MEMBERSHIP PROBLEMS FOR FINITE ENTROPIC GROUPOIDS

J. JEŽEK AND M. MARÓTI

ABSTRACT. In the first part of this paper we find an algorithm, deciding for any finite groupoid, whether it satisfies all entropic equations. In the second part we prove that there is no algorithm, deciding the same for any finite partial groupoid. Satisfaction of an equation in a partial groupoid means that if both sides are defined, with respect to some interpretation, then the values of the sides must be equal.

1. Introduction

By a *medial* groupoid we mean a groupoid satisfying the equation $(xy)(uv) \approx (xu)(yv)$. *Entropic* groupoids are homomorphic images of medial cancellation groupoids. The class of entropic groupoids is a variety. This variety has been introduced in [2]; the paper (some parts can be also found in [3],[4] and [5]) contains several equivalent definitions. The variety is not finitely based.

In [1], the following problem has been raised: Does there exist an algorithm, deciding for any finite groupoid whether it is entropic? In the first part of this paper we are going to present such an algorithm. On the other hand, in the second part (the sections starting with section 5) we will show that there is no algorithm deciding for any finite partial groupoid whether it satisfies all the equations of entropic groupoids. By satisfaction of an equation in a partial groupoid we mean that if both sides are defined, with respect to some interpretation, then the values of the two sides must be equal.

The algorithm that we are going to present in the first part is based on Theorem 2. It works as follows: Given a groupoid with N elements, check if it satisfies all the basic entropic equations of depth up to $5N^{18}$. If it does, the groupoid is entropic according to the theorem; if it does not, then of course it is not entropic. This algorithm is of no practical value: even for $N = 2$, the number of equations to be considered is too big. For $N = 2$, however, one can do much better: it is easy to see that a two-element groupoid is entropic

1991 *Mathematics Subject Classification.* 20N02, 68W05, 68Q05, 03D35, 03D10.

This material is based upon work supported by the Grant Agency of the Czech Republic, grant No. 201/96/0312 and by the National Science Foundation under grants no. DMS 9941276 and DMS 9596043.

if and only if it is medial, and this is easy to check. The following problem remains open: Can the membership problem for finite entropic groupoids be decided by an algorithm working in a reasonable time for groupoids with, say, at most 26 elements? Is there an algorithm, working in polynomial time?

For the terminology and basic notions of equational logic, helpful for understanding the following text, the reader is referred to [7].

In order to be able to describe the equational theory of entropic groupoids, we need to introduce the following notation. Given a term t (we mean a term in the similarity type containing just one binary operation symbol, for multiplication) and an occurrence o of a variable x in t, the *weight* of o (in t) is the ordered pair (i, j), where i is the number of southwest turns and j is the number of southeast turns in the downward path connecting the top of the term's tree with the occurrence o; the sum $i + j$ is called the *depth* of o. For example, the weight of the (single) occurrence of z in $(x(yz))(xy)$ is $(1, 2)$, and the depth is 3. Now, an equation $t \approx u$ belongs to the equational theory of entropic groupoids if and only if for any variable x and any ordered pair (i, j) of nonnegative integers, the number of occurrences of x of weight (i, j) in t is the same as the number of occurrences of x of weight (i, j) in u (see [3]). For example, the medial law, and also the equation $(x(yz))((uv)w) \approx (x(yv))((uz)w)$ belong to the equational theory.

The paper [8] contains a construction of an infinite independent base for the equations of entropic groupoids. In this paper we will need the following consequence (which is, however, also easy to prove without relying on [8]).

By a *slim* term we mean a term t such that whenever uv is a subterm of t, then either u or v is a variable. By a *linear* term we mean a term containing no variable more than once. Let t, u be two slim terms such that the term tu is linear; let x be a variable in t and y be a variable in u, such that the weights of x and y in tu are the same, and there is no variable in tu of greater depth. Denote by t' the term obtained from t by replacing x with y, and by u' the term obtained from u by replacing y with x. Equations $tu \approx t'u'$, obtained in this way, will be called *basic entropic equations.*

Lemma 1. *The set of basic entropic equations is a base for the equational theory of entropic groupoids.* □

By the depth of a term t we mean the maximum of the depths of occurrences of variables in t, and by the depth of an equation $t \approx u$ we mean the maximum of the depths of t and u. The aim of the next three sections is to prove the following theorem, yielding the decidability of the membership problem for finite entropic groupoids.

Theorem 2. *Let G be a finite groupoid with N elements $(N \geq 2)$. Then G is entropic if and only if it satisfies all the basic entropic equations of depth at most $5N^{18}$.*

2. Proof of Theorem 2: shifting around

Let G be a finite groupoid with N elements ($N \geq 2$). Let us fix two symbols α and β (they can be thought of as symbols for the southwest and the southeast direction in trees of terms, respectively). For any positive integer n denote by E_n the set of finite sequences $e = (e_1, a_1, \dots, e_{n-1}, a_{n-1}, e_n)$, where $e_i \in \{\alpha, \beta\}$ and $a_i \in G$. The elements of E_n will be called *paths* (of length n).

For $a \in G$, $e = (e_1, a_1, \dots, e_n) \in E_n$ and $i \in \{0, \dots, n-1\}$ define an element $(a * e)_i$ of G as follows: $(a * e)_0 = a$; if $e_i = \alpha$, then $(a * e)_i = (a * e)_{i-1} a_i$; if $e_i = \beta$, then $(a * e)_i = a_i (a * e)_{i-1}$.

For $i \in \{0, \dots, n\}$ put $\mathrm{w}_e^\alpha(i) = |\{j : 1 \leq j \leq i,\ e_i = \alpha\}|$, $\mathrm{w}_e^\beta(i) = |\{j : 1 \leq j \leq i,\ e_i = \beta\}|$ and $\mathrm{w}_e(i) = (w_e^\alpha(i), \mathrm{w}_e^\beta(i))$. The ordered pair $\mathrm{w}_e(i)$ will be called the *e-weight* of i (it would be also possible to call it the weight of the i-th position in the path e, with respect to the paths's bottom). The e-weight of n will be called the *weight* of the path e.

Let $(a, b) \in G^2$ be fixed. Also, for most of the time, the positive integer n will be fixed.

For $e \in E_n$ we define a mapping κ_e of $\{0, \dots, n-1\}$ by $\kappa_e(i) = ((a*e)_i, (b*e)_i)$.

Two paths $e = (e_0, a_1, \dots, e_n)$ and $f = (f_0, b_1, \dots, f_n)$ of the same length n are said to be *similar* if $e_n = f_n$, $\kappa_e(n-1) = \kappa_f(n-1)$ and there is a permutation π of $\{0, \dots, n-1\}$ such that $f_i = e_{\pi(i)}$ and $b_i = a_{\pi(i)}$ for all $i = 1, \dots, n-1$.

For $0 \leq i < j \leq n$ put $[i, j] = \{i, i+1, \dots, j\}$. These sets will be called *segments*. The number $j - i$ is called the *length* of $[i, j]$. (By definition, the length is always positive.) Two segments $[i, j]$ and $[k, l]$ are said to be *nonoverlapping* if either $j \leq k$ or $l \leq i$. By the *total length* of a set S of pairwise nonoverlapping segments we mean the sum of the lengths of all segments in S. A segment $[i, j]$ is called *regular* if $j < n$. For a regular segment $[i, j]$, the two ordered pairs, $\kappa_e(i)$ and $\kappa_e(j)$, will be called the *lower* and the *upper e-value* of $[i, j]$, respectively; if they are the same, we say that the segment is *e-valued* and we call $\kappa_e(i)$ the e-value of $[i, j]$. A segment is called *e-correct* if it is e-valued and of length at most N^2 (in particular, it must be regular). Since the range of κ has at most N^2 elements, it is easy to see that for a given e, every regular segment of length at least N^2 contains at least one e-correct subsegment. A regular segment $[i, j]$ is called *e-correctly glued* if there is a sequence $i = p_0 < p_1 < \dots < p_r = j$ such that $[p_{k-1}, p_k]$ is e-correct for any $k = 1, \dots, r$. Of course, every e-correctly glued segment is e-valued.

By an *e-assembly* we will mean a set of pairwise disjoint, e-correctly glued segments with pairwise different e-values. (Clearly, an e-assembly contains at most N^2 sets.) By a *gap* in C we mean any regular segment $[i, j]$ such that i is either 0 or the last element of a segment in C, j is either $n-1$ or the first element of a segment in C, and there is no segment in C contained in $[i,j]$. Clearly, there are at most N^2+1 gaps in C, and the sum of the lengths of all gaps and of all segments in C gives $n-1$ precisely. By a *maximal e-assembly* we will mean an e-assembly C such that for any path e' similar to e, any e'-assembly has total length less or equal to the total length of C.

