Coherence modulo relations

Benjamin Dupont and Philippe Malbos

Institut Camille Jordan, University of Lyon, France, {bdupont,malbos}@math.univ-lyon1.fr

Abstract – The computation of minimal convergent presentations for monoids, categories or higher-dimensional categories appear in low-dimensional combinatorial problems on these structures, such as coherence problems. A method to compute coherent presentations using convergent string rewriting systems was developed following works of Squier. In this approach, coherence results are formulated in terms of confluence diagrams of critical pairs. This work proposes an extension of these methods to string rewriting systems modulo.

1 Introduction

The computation of minimal convergent presentations for monoids, categories or higher-dimensional categories appear in low-dimensional combinatorial problems on these structures, such as coherence problems. A method to compute coherent presentations using convergent string rewriting systems was developed following works of Squier, see [9, 10]. In this approach, coherence results are formulated in terms of confluence diagrams of critical pairs. This work proposes an extension of these methods to algebraic or categorical structures having additional algebraic axioms, such as commutation, linearity or inverses. Using a notion of rewriting modulo, we show how to compute coherent presentations modulo algebraic axioms.

Rewriting modulo was developed in several approaches. The rules are split into two parts: oriented rules in a set R and non-oriented equations in a set E. The most naive approach of rewriting modulo is to use a rewriting system R/E consisting in rewriting on congruence classes modulo E, but this appears unefficient for analysis of confluence, see [1, Chapter 11]. Another approach of rewriting modulo has been considered by Huet in [11] where rewriting paths does not involve equivalence steps, and confluence is formulated modulo equivalence. Jouannaud and Kirchner enlarged this approach in [12] by providing completion methods for any rewriting system between R and R/E. Several other approaches have also been developed for term rewriting systems modulo to deal with various equational theories, see [2, 14, 17].

In this work, we extend Huet's approach to prove coherence results modulo algebraic relations, e.g. inverses for rewriting in groups, or commutation for linear rewriting. Indeed, in most cases, algebraic relations such as commutation cannot be oriented in a terminating way. Moreover, rewriting modulo can be used to delete some critical branchings that should not be considered in the analysis of coherence. This is the case for the computation of coherent presentations for algebraic structures such as groups or algebras.

Known approaches of rewriting for groups are mainly based on a presentation of groups as monoids with explicit inverses and explicit rules for inverse axioms. The SRS is thus defined on the set of generators of the group, their formal inverses, and the explicit rules for inverses, [3–6, 15]. However, coherent presentations of groups have to take into account that the presentation is modulo these inverse relations. The objective is to study confluence modulo the confluence diagrams induced by these relations, and to consider rewriting steps in the free group. This approach corresponds to rewriting on congruence classes modulo the equivalence given by the inverse relations, and thus is not suitable to study confluence. For this reason, we consider the weaker theory of rewriting modulo introduced by Huet. One of the main applications is to extend the coherent results obtained by rewriting methods on Artin-Tits monoids in [7] to Artin-Tits groups.

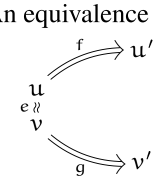
The second example of algebraic structure we consider is the axiomatic of vector spaces in the theory of linear rewriting developed in [8]. Actually, the critical branchings in linear rewriting are defined modulo the axioms of vector spaces, namely abelian groups and distributivity of the multiplicative law. For instance, if we denote by E the set of axiomatic rules of vector spaces and R is a linear rewriting system with two rules $3x \Rightarrow 2y$ and $2x \Rightarrow z$, then the branching $2y \Leftarrow 3x = 2x + x \Rightarrow z + x$ can be interpreted as a branching of R modulo E. In this way, the coherence result obtained on algebras in [8] can be formulated in terms of rewriting modulo.

This work presents a construction of coherent extensions of SRS modulo. In a first part, we recall the notion of confluence modulo as introduced by Huet, [11]. Then, we introduce the notion of coherent extension modulo, that corresponds to homotopy bases of SRS as defined in [10] when the set of axioms is empty. It is defined by a set of 3-cells modulo tiling all the spheres created by rewriting paths which are parallel modulo the axioms. In the last section, we enounce a generalization of Squier's coherence theorem to confluent SRS modulo. The proof of this result is given in Appendix A as well as some recalls on the categorical language on SRS used in this work in Appendix B.

