#### ALGEBRAIC GRAPH REWRITING USING A SINGLE PUSHOUT

P.M. van den Broek
Department of Computer Science, Twente University
P.O.Box 217, 7500 AE Enschede, the Netherlands

#### ABSTRACT

We show how graph rewriting can be described with a single pushout in a suitable category of graphs, and compare our result with the conventional approach which uses double pushouts.

#### 1 Introduction

The conventional algebraic approach of graph rewriting is originally due to Ehrig, Pfender and Schneider [1], and is reviewed by Ehrig [2]; it will be named after Ehrig in this paper. It uses double pushouts in a category of graphs to describe graph rewritings.

It would conceptually be much simpler if single pushouts could be used instead of double pushouts. Single pushouts have been used by Raoult [3] and Kennaway [4] to describe graphical term rewriting. A comparison of their approach with Ehrig's approach has been given in Van den Broek [5].

In this paper we present an algebraic approach of graph rewriting which is as general as Ehrig's approach, and which uses single pushouts. This will be done by giving a category of graphs and graph morphisms such that the following pushout diagram

$$\begin{array}{cccc}
L & \xrightarrow{f} & R \\
g & & & \\
G & \xrightarrow{h} & H & diagram 1
\end{array}$$

represents the graph rewriting  $G \longrightarrow H$  via the rewrite rule  $L \longrightarrow R$  based on the occurrence  $L \longrightarrow G$ .

In Ehrig's category of graphs, graph morphisms are total functions on arcs and nodes, so each item (arc or node) of G has an image in H.

So in this category the diagram above can only describe rewritings where items are added to G but none are deleted. Therefore Ehrig's method uses double pushouts, where both G and H arise as pushouts, using a context graph which consists of the items which G and H have in common.

The basic idea behind our approach is to accept that rewriting can only add items to a graph, but not remove any. Instead of removing items, items will be marked. So our graphs will consist of items of two kinds: those which are marked and those which are not. The marked items are to be considered as garbage.

The key difference between the approach of Raoult and Kennaway and our approach is that our graph morphisms, contrary to those of Raoult and Kennaway, preserve the graph structure; they only may violate the markedness of items.

In section 2 we introduce formally our category of graphs and graph morphisms, and the concepts of rewrite rule, occurrence and graph rewriting in this category. In section 3 we establish a bijection between the rewrite rules in this category and Ehrig's rewrite rules corresponding rewrite rules have the same interpretation. In section 4 we show that corresponding rewrite rules give the same rewrite results in both formalisms when the rewrite result in Ehrig's formalism exists. A sufficient condition for the existence of a graph rewriting in our category, given a rewrite rule and an occurrence of its left-hand side in a graph, is given in section 5. Here we also give an example of a graph rewriting in our category for which a corresponding graph rewriting formalism does not exist. In section 6 we define parallel independence of graph rewritings in our category, and show that the commutativity property holds for parallel independent graph rewritings. In section 7 sketch the proofs of several propositions from the earlier sections, after we have introduced a suitable notation.

# 2 The category of marked graphs

The objects in our category are marked graphs, which are graphs with items of two kinds: marked items and unmarked items.

**Definition** A marked graph G is an 8-tuple  $<N_G, M_G, A_G, B_G, s_G, t_G, m_{G,1}, m_{G,2}>$ . Here  $N_G$  denotes a set of unmarked nodes,  $M_G$  a set of marked nodes,  $A_G$  a set of unmarked arcs,  $B_G$  a set of marked arcs;  $s_G: A_G \cup B_G \to N_G \cup M_G$  and  $t_G: A_G \cup B_G \to N_G \cup M_G$  are mappings which map

arcs to their sources and targets respectively;  $m_{G,1}$  and  $m_{G,2}$  are mappings which map nodes and arcs to fixed alphabets of node colours and arc colours respectively.

**Definition** Given two marked graphs G and H, a morphism  $f:G\to H$  is a pair of maps  $f_1:N_G\cup M_G\to N_H\cup M_H$  and  $f_2:A_G\cup B_G\to A_H\cup B_H$  which preserve sources, targets and colours, i.e.

$$f_1.s_G = s_H.f_2,$$
  $f_1.t_G = t_H.f_2,$   $m_{G,1} = m_{H,1}.f_1,$   $m_{G,2} = m_{H,2}.f_2,$ 

and which map marked items onto marked items, i.e.

$$f_1M_G \subseteq M_H$$
,  $f_2B_G \subseteq B_H$ .

