

STRING DATA STRUCTURES FOR CHINESE MONOIDS

NOHRA HAGE ¹

Abstract – The structure of Chinese monoid appeared in the classification of monoids with the growth function coinciding with that of the plactic monoid. In this work, we deal with the presentations of the Chinese monoid from the rewriting theory perspective using the notion of string data structures. We define a string data structure associated to the Chinese monoid using the insertion algorithm on Chinese staircases. As a consequence, we construct a finite semi-quadratic convergent presentation of the Chinese monoid and we extend it into a finite coherent presentation of this monoid.

1. INTRODUCTION

The Chinese monoid was discovered by Duchamp and Krob during their classification of monoids with growth similar to that of the plactic monoid, [6]. The latter monoid was introduced by Lascoux and Schützenberger [13] and it has found several applications in algebraic combinatorics and representation theory [7]. Plactic monoids were also investigated using rewriting methods by constructing finite convergent presentations for classical types [1, 2, 10] and by computing coherent presentations for plactic monoids of type A, [11]. The *Chinese monoid* of rank n is the monoid generated by the ordered set $\{1 < \dots < n\}$ and submitted to the relations $zyx = zxy = yzx$, for $x \leq y \leq z$. Since Young tableaux play an important role in the structure of plactic monoids, a similar notion of *Chinese staircases* was found for the Chinese monoid, [4]. Moreover, a right insertion and a left insertion similar to Schensted’s insertions were also introduced on the structure of Chinese staircases yielding a cross-section property for the Chinese monoid, [3, 4]. Recently, the Chinese monoid has motivated a wide range of other interesting work in rewriting theory [5, 9] including computing a finite convergent presentation [3] by adding *column generators*. However, the latter presentation is not semi-quadratic in the sense that the targets of its rewriting rules can contain more than two columns generators, and thus it is difficult to extend this presentation into a coherent presentation for the Chinese monoid. In this work, we construct a finite semi-quadratic convergent presentation of the Chinese monoid and we extend it into a finite coherent presentation using the notion of *string data structures* on the Chinese staircases.

Recall from [12] that a *string data structure* (SDS) \mathbb{S} on a totally ordered alphabet A is defined by a set of combinatorial objects D_A together with an insertion map constructing the elements of D_A and a reading map describing the elements of D_A by words on A . The insertion map defines a product $\star_{\mathbb{S}}$ on D_A . The *structure monoid* associated to \mathbb{S} is presented by the 2-polygraph whose set of 1-cells is D_A and whose 2-cells rewrite every two elements in D_A into their product by $\star_{\mathbb{S}}$. One shows that the associativity of $\star_{\mathbb{S}}$, the fact that D_A satisfies the cross-section property for the structure monoid, and the confluence of 2-polygraph presenting the structure monoid are equivalent properties, see [12]. Note that the associativity of $\star_{\mathbb{S}}$ can be also deduced from the existence of two insertion algorithms that commute. Moreover, one can compute finite coherent presentation of the structure monoid made of a *generating set* of the set of combinatorial objects, rewriting rules describing the insertion on words and relations among the insertion algorithms of the data structure.

¹Ecole supérieure d’ingénieurs de Beyrouth (ESIB), Université Saint-Joseph de Beyrouth (USJ), Liban.
nohra.hage@usj.edu.lb

2. String data structures, cross-section and coherent presentations

We define a string data structure presenting the Chinese monoid made by the set of Chinese staircases together with the right insertion of the Chinese monoid and a reading map that transforms Chinese staircases into words on $\{1, \dots, n\}$. We show that the right insertion and the left insertion for the Chinese monoid commute which gives the associativity of the product defined by this SDS. We obtain as a consequence that the set of Chinese staircases satisfies the cross-section property for the Chinese monoid. Finally, using the SDS on the Chinese staircases, we construct a finite coherent convergent presentation of the Chinese monoid.

2. STRING DATA STRUCTURES, CROSS-SECTION AND COHERENT PRESENTATIONS

We recall in this section the notion of string data structures and the construction given in [12].