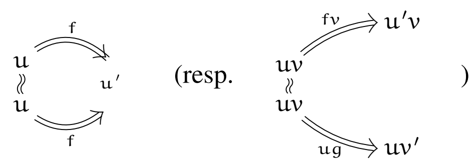
2 Rewriting modulo

Let us recall the notion of rewriting modulo a set of relations. In the sequel, all the SRS considered are defined on a same alphabet X. Given two SRS R and E, a rewiting with respect to R modulo relations defined by E consists in rewriting using rules of R on congruence classes modulo E. This corresponds to studying the rewriting system R/E defined by $u \Rightarrow_{R/E} v$ if and only if there exists strings u' and v' on X such that $u =_E u' \Rightarrow_R v' =_E v$. However, studying confluence of this rewriting system is complicated as explained in [1], so we use a weaker notion of confluence modulo as introduced by Huet in [11]. Whenever it exists, we denote by $\hat{u}$ a normal form of a string u on X with respect to R.

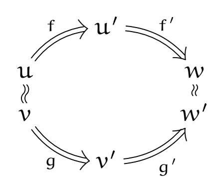
Equivalence modulo. Let consider the free $(2,1)$-category $E^\top$ generated by E (see Appendix B for categorical constructions). The 2-cells of $E^\top$ will be called *equivalences modulo* E. An equivalence modulo E of length equals to 1 is called a *one-step equivalence*. We denote by $\approx_E$ the equivalence relation generated by E. A *branching modulo* E of the SRS R is a pair (f, g) of 2-cells of the free 2-category R^* such that $s_1(f) \approx_E s_1(g)$ as depicted by the diagram on the side. We do not distinguish the branchings (f, g) and (g, f). Such a branching (f, g) is *local* if $\ell(f), \ell(g), \ell(e) \leqslant 1$ and $\ell(f) + \ell(g) + \ell(e) = 2$.



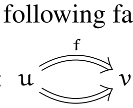
An *aspherical (resp. Peiffer) branching modulo* E of R is a pair (f, f) (resp. (fv, ug)) of 2-cells of R^* depicted by



A branching (f, g) is *confluent modulo* E if there exists 2-cells f' and g' in R^*, as in the diagram on the side. (f', g') is called a confluence modulo E. We will denote this by $u' \overset{E}{\vee} v'$. We say that R is *confluent modulo* E if all of its branchings are confluent modulo E.



Local branching. Local branchings belong to one of the following families:

i) *local aspherical* branchings, for a rewriting step f of R: 

ii) *local Peiffer* branchings: for rewriting steps f, g of R (resp. for a rewriting step f of R and a one-step equivalence e of E):

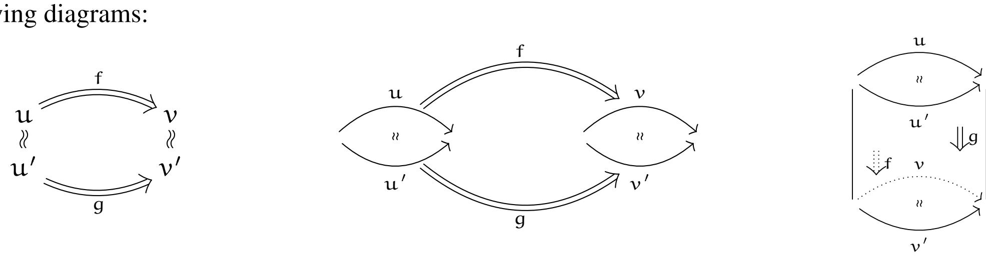
iii) *overlapping* branchings are the remaining local branchings, in which we distinguish two families:

A *critical branching modulo* E is an overlapping local branching that is minimal for the order generated by the relations $(f, g) \preccurlyeq (ufv, ugv)$.

Local Confluence modulo. A SRS R is *locally confluent modulo* a SRS E if any of its local branchings is confluent modulo E. Note that any aspherical and Peiffer branching being confluent modulo E, local confluence modulo E is equivalent to the confluence of overlappings modulo E. Huet show that under the assumption that the composite $\Rightarrow_R \cdot \approx_E$ is terminating, then R is confluent modulo E if and only if any overlapping branching of R modulo E is confluent modulo E, [11]. Under the same assumption, he shows that a SRS R is locally confluent modulo a SRS E if and only if any critical branching of R modulo E is confluent modulo E, [11].

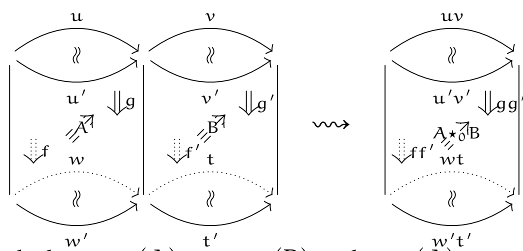
3 Coherent extensions modulo

Given two SRS R and E, a 2-*sphere modulo* E in the free $(2,1)$-category $R^\top$ is a pair (f, g) of 2-cells in $R^\top$ which are *parallel modulo* E, that is, $s_1(f) \approx_E s_1(g)$ and $t_1(f) \approx_E t_1(g)$ and such that f or g is not trivial. Note that the case f and g trivial produce a 2-sphere in $E^\top$. These 2-spheres do not fit in the construction of coherence extensions modulo E. A 2-sphere modulo E will be pictured by one of the following diagrams:

