



**HAL**  
open science

# Lindenmayer graph languages, first-order theories and expanders

Teodor Knapik

► **To cite this version:**

Teodor Knapik. Lindenmayer graph languages, first-order theories and expanders. 2024. hal-04580653

**HAL Id: hal-04580653**

**<https://hal.science/hal-04580653v1>**

Preprint submitted on 20 May 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

# Lindenmayer graph languages, first-order theories and expanders

Teodor Knapik<sup>1</sup>

<sup>1</sup>ISEA, Université de la Nouvelle Calédonie,

May 20, 2024

## Abstract

Combinatorial generation of expander families and Lindenmayer-style development models are both parallel in nature. Both can be handled within proposed parallel graph grammar formalism. Their first-order properties can then be checked by encompassing the generated graph language into an appropriate automatic structure.

**Keywords:** parallel graph grammars, 1st order model checking, automatic structures, expanders.

## 1 Introduction

Deeply concerned about the ongoing destructive processes that threaten the mankind and the nature, we advocate for an increased effort toward applications of relevant theories beyond computer science and mathematics. Formal languages, graphs and automata cover several themes in line with the current emergency. Lindenmayer systems, also known as L-systems, and expanders show promising potential for such applications as generative mechanisms for modelling structures of interest in materials science. Their connections to logic in computer science include the following question: “Which properties of expander families can be algorithmically checked?”

### Our contribution

To answer the question above, we introduce 0L graph grammars<sup>1</sup> and demonstrate how to implement the most important constructs for the combinatorial generation of expander families within our formalism. To remain credible, we design it to capture grids, meshes, and similar multidimensional structures that are essential in physics, chemistry, and biology. This application-oriented requirement excludes the possibility of checking monadic second-order definable properties. We only consider here those definable at the first order, nevertheless retaining the quantitative logics for a more in-depth future investigation.

We identify three features of 0L graph grammars that shape the border of the decidable: edge-determinism, vertex-determinism and completeness. After defining the structure encompassing the graph language generated by a 0L graph grammar, we show that its word-automatic

---

<sup>1</sup>In the classification of L-systems “0” stands for “context-free” and “L” for Lindenmayer.

presentation can be constructed in quadratic time. Thanks to our another construction that makes complete any 0L graph grammar, the completeness is not a restriction at this stage.

Given a first-order sentence in graphs signature [7] and a 0L graph grammar, we ask if the corresponding 0L graph language includes a graph satisfying the sentence. By checking the associated automatic structure for an adequate translation of the input sentence, we have an answer in the case of deterministic 0L graph grammars. On the other side of the decidability border, we reduce the acceptance of the empty word by a deterministic Turing machine to the latter question. The reduction yields a 0L graph grammar that is edge-deterministic and possesses a unique derivation. With a few satellite undecidability results, we can conclude that the challenge lies in identifying individual graphs within the structure encompassing the entire 0L graph language.

## Significance

Introduced in 1968 by Aristid Lindenmayer [25] and later generalised as graph grammars [20], L-systems remain an active research area [11]. Their extensions have been widely used for modelling the growth of living organisms or plant structure [33, 27], often using dedicated programming languages [31]. However, their applications go beyond life sciences and affect areas as varied as image synthesis [35, 33], multi-agent simulation [28] or engineering [3].

Later named by Pinsker [32] but introduced by Kolmogorov and Barzdin as early as 1967, expanders have their origins in Kolmogorov's interest in the structure of axonal connections in the brain [2]. Applications of expanders span many areas such as communication networks, error-correcting codes, pseudorandomness, cryptography, computational complexity including the PCP theorem [8], derandomisation, computational group theory or number theory to name a few (see [19] and [26] for more details).

To our knowledge, parallel graph grammars and expanders<sup>2</sup> have been absent from the field of logic. This paper establishes their link with finite automata and finite [15, 24] as well as metafinite [14, 4] model theory.

