Equivalence of Match-, Change- and Inverse Match-Boundedness for Length-Preserving String Rewriting

Dieter Hofbauer
Berufsakademie Nordhessen
Eichlerstraße 25, D-34537 Bad Wildungen, Germany.

Johannes Waldmann
Hochschule für Technik, Wirtschaft und Kultur (FH) Leipzig
Fb IMN, PF 30 11 66, D-04251 Leipzig, Germany.

1 Introduction

String rewriting is a model of non-deterministic, step-wise computation. To get more detailed information on derivations, positions in strings can be annotated by natural numbers, called heights. That means switching from an alphabet Σ to the annotated alphabet $\Sigma \times \mathbb{N}$. For each rewrite derivation, there is a corresponding annotated derivation. It starts with all annotations equal to zero, and proceeds by computing annotations in the contractum from annotations in the redex (and leaving annotations unchanged elsewhere).

One example of annotated rewriting was given by Ravikumar [Rav04], where the annotations in the contractum are obtained as successors of the annotations at corresponding positions in the redex. These annotations are called *change heights*. This definition makes sense only for length-preserving systems, i.e., rewriting systems where for each rule, the left- and the right-hand side have equal length.

As an example, take the rewriting system $R = \{abb \to bba, bbb \to aaa\}$ over the alphabet $\Sigma = \{a, b\}$, and the R-derivation $\underline{abb}bb \to bb\underline{abb} \to b\underline{bbb}a \to baaaa$, where for each step, the redex is underlined. If we annotate this derivation with change heights, we get

$$\underline{a_0 b_0 b_0} b_0 b_0 \to b_1 b_1 \underline{a_1 b_0 b_0} \to b_1 \underline{b_1 b_2 b_1} a_1 \to b_1 a_2 a_3 a_2 a_1.$$

The concept of change heights has been generalized by Geser and the present authors [GHW04] to *match heights*: here the annotations in the contractum are defined as the successor of the minimal annotation in the redex. If we annotate the above derivation with match heights, we get

$$\underline{a_0b_0b_0}b_0b_0 \to b_1b_1\underline{a_1b_0b_0} \to b_1\underline{b_1b_1b_1}a_1 \to b_1a_2a_2a_2a_1.$$

If there is an upper bound on the heights of all annotated R-derivations that start from strings annotated by zero, then the string rewriting system R is called annotation-bounded (e.g., change-bounded, match-bounded).

Match-bounded string rewriting systems effectively preserve regularity of languages, and they are terminating. This concept can be used for automated termination provers, also for term rewriting [GHWZ07, KM07].

The inverse of a rewrite system flips all its arrows; in the example, $R^- = \{bba \to abb, aaa \to bbb\}$. Then the above R-derivation corresponds to the R^--derivation $b\underline{aaa}a \to bb\underline{bba} \to \underline{bba}bb \to abbbb$. If we annotate this inverted derivation with match heights, we get

$$b_0\underline{a_0a_0a_0}a_0 \to b_0b_1\underline{b_1b_1a_0} \to \underline{b_0b_1a_1}b_1b_1 \to a_1b_1b_1b_1b_1.$$

These match heights of the inverse derivation are called *inverse match heights* of the original derivation. Systems that are inverse match-bounded have been treated in [GHW05], where it is shown that they effectively preserve context-freeness, and that the termination and the uniform termination problem are both decidable for this class. There, the relation between match-boundedness of the original system and the inverse system has been stated as an open question. Here, we give a partial answer: for length-preserving systems R, the following statements are equivalent:

1. R is match-bounded,
2. R is change-bounded,
3. R is inverse change-bounded,
4. R is inverse match-bounded.

The following implications are obvious: (2) $\Rightarrow$ (1) and (2) $\Leftrightarrow$ (3). So we have to show (1) $\Rightarrow$ (2), and the rest follows by symmetry.

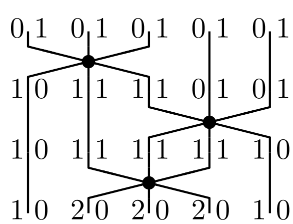
The proof uses linear algebra in the (min,plus) semi-ring to compute match heights. (A similar observation is that the Tetris computer game can be modelled by (max,plus) matrices [GP97].) For lack of space, we omit the proof here, and only illustrate the claim.