It is easily verified that the marked graphs and morphisms, with the usual composition (denoted by the infix symbol • ) and identity, form a category.

**Definition** A marked graph G is called a *graph* if and only if  $M_G = \emptyset$  and  $B_G = \emptyset$ , i.e. if it has no marked items.

With this definition of graphs Ehrig's category of graphs is a subcategory of the category of marked graphs. Note that the category of marked graphs contains in fact too many objects. Considerig marked items as garbage implies that only a marked graph with the property that it remains a marked graph when marked items are deleted is a useful object. Such a marked graph will be called proper:

**Definition** A marked graph G is a proper marked graph if and only if  $s_GA_G \subseteq N_G$  and  $t_GA_G \subseteq N_G$ .

**Definition** The *unmarked subgraph* of a proper marked graph is the graph which is obtained from it by deleting all its marked items.

Proposition 1 The category of marked graphs has pushouts.

**Proof** It is easily verified that one obtains a pushout in the category of marked graphs by marking as few as possible items of the corresponding pushout in the category of graphs such that the morphisms map marked items onto marked items.

As in the category of graphs, the pushout can be constructed by a kind of gluing procedure. Consider diagram 1 as a pushout diagram in the category of marked graphs. Let ~ be the relation on  $N_G \cup M_G \cup N_R \cup M_R$  given by

$$f_1x \sim g_1x$$
  $\forall x \in N_L \cup M_L$ 

and  $\sim$  the equivalence relation which is the reflexive, symmetric and transitive closure of  $\sim$ . The nodes of H are the equivalence classes of  $\approx$ ; so a node of H can be considered to be a set of nodes of G and R which are glued together. A node of H is marked if and only if it contains a marked node of G or R. The mappings  $h_1$  and  $k_1$  map nodes of G and R respectively onto their equivalence classes. The same procedure holds for arcs.

Let L,R and G be proper marked graphs and let  $f:L\to R$  and  $g:L\to G$  be morphisms.

**Definition** (f,L,R) is a marked rewrite rule if and only if

- L is a graph (there is no use for marked items in the left-hand side of a rewrite rule),
- f<sub>1</sub> and f<sub>2</sub> are injective mappings, and
- $M_R \subseteq f_1N_L$  and  $B_R \subseteq f_2A_L$  (there is no use for marked items in the right-hand side of a rewrite rule which are not in the image of the left-hand side).

**Definition** If (f,L,R) is a marked rewrite rule then (g,L,G) is a marked occurrence if and only if  $g_1N_L \subseteq N_G$  and  $g_2A_L \subseteq A_G$  (marked items of G should not take part in the rewriting).

**Definition** If (f,L,R) is a marked rewrite rule and (g,L,G) is a marked occurrence then the pushout construction (diagram 1) gives a marked graph rewriting  $G \longrightarrow H$  via (f,L,R) and based on (g,L,R) if and only if H is a proper marked graph.

Note that the pushout will always exist, but the marked graph rewriting will not exist unless H is proper.

## 3 Connection between the rewrite rules in both formalisms

In this section we will show how one can construct a marked rewrite rule from a rewrite rule in Ehrig's approach which has intuitively the same meaning, and vice versa. These constructions are the inverse of each other, so we obtain a meaning preserving bijection between the rewrite rules in both formalisms.

A rewrite rule in Ehrig's formalism (just called rewrite rule in the sequel) is a 5-tuple (b,k,L,K,R) where L,K and R are graphs and b :  $K \to L$  and k :  $K \to R$  are injective morphisms. The graph K is the

common part of the left-hand side L and the right-hand side R; b and k are the embeddings of K in L and R respectively.

For each rewrite rule a marked rewrite rule can be constructed as follows. First construct the graph Q and the morphism  $f:L\to Q$  by the following pushout:

So Q is the graph which is obtained by gluing from L and R their common part together. Let P be the proper marked graph which is obtained from Q by marking the nodes in  $f_1(N_L\backslash b_1N_K)$  and the arcs in  $f_2(A_L\backslash b_2A_K)$ , and let  $e:Q\to P$  be the morphism with the property that  $e_1$  and  $e_2$  are the identity mapping. Then  $(e\cdot f,L,P)$  is a marked rewrite rule which has the same informal interpretation as the rewrite rule (b,k,L,K,R).

