ОБОБЩЕННАЯ ВЫЧИСЛИМОСТЬ И ОПРЕДЕЛИМОСТЬ
(Вычислительные системы)
1998 год Выпуск 161

УДК 519.685

СЕМАНТИЧЕСКАЯ ФОРМАЛИЗАЦИЯ ОБЪЕКТНОЙ МОДЕЛИ ДАННЫХ[1]

О.Г. Юрченко

Введение

В основу концепции семантического программирования [1,2] положено понятие истинности на модели и процесс исполнения семантической программы представляет собой выборку тех элементов модели, которые удовлетворяют условию запроса. Поэтому практически на любой язык запросов можно смотреть как на язык семантического программирования и одним из возможных применений последнего является формальное описание семантики распространенных языков запросов. Другим возможным применением семантического программирования в области баз данных является его использование для описания результатов концептуального проектирования баз данных и формальной спецификации ограничений.

В данной работе описывается семантическая формализация объектной модели данных и определяется формальный объектно-ориентированный язык запросов, который может быть использован для описания семантики OQL (вариант, описанный в стандарте ODMG 93 v 1.2

[1]Работа выполнена при поддержке гранта РФФИ № 96-01-00097.

[7]). Кроме того, он может служить формальной основой для языка описания ограничений при проектировании баз данных в рамках объектной модели.

Большое влияние на данную работу оказали идеи, изложенные в [3,5,6], кроме того существенно использованы результаты, полученные в [8,9].

Формализация объектной модели данных

Пусть $\langle Class, \sqsubseteq \rangle$ — конечное частично упорядоченное множество классов, $\sqsubseteq$ — отношение наследования. Полагаем, что $Class = TClass \bigcup PClass \bigcup \{set(c) | c \in TClass \bigcup \bigcup PClass\}$, причем $TClass \bigcap PClass = \emptyset$.

Под $TClass$ мы понимаем множество базовых классов, объекты которых являются временными (transient), а $PClass$ обозначает множество классов, объекты которых хранятся в базе данных (persistent objects). Объекты классов из $TClass$ будут рассматриваться исключительно как строительный материал для объектов классов из $PClass$. Примерами классов из $TClass$ могут быть Integer, Float и т.д. При этом мы будем придерживаться естественного предположения, что если $c_1, c_2 \in Class$ и $c_1 \sqsubseteq c_2$, то либо $c_1, c_2 \in PClass$, либо $c_1, c_2 \in TClass$. Также предполагаем, что если $c_1, c_2 \in Class, c_1 \sqsubseteq c_2$, то $set(c_1) \sqsubseteq set(c_2)$.

Конечную последовательность классов $c_1; \ldots; c_n$, $n > 0$, будем называть типом над $Class$. Естественным образом определим понятие подтипа: $c_1; \ldots; c_n \sqsubseteq_T \sqsubseteq_T t_1; \ldots; t_m \Leftrightarrow n = m \bigwedge \forall 1 \leq i \leq n \ c_i \sqsubseteq t_i$ Множество всех типов будем обозначать через $Type$.

Обозначим через σ следующую сигнатуру $\sigma = \langle Class, \sqsubseteq, Attr, Meth, Rel, Ext, Const \rangle$, где $Const$ — множество символов констант, $Attr, Meth, Rel$ — некоторые семейства конечных множеств символов, а Ext — множество символов:

$Attr = \{Attr_c | c \in Class\}$

$Meth = \{Meth_c | c \in Class\}$

$Rel = \{Rel_t | t \in Type\}$

$Ext = \{Ext_c | c \in PClass\}$,

где $Attr_c$ обозначает множество атрибутов класса c,

$Meth_c$ — множество методов класса c, Ext_c — множество хранимых объектов класса c (extent), Rel_t — семейство множеств отношений между классами.

Каждому символу из $Attr_c, Meth_c, Rel_t, Ext_c$ ($c \in Class$, $t \in Type$) припишем тип над $Class$, причем предполагаются выполненными следующие условия.

1. Если $A \in Attr_c$, то $type(A) = c'$ для некоторого $c' \in Class$. Причем, если $c \in TClass$, для каждого $A \in Attr_c$ $type(A) \in TClass$. Это означает, что в данной работе нас интересует, прежде всего, концептуальное представление информации, лежащее на уровне объектно-орентированной СУБД, и нас в данный момент совершенно не интересуют временные (transient) объекты, способные оперировать хранимыми (persistent) объектами (business objects).

2. Если $f \in Meth_c$, то $type(f) = c_1; \ldots; c_n$ для некоторого типа $c_1; \ldots; c_n \in Type$.

3. Если $P \in Rel_t$, то $type(P) = t$.

4. $type(Ext_c) = c$.

Далее опишем свойства, которым должны удовлетворять $Attr, Meth, Rel$:

1) если $c_1, c_2 \in Class$, $c_1 \sqsubseteq c_2$, то $Attr_{c_1} \bigcap Attr_{c_2} = \emptyset$;.