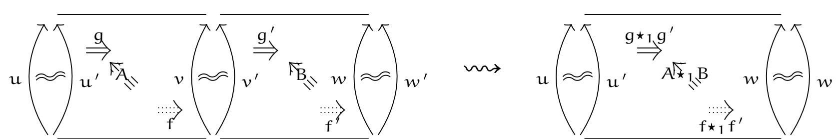


We will denote by $\mathrm{Sph}_E(R)$ the set of 2-spheres modulo E in $R^\top$. A *cellular extension of* $R^\top$ *modulo* E is a set Γ equipped with a map $\gamma : \Gamma \longrightarrow \mathrm{Sph}_E(R)$, whose elements are called 3-*cells modulo* E. We say that Γ is *coherent* if the map γ is surjective. A 3-cell A modulo E filling a 2-sphere (f, g) modulo E will be denoted by $A : f \Rrightarrow_E g$. We say that f (resp. g) is the 2-source (resp. 2-target) of A and we denote it by $s_2(A)$ (resp. $t_2(A)$). We define formal compositions $\star_0, \star_1, \star_2$ of 3-cells modulo E in a cellular extension Γ as pasting operations defined as follows:

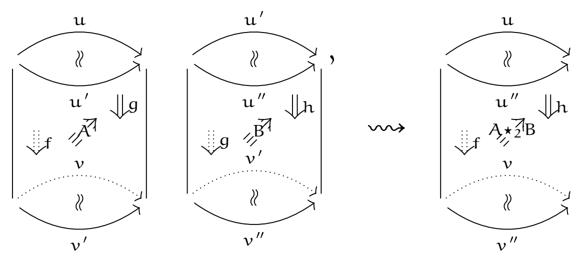
i) Given A and B in Γ, one defines $A \star_0 B$ as the 3-cell modulo E tiling the cylinder as follows:



ii) Given A and B in Γ such that $t_1 s_2(A) = s_1 s_2(B)$ and $t_1 t_2(A) = s_1 t_2(B)$, one defines a 3-cell modulo E denoted by $A \star_1 B$ tiling the following composite cylinder:



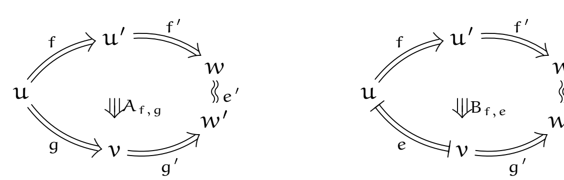
iii) Given A and B in Γ such that $t_2(A) = s_2(B)$, one defines a 3-cell modulo E denoted by $A \star_2 B$ tiling the cylinder obtained as follows:



Let Γ be a cellular extension of $R^\top$ modulo E. We will denote by $\mathcal{C}(\Gamma)$ the closure of Γ with respect to compositions $\star_0$, $\star_1$ and $\star_2$ of 3-cells of Γ and their formal inverses A^{-1} for $A \in \Gamma$ quotiented by the exchange relations $(A \star_i B) \star_j (A' \star_i B') = (A \star_j A') \star_i (B \star_j B')$ for any $0 \leqslant i < j \leqslant 2$ and the inverse relations $A \star_i A^- = 1_{s_i(A)}$ for any A in Γ and $i = 0, 1, 2$. When $\mathcal{C}(\Gamma)$ is a coherent extension of $R^\top$ modulo E, we say that Γ is *acyclic modulo* E.

4 Coherence from confluence modulo

Squier's completion. Suppose that R is a confluent SRS modulo a SRS E. A *Squier's completion modulo* E of R is a cellular extension modulo E of $R^\top$ whose elements are the 3-cells

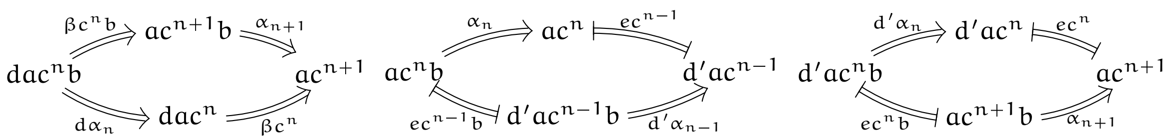


for any critical branching (f, g) and (f, e) of R modulo E, where f, g are rewriting steps of R and e is a one-step equivalence of E. Note that such a completion is not unique in general and depends on the rewriting sequences f', g' and the equivalence e' used to obtain the confluence diagrams.

