$(\lambda x.\lambda y.(x\ z))[y/z] \rightarrow (\lambda x.\lambda u.(x\ z))[y/z] \rightarrow (\lambda x.\lambda u.(x\ y))$

Auswertungsregeln (Reduktionsregeln):

$$\text{Beta} \quad \frac{((\lambda x.t) s)}{t[s/x]}$$

$$\text{Case} \quad \frac{(\text{case}_T (c t_1 \dots t_n) \text{ of } \{\dots; c x_1 \dots x_n \rightarrow s; \dots\})}{s[t_1/x_1, \dots, t_n/x_n]}$$

Zunächst darf man diese überall verwenden:
in allen Programmkontexten

Redex (reducible expression) nennt man
den Unterausdruck s , der **unmittelbar reduziert** wird.

$(\lambda x.(\lambda y.x (y x))) (\lambda z.z)$

→

$(\lambda y.(\lambda z.z) (y (\lambda z'.z')))$

$(\text{case}_{\text{List}} (\text{Cons } 1 (\text{Cons } 2 \text{ Nil})) \{\text{Cons } y \text{ ys } \rightarrow \text{ys}; \text{Nil } \rightarrow \text{Nil}\})$

→

$(\text{Cons } 2 \text{ Nil})$

Die Anwendung der Auswertungsregeln ist noch nicht eindeutig:

$(\lambda x .x)((\lambda y .y)z)$ hat zwei Reduktionsmöglichkeiten.

Ziel: die richtige deterministischen Auswertung

Definition

Reduktionskontexte:

$$R ::= [] \mid (R \ e) \mid (\text{case}_T \ R \ \text{of} \ \{p_1 \rightarrow t_1; \dots; p_n \rightarrow t_n\})$$

Beobachtung: das Loch eines Reduktionskontextes ist:

- nicht in Abstraktionen
- nicht in einer Konstruktor-Anwendung
- nicht im Argument einer Anwendung.
- nicht in case-Alternativen.

Normalordnungs-Reduktion (*normal-order-Reduktion*):

Der Ausdruck in einem Reduktionskontext wird reduziert:

$R[s] \rightarrow R[t]$ *Ein-Schritt-Normalordnungsreduktion*

R ein Reduktionskontext,

s keine WHNF

und s reduziert unmittelbar zu t

Der Unterterm s zusammen mit seiner Position
ist der *Normalordnungsredex*

Bezeichnungen

\xrightarrow{n} eine Normalordnungsreduktion.

$\xrightarrow{n,+}$ bzw. $\xrightarrow{n,*}$ die transitive und reflexiv-transitive Hülle von \xrightarrow{n}

$\xrightarrow{n,*}$ ist die **Normalordnungsrelation**
oder **Auswertung** eines Terms.

Sei t geschlossener Term.

Wenn ein t' existiert mit $t \xrightarrow{n,*} t'$ und t' ist WHNF,

dann: t **terminiert** (auch: **konvergiert**) Bezeichnung: $t\Downarrow$.

Wenn nicht, dann: t **divergiert** Bezeichnung: $t\Uparrow$.

Wenn $t \Downarrow$, sagen wir auch: *hat eine WHNF*.

Es gilt: Wenn $t \xrightarrow{n,*} t'$ und t' ist WHNF,
dann ist t' **eindeutig**

Aber: $t \xrightarrow{*,\text{keine n.o.}} t''$ und t'' ist WHNF,
kann für viele verschiedene t'' gelten.

Beispiel `(cons (($\lambda x.x$) True)) Nil` ist WHNF,
aber reduziert zu `(cons True Nil)`
(das ist **keine** Normalordnung.)