2) если $c_1, c_2, c_3 \in Class$ и $c_1 \sqsubseteq c_2$, $c_1 \sqsubseteq c_3$ и $c_2 \neq c_3$, то $Attr_{c_2} \bigcap Attr_{c_3} = \emptyset$;

3) для любых $c_0, c_1 \in Class$ ($c_0 \sqsubseteq c_1$) и $t \in Meth_{c_1}$ существует наименьший c, такой, что $t \in Meth_c$ и $c_0 \sqsubseteq c$;

4) если $m \in Meth_{c_1} \bigcap Meth_{c_2}$ и $c_1 \sqsubseteq c_2$, то m из $Meth_{c_1}$ и m из $Meth_{c_2}$ имеют один и тот же тип;

5) будем полагать,что множества $Attr_c \bigcup Meth_c, Rel_t$, Ext_c попарно не пересекаются для любых $c \in Class$, $t \in Type$;

6) $Attr_{c_1} \bigcap Meth_{c_2} = Attr_{c_2} \bigcap Meth_{c_1} = \emptyset$ для любых $c_1, c_2 \in Class$ таких, что $c_1 \sqsubseteq c_2$;

7) считаем, что для каждого $c \in Class$ для $t = c; c$ в Rel_t существует символ равенства $=$, для $t = c; set(c)$ в Rel_t существует символ $\in$;

8) будем полагать, что если $c \in TClass$, то для каждого $A \in Attr_c$ $type(A) \in TClass$.

ОПРЕДЕЛЕНИЕ 1. Моделью сигнатуры σ будем называть пару $\mathcal{M} = \langle (M_c)_{c \in Class}, \nu \rangle$, где $\nu = \{\nu_t\}_{t \in Type}$ — семейство интерпретаций сигнатурных символов, M_c — носитель класса $c \in Class$,

$$M_c = \{f : \bigcup_{c_0 \sqsubseteq c} Attr_{c_0} \to \bigcup_{p \in Class} M_p | f(A) \in M_{type(A)}\}.$$

Предполагаем, что носитель модели, множество $(M_c)_{c \in Class}$, удовлетворяет следующим условиям:

1) $c_1 \sqsubseteq c_2 \iff M_{c_1} \subseteq M_{c_2}$, т.е. каждый элемент из M_{c_1} является (может рассматриваться) элементом M_{c_2};

2) если $M_{c_1} \bigcap M_{c_2} \neq \emptyset$, то существуют $c^0, \ldots, c^m$ такие, что $M_{c_1} \bigcap M_{c_2} = M_{c^0} \bigcup \ldots \bigcup M_{c^m}$;

3) $M_{set(c)}$ — множество подмножеств над M_c.

Условия на ν:

1) для каждого $c \in PClass$ $\nu_c(Ext_c)$ — конечное подмножество M_c;

2) если $c \in Class, f \in Meth_c, type(f) = c_1; \ldots; c_n$, то $\nu_c(f) : M_c \times M_{c_1} \times \ldots \times M_{c_{n-1}} \longrightarrow M_{c_n}$;

3) если $R \in Rel_t, t = c_1; \ldots; c_n$, то $\nu_t(R) \subseteq M_{c_1} \times \ldots \times M_{c_n}$.

Множество $\bigcup_{c \in Class} M_c$ в дальнейшем будем обозначать через $|\mathcal{M}|$.

Теперь расширим сигнатуру $\sigma = \langle Class, \sqsubseteq, Attr, Meth, Rel, Ext, Const \rangle$ до сигнатуры $\sigma^+ = \langle Class^+, \sqsubseteq, Attr^+, Meth^+, Rel^+, Ext^+, Const \rangle$ следующим образом.

Добавим к $Class$ бесконечное множество $QClass$ новых классов. Можно считать (и это вполне естественно), что вместе с $QClass$ к $Class$ добавился элемент $Query$ и каждый класс из $QClass$ является производным от класса $Query$ (т.е. если $q \in QClass$, то $q \sqsubseteq Query$). Но в данной работе мы ограничимся только добавлением множества $QClass$, причем для большей простоты будем полагать, что если $q_1, q_2 \in QClass$, то $q_1 \not\sqsubseteq q_2$, а также если $q_1 \in QClass, q_2 \in Class$, то $q_1 \not\sqsubseteq q_2$ и $q_2 \not\sqsubseteq q_1$. Последнее предположение нуждается в дополнительном пояснении. Классы из $QClass$ являются определяемыми классами. Объекты этих классов являются вычисляемыми. То, что мы запрещаем наследовать классы из $QClass$ от классов из $PClass$, объясняется тем, что класс, производный от класса из $PClass$, множество хранимых объектов которого является вычисляемым, соответствует представлению (View) из реляционной модели данных и проводя аналогию кажется уместным считать, что этот класс сам содержится в $PClass$.

Полученное множество классов обозначим через $Class^+$.

$Meth^+ = Meth \bigcup \{Meth_q | q \in QClass\}$, где все добавленные множества $Meth_q$ — пустые.

Далее, для каждого $q \in QClass$ добавим в $Attr$ непустое множество $Attr_q$ и в Ext добавим элемент Ext_q. Полученные множества обозначим через $Attr^+$ и Ext^+.