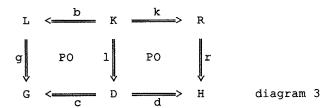
Now the other way around. Let (f,L,P) be a marked rewrite rule. Let R be the unmarked subgraph of P. Let K be the subgraph of L which contains the items of L which are mapped by  $f_1$  or  $f_2$  onto items of R. Let b be the embedding of K in L and k the restriction of f to K. Then (b,k,L,K,R) is a rewrite rule and it has the same informal meaning as the marked rewrite rule (f,L,P).

**Proposition 2** The two constructions described above are each others inverse.

So we have obtained a meaning preserving bijection between rewrite rules and marked rewrite rules.

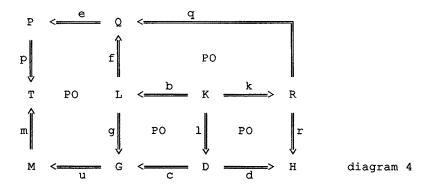
# 4 Comparison of rewrite results in both formalisms

In this section we will show that corresponding rewrite rules give the same rewrite results in both formalisms when the rewrite result in Ehrig's formalism exists. First we introduce a graph rewriting in Ehrig's formalism. Consider diagram 3.



The upper line denotes a rewrite rule (b,k,L,K,R). The leftmost vertical line denotes an occurrence (g,L,G). There exists a graph rewriting  $G \longrightarrow H$  via (b,k,L,K,R) based on (g,L,G) if and only if a graph D and morphisms c and l exist such that the left square is a pushout. Then H is determined by the right pushout square.

Diagram 3 can be extended to diagram 4:



The right upper rectangle is a pushout and P is obtained from Q as described in the previous section. M is any proper marked graph whose unmarked subgraph is G, and u is the embedding of G in M. The diagram is then completed by the pushout of the leftmost rectangle. Now  $(e \circ f, L, P)$  is a marked rewrite rule corresponding to the rewrite rule (b, k, L, K, R) as in the previous section, and  $(u \circ g, L, M)$  is a marked occurrence compatible with the occurrence (g, L, G). There exists a marked graph rewriting  $M \longrightarrow T$  via this marked rewrite rule and based on this marked occurrence if and only if T is proper.

**Proposition 3** In the situation described above T is a proper marked graph and its unmarked subgraph is isomorphic to H.

**Proof** The proof is given in section 7.

So, when an Ehrig rewriting  $G \longrightarrow H$  via some rewrite rule exists, there also exists a marked rewriting  $M \longrightarrow T$  via the corresponding marked rewrite rule for each M whose unmarked subgraph is G, such that

the unmarked subgraph of T is H. Proposition 3 also shows that, when a series of marked rewritings is performed on a graph, it is not necessary to delete the marked items after each step.

## 5 The marked gluing condition

Given the rewrite rule and the occurrence of diagram 3, the rewriting of this diagram exists if and only if Ehrig's gluing condition holds, which reads

$$Da \subseteq b_1N_K & Id_1 \subseteq b_1N_K & Id_2 \subseteq b_2A_K,$$

where

Da = 
$$\{x \in N_L | \exists a \in (A_G \setminus g_2 A_L) \quad g_1 x = s_G a \text{ or } g_1 x = t_G a\}$$
  
 $Id_1 = \{x \in N_L | \exists y \in N_L \quad x \neq y \& g_1 x = g_1 y\}$   
 $Id_2 = \{x \in A_L | \exists y \in A_L \quad x \neq y \& g_2 x = g_2 y\}$ 

Deleting from G an image of a node of Da (dangling) will certainly leave a dangling arc in G. Therefore Da should belong to the image of b. Deleting from G an image of an item of  $\mathrm{Id}_1\cup\mathrm{Id}_2$  (identification) runs the same risk; the left pushout of diagram 3 will however certainly not exist unless  $\mathrm{Id}_1\cup\mathrm{Id}_2$  also belongs to the image of b.

This gluing condition is easily expressed for marked rewrite rules and marked occurrences, using the correspondence of rewrite rules and marked rewrite rules of the previous section. Let (f,L,R) be a marked rewrite rule and (g,L,G) a marked occurrence. Then Ehrig's gluing condition can be written as

$$f_1$$
 Da  $\subseteq$   $N_R$  &  $f_1$  Id $_1 \subseteq N_R$  &  $f_2$  Id $_2 \subseteq A_R$ 

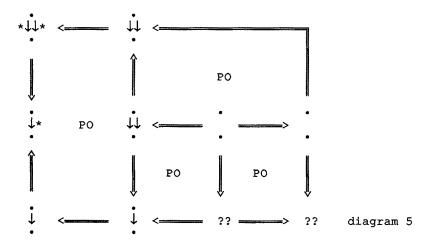
From the result of section 3 it follows that this condition is a sufficient condition for the existence of a marked graph rewriting  $G \longrightarrow H$  via (f,L,R) and based on (g,L,G). We have however a stronger result.