Lemma 3. *Let $e \in E_n$ and C be a maximal e-assembly. Then the total length of C is at least $n-N^4$.*

Proof. Suppose, on the contrary, that the total length of C is smaller than $n-N^4$. This is the same as to say that the sum of the lengths of the gaps in C is at least N^4. There are at most N^2+1 gaps. If each of them were of length at most N^2-1, then the sum of their lengths would be at most $(N^2+1)(N^2-1)=N^4-1$, a contradiction. So, there is at least one gap of length at least N^2. But then, there is an e-correct segment $[u,v]$ contained in that gap.

Suppose there is no segment in C having the same e-value as $[u,v]$. Then $C \cup \{[u,v]\}$ is an e-assembly of greater total length compared to that of C, a contradiction.

So, there is precisely one segment $[k,h] \in C$ with the same e-value as $[u,v]$. We have either $v \le k$ or $h \le u$. Let us consider the first case.

Where $e=(e_1,a_1,\dots,e_n)$, let

$$e'=(e_1,a_1,\dots,e_u,a_u,e_{v+1},a_{v+1},\dots,e_k,a_k,e_{u+1},a_{u+1},\dots,e_v,a_v,\\ e_{k+1},a_{k+1},\dots).$$

Let C' be the set obtained from C by replacing $[k,h]$ with $[k-(v-u),h]$ and any segment $[i,j]\in C$, contained in $[v,k]$, with $[i-(v-u),j-(v-u)]$. It is easy to see that e' is similar to e and C' is an e'-assembly with total length larger than the total length of C, a contradiction.

In the second case, if $h \le u$, the segment $[u,v]$ could be shifted to hang at the position h and joined to $[k,h]$ in a similar way, yielding a contradiction as well. □

Lemma 4. *For every path $e \in E_n$ there exists a path e' similar to e such that there is a set S of pairwise nonoverlapping, e'-correct segments of total length at least $n-N^4$.*

Proof. It is an immediate consequence of Lemma 3. □

3. Proof of Theorem 2 continued: slopes

Throughout this section let a pair $(a,b)\in G^2$ and a path $e\in E_n$ be fixed. We will assume that there exists a set S of pairwise nonoverlapping, e-correct segments of total length at least $n-N^4$, and we will keep S fixed.

Lemma 5. *Let m be a positive integer and let $(i,j),(k,l)$ be two pairs of nonnegative integers such that $0 < i+j \leq m$, $0 < k+l \leq m$ and $\frac{i}{i+j} \neq \frac{k}{k+l}$. Then $|\frac{i}{i+j} - \frac{k}{k+l}| \geq \frac{1}{m^2}$.*

Proof. We have $|\frac{i}{i+j} - \frac{k}{k+l}| = |\frac{c}{(i+j)(k+l)}|$ for an integer c. Since the fraction is nonzero, we have $|c| \geq 1$ and hence $|\frac{c}{(i+j)(k+l)}| \geq \frac{1}{m^2}$. □

For each segment $[i,j]$ put $\lambda_e[i,j] = \frac{\mathrm{w}_e^\alpha(j) - \mathrm{w}_e^\alpha(i)}{j-i}$. This is a rational number between 0 and 1; it will be called the *e-slope* (or just slope, if e is clear from context) of $[i,j]$. Since

$$\lambda_e[i,j] = \frac{w_e^\alpha(j) - \mathrm{w}_e^\alpha(i)}{w_e^\alpha(j) - \mathrm{w}_e^\alpha(i) + w_e^\beta(j) - \mathrm{w}_e^\beta(i)},$$

it follows from Lemma 5 that if λ_1 and λ_2 are two different slopes of two segments of length at most N^2, then $|\lambda_1 - \lambda_2| \geq \frac{1}{N^4}$.

Put $\Lambda_e = \lambda_e[0,n] = \frac{\mathrm{w}_e^\alpha(n)}{n}$.

A rational number r will be called *large* (with respect to e) if $r \geq \Lambda_e + \frac{1}{2N^4}$; it will be called *small* if $r \leq \Lambda_e - \frac{1}{2N^4}$; and *middle* if $|r - \Lambda_e| < \frac{1}{2N^4}$.

Lemma 6. *There is at most one middle rational number r with the property that there is a segment of length at most N^2 with e-slope equal to r.*

Proof. It follows from Lemma 5 and the definitions. □

If it exists, the unique middle rational number from Lemma 6 will be denoted by Λ_e'. If it does not exist, we put Λ_e'=Λ_e.

The set S is the disjoint union $S_{-1} \cup S_0 \cup S_1$, where S_{-1}, S_0 and S_1 denote the set of the segments in S with small, middle and large slopes, respectively.

For $k \in \{-1,0,1\}$ put $d_k = \sum_{[i,j]\in S_k} (\lambda_e[i,j] - \Lambda_e)(j-i)$.

Lemma 7. *We have*

(1) $|d_{-1} + d_0 + d_1| \leq N^4$,
(2) $-|S_{-1}|N^2 \leq d_{-1} \leq -\frac{|S_{-1}|}{2N^4}$,
(3) $|d_0| \leq \frac{|S_0|}{2N^2}$,
(4) $\frac{|S_1|}{2N^4} \leq d_1 \leq |S_1|N^2$.

Proof. For each $i = 0,\dots,n$ put $\delta_e(i) = \mathrm{w}_e^\alpha(i) - i\Lambda_e$. (This rational number could be called the distance of the i-th position on the branch e from the line connecting the top of e with its bottom.) Clearly, $\delta_e(0) = \delta_e(n) = 0$.

It is easy to check that for any segment $[i,j]$ we have $\delta_e(j) - \delta_e(i) = (\lambda_e[i,j] - \Lambda_e)(j-i)$. Denote by S' the set of all the segments of length 2 that are not contained in any segment from S, so that the total length of $S \cup S'$ is precisely n

and the total length of S' is at most N^4. We have

$$0 = \delta_e(n) - \delta_e(0) = \sum_{[i,j]\in S\cup S'} (\delta_e(j) - \delta_e(i)) = \sum_{[i,j]\in S\cup S'} (\lambda_e[i,j] - \Lambda_e)(j-i)$$
$$= d_{-1} + d_0 + d_1 + \sum_{[i-1,i]\in S'} (\lambda_e[i-1,i]).$$

The last sum is in absolute value at most N^4, so $|d_{-1} + d_0 + d_1| \leq N^4$. We have proved (1).

In order to prove (2), (3) and (4), observe that $1 \leq j-i \leq N^2$ and $|\lambda_e[i,j] - \Lambda_e| < \frac{1}{2N^4}$ in the case (3), while $\frac{1}{2N^4} \leq |\lambda_e[i,j] - \Lambda_e| \leq 1$ in cases (2) and (4). □

Lemma 8. *If $n > 5N^{18}$, then at least one of the following two cases takes place: either $|S_0| \geq 2N^{10}$ or both $|S_{-1}| \geq N^{10}$ and $|S_1| \geq N^{10}$.*

Proof Let $|S_0| < 2N^{10}$. Since the total length of S is at least $n - N^4 > 5N^{18} - N^4$ and each segment in S is of length at most N^2, we have $|S_{-1}| + |S_0| + |S_1| = |S| > \frac{5N^{18}-N^4}{N^2} = 5N^{16} - N^2$. Hence $|S_{-1}| + |S_1| > 5N^{16} - N^2 - 2N^{10}$. Then at least one of the two sets, either S_{-1} or S_1, has more than $\frac{5N^{16}-N^2-2N^{10}}{2}$ elements. By symmetry, it is sufficient to consider the case $|S_{-1}| > \frac{5N^{16}-N^2-2N^{10}}{2}$. This number is larger than N^{10}, so it remains to prove that also S_1 has at least N^{10} elements. By Lemma 7, $|d_{-1}| > \frac{5N^{16}-N^2-2N^{10}}{4N^4}$, so that $d_1 \geq |d_{-1}| - |d_0| - N^4 > \frac{5N^{16}-N^2-2N^{10}}{4N^4} - \frac{2N^{10}}{2} - N^4$. We can again apply Lemma 7 to see that $|S_1| \geq \frac{d_1}{N^2} > \frac{5N^{16}-N^2-2N^{10}}{4N^6} - N^6 - N^2$. However, it is easy to check that this number is larger than N^{10}. □

Lemma 9. *If $n > 5N^{18}$, then there are two disjoint sets P_1, P_2 of pairwise nonoverlapping, e-correct segments and two ordered pairs (p_1, q_1), (p_2, q_2) of nonnegative integers such that $|P_1| \geq N^6$, $|P_2| \geq N^6$, $(\mathrm{w}_e^\alpha(j) - \mathrm{w}_e^\alpha(i), \mathrm{w}_e^\beta(j) - \mathrm{w}_e^\beta(i)) = (p_1, q_1)$ for all $[i,j] \in P_1$, $(\mathrm{w}_e^\alpha(j) - \mathrm{w}_e^\alpha(i), \mathrm{w}_e^\beta(j) - \mathrm{w}_e^\beta(i)) = (p_2, q_2)$ for all $[i,j] \in P_2$, and $\frac{p_1}{p_1+q_1} \leq \Lambda'_e \leq \frac{p_2}{p_2+q_2}$.*