2 Match Heights and Inverse Match Heights

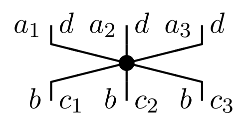
For our considerations, we can completely ignore the rules of R and the letters of Σ and consider *only* their height annotations instead. Then an annotated derivation is just a sequence of strings over $\mathbb{N}$. In our running example, rules have width 3 and we apply them on a string of length 5 at positions 0, 2, and 1 successively. We get the following sequences (redexes underlined):

Derivation:	$\underline{abb}bb \to bb\underline{abb} \to b\underline{bbb}a \to baaaa,$
Change heights:	$\underline{000}00 \to 11\underline{100} \to 1\underline{121}1 \to 12321,$
Match heights:	$\underline{000}00 \to 11\underline{100} \to 1\underline{111}1 \to 12221,$
Inverse match heights:	$11111 \leftarrow \underline{011}11 \leftarrow 01\underline{110} \leftarrow 0\underline{000}0.$

We now visualize this rewrite sequence, together with its match heights and inverse match heights, as a (directed) graph. Each rewrite step corresponds to a node, each edge (orientet top-down) stands for an (annotated) position in a string. Edges are annotated by the pair of match height and inverse match height at the corresponding position. Note that match height annotations start at zero in the top (north) row, and inverse match height annotations start at zero in the bottom (south) row.



general pattern for one node:



where $b = 1+\min\{a_1, a_2, a_3\}$ and $d = 1 + \min\{c_1, c_2, c_3\}$

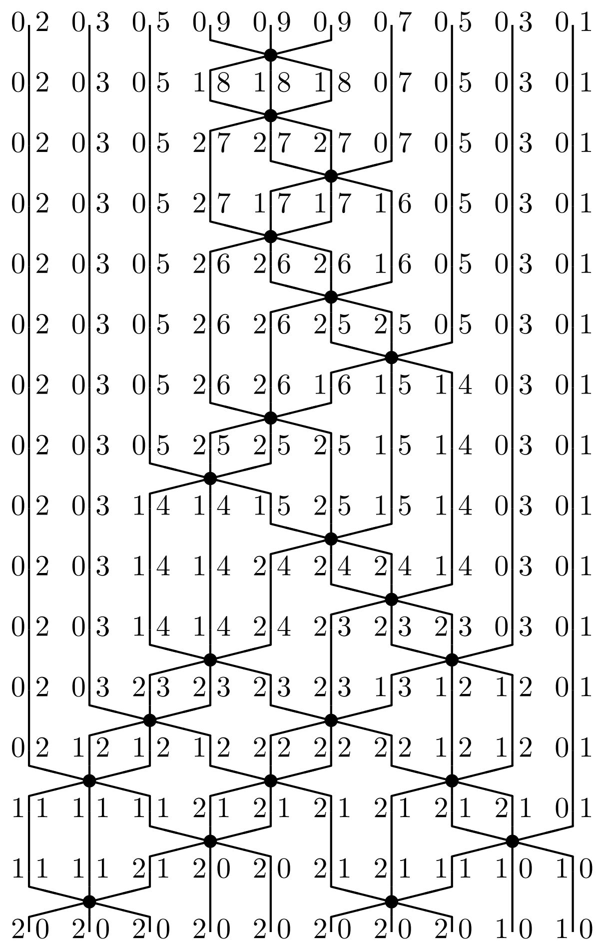
The question about the relation between bounds for match heights and inverse match heights can be rephrased as a problem on planar directed acyclic graphs as in the picture. Assume that all nodes have equal in- and outdegree w. If each node has distance at most M from the top, is there a bound M' on the distances from the bottom?

The answer is yes, and we can prove $M' \leq (2Mw + 1)(w + 1)^M$. We do not know how sharp this bound is. Via computer experiments we found derivations of the following inverse match heights M' (for given match height

M and rule width w):

M'	$w = 1$	2	3	4	5
$M = 1$	1	2	2	2	2
2	2	6	9	12	15
3	3	12	20		
4	4	19			
5	5	26			

The following example illustrates rule width $w = 3$, match-bound $M = 2$ and inverse match-bound $M' = 9$.



References

[GHW04] Alfons Geser, Dieter Hofbauer, and Johannes Waldmann. Match-bounded string rewriting systems. *Appl. Algebra Eng. Commun. Comput.*, 15(3-4):149–171, 2004.

[GHW05] Alfons Geser, Dieter Hofbauer, and Johannes Waldmann. Termination proofs for string rewriting systems via inverse match-bounds. *J. Autom. Reasoning*, 34(4):365–385, 2005.

[GHWZ07] Alfons Geser, Dieter Hofbauer, Johannes Waldmann, and Hans Zantema. On tree automata that certify termination of left-linear term rewriting systems. *Inf. Comput.*, 205(4):512–534, 2007.

[GP97] Stephane Gaubert and Max Plus. Methods and applications of (max, +) linear algebra. In Rüdiger Reischuk and Michel Morvan, editors, *STACS*, volume 1200 of *Lecture Notes in Computer Science*, pages 261–282. Springer, 1997.

[KM07] Martin Korp and Aart Middeldorp. Proving termination of rewrite systems using bounds. In Franz Baader, editor, *RTA*, volume 4533 of *Lecture Notes in Computer Science*, pages 273–287. Springer, 2007.

[Rav04] Bala Ravikumar. Peg-solitaire, string rewriting systems and finite automata. *Theor. Comput. Sci.*, 321(2-3):383–394, 2004.