**Definition** The marked gluing condition for a marked rewrite rule (f,L,R) and a marked occurrence (g,L,G) is  $f_1$  Da  $\subseteq N_R$  &  $f_1$  Id<sub>1</sub>  $\subseteq N_R$   $\Diamond$ 

**Proposition 4** The marked gluing condition for the marked rewrite rule (f,L,R) and the marked occurrence (g,L,G) is a sufficient condition for the existence of a marked graph rewriting  $G \longrightarrow H$  via (f,L,R) and based on (g,L,G).

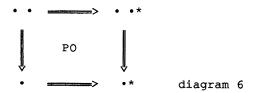
**Proof** The proof is given in section 7.

A simple example where the gluing condition does not hold but the marked gluing condition holds is shown in diagram 5.



This diagram has the same structure as diagram 4, but the marked graphs are shown explicitely; marked items are shown with an asterisk. The (marked) rewrite rule means that double arcs in a graph may be removed, and it is tried to use this rule to remove a single arc.

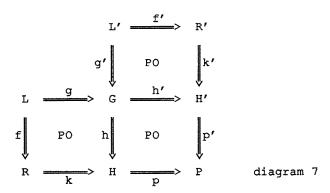
An example which shows that the marked gluing condition is not a necessary condition for the existence of a marked graph rewriting is the following.



The upper line is a marked rewrite rule which says that from each two nodes one node may be removed. This rule is applied to a graph consisting of one single node. The rewriting succeeds, even though the marked gluing condition does not hold.

### 6 Parallel independence and the commutativity property

In this section we will introduce the concept of parallel independence of marked graph rewritings and show that the commutativity property (called Church-Rosser property by Ehrig) holds for parallel independent marked graph rewritings.



Consider diagram 7. Here (f,L,R) and (f',L',R') are marked rewrite rules, (g,L,G) and (g',L',G) are marked occurrences and the upper and the left pushout squares give the marked graph rewritings  $G \longrightarrow H'$  and  $G \longrightarrow H$  respectively. The diagram is completed by the lower right pushout square.

**Definition** The marked graph rewritings  $G \longrightarrow H'$  and  $G \longrightarrow H$  of diagram 7 are parallel independent if no items of G which are in the image of G under G are mapped onto marked items of G which are in the image of G under G are mapped onto marked items of G which are in the image of G under G are mapped onto marked items of G by G by G are mapped onto marked items of G by G by G are formally this condition reads

$$g_1N_L \cap g'_1.f'_1^{-1} M_{R'} = \emptyset$$
 $g_2A_L \cap g'_2.f'_2^{-1} B_{R'} = \emptyset$ 
 $g'_1N_{L'} \cap g_1.f_1^{-1} M_{R} = \emptyset$ 
 $g'_2A_{L'} \cap g_2.f_2^{-1} B_{R} = \emptyset$ 

In case there exist corresponding graph rewritings in Ehrig's formalism, i.e. if the gluing condition is satisfied for both marked graph rewritings, this definition of parallel independence coincides with the definition by Ehrig. The following proposition states that parallel independent marked rewritings are commutative:

**Proposition 5** Let  $G \longrightarrow H$  and  $G \longrightarrow H'$  be parallel independent marked graph rewritings via marked rewrite rules r and r'

respectively. Then there exist a marked graph P and marked graph rewritings  $H \longrightarrow P$  and  $H' \longrightarrow P$  via r' and r respectively.

**Proof** Note that parallel independence of  $G \longrightarrow H$  and  $G \longrightarrow H'$  is necessary and sufficient for  $(h \cdot g', L', H)$  and  $(h' \cdot g, L, H')$  to be marked occurrences. A standard result from category theory gives that the two rectangles LRPH' and L'R'PH are also pushouts. The proposition has been proven if we show that  $H \longrightarrow P$  and  $H' \longrightarrow P$  are two marked graph rewritings, i.e. that P is a proper marked graph. This will be shown in section 7.