4.1. Theorem (Coherence modulo). *Let* R *be a SRS confluent modulo a SRS* E *such that* $\Rightarrow_{\mathrm{R}} \cdot \approx_{\mathrm{E}}$ *is terminating, then any Squier's completion of* R *modulo* E *is acyclic.*

A proof of this result is given in Appendix A. As a consequence of Theorem 4.1, with the same hypothesis, one can prove that if Γ is an acyclic extension of $\mathrm{E}^{\top}$ then $\mathcal{C}(\mathcal{S}(\mathrm{R},\mathrm{E})) \sqcup \Gamma$ is a coherent extension of $(\mathrm{R} \sqcup \mathrm{E})^{\top}$. In particular when E is convergent, we fix a Squier completion $\mathcal{S}(\mathrm{E})$ and we get that $\mathcal{C}(\mathcal{S}(\mathrm{R},\mathrm{E})) \sqcup \mathcal{S}(\mathrm{E})$ is a coherent extension of $(\mathrm{R} \sqcup \mathrm{E})^{\top}$.

Example. Let R be the SRS on $\mathrm{X} = \{a, b, c, d, d'\}$ defined by the rules $ab \overset{\alpha_0}{\Rightarrow} a$ and $da \overset{\beta}{\Rightarrow} ac$. We complete the SRS R into a confluent SRS by adding the rules $ac^n b \overset{\alpha_n}{\Rightarrow} ac^n$ for all n in $\mathbb{N}$ in R. Let us consider the SRS E defined on X with the rule $d'a \overset{e}{\Rightarrow} ac$. A Squier's completion of R modulo E is then given by a family of 3-cells modulo E tiling the following confluence modulo diagrams:



Note that up to a diagram rotation, the last two families of confluence diagrams are the same, so the coherent extension of R modulo E consists in the two families of 3-cells given by the first two confluence modulo diagrams. We recover the coherent extension of the SRS $\mathrm{R} \sqcup \mathrm{E}$ given in [13].

Finiteness conditions modulo. In the case where E is empty, Theorem 4.1 is the Squier's theorem for SRS, [16], see [10] for a polygraphic proof. From Squier's result, it follows the homotopical finiteness from convergence: if a monoid admits a presentation by a finite convergent SRS, then it has finite derivation type (FDT). From Theorem 4.1, we deduce a new finiteness condition modulo. If a monoid admits a presentation by a finite convergent SRS R modulo E, then it has *FDT modulo* E, that is, R admits a finite cellular extension Γ modulo E such that $\mathcal{C}(\Gamma)$ is acyclic. If the SRS R modulo E has FDT modulo E and E has FDT, then the SRS $\mathrm{R} \sqcup \mathrm{E}$ has FDT. In particular, if R is a finite confluent SRS modulo a finite convergent SRS such that $\Rightarrow_{\mathrm{R}} \cdot \approx_{\mathrm{E}}$ terminating, then $\mathrm{R} \sqcup \mathrm{E}$ has FDT.

5 Conclusion and work in progress

In this work, we have presented a coherence result for SRS modulo a set of axioms. This result is based on the notion of confluence modulo introduced by Huet. However, completion procedures for such SRS are missing. We expect that some completion methods given in [2, 12, 17] could be adapted to compute Squier's completion of non confluent SRS modulo. In particular, with the axioms of group, naive completion modulo induces infinitely many completion steps due to overlapping branchings between a rule and an equivalence obtained by adding elements of the form xx^{-1}. The objective is to define an appropriate completion procedure allowing to avoid this infinite completion. One approach is to consider a restriction of local obstruction of confluence by considering rewriting step conditioned by algebraic context.

REFERENCES

[1] Franz Baader and Tobias Nipkow. *Term rewriting and all that*. Cambridge University Press, 1998.

[2] Leo Bachmair and Nachum Dershowitz. Completion for rewriting modulo a congruence. *Theoretical Computer Science*, 67(2):173 – 201, 1989.

[3] Fabienne Chouraqui. Rewriting systems and embedding of monoids in groups. *Groups Complex. Cryptol.*, 1(1):131–140, 2009.

[4] Fabienne Chouraqui. The Knuth-Bendix algorithm and the conjugacy problem in monoids. *Semigroup Forum*, 82(1):181–196, 2011.

[5] Volker Diekert, Andrew Duncan, and Alexei G. Myasnikov. Cyclic rewriting and conjugacy problems. *Groups Complex. Cryptol.*, 4(2):321–355, 2012.

[6] Volker Diekert, Andrew J. Duncan, and Alexei G. Myasnikov. Geodesic rewriting systems and pregroups. In *Combinatorial and geometric group theory*, Trends Math., pages 55–91. Birkhäuser/Springer Basel AG, Basel, 2010.

