ՀԱՅԱՍՏԱՆԻ ԳԻՏՈՒԹՅՈՒՆՆԵՐԻ ԱԶԳԱՅԻՆ ԱԿԱԴԵՄԻԱ
НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК АРМЕНИИ
NATIONAL ACADEMY OF SCIENCES OF ARMENIA
ДОКЛАДЫ ԶԵԿՈՒՅՑՆԵՐ REPORTS

Հատոր Том 109 2009 № 1

МАТЕМАТИКА

УДК 519.68:510

Г. Г. Грачян

Комбинаторы неподвижной точки и основная семантика бестиповых функциональных программ

(Представлено чл.-кор. НАН РА И.Д. Заславским 17/XI 2008)

Ключевые слова: *неподвижная точка, комбинатор неподвижной точки, семантика, лямбда-исчисление*

Бестиповая функциональная программа представляет собой систему уравнений с отделяющимися переменными в бестиповом λ-исчислении [1, 2]. Основная семантика таких программ обычно определяется с помощью комбинатора неподвижной точки Y. В данной работе доказывается, что семантики, определяемые с помощью разных комбинаторов неподвижной точки, эквивалентны в том смысле, что если применить их к одним и тем же данным, то получим один и тот же результат.

В [2] определяются и исследуются алгоритмы интерпретации бестиповых функциональных программ. Процедурные семантики, основанные на этих алгоритмах, сравниваются с основной семантикой, определяемой с помощью комбинатора неподвижной точки Y. Из настоящей работы следует, что результаты, полученные в [2], верны для семантик, определяемых с помощью произвольного комбинатора неподвижной точки.

1. Используемые определения и результаты. Пусть V — счетное множество переменных и Λ — множество бестиповых термов, определяемых обычным образом (см. [1]). Традиционным образом вводятся понятия свободного и связанного вхождения переменной в терм и понятие свободной переменной терма. Множество всех свободных переменных терма M условимся обозначать $FV(M)$. Терм, не содержащий свободных переменных, назовем замкнутым. Множество всех замкнутых термов обозначим Λ^0. Через $M[x_1, \dots, x_m]$ условимся обозначать терм M с указанием интересующих нас попарно различных переменных $x_1, \ldots, x_m$, $m \geq 1$. Через $M[N_1, \ldots, N_m]$ обозначим терм, полученный в результате одновременной подстановки термов $N_1, \ldots, N_m$ в терм M вместо всех свободных вхождений переменных $x_1, \ldots, x_m$, соответственно. Подстановка называется допустимой, если ни одна свободная переменная подставляемого терма не связывается в результате подстановки. Далее мы будем рассматривать только допустимые подстановки. Термы M и N называются конгруэнтными (обозначим $M \equiv N$), если один терм можно получить из другого переименованием связанных переменных. Далее мы не будем различать конгруэнтные термы.

Обычным образом определяются понятие β-редекса, отношения одношаговой β-редукции ($\rightarrow_\beta$), β-редукции ($\twoheadrightarrow_\beta$), β-равенства ($=_\beta$) и понятие β-нормальной формы. Так как в дальнейшем мы будем иметь дело только с β-редукцией, условимся далее β-редекс называть просто редексом, отношение одношаговой β-редукции — просто одношаговой редукцией, отношение β-редукции — просто редукцией, а β-нормальную форму — просто нормальной формой. Будем также опускать символ β в соответствующих обозначениях, т.е. отношение $\rightarrow_\beta$ обозначать $\rightarrow$, отношение $\twoheadrightarrow_\beta$ — $\twoheadrightarrow$, а отношение $=_\beta$ — $=$. Множество всех нормальных форм условимся обозначать NF, а множество всех замкнутых нормальных форм — NF^0.

Если во время некоторой одношаговой редукции свертывается (заменяется сверткой) самый левый редекс, то эта одношаговая редукция называется левой одношаговой редукцией. Редукция, состоящая из левых одношаговых редукций, называется левой редукцией. Левую одношаговую редукцию $M \rightarrow N$ обозначим $M \underset{L}{\rightarrow} N$, а левую редукцию $M \twoheadrightarrow N$ обозначим $M \underset{L}{\twoheadrightarrow} N$. Известно, что если терм M имеет нормальную форму N, то $M \underset{L}{\twoheadrightarrow} N$, т.е. к нормальной форме можно "дойти" с помощью левой редукции. Доказательство следующих теорем можно найти в [1].