Пусть $Type$ — множество типов над $Class^+$. Поставим в соответствие введенным символам типы из $Type$.

1. Если $A \in Attr_q$, то $type(A) = c$ для некоторого $c \in Class$.

2. $type(Ext_q) = q$ для всех $q \in QClass$.

Моделью сигнатуры σ^+ будет $\mathcal{M}^+ = \langle (M_c)_{c \in Class^+}, \nu, EXT, \mu \rangle$, где $\mathcal{M} = \langle (M_c)_{c \in Class}, \nu \rangle$ — модель сигнатуры σ, если $c \in QClass$, то $M_c = \{f : Attr_c \longrightarrow \bigcup\limits_{p \in Class} M_p | \ f(A) \in \in M_{type(A)}\}$, $EXT = \{EXT_q | q \in QClass\}$, где $EXT_q = = \mathcal{P}(M_q)$ — множество подмножеств M_q, μ — интерпретация символов из $\{Ext_q | q \in QClass\}$ в EXT такая,что $\mu(Ext_q) \in EXT_q$.

Объектно-ориентированный язык запросов

Введем в рассмотрение для каждого класса c из $Class^+$ бесконечное множество переменных $X(c)$ этого класса. Считаем, что $X(c)$ и $X(t)$ для различных c и t не пересекаются. Пусть X — множество всех предметных переменных. Введем следущие обозначения: $X(TClass) = = \bigcup\limits_{c \in TClass} X(c)$, $X(PClass) = \bigcup\limits_{c \in PClass} X(c)$, $X(QClass) = = \bigcup\limits_{c \in QClass} X(c)$.

Введем понятие терма сигнатуры σ^+.

ОПРЕДЕЛЕНИЕ 2.

1. Предметная переменная или константа класса c является термом этого класса.

2. Если t — терм класса c, $c \sqsubseteq s$, $A \in Attr_s$, $type(A) = s'$, то $t.A$ — терм класса s'.

3. Если $f \in Meth_c$, $type(f) = c_1;\ldots;c_n$, $c' \sqsubseteq c$, $c'_1 \sqsubseteq c_1,\ldots,c'_{n-1} \sqsubseteq c_{n-1}$, $t,t_1,\ldots,t_{n-1}$ — термы классов $c',c'_1,\ldots,c'_{n-1}$ соответственно, то $t.f(t_1,\ldots,t_{n-1})$ и $t.c' :: f(t_1,\ldots,t_{n-1})$ — термы класса c_n.

Множество переменных, входящих в терм t, обозначим через $FV(t)$.

Определим понятие формулы сигнатуры σ^+.

1. Если $P \in Rel_t$, $t = c_1;\ldots;c_n$, $c'_1 \sqsubseteq c_1,\ldots,c'_n \sqsubseteq c_n$, $t_1,\ldots,t_n$ — термы классов $c'_1,\ldots,c'_n$ соответственно, то $P(t_1,\ldots,t_n)$ и $t :: P(t_1,\ldots,t_n)$ — формулы сигнатуры σ^+.

2. Если t — терм класса c, $c \in PClass \bigcup QClass$, то $Ext_c(t)$ — формула.

Формулы вида 1 и 2 из данного определения будем называть атомарными.

3. Если Φ_1 и Φ_2 — формулы, то $\neg\Phi_1$, $(\Phi_1 \wedge \Phi_2)$, $(\Phi_1 \vee \Phi_2)$ — формулы.

4. Если Φ — формула сигнатуры σ^+, $x \in X$, то $\exists x\Phi$, $\forall x\Phi$ — формулы сигнатуры σ^+.

ОПРЕДЕЛЕНИЕ 3. Для каждой формулы Φ сигнатуры σ^+ определим множество $FV(\Phi)$ свободных переменных формулы Φ следующим образом:

1) если Φ — атомарная формула вида $P(t_1,\ldots,t_n)$ или $t :: P(t_1,\ldots,t_n)$, то $FV(\Phi) = FV(t_1)\bigcup\ldots\bigcup FV(t_n)$;

2) если Φ — атомарная формула вида $Ext_c(t)$, то $FV(\Phi) = FV(t)$;

3) если Φ — формула вида $\neg\Psi$, то $FV(\Phi) = FV(\Psi)$;

4) если Φ — формула вида $(\Phi_1 \wedge \Phi_2)$ или $(\Phi_1 \vee \Phi_2)$, то $FV(\Phi) = FV(\Phi_1)\bigcup FV(\Phi_2)$;

5) если Φ — формула вида $\exists x\Psi$ или $\forall x\Psi$, то $FV(\Phi) = FV(\Psi) \setminus \{x\}$.

Вхождение η переменной x в формулу Φ сигнатуры σ^+ будем называть связанным, если η лежит в области действия некоторого вхождения квантора $\forall$ или $\exists$, за которым сразу следует символ x. Если вхождение η переменной x в формулу Φ не является связаным, то будем называть его свободным.

