DECORATED LINEAR ORDER TYPES AND THE THEORY OF CONCATENATION

VEDRAN ČAČIĆ, PAVEL PUDLÁK, GREG RESTALL, ALASDAIR URQUHART, AND ALBERT VISSER

Abstract. We study the interpretation of Grzegorczyk's Theory of Concatenation TC in structures of decorated linear order types satisfying Grzegorczyk's axioms. We show that TC is incomplete for this interpretation. What is more, the first order theory validated by this interpretation interprets arithmetical truth. We also show that every extension of TC has a model that is not isomorphic to a structure of decorated order types. We provide a positive result, to wit a construction that builds structures of decorated order types from models of a suitable concatenation theory. This construction has the property that if there is a representation of a certain kind, then the construction provides a representation of that kind.

§**1. Introduction.** In his paper [2], Andrzej Grzegorczyk introduces a theory of concatenation TC. The theory has a binary function symbol $*$ for concatenation and two constants $\mathtt{a}$ and $\mathtt{b}$. The theory is axiomatized as follows.

TC1. $\vdash (x * y) * z = x * (y * z)$
TC2. $\vdash x * y = u * v \rightarrow ((x = u \wedge y = v) \vee \exists w\ ((x*w = u \wedge y = w*v) \vee (x = u*w \wedge y*w = v)))$
TC3. $\vdash x * y \neq \mathtt{a}$
TC4. $\vdash x * y \neq \mathtt{b}$
TC5. $\vdash \mathtt{a} \neq \mathtt{b}$

Axioms TC1 and TC2 are due to Tarski [7]. Grzegorczyk calls axiom TC2 the *editor axiom*. We will consider two weaker theories. The theory TC_0 has the signature with just concatenation, and is axiomatized by TC1,2. The theory TC_1 is axiomatized by TC1,2,3. We will also use TC_2 for TC.

Some of the results of this note were obtained during the Excursion to mountain Ślęża of the inspiring Logic Colloquium 2007 in Wrocław and, in part, in the evening after the Excursion. We thank the organizers for providing this wonderful opportunity. We thank Dana Scott for his comments, insights and questions. We are grateful to Vincent van Oostrom for some perceptive remarks.

Pavel Pudlák was supported by grants A1019401 and 1M002162080.

Logic Colloquium '07
Edited by Françoise Delon, Ulrich Kohlenbach, Penelope Maddy, and Frank Stephan
Lecture Notes in Logic, 35

The theories we are considering have various interesting interpretations. First they are, of course, theories of strings with concatenation; in other words, theories of free semigroups. Secondly they are theories of wider classes of structures, to wit structures of *decorated linear order types*, which will be defined below[1].

The theories TC_i are theories for concatenation without the empty string, i.e., without the unit element. Adding a unit ε one obtains another class of theories $\mathsf{TC}_i^\varepsilon$, theories of free monoids, or theories of structures of decorated linear order types including the empty linear decorated order type. The basic list of axioms is as follows.

$\mathsf{TC}^\varepsilon 1.\ \vdash \varepsilon * x = x \wedge x * \varepsilon = x$
$\mathsf{TC}^\varepsilon 2.\ \vdash (x * y) * z = x * (y * z)$
$\mathsf{TC}^\varepsilon 3.\ \vdash x*y = u*v \rightarrow \exists w\ ((x*w = u \wedge y = w*v) \vee (x = u*w \wedge w*y = v))$
$\mathsf{TC}^\varepsilon 4.\ \vdash \mathtt{a} \neq \varepsilon$
$\mathsf{TC}^\varepsilon 5.\ \vdash x * y = \mathtt{a} \rightarrow (x = \varepsilon \vee y = \varepsilon)$
$\mathsf{TC}^\varepsilon 6.\ \vdash \mathtt{b} \neq \varepsilon$
$\mathsf{TC}^\varepsilon 7.\ \vdash x * y = \mathtt{b} \rightarrow (x = \varepsilon \vee y = \varepsilon)$
$\mathsf{TC}^\varepsilon 8.\ \vdash \mathtt{a} \neq \mathtt{b}.$

We take $\mathsf{TC}_0^\varepsilon$ to be the theory axiomatized by $\mathsf{TC}^\varepsilon 1, 2, 3$, $\mathsf{TC}_1^\varepsilon$ to be $\mathsf{TC}_0^\varepsilon + \mathsf{TC}^\varepsilon 4, 5$ and $\mathsf{TC}^\varepsilon := \mathsf{TC}_2^\varepsilon$ to be $\mathsf{TC}_1^\varepsilon + \mathsf{TC}^\varepsilon 6, 7, 8$.