Теорема Черча-Россера.

a) Если $M \twoheadrightarrow M_1$ и $M \twoheadrightarrow M_2$, то существует такой терм N, что $M_1 \twoheadrightarrow N$ и $M_2 \twoheadrightarrow N$.

b) Если $M_1 = M_2$, то существует такой терм N, что $M_1 \twoheadrightarrow N$ и $M_2 \twoheadrightarrow N$.

Следствие теоремы Черча-Россера.

a) Если $M = N$, где N — нормальная форма, то $M \twoheadrightarrow N$.

b) Если $M = N_1$ и $M = N_2$, где N_1 и N_2 — нормальные формы, то $N_1 \equiv N_2$.

Теорема о неподвижной точке. *Существует такой терм Z, что если его применить к произвольному терму M, то получим неподвижную точку терма M, т.е. $M(ZM) = ZM$. Такие термы называются комбинаторами неподвижной точки.*

Например, терм $Y \equiv \lambda h.(\lambda x.h(xx))(\lambda x.h(xx))$ является комбинатором неподвижной точки.

Для формулировки теоремы о кратной неподвижной точке введем обозначения для некоторых термов:

$$< M_1, \ldots, M_n > \equiv \lambda x.xM_1 \ldots M_n,$$

где $x \in V$, $M_i \in \Lambda$, $x \notin FV(M_i)$, $i = 1, \ldots n$, $n \geq 1$;

$$U_i^n \equiv \lambda x_1 \ldots x_n.x_i,$$

где $x_j \in V$, $j \neq k \Rightarrow x_j \not\equiv x_k$, $k, j = 1, \ldots n$, $1 \leq i \leq n$, $n \geq 1$;

$$P_i^n \equiv \lambda x.xU_i^n,$$

где $x \in V$, $1 \leq i \leq n$, $n \geq 1$;

Теорема о кратной неподвижной точке. *Пусть $M_1, \ldots, M_n$ ($n \geq 1$) — произвольные термы, Z — комбинатор неподвижной точки и*

$$L_i^Z \equiv P_i^n(Z(\lambda x. < M_1(P_1^n x) \ldots (P_n^n x), \ldots, M_n(P_1^n x) \ldots (P_n^n x) >)),$$

$i = 1, \ldots, n$. *Тогда* $M_i L_1^Z \ldots L_n^Z = L_i^Z$, $i = 1, \ldots, n$.

2. **Инвариантность основной семантики бестиповых функциональных программ относительно комбинатора неподвижной точки.** Бестиповая функциональная программа P (см. [2]) представляет собой систему уравнений следующего вида:

$$\begin{aligned} f_1 &= M_1[f_1, \ldots, f_n], \\ &\ldots \\ f_n &= M_n[f_1, \ldots, f_n], \end{aligned} \tag{1}$$

где $f_i \in V$, $i \neq j \Rightarrow f_i \not\equiv f_j$, $M_i[f_1, \ldots, f_n] \in \Lambda$, $FV(M_i[f_1, \ldots, f_n]) \subseteq \{f_1, \ldots, f_n\}$, $i, j = 1, \ldots, n$, $n \geq 1$. Главным уравнением программы (1) будем считать первое уравнение системы. Зафиксируем комбинатор неподвижной точки Z и рассмотрим следующее решение указанной системы:

$$(L_1^Z, \ldots, L_n^Z),$$

где $L_i^Z \equiv P_i^n(Z(\lambda x. < M_1[P_1^n x, \ldots, P_n^n x], \ldots, M_n[P_1^n x, \ldots, {}_n^n x] >))$, $i = 1, \ldots, n$. Главной компонентой указанного решения назовем терм L_1^Z, который и является основной семантикой неподвижной точки программы P, соответствующей комбинатору неподвижной точки Z. Определим множество $Fix(P, Z)$, соответствующее семантике L_1^Z следующим образом:

$$Fix(P, Z) = \{(Q_1, \ldots, Q_k, M_0) | L_1^Z Q_1, \ldots, Q_k \twoheadrightarrow M_0, Q_1, \ldots, Q_k, M_0 \in NF^0, k \geq 0\}.$$

Основным результатом данной работы является следующая теорема об инвариантности основной семантики.