$(\lambda x .x)((\lambda y .y)z)$ hat zwei Reduktionsmöglichkeiten,

aber nur eine Normalordnungsreduktion:

$$(\lambda x .x)((\lambda y .y)z) \xrightarrow{n} ((\lambda y .y)z).$$

Der Reduktionskontext R ist: $([] ((\lambda y .y)z))$

$$((\lambda x.x) \text{Nil}) \xrightarrow{n} \text{Nil}$$

$$((\lambda x.\lambda y.x) s t) \xrightarrow{n} ((\lambda y'.s) t) \xrightarrow{n} s.$$

Umbenennung $[y'/y]$ nach dem ersten Reduktionsschritt;
 y' kommt nicht in s vor

Lemma

- Jede unmittelbare Reduktion in einem Reduktionskontext ist eine Normalordnungsreduktion.
- Der Normalordnungsredex und die Normalordnungsreduktion sind eindeutig.
- Eine WHNF hat keinen Normalordnungsredex und erlaubt keine Normalordnungsreduktion.

Gordon Plotkin:

Eine Programmiersprache hat vier wichtige Komponenten:

- Ausdrücke
- Kontexte
- Werte
- eine Auswertungsrelation auf Ausdrücken.

Das definiert dann u.a. $t \Downarrow$.

Die Theorie der korrekten Transformationen kann man darauf aufbauen:

s, t sind **gleich**, wenn für alle Kontexte C : $C[s] \Downarrow \Leftrightarrow C[t] \Downarrow$

(siehe zweite Semesterhälfte)

- Für case:

$(\text{case}_T e \text{ of } \{alt_1; \dots; alt_m\})$ ist *direkt dynamisch ungetypt*

wenn e eine der folgenden Formen hat:

- $(c a_1 \dots a_{arc})$ und c ist ein Konstruktor, der nicht zum Typ T gehört,
- $\lambda x . t.$

- für Constructoren:

$((c a_1 \dots a_{arc}) e)$ ist *direkt dynamisch ungetypt*.

e ist *dynamisch ungetypt*,
wenn $\exists e' : e \xrightarrow{n,*} e'$ und e' ist direkt dynamisch ungetypt

e ist *dynamisch wohlgetypt* wenn e nicht dynamisch ungetypt ist.

Haskells Typisierung ist restriktiver als
“es tritt kein dynamischer Typfehler auf”

Algorithmus zur Bestimmung des Normalordnungsredex
geschrieben als eine Menge von Regeln,
die ein Label (R) verschieben:

Start mit t^R ,

$$\begin{aligned} C[(s\ t)^R] &\rightarrow C[(s^R\ t)] \\ C[(\text{case } s\ \text{alts})^R] &\rightarrow C[(\text{case } s^R\ \text{alts})] \end{aligned}$$

Regel-Anwendung solange, bis nicht mehr möglich.

Danach die Normalordnungs-Reduktion: (falls möglich)

$$\text{Beta} \quad C[((\lambda x.t)^R s)] \rightarrow C[t[s/x]]$$

$$\text{Case} \quad C[(\text{case}_T (c t_1 \dots t_n)^R \text{ of } \{\dots; c x_1 \dots x_n \rightarrow s; \dots\})] \\ \rightarrow C[s[t_1/x_1, \dots, t_n/x_n]]$$

wenn keine möglich ist, dann gibt es zwei Möglichkeiten:
WHNF oder dynamischer Typfehler.

WHNF: genau dann, wenn R -Label des Top-Terms sich durch die Regeln nicht verschieben lässt.

Normalordnung reduziert einen R -markierten Unterterm,
bis dieser in WHNF ist.

Dies ergibt einen rekursiven Auswertungsalgorithmus

(leider einen ineffizienten, den man so nicht implementiert)

$\text{bot} := (\lambda x.(x\ x)) (\lambda y.(y\ y))$ terminiert nicht:

Die (beta)-Reduktion ergibt:

$$(\lambda x.(x\ x))^R (\lambda y.(y\ y)) \xrightarrow{n} ((\lambda y.(y\ y)) (\lambda y.(y\ y))) \xrightarrow{n} \dots$$