Proof. It follows easily from Lemma 8, since any set of N^{10} segments of length at most N^2 contains necessarily a subset of N^6 segments $[i,j]$ with identical pairs $(\mathrm{w}_e^\alpha(j) - \mathrm{w}_e^\alpha(i), \mathrm{w}_e^\beta(j) - \mathrm{w}_e^\beta(i))$. (These are ordered pairs (r,s) of nonnegative integers with $0 < r+s \leq N^2$, and one can easily see that the number of such ordered pairs is at most N^4.) □

4. Proof of Theorem 2 completed

Lemma 10. *The following two conditions are equivalent for a given quadruple of ordered pairs $(c_i, d_i) \neq (0,0)$ $(i = 1,2,3,4)$ of nonnegative integers such that $\frac{c_1}{c_1+d_1} \leq \frac{c_2}{c_2+d_2}$ and $\frac{c_3}{c_3+d_3} \leq \frac{c_4}{c_4+d_4}$:*

(1) *there exists a quadruple* $(n_1, n_2, n_3, n_4) \neq (0, 0, 0, 0)$ *of nonnegative integers such that* $n_1c_1 + n_2c_2 = n_3c_3 + n_4c_4$ *and* $n_1d_1 + n_2d_2 = n_3d_3 + n_4d_4$*;*
(2) *there is a rational number* r *such that* $\frac{c_1}{c_1+d_1} \leq r \leq \frac{c_2}{c_2+d_2}$ *and* $\frac{c_3}{c_3+d_3} \leq r \leq \frac{c_4}{c_4+d_4}$.

If (2) is satisfied, then the integers n_1, n_2, n_3, n_4 *can be always selected to be less or equal* m^3*, where* m *is the maximum of the numbers* c_i *and* d_i*.*

Proof. If (1) is satisfied, we can put

$$r = \frac{n_1c_1 + n_2c_2}{n_1c_1 + n_2c_2 + n_1d_1 + n_2d_2} = \frac{n_3c_3 + n_4c_4}{n_3c_3 + n_4c_4 + n_3d_3 + n_4d_4}.$$

Let (2) be satisfied. If $c_1d_4 = c_4d_1$, we can take either $(c_4, 0, 0, c_1)$ or $(d_4, 0, 0, d_1)$ for (n_1, n_2, n_3, n_4); at least one of the two quadruples is different from $(0, 0, 0, 0)$. Similarly, if $c_2d_3 = c_3d_2$, we can take either $(0, c_3, c_2, 0)$ or $(0, d_3, d_2, 0)$. If $c_1 = c_2 = c_3 = c_4$, we can take $(n_1, n_2, n_3, n_4) = (d_3, d_4, d_1, d_2)$. In all other cases we can take $n_1 = c_4(c_2d_3 - c_3d_2)$, $n_2 = c_3(c_4d_1 - c_1d_4)$, $n_3 = c_2(c_4d_1 - c_1d_4)$, $n_4 = c_1(c_2d_3 - c_3d_2)$; it follows from (2) that these numbers are nonnegative. □

In order to prove Theorem 2, it is obviously sufficient to show that for any positive integer n, any $(a, b) \in G^2$ and any $e, f \in E_n$ with the same weights and such that $e_n = \alpha$ and $f_n = \beta$,

$$(a * e)_{n-1}(b * f)_{n-1} = (b * e)_{n-1}(a * f)_{n-1}.$$

Suppose that this is not true and let n be the least positive integer for which there exist $(a, b) \in G^2$ and $e, f \in E_n$ giving a contradiction. According to the assumption, $n > 5N^{18}$.

By Lemma 4, there exist paths e' and f' similar to e and f respectively, such that there are a set S of pairwise nonoverlapping, e'-correct segments and a set T of pairwise nonoverlapping, f'-correct segments, both S and T of total length at least $n - N^4$.

We have $e_n = e'_n = \alpha$ and $f_n = f'_n = \beta$. Since $\mathrm{w}_e(n) = \mathrm{w}_f(n) = w_{e'}(n) = \mathrm{w}_{f'}(n)$, we have $\Lambda_e = \Lambda_f = \Lambda_{e'} = \Lambda_{f'}$, the four sets of middle rational numbers are the same for all these four paths, and also $\Lambda'_e = \Lambda'_f = \Lambda'_{e'} = \Lambda'_{f'}$; let us denote this number by Λ.

Now Lemma 9, applied to the path e', produces two sets P_1, P_2 of cardinalities at least N^6 and two ordered pairs $(p_1, q_1), (p_2, q_2)$; and applied to f', it similarly produces two sets P_3, P_4 and two ordered pairs $(p_3, q_3), (p_4, q_4)$. We have $0 < p_i + q_i \leq N^2$ $(i = 1, 2, 3, 4)$ and we have both $\frac{p_1}{p_1+q_1} \leq \Lambda \leq \frac{p_2}{p_2+q_2}$ and $\frac{p_3}{p_3+q_3} \leq \Lambda \leq \frac{p_4}{p_4+q_4}$. It follows by Lemma 10 that there is a quadruple $(n_1, n_2, n_3, n_4) \neq (0, 0, 0, 0)$ of nonnegative integers such that $n_1p_1 + n_2p_2 = n_3p_3 + n_4p_4$, $n_1q_1 + n_2q_2 = n_3q_3 + n_4q_4$ and $n_i \leq N^6$ $(i = 1, 2, 3, 4)$. Take n_1 segments $[r_i, s_i]$ in P_1 $(i = 1, \dots, n_1)$ and n_2 segments $[r_i, s_i]$ in P_2 $(i = n_1 + 1, \dots, n_1 + n_2)$ and denote by e'' the path obtained from e' by deleting all the members with indexes in one of the sets $\{r_{i+1},\dots,s_i\}$ $(i=1,\dots,n_1+n_2)$. This new path is of a length $m<n$, and its weight is $w_e(n)-(n_1p_1+n_2p_2,n_1q_1+n_2q_2)$. We can similarly obtain a path f'' from f'; its weight is $w_f(n)-(n_3p_3+n_4p_4,n_3q_3+n_4q_4)$, and we see that e'' and f'' are of the same weight. (In particular, f'' is of the same length $m<n$ as e''.) By the minimality of n, $(a*e'')_{m-1}(b*f'')_{m-1}=(b*e'')_{m-1}(a*f'')_{m-1}$. Since all the segments that have been 'squeezed to one point' during this process were correct (with respect to the appropriate paths), we have $(a*e)_{n-1}=(a*e')_{n-1}=(a*e'')_{m-1}$, $(b*e)_{n-1}=(b*e')_{n-1}=(b*e'')_{m-1}$, $(a*f)_{n-1}=(a*f')_{n-1}=(a*f'')_{m-1}$ and $(b*f)_{n-1}=(b*f')_{n-1}=(b*f'')_{m-1}$. It follows that $(a*e)_{n-1}(b*f)_{n-1}=(b*e)_{n-1}(a*f)_{n-1}$.

This completes the proof of Theorem 2.

5. An equivalent concept of Turing machine

The aim of the rest of this paper is to prove the following theorem.

Theorem 11. *There is no algorithm deciding for any finite partial groupoid whether it satisfies all equations of entropic groupoids.*

The basic idea is to encode an arbitrary Turing machine $\mathcal{T}$ into a finite partial groupoid $\mathbf{G}(\mathcal{T})$ in such a way that $\mathbf{G}(\mathcal{T})$ satisfies all entropic equations if and only if $\mathcal{T}$ does not halt. Since there is no algorithm deciding whether a Turing machine halts, we will get the desired undecidability. Firstly, we give an abstract (and slightly modified) definition of the Turing machine, and give a local characterization of the computation of $\mathcal{T}$. Then we set up $\mathbf{G}(\mathcal{T})$ carefully. Our ultimate goal is to show that if an entropic equation $p\approx q$ fails in $\mathbf{G}(\mathcal{T})$ then we can recover a halting computation in the term-trees. Failure of the equation means that we have some evaluation where both sides are defined but the values of the sides are not equal. These two facts will guarantee that the term-trees satisfy the local characterization of being a halting computation of $\mathcal{T}$. The local checkings are performed by the partial multiplication, yielding an undefined value whenever some error has been found. The rest of the paper will carry out this outline.

We need a slight modification of the concept of Turing machine, summarized as follows. We require that at each even step the machine's head (after writing the new value to the cell) either moves to the right or does not move, and at each odd step it either moves to the left or does not move. Also, the machine should never enter the initial state again after the start. So we have the following formal definitions.

Definition 12. A *Turing machine* $\mathcal{T}=(S,0,1,T)$ is a set S of states with two distinguished elements $0,1\in S$ together with a mapping $T:S\times\{0,1\}\times\{0,1\}\to(S\setminus\{1\})\times\{0,1\}\times\{0,1\}$. We call 0 the *halting state*, and 1 the *initial state*.