ОПРЕДЕЛЕНИЕ 4. Пусть $\mathcal{M}^+$ — модель сигнатуры σ^+ и пусть $V \subseteq X$. Интерпретацией переменных V в $\mathcal{M}^+$ будем называть отображение $\xi : V \longrightarrow \bigcup\limits_{c \in Class^+} M_c$, сохраняющее типы, т.е. если $x \in V$ и x — переменная класса c, то $\xi(x) \in M_c$.

Если $FV(t) \subseteq V$ для терма t сигнатуры σ^+, то значение $t[\xi]$ терма t в $\mathcal{M}^+$ определяется следующим образом.

ОПРЕДЕЛЕНИЕ 5.

1) Если $t = x$, $x \in V$, то $t[\xi] = \xi(x)$;

2) если $t = c$, c — константа, то $t[\xi] = \nu(c)$;

3) если $t = t_0.f(t_1, \ldots, t_n)$, то $t[\xi] = \nu_c(f)(t_0[\xi], t_1[\xi], \ldots, t_n[\xi])$, где c — класс с наименьшим носителем M_c, содержащим $t_0[\xi]$ и $f \in Meth_c$ (по определению модели такой класс существует);

4) если $t = t_0.c :: f(t_1, \ldots, t_n)$, то $t[\xi] = \nu_c(f)(t_0[\xi], t_1[\xi], \ldots$ $\ldots, t_n[\xi])$;

5) если $t = t_0.A$, где $A \in Attr_c$, то $t[\xi] = t_0[\xi](A)$.

Для модели $\mathcal{M}^+$ сигнатуры σ^+, интерпретации $\gamma : V \longrightarrow |\mathcal{M}^+|$ и формулы Φ сигнатуры σ^+, для которой $FV(\Phi) \subseteq V$, определим отношение истинности на модели $\mathcal{M}^+ \models \Phi[\gamma]$ следующим образом.

ОПРЕДЕЛЕНИЕ 6.

1) Если $\Phi = t :: P(t_1, \ldots, t_n)$, $P \in Rel_t$, то $\mathcal{M}^+ \models \Phi[\gamma]$ эквивалентно $\langle t_1[\gamma], \ldots, t_n[\gamma] \rangle \in \nu_t(P)$;

2) если $\Phi = P(t_1, \ldots, t_n)$, $P \in Rel_t$, $t = c_1; \ldots; c_n$, $c'_1, \ldots$ $\ldots, c'_n$ — наименьшие сорта такие, что $c'_1 \sqsubseteq c_1, \ldots, c'_n \sqsubseteq c_n$ и $t_1[\gamma] \in M_{c'_1}, \ldots, t_n[\gamma] \in M_{c'_n}$, то $\mathcal{M}^+ \models \Phi[\gamma]$ эквивалентно $\langle t_1[\gamma], \ldots, t_n[\gamma] \rangle \in \nu_k(P)$, где k — наименьший тип такой, что $P \in Rel_k$ и $c'_1; \ldots; c'_n \sqsubseteq_T k$;

3) если $\Phi = \neg\Psi$, то $\mathcal{M}^+ \models \Phi[\gamma]$ тогда и только тогда, когда неверно, что $\mathcal{M}^+ \models \Psi[\gamma]$;

4) если $\Phi = (\Phi_1 \vee \Phi_2)$, то $\mathcal{M}^+ \models \Phi[\gamma]$ эквивалентно $\mathcal{M}^+ \models \Phi_1[\gamma]$ или $\mathcal{M}^+ \models \Phi_2[\gamma]$;

5) если $\Phi = (\Phi_1 \wedge \Phi_2)$, то $\mathcal{M}^+ \models \Phi[\gamma]$ эквивалентно $\mathcal{M}^+ \models \Phi_1[\gamma]$ и $\mathcal{M}^+ \models \Phi_2[\gamma]$;

6) если $\Phi = \exists x\Psi$, то $\mathcal{M}^+ \models \Phi[\gamma] \iff$ существует интерпретация $\gamma_1 : V_1 \longrightarrow |\mathcal{M}^+|$, для которой $x \in V_1$, $FV(\Phi) \subset V_1$, $\mathcal{M}^+ \models \Psi[\gamma_1]$, и для всех $x \in FV(\Phi)$ $\gamma_1(x) = \gamma(x)$;

7) если $\Phi = \forall x\Psi$, то $\mathcal{M}^+ \models \Phi[\gamma] \iff$ для любой интерпретации $\gamma_1 : V_1 \longrightarrow |\mathcal{M}^+|$, для которой $x \in V_1$, $FV(\Phi) \subset V_1$, и для всех $x \in FV(\Phi)$ $\gamma_1(x) = \gamma(x)$ имеет иесто $\mathcal{M}^+ \models \Psi[\gamma_1]$;

8) если $\Phi = Ext_c(t)$, то $\mathcal{M}^+ \models \Phi[\gamma] \iff t[\gamma] \in \nu_c(Ext_c)$.