Теорема 1 (об инвариантности основной семантики). *Для любых комбинаторов неподвижной точки Z и Z' и любой бестиповой функциональной программы P имеем*

$$Fix(P, Z) = Fix(P, Z'),$$

т.е. множество, соответствующее семантике неподвижной точки, инвариантно относительно комбинатора неподвижной точки.

Справедливость теоремы 1 вытекает из следующей основной леммы 1.

Лемма 1 (основная). *Пусть Z, Z' — произвольные комбинаторы неподвижной точки и M — терм, в котором зафиксировано некоторое вхождение подтерма ZL, $L \in \Lambda$. Обозначим через M' терм, который получается из M заменой указанного вхождения подтерма ZL термом $Z'L$. Предположим также, что свободные вхождения переменных в соответствующих вхождениях подтермов ZL и $Z'L$ не связываются в термах M и M', соответственно. Тогда M_0 является нормальной формой терма M в том и только том случае, когда M_0 является нормальной формой терма M'.*

Приведем **доказательство** теоремы 1, используя основную лемму 1. Пусть P — произвольная бестиповая функциональная программа вида (1), а Z и Z' — произвольные комбинаторы неподвижной точки. Тогда,

$$Fix(P, Z) = \{(Q_1, \ldots, Q_k, M_0) | L_1^Z Q_1 \ldots Q_k \twoheadrightarrow M_0, Q_1, \ldots, Q_k, M_0 \in NF^0,\ k \geq 0\},$$

где

$$L_1^Z \equiv P_1^n(Z(\lambda x. < M_1[P_1^n x, \ldots, P_n^n x], \ldots, M_n[P_1^n x, \ldots, P_n^n x] >))$$

и

$$Fix(P, Z') = \{(Q_1, \ldots, Q_k, M_0) | L_1^{Z'} Q_1 \ldots Q_k \twoheadrightarrow M_0, Q_1, \ldots, Q_k, M_0 \in NF^0,\ k \geq 0\},$$

где

$$L_1^{Z'} \equiv P_1^n(Z'(\lambda x. < M_1[P_1^n x, \ldots, P_n^n x], \ldots, M_n[P_1^n x, \ldots, P_n^n x] >)).$$

Легко видеть, что для произвольных термов $Q_1, \ldots, Q_k, M_0 \in N^0$ $(k \geq 0)$, термы $L_1^Z Q_1 \ldots Q_k$ и $L_1^{Z'} Q_1 \ldots Q_k$ удовлетворяют условиям леммы 1 и, следовательно, $L_1^Z Q_1 \ldots Q_k \twoheadrightarrow M_0$ тогда и только тогда, когда $L_1^{Z'} Q_1 \ldots Q_k \twoheadrightarrow M_0$. Теорема 1 доказана.

Остается доказать основную лемму 1. Прежде всего введем понятие контекста неподвижной точки, которое является ключевым понятием при доказательстве основной леммы 1.

Терм $C[f]$ назовем контекстом неподвижной точки относительно переменной f, если он удовлетворяет условию $C[f] = fC[f]$ и не имеет вид $fC'[f]$,

$C'[f] \in \Lambda$. Множество всех контекстов неподвижной точки относительно переменной f обозначим через $FP(f)$. Далее, когда будет ясно, относительно какой переменной рассматривается контекст неподвижной точки, мы не будем явно указывать эту переменную. Например, контекстами неподвижной точки относительно переменной f являются термы $Yf \equiv (\lambda h.h(xx))(\lambda x.h(xx))f$ и $W[f] \equiv (\lambda x.f(xx))(\lambda x.f(xx))f$. Однако термы $f(Yf)$ и $fW[f]$ не являются контекстами неподвижной точки, несмотря на то, что $f(Yf) = f(f(Yf))$ и $fW[f] = f(fW[f])$. Нетрудно заметить, что для любого комбинатора неподвижной точки Z терм Zf, где $f \in V$, является контекстом неподвижной точки и что для любого контекста неподвижной точки $C[f]$ терм $\lambda f.C[f]$ является комбинатором неподвижной точки.