One can show that TC is *bi-interpretable* with TC^ε, in which a unit ε is added via one dimensional interpretations without parameters. The theory TC_1 is bi-interpretable with $\mathsf{TC}_1^\varepsilon$ via two-dimensional interpretations with parameters. The situation for TC_0 seems to be more subtle. See also [10]. In Section 6, we will study an extension of $\mathsf{TC}_0^\varepsilon$.

Andrzej Grzegorczyk and Konrad Zdanowski have shown that TC is essentially undecidable. This result can be strengthened by showing that Robinson's Arithmetic Q is mutually interpretable with TC. Note that TC_0 is undecidable —since it has an extension that parametrically interprets TC— but that TC_0 is not essentially undecidable: it is satisfied by a one-point model. Similarly TC_1 is undecidable, but it has as an extension the theory of finite strings of a's, which is a notational variant of Presburger Arithmetic and, hence, decidable.

We will call models of TC_0 *concatenation structures*, and we will call models of TC_i *concatenation i-structures*. The relation of isomorphism between concatenation structures will be denoted by $\cong$. We will be interested in concatenation structures, whose elements are decorated linear order types with the operation *concatenation of decorated order types*. Let a non-empty class A be given. An A-decorated linear ordering is a structure $\langle D, \leq, f \rangle$, where D is a non-empty domain, $\leq$ is a linear ordering on D, and f is a function from

[1]A special case of decorated linear order types is addition of sets as discovered by Tarski (see [8]). It is shown by Laurence Kirby in [3] that the structure of addition on sets is isomorphic to addition of well-founded order types with a proper class of decorating objects.

D to A. A mapping ϕ is an *isomorphism* between A-decorated linear order types $\langle D, \leq, f \rangle$ and $\langle D', \leq', f' \rangle$ iff it is a bijection between D and D' such that, for all d, e in D, $d \leq e \Leftrightarrow \phi d \leq' \phi e$, and $fd = f'\phi d$. Our notion of isomorphism gives us a notion of A-decorated linear order type. We have an obvious notion of concatenation between A-decorated linear orderings which induces a corresponding notion of concatenation for A-decorated linear order types. We use $\alpha, \beta, \ldots$ to range over such linear order types. Since, linear order types are classes we have to follow one of two strategies: either to employ Scott's trick to associate a set object to any decorated linear order type or to simply refrain from dividing out isomorphism but to think about decorated linear orderings modulo isomorphism. We will employ the second strategy.

We will call a concatenation structure whose domain consists of (representatives of) A-decorated order types, for some A, and whose concatenation *is* concatenation of decorated order types: a *concrete concatenation structure*. It seems entirely reasonable to stipulate that e.g. the interpretation of a in a concrete concatenation structure is a decorated linear order type of a one element order. However, for the sake of generality we will refrain from making this stipulation.

Grzegorczyk conjectured that every concatenation 2-structure is isomorphic to a concrete concatenation structure. We prove that this conjecture is false. (i) Every extension of TC_1 has a model that is not isomorphic to a concrete concatenation 1-structure and (ii) the set of principles valid in all concrete concatenation 2-structures interprets arithmetical truth.

The plan of the paper is as follows. We show, in Section 2, that we have, for all decorated order types α, β and γ, the following principle:

$$(\dagger) \qquad \beta * \alpha * \gamma = \alpha \Rightarrow \beta * \alpha = \alpha * \gamma = \alpha.$$

This fact was already known. It is due to Lindenbaum, credited to him in Sierpiński's book [6] on p. 248. It is also problem 6.13 of [4].

It is easy to see that every group is a concatenation structure and that $(\dagger)$ does not hold in the two element group. We show, in Section 5, that every concatenation structure can be extended to a concatenation structure with any number of atoms. It follows that there is a concatenation structure with at least two atoms in which $(\dagger)$ fails. Hence, TC is incomplete for concrete concatenation structures. In Section 3, we provide a counterargument of a different flavour. We provide a tally interpretation that defines the natural numbers (with concatenation in the role of addition) in every concrete concatenation 2-structure. It follows that every extension of TC_1 is satisfied by a concatenation 1-structure that is not isomorphic to any concrete concatenation 1-structure, to wit any model of that extension that contains a non-standard element. In Section 4, we strengthen the result of Section 3, by showing that in concrete concatenation 2-structures we can add multiplication to the natural numbers.

It follows that the set of arithmetically true sentences is interpretable in the concretely valid consequences of TC_2.

Finally, in Section 6 we prove a positive result. We provide a mapping from arbitrary models of a variant of an extension of TC_0 to structures of decorated order types. As we have shown such a construction cannot always provide a representation. We show that, for a restricted class of representations, we do have: if a model has a representation in the class, then the construction yields such a representation.

