Optimizing Multiset Metagrammars. Formal Definitions

Igor Sheremet¹ and Igor Zhukov²

¹ Financial University under the Government of Russian Federation, Moscow, Russian Federation
² National Research Nuclear University Moscow Engineering Physical Institute, Moscow, Russian Federation

¹iasheremet@fa.ru, ²i.zhukov@inbox.ru

ABSTRACT
A multiset-based approach to optimization problems formalizing and solving is considered. Optimizing multiset metagrammars (OMMG) are described as specific knowledge representation model developed precisely for this purposes. OMMG provide convergence of classical optimization theory and modern knowledge engineering combining the best features of mathematical and logical programming. Paper is dedicated to the formal definition of OMMG syntax and semantics as well as to the multigrammatical representation of the most well-known classical optimization problems.

Keywords: Multisets, Multiset grammars and metagrammars, Optimizing multiset metagrammars, Multigrammatical representation of optimization problems, Unconventional programming paradigms, Knowledge representation, Mathematical programming, Logical programming.

1. INTRODUCTION
Multisets (MS) for many years are one of the most interesting and promising mathematical objects which application to various areas of modern computer science is investigated thoroughly [1-13]. One of the most useful directions of multisets processing theory and technologies development is rule-based multiset programming paradigm [14-16], which ideology is exploited widely in DNA and membrane computing [17-19] as well as in Gamma-like programming languages [10,20,21], and chemical programming, which is operating macrosets, or generalized multisets [22-25].

As for the state of the art practical applications, the most valuable are constraint multiset grammars (CMSG) developed mainly for visual objects (first of all handwritten symbols) recognition [26-28]. CSMG inherit key features of constraints logical programming [29-31], and are convenient tool for the description of planar objects with complicated recursive structure.

Author’s contribution to the area considered are so called optimizing multiset metagrammars (OMMG) [32, 33] developed for problem solving in system analysis and optimization area (first of all, analysis and construction of the large-scale sociotechnical systems like business structures, manufacturing facilities, macroeconomic objects etc.) From the practical point of view OMMG are specific knowledge representation model providing convergence of classical optimization theory and modern knowledge engineering. OMMG integrate basic features of mathematical programming (optimal solutions search in solutions space which is strictly defined by goal functions and constraints) and logical programming (natural and easily modified top-down recursive representation of complex objects and processed as well as systems operation logic).

Presented paper is dedicated to further development of OMMG pragmatics by constructing multigrammatical representation (formulation) of classical optimization problems.

The first part of the paper contains formal definitions of OMMG family and its various subsets, while main content of the second part is mentioned representations description and investigation in order to establish some standard techniques useful for practical problems solving and OMMG application hardware implementation.

2. BASIC NOTIONS AND DEFINITIONS
Theory of multisets background is notion of multiset which is understood mainly as set of so called multiojbjects composed of indistinguishable elements (objects). Multiojbject containing \( n \) objects of type \( a \) is denoted as \( n \cdot a \), where \( n \) is called multiplicity of object \( a \) (\( \cdot \) is composing symbol). Record

\[ \nu = \{n_1 \cdot a_1, \ldots, n_m \cdot a_m\} \tag{1} \]

means that multiset \( \nu \) contains \( n_1 \) objects of type \( a_1 \), \ldots, \( n_m \) objects of type \( a_m \). We shall use one and the same symbol \( \in \) for denotation of object and multiojbject entering multiset \( \nu \), so \( a_i \in \nu \) and \( n_i \cdot a_i \in \nu \) means that object \( a_i \)
enters \( v \) as well as multiobject \( n_1 \cdot a_i \). From the substantial point of view set \( \{a_1, ..., a_m\} \) and multiset \( \{1 \cdot a_1, ..., 1 \cdot a_m\} \) are equivalent as well as object \( a \) and multiobject \( 1 \cdot a \). Empty multiset and empty set are denoted as \( \{\} \). If object \( a \) multiplicity is zero, it is equivalent to the absence of \( a \) in the multiset \( v \), what is recorded, as usual, \( a \not\in v \).

We shall say that multiset
\[
\begin{align*}
v &= \{n_1 \cdot a_1, ..., n_m \cdot a_m\} \\
v' &= \{n'_1 \cdot a'_1, ..., n'_m \cdot a'_m\}
\end{align*}
\]
is included to multiset
\[
\begin{align*}
v' &= \{n'_1 \cdot a'_1, ..., n'_m \cdot a'_m\}
\end{align*}
\]
what is denoted as \( v \subseteq v' \), if
\[
\left( \forall \ n \cdot a \in v \right) \left( \exists n' \cdot a \in v' \right) \ n \leq n',
\]
and \( v \) is strictly included to \( v' \), what is denoted \( v \varsubsetneq v' \), when \( v \subseteq v' \) and \( v \neq v' \).

We shall use small "a" with lower indices as well as small "a", "b" etc. for objects denotation; multisets will be denoted by small "v" with or without indices.

**Example 1.** Multiset \( v = \{3 \cdot a, 4 \cdot b, 1 \cdot c\} \) is strictly included to multiset \( v' = \{4 \cdot a, 4 \cdot b, 1 \cdot c, 1 \cdot d\} \).

There are five basic operations on multisets: addition, subtraction, multiplication by integer number, join and intersection.