Пусть заданы комбинаторы неподвижной точки Z, Z' и термы L, M и M', удовлетворяющие условиям основной леммы 1. Заметим, что термы M и M' отличаются только вхождениями подтермов ZL и $Z'L$, которые принадлежат множеству $FP(L)$. Основная лемма утверждает, что M_0 является нормальной формой терма M в том и только в том случае, если M_0 является нормальной формой терма M'. Введем понятие C-эквивалентности и сформулируем лемму 2 о C-эквивалентности, которая обобщит утверждение основной леммы 1 для термов, отличающихся несколькими вхождениями произвольных подтермов из множества $FP(L)$.

C-эквивалентность. Термы M и M' назовем C-эквивалентными и напишем $M \sim_C M'$, если они отличаются только вхождениями термов из множества $FP(L)$, т.е. M' получается из M путем замены некоторых вхождений подтермов $C_1[L], \ldots, C_k[L] \in FP(L)$, $k \geq 0$, произвольными (возможно теми же самыми) термами $C'_1[L], \ldots, C'_k[L] \in FP(L)$ из множества $FP(L)$ (рис. 1). Заметим, что если $k = 0$, то $M \equiv M'$.

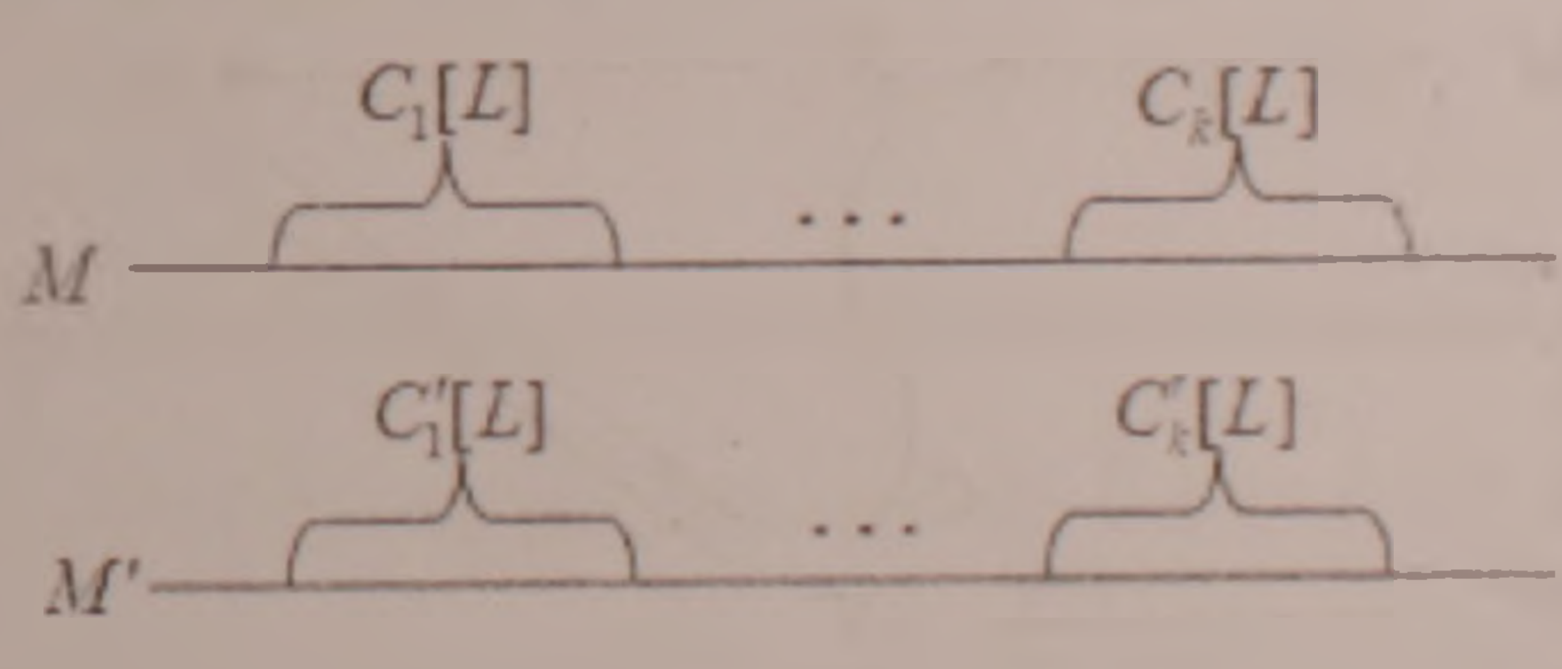

Рис. 1

Лемма 2 (о C-эквивалентности). *Если $M \sim_C M'$, то M_0 является нормальной формой терма M тогда и только тогда, когда M_0 является нормальной формой терма M'.*

Для доказательства леммы 2 о C-эквивалентности приведем без доказательства следующую лемму 3 о левой редукции C-эквивалентных термов.

Лемма 3 (о левой редукции C-эквивалентных термов). *Пусть $M \sim_C M'$ и $M \notin NF$. Тогда существуют такая непустая левая редукция $M \underset{L}{\twoheadrightarrow} N$ и такой терм $N' = M'$, что $N' \sim_C N$.*