[7] Stéphane Gaussent, Yves Guiraud, and Philippe Malbos. Coherent presentations of Artin monoids. *Compos. Math.*, 151(5):957–998, 2015.

[8] Yves Guiraud, Eric Hoffbeck, and Philippe Malbos. Convergent presentations and polygraphic resolutions of associative algebras. arXiv:1406.0815v2, December 2017.

[9] Yves Guiraud and Philippe Malbos. Coherence in monoidal track categories. *Math. Structures Comput. Sci.*, 22(6):931–969, 2012.

[10] Yves Guiraud and Philippe Malbos. Polygraphs of finite derivation type. *Math. Structures Comput. Sci.*, 28(2):155–201, 2018.

[11] Gérard Huet. Confluent reductions: abstract properties and applications to term rewriting systems. *J. Assoc. Comput. Mach.*, 27(4):797–821, 1980.

[12] Jean-Pierre Jouannaud and Helene Kirchner. Completion of a set of rules modulo a set of equations. In *Proceedings of the 11th ACM SIGACT-SIGPLAN Symposium on Principles of Programming Languages*, POPL '84, pages 83–92, New York, NY, USA, 1984. ACM.

[13] Yves Lafont. A new finiteness condition for monoids presented by complete rewriting systems (after Craig C. Squier). *J. Pure Appl. Algebra*, 98(3):229–244, 1995.

[14] Claude Marché. Normalized rewriting: an alternative to rewriting modulo a set of equations. *J. Symbolic Comput.*, 21(3):253–288, 1996.

[15] John Pedersen and Margaret Yoder. Term rewriting for the conjugacy problem and the braid groups. *J. Symbolic Comput.*, 18(6):563–572, 1994.

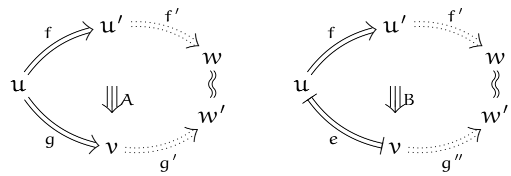
[16] Craig C. Squier, Friedrich Otto, and Yuji Kobayashi. A finiteness condition for rewriting systems. *Theoret. Comput. Sci.*, 131(2):271–294, 1994.

[17] Patrick Viry. Rewriting modulo a rewrite system. Technical report, 1995.

A. Proof of Theorem 4.1

Proof. Let us consider a Squier's completion $\mathcal{S}(R, E)$ of R modulo E.
Step 1. Prove that, for every local branching (f, g) and (f, e) of R modulo E with f, g in R and e in E,

there exist 2-cells f' and g' in R^* and 3-cells $A : f \star_1 f' \Rrightarrow g \star_1 g'$ and $B : f \star_1 f' \Rrightarrow e \star_1 g'$ modulo E in $\mathcal{C}(\mathcal{S}(R, E))$, as in the following diagram:



In the case of a local aspherical branching, we set A as an identity. For a local Peiffer branching (f, g) with f, g in R, we can choose f' and g' such that $f \star_1 f' = g \star_1 g'$ and we set A an identity. For a local Peiffer branching (f, e) with f in R and e in E, we can choose f' as the empty 2-cell, $g'' = f$ and the right equivalence being e so that B is also an identity. Moreover, if we have an overlapping branching (f, g) (resp. (f, e)) that is not critical, we have $(f, g) = (uhv, ukv)$ (resp. $(f, e) = (uhv, ue'v)$) for some u, v in X^* such that both (h, k) and (h, e') are critical. We consider the 3-cells $A' : f \star_1 f' \Rrightarrow_E g \star_1 g'$ and $B' : f \star_1 f' \Rrightarrow_E e \star_1 g''$ corresponding respectively to the critical branchings (h, k) and (h, e'). We conclude by setting $f' = uh'v \quad g' = uk'v \quad g'' = ue'v \quad A' = u \star_0 A' \star_0 v \quad B = u \star_0 B' \star_0 v$.