Result of \( v \) and \( v' \) multisets addition denoted as bold symbol "+" is multiset
\[
\begin{align*}
v + v' &= (\bigcup_{a \in \{a_1, ..., a_m\}} \{n \cdot a\}) \cup \bigcup_{a \in \{a'_1, ..., a'_m\}} \{n' \cdot a\}
\end{align*}
\]
result of \( v \) subtraction from \( v' \) denoted as bold symbol "-" is defined as follows:
\[
\begin{align*}
v - v' &= \bigcup_{a \in \{a_1, ..., a_m\}} \{n \cdot a\} \cup \bigcup_{a \in \{a'_1, ..., a'_m\}} \{n' \cdot a\}
\end{align*}
\]
result of multiplicity multiplication by integer number, what is denoted as \( v \ast n \), is multiset:
\[
\begin{align*}
v \ast n &= (\bigcup_{a \in \{a_1, ..., a_m\}} \{n \cdot a\}) \cup \bigcup_{a \in \{a'_1, ..., a'_m\}} \{n' \cdot a\}
\end{align*}
\]
result of join and intersection denoted as bold symbol "\( \cup \)" and "\( \cap \)" is multiset:
\[
\begin{align*}
v \cup v' &= (\bigcup_{a \in \{a_1, ..., a_m\}} \{n \cdot a\}) \cup (\bigcup_{a \in \{a'_1, ..., a'_m\}} \{n' \cdot a\})
\end{align*}
\]
result of multiplication by integer number, what is denoted as \( v \cdot n \), is multiset:
\[
\begin{align*}
v \cdot n &= (\bigcup_{a \in \{a_1, ..., a_m\}} \{n \cdot a\}) \cup (\bigcup_{a \in \{a'_1, ..., a'_m\}} \{n' \cdot a\})
\end{align*}
\]
Multisets may be used for generating another multisets by means of multiset grammars. By analogy with classical grammars operating on strings of symbols \([34,35]\), we shall define multigrammar as a couple 
\[ S = \langle \nu_0, R \rangle, \]
where \(\nu_0\) is a multiset called kernel, while \(R\) called scheme is a finite set of so-called rules which are used for generation of new multisets from already generated. Rule has form
\[ v \rightarrow v', \]
where \(v\) and \(v'\), called left and right parts of the rule correspondingly, are multisets, and \(v \neq \{\emptyset\}\). Semantics of rule is defined as follows. Let \(\overline{v}\) be a multiset. We shall say that rule (11) is applicable to \(\overline{v}\), if
\[ v \subseteq \overline{v}. \]
Then result of application is a multiset
\[ \overline{v}' = \overline{v} \cdot v + v', \]
(13)
and is strictly defined as follows:
\[ v_0 \quad \overset{R}{\Rightarrow} \quad v, \]
Set of multisets generated by MG \(S = \langle \nu_0, R \rangle\) is denoted as
\[ V_S = \{v \mid v_0 \overset{R}{\Rightarrow} v\}. \]
(17)
MS \(v\) is called terminal multiset (TMS) if there is no one rule \(r \in R\) which is applicable to \(v\). Set of terminal sets (STMS) will be denoted \(V_T\). Iterative representation of MG semantics, i.e. SMS \(V_S\) generated by MG \(S = \langle \nu_0, R \rangle\), is following:
\[ V_S(0) = \{v_0\}, \]
(18)
\[ V_S(i+1) = V_S(i) \cup \left( \bigcup_{r \in R} \bigcup_{\overline{v} \in V_S(i)} \pi(\overline{v}, r) \right), \]
(19)
\[ V_S = V_S(\infty), \]
(20)
where
\[ \pi(\overline{v}, v \rightarrow v') = \begin{cases} \{\overline{v} \cdot v + v'\}, & \text{if } v \subseteq \overline{v}, \\ \{\emptyset\}, & \text{otherwise}. \end{cases} \]
(21)
As seen, (21) implements application of rule \(v \rightarrow v'\) as defined by (12) – (13).
SMS \(V_S\) is fixed point of the described process, i.e. \(V_S = V_S(i)\) with \(i \rightarrow \infty\). If for some finite \(i\) \(V_S(i) = V_S(i+1)\), then \(V_S = V_S(i)\), and \(V_S\) is also finite. For SMS \(v\)
\[ \pi(v, r) = \{\emptyset\} \]
(22)
for all rules \(r \in R\), i.e. no one multiset may be generated from TMS, and
\[ \overline{V}_S = \{v \mid v \in V_S \& (\forall r \in R) \pi(v, r) = \{\emptyset\}\}. \]

By analogy with classical string-generating grammars, multiset grammars may be general, or context-sensitive (CS), and context-free (CF). In the last, left parts of all rules \(r \in R\) are multisets of a form \(\{1 \cdot a\}\) while rules in CS MG contain arbitrary left and right parts (the only limitation already mentioned is that left part must be non-empty).

**Example 3.** Consider context-free multigrammar
\[ S = \{1 \cdot a, 2 \cdot b\}, \{r_1, r_2, r_3, r_4\} >, \]
where
\[ r_1 = \{1 \cdot a\} \rightarrow \{3 \cdot b, 2 \cdot c\}, \]
\[ r_2 = \{1 \cdot a\} \rightarrow \{1 \cdot b, 3 \cdot c\}, \]
\[ r_3 = \{1 \cdot b\} \rightarrow \{2 \cdot c\}, \]
\[ r_4 = \{1 \cdot b\} \rightarrow \{1 \cdot c\}, \]
According to (18) – (19), there are following generation chains:
\[ \{1 \cdot a, 2 \cdot b\} \rightarrow [5 \cdot b, 2 \cdot c] \rightarrow [7 \cdot c], \]
\[ \{1 \cdot a, 2 \cdot b\} \rightarrow [3 \cdot b, 3 \cdot c] \rightarrow [9 \cdot c], \]
\[ \{1 \cdot a, 2 \cdot b\} \rightarrow [3 \cdot b, 3 \cdot c] \rightarrow [6 \cdot c], \]
and
\[ V_S = \{1 \cdot b, 2 \cdot c\}, \{5 \cdot b, 2 \cdot c\}, \{3 \cdot b, 3 \cdot c\}, \{6 \cdot c\}, \{7 \cdot c\}, \{9 \cdot c\}, \{12 \cdot c\}\].