#### 7 Proofs

In this section we will present a number of proofs of propositions given in the previous sections. We have postponed these proofs to this section since for these proofs we will introduce a special notation for graphs. The reason for this is the following. A typical situation is that we have a diagram consisting of pushouts, e.g. diagram 4 or diagram 7, and that a proof of a proposition consists of a straightforward verification, using the properties of pushouts extensively. This leads in general to a case-analysis with many cases, and to proofs which are boring to read.

In this section a graph will be denoted as a partition of its items. Sets of items will be denoted by small numbers. The items of a set will either be all marked or all unmarked; sets of marked items will be shown underlined. Consider for example diagram 2. Let the items of K all belong to the same set: K = (1). Since k is injective the items of R can be particulated into 2 disjoint sets: R = (1,2). Note that K is denoted as a subgraph of R. Since b is also injective we may take L = (1,3). The graph Q, obtained via the gluing procedure, can now be denoted by (1,2,3). Note that the graph structure is not present in this notation. This is no restriction however, since the graph structure is preserved by the morphisms.

**Proof of Proposition 2.** Let, as above, K = (1), R = (1,2), L = (1,3) and Q = (1,2,3). The marked graph P was obtained from Q by marking the items in set 3, so P = (1,2,3). The marked rewrite rule which corresponds to the rewrite rule (b,k,L,K,R) then is (f,L,P).

From the definition of marked rewrite rule it follows that for each marked rewrite rule (f,L,P) we can take L=(1,3) and  $P=(1,2,\underline{3})$ . The construction of the graphs K and R from L and P is as follows: R is

the unmarked subgraph of P, so R = (1,2) (recall that P is a proper marked graph by definition), and K consists of those items of L which are mapped by f onto items of R, hence K = (1).

It is now obvious that we established a 1-1 correspondence between rewrite rules and marked rewrite rules.

**Proof of Proposition 3** Consider diagram 4. Let, as above, K = (1), R = (1,2), L = (1,3), Q = (1,2,3) and P = (1,2,3). Suppose the items of K are mapped by 1 onto the items in set 1' of D. Then we may write D = (1',4). Since K is injective, K is injective too, so we may write K is injective, K is obtained from K by adding marked items to K, so K is K in K. So K is a proper marked graph whose unmarked subgraph is K.

**Proof of Proposition 4** Consider diagram 1, and assume that (f,L,R) is a marked rewrite rule and (g,L,G) is a marked occurrence. We may write L = (1,2) and R = (1,2,3). Let 1' and 2' be the images under g of 1 and 2 respectively. Since 1' and 2' need not be disjoint we write  $G = (1' \setminus 2', 2' \setminus 1', 1' \cap 2', 4, \underline{5})$ . The gluing construction then gives  $H = (1' \setminus 2', 2' \setminus 1', 1' \cap 2', 3, 4, \underline{5})$ .

Now suppose that n is a marked node in set x of H and a is an unmarked arc in set y of H which is adjacent to n. For  $\langle x,y \rangle$  there are the following 9 possibilities:  $\langle 2' \backslash 1', 1' \backslash 2' \rangle$ ,  $\langle 2' \backslash 1', 3 \rangle$ ,  $\langle 2' \backslash 1', 4 \rangle$ ,  $\langle 1' \cap 2', 1' \backslash 2' \rangle$ ,  $\langle 1' \cap 2', 3 \rangle$ ,  $\langle 1' \cap 2', 4 \rangle$ ,  $\langle 5, 1' \backslash 2' \rangle$ ,  $\langle 5, 3 \rangle$  and  $\langle 5, 4 \rangle$ . Our task is to show that the marked gluing condition excludes all 9 possibilities.

We can exclude 5 of these possibilities without using the marked gluing condition: <5,3> is excluded since a would belong to R while n does not;  $<2'\setminus1',1'\setminus2'>$  and  $<2'\setminus1',3>$  are excluded since R is a proper marked graph;  $<5,1'\setminus2'>$  and <5,4> are excluded since G is a proper marked graph.

Since Da consists of those nodes from 1 and 2 whose image under g have an adjacent arc in 4 and  $N_R$  consists of the nodes in 1 and 3, the condition  $f_1$  Da  $\subseteq$   $N_R$  is equivalent with the exclusion of  $<2'\setminus 1'$ , 4> and  $<1'\cap 2'$ , 4>.