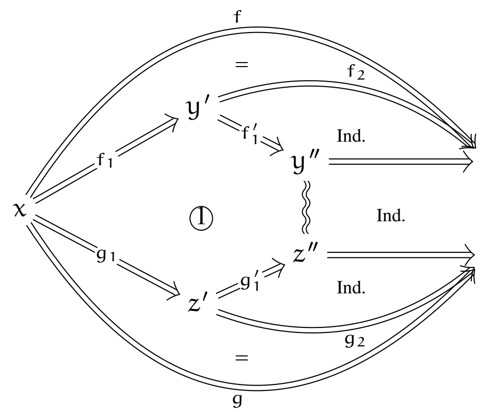
Step 2. We prove that, for any 2-cells $f : x \Rightarrow y$ and $g : x \Rightarrow z$ of R^*, there exists a 3-cell modulo E from f to g in $\mathcal{C}(\mathcal{S}(R, E))$. To do this, we decompose the 2-cells f and g into $f = f_1 \star_1 f_2$ and $g = g_1 \star_1 g_2$ where $f_1 : x \Rightarrow y'$ and $g_1 : x \Rightarrow z'$ are rewriting steps of R, and f_2,g_2 are in R^*. Then (f_1, g_1) is a local branching of R modulo E and we use local confluence modulo E to get 1-cells y'' and z'' in X^* and 2-cells $f_1' : y' \Rightarrow y''$ and $g_1' : z' \Rightarrow z''$ in R^* with $y'' \approx_E z''$. Using Step 1, we get a 3-cell $A : f_1 \star_1 f_1' \Rrightarrow_E g_1 \star_1 g'1$. We construct a 3-cell modulo E from f to g using Noetherian induction principle from [11] defined as follows: we fix an auxiliary string rewriting system R^{aux} with only one 0-cell, whose 1-cells are the pairs (x, y) of elements of X^*. R^{aux} contains a 2-cell $(x, y) \Rightarrow_{R^{aux}} (x', y')$ in any of the following situation:

- there exist 2-cells $x \Rightarrow x'$ and ($y \Rightarrow y'$ or $x \Rightarrow y'$) in R^*;
- there exists a 2-cell $y \Rightarrow y'$ in R and an equivalence $x \approx_E x'$ in E;
- there exist 2-cells $y \Rightarrow x'$ and $y \Rightarrow y'$ in R^*;
- $x \overset{e}{\approx}_E y \approx_E x' \overset{e'}{\approx}_E y'$ and $\ell(e) < \ell(e')$.

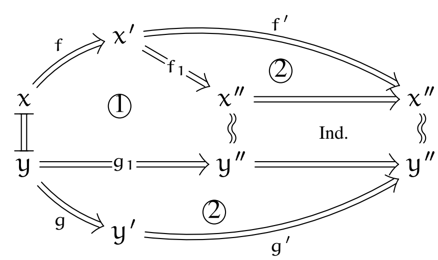
Following [11, Proposition 2.2], if $\Rightarrow_R \cdot \approx_E$ is terminating, then so is R^{aux}. Let us apply Noetherian induction on R^{aux} with the following property:

$$\mathcal{P}(x, y) : \quad x \approx_E y \;\Rightarrow\; \forall\, x', y' \mid x \overset{*}{\Rightarrow}_P x' \;\&\; y \overset{*}{\Rightarrow}_R \Rightarrow x' \overset{E}{\vee} y'$$

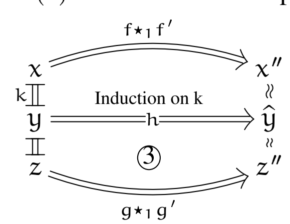
This leads to the diagram on the right which enables to construct a 3-cell $A : f \Rrightarrow_E g$ in $\mathcal{C}(\mathcal{S}(R, E))$.



Step 3. We now prove that for each rewriting steps $f : x \Rightarrow x'$ and $g : y \Rightarrow y'$ in R such that $x \overset{e}{\approx}_E y$, there exist 2-cells $f' : x' \Rightarrow x''$, $g' : y' \Rightarrow y''$ in R^* and a 3-cell modulo E from $f \star_1 f'$ to $g \star_1 g'$. We will prove the result by induction on $\ell(e)$. If $\ell(e) = 0$, this is Step 1. Suppose that $\ell(e) = 1$, that is $x \vDash\!\dashv y$. By local confluence of R modulo E, looking at the local branching (f, e), we get the existence of 2-cells $f_1 : x' \Rightarrow x''$, $g_1 : y \Rightarrow y''$ in R^* with $x'' \approx_E y''$. By Step 1, there exists a 3-cell modulo E in $\mathcal{C}(\mathcal{S}(R, E))$ from $f \star_1 f_1$ to g_1. We construct the 3-modulo E from $f \star_1 f'$ to $g \star_1 g'$ using Noetherian's induction as illustrated by the following diagram.