**3. UNITARY MULTIGRAMMARS**

One practice-oriented and useful modification of CF multigrammars is called unitary multigrammars (UMG) having slightly different syntax and semantics of basic constructions.

UMG is couple \(S = \langle a_0, R \rangle\), where \(a_0\) is called title object, and \(R\) is, as higher, scheme, being here set of so-called unitary rules (UR) having form
\[ a \rightarrow n_1 \cdot a_1, ..., n_m \cdot a_m, \]
(24)
where \(a\) is called head while unordered list of multiobjects \(n_1 \cdot a_1, ..., n_m \cdot a_m\) is called body of the unitary rule ("unordered" means all permutations of \(n_1 \cdot a_1, ..., n_m \cdot a_m\) represent one and the same body, or \(n_1 \cdot a_1, ..., n_m \cdot a_m\) is interpreted as multiset \(\{n_1 \cdot a_1, ..., n_m \cdot a_m\}\)). Application of unitary rule (24) to multiset \(\overline{v}\) containing multiobject \(n \cdot a\) is, as higher, called generation step, and is defined by the following equality not contradicting to (12) – (13) and (21):
\[ \vec{v}' = \vec{v} \cdot \{n \cdot a\} \ast n \cdot \{n_1 \cdot a_1, ..., n_m \cdot a_m\}, \]  
(25)
what corresponds to \(n\) times application of CF rule \(\{1 \cdot a\} \rightarrow \{n_1 \cdot a_1, ..., n_m \cdot a_m\}\).  
(26)
However, record is not the only difference between context-free and unitary multigrams. Being applied to practical problems, unitary rules are most frequently interpreted as descriptions of complex objects through their elements (which, in turn, may be decomposed, and so on until non-divided entities); \(24\) means that object (of type \(a\)) consists of \(n_1\) objects (of type) \(a_1, ..., n_m\) objects (of type) \(a_m\). So if there are \(m > 1\) alternative ways of object \(a\) structuring – for example, along with \(24\) there is one more UR with head \(a\) –
\[ a \rightarrow n_1' \cdot a_1', ..., n_m' \cdot a_m', \]  
(27)
and we apply unitary rule \(24\) once in the generation chain, then for all following generation steps, if they occur, we shall use the same UR \(24\). For short, object \(a\) is postulated constant within all generation chain.  

As seen, UMG semantics is based on “variable-like” interpretation of objects, which differ by this feature from non-terminals in classical CF string-generating grammars, which semantics allows arbitrary alternatives of any non-terminal substitution at every generation step where this non-terminal may be replaced.  

Iterative representation of UMG \(S = a_0, R >\) semantics is following:
\[ X_{(i)} = \{\{1 \cdot a_0\}, R\} >, \]  
(28)
\[ X_{(i+1)} = \bigcup_{\vec{v}, R > \in X_{(i)}} \{\vec{v}, R, r\} \in R, \]  
(29)
\[ V_S = \{\vec{v} | \vec{v} < \vec{v}, R > \in X_{(0)}\}, \]  
(30)
where
\[ \vec{v}, \vec{R}, r = \{\vec{v} \ast \{n \cdot a\} \ast n \ast \{n_1 \cdot a_1, ..., n_m \cdot a_m, \{\{0\}\} \text{ otherwise,} \]  
(31)
\[ alt(a \rightarrow w) = \{a \rightarrow w' | a \rightarrow w' \in R \ast w' \neq w\}. \]  
(32)
where \(w\) and \(w'\) designate \(n_1 \cdot a_1, ..., n_m \cdot a_m\) and \(n_1' \cdot a_1', ..., n_m' \cdot a_m'\) correspondingly, and \(r\) is unitary rule \(a \rightarrow n_1 \cdot a_1, ..., n_m \cdot a_m\). Inequality \(w \neq w'\) in \(32\) is interpreted as if \(w\) and \(w'\) were multisets.

In \(28 - 30\) \(X_{(i)}\) is set of couples of the form \(\vec{v}, \vec{R}\), where \(\vec{v}\) is multiset created by already executed generation steps, and \(\vec{R}\) is subset of \(R\) scheme including all unitary rules which may be applied to \(\vec{v}\) at the next such step. As seen, difference between \(21\) and \(31\) is that in the last case \(a \rightarrow w\) application to \(\vec{v}\) implemented by \(\vec{v}\) results not only in substitution of \(n\) objects \(a\) having place in \(\vec{v}\) by \(n \times n_1\) objects \(a_1, ..., n \times n_m\) objects \(a_m\), but also in elimination from \(\vec{R}\), which is current \(\vec{R}\) subset of applicable unitary rules, all URs \(a \rightarrow w'\) with the same head \(a\) and alternative bodies \(w' \neq w\). This prevents from applying \(a \rightarrow w'\) in future generation steps, that strictly corresponds to URs semantics verbally described higher; so object \(a\) will be “detailed” by the same only UR \(a \rightarrow w\), as at the executed step. As seen, result of function \(alt\) application is set of all URs with the same head \(a\) and bodies different from \(w\) having place in the argument \(a \rightarrow w\).