The mapping T describes the commands of $\mathcal{T}$ in the following way. Given any state $s \in S$, a digit $\epsilon \in \{0,1\}$ (the digit in the cell the head is pointing to) and the parity $\pi \in \{0,1\}$ of the step; the triple $T(s,\epsilon,\pi) = (s',\epsilon',\mu')$ contains the new state $s' \in S \setminus \{0\}$, the new digit $\epsilon' \in \{0,1\}$ (replacing ϵ in the cell the head was pointing to), and a number μ' describing the head's movement in the following way. For $\pi = 0$, the value $\mu' = 0$ means "stand," while $\mu' = 1$ means "move to the right." For $\pi = 1$, $\mu' = 0$ means "move to the left," while $\mu' = 1$ means "stand."

Definition 13. A *configuration* (t,p,s,π) of $\mathcal{T}$ is an element of $\{0,1\}^{\mathbb{Z}} \times \mathbb{Z} \times S \times \{0,1\}$. We call $t : \mathbb{Z} \to \{0,1\}$ the *tape*, $p \in \mathbb{Z}$ the *position*, $s \in S$ the *state* and $\pi \in \{0,1\}$ the *parity* of the configuration. The *initial configuration* is the quadruple $(\bar{0},0,1,0)$ with the constant 0 tape ($\bar{0}$ stands for the mapping of $\mathbb{Z}$ onto $\{0\}$.)

For any configuration the Turing machine $\mathcal{T}$ uniquely determines (computes) the next configuration. By iteration, starting from the initial configuration, we obtain a sequence of configurations, which will be called the computation of $\mathcal{T}$. We will be interested in whether the halting state $0 \in S$ appears in this sequence.

Definition 14. The *processor* for $\mathcal{T}$ is the mapping of the set of configurations into itself, denoted by $\mathcal{T}^*$ and defined as

$$\mathcal{T}^* : (t,p,s,\pi) \mapsto (t', p + \mu' - \pi, s', 1 - \pi),$$

where $(s',\epsilon',\mu') = T(s,t(p),\pi)$ and

$$t'(n) = \begin{cases} t(n) & \text{if } n \neq p, \\ \epsilon' & \text{otherwise.} \end{cases}$$

Definition 15. The *computation* of $\mathcal{T}$ is the mapping $\bar{\mathcal{T}}$ of $\mathbb{N}$ into the set of configurations defined recursively by

$$\bar{\mathcal{T}}(m) = \begin{cases} (\bar{0},0,1,0) & \text{if } m = 0, \\ \mathcal{T}^*(\bar{\mathcal{T}}(m-1)) & \text{if } m > 0. \end{cases}$$

We say that the Turing machine $\mathcal{T}$ *halts* if there exists a natural number m such that $\bar{\mathcal{T}}(m) = (t,p,0,\pi)$ for some t,p and π.

It is not hard to see that our concept of Turing machine is equivalent to the usually accepted definition. It is well known that there is no algorithm deciding whether a Turing machine halts. We will use this fact, and for each Turing machine $\mathcal{T}$ we construct a finite partial groupoid $\mathbf{G}(\mathcal{T})$ which satisfies all entropic equations if and only if $\mathcal{T}$ does not halt. More precisely, we will show that an entropic equation does not hold in $\mathbf{G}(\mathcal{T})$ if and only if we can

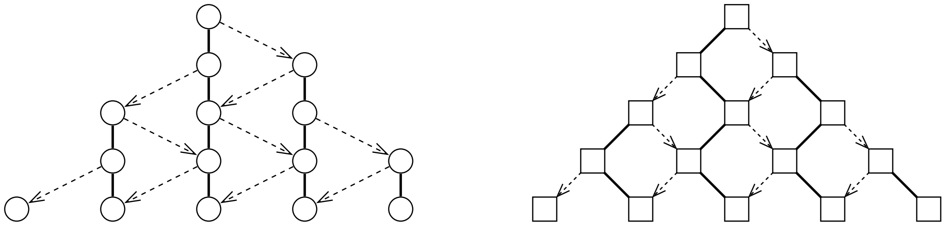

Figure 1. The triangle of computation of a Turing machine

find the halting computation encoded into the equation. To achieve this goal, we transform the computation into a tree-like shape.

We can write the consecutive configurations below each other, at each step taking into account only those cells of the tape to which any possible Turing machine's head can point. In this way, for $\mathcal{T}$ starting on the empty tape, we obtain a "triangle" as sketched on the left hand side of Figure 1. The circles connected by solid lines form the life-cycles of the cells. The dashed lines indicate the possible movements (right at even and left at odd steps). We can give a better shape to this "triangle" if we shift each odd row to the left a little bit (as pictured on the right hand side). In this way the life-cycle of a cell appears as a zig-zag, and the value $\mu' = 0$ (or 1) means that the head appears to move to the left (or right).

Definition 16. A *box* is an element (s, μ, ϵ, π) of the set $B = S \times \{0, 1\} \times \{0, 1\} \times \{0, 1\}$. A box (s, μ, ϵ, π) is called *tape-box* if $(s, \mu) = (1, 0)$, otherwise it is a *head-box* with state s.

A box (s, μ, ϵ, π) represents a box at the right hand side of Figure 1. The parity of the row is stored in π, while ϵ contains the digit of the corresponding cell. In most of the cases $(s, \mu) = (1, 0)$, which means that the head is not pointing to this cell (tape-box). If the head is pointing to the cell (head-box), then s stores the current state of $\mathcal{T}$ and μ stores the direction of the previous step. But for the initial state $s = 1$ there is no previous step, so we use $(s, \mu) = (1, 1)$ in this case (and that is why we can use the unused value $(1, 0)$ for tape-boxes).

Definition 17. Let Δ be $\mathbb{N} \times \mathbb{N}$ with the dual order. Thus $(0, 0)$ is the largest element in Δ, and $(i, j) \leq (k, l)$ if and only if $i \geq k$ and $j \geq l$. A *triangle* for $\mathcal{T}$ is a mapping of a finite upset $D \subset \Delta$ (i.e., a finite union of principal filters) into the set of boxes B.

As an example, a triangle is sketched on the right hand side of Figure 1. The finite upset $D \subset \Delta$ gives the shape of the triangle, while an element $(i, j) \in D$ identifies the unique box in the triangle with i many left turns and j many right turns.

Definition 18. The *number of initial states* in a triangle $A : D \to B$ is the cardinality of the set $\{(i,j) \in D \mid A(i,j) = (s,\mu,\epsilon,\pi) \text{ and } (s,\mu) = (1,1)\}$. The *number of halting states* is the cardinality of the set $\{(i,j) \in D \mid A(i,j) = (s,\mu,\epsilon,\pi) \text{ and } s = 0\}$.

Definition 19. A *valid triangle* for $\mathcal{T}$ is a triangle $A : D \to B$ satisfying the following ten conditions for any $(i,j) \in D$. Put $(s,\mu,\epsilon,\pi) = A(i,j)$ and $(s',\epsilon',\mu') = T(s,\epsilon,\pi)$. Put $(s_0,\mu_0,\epsilon_0,\pi_0) = A(i+1,j)$ whenever $(i+1,j) \in D$, and $(s_1,\mu_1,\epsilon_1,\pi_1) = A(i,j+1)$ whenever $(i,j+1) \in D$. The conditions are:

(1) $\pi = (i+j) \bmod 2$,
(2) if $i = 0$ or $j = 0$ then $(s,\mu) = (1,0)$ and $\epsilon = 0$,
(3) if $(s,\mu) = (1,0)$ then $\epsilon_\pi = \epsilon$,
(4) if $(s,\mu) \neq (1,0)$ then $\epsilon_\pi = \epsilon'$,
(5) if $(s,\mu) \neq (1,0)$ then $(s_{\mu'},\mu_{\mu'}) = (s',\mu')$,
(6) if $s_0 \neq 1$ and $\mu_0 = 0$ then $(s,\mu) \neq (1,0)$ and $\mu' = 0$,
(7) if $s_0 \neq 1$ and $\mu_0 = 1$ then $(s,\mu) = (1,0)$,
(8) if $s_1 \neq 1$ and $\mu_1 = 1$ then $(s,\mu) \neq (1,0)$ and $\mu' = 1$,
(9) if $s_1 \neq 1$ and $\mu_1 = 0$ then $(s,\mu) = (1,0)$,
(10) if $(s,\mu) = (1,1)$ then $\pi = 0$.

Note that $A(i+1,j)$ and $A(i,j+1)$ are not necessarily defined, and any condition involving undefined variables is considered as fulfilled.

It is easy to see that there are valid triangles. For example, take the almost constant triangle with D arbitrary and $A(i,j) = (1,0,0,\pi)$ for all $(i,j) \in D$ (the parity π must vary). Beside this trivial example, there are triangles containing starting pieces of the computation of $\mathcal{T}$. Basically, we can put the initial state to, or "start" the machine at, any position $(k,l) \in D$ which is not on the side and for which $\pi = 0$.