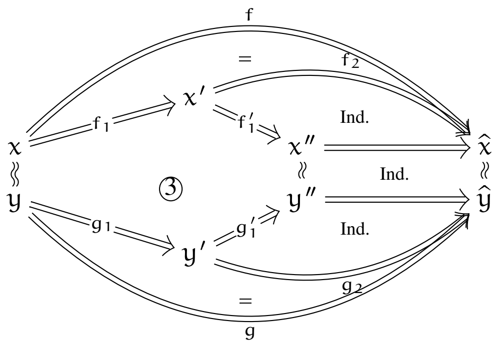


Hence the result is proved for $\ell(e) = 1$. Suppose the result proved for $\ell(e) = k > 1$ and let us prove the result for $\ell(e) = k + 1$. Suppose that $x \overset{k+1}{\vDash\!\dashv} y$, we decompose this reduction by $x \overset{k}{\vDash\!\dashv} z \vDash\!\dashv y$. We fix a 2-cell $h : y \Rightarrow \hat{y}$ in R^*. By confluence modulo E, there exists 2-cells $f' : x' \Rightarrow x''$ and $g' : z' \Rightarrow z''$ in R^* such that $x'' \approx_E \hat{y} \approx_E z''$. We construct a 3-cell modulo E between $f \star_1 f'$ and $g \star_1 g'$ as depicted on the diagram on the right.

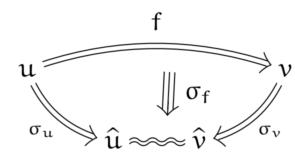


Step 4. Now, let us prove that for any 2-cells $f : x \Rightarrow \hat{x}$ and $g : y \Rightarrow \hat{y}$ with $x \overset{e}{\approx}_E y$, there exists a 3-cell $A : f \Rrightarrow_E g$ modulo E in $\mathcal{C}(\mathcal{S}(R, E))$. Let us first write $f = f_1 \star_1 f_2$ and $g = g_1 \star_1 g_2$ where $f_1 : x \Rightarrow x'$ and $g_1 : y \Rightarrow y'$ are rewriting steps in R and $f_2 : x' \Rightarrow \hat{x}$, $g_2 : y' \Rightarrow \hat{y}$ are 2-cells in R^*. Using the confluence modulo E on the triple (f_1, e, g_1), we get the existence of 1-cells x'', y'' and 2-cells $f'_1 : x' \Rightarrow x''$ and $g'_1 : y' \Rightarrow y''$ such that $x'' \approx_E y''$. According to Step 3, there exists a 3-cell modulo E in $\mathcal{C}(\mathcal{S}(R, E))$ from $f_1 \star_1 f'_1$ to $g_1 \star_1 g'_1$. By Noetherian induction principle, we get the

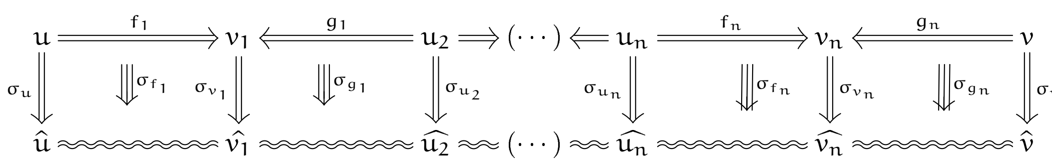
following diagram allowing us to construct the 3-cell A:



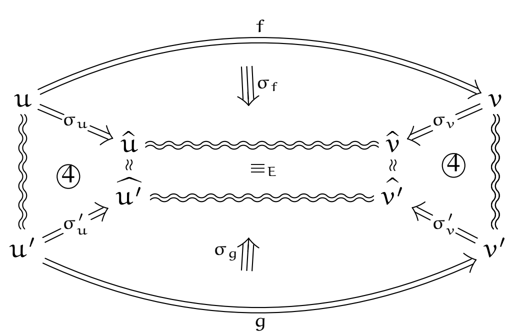
Step 5. We prove that every 2-sphere modulo E of $R^{\top}$ is the boundary of a 3-cell modulo E of $\mathcal{C}(\mathcal{S}(R,E))$. First, let us consider a 2-cell $f : u \Rightarrow v$ in R^*. Using confluence modulo E of R, there exist 2-cells in R^* $\sigma_u : u \Rightarrow \hat{u}$ and $\sigma_v : v \Rightarrow \hat{v}$ in R^* such that $\hat{u} \approx_E \hat{v}$. By construction, the 2-cells $f \star_1 \sigma_v$ and σ_u are parallel modulo E and their respective targets are normal forms. By Step 4, there exists a 3-cell modulo E in $\mathcal{C}(\mathcal{S}(R,E))$ from $f \star_1 \sigma_v$ to σ_u as in the diagram on the right.