Lower we shall designate by \(\beta(v)\) set of all objects having place in multiset \(v\) (it is called multiset basis \([11, 12]\)):
\[ \beta(v) = \{a \mid a \in v\}. \]  
(33)
By \(A_2\) we shall designate set of all objects having place in UMG \(S = a_0, R >\), while by \(A_2\) – set of all terminal objects having place in \(S\), i.e. objects which are not heads of unitary rules \(r \in R\) and present only in URs bodies. As seen,
\[ A_2 \subseteq A_2. \]  
(34)

Multiset \(v \in V_2\) such that all objects of \(v\) are terminal is, as higher, called terminal multiset; this notion fully corresponds to the defined one in relation to multigrams. Formally,
\[ V_2 = \{v | v \in V_2 \ast \beta(v) \subseteq A_2\}. \]  
(35)
Let us present an example illustrating main semantical feature of unitary multigrams in comparison with context-free multigrams.

**Example 4.** Consider UMG \(S = a_0, R >\), where \(R\) contains following unitary rules (recorded for unambiguity in angle brackets):
\[ r_1 = (a \rightarrow 2 \cdot b, 3 \cdot c), \]  
\[ r_2 = (a \rightarrow 3 \cdot b, 2 \cdot c), \]  
\[ r_3 = (b \rightarrow 1 \cdot d), \]  
\[ r_4 = (b \rightarrow 2 \cdot d), \]  
\[ r_5 = (c \rightarrow 1 \cdot b, 2 \cdot d), \]  
\[ r_6 = (c \rightarrow 3 \cdot b, 1 \cdot d). \]

There are following generation chains:
\[ a \Rightarrow \{2 \cdot b, 3 \cdot c\} \Rightarrow \{3 \cdot c, 2 \cdot d\} \Rightarrow \{3 \cdot b, 6 \cdot d\} \Rightarrow \{9 \cdot d\}, \]
\[ a \Rightarrow \{2 \cdot b, 3 \cdot c\} \Rightarrow \{3 \cdot c, 2 \cdot d\} \Rightarrow \{9 \cdot b, 3 \cdot d\} \Rightarrow \{12 \cdot d\}, \]
\[ a \Rightarrow \{2 \cdot b, 3 \cdot c\} \Rightarrow \{3 \cdot c, 4 \cdot d\} \Rightarrow \{3 \cdot b, 10 \cdot d\} \Rightarrow \{16 \cdot d\}, \]
\[ a \Rightarrow \{2 \cdot b, 3 \cdot c\} \Rightarrow \{3 \cdot c, 4 \cdot d\} \Rightarrow \{9 \cdot b, 7 \cdot d\} \Rightarrow \{25 \cdot d\}, \]  
etc. (generation chains beginning from \(r_2\) unitary rule may be constructed by the concerned reader independently). As seen, there is no one generation chain which contains \(r_3\) and \(r_4\) simultaneously: application of \(r_3\) at the second generation step excludes application of \(r_4\) at the fourth one, and vice versa.

**4. FILTERS AND FILTERING UNITARY MULTIGRAMMARS**

Filters are constructions which provide selection of subsets of sets of generated TMS.

One may consider filters as query language with special features used for selection of solutions from solutions space.

Filters consist from atomary constructions called conditions. There are two types of conditions: boundary
and optimizing. The first, in turn, may be elementary and chain.

Elementary boundary condition (EBC) may have form $apn$, $npa$ and $apa'$, where $a$ and $a'$ are objects, $n$ is integer number, and $\rho \in \{<,=,\leq\}$ is symbol of relation.

EBC semantics is following.

Let $V$ be set of multisets. Multiset $v \in V$ satisfies EBC $\bar{npa}$, if $n \cdot a \in v$, and $\bar{npn}$ is true. Similarly, $v \in V$ satisfies EBC $apn$, if $npn$ is true. And, at last, $v \in V$ satisfies EBC $apa'$, if $\bar{a} \cdot a \in v$, $\bar{a} \cdot a' \in v$ and $\bar{npn}$ is true.

**Example 5.** Let $V = \{v_1, v_2, v_3\}$, where

\[
\begin{align*}
v_1 &= \{3 \cdot a, 2 \cdot b, 4 \cdot c\}, \\
v_2 &= \{1 \cdot a, 3 \cdot b\}, \\
v_3 &= \{1 \cdot b, 5 \cdot c\}.
\end{align*}
\]

Then result of application of EBC $a < 2$ to $v$ is set $\{v_2, v_3\}$ (note, that $v_3$ is included to this set because, as said higher, absence of object in the multiset is equivalent to zero multiplicity of this object). If EBC is $b > 3$, then result is empty set, if $c = 5$, then result is $\{v_3\}$; if $b > c$, then result is $\{v_2\}$.

Optimizing condition has form $a = opt$, where $a$ is object, and $opt \in \{min, max\}$.

Multiset $v \in V$ satisfies optimizing condition $a = min$ if for every $v' \in V$, such that $v' \neq v$, multiplicity $n$ in multiset $n \cdot a \in v$ is not less than multiplicity $n'$ in multiset $n' \cdot a \in v'$, i.e. $n \geq n'$. As higher, case $a \not\in v$ ($a' \not\in v$) is equivalent to $0 \cdot a \in v (0 \cdot a' \in v')$.