ОПРЕДЕЛЕНИЕ 7. Формулы Φ и Ψ сигнатуры σ^+ такие, что $FV(\Phi) = FV(\Psi)$, будем называть эквивалентными, если для любой модели $\mathcal{M}^+$ сигнатуры σ^+ и любой интерпретации $\gamma : V \longrightarrow |\mathcal{M}^+|$, для которой $FV(\Phi) \subseteq V$, $\mathcal{M}^+ \models \Phi[\gamma]$ тогда и только тогда, когда $\mathcal{M}^+ \models \Psi[\gamma]$.

В силу определения истинности нетрудно заметить, что имеют место все классические эквивалентности.

Будем говорить, что формула Φ сигнатуры σ^+ находится в пренексной нормальной форме, если она имеет вид $Q_1x_1, \ldots, Q_nx_n\Phi_1$, где $Q_i \in \{\forall, \exists\}$, а Φ_1 находится в дизъюнктивной нормальной форме, т.е. имеет вид $\bigvee\limits_{i=0}^{m} (\Phi_0^i \wedge \ldots \wedge \Phi_{m_i}^i)$, где Φ_j^i — атомарные формулы или отрицания атомарных, а $\Phi_0^i \wedge \ldots \wedge \Phi_{m_i}^i$ — обычные обобщенные обозначения дизъюнктивных членов Φ_1.

Очевидно, что для любой формулы Φ сигнатуры σ^+ существует формула Ψ сигнатуры σ^+, находящаяся в пренексной нормальной форме и эквивалентная Φ.

Будем говорить, что вхождение атомарной подформулы Φ в формулу Ψ, находящуюся в пренексной нормальной форме, *позитивно*, если в данном вхождении перед Φ не стоит знак отрицания $\neg$. Подформула Φ позитивна в формуле Ψ, если все вхождения Φ в пренексную нормальную форму Ψ позитивны.

Введем следующие обозначения. Будем обозначать формулу $\forall x(\neg Ext_c(x) \vee \Psi)$ через $\forall x \in Ext_c\Psi$, а формулу $\exists x(Ext_c(x) \wedge \Psi)$ через $\exists x \in Ext_c\Psi$. Кванторы $\forall x \in Ext_c$ и $\exists x \in Ext_c$ будем называть ограниченными.

ОПРЕДЕЛЕНИЕ 8. Формулу Φ сигнатуры σ^+ будем называть допустимой, если она удовлетворяет следующим условиям:

1) в запись формулы Φ сигнатуры σ^+ связанно входят только переменные из $\bigcup\limits_{c \in PClass \cup QClass} X(c)$;

2) в пренексной нормальной форме формулы Φ, имеющей вид $Q_1 x_1, \ldots, Q_n x_n \bigvee\limits_{i=0}^{m} (\Phi_0^i \wedge \ldots \wedge \Phi^i{}_{m_i} \wedge \Theta_0^i \wedge \ldots \wedge \Theta^i{}_{n_i})$, где Φ_j^i — атомарные формулы, а Θ_j^i — отрицания атомарных формул, для каждого i имеет место включение $\bigcup\limits_{j=0}^{n_i} FV(\Theta_j^i) \subseteq \bigcup\limits_{j=0}^{m_i} FV(\Phi_j^i)$;

3) для Φ существует эквивалентная ей формула Ψ такая, что все связанные вхождения переменных имеют ограниченные кванторы.

Пусть $Q \in QClass$, q — предметная переменная класса Q, Φ -допустимая формула, удовлетворяющая условиям:

1) q входит свободно в Φ;

2) нет связанных вхождений предметных переменных класса Q в Φ;

3) все вхождения в Φ подформул вида $Q(t)$ позитивны.

Выражение вида $Ext_Q(q)$ def Φ будем называть определением класса Q.

Набором параметров данного определения будем называть $FV(\Phi) \setminus \{q\}$. Формулу Φ будем называть правой частью данного определения.

ОПРЕДЕЛЕНИЕ 9. Графом зависимостей набора определений

$$\left.\begin{array}{l} Ext_{Q_1}(q_1) \ def \ \Phi_1, \\ \ldots \\ Ext_{Q_n}(q_n) \ def \ \Phi_n \end{array}\right\} \qquad (1)$$

(здесь и далее $Q_1, \ldots, Q_n$ попарно различны) назовем ориентированный граф $G(V, E)$ с множеством вершин $V = \{Q_1, \ldots, Q_n\}$ и множеством дуг $E = \{(Q_i, Q_j) | Ext_{Q_j}$ входит в запись $\Phi_i\}$.

ОПРЕДЕЛЕНИЕ 10. Если в графе зависимостей набора определений (1) существуют пути из Q_i в Q_j и из Q_j в Q_i, то определения $Ext_{Q_i}(q)\ def\ \Phi_i$ и $Ext_{Q_j}(q)\ def\ \Phi_j$ будем называть взаимно рекурсивными.

Если в E содержится (Q_i, Q_i), то определение класса Q_i будем называть рекурсивным.