String data structures. A *string data structure*, SDS for short, \mathbb{S} on a totally ordered alphabet A is a quadruple (D_A, ℓ, I, R) made of a set D_A , a reading ℓ of words on A , a *one-element insertion map* I and a *reading map* R defined as follows:

- i) the inclusions $A \subseteq D_A \subseteq A^*$ hold, where A^* denotes the free monoid on A ,
- ii) the map $\ell : A^* \rightarrow A^*$ sends each word $x_1 \dots x_k$ in A^* on a word $x_{\sigma(1)} \dots x_{\sigma(k)}$ in A^* , where σ is a permutation on $\{1, \dots, k\}$,
- iii) $I : D_A \times A \rightarrow D_A$ inserts an element of A into an element of D_A such that any restriction $I(-, x)$ is injective for any x in A . By iteration, one defines an *insertion map* $I^* : D_A \times A^* \rightarrow D_A$ that inserts a word in A^* into an element of D_A wrt ℓ , that is $I^*(d, x_1 \dots x_n) = I^*(I(d, y_1), y_2 \dots y_n)$, for every $d \in D_A$ and $x_1 \dots x_n \in A^*$, where $y_1 \dots y_n = \ell(x_1 \dots x_n)$,
- iv) $R : D_A \rightarrow A^*$ is an injective map satisfying $I^*(\emptyset, \ell(-))R = \text{Id}_{D_A}$ and $R(\emptyset)$ is the empty word.

The map $I^*(\emptyset, \ell(-)) : A^* \rightarrow D_A$, denoted by $C_{\mathbb{S}}$, is called the *constructor* of the SDS \mathbb{S} . We will use the *right-to-left* (resp. *left-to-right*) reading of words denoted by ℓ_r (resp. ℓ_l). A *right* (resp. *left*) *insertion SDS* is an SDS whose insertion map is said *right* (resp. *left*), that is inserting a word into an element of D_A with respect to ℓ_l (resp. ℓ_r). Two one-element insertion maps $I, J : D_A \times A \rightarrow D_A$ commute if the relation $J(I(d, x), y) = I(J(d, y), x)$ holds for every $d \in D_A$ and $x, y \in A$. An *opposite* of a right (resp. left) insertion SDS (D_A, ℓ_l, I, R) (resp. (D_A, ℓ_r, I, R)) is a left (resp. right) insertion SDS (D_A, ℓ_r, J, R) (resp. (D_A, ℓ_l, J, R)) such that I and J commute. An SDS $\mathbb{S} = (D_A, \ell, I, R)$ is *associative* if the product $\star_{\mathbb{S}} : D_A \times D_A \rightarrow D_A$ defined by $d \star_{\mathbb{S}} d' = I^*(d, \ell(R(d')))$, for every $d, d' \in D_A$ is associative.

Let \mathbb{S} be a right (resp. left) insertion SDS. If there is a left (resp. right) insertion SDS \mathbb{T} opposite to \mathbb{S} , then the SDS \mathbb{S} and \mathbb{T} commute, that is $d \star_{\mathbb{S}} d' = d' \star_{\mathbb{T}} d$, for any $d, d' \in D_A$, and are associative, [12].

Structure monoid. Let $\mathbb{S} = (D_A, \ell, I, R)$ be an SDS. Denote by $|$ the product of the free monoid on D_A . The *structure monoid* associated to the SDS \mathbb{S} is the monoid, denoted by $\mathbf{M}(\mathbb{S})$, and presented by the 2-polygraph $\mathcal{R}(\mathbb{S})$, called the *standard presentation* induced by \mathbb{S} , whose set of 1-cells is D_A and whose 2-cells are $\gamma_{d,d'} : d|d' \Rightarrow d \star_{\mathbb{S}} d'$, for any $d, d' \in D_A$. Since every application of a 2-cell of $\mathcal{R}(\mathbb{S})$ yields a strictly smaller preceding word with respect to the deglex order on D_A^* , $\mathcal{R}(\mathbb{S})$ is terminating. Moreover, if \mathbb{S} is associative, then $\mathcal{R}(\mathbb{S})$ is convergent.