**Example 6.** Let $V$ be the same as in the previous example, and optimizing condition is $a = max$. Then result of its application to $v$ is set $\{v_1\}$. If optimizing condition is $b = min$, then result is $\{v_3\}$.

Filter $F$ is set of conditions and may be represented as join of boundary $F_\varnothing$ and optimizing $F_{opt}$ subfilters:

\[
F = F_\varnothing \cup F_{opt},
\]

where $F_\varnothing$ is set of EBC and $F_{opt}$ is set of optimizing conditions.

Result of filtration of MS $V$ by filter $F$ is denoted $V \downarrow F$ and is defined as follows:

\[
V \downarrow F = (V \downarrow F_\varnothing) \downarrow F_{opt},
\]

i.e. $V$ is filtered by boundary subfilter, and obtained result is filtered by optimizing subfilter:

\[
F_\varnothing = \{bc_1, ..., bc_k\}, \quad F_{opt} = \{opt_1, ..., opt_l\},
\]

\[
v \downarrow F_\varnothing = \bigcap_{i=1}^k (v \downarrow \{bc_i\}) = v',
\]

\[
v' \downarrow F_{opt} = \bigcap_{j=1}^l (v' \downarrow \{opt_j\}).
\]

**Example 7.** Let $V$ is the same as in two previous examples, and filter $F = \{b > 1, b \leq 3, c = min\}$. Then

\[
V \downarrow F = (V \downarrow F_\varnothing) \downarrow F_{opt} = (V \downarrow \{b > 1, b \leq 3\}) \downarrow \{c = min\} = ([v_1, v_2] \cap [v_1, v_2, v_3]) \downarrow \{c = min\} = [v_1, v_2] \downarrow \{c = min\} = [v_2],
\]

because $c \not\in v_2$, so object $c$ multiplicity is zero.

Note that due to commutativity of set-theoretical join and intersection operations filtration inside subfilters may be executed in arbitrary order.

Let us now define syntax and semantics of so called filtering multigrams (FMG) and filtering unitary multigrams (FUMG).

FMG $S = \langle v_0, R, F >$, where $v_0$ and $R$ have the same sense as in MG, and $F$ is filter, defines set of terminal multisets in the following way:

\[
V_s = V_{s'} \downarrow F,
\]

where $S' = \langle v_0, R >$, i.e. STMS, generated by MG $S'$, is filtered by $F$, and result is defined as generated by FMG $S$.

Similarly, FUMG $S = \langle a_0, R, F >$, where $a_0$ and $R$ have the same sense as in UMG, and $F$ is filter, defines set of terminal multisets as follows:

\[
V_s = V_{s'} \downarrow F,
\]

where $S' = \langle a_0, R >$.

**Example 8.** Let $S = \langle a, R, F >$, where $R = \{r_1, r_2, r_3, r_4\}$, and

\[
\begin{align*}
r_1 &= \langle a \rightarrow 3 \cdot b, 2 \cdot c >, \\
r_2 &= \langle a \rightarrow 2 \cdot b, 5 \cdot c >, \\
r_3 &= \langle b \rightarrow 2 \cdot c, 1 \cdot d >, \\
r_4 &= \langle b \rightarrow 1 \cdot c, 3 \cdot d >,
\end{align*}
\]

and $F = \{c \leq 8, d > 3, c = max, d = min\}$. Then $S' = \langle a, R >,$

\[
\begin{align*}
V_{s'} &= \{8 \cdot c, 3 \cdot d, \{5 \cdot c, 9 \cdot d\}, \{9 \cdot c, 2 \cdot d\}, \{7 \cdot c, 6 \cdot d\}\}, \\
V_s &= V_{s'} \downarrow F = (V_{s'} \downarrow \{c \leq 8, d > 2\}) \downarrow \{c = max, d = min\} = \\
&= \{8 \cdot c, 3 \cdot d\}. 
\end{align*}
\]

Type equation here.

There may be more sophisticated boundary subfilters containing not only elementary but also so called chain boundary condition (CBC). Each CBC is constructed from EBC by writing them sequentially:

\[
e_1 \rho_1 e_2 \rho_2 e_3 ... e_i \rho_i e_{i+1} ... e_m \rho_m e_{m+1},
\]

where $e_1, ..., e_{m+1}$ are objects or non-negative integers while $\rho_1, ..., \rho_m$ are symbols of relations ($<, =, \leq$).

CBC semantics is defined as follows. CBC (44) is replaced by boundary filter

\[
\{e_1 \rho_1 e_2, e_2 \rho_2 e_3, ..., e_i \rho_i e_{i+1}, ..., e_m \rho_m e_{m+1}\}.
\]
5. UNITARY MULTISET METAGRAMMARS

Unitary multigrams are useful for the description of objects which structure variants number is relatively little, and this variety may be simply reflected by explicit representation of alternative structures as unitary rules with the same head and different bodies. But in practically interesting cases this number may be such that explicit representation of sets of alternatives for all objects, which structure if a priori unknown, is really impossible. So there is necessary such UMG extension which would allow compact description of arbitrary number of alternative variants of object structure. This extension is called unitary multiset metagrammars (multimetagrammars), or, for short, UMMG.

Unitary multiset metagrammar \( S \) is triple \(< a_0, R, F >\), where \( a_0 \) and \( F \) are, as higher, title object and filter correspondingly, while \( R \) is scheme containing unitary rules and so called unitary metarules (UMR).