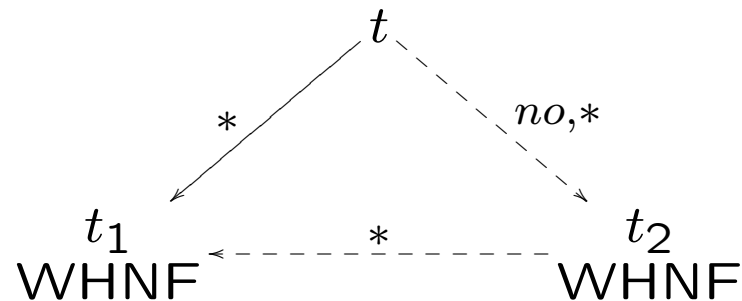
aussagenlogische Verknüpfungen **und**, **oder**, **nicht**:

```
und    =  $\lambda x . \lambda y . \text{case}_{\text{Boo1}} x \text{ of } \{\text{True} \rightarrow y; \text{False} \rightarrow \text{False}\}$   
oder   =  $\lambda x . \lambda y . \text{case}_{\text{Boo1}} x \text{ of } \{\text{True} \rightarrow \text{True}; \text{False} \rightarrow y\}$   
nicht  =  $\lambda x . \text{case}_{\text{Boo1}} x \text{ of } \{\text{True} \rightarrow \text{False}; \text{False} \rightarrow \text{True}\}$ 
```

In Haskell sind die logischen Verknüpfungen bereits im Prelude definiert als `&&`, `||`, `not`.

Satz (Standardisierung). Sei t ein (evtl. offener) Ausdruck und $t \xrightarrow{*} t_1$ mit (Beta) und (case)-Reduktionen, wobei t_1 eine WHNF ist.

Dann existiert eine WHNF t_2 , so dass $t \xrightarrow{no,*} t_2$, und $t_2 \xrightarrow{*} t_1$.



Folgerung: es ist korrekt, an beliebigen Positionen zu reduzieren, bis eine WHNF erreicht ist.

Wenn die erreichte WHNF eine Konstruktorkonstante ist, dann ist sie sogar eindeutig.

Wir geben die [Syntax von KFP](#) hier zur Information schon mal an.
Sie wird in der zweiten Semesterhälfte genauer besprochen

KFP ist analog zu Kernsprachen von Compilern funktionaler Programmiersprachen.

Einfache Syntax, möglichst wenig vordefiniert

Syntax: von KFP:

Es gibt N *Konstruktoren* $c_i, i = 1, \dots, N$
mit jeweils Stelligkeit $ar(c_i)$

CFG für KFP-Ausdrücke:

$EXP ::= V \quad V \text{ sind Variablen}$
| $\lambda V . EXP$
| $(EXP EXP)$
| $(c EXP_1 \dots EXP_n)$ wobei c Konstruktor und $n = ar(c)$
| $(\text{case } EXP \{Pat_1 \rightarrow Exp_1; \dots; Pat_{N+1} \rightarrow Exp_{N+1}\})$
 Pat_i ist Pattern zum Konstruktor i , und
 Pat_{N+1} das Pattern *lambda*.
 $Pat ::= (c V_1 \dots V_{ar(c)}) \mid \text{lambda}$

KFP: Unterschied zu KFPT:

- KFP hat ein ungetyptes case-Konstrukt
- zu jedem Konstruktoren im case ex. eine Alternative
- Abstraktionen haben Pattern: `lambda`

⇒: Dadurch kann man `seq` in KFP definieren

Wir erweitern KFPT um Haskell's **rekursive Superkombinatoren**.

Die Erweiterung nennen wir **KFPTS**

⇒: Programmierung in KFPTS wie in einem vereinfachten Haskell

Ausdrücke wie in KFPT, nur um Superkombinatornamen erweitert.

Es gibt Funktionsnamen (Superkombinatornamen) als neue Konstanten und eine Menge von Definitionen der Form

$Superkombinatorname \ V_1 \dots V_n = Exp$

Bedingungen:

- $V_1 \dots V_n$ sind verschieden;
In Exp kommen nur diese Variablen frei vor.
- Die Namensräume für Variablen, Konstruktoren, Kombinatornamen sind disjunkt.
- Jeder Superkombinator wird höchstens einmal definiert

Das eröffnet die Möglichkeit, [rekursive Funktionen](#) zu definieren da Namen von Funktionen beliebig verwendet werden können.