The *reading of the standard presentation* of the SDS \mathbb{S} is the 2-polygraph $\mathcal{R}(A, \mathbb{S})$ whose set of 1-cells is A and whose 2-cells are $\gamma_{d,d'} : R_{\mathbb{S}}(d)R_{\mathbb{S}}(d') \Rightarrow R_{\mathbb{S}}(d \star_{\mathbb{S}} d')$, for any $d, d' \in D_A$. If \mathbb{S} is associative,

2. String data structures, cross-section and coherent presentations

then the 2-polygraph $\mathcal{R}(A, \mathbb{S})$ is locally confluent, [12]. Recall that an associative SDS $\mathbb{S} = (D_A, \ell, I, R)$ is *compatible* with an equivalence relation \sim on A^* if for any d in D_A and w, w' in A^* , $w \sim w'$ implies $I^*(d, w) = I^*(d, w')$, and for any $w \in A^*$, one has $RC(w) \sim w$. Let K be a set and let \sim be an equivalence relation on K^* . A subset $S \subset K^*$ *satisfies the cross-section property for the monoid K^*/\sim* if each equivalence class with respect to \sim contains exactly one element of S . Let \mathbb{S} be a right insertion associative SDS compatible with the equivalence relation $\sim_{\mathbb{S}}$ induced by $\mathcal{R}(A, \mathbb{S})$. One shows that the monoids $(D_A, \star_{\mathbb{S}})$ and $A^*/\sim_{\mathbb{S}}$ are isomorphic, [12]. In particular, if $\mathcal{R}(A, \mathbb{S})$ is terminating, then the set of normal forms wrt $\mathcal{R}(\mathbb{S})$ satisfies the cross-section property for $\mathbf{M}(\mathbb{S})$ if and only if the set of normal forms wrt $\mathcal{R}(A, \mathbb{S})$ satisfies the cross-section property for $\mathbf{M}(\mathbb{S})$.

A reduced presentation. Let $\mathbb{S} = (D_A, \ell, I, R)$ be an SDS. An *internal composition* for \mathbb{S} is a binary relation $|$ on D_A , such that $R(d|d') = R(d)R(d')$, for any $d, d' \in D_A$, where $d|d'$ denotes $(d, d') \in |$. A *generating set* with respect to an internal composition $|$ for \mathbb{S} is a subset Q of D_A such that $A \subseteq Q$, and any element d in D_A can be written $d = c_1|c_2|\dots|c_k$, with $c_1, \dots, c_k \in Q$. From a generating set Q of \mathbb{S} with respect to an internal composition $|$, one can define an SDS $\mathbb{S}_Q = (D_A, \ell_Q, I_Q, R_Q)$ on Q , where

- i) the map $\ell_Q : Q^* \rightarrow Q^*$ induces a permutation on the letters of each words on Q ,
- ii) $I_Q : D_A \times Q \rightarrow D_A$ is an one element insertion map defined by $I_Q(d, c) = I^*(d, R(c))$, for any $c \in Q$ and $d \in D_A$, that induces an insertion map $I_Q^* : D_A \times Q^* \rightarrow D_A$ wrt ℓ_Q ,
- iii) $R_Q : D_A \rightarrow Q^*$ is the reading map associated to the composition $|$, that is, for any d in D_A , $R_Q(d) = c_1|c_2|\dots|c_k$ is the decomposition of d with respect to $|$.

Consider an SDS $\mathbb{S} = (D_A, \ell, I, R)$ and a generating set Q of \mathbb{S} wrt an internal composition $|$. The *reduced 2-polygraph of \mathbb{S}* is the 2-polygraph, denoted by $\mathcal{R}(Q, D_A, \mathbb{S})$, whose set of 1-cells is Q and whose 2-cells are of the form $\gamma_{c, c'} : c|c' \Rightarrow R_Q(c \star_{\mathbb{S}} c')$, for any c, c' in Q such that $c|c' \notin D_A$.

Recall that a normalization strategy σ of $\mathcal{R}(Q, D_A, \mathbb{S})$ *computes* the constructor $C_{\mathbb{S}}$ if it is normalizing and it reduces any 1-cell $c_1|c_2|\dots|c_n$ in Q^* to $R_Q(c_1 \star_{\mathbb{S}} c_2 \star_{\mathbb{S}} \dots \star_{\mathbb{S}} c_n)$.