§2. **A principle for decorated order types.** In this section we prove a universal principle that holds in all concatenation structures, which is not provable in TC. There is an earlier proof of this principle [4, p. 187]. Our proof, however, is different from that of Komjáth and Totik.

THEOREM 1. *Let* $\alpha_0, \alpha_1, \alpha_2$ *be decorated order types. Suppose that* $\alpha_1 = \alpha_0 * \alpha_1 * \alpha_2$. *Then,* $\alpha_1 = \alpha_0 * \alpha_1 = \alpha_1 * \alpha_2$.

PROOF. Suppose $\alpha_1 = \alpha_0 * \alpha_1 * \alpha_2$. Consider a decorated linear ordering $\mathcal{A} := \langle A, \leq, f \rangle$ of type α_1, By our assumption, we may partition A into A_0, A_1, A_2, such that:

$$\langle A, \leq, f \rangle = \langle A_0, \leq\restriction A_0, f \restriction A_0 \rangle * \langle A_1, \leq\restriction A_1, f \restriction A_1 \rangle * \langle A_2, \leq\restriction A_2, f \restriction A_2 \rangle,$$

where $\mathcal{A}_i := \langle A_i, \leq\restriction A_i, f \restriction A_i \rangle$ is an instance of α_i, Let $\phi : \mathcal{A} \to \mathcal{A}_1$ be an isomorphism.

Let $\phi^n \mathcal{A}_{(i)} := \langle \phi^n[A_{(i)}], \leq\restriction \phi^n[A_{(i)}], f \restriction \phi^n[A_{(i)}] \rangle$. We have: $\phi^n \mathcal{A}_i$ is of order type α_i and $\phi^n \mathcal{A}$ is of order type α_1.

Clearly, $\phi\mathcal{A}_0$ is an initial substructure of $\phi\mathcal{A} = \mathcal{A}_1$. So, $\mathcal{A}_0$ and $\phi\mathcal{A}_0$ are disjoint and $\phi\mathcal{A}_0$ adjacent to the right of $\mathcal{A}_0$. Similarly, for $\phi^n\mathcal{A}_0$ and $\phi^{n+1}\mathcal{A}_0$. Take $A_0^\omega := \bigcup_{i \in \omega} \phi^i A_0$. We find that $\mathcal{A}_0^\omega := \langle A_0^\omega, \leq\restriction A_0^\omega, f \restriction A_0^\omega \rangle$ is initial in $\mathcal{A}$ and of decorated linear order type α_0^ω. So $\alpha_1 = \alpha_0^\omega * \rho$, for some ρ. It follows that $\alpha_0 * \alpha_1 = \alpha_0 * \alpha_0^\omega * \rho = \alpha_0^\omega * \rho = \alpha_1$. The other identity is similar. ⊣

So, every concrete concatenation structure validates that $\alpha_1 = \alpha_0 * \alpha_1 * \alpha_2$ implies $\alpha_1 = \alpha_0 * \alpha_1 = \alpha_1 * \alpha_2$. We postpone the proof that this principle is not provable in TC to Section 5.

§3. **Definability of the natural numbers.** In this section, we show that the natural numbers can be defined in every concrete concatenation 1-structure. We define:

- $x \subseteq y :\leftrightarrow x = y \vee \exists u\ (u * x = y) \vee \exists v\ (x * v = y) \vee \exists u, v\ (u * x * v = y)$.
- $x \subseteq_{\mathsf{ini}} y :\leftrightarrow x = y \vee \exists v\ (x * v = y)$.
- $x \subseteq_{\mathsf{end}} y :\leftrightarrow x = y \vee \exists u\ (u * x = y)$.
- $(n : \widetilde{\mathsf{N}}_{\mathsf{a}}) :\leftrightarrow \forall m \subseteq_{\mathsf{ini}} n\ (m = \mathsf{a} \vee \exists k\ (k \neq m \wedge m = k * \mathsf{a}))$.

The use of ':' in $n : \widetilde{\mathsf{N}}_{\mathsf{a}}$ is derived from the analogous use in type theory. We could read it as: n is of sort N_{a}. We write $m, n : \widetilde{\mathsf{N}}_{\mathsf{a}}$ for: $(m : \widetilde{\mathsf{N}}_{\mathsf{a}}) \wedge (n : \widetilde{\mathsf{N}}_{\mathsf{a}})$, etc. In the context of a structure we will confuse $\widetilde{\mathsf{N}}_{\mathsf{a}}$ with the extension of $\widetilde{\mathsf{N}}_{\mathsf{a}}$ in that structure.