Example 20. For any given finite upset $D \subset \Delta$ and any pair of elements $(1,1) \geq (k,l) \in D$ such that $k + l = 0 \bmod 2$ we define a valid triangle $A_{k,l} : D \to B$ as

$$A_{k,l}(i,j) = \begin{cases} (1,0,0,\pi) & \text{if } (i,j) \not\leq (k,l), \\ (1,0,t(p),\pi) & \text{if } (i,j) \leq (k,l) \text{ and } p \neq n, \\ (1,1,0,\pi) & \text{if } (i,j) = (k,l) \text{ and } p = n, \\ (s,\mu,t(p),\pi) & \text{if } (i,j) < (k,l) \text{ and } p = n, \end{cases}$$

where $\pi = (i+j) \bmod 2$, $m = (i-k)+(j-l)$, $n = \lceil((j-l)-(i-k))/2\rceil$, $\bar{\mathcal{T}}(m) = (t,p,s,-)$, $\bar{\mathcal{T}}(m-1) = (t',p',s',\pi')$, and $T(s',t'(p'),\pi') = (s,t(p'),\mu)$. Note that these variables are not necessarily defined in all the cases, but they are defined in those cases when we need them in the definition. (It may be helpful to compare this definition with the illustration in Figure 2.)

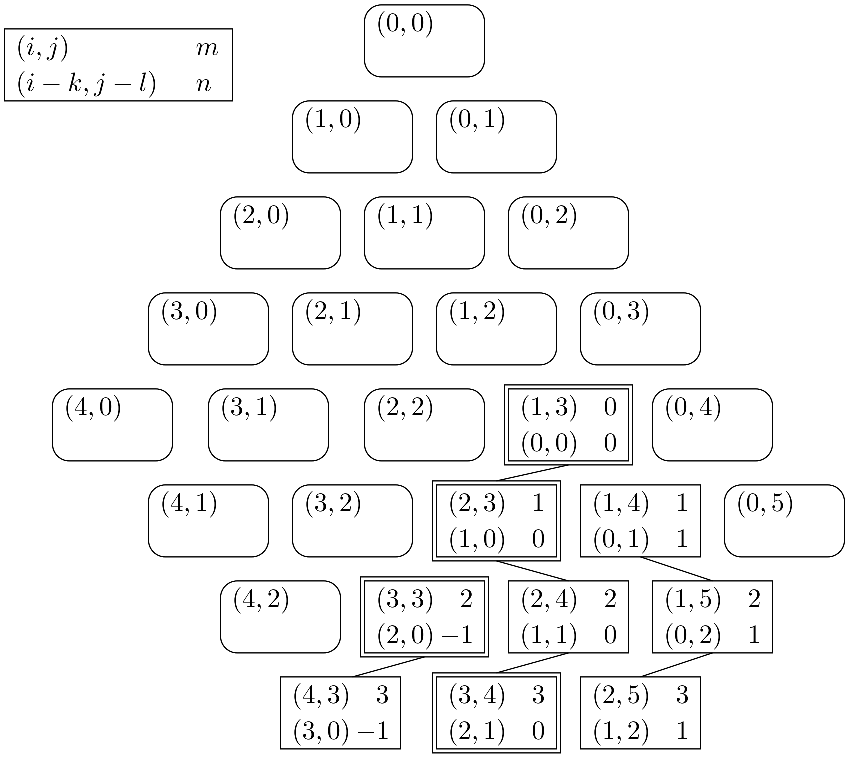

FIGURE 2. Various indices of boxes in a triangle when $(k, l) = (1, 3)$.

It is not very hard to see that the (local character) conditions (1) – (9) of Definition 19 guarantee that every valid triangle is one of these, with the exception that it may contain no initial state or more than one:

Lemma 21. *A triangle $A : D \to B$ with exactly one initial state is valid if and only if it is one of the valid triangles defined in Example 20.*

Proof. This is not difficult to verify, and we just indicate the proof. It may help to understand the intuitive idea if we translate each item of Definition 19 into human language:

(1) the parity π is correct everywhere;
(2) the head never points to the side of the triangle, and the sides contain only the digit 0;
(3) if this is a tape-box, then the next box along the life-cycle (zig-zag) of this cell contains the same digit (it cannot be overwritten);
(4) if this is a head-box, then the next box along the life-cycle (zig-zag) of this cell contains the digit what the Turing machine produces;
(5) if this is a head-box, then the Turing machine uniquely determines the next state and the direction of the movement, moreover the head moves either to the left or to the right according to the direction;

(6) if the south-west neighbor of a box is a head-box containing a non-initial state and a direction which indicates that the head came from this box, then this box must be a head-box, as well;

(7) if the south-west neighbor of a box is a head-box containing a non-initial state and a direction which indicates that the head did not come from this box, then this box must be a tape-box;

(8) the same as point (6) but for the other direction;

(9) the same as point (7) but for the other direction;

(10) every head-box with the initial state must be in an even row.

Items $(2), (6)$ and (8) guarantee that every head-box is connected up to the unique head-box containing the initial state. Items $(6), (7), (8)$ and (9) guarantee that the head-boxes form a descending chain. Now, by $(1), (3), (4)$ and (5), the triangle must contain the computation of $\mathcal{T}$ starting from the unique head-box with the initial state. Finally, $\mathcal{T}$ moves to the left or stays (stays or moves to the right) in odd steps (in even steps) according to (10). □

6. The partial groupoid $\mathbf{G}(\mathcal{T})$

Now we are ready to define the partial groupoid $\mathbf{G}(\mathcal{T})$ for any given Turing machine $\mathcal{T}$. First of all, we will describe the elements of the groupoid, then define the partial multiplication. We need the following variables for indexing:

$s \in S$	the state in the box
$\mu \in \{0, 1\}$	the direction of the last step
$\epsilon \in \{0, 1\}$	the digit in the box
$\pi \in \{0, 1\}$	the parity of the row
$\iota \in \{-1, 0, 1\}$	a flag indicating the presence of the initial state
$\chi \in \{-1, 0, 1\}$	a flag indicating the presence of the halting state
$\delta \in \{-1, 0, 1\}$	the direction of the "checking"

Each element in $\mathbf{G}(\mathcal{T})$ belongs to one of the following classes denoted by letters: $i, t, l_1, l_2, l_3, l_4, l_5, l_6, r_1, r_2, r_3, r_4, r_5, r_6, a, b, c, d, e, f, g$. In each class the elements are indexed by values of the variables $(s, \mu, \epsilon, \pi, \iota, \chi, \delta)$ defined above. Many classes do not use all of these variables for indexing. For example, class t has only one index δ, and we denote the two elements of this class by $t(-1)$ and $t(1)$. Note that, as you can see in the following list, δ cannot be 0 in this class. In class i there is only one element (which will be irreducible), so it has no index at all. As another example, class l_1 contains 16 elements. Here are the elements:

i	
$t(\delta)$	$\delta \in \{-1, 1\}$
$l_n(\pi, \iota, \chi, \delta)$	$\iota, \chi, \delta \in \{0, 1\}$, $n = 1, 2, 3, 4, 5$
$l_6(\pi)$	

$r_n(\pi, \iota, \chi, \delta)$	$\iota, \chi, \delta \in \{-1, 0\}$, $n = 1, 2, 3, 4, 5$
$r_6(\pi)$	
$a(s, \mu, \epsilon, \pi, \iota, \chi, \delta)$	$\iota, \chi, \delta \in \{-1, 0, 1\}$
$b(s, \mu, \epsilon, \pi, \iota, \chi, \delta)$	$\iota, \chi, \delta \in \{-1, 0\}$
$c(s, \mu, \epsilon, \pi, \iota, \chi, \delta)$	$\iota, \chi, \delta \in \{0, 1\}$
$d(s, \mu, \epsilon, \pi, \iota, \chi, \delta)$	$\iota, \chi, \delta \in \{-1, 0\}$
$e(s, \mu, \epsilon, \pi, \iota, \chi, \delta)$	$\iota, \chi, \delta \in \{0, 1\}$
$f(s, \mu, \epsilon, \pi)$	
$g(s, \mu, \epsilon, \pi)$	

Now we are ready to define the partial multiplication of $\mathbf{G}(\mathcal{T})$. We will list only the defined products, and sort them accordingly to the class of the resulting element. In this way we can easily see to which products can the elements be reduced. We encourage the reader to locate the list of defined products in Figure 3. Since our variables for indexing have overlapping ranges, and many classes have various indexes, it may be helpful to indicate the variable beside the value. So the expression $a_{:b}$ stands for the value a and means that a is to be considered as a value of the indexing variable b.