Unitary metarule has form

\[
a \rightarrow \mu_1 \cdot a_1, ..., \mu_m \cdot a_m.
\]

(46)

where \( \mu_i \) is positive integer number, as in sections 2–3, or variable \( \gamma \in \Gamma \), where \( \Gamma \) is universum of variables. When \( \mu_i \) is variable then it is called multiplicity-variable (MV). As seen, unitary rule is partial case of unitary metarule, which multiplicities \( \mu_1, ..., \mu_m \) having place in (46) are constants. As in (24), object \( a \) is called head, while unordered (in the same sense as in UR definition) list \( \mu_1 \cdot a_1, ..., \mu_m \cdot a_m \) – body of UMR.

Filter \( F \) of UMMG \( S = < a_0, R, F > \) is set of conditions, which may be boundary and optimizing as in UMG, i.e. having form \( a_0 \cdot n_pa_0pa' \) and \( \gamma = a \cdot \alpha t \), where \( \rho \in \{<, =, \leq\} \), \( \alpha t \in \{\min, \max\} \). At the same time \( F \) contains chain boundary conditions of the form

\[
\gamma \leq n \leq n'.
\]

(47)

which contain variables not objects \((n \geq 0)\); there is one and only one CBC (47) for every variable \( \gamma \) having place in unitary metarules entering scheme \( R \). This CBC is called variable declaration and has quite evident semantics.

If \( F \) includes subfilter \( F_T = \{n_1 \leq \gamma_1 \leq n_1', ..., n_l \leq \gamma_l \leq n_l'\} \) containing all variables declarations, then every tuple \((\overline{n}_1, ..., \overline{n}_l)\) such that \( \overline{n}_i \in [n_i, n_i'] \) for \( i = 1, ..., l \) provides creation of one unitary multigrammar by substitution of \( \overline{n}_i \) to all unitary metarules having place in \( R \) scheme instead of multiplicities-variables being in their bodies. Unitary rules entering \( R \) are transferred to new scheme denoted \( R \circ \langle \overline{n}_1, ..., \overline{n}_l \rangle \) without any transformations. Every such UMG generates set of terminal multisets by the help of filter \( F = F - F_T \), which contains all “usual” (without variables) conditions (both boundary and optimizing). As may be seen from this short informal description, UMMG \( S = \langle a_0, R, F \rangle \) is simple unified tool for compact representation of set of FUMG denoted \( S^* \): for practically valuable problems number of such FUMG, i.e. \( S^* \) cardinality, may reach extremely large numbers. It is easy to see that this number is not greater than product of all variables domains cardinalities:

\[
|S^*| \leq \prod_{j=1}^l (n_j' - n_j + 1)
\]

(48)

(if there are no identical UMG after variables values substitution then (48) becomes equality).

Let us now give strict definition of unitary multiset metagrammars semantics. (From the described it is obvious nature of “multiset metagrammar” notion – as “metalanguage” is used for description of another language, so “multiset metagrammar” is “unitary-like” multigrammar used for description of other UMG by means of unitary metarules, variables-multiplicities as well as by boundary conditions defining their domains).

UMMG \( S = \langle a_0, R, F \rangle \) defines set of terminal multisets \( \mathcal{V}_S \) in such a way:

\[
\mathcal{V}_S = \left( \bigcup_{S \in S^*} \mathcal{V}_S \right) \downarrow \mathcal{F},
\]

(49)

\[
S^* = \bigcup_{\overline{n}_i \in [n_i, n_i']} \bigcup_{\overline{n}_j \in [n_j, n_j'] \circ \langle \overline{n}_1, ..., \overline{n}_l \rangle} \{\langle a_0, R \}
\]

(50)
R \circ \langle \vec{n}_1, ..., \vec{n}_l \rangle = \{ r \circ \langle \vec{n}_1, ..., \vec{n}_l \rangle | r \in R \}.

(51)

\mathcal{F} = F - F_\Gamma.

(52)

F_\Gamma = \bigcup_{j=1}^{l} \{ \vec{n}_j | y_j' \leq y_j \}.

(53)

and, if r is a \leftarrow \mu_1 \cdot a_1, ..., \mu_m \cdot a_m, then \ r \circ \langle \vec{n}_1, ..., \vec{n}_l \rangle is unitary rule

a \leftarrow (\mu_1 \circ \langle \vec{n}_1, ..., \vec{n}_l \rangle) \cdot a_1, ..., (\mu_m \circ \langle \vec{n}_1, ..., \vec{n}_l \rangle) \cdot a_m.

(54)

where

\mu_i \circ \langle \vec{n}_1, ..., \vec{n}_l \rangle = \begin{cases} \mu_i, & \text{if } \mu_i \in [1, \infty], \\ \vec{n}_i, & \text{if } \mu_i \in \Gamma. \end{cases}

(55)

As seen, according to (54) – (55), all multiplicities-variables of UMR a \leftarrow \mu_1 \cdot a_1, ..., \mu_m \cdot a_m are substituted by their corresponding values from tuple \langle \vec{n}_1, ..., \vec{n}_l \rangle while all multiplicities-constants remain unchanged. Evidently, if all \mu_1, ..., \mu_m are constants, i.e. UMR is UR, it remains unchanged.

Note, that multiplicities-variables area of actuality is whole \ R scheme, i.e. if there are \ n > 1 \ occurrences of one and the same variable \ y \ in different unitary metarules, they all are substituted by one and the same value from the applied tuple \langle \vec{n}_1, ..., \vec{n}_l \rangle.