We are left with the 2 possibilities  $<1'\cap2',1'\setminus2'>$  and  $<1'\cap2',3>$ . Since  $\mathrm{Id}_1$  consists of the nodes of 1 and 2 which have the same image under g as some other node of 1 and 2, the condition  $f_1$   $\mathrm{Id}_1\subseteq N_R$  implies that  $1'\cap2'$  contains no nodes. This excludes the last two possibilities, so our proposition has been proven.

Note that from the proof above we also obtain a necessary and sufficient condition for the existence of a marked rewriting: the conjunction of  $f_1$  Da  $\subseteq$  N<sub>R</sub> and a necessary and sufficient condition for the exclusion of the possibilities <1'\times2',1'\2'> and <1'\times2',3>. This last condition can be formulated as: nodes of 1 which are identified by g with a node in 2 should have no adjacent arcs in 3, nor adjacent arcs in 1 which are not identified by g with an arc in 2.

Completion of the proof of proposition 5 Consider diagram 7. Our task is to show that P is a proper marked graph if the marked rewritings  $G \longrightarrow H$  and  $G \longrightarrow H'$  are parallel independent. Let L = (1,2)and R = (1, 2, 3). Let 1' and 2' be the images under g of 1 and 2 respectively. Then we may write  $G = (1' \setminus 2', 2', 4, 5)$ . Analogously we may write L' = (6,7), R' = (6,7,8) and  $G = (6' \setminus 7',7',9,5)$ . A notation for G which is useful for both of its rewritings is  $G = ((1'\setminus 2') \cap (6'\setminus 7'), (1'\setminus 2') \cap 7', (1'\setminus 2') \cap 9, 2' \cap (6'\setminus 7'), 2' \cap 7', 2' \cap 9,$  $4\cap(6'\setminus7'),4\cap7',4\cap9,\underline{5})$ . At this point we can use the parallel independence of the marked rewritings  $G \longrightarrow H$  and  $G \longrightarrow H'$ , since it implies that  $(1'\backslash 2') \cap 7' = 2' \cap (6'\backslash 7') = 2' \cap 7' = \emptyset$ , by simplifying the notation of G to G =  $((1'\2')\cap(6'\7'), (1'\2')\cap9, 2'\cap9, 4\cap(6'\7'), 4\cap7',$  $4\cap 9, 5$ ). We may now use the gluing procedure to obtain H, H' and P, giving  $H = ((1'\backslash 2') \cap (6'\backslash 7'), (1'\backslash 2') \cap 9, 2'\cap 9, 4\cap (6'\backslash 7'), 4\cap 7', 4\cap 9, 3, 5),$  $((1'\2')\cap(6'\7'), (1'\2')\cap9, 2'\cap9, 4\cap(6'\7'), 4\cap7', 4\cap9, 8, 5)$  $P = ((1'\2')\cap(6'\7'), (1'\2')\cap9, 2'\cap9, 4\cap(6'\7'), 4\cap7', 4\cap9, 3, 8, 5).$ suppose that n is a marked node in set x of P and a is an unmarked arc set y of P which is adjacent to n. For <x,y> there are 18 possibilities. It is easy to check that the fact that H and H' are proper marked graphs excludes 16 of these possibilities, leaving only the possibilities  $\langle 2' \cap 9, 8 \rangle$  and  $\langle 4 \cap 7', 3 \rangle$ . If a belongs to 8 then a is an arc of R'; then n should be a node of R'. This excludes the possibility <2'∩9,8>; analogously the possibility excluded. So we conclude that n and a cannot exist; thus P is a proper marked graph.

### References

- [1] H. Ehrig, H. Pfender and H.J. Schneider, Graph grammars: an algebraic approach, in: Proc. 14th Ann. Conf. on Switching and Automata Theory (1973) 167-180.
- [2] H. Ehrig, Introduction to the algebraic theory of graph grammars

- (a survey), in: V. Claus, H. Ehrig and G. Rozenberg, eds., Graph grammars and their application to Computer Science and Biology, Lecture Notes in Computer Science 73 (Springer, Berlin, 1979) 1-69.
- [3] J.C. Raoult, On graph rewritings, Theoret. Comput. Sci. 32 (1984) 1-24.
- [4] R. Kennaway, On "On graph rewritings", Theoret. Comput. Sci. 52 (1987) 37-58.
- [5] P.M. van den Broek, Comparison of two graph-rewrite systems, Theoret. Comput. Sci. 61 (1988) 67-81.