The element i is irreducible. The open circles in Figure 3 are meant to be labeled by this element. Both elements of class t have a unique decomposition:

$$\begin{gathered} t(-1_{:\delta}) = l_1(0_{:\pi}, 0_{:\iota}, 0_{:\chi}, 0_{:\delta}) \cdot r_1(0_{:\pi}, -1_{:\iota}, -1_{:\chi}, -1_{:\delta}), \\ t(1_{:\delta}) = l_1(0_{:\pi}, 1_{:\iota}, 1_{:\chi}, 1_{:\delta}) \cdot r_1(0_{:\pi}, 0_{:\iota}, 0_{:\chi}, 0_{:\delta}). \end{gathered}$$

Every element in classes l_1, l_2, l_4 and r_1, r_2, r_4 has a unique decomposition, where we preserve the indexes:

$$\begin{gathered} l_n(\pi, \iota, \chi, \delta) = l_{n+1}(\pi, \iota, \chi, \delta) \cdot i \qquad \text{for } n = 1, 2, \\ l_4(\pi, \iota, \chi, \delta) = i \cdot l_5(\pi, \iota, \chi, \delta), \\ r_n(\pi, \iota, \chi, \delta) = i \cdot r_{n+1}(\pi, \iota, \chi, \delta) \qquad \text{for } n = 1, 2, \\ r_4(\pi, \iota, \chi, \delta) = r_5(\pi, \iota, \chi, \delta) \cdot i. \end{gathered}$$

Elements in class l_3 (and r_3) can have one or more decompositions. The decompositions are always products of elements from l_1 and l_4 (r_4 and r_1), where the elements from l_1 (r_1) will always have the other parity. Here we combine the information from the factors. The product will be not defined if $\iota_0 = \iota_1 \neq 0$.

$$\begin{gathered} l_3(\pi, \iota_0 + \iota_1, \max(\chi_0, \chi_1), \delta) = l_1(1 - \pi, \iota_0, \chi_0, 0_{:\delta} \text{ or } \delta) \cdot l_4(\pi, \iota_1, \chi_1, \delta) \\ r_3(\pi, \iota_0 + \iota_1, \min(\chi_0, \chi_1), \delta) = r_4(\pi, \iota_0, \chi_0, \delta) \cdot r_1(1 - \pi, \iota_1, \chi_1, 0_{:\delta} \text{ or } \delta) \end{gathered}$$

In class l_5 (and r_5) we have two types of decompositions. If $\delta = 0$ (no direction), then $l_5(\pi, \iota, \chi, 0_{:\delta})$ is reducible if and only if $\iota = 0$ and $\chi = 0$ (no initial

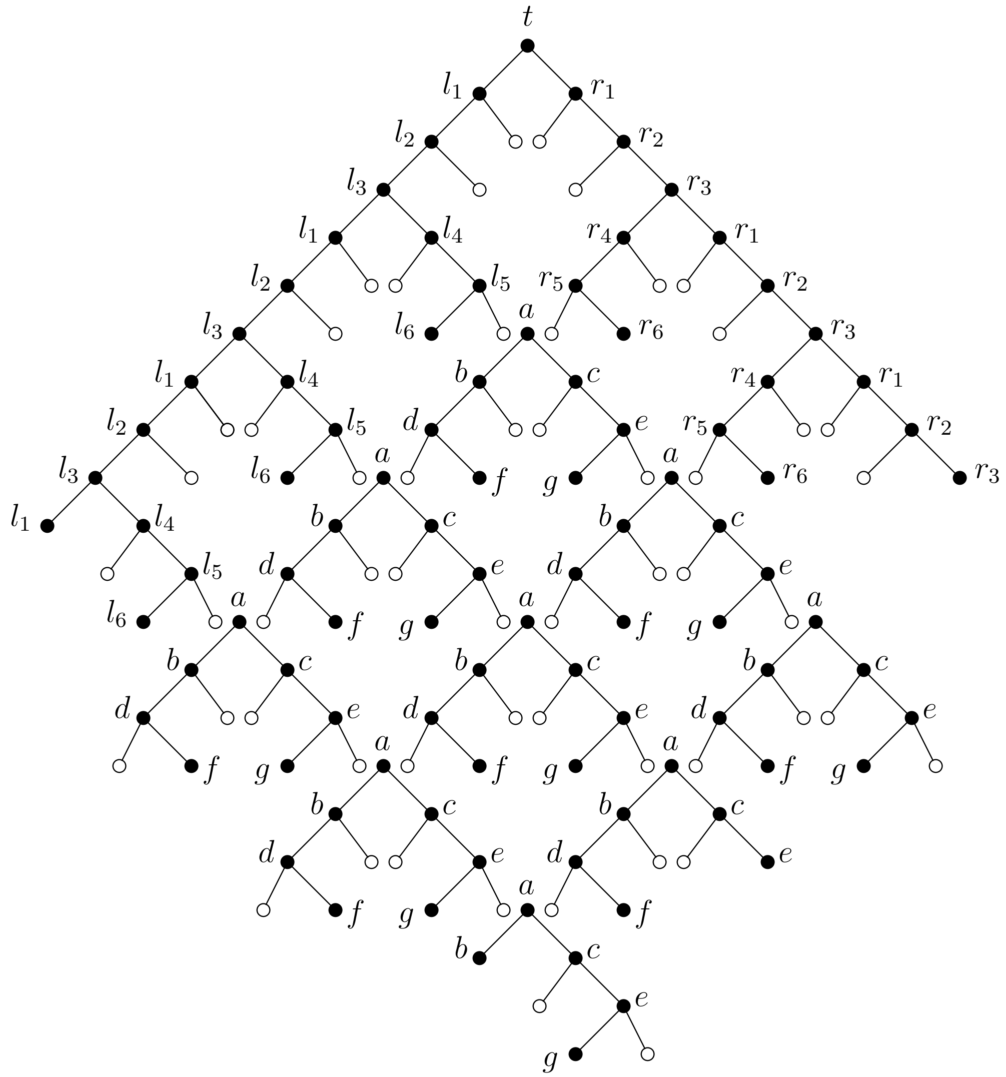

Figure 3. Decompositions of elements

or halting state):

$$l_5(\pi, 0_{:\iota}, 0_{:\chi}, 0_{:\delta}) = l_6(\pi) \cdot i,$$
$$r_5(\pi, 0_{:\iota}, 0_{:\chi}, 0_{:\delta}) = i \cdot r_6(\pi).$$

If $\delta \neq 0$, then the element is always reducible to a product of elements from l_6 and a in two different ways. Note that, according to our assumptions, we must have the digit 0 and the no-state along the sides of the triangle of computation.

$$l_5(\pi, \iota, \chi, 1_{:\delta}) = l_6(\pi) \cdot a(1_{:s}, 0_{:\mu}, 0_{:\epsilon}, 1 - \pi, \iota, \chi, 0_{:\delta} \text{ or } 1_{:\delta}),$$
$$r_5(\pi, \iota, \chi, -1_{:\delta}) = a(1_{:s}, 0_{:\mu}, 0_{:\epsilon}, 1 - \pi, \iota, \chi, 0_{:\delta} \text{ or } -1_{:\delta}) \cdot r_6(\pi).$$

Note that there are more possible values of ι and χ for class a than for classes l_5 and r_5. Thus the products above are not defined if the value of ι or χ is not in the proper range.
The elements of l_6 and r_6 are irreducible. So up to this point we have defined all products that yield elements on the "sides" (classes t, $l_1, \dots, l_6$, $r_1, \dots, r_6$).

Now we are ready to define the decomposition of elements in classes $a, \dots, g$. For class a the following products are defined:

$$\begin{aligned} a(s,\mu,\epsilon,\pi,\iota,\chi,\delta) = b(s,\mu,\epsilon,\pi,\min(\iota,0),\min(\chi,0),\min(\delta,0)) \cdot \\ c(s,\mu,\epsilon,\pi,\max(\iota,0),\max(\chi,0),\max(\delta,0)). \end{aligned}$$

Every element in classes b and c has a unique decomposition:

$$\begin{aligned} b(s,\mu,\epsilon,\pi,\iota,\chi,\delta) &= d(s,\mu,\epsilon,\pi,\iota,\chi,\delta) \cdot i, \\ c(s,\mu,\epsilon,\pi,\iota,\chi,\delta) &= i \cdot e(s,\mu,\epsilon,\pi,\iota,\chi,\delta). \end{aligned}$$

In class d (and e) we have two types of decompositions. If $\delta = 0$ (no direction) then $d(s,\mu,\epsilon,\pi,\iota,\chi,\delta)$ $(e((s,\mu,\epsilon,\pi,\iota,\chi,\delta))$ is reducible if and only if $\iota = 0$ and $\chi = 0$:

$$\begin{aligned} d(s,\mu,\epsilon,\pi,0_{:\iota},0_{:\chi},0_{:\delta}) &= i \cdot f(s,\mu,\epsilon,\pi), \\ e(s,\mu,\epsilon,\pi,0_{:\iota},0_{:\chi},0_{:\delta}) &= g(s,\mu,\epsilon,\pi) \cdot i. \end{aligned}$$

If $\delta \neq 0$ then the product

$$\begin{aligned} & d(s,\mu,\epsilon,\pi,\iota+\iota',\min(\chi,\chi'),-1_{:\delta}) = \\ & \qquad a(s_0,\mu_0,\epsilon_0,1-\pi,\iota,\chi,0_{:\delta} \text{ or } -1_{:\delta}) \cdot f(s,\mu,\epsilon,\pi) \end{aligned}$$

is defined whenever

$$\iota' = \begin{cases} -1 & \text{if } (s_0,\mu_0) = (1,1), \\ 0 & \text{otherwise;} \end{cases} \qquad \chi' = \begin{cases} -1 & \text{if } s_0 = 0, \\ 0 & \text{otherwise;} \end{cases}$$

and the following conditions hold:

(1) $-1 \leq \iota, \chi \leq 0$,
(2) $\iota \neq -1$ or $\iota' \neq -1$,
(3) if $\iota' = -1$ then $\pi = 1$,
(4) the conditions (3) – (10) in Definition 19 where $T(s,\epsilon,\pi) = (s',\epsilon',\mu')$.