## Related work

Parallel graph rewriting has a rich literature (see [10]) from pioneering work of [9] to more recent developments [37, 36] but all these approaches are context-sensitive. With an undecidable emptiness problem, an algorithmic checking of their properties, which is our main concern, seems very unlikely. Loosely related to that stream of work, our contribution can be seen as a simplification of formalisms of [20] and [21]. Graphs are considered there up to isomorphism which makes difficult referencing vertices or edges, requires long definitions and quite complicated proofs. With its first order logic-based core, our work extends both approaches. Additionally, it allows for multiple edges and loops. Several results, extending those of [20], are a by-product of the (un)decidability of model or language checking problems for 0L graph languages. Drawing our inspiration in the rich theory of monadic second order logic of graphs [7], we aim to stimulate an exploration of the connections between parallel graph languages and logics.

## Organisation of this paper

After some preliminaries, in Sect. 3 we review the basics of automatic structures. In Sect. 4 we introduce 0L graph grammars and state their typical language-theoretic decidability results. In

---

<sup>2</sup>except computation theory

Sect. 5 we review the most prominent combinatorial constructions of expanders and show how the corresponding deterministic or vertex-deterministic 0L graph grammars can be constructed. Logic related decidability results are stated in Sect. 6 together with the discussion about the central problem of this paper: the language checking problem. Most of our positive results follow from the construction of Sect. 7, namely the embedding of any complete 0L graph grammar into an automatic structure. In Sect. 8 we show how to simulate any deterministic Turing machine by an edge-deterministic 0L graph grammar. Short conclusion closes the main part of the paper and all missing proofs can be found in subsequent appendices.

## 2 Basic notations

Throughout the paper, placeholder “ $\_$ ” is often used to avoid specifying inessential parts in a notation like e.g. writing “let  $x \xrightarrow{a} \_$ ” in a more compact way than “let  $x \xrightarrow{a} y$  for some  $y$ ” while  $y$  does not matter. As usual  $[n]$  stands for  $\{1, 2, \dots, n\}$  with  $[0] = \emptyset$ . A tuple of length  $n \in \mathbb{N}$  (also called an  $n$ -tuple) on a set  $D$  is a map  $\bar{t} : [n] \rightarrow D$ . When  $\bar{t}(i) = e_i$ , for  $i \in [n]$ , a usual notation  $\bar{t} = (e_1, \dots, e_n)$  is often used. The 0-tuple is written  $()$ . Instead of  $D^{[n]}$ , the set of all  $n$ -tuples on  $D$  is simply written  $D^n$ . A tuple  $w$  of symbols from a finite *alphabet*  $\Gamma$  is called a *word* and is written  $w = a_1 a_2 \dots a_n$  where  $w(i) = a_i$ ,  $a_i \in \Gamma$  and  $w(\leq i) = a_1 a_2 \dots a_i$ , for  $i \in [n]$ . The length of  $w$ , written  $|w|$ , is  $n$  and the word of length 0, called the *empty word*, is written  $\varepsilon$  with  $\Gamma^0 = \{\varepsilon\}$ . The set of all words is written  $\Gamma^*$ , where  $\Gamma^* := \bigcup_{n \in \mathbb{N}} \Gamma^n$ .

Instead of being indexed by numbers in  $[n]$ , tuples may also be indexed by elements of a set  $\mathcal{X}$ . The family of subsets of  $\mathcal{X}$  of size  $k$  is written  $\wp_k(\mathcal{X})$ . Given  $\mathcal{X} \in \wp_n(\mathcal{X})$ , an  $\mathcal{X}$ -tuple  $\bar{t}$  of a set  $D$  is an element of  $D^{\mathcal{X}}$ . An  $\mathcal{X}$ -relation on  $D$  is a set of  $\mathcal{X}$ -tuples. Operations on such relations are similar to those of relational algebra [22]. Let  $\mathcal{X} \in \wp_n(\mathcal{X})$  and  $\mathcal{Y} \in \wp_m(\mathcal{X})$ . For a (possibly partial) function  $\sigma : \mathcal{X} \rightarrow \mathcal{Y}$ , define an *upturn (along  $\sigma$ )* of  $\bar{t} \in D^{\mathcal{Y}}$  (resp.  $\mathcal{R} \subseteq D^{\mathcal{Y}}$ ) as

$$\begin{aligned} \sigma^{-1}(\bar{t}) &:= \{\bar{s} \in D^{\mathcal{X}} : \forall x \in \text{dom}(\sigma) \bar{s}