We prove the main theorem of this section.

THEOREM 2. *In any concrete concatenation* 1*-structure, we have*:

$$\widetilde{\mathsf{N}}_{\mathsf{a}} = \{\mathsf{a}^{n+1} \mid n \in \omega\}.$$

In other words, $\widetilde{\mathsf{N}}_{\mathsf{a}}$ *is precisely the class of natural numbers in tally representation* (*starting with* 1). *Note that* $*$ *on this set is addition.*

PROOF. Consider any concrete concatenation 1-structure $\mathfrak{A}$. It is easy to see that every a^{n+1} is in $\widetilde{\mathsf{N}}_{\mathsf{a}}$. Clearly, every element x of $\widetilde{\mathsf{N}}_{\mathsf{a}}$ is either a or it has a predecessor, i.e., there is a y such that $x = y * \mathsf{a}$. The axioms of TC_1 guarantee that this predecessor is unique. This justifies the introduction of the partial predecessor function pd on $\widetilde{\mathsf{N}}_{\mathsf{a}}$. Let α be the order type corresponding to a. Let β_0 be any element of $\widetilde{\mathsf{N}}_{\mathsf{a}}$. If, for some n, $\mathsf{pd}^n\beta_0$ is undefined, then β_0 is clearly of the form α^{k+1}, for k in ω.

We show that the other possibility cannot obtain. Suppose $\beta_n := \mathsf{pd}^n\beta_0$ is always defined. Let $\mathcal{A}$ be a decorated linear ordering of type α and let $\mathcal{B}_i$ be a decorated linear ordering of type β_i. We assume that the domain A of $\mathcal{A}$ is disjoint from the domains B_i of the $\mathcal{B}_i$. Thus, we may implement $\mathcal{B}_{i+1} * \mathcal{A}$ just by taking the union of the domains.

Let ϕ_i be isomorphisms from $\mathcal{B}_{i+1} * \mathcal{A}$ to $\mathcal{B}_i$. Let $\mathcal{A}_i := (\phi_0 \circ \cdots \circ \phi_i)(\mathcal{A})$. Then, the $\mathcal{A}_i$ are all of type α and, for some $\mathcal{C}$, we have $\mathcal{B}_0 \cong \mathcal{C} * \cdots * \mathcal{A}_1 * \mathcal{A}_0$. Similarly $\mathcal{B}_1 \cong \mathcal{C} * \cdots * \mathcal{A}_2 * \mathcal{A}_1$. Let $\breve{\omega}$ be the opposite ordering of ω. It follows that $\beta_0 = \gamma * \alpha^{\breve{\omega}} = \beta_1 = \mathsf{pd}(\beta_0)$. Hence, β_0 is not in $\widetilde{\mathsf{N}}_{\mathsf{a}}$[2]. A contradiction. ⊣

We call a concatenation structure *standard* if $\widetilde{\mathsf{N}}_{\mathsf{a}}$ defines the tally natural numbers. Since, by the usual argument, any extension of TC_1 has a model with non-standard numbers, we have the following corollary.

COROLLARY 3. *Every extension of* TC_1 *has a model that is not isomorphic to a concrete concatenation* 1*-structure. In a different formulation*: *for every concatenation* 1*-structure there is an elementarily equivalent concatenation* 1*-structure that is not isomorphic to a concrete concatenation* 1*-structure.*

Note that the non-negative tally numbers with addition form a concrete concatenation 1-structure. Thus, the concretely valid consequences of $\mathsf{TC}_1 + \forall x\ (x : \widetilde{\mathsf{N}}_{\mathsf{a}})$, i.e., the principles valid in every concrete concatenation 1-structure satisfying $\forall x\ (x : \widetilde{\mathsf{N}}_{\mathsf{a}})$ are decidable.

[2]Note that we are not assuming that γ is in $\mathfrak{A}$.

§4. **Definability of multiplication.** If we have two atoms to work with, we can add multiplication to our tally numbers. This makes the set of concretely valid consequences of TC non-arithmetical. The main ingredient of the definition of multiplication is the theory of relations on tally numbers. In TC, we can develop such a theory. The development has some resemblance to the construction in the classic paper of Quine [5]. However, the ideas here are somewhat more intricate, since we are working in a more general context than that of [5]. We represent the relation $\{\langle x_0, y_0\rangle, \dots, \langle x_{n-1}, y_{n-1}\rangle\}$, by:

$$\mathsf{bb} * x_0 * \mathsf{b} * y_0 * \mathsf{bb} * x_1 * \dots \mathsf{bb} * x_{n-1} * \mathsf{b} * y_{n-1} * \mathsf{bb}.$$

We define:

- $r : \mathsf{REL} :\leftrightarrow \mathsf{bb} \subseteq_{\mathsf{end}} r$,
- $\emptyset := \mathsf{bb}$,
- $x[r]y :\leftrightarrow x, y : \widetilde{\mathsf{N}}_{\mathsf{a}} \wedge \mathsf{bb} * x * \mathsf{b} * y * \mathsf{bb} \subseteq r$.
- $\mathsf{adj}(r, x, y) := r * x * \mathsf{b} * y * \mathsf{bb}$.