Coherent presentations. Recall from [8] the notion of coherent presentation. Let \mathcal{R} be a 2-polygraph and let \mathbf{M} be a monoid. For every 2-cell β in \mathcal{R} we will denote respectively by $s_1(\beta)$ and $t_1(\beta)$ the source and the target of β . We will denote by \mathcal{R}^{\top} the $(2, 1)$ -category freely generated by the 2-polygraph \mathcal{R} , that is the free 2-category enriched in groupoid generated by the set of 2-cells in \mathcal{R} . An *extended presentation* of \mathbf{M} is a pair $(\mathcal{R}, \mathcal{R}_3)$ made of a 2-polygraph \mathcal{R} that presents \mathbf{M} and a globular extension \mathcal{R}_3 of the $(2, 1)$ -category \mathcal{R}_2^{\top} freely generated by \mathcal{R} , that is a set of 3-cells $A : f \Rightarrow g$ relating 2-cells f and g in \mathcal{R}_2^{\top} , respectively denoted by $s_2(A)$ and $t_2(A)$ and satisfying the globular relations $s_1 s_2(A) = s_1 t_2(A)$ and $t_1 s_2(A) = t_1 t_2(A)$. We will denote by \mathcal{R}_3^{\top} the free $(3, 1)$ -category generated by an extended presentation $(\mathcal{R}, \mathcal{R}_3)$ of \mathbf{M} , that is, the 3-category generated by $(\mathcal{R}, \mathcal{R}_3)$ whose 2-cells and 3-cells are invertible. A *2-sphere* of \mathcal{R}_2^{\top} is a pair (f, g) of 2-cells of \mathcal{R}_2^{\top} such that $s_1(f) = s_1(g)$ and $t_1(f) = t_1(g)$. A *coherent presentation of \mathbf{M}* is an extended presentation $(\mathcal{R}, \mathcal{R}_2)$ of \mathbf{M} such that the cellular extension \mathcal{R}_3 is a *homotopy basis* of the $(2, 1)$ -category \mathcal{R}_2^{\top} , that is, for every 2-sphere γ of \mathcal{R}_2^{\top} , there exists a 3-cell in \mathcal{R}_3^{\top} with boundary γ .

Let \mathbb{S} be an associative SDS and Q be a generating set of \mathbb{S} such that the 2-polygraph $\mathcal{R}(Q, D_A, \mathbb{S})$ is terminating. If there exists a normalization strategy of $\mathcal{R}(Q, D_A, \mathbb{S})$ that computes $C_{\mathbb{S}}$, then the set of normal forms wrt $\mathcal{R}(Q, D_A, \mathbb{S})$ satisfies the cross-section property for $\mathbf{M}(\mathbb{S})$. In particular, if the leftmost normalization strategy σ^{\top} computes $C_{\mathbb{S}}$, then the 2-polygraph $\mathcal{R}(Q, D_A, \mathbb{S})$ can be extended into

3. String data structures for Chinese monoids

a coherent convergent presentation by adjunction of, [12]:

$$\begin{array}{ccc}
 c|c'|c''| & \xrightarrow{\sigma_{cc'c''}^\top} & R_Q(c \star_S c' \star_S c'') \\
 \searrow c|\gamma_{c',c''} & & \nearrow \sigma_{c|R_Q(c' \star_S c'')}^\top \\
 & & c|R_Q(c' \star_S c'')
 \end{array}
 \quad \text{for every } c, c', c'' \text{ in } Q.$$

3. STRING DATA STRUCTURES FOR CHINESE MONOIDS

Chinese SDS. The *Chinese monoid of rank* n , denoted by \mathbf{C}_n , [4], is presented by the *Chinese presentation* generated by the set $[n]$ and subject to the relations $zyx = zxy = yzx$, for $x \leq y \leq z$. A *Chinese staircase* on $[n]$ is a collection of boxes in right-justified rows filled with non-negative integers, where the rows (resp. the columns) are indexed with an initial segment of $[n]$ from top to bottom (resp. from right to left). We denote by t_{ij} (resp. t_i) the contents of the box in row i and column j (resp. row i and column i). Denote by Ch_n the set of Chinese staircases on $[n]$. We denote by R_r the map $\text{Ch}_n \rightarrow [n]^*$ that reads the rows of a staircase from right to left and from top to bottom, where every k -th row is reading as follows $(k1)^{t_{k1}}(k2)^{t_{k2}} \dots (k(k-1))^{t_{k(k-1)}}(k)^{t_k}$. We will call $\text{Im}(R_r)$ the set of *Chinese words*. For instance, a Chinese staircase t on $[3]$ is represented as on the right and $R_r(t) = 1^{t_1}(21)^{t_{21}}(2)^{t_2}(31)^{t_{31}}(32)^{t_{32}}(3)^{t_3}$.