Доказательство леммы 2. Пусть $M \sim_C M'$. Надо доказать, что M_0 является нормальной формой терма M тогда и только тогда, когда M_0 является нормальной формой терма M'. Предположим, что M имеет нормальную форму M_0, и докажем, что M_0 является нормальной формой и терма M', т.е. что $M' = M_0$. Обратная сторона леммы доказывается аналогично. Так как M_0 — нормальная форма M, то существует левая редукция $M \underset{L}{\twoheadrightarrow} M_0$. Докажем индукцией по длине редукции $M \underset{L}{\twoheadrightarrow} M_0$, которую обозначим через m, что из $M \sim_C M'$ и $M \underset{L}{\twoheadrightarrow} M_0 \in NF$ следует $M' = M_0$. Если $m = 0$, то $M \equiv M_0$ и M является нормальной формой. Следовательно, все подтермы терма M тоже нормальные формы. Так как M и M' могут отличаться только вхождениями подтермов из множества $FP(L)$ и так как последние не могут быть нормальными формами, то $M \sim_C M'$ означает $M' \equiv M \equiv M_0$. Теперь докажем шаг индукции. Предположим, что $m > 0$ и что для любых термов N и N', если $N \sim_C N'$ и левая редукция $N \underset{L}{\twoheadrightarrow} N_0 \in NF$ имеет длину меньше m, $N' = N_0$. Так как $m > 0$, то редукция $M \underset{L}{\twoheadrightarrow} M_0$ непуста и, следовательно, $M \notin NF$. Тогда согласно лемме 3 о левой редукции C-эквивалентных термов существуют такая непустая левая редукция $M \underset{L}{\twoheadrightarrow} N$ и такой терм $N' = M'$, что $N' \sim_C N$. Так как левая редукция однозначна и $M \underset{L}{\twoheadrightarrow} M_0$ есть максимально длинная левая редукция терма M, то из $M \underset{L}{\twoheadrightarrow} N$ следует $M \underset{L}{\twoheadrightarrow} N \underset{L}{\twoheadrightarrow} M_0$. Поскольку редукция $M \underset{L}{\twoheadrightarrow} N$ непуста, то длина редукции $N \underset{L}{\twoheadrightarrow} M_0$ меньше m. Следовательно, мы можем применить предположение индукции для терма N. Согласно предположению индукции из $N \underset{L}{\twoheadrightarrow} M_0$ и $N' \sim_C N$ следует $N' = M_0$ и поскольку $M' = N'$, то $M' = N' = M_0$ (рис 2). Лемма 2 доказана.