Clearly, we have: $\mathsf{TC} \vdash \forall u, v\ \neg u[\emptyset]v$. To verify that this coding works we need the adjunction principle.

THEOREM 4. *We have*:

$$\mathsf{TC} \vdash (r : \mathsf{REL} \wedge x, y, u, v : \widetilde{\mathsf{N}}_{\mathsf{a}}) \rightarrow \big(u[\mathsf{adj}(r, x, y)]v \leftrightarrow \big(u[r]v \vee (u = x \wedge v = y)\big)\big).$$

We can prove this result by laborious and unperspicuous case splitting. However, it is more elegant to do the job with the help of a lemma. Consider any model of TC_0. Fix an element w. We call a sequence $(w_0, \dots, w_k)$ a *partition* of w if we have that $w_0 * \dots * w_k = w$. The partitions of w form a category with the following morphisms. $f : (u_0, \dots, u_n) \rightarrow (w_0, \dots, w_k)$ iff f is a surjective and weakly monotonic function from $n+1$ to $k+1$, such that, for any $i \leq k$, $w_i = u_s * \dots * u_\ell$, where $f(j) = i$ iff $s \leq j \leq \ell$. We write $(u_0, \dots, u_n) \leq (w_0, \dots, w_k)$ for: $\exists f\ f : (u_0, \dots, u_n) \rightarrow (w_0, \dots, w_k)$. In this case we say that $(u_0, \dots, u_n)$ is *a refinement* of $(w_0, \dots, w_k)$.

LEMMA 5. *Consider any concatenation structure. Let w be an element of the structure. Then, any two partitions of w have a common refinement.*

PROOF. Fix any concatenation structure. We first prove that, for all w, all pairs of partitions $(u_0, \dots, u_n)$ and $(w_0, \dots, w_k)$ of w have a common refinement, by induction of $n + k$.

If either n or k is 0, this is trivial. Suppose that $(u_0, \dots, u_{n+1})$ and $(w_0, \dots, w_{k+1})$ are partitions of w. By the editor axiom, we have either (a) $u_0 * \dots * u_n = w_0 * \dots * w_k$ and $u_{n+1} = w_{k+1}$, or there is a v such that (b) $u_0 * \dots * u_n * v = w_0 * \dots * w_k$ and $u_{n+1} = v * w_{k+1}$, or (c) $u_0 * \dots * u_n = w_0 * \dots * w_k * v$ and $v * u_{n+1} = w_{k+1}$. We only treat case (b), the other cases being easier or similar. By the induction hypothesis, there is a common refinement $(x_0, \dots x_m)$ of $(u_0, \dots, u_n, v)$ and $(w_0, \dots, w_n)$. Let this be witnessed by f, resp. g. It is easily seen that $(x_0, \dots x_m, w_{k+1})$ is the desired refinement with witnessing functions f' and g', where $f' := f[m+1 \mapsto n+1]$, $g' := g[m+1 \mapsto k+1]$. Here $f[m+1 \mapsto n+1]$ in the result of extending f to assign $n+1$ to $m+1$. ⊣

We turn to the proof of Theorem 4. The verification proceeds more or less as one would do it for finite strings.

PROOF. Consider any concatenation 2-structure. Suppose $\mathsf{REL}(r)$. The right-to-left direction is easy, so we treat left-to-right. Suppose x, y, u and v are tally numbers. and $u[\mathsf{adj}(r, x, y)]v$. There are two possibilities. Either $r = \mathtt{bb}$ or $r = r_0 * \mathtt{bb}$. We will treat the second case. Let $s := \mathsf{adj}(r, x, y)$. One the following four partitions is a partition of s: (i) $(\mathtt{b}, \mathtt{b}, u, \mathtt{b}, v, \mathtt{b}, \mathtt{b})$, or (ii) $(w, \mathtt{b}, \mathtt{b}, u, \mathtt{b}, v, \mathtt{b}, \mathtt{b})$, or (iii) $(\mathtt{b}, \mathtt{b}, u, \mathtt{b}, v, \mathtt{b}, \mathtt{b}, z)$, or (iv) $(w, \mathtt{b}, \mathtt{b}, u, \mathtt{b}, v, \mathtt{b}, \mathtt{b}, z)$. We will treat cases (ii) and (iv).