Example 10. Consider UMMG \ S = \langle a, R, F \rangle, where \ R contains following three unitary metarules:

r_1 = (a \rightarrow 3 \cdot b, y_1 \cdot c),

r_2 = (b \rightarrow 1 \cdot c, y_2 \cdot d),

r_3 = (b \rightarrow y_1 \cdot c, y_2 \cdot d),

while filter \ F contains following conditions:

2 \leq c \leq 4,

1 \leq y_1 \leq 2,

0 \leq y_2 \leq 1,

d = max.

According to (49) – (55), \ S defines four UMG:

S_1 = \langle a, R \circ (1,0) \rangle,

S_2 = \langle a, R \circ (1,1) \rangle,

S_3 = \langle a, R \circ (2,0) \rangle,

S_4 = \langle a, R \circ (2,1) \rangle,

which correspond to \ 4 = 2 \times 2 \ sets of two-values domains of \ y_1 \ and \ y_2 \ variables.

As seen,

R_1 = R \circ (1,0) = \{ (a \rightarrow 3 \cdot b, 1 \cdot c), (b \rightarrow 1 \cdot c) \},

\mathcal{V}_{T_1} = \{4 \cdot c\},

R_2 = R \circ (1,1) = \{ (a \rightarrow 3 \cdot b, 1 \cdot c), (b \rightarrow 1 \cdot c, 1 \cdot d) \},

\mathcal{V}_{T_2} = \{4 \cdot c, 3 \cdot d\},

R_3 = R \circ (2,0) = \{ (a \rightarrow 3 \cdot b, 2 \cdot c), (b \rightarrow 1 \cdot c) \},

\mathcal{V}_{T_3} = \{5 \cdot c\},

R_4 = R \circ (2,1)

= \{ (a \rightarrow 3 \cdot b, 2 \cdot c), (b \rightarrow 1 \cdot c, 1 \cdot d) \},

\mathcal{V}_{T_4} = \{5 \cdot c, 3 \cdot d\},

\{8 \cdot c, 3 \cdot d\}.

(56)

(57)

(58)

(note, that after variables values substitution there may occur identical URs as in \ R_1, R_2 and \ R_3 \ cases, that’s why number of URs in constructed UMG may be less than number of UMRs in initial UMMG). After that,

\mathcal{V}_e = \langle \mathcal{V}_{T_1} \cup \mathcal{V}_{T_2} \cup \mathcal{V}_{T_3} \cup \mathcal{V}_{T_4} \rangle

\downarrow \{ 2 \leq c \leq 4, d = max \} =

\{ \{4 \cdot c\}, \{4 \cdot c, 3 \cdot d\}, \{5 \cdot c\}, \{5 \cdot c, 3 \cdot d\}, \{8 \cdot c, 3 \cdot d\} \} \downarrow \{ 2 \leq c \leq d \})

\downarrow \{ d = max \} = \{ \{4 \cdot c\}, \{4 \cdot c, 3 \cdot d\} \}

\downarrow \{ d = max \} = \{ \{4 \cdot c\}, \{4 \cdot c, 3 \cdot d\} \}.

As seen, UMMG filters contain boundary conditions concerning objects and variables but optimizing conditions concerning only objects. That’s why from both theoretical and practical points of view it is reasonable to extend mentioned filters by optimizing conditions relating variables and having form

\ y = opt.

(56)

This form defines generated TMS optimality through multiplicities-variables values used while these terminal multisets generation.

Semantics of optimizing condition (56) is quite clear: select those TMS which are generated by the help of \ y \ value which is optimal (minimal, maximal) among all other TMS generated by the help of all other values from \ y \ domain. By this verbal description we extend TMS optimality notion from multiplicities-constants to also multiplicities-variables having place in unitary metarules applied while TMS generation.

Formal definition of the sense of (56) verbally described higher is as follows.

Let us introduce \ l \ auxiliary terminal objects \ \vec{y}_1, ..., \vec{y}_l \ corresponding to variables \ y_1, ..., y_1 \ having place in UMMG \ S = \langle a_0, R, F \rangle, i.e. in unitary metarules and boundary conditions defining multiplicities-variables domains. After that let us add one new unitary metarule

\vec{a}_0 \rightarrow 1 \cdot a_0, y_1 \cdot \vec{y}_1, ..., y_l \cdot \vec{y}_l

(57)

to \ R \ scheme thus creating new \ \tilde{R} \ scheme which contains (57) and all elements (URs and UMRs) of \ R. After that we shall replace all optimizing conditions of the form \ y = opt, having place in \ F \ filter, by \ \vec{y} = opt, thus converting them to the “canonical” object-containing form; let us repeat, that \ \vec{y} \ is object not variable and, more, terminal object because there is no any UR or UMR with the head \ \vec{y} \ in \ R. Constructed filter will be denoted lower as \ \tilde{F}.
As seen now, UMMG \( \bar{S} = (\bar{a_0}, \bar{R}, \bar{F}) \) generates terminal multisets of the form

\[
\{n_{i_1} \cdot a_{i_1}, ..., n_{i_k} \cdot a_{i_k}, n_{\bar{y}_1}, ..., n_{\bar{y}_l}\},
\]

which are selected to \( V_{\bar{S}} \), if and only if TMS

\[
\{n_{i_1} \cdot a_{i_1}, ..., n_{i_k} \cdot a_{i_k}\}
\]

satisfies all conditions entering \( F \) and concerning terminal objects having place in \( R \) scheme, as well as TMS

\[
\{n_{\bar{y}_1}, ..., n_{\bar{y}_l}\}
\]

satisfies all optimizing conditions of the form \( \bar{y}_j = opt_i \).