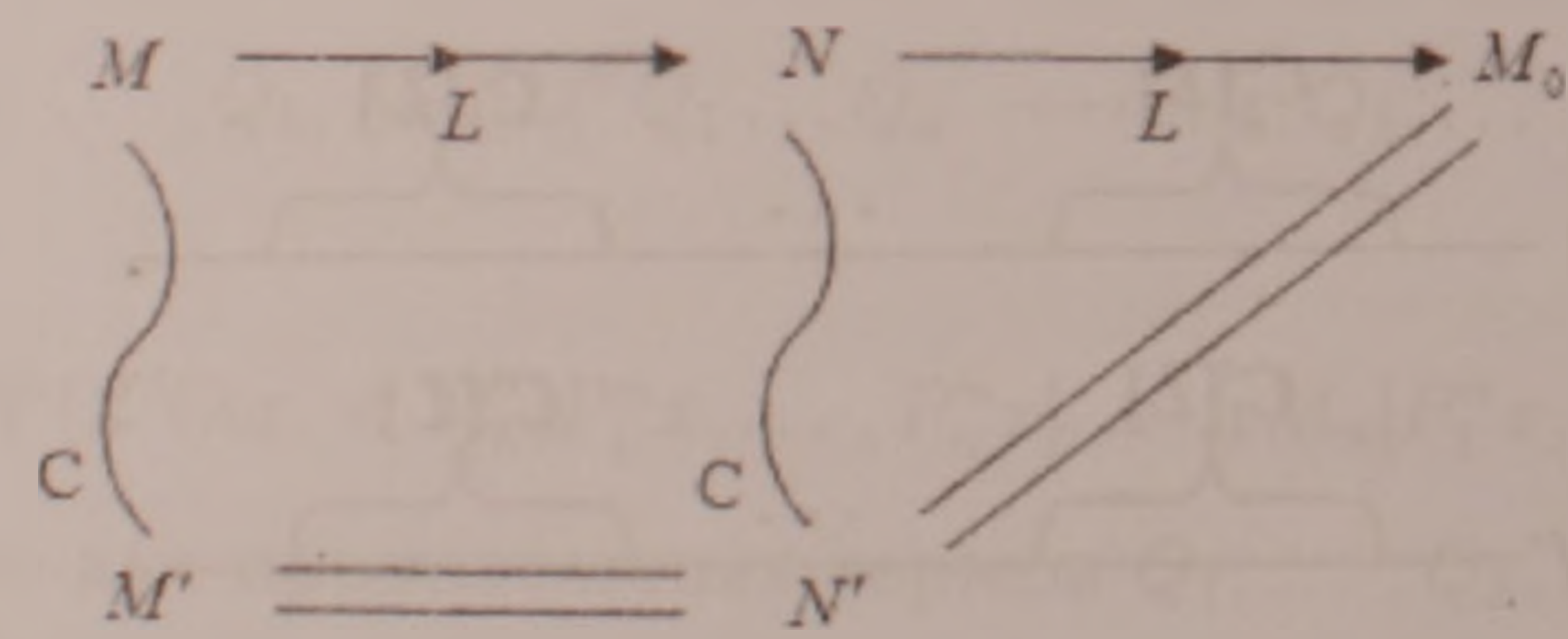

Рис. 2

3. Основные семантики бестиповых функциональных программ при различных комбинаторах неподвижной точки и β-равенство. Итак, если L_1 и L'_1 — семантики бестиповой функциональной программы P, определяемые с помощью разных комбинаторов неподвижной точки, то согласно теореме 1 они эквивалентны в следующем смысле:

$$L_1 Q_1 \dots Q_k \twoheadrightarrow M_0 \Leftrightarrow L'_1 Q_1 \dots Q_k \twoheadrightarrow M_0,$$

где $Q_1, \dots, Q_k$, $M_0 \in NF^0$. Однако, как выясняется, не всегда $L_1 = L'_1$, т.е. семантики одной и той же бестиповой функциональной программы, определяемые с помощью разных комбинаторов неподвижной точки, могут не быть β-равными.