Suppose $\sigma := (w, \mathtt{b}, \mathtt{b}, u, \mathtt{b}, v, \mathtt{b}, \mathtt{b})$ is a partition of s. We also have that $\tau := (r_0, \mathtt{b}, \mathtt{b}, x, \mathtt{b}, y, \mathtt{b}, \mathtt{b})$ is a partition of s. Let $(t_0, \dots, t_k)$ be a common refinement of σ and τ, with witnessing functions f and g. The displayed b's in these partitions must have unique places among the t_i. We define m_σ to be the unique i such that $f(i) = m$, provided that $\sigma_m = \mathtt{b}$. Similarly, for m_τ. (To make this unambiguous, we assume that if $\sigma = \tau$, we take σ as the common refinement with f and g both the identity function.)

We evidently have $7_\sigma = 7_\tau = k$ and $6_\sigma = 6_\tau = k - 1$. Suppose $4_\sigma < 4_\tau$. It follows that $\mathtt{b} \subseteq v$. So, v would have an initial subsequence that ends in b, which is impossible. So, $4_\sigma \not< 4_\tau$. Similarly, $4_\tau \not< 4_\sigma$. So $4_\sigma = 4_\tau$. It follows that $v = y$. Reasoning as in the case of 4_σ and 4_τ, we can show that $2_\sigma = 2_\tau$ and, hence $u = x$.

Suppose $\rho := (w, \mathtt{b}, \mathtt{b}, u, \mathtt{b}, v, \mathtt{b}, \mathtt{b}, z)$ is a partition of s. We also have that $\tau := (r_0, \mathtt{b}, \mathtt{b}, x, \mathtt{b}, y, \mathtt{b}, \mathtt{b})$ is a partition of s. Let $(t_0, \dots, t_k)$ be a common refinement of ρ and τ, with witnessing functions f and g. We consider all cases, where $1_\tau < 6_\rho$. Suppose $6_\rho = 1_\tau + 1 = 2_\tau$. Note that $7_\rho = 6_\rho + 1$, so we find: $\mathtt{b} \subseteq x$, quod non, since x is in $\widetilde{\mathsf{N}}_{\mathtt{a}}$. Suppose $2_\tau < 6_\rho < 4_\tau$. In this case we have a b as substring of x. Quod non. Suppose $6_\rho = 4_\tau$. Since $7_\rho = 6_\rho + 1$, we get a b in y. Quod non. Suppose $4_\tau < 6_\rho < 6_\tau$. In this case, we get a b in y. Quod impossible. Suppose $6_\rho \geq 6_\tau = k - 1$. In this place there is no place left for z among the t_i. So, in all cases, we obtain a contradiction. So the only possibility is $6_\rho \leq 1_\tau$. Thus, it follows that $u[r]v$. ⊣

We can now use our relations to define multiplication of tally numbers in the usual way. See e.g. Section 2.2 of [1]. In any concrete concatenation 2-structure, we can use induction to verify the defining properties of multiplication as defined. It follows that we can interpret all arithmetical truths in the set of concretely valid consequences of TC.

COROLLARY 6. *We can interpret true arithmetic in the set of all principles valid in concrete concatenation* 2*-structures.*

§5. **The sum of concatenation structures.** In this section we show that concatenation structures are closed under sums. This result will make it possible to verify the claim that the universal principle of Section 2 is not provable in TC. The result has some independent interest, since it provides a good closure property of concatenation structures.

Consider two concatenation structures $\mathfrak{A}_0$ and $\mathfrak{A}_1$. We write $\star$ for concatenation in the $\mathfrak{A}_i$. We may assume, without loss of generality, that the domains of $\mathfrak{A}_0$ and $\mathfrak{A}_1$ are disjoint. We define the sum $\mathfrak{B} := \mathfrak{A}_0 \oplus \mathfrak{A}_1$ as follows.

- The domain of $\mathfrak{B}$ consists of non-empty sequences $w_0 \cdots w_{n-1}$, where the w_j are alternating between elements of the domains of $\mathfrak{A}_0$ and $\mathfrak{A}_1$. In other words, if w_j is in the domain of $\mathfrak{A}_i$, then w_{j+1}, if it exists, is in the domain of $\mathfrak{A}_{1-i}$.
- The concatenation $\sigma * \tau$ of $\sigma := w_0 \cdots w_{n-1}$ and $\tau := v_0 \cdots v_{k-1}$ is $w_0 \cdots w_{n-1} v_0 \cdots w_{k-1}$, in case w_{n-1} and v_0 are in the domains of different structures $\mathfrak{A}_i$. The concatenation $\sigma * \tau$ is $w_0 \cdots (w_{n-1} \star v_0) \cdots w_{k-1}$, in case w_{n-1} and v_0 are in in the same domain.