It is not difficult to define \( V_{\bar{S}} \) by subtracting multisets of the form (60) from all TMS \( v \in V_{\bar{S}} \), but from the practical point of view it is more useful to consider \( V_{\bar{S}} \) not \( V_{\bar{S}} \) as a result of \( S \) application. It’s clear that all TMS \( v \in V_{\bar{S}} \) contain variables \( y_1, ..., y_l \) values as multiplicities-constants of terminal objects \( \bar{y}_1, ..., \bar{y}_l \), which computation is usually main purpose of the mentioned unitary multimetagrammar application.

**Example 11.** Consider UMMG \( S = (a, R, F) \), where \( R \) contains the same unitary metarules as in the previous example while filter \( F \) contains all conditions from this example and one more optimizing condition \( y_1 = max \).

Then

\( R = R \cup \{r_0\} \)

where

\( r_0 = (\bar{a} \rightarrow 1 \cdot a, \bar{y}_1, \bar{y}_2) \),

and \( \bar{S} = (\bar{a}, R, \bar{F}) \), where

\( \bar{F} = F \cup \{\bar{y}_1 = \text{max}\} \).

According to (49) – (55), UMMG \( S \) defines four UMMGs

\( S_1 = (a, R \circ (1,0)) \),
\( S_2 = (a, R \circ (1,1)) \),
\( S_3 = (a, R \circ (2,0)) \),
\( S_4 = (a, R \circ (2,1)) \),

which, as in the Example 10, correspond to four sets of variables \( y_1 \) and \( y_2 \) values. Then, as higher,

\[
R_1 = R \circ (1,0) = \{(a_0 \rightarrow 1 \cdot a, 1 \cdot \bar{y}_1), (a \rightarrow 3 \cdot b, 1 \cdot c), (b \rightarrow 1 \cdot c)\},
\]

\[
V_{s_1} = \{\{4 \cdot c, 1 \cdot \bar{y}_1\}\}
\]

\[
R_2 = R \circ (1,1) = \{(a_0 \rightarrow 1 \cdot a, 1 \cdot \bar{y}_1, 1 \cdot \bar{y}_2), (a \rightarrow 3 \cdot b, 1 \cdot c), (b \rightarrow 1 \cdot c)\},
\]

\[
V_{s_2} = \{\{4 \cdot c, 3 \cdot d, 1 \cdot \bar{y}_1, 1 \cdot \bar{y}_2\}\}
\]

\[
R_3 = R \circ (2,0) = \{(a_0 \rightarrow 1 \cdot a, 2 \cdot \bar{y}_1), (a \rightarrow 3 \cdot b, 2 \cdot c), (b \rightarrow 1 \cdot c)\},
\]

\[
V_{s_3} = \{\{5 \cdot c, 2 \cdot \bar{y}_1\}\}
\]

\[
R_4 = R \circ (2,1) = \{(a_0 \rightarrow 1 \cdot a, 2 \cdot \bar{y}_1, 1 \cdot \bar{y}_2), (a \rightarrow 3 \cdot b, 2 \cdot c), (b \rightarrow 1 \cdot c, 1 \cdot d)\},
\]

\[
V_{s_4} = \{\{5 \cdot c, 3 \cdot d, 2 \cdot \bar{y}_1, 1 \cdot \bar{y}_2\}, \{8 \cdot c, 3 \cdot d, 2 \cdot \bar{y}_1, 1 \cdot \bar{y}_2\}\}.
\]

After that

\[
\bar{V}_s = (\bar{V}_{s_1} \cup \bar{V}_{s_2} \cup \bar{V}_{s_3} \cup \bar{V}_{s_4})
\]

\[
\downarrow \{2 \leq c \leq 4, d = max, \bar{y}_1 = max\} = \{\{4 \cdot c, 1 \cdot \bar{y}_1\}, \{4 \cdot c, 3 \cdot d, 1 \cdot \bar{y}_1, 1 \cdot \bar{y}_2\}\}
\]

\[
\downarrow \{d = max, \bar{y}_1 = max\} = \{\{4 \cdot c, 3 \cdot d, 1 \cdot \bar{y}_1, 1 \cdot \bar{y}_2\}\}.
\]

i.e. set of terminal multisets generated by UMMG \( S \) contains the only TMS \( \{4 \cdot c, 3 \cdot d, 1 \cdot \bar{y}_1, 1 \cdot \bar{y}_2\} \), which corresponds to variables values \( y_1 = 1 \) and \( y_2 = 1 \).

As seen, it is very practically useful to consider namely \( V_s \) as a result of UMMG \( S = (a_0, R, F) \) application (even in the cases where \( R \) does not contain optimizing conditions of the form \( y = opt \)). The only inconvenience which attentive reader may notice due to the last example is the absence of multiobjects of the form \( n \cdot \bar{y} \), where \( n = 0 \). This inconvenience, however, may be eliminated simply by imperative inclusion of multiobjects of the form \( 0 \cdot \bar{y} \) to the generated multiset.

UMMG, which filter contains at least one optimizing condition, is called *optimizing multimetagrammar*. As may be seen, optimizing multimetagrammars are special form of constraints logic programming language, oriented to systems analysis and optimization applications. Mutual interconnections of different classes of unitary multiset grammars are presented at Fig.1.

**Fig. 1.**

OMMG algorithms as well as OMMG family subsets and formal features (internal alternativity, integrity, cyclicity, finitarity) are described in details in [32,33], while multigrammatical representation of classical optimization problems – in the second part of this paper.

**REFERENCES**