		t_1	1
	t_2	t_{21}	2
t_3	t_{32}	t_{31}	3
3	2	1	

Consider the *right insertion* I_{Ch_n} , [4], and the *left insertion* J_{Ch_n} , [3], that insert an element of $[n]$ into an element of Ch_n , see Appendix A. Let $I_{\text{Ch}_n}^*$ (resp. $J_{\text{Ch}_n}^*$) be the insertion map $\text{Ch}_n \times [n]^* \rightarrow \text{Ch}_n$ associated to I_{Ch_n} (resp. J_{Ch_n}) wrt to the reading l_l (resp. l_r). This defines two SDSs $\mathbb{C}_n = (\text{Ch}_n, l_l, I_{\text{Ch}_n}, R_r)$ and $\mathbb{C}_n^o = (\text{Ch}_n, l_r, I_{\text{Ch}_n}, R_r)$ on the structure of staircases.

Reduced presentation of \mathbf{C}_n . We construct a finite semi-quadratic convergent presentation of the monoid \mathbf{C}_n by adding columns generators c_{yx} , for all $1 \leq x < y \leq n$, and square generators c_{xx} , for all $1 < x < n$. We will denote by Q_n the set defined by

$$Q_n = \{c_{yx} \mid 1 \leq x < y \leq n\} \cup \{c_{xx} \mid 1 < x < n\} \cup \{c_1, \dots, c_n\},$$

where c_1, \dots, c_n represent the initial generators $1, \dots, n$. The reduced 2-polygraph $\mathcal{R}(Q_n, \text{Ch}_n, \mathbb{C}_n)$ of the SDS \mathbb{C}_n is the 2-polygraph whose set of 1-cells is Q_n and whose 2-cells are $\gamma_{u,v} : c_u c_v \Rightarrow R_r(c_u \star_{\mathbb{C}_n} c_v)$ such that $c_u c_v$ is not a Chinese word.

3.1. Theorem. *The SDS \mathbb{C}_n satisfies the following properties*

- i) *the 2-polygraph $\mathcal{R}(Q_n, \text{Ch}_n, \mathbb{C}_n)$ is a finite semi-quadratic presentation of the monoid \mathbf{C}_n ,*
- ii) *the SDS \mathbb{C}_n is associative,*
- iii) *the 2-polygraph $\mathcal{R}(Q_n, \text{Ch}_n, \mathbb{C}_n)$ is a convergent,*
- iv) *the set of Chinese words satisfies the cross-section property for the monoid \mathbf{C}_n .*