In case $\sigma * \tau$ is obtained via the first case, we say that σ and τ are *glued together*. If the second case obtains, we say that σ and τ are *clicked together*.

THEOREM 7. *The structure* $\mathfrak{B} = \mathfrak{A}_0 \oplus \mathfrak{A}_1$ *is a concatenation structure.*

PROOF. Associativity is easy. We check the editor property TC2. Suppose $\sigma_0 * \sigma_1 = z_0 \cdots z_{m-1} = \tau_0 * \tau_1$. We distinguish a number of cases.

CASE 1. Suppose both of the pairs σ_0, σ_1 and τ_0, τ_1 are glued together. Then, for some $k, n > 0$, we have $\sigma_0 = z_0 \cdots z_{k-1}$, $\sigma_1 = z_k \cdots z_{m-1}$, $\tau_0 = z_0 \cdots z_{n-1}$, and $\tau_1 = z_n \cdots z_{m-1}$.

So, if $k = n$, we have $\sigma_0 = \tau_0$ and $\sigma_1 = \tau_1$.

If $k < n$, we have $\tau_0 = \sigma_0 * (z_k \cdots z_{n-1})$ and $\sigma_1 = (z_k \cdots z_{n-1}) * \tau_1$. The case that $n < k$ is similar.

CASE 2. Suppose σ_0, σ_1 is glued together and that τ_0, τ_1 is clicked together. So, there are $k, n > 0$, u_0, and u_1 such that $\sigma_0 = z_0 \cdots z_{k-2} u_0$, $\sigma_1 = u_1 z_k \cdots z_{m-1}$, $u_0 \star u_1 = z_{k-1}$, $\tau_0 = z_0 \cdots z_{n-1}$, and $\tau_1 = z_n \cdots z_{m-1}$.

Suppose $k \le n$. Then, $\tau_0 = \sigma_0 * (u_1 z_k \cdots z_{n-1})$ and $\sigma_1 = (u_1 z_k \cdots z_{n-1}) * \tau_1$. Note that, in case $k = n$, the sequence $z_k \cdots z_{n-1}$ is empty. The case that $k \ge n$ is similar.

CASE 3. This case, where σ_0, σ_1 is clicked together and τ_0, τ_1 is glued together, is similar to Case 2.

CASE 4. Suppose that σ_0, σ_1 and τ_0, τ_1 are both clicked together. So, there are $k, n > 0$, u_0, u_1, v_0, v_1 such that $\sigma_0 = z_0 \cdots z_{k-2} u_0$, $\sigma_1 = u_1 z_k \cdots z_{m-1}$, $u_0 \star u_1 = z_{k-1}$, $\tau_0 = z_0 \cdots z_{n-2} v_0$, $\tau_1 = v_1 z_n \cdots z_{m-1}$ and $v_0 \star v_1 = z_{n-1}$.

Suppose $k = n$. We have $u_0 \star u_1 = z_{k-1} = v_0 \star v_1$. So, we have either (a) $u_0 = v_0$ and $u_1 = v_1$, or, for some w, either (b) $u_0 \star w = v_0$ and $u_1 = w \star v_1$, or (c) $u_0 = v_0 \star w$ and $w \star u_1 = v_1$. In case (b), we have: $\sigma_0 * w = \tau_0$ and $\sigma_1 = w * \tau_1$. We leave (a) and (c) to the reader.

Suppose $k < n$. We have:

$$\sigma_0 * (u_1 z_k \cdots z_{n-2} v_0) = \tau_0 \text{ and } \sigma_1 = (u_1 z_k \cdots z_{n-2} v_0) * \tau_1.$$

The case that $k > n$ is similar. ⊣

It is easy to see that $\oplus$ is a sum or coproduct in the sense of category theory. The following theorem is immediate.

THEOREM 8. *If a is an atom of $\mathfrak{A}_i$, then a is an atom of $\mathfrak{A}_0 \oplus \mathfrak{A}_1$.*

Finally, we have the following theorem.