The decomposition of $e(s,\mu,\epsilon,\pi,\iota,\chi,1)$ is defined analogously:

$$\begin{aligned} & e(s,\mu,\epsilon,\pi,\iota+\iota',\max(\chi,\chi'),1_{:\delta}) = \\ & \qquad g(s,\mu,\epsilon,\pi) \cdot a(s_1,\mu_1,\epsilon_1,1-\pi,\iota,\chi,0_{:\delta} \text{ or } 1_{:\delta}) \end{aligned}$$

whenever

$$\iota' = \begin{cases} 1 & \text{if } (s_1,\mu_1) = (1,1), \\ 0 & \text{otherwise;} \end{cases} \qquad \chi' = \begin{cases} 1 & \text{if } s_1 = 0, \\ 0 & \text{otherwise;} \end{cases}$$

and the following conditions hold:

(1) $0 \leq \iota, \chi \leq 1$,
(2) $\iota \neq 1$ or $\iota' \neq 1$,
(3) if $\iota' = 1$ then $\pi = 1$,
(4) the conditions $(3) - (10)$ in Definition 19 where $T(s, \epsilon, \pi) = (s', \epsilon', \mu')$.

And finally, every element in classes f and g is irreducible. Up to this point we have listed all defined products, and finished the definition of the partial groupoid $\mathbf{G}(\mathcal{T})$.

7. Facts

In this section we give a series of facts, each of which can be proved easily from the definition of $\mathbf{G}(\mathcal{T})$. Then we prove our main theorem.

Fact 22. *In each defined product the classes of the factors uniquely determine the class of the product. Moreover, the values of the factors' indexing variables determine the values of the product's indexes. Thus the partial multiplication of* $\mathbf{G}(\mathcal{T})$ *is well defined.* □

Fact 23. *The class of any defined product uniquely determines the classes of the two factors, except for classes* l_5*,* r_5*,* d *and* e*. If the product lies in one of these classes, then either both factors are irreducible or one is irreducible and the other is in class* a*.* □

Definition 24. A *decomposition tree* of an element $r \in G(\mathcal{T})$ is a finite binary tree labeled by elements of $\mathbf{G}(\mathcal{T})$ such that the root is labeled r and each non-leaf node is labeled by the product of its children's labels.

Definition 25. Let R be any decomposition tree. Remove all nodes labeled by i, $t(-1)$ or $t(1)$ together with the connecting edges. Also, remove all edges between nodes l_3–l_1, r_3–r_1, l_5–a, r_5–a, d–a and e–a (but not the nodes). An *l-block* (*r-block*) is a component of the rest of the graph labeled only by elements from classes $l_1, \ldots, l_6$ $(r_1, \ldots, r_6)$. Every other block is an *a-block* labeled by elements from classes $a, \ldots, g$. An l or r-block is *full* if it has 6 nodes. For an a-block to be full we require 7 nodes.

Fact 26. *In each block the nodes have the same parity (the* π*-index of the labeling elements). In each* a*-block the nodes have the same* s*,* μ *and* ϵ*-indexes.* □

Fact 27. *The parities of any two adjacent blocks (connected by a single edge in the tree) are different.* □

Fact 28. *Let* r *be an element of* $\mathbf{G}(\mathcal{T})$ *which has a* χ*-index* χ*, and* R *be a decomposition tree of* r*. If* $\chi = 0$ *then every element in* R *has* χ*-index* 0*, and no element has* s*-index* 0*. If* $\chi \neq 0$ *then no element has* χ*-index* $-\chi$*, and there is either a node labeled by* $a(0_{:s}, -, -, -, -, -, -)$ *or a leaf with* χ*-index* χ*.* □

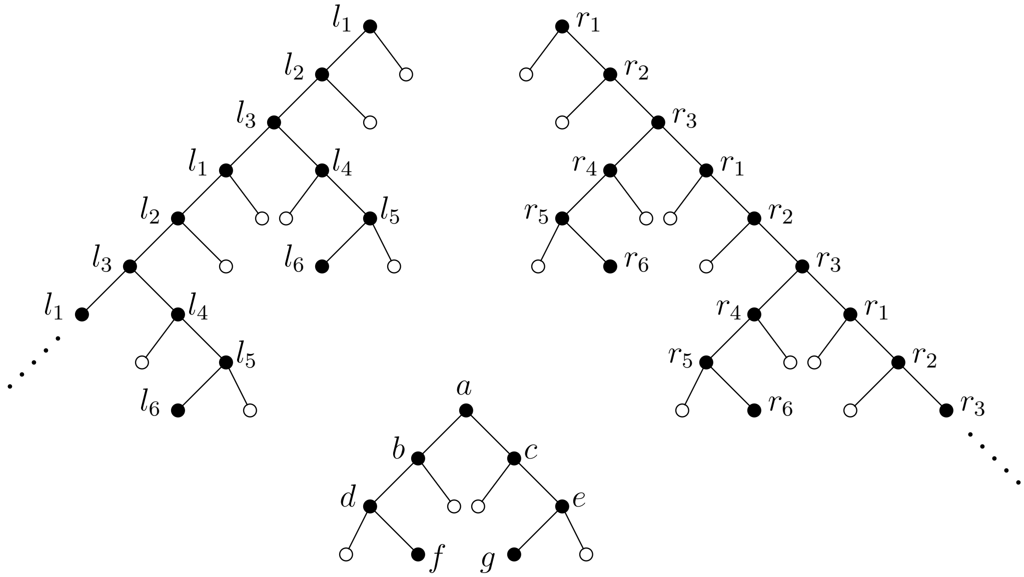

Figure 4. Decompositions of elements with δ-index 0

Fact 29. *Let r be an element of $\mathbf{G}(\mathcal{T})$ which has a ι-index ι, and R be a decomposition tree of r. If $\iota = 0$ then every element in R has ι-index 0, and no element has (s, μ)-index $(1, 1)$. If $\iota \neq 0$ then no element has ι-index $-\iota$, and there is either exactly one node labeled by $a(1_{:s}, 1_{:\mu}, -, -, -, -, -)$ or there is exactly one leaf with ι-index ι.* □

Fact 30. *In each l or r-block the nodes have the same δ-indexes. In an a-block if top node (class a) has δ-index 0, then all nodes in the block have δ-index 0. If the top node has δ-index 1, then the nodes in classes c and e have δ-index 1, and in classes b and d index 0 (f and g have no δ-index). Moreover, the node in class e decomposes into a product of a g-node and an a-node. A similar statement is true for δ-index -1.* □

Fact 31. *Let r be an element of $\mathbf{G}(\mathcal{T})$ which has a δ-index δ, and R be a decomposition tree of r. Then every element in R has δ-index 0 or δ.* □

Fact 32. *Let R be a decomposition tree for an element $r \in G(\mathcal{T})$. If R contains an edge between nodes l_5–a or e–a, then r has δ-index 1. If R contains an edge between nodes r_5–a or d–a, then r has δ-index -1.* □

Fact 33. *Let $r \in G(\mathcal{T})$ be an element which has δ-index 0, and R be a decomposition tree for r. Then R is a subtree of one of the three trees in Figure 4.* □

Fact 34. *Let $r \in G(\mathcal{T})$ be an element which has δ-index 1, and R be a decomposition tree for r. Then R is a subtree of the tree in Figure 5. Furthermore,*