REFERENCES

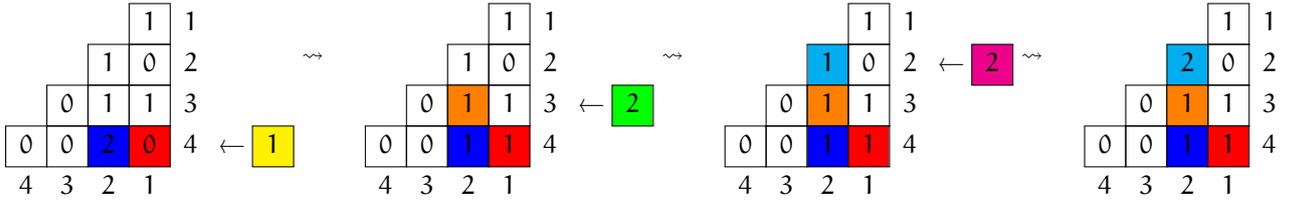
- [3] A. J. Cain, R. D. Gray, and A. Malheiro. Rewriting systems and biautomatic structures for Chinese, hypoplactic, and Sylvester monoids. *Internat. J. Algebra Comput.*, 25(1-2):51–80, 2015.
- [4] J. Cassaigne, M. Espie, D. Krob, J.-C. Novelli, and F. Hivert. The Chinese monoid. *Int. J. Algebra Comput.*, 11(3):301–334, 2001.
- [5] Y. Chen and J. Qiu. Gröbner-Shirshov basis for the Chinese monoid. *J. Algebra Appl.*, 7(5):623–628, 2008.
- [6] G. Duchamp and D. Krob. Plactic-growth-like monoids. In *Words, languages and combinatorics, II (Kyoto, 1992)*, pages 124–142. World Sci. Publ., River Edge, NJ, 1994.
- [7] W. Fulton. *Young tableaux*, volume 35 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1997. With applications to representation theory and geometry.
- [8] S. Gaussent, Y. Guiraud, and P. Malbos. Coherent presentations of Artin monoids. *Compos. Math.*, 151(5):957–998, 2015.
- [9] E. Güzel Karpuz. Complete rewriting system for the Chinese monoid. *Appl. Math. Sci. (Ruse)*, 4(21-24):1081–1087, 2010.
- [10] N. Hage. Finite convergent presentation of plactic monoid for type C. *Internat. J. Algebra Comput.*, 25(8):1239–1263, 2015.
- [11] N. Hage and P. Malbos. Knuth’s coherent presentations of plactic monoids of type A. *Algebras and Representation Theory*, 20(5):1259–1288, Oct 2017.
- [12] N. Hage and P. Malbos. Coherence of monoids by insertions. preprint, 2018.
- [13] A. Lascoux and M.-P. Schützenberger. Sur une conjecture de H. O. Foulkes. *C. R. Acad. Sci. Paris Sér. A-B*, 286(7):A323–A324, 1978.

A. INSERTION'S ALGORITHMS

The right insertion. The right insertion algorithm introduced in [4, Subsection 2.2] is the insertion map $I_{\text{Ch}_n} : \text{Ch}_n \times [n] \rightarrow \text{Ch}_n$ that sends (t, x) to the Chinese staircase $I_{\text{Ch}_n}(tx)$ as follows. Write $t = (t', R_1)$, where R_1 is the bottom row of t , and t' is the staircase obtained by the remaining rows of t . Let z be the greatest index of t . If $x > z$, then $I_{\text{Ch}_n}(tx) = t$. If $x = z$, then $I_{\text{Ch}_n}(tx) = (t', R'_1)$, where R'_1 is obtained from R_1 by adding 1 to t_z . If $x < z$, let y be maximal such that the entry in column y of R_1 is non-zero or if such a y does not exist, set $y = x$. Three cases appear:

- i) If $x \geq y$, then $I_{\text{Ch}_n}(tx) = (I_{\text{Ch}_n}(t'x), R_1)$.
- ii) If $x < y < z$, then $I_{\text{Ch}_n}(tx) = (I_{\text{Ch}_n}(t'y), R'_1)$, where R'_1 is obtained from R_1 by subtracting 1 from t_{zy} and adding 1 to t_{zx} .
- iii) If $x < y = z$, then $I_{\text{Ch}_n}(tx) = (t', R'_1)$, where R'_1 is obtained from R_1 by subtracting 1 from t_z and adding 1 to t_{zx} .

For instance, consider the following example



The left insertion. The left insertion algorithm introduced in [3, Algorithm 3.5] is the insertion map $J_{\text{Ch}_n} : \text{Ch}_n \times [n] \rightarrow \text{Ch}_n$ that sends (t, x) to the Chinese staircase $J_{\text{Ch}_n}(xt)$ as follows. Take an element y from $[n] \cup \{\infty\}$, initially set to ∞ . There are two steps. In the first step, for $i = 1, \dots, x - 1$, iterate the following. If every entry in the row i is empty, do nothing. Otherwise, let z be minimal such that $t_{iz} > 0$. Then:

- i) If $y = \infty$, then
 - (a) If $z < i$, decrement t_{iz} by 1, increment t_i by 1, and set $y = z$.
 - (b) If $z = i$, decrement t_i by 1, and set $y = z$.
- ii) If $y \neq \infty$, then
 - (a) If $z < y$, decrement t_{iz} by 1, increment t_{iy} by 1, and set $y = z$.
 - (b) If $z \geq y$, do nothing.

In the second step, for $i = x$, if $y = \infty$, then increment t_i by 1. Otherwise, decrement t_{iy} by 1. Finally, output the current staircase.