Теорема 2. *Существует бестиповая функциональная программа P_0 и такие комбинаторы неподвижной точки Z и Z', что если L — семантика программы P_0, определяемая с помощью комбинатора неподвижной точки Z, а L' — семантика программы P_0, определяемая с помощью комбинатора неподвижной точки Z', то $L \neq L'$.*

Приведем соответствующий пример программы P_0 и комбинаторов неподвижной точки Z и Z', удовлетворяющих условиям теоремы 2. Прежде всего приведем несколько общепринятых обозначений. Термами $T \equiv \lambda xy.x$, $F \equiv \lambda xy.y$ принято представлять истинностные значения $True$ и $False$, соответственно. Пусть B, M и N — произвольные термы, тогда выражением $if\ B\ then\ M\ else\ N$ обозначим терм BMN. Нетрудно заметить, что в случае $B = T$ имеет место $BMN = M$, а в случае $B = F$ — $BMN = N$. Определим термы для представления чисел. Для каждого числа $n \geq 0$ определим терм $\overline{n}$ следующим образом:

$$\overline{0} \equiv \lambda x.x,\ \overline{n+1} \equiv \lambda x.xF\overline{n}.$$

Будем использовать также следующие термы:

$$S^+ \equiv \lambda xy.yFx,\ P^- \equiv \lambda x.xF,\ Zero \equiv \lambda x.xT \text{ и } I \equiv \lambda x.x\,.$$

Нетрудно убедиться, что для любого числа $n \geq 0$ имеем: $S^+\overline{n} = \overline{n+1}$, $P^-\overline{n+1} = \overline{n}$, $Zero\,\overline{n} = T$, если $n = 0$, и $Zero\,\overline{n} = F$, если $n > 0$, и $IM = M$ для любого терма M.

Пусть P_0 — следующая бестиповая функциональная программа:

$$F_1 = M_1[f_1] \equiv \lambda n.if\ Zero\,n\ then\ \overline{0}\ else\ f_1(P^- n),$$

$Z \equiv Y \equiv \lambda h.(\lambda x.h(xx))(\lambda x.h(xx))$ и $Z' \equiv \lambda h.(\lambda xy.h(xxy))(\lambda xy.h(xxy))T$. Можно доказать, что если L — семантика программы P_0, определяемая с помощью комбинатора неподвижной точки Z, а L' — семантика той же программы, определяемая с помощью комбинатора неподвижной точки Z', то $L \neq L'$.

Ереванский государственный университет

Г. Г. Грачян

Комбинаторы неподвижной точки и основная семантика бестиповых функциональных программ

Бестиповая функциональная программа представляет собой систему уравнений с отделяющимися переменными в бестиповом λ-исчислении. Основная семантика таких программ обычно определяется с помощью комбинатора неподвижной точки Y. Доказывается, что семантики, определяемые с помощью разных комбинаторов неподвижной точки, эквивалентны в том смысле, что если применить их к одним и тем же данным, то получим один и тот же результат.

Գ. Գ. Հրաչյան

Անշարժ կետի կոմբինատորները և առանց տիպերի ֆունկցիոնալ ծրագրերի հիմնական սեմանտիկան

Առանց տիպերի ֆունկցիոնալ ծրագիրը իրենից ներկայացնում է անջատվող փոփոխականներով հավասարումների համակարգ առանց տիպերի λ-հաշվում: Նման ծրագրերի հիմնական սեմանտիկան սովորաբար սահմանվում է Y անշարժ կետի կոմբինատորի միջոցով: Աշխատանքում ցույց է տրվում, որ անշարժ կետի տարբեր կոմբինատորներով սահմանված սեմանտիկաները համարժեք են այն իմաստով, որ եթե դրանք կիրառենք նույն տվյալների համար ապա կստանանք նույն արդյունքը:

G. G. Hrachyan

Fix-Point Combinators and the Main Semantics of Type-Free Functional Programs

The type-free functional program is a system of equations with separated variables in the type-free λ-calculus. The main semantics of such programs is usually defined with the fixpoint combinator Y. The current paper shows that the semantics defined using the different fixpoint combinators are equivalent in the case if we apply these semantics to the same data, we will get the same result.

Литература

1. *Barendregt H.P.* The Lambda Calculus, Its Syntax and Semantics, North-Holland Pub. Comp., 1981 (рус. пер. *Барендрегт Х.* Лямбда исчисление, его синтаксис и семантика. М. Мир. 1985).

2. *Нигиян С.А., Аветисян С.А.* - Программирование. 2002. N 3. C. 5-14.