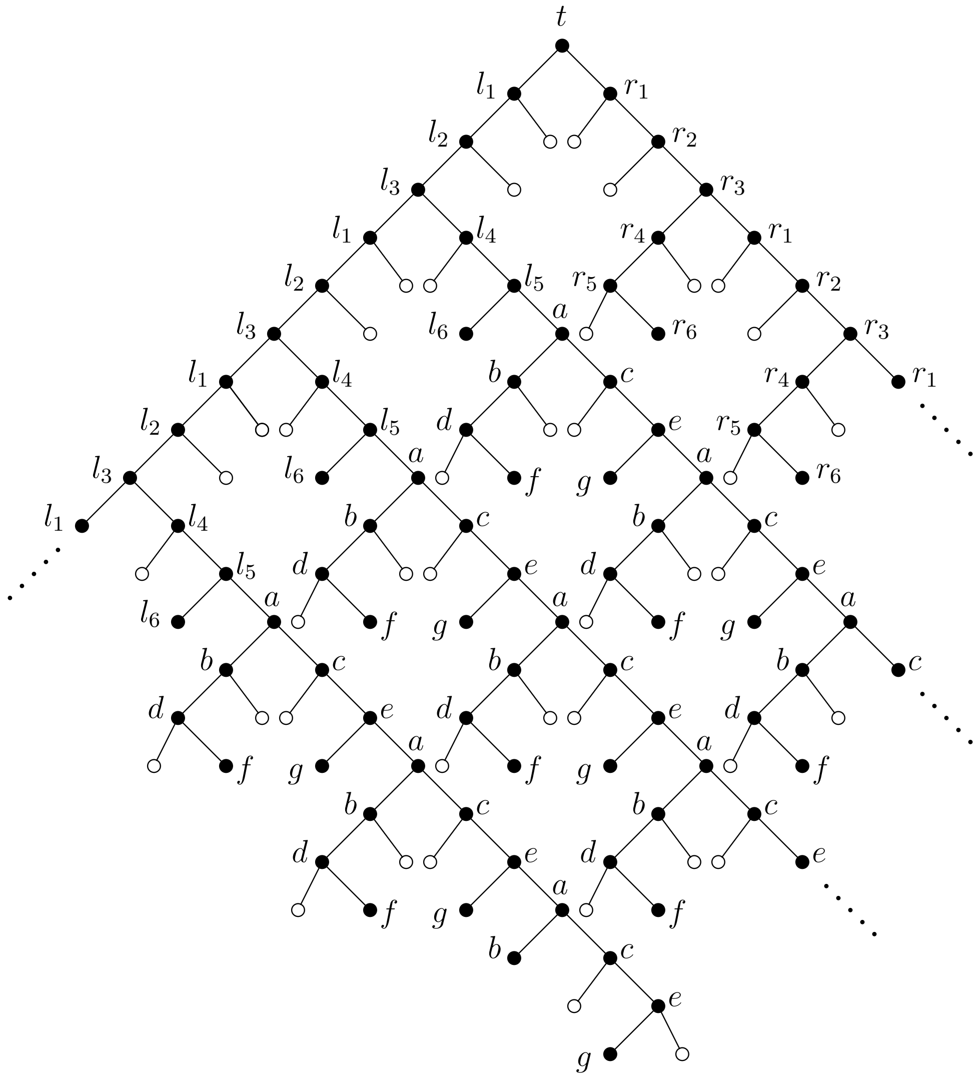

FIGURE 5. Decompositions of elements with δ-index 1

the nodes with δ-index 1 *form a subtree of* R *(top part), and the other pieces contain only nodes with δ-index* 0*. A similar statement is true for δ-index* -1. □

Fact 35. *Let R be any decomposition tree. Let us change pairs of nodes in R which have the same weight, in such a way that all products are defined in the obtained decomposition tree R'. Then R' has the same shape as R, and every non-leaf node in R' is in the same class as in R, and has the same s, μ, ϵ and π indexes.* □

Now we restate our main theorem as

Theorem 36. *$\mathcal{T}$ halts if and only if $\mathbf{G}(\mathcal{T})$ satisfies all entropic equation.*

Proof. Suppose that an entropic equation $p \approx q$ in n variables fails in $\mathbf{G}(\mathcal{T})$. Thus there exists an evaluation $\bar{u} \in G(\mathcal{T})^n$ such that both $p(\bar{u})$ and $q(\bar{u})$ are defined but $p(\bar{u}) \neq q(\bar{u})$. Put $r = p(\bar{u})$, $r' = q(\bar{u})$, and let R, R' be the decomposition trees of r and r' corresponding to p and q, respectively.

Claim 1. *The elements r and r' are exactly the two elements of the class t.*

If r has δ-index 0, then by Fact 33 R is a subtree of one of the trees in Figure 4. But in all of these trees there are no two different nodes with the same weight (number of left-right turns), except for adjacent leaves (variables) both evaluated with the same irreducible element i. Clearly, interchanging these variables does not yield a different final result, which contradicts to $p(\bar{u}) \neq q(\bar{u})$. Thus r cannot have δ-index 0.

Now suppose that r has δ-index 1 and is not in class t. By Fact 34 R must be a subtree of the tree in Figure 5. This implies that if R contains an r-block, then the whole tree must consist only of r-blocks. But the same argument as for δ-index 0 shows that this cannot happen. Similarly, we see that R must contain at least one a-block. Let us take the right-most a-node in R, and denote it by o. This node must be connected to the left with either an l_5–a or e–a edge. For o there is no other node with the same weight. Thus o cannot be interchanged with another node in R'. But in R' the node o must also be an a-node by Fact 35, so it must be connected to the left with either an l_5–a or e–a edge, as well. Then by Fact 32 every a-node in both R and R' must be connected to the left. This shows that we cannot interchange different nodes with the same weight, except for the adjacent leaves which are evaluated with the same i anyway. This contradicts again with $r \neq r'$.

A similar argument works for δ-index -1. This proves that both r and r' must be in class t.

Claim 2. *Every element in $\bar{u}$ has zero δ, χ and ι-indices.*

Suppose the contrary. Then by Facts 28, 29 and 31 the same non-zero index must appear in the decompositions of both r and r'. These two elements are the elements of class t, and they decompose into elements with indexes 1 and -1. But a non-zero index cannot be 1 and -1 at the same time.

Suppose that $r = t(1_{:\delta})$ and $r' = t(-1_{:\delta})$. Now the following claim is a corollary of Facts 30, 31, 32 and Claim 2:

Claim 3. *Every a-node in R is connected to the left with either an l_5–a or e–a edge. Every a-node in R' is connected to the right with either an r_5–a or d–a edge. Thus the a-blocks form a finite upset of Δ.*

Now we can see how to transform the decomposition tree R into a valid triangle $A : D \to B$. The a-blocks form an upset of Δ, and we assign the s,

μ, ϵ and π-indexes of an a-box, which are constant in the block by Fact 26, to the corresponding box. By Fact 35 we know that this assignment is the same for both sides. By Fact 29 the ι-index 1 in

$$t(1_{:\delta}) = l_1(0_{:\pi}, 1_{:\iota}, 1_{:\chi}, 1_{:\delta}) \cdot r_1(0_{:\pi}, 0_{:\iota}, 0_{:\chi}, 0_{:\delta}).$$

shows that there is exactly one initial state in the triangle. The χ-index 1 of r_1 shows that there is at least one halting state in the triangle. The l-classes on the left hand side of R guarantee that every a-class on the left-hand side is a tape-box containing the empty digit. The r-classes of R' do the same job on the other side. Since every a-block is connected to the north-west (in R) and to the north-east neighbor (in R') all the local checkings are carried out in the decomposition of elements in classes d and e. All these together prove that the defined triangle is valid and contains exactly one initial state. Now by Lemma 21 we have a halting computation of T.

The converse statement, that is, if $\mathcal{T}$ halts then we have some entropic equation $p \approx q$ which fails in $\mathbf{G}(\mathcal{T})$, is quite obvious. We just have to construct a large enough tree like the one in Figure 5 for p. We put different variables to each leaves of the tree. We get q if we interchange each occurrence of a with the unique i which has the same weight. Now the evaluation we want to check is the one obtained from Example 20 with ι, χ and δ-index 0. □

With the same construction and a slightly more subtle argument one can prove the following theorem, as well:

Theorem 37. *There is no algorithm deciding for any finite partial groupoid whether it can be embedded into an entropic groupoid.*

References

[1] J. Ježek, *A decidable equational theory with undecidable membership problem for finite algebras.* (To appear in Algebra Universalis.)

[2] J. Ježek and T. Kepka, *Medial groupoids.* Rozpravy ČSAV, Řada mat. a přír. věd **93/2** (1983), 93 pp.

[3] J. Ježek and T. Kepka, *Equational theories of medial groupoids.* Algebra Universalis **17** (1983), 174–190.

[4] J. Ježek and T. Kepka, *Free entropic groupoids.* Commentationes Math. Univ. Carolinae **22** (1981), 223–233.

[5] J. Ježek and T. Kepka, *Semigroup representations of medial groupoids.* Commentationes Math. Univ. Carolinae **22** (1981), 513–524.

[6] J. Ježek and T. Kepka, *Medial division groupoids.* Proceedings of the American Mathematical Society **119** (1993), 423–426.

[7] R. McKenzie, G. McNulty and W. Taylor, *Algebras, Lattices, Varieties, Volume I.* Wadsworth & Brooks/Cole, Monterey, CA, 1987.

[8] G. Pollák and Á. Szendrei, *Independent basis for the identities of entropic groupoids.* Commentationes Math. Univ. Carolinae **22** (1981), 71–85.

Charles University, Praha, Czech Republic

Vanderbilt University, Nashville, TN 37240