THEOREM 9. *Let A be any set and let $\mathfrak{B} := \langle B, * \rangle$ be any concatenation structure. We assume that A and B are disjoint. Then, there is an extension of $\mathfrak{B}$ with at least A as atoms.*

PROOF. Let A^* be the free semi-group generated by A. We can take as the desired extension of $\mathfrak{B}$, the structure $A^* \oplus \mathfrak{B}$. ⊣

REMARK 5.1. The whole development extends with only minor adaptations, when we replace axiom TC2 by:

- $\vdash x * y = u * v \rightarrow ((x = u \wedge y = v) \dot{\vee} (\exists! w\ (x * w = u \wedge y = w * v) \vee \exists! w\ (x = u * w \wedge y * w = v))$

Here $\dot{\vee}$ is *exclusive or*.

§6. A canonical construction. Although we know that not every concatenation structure can be represented by decorated linear orderings, i.e., as a concrete concatenation structure, there may exist a canonical construction of a concrete concatenation structure which is a representation whenever there exists any concrete representation. In this section we shall propose such a construction, but we can only show that it is universal in a restricted subclass of all concrete representations.

It will be now more convenient to work with a theory for monoids, rather than for semigroups, as we did in the previous sections. We will work in the theory $\mathsf{TC}^{\varepsilon}_{+}$, which is $\mathsf{TC}^{\varepsilon}_{0}$ plus the following axiom.

$\mathsf{TC}^{\varepsilon}9$. $\vdash\ x * y * z = y\ \rightarrow\ (x = \varepsilon \wedge z = \varepsilon)$.

We do not postulate the existence of irreducible elements, as they do not play any role in what follows, but they surely can be present. We shall call elements of a model $\mathcal{M}$ of $\mathsf{TC}^{\varepsilon}_{+}$: words. When possible, the concatenation symbol $*$ will be omitted.

When considering representations of structures with a unit element ε by decorated order types, one has to allow an empty decorated order structure.

Thus a representation of a model $\mathcal{M}$ of $\mathsf{TC}_+^\varepsilon$ is a mapping ρ that assigns a decorated order structure to every $w \in \mathcal{M}$ so that

1. $\rho(\varepsilon) = \emptyset$,
2. $\rho(uv) = \rho(u)\rho(v)$ and
3. $\rho(w) \cong \rho(v)$ implies $w = v$.

LEMMA 10. *In a model $\mathcal{M}$ of $\mathsf{TC}_+^\varepsilon$ the binary relation $\exists u(xu = y)$ defines an ordering on the elements of $\mathcal{M}$.*

DEFINITION 6.1. A *k-partition* of a word w is a k-tuple $(w_1, \dots, w_k)$ such that $w_1 \dots w_k = w$; we shall often abbreviate it by $w_1 \dots w_k$. An ordering relation is defined on 3-partitions of w by:

- $u_1u_2u_3 \leq v_1v_2v_3 \;\; :\Leftrightarrow \;\; \exists x_1, x_3\,(v_1x_1 = u_1 \wedge x_3v_3 = u_3)$.

The axioms ensure that for any two partitions there is a unique common refinement.

DEFINITION 6.2 (Word Ultrafilters). Let w be a word and S a set of 3-partitions of w. We shall call S a word ultrafilter on w if

1. $\varepsilon w \varepsilon \in S$
2. $x\varepsilon y \notin S$ for any x, y
3. if $U \in S$, V is a 3-partition of w and $U \leq V$, then $V \in S$
4. if $xyz \in S$ and $y = y_1y_2$, then exactly one of the following two cases holds: $(x, y_1, y_2z) \in S$ or $(xy_1, y_2, z) \in S$.

Let S be a word ultrafilter on w and $xyz \in S$. Then we define the natural restriction of S to y which is a word ultrafilter S_y on y defined by:

- $(r, s, t) \in S_y \;\; :\Leftrightarrow \;\; (xr, s, tz) \in S$.

We shall define an ordering on word ultrafilters on a fixed w and an equivalence on word ultrafilters on all words of M. Let S and T be word ultrafilters on w, then we define:

- $S < T \;\; :\Leftrightarrow \;\; \exists u, v\; ((\varepsilon, u, v) \in S \wedge (u, v, \varepsilon) \in T)$.

Let S and T be word ultrafilters on possibly different words, then we define:

- $S \sim T \; :\Leftrightarrow \; \exists x, x', y, z, z'\; (xyz \in S \wedge x'yz' \in T \wedge S_y = T_y)$.

Notice that $<$ is a strict ordering on word ultrafilters on w, but for $S < T$ it still may be $S \sim T$.

DEFINITION 6.3. Let w be a word. The *canonical decorated ordering* associated with w is the ordering of all word ultrafilters on w, where each word ultrafilter S is decorated by $[S]_\sim$, the equivalence class of $\sim$ containing S. This decorated ordering will be denoted by $C(w)$.

Here are some basic properties of $C(w)$.

- The topological space determined by the ordering is compact and totally disconnected. In particular, it has largest and smallest elements.