ОПРЕДЕЛЕНИЕ 11. Набор определений (1) назовем схемой, если он удовлетворяет следующим условиям:

1) q_i не входит свободно в Φ_j для $i \neq j$;

2) если определения $Ext_{Q_i}(q)\ def\ \Phi_i$, $Ext_{Q_j}(q)\ def\ \Phi_j$ взаимно-рекурсивны, то

а) Φ_j не содержит связанных вхождений переменных из $X(Q_i)$,

б) все подформулы Φ_j вида $Ext_{Q_i}(t)$ входят в Φ_j позитивно;

3) $\bigcup\limits_{i=1}^{n} FV(\Phi_i) \setminus \{q_1, \ldots, q_n\} \subseteq X(TClass) \bigcup X(PClass)$.

Будем называть $\bigcup\limits_{i=1}^{n} FV(\Phi_i) \setminus \{q_1, \ldots, q_n\}$ набором предметных параметров схемы (1). Набор символов из $\{Ext_q \mid q \in QClass \setminus \{Q_1, \ldots, Q_n\}\}$ и входящих в правые части определений из схемы будем называть набором классовых параметров схемы; $\{Q_1, \ldots, Q_n\}$ будем называть сигнатурой схемы.

ОПРЕДЕЛЕНИЕ 12. Запросом сигнатуры σ^+ над моделью $\mathcal{M}^+$ будем называть тройку $\langle Sch, \xi, \Phi \rangle$, где Sch — схема с набором предметных параметров $\overline{v}$, ξ — интерпретация переменных в $|\mathcal{M}^+|$ с областью определения, содержащей $\overline{v}$, Φ — допустимая формула сигнатуры σ^+.

Денотационная семантика

ОПРЕДЕЛЕНИЕ 13. Обобщенной моделью сигнатуры σ^+ будем называть тройку $\mathcal{M}_b^+ = \langle (M_o)_{o \in Class^+}, \nu, EXT \rangle$.

Интерпретацией определяемых классов будем называть отображение μ из $\{Ext_q \mid q \in QClass\}$ в EXT, такое, что $\mu(Ext_q) \in EXT_q$.

Если ξ — интерпретация предметных переменных из множества свободных переменных формулы Φ сигнатуры σ^+, μ — интерпретация определяемых классов, $\mathcal{M}_b^+$ — обобщенная модель сигнатуры σ^+, то обозначим $\mathcal{M}_b^+ \models$ $\models \Phi[\xi, \mu] \Longleftrightarrow \mathcal{M}^+ \models \Phi[\xi]$, где $\mathcal{M}^+ = \langle (M_c)_{c \in Class^+}, \nu, EXT, \mu \rangle$.

Пусть Φ — допустимая формула сигнатуры σ^+, $FV(\Phi) = \{q\} \cup \overline{y}$, где $q \notin \overline{y}$, $type(q) = Q$. Пусть $Ext_{Q_1}, \ldots$ $\ldots, Ext_{Q_m}, Ext_{R_1}, \ldots, Ext_{R_k}$ — все символы из $\{Ext_c \mid c \in$ $\in QClass\}$, входящие в Φ, причем Φ не содержит связанных вхождений переменных классов $Q_1, \ldots, Q_m$ и $Ext_{Q_1}, \ldots$ $\ldots, Ext_{Q_m}$ входят позитивно в Φ.

Пусть $\mathcal{M}_1^+ = \langle (M_c)_{c \in Class^+}, \nu, EXT, \mu_1 \rangle$, $\mathcal{M}_2^+ =$ $= \langle (M_c)_{c \in Class^+}, \nu, EXT, \mu_2 \rangle$ — две модели сигнатуры σ^+, различающиеся лишь интерпретацией символов из $\{Ext_q \mid q \in QClass\}$, причем $\mu_1(Ext_{Q_i}) \subseteq \mu_2(Ext_{Q_i})$, $\mu_1(Ext_{R_i}) = \mu_2(Ext_{R_i})$.

Пусть ξ — интерпретация переменных из $FV(\Phi)$. Имеет место следующая

ЛЕММА 1. *Если* $\mathcal{M}_1^+ \models \Phi[\xi]$, *то* $\mathcal{M}_2^+ \models \Phi[\xi]$.

ДОКАЗАТЕЛЬСТВО. Для атомарной формулы Φ утверждение очевидно. Будем считать, что Φ находится в пренексной нормальной форме, т.е. $\Phi = K_1 x_1, \ldots, K_n x_n \Phi_1$, где $K_i \in \{\forall, \exists\}$, $\Phi_1 = \bigvee_{i=0}^{k} (\Phi_0^i \wedge \ldots \wedge \Phi_{m_i}^i \wedge \Psi_0^i \wedge \ldots \wedge \Psi_{n_i}^i)$, причем Φ_j^i — атомарная формула, Ψ_j^i — отрицание атомарной формулы.

Поскольку Ext_{Q_l} не содержится в Ψ_j^i, ясно, что для любой интерпретации ξ_1 переменных из $FV(\Phi_1)$ если $\mathcal{M}_1^+ \models \Phi_1[\xi_1]$, то $\mathcal{M}_2^+ \models \Phi_1[\xi_1]$.

Далее, из определения истинности формулы на модели следует, что если утверждение леммы верно для формулы Ψ, то оно также останется в силе и для формул $\exists\Psi$ и $\forall\Psi$. Таким образом лемма 1 доказана.