Ein *KFPTS-Programm* besteht aus:

1. Einer Menge von Typen und Konstruktoren wie in KFPT,
2. einer Menge von Kombinator-Definitionen
3. und aus einem Ausdruck.
definiert als $\boxed{\text{main} = \text{Exp}}$ (main-Konvention)

Reduktionsregeln:

Beta
$$\frac{((\lambda x.t) s)}{t[s/x]}$$

SK-Beta
$$\frac{(f a_1 \dots a_n)}{r[a_1/x_1, \dots, a_n/x_n]}$$
 wenn f definiert ist als: $f x_1 \dots x_n = r$
d.h. wenn $ar(f) = n$

Case
$$\frac{(\text{case}_T (c t_1 \dots t_n) \text{ of } \{ \dots ; c x_1 \dots x_n \rightarrow s ; \dots \})}{s[t_1/x_1, \dots, t_n/x_n]}$$

wobei $n = ar(c)$ und c ist T -Konstruktor

Unterschied zu Beta in KFPT:

SK-Beta reduziert **nur bei ausreichend vielen** Argumenten.

```
map f xs = caseList xs of {(Cons y ys) → (Cons (f y) (map f ys));  
                           Nil → Nil}  
main     = map not (Cons True Nil)
```

Reduktionskontexte in KFPTS wie in KFPT

Der Wert des Programms
ist der Wert des `main`-Ausdrucks nach Auswertung.

KFPTS-Normalordnungsreduktion:
innerhalb eines KFPTS-Reduktionskontexts wird
unmittelbar mit Case, Beta oder SK-Beta reduziert.

Normal-Ordnungs Reduktionen mit R-Label:

Beta $C[((\lambda x.t)^R s)] \rightarrow C[t[s/x]]$

Case $C[(\text{case}_T (c t_1 \dots t_n)^R \text{ of } \{ \dots ; c x_1 \dots x_n \rightarrow s ; \dots \})] \rightarrow C[s[t_1/x_1, \dots, t_n/x_n]]$

Neu :

SK-Beta $C[(SK^R s_1 \dots s_{ar(SK)})] \rightarrow C[r_{SK}[s_1/V_1, \dots, s_n/V_n]]$

Eine **KFPTS-WHNF** ist entweder

- ein Ausdruck $(c\ t_1 \dots t_n)$; oder
- ein Ausdruck $(f\ t_1 \dots t_n)$, wobei f Kombinator und $n < ar(f)$; oder
- ein Ausdruck $\lambda x.s$

Beispiel: `map` und `(map not)` sind WHNFs

(strict f t) macht f strikt im ersten Argument:

$$\frac{\text{Wert}(f\ t) = a; t\Downarrow}{\text{Wert}(\text{strict } f\ t) = a} \quad \text{und} \quad \frac{t\Uparrow}{(\text{strict } f\ t)\Uparrow}$$

(seq s t) wertet s und t aus und gibt den Wert von t zurück:

$$\frac{\text{Wert}(t) = a; s\Downarrow}{\text{Wert}(\text{seq } s\ t) = a} \quad \text{und} \quad \frac{s\Uparrow}{(\text{seq } s\ t)\Uparrow}$$

Beide sind gleichwertig:

$$\text{strict } f \ t = \text{seq } t \ (f \ t) \quad \text{seq } s \ t = \text{strict } (\backslash x \rightarrow t) \ s$$

In KFP sind `strict` und `seq` definierbar:

$$\text{strict } f = \backslash x . \text{ case } x \text{ of } \{p_1 \rightarrow f \ x; \dots ; p_N \rightarrow f \ x; p_{\{N+1\}} \rightarrow f \ x\}$$

Beachte: In KFPT, KFPTS sind `strict` und `seq` **nicht** definierbar.

KFP	Abstraktionen, Anwendungen, Konstruktoranwendungen jedes case hat alle Pattern und ein extra Pattern lambda
KFPT	Abstraktionen, Anwendungen, Konstruktoranwendungen jedes case hat nur Pattern eines Typs
KFPTS	Wie KFPT; Superkombinatoren sind erlaubt.
KFPT + seq	KFPT und seq wird als definiert angenommen
KFPTS + seq	KFPTS und seq wird als definiert angenommen
KFPTSP	KFPTS + seq und polymorphe Typisierung