Now, let us consider a 2-cell $f : u \Rightarrow v$ of $R^{\top}$. By construction of $R^{\top}$, the 2-cell f can be decomposed in a non unique way into a zigzag sequence $f_1 \star g_1^{-1} \star_1 \cdots \star_1 f_n \star_1 g_n^{-1}$ with source u and target v where each f_i and g_i is a 2-cell of R^*. We define a 3-cell modulo $\sigma_f : f \star_1 \sigma_v \Rrightarrow_E \sigma_u$ in $\mathcal{C}(\mathcal{S}(R,E))$ as the following composition:



Proceeding similarly for any other 2-cell $g : u' \Rightarrow v'$ of $R^{\top}$, we get a 3-cell $\sigma_g : g \star_1 \sigma_{v'} \Rrightarrow_E \sigma_u$ in $\mathcal{C}(\mathcal{S}(R,E))$. In this way, for any 2-sphere (f,g) modulo E in $R^{\top}$, there exists a 3-cell modulo $f \Rrightarrow_E g$ in $\mathcal{C}(\mathcal{S}(R,E))$ given by the following composition:



□

B. Categorical formulation of string rewriting systems

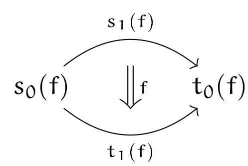
In this work, the constructions on SRS are formulated in categorical language. In this part, we recall the notions used in the text for a reader unfamiliar to this language. We refer to [10] for a deeper presentation of categorical formulation of SRS.

1-categories of strings. Given an alphabet X, we denote by X^* the free monoid generated by X. This monoid can be seen as the free 1-category generated by X, that is a 1-category with only one 0-cell and whose 1-cells are strings made of elements of X. Having only one 0-cell, any two 1-cells of X^* are composable and the composition corresponds to concatenation of strings. This concatenation is associative and unitary with the empty string as unit.

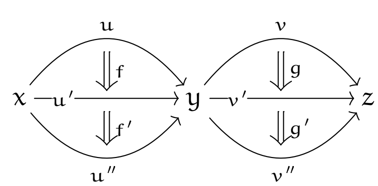
2-categories of rewriting steps. Recall that a 2-category $\mathcal{C}$ is defined by a set $\mathcal{C}_0$ of 0-cells, a set $\mathcal{C}_1$ of 1-cells and a set $\mathcal{C}_2$ of 2-cells and equipped with two compositions $\star_0$ for 1-cells and 2-cells and $\star_1$ for 2-cells. A 2-category is equipped with source and target maps making it a 2-graph, that is a digram in the category of sets:

$$\mathcal{C}_0 \overset{s_0}{\underset{t_0}{\leftleftarrows}} \mathcal{C}_1 \overset{s_1}{\underset{t_1}{\leftleftarrows}} \mathcal{C}_2$$

where the maps satisfy the globular relations: $s_0 s_1 = s_0 t_1$ and $t_0 s_1 = t_0 t_1$. For any $1 \leqslant i < j \leqslant 2$, the i-cell $s_i(f)$ (resp. $t_i(f)$) is called the i-*source* (resp. i-*target*) of a j-cell f. A 2-cell f in $\mathcal{C}$ can be pictured by

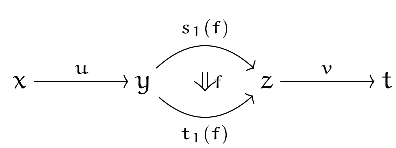


The composition $\star_0$ and $\star_1$ are associative and unitary and compatible with source and target maps. They also satisfy the exchange law, that is, for any situation



the equality $(f \star_0 g) \star_1 (f' \star_0 g') = (f \star_1 f') \star_0 (g \star_1 g')$ holds.

Given a SRS R on an alphabet X, one can construct the free 2-category generated by R, denoted by R^* and defined as follows. It has only one 0-cell, its 1-cells are strings on X and its 2-cells are rewriting paths of R. The $\star_0$-composition in R^* corresponds to concatenation of strings, and the $\star_1$-composition is the sequential composition of rewritings of R. Each 2-cell f of R^* can be decomposed into a sequence $f = f_1 \star_1 f_2 \star_1 \ldots \star_1 f_k$, where each f_i is a 2-cell corresponding to a rewriting step of the form:



that we will denote by ufv. The length of a 2-cell f in R^*, denoted by $\ell(f)$ is the minimal number of rewriting steps in any $\star_1$-decomposition of f. We denote by $u \Rightarrow_R v$ if there exists a 2-cell in R^* of length 1, that is u rewrites to v in one R-step.

$(2,1)$**-categories of equivalence.** Let E be a SRS on an alphabet X. The free $(2,1)$-category generated by E, denoted by $E^\top$, is the free 2-category on E in which all the 2-cells are invertible with respect to the $\star_1$-composition. That is its 0-cells, 1-cells and 2-cells are those of E^*, and any 2-cell f of $E^\top$ has an inverse $f^- : t_1(f) \Rightarrow s_1(f)$ with respect the $\star_1$-composition satisfying the relations $f \star_1 f^- = 1_{s_1(f)}$ and $f^- \star_1 f = 1_{t_1(f)}$. The 2-cells of the $(2,1)$-category $E^\top$ corresponds to elements of the equivalence relation generated by E, that we will denote by $\approx_E$.