Обозначим через $f_{\xi,\mu}(\Phi)$ следующую функцию ($\mathcal{P}(M_Q)$ — множество подмножеств M_Q):

$$f_{\xi,\mu}(\Phi) : \mathcal{P}(M_{Q_1}) \times \ldots \times \mathcal{P}(M_{Q_m}) \longrightarrow \mathcal{P}(M_Q),$$

где $f_{\xi,\mu}(\Phi)(A_1, \ldots, A_m) = \{a \in M_Q \mid \mathcal{M}_b^+ \models \Phi[\xi_1, \mu_1]$ для некоторых ξ_1, μ_1, удовлетворяющих приведенным ниже условиям 1–4:

1) если $y \in \bar{y}$, то $\xi_1(y) = \xi(y)$,

2) $\xi_1(q) = a$,

3) для каждого $i \in \{1, \ldots, k\}$ $\mu_1(Ext_{R_i}) = \mu(Ext_{R_i})$,

4) для каждого $i \in \{1, \ldots, m\}$ $\mu_1(Ext_{Q_i}) = A_i$.

Как следствие леммы 1 получаем следующий результат.

ЛЕММА 2. *$f_{\xi,\mu}(\Phi)$ монотонна, т.е. если $A_1 \subseteq B_1, \ldots$ $\ldots, A_m \subseteq B_m$, то $f_{\xi,\mu}(\Phi)(A_1, \ldots, A_m) \subseteq f_{\xi,\mu}(\Phi)(B_1, \ldots, B_m)$.*

ОПРЕДЕЛЕНИЕ 14. Пусть

$$\left.\begin{array}{l} Ext_{Q_1}(q_1)\ def\ \Phi_1 \\ \ldots \\ Ext_{Q_n}(q_n)\ def\ \Phi_n, \end{array}\right\} \tag{2}$$

— схема Sch с наборами предметных параметров $\bar{v}$ и классовых параметров $\overline{V}$. Будем называть Sch простой, если:

1) переменные классов $Q_1, \ldots, Q_n$ не имеют связанных вхождений в $\Phi_1, \ldots, \Phi_n$;

2) каждая Φ_i не содержит негативных вхождений подформул вида $Ext_{Q_j}(t)$.

Пусть Sch — простая схема (2), ξ, μ — означивания параметров из $\bar{v}$ и $\overline{V}$ соответственно.

Определим оператор

$\Gamma_{\xi,\mu}(Sch) : \mathcal{P}(M_{Q_1}) \times \ldots \times \mathcal{P}(M_{Q_n}) \longrightarrow \mathcal{P}(M_{Q_1}) \times \ldots$ $\ldots \times \mathcal{P}(M_{Q_n})$,

$\Gamma_{\xi,\mu}(Sch) = \langle f_{\xi,\mu}(\Phi_1), \ldots, f_{\xi,\mu}(\Phi_n)\rangle$.

Очевидно $\Gamma_{\xi,\mu}(Sch)$ монотонный, а $\mathcal{P}(M_{Q_1}) \times \ldots$ $\ldots \times \mathcal{P}(M_{Q_n})$ — полная решетка.

Имеет место известная

ТЕОРЕМА 1 [3,4]. *Монотонное преобразование T на полной решетке $\langle V, \leq \rangle$ имеет непустое множество неподвижных* *точек. В частности, T имеет наименьшую неподвижную точку $lfp(T)$ такую, что $lfp(T) = inf(x \in V \mid T(x) \leq x)$.*

ОПРЕДЕЛЕНИЕ 15. Денотационной семантикой простой схемы Sch (2) назовем отображение $Den(Sch)$ из $\{Ext_{Q_1}, \ldots, Ext_{Q_n}\}$ в набор отношений $lfp(\Gamma_{\xi,\mu}(Sch))$, при этом Ext_{Q_i} ставится в соответствие i–я компонента $lfp(\Gamma_{\xi,\mu}(Sch))$.

Теперь пусть Sch — произвольная схема (2). Пусть ξ, μ — означивания параметров схемы. Определим расширеный граф зависимостей G^* схемы Sch. Граф G^* получается из графа зависимостей G схемы Sch пометкой знаком $*$ тех дуг (Q_i, Q_j), для которых либо существует связанное вхождение переменной класса Q_j в Φ_i, либо существует негативное вхождение подформулы вида $Ext_{Q_j}(t)$ в Φ_i.

Разбиением сигнатуры схемы Sch называется разбиение множества $\{Q_1, \ldots, Q_n\}$ на подмножества $\{\overline{Q}_1, \ldots, \overline{Q}_m\}$, удовлетворяющие следующим условиям:

а) если $P \in \overline{Q}_i$, $Q \in \overline{Q}_j$ и (P, Q) — дуга G^*, то $i \geq j$,

б) если $P \in \overline{Q}_i$, $Q \in \overline{Q}_j$ и (P, Q) — дуга G^*, помеченная $*$, то $i > j$.

ТЕОРЕМА 2. *Для любой схемы существует разбиение.*

Ниже приведем способ нахождения разбиения.

Алгоритм разбиения

INPUT: схема Sch, G^*.

OUTPUT: Разбиение для сигнатуры Sch.

METHOD:

Выполнить следующие шаги:

1. На основе графа G^* построить граф G^*_{tr} следующим образом. Для каждой пары вершин Q_i и Q_j в G^*, если существует путь из Q_i в Q_j, содержащий дуги, помеченные знаком $*$, добавить дугу (Q_i, Q_j), помеченную знаком $*$, в результирующий граф (если такая дуга еще не существует).

2. $i := 1$.

3. Определить множество K всех вершин графа G^*_{tr}, из которых не выходят дуги, помеченные $*$.

4. $\overline{Q}_i = K$.

5. Исключить все вершины множества K вместе с соответствующими им дугами из $G^*_{t_r}$.

6. Если в $G^*_{t_r}$ еще остались вершины, то $i := i + 1$ и перейти к шагу 3, иначе закончить.

ENDMETHOD.

Пусть Sch_i — подсхема Sch с сигнатурой $\overline{Q}_i$, множеством предметных параметров $\overline{v}$ и множеством классовых параметров $\overline{V} \bigcup \bigcup\limits_{j<i} \{Ext_q \mid q \in \overline{Q}_j\}$.

Имеет место

ТЕОРЕМА 3. Sch_i — *простая схема.*

ДОКАЗАТЕЛЬСТВО теоремы вытекает из определения разбиения и определения схемы.

ОПРЕДЕЛЕНИЕ 16. Денотационной семантикой схемы Sch назовем отображение $Den(Sch)$ множества $\{Ext_{Q_1}, \dots$ $\dots, Ext_{Q_n}\}$ такое, что если $P \in \overline{Q}_i$, то $Den(Sch)(P) =$ $= Den(Sch_i)(P)$.

ОПРЕДЕЛЕНИЕ 17. Модель $\mathcal{M}^+ = \langle (M_c)_{c \in Class^+}, \nu,$ $EXT, \mu\rangle$ сигнатуры σ^+ будем называть определяемой схемой Sch (2), если $\nu(Ext_{Q_i}) = Den(Sch)(Q_i)$.

ОПРЕДЕЛЕНИЕ 18. Денотационной семантикой запроса $\langle Sch, \xi, \Phi\rangle$ $(FV(\Phi) = \{x_1, \dots, x_m\})$ сигнатуры σ^+ над моделью $\mathcal{M}^+ = \langle (M_c)_{c \in Class^+}, \nu, EXT, \mu\rangle$, определяемой схемой Sch, при интерпретации предметных параметров схемы ξ будем называть $\{(a_1, \dots, a_m) \mid a_i \in M_{type(x_i)},$ $\mathcal{M}^+ \models \Phi[\xi_1]$, где $\xi_1(x_i) = a_i$ и на $\overline{v}$ значения ξ_1 и ξ совпадают$\}$.

З а к л ю ч е н и е

Перспективным направлением исследований, проводимых в рамках концепции семантического программирования и связанных с описанием семантики OQL, является повышение выразительной силы описанного в данной работе языка за счет введения в язык понятия подзапроса, т.е. запроса, который можно рассматривать в определенном контексте как терм типа множество объектов некоторого класса. Это позволит значительно упростить формализацию GROUP BY оператора и подзапросов языка OQL.

Литература

1. ГОНЧАРОВ С.С., СВИРИДЕНКО Д.И. Σ-программы и их семантики // Логические методы в программировании.— Новосибирск, 1987. — Вып. 120: Вычислительные системы. — С. 24–51.

2. ГОНЧАРОВ С.С., ЕРШОВ Ю.Л., СВИРИДЕНКО Д.И. Методологические аспекты семантического программирования // Научное знание: логика, понятия, структура. — Новосибирск: Наука, 1987. — С. 154–184.

3. ЧЕРИ С., ГОТЛОВ Г., ТАНКА Л. Логическое программирование и базы данных. — М.: Мир, 1992.

4. TARSKI A. A lattice theoretical fixpoint theorem and its applications // Pacific J. Math. — 1955. — № 5.

5. GOGOLLA M. An extended entity-relationship model. Fundamentals and pragmatics. — Springer–Verlag. Lecture Notes in Comput. Sci. — 1994. — Vol. 767.

6. GOGUEN J.A., MESEGUER J. Models and equality for logical programming. — Springer–Verlag. Lecture Notes in Comput. Sci., 1987. — Vol. 250. — P.1–22.

7. The Object Database Standard: ODMG–93 Release 1.2, editted by R.G.G. Cattel–Morgan Kaufmann Publishers, Inc., CA: 1995.

8. ЮРЧЕНКО О.Г. S-программы // Теория вычислений и языки спецификаций. — Новосибирск, 1995. — Вып. 152: Вычислительные системы. — С. 20–37.

9. URCHENKO O. Towards a semantic query language // Siberian Adv. Math. - Allerton Press Inc. — 1996. — Vol. 6, №3. — P. 81–95.

Поступила в редакцию
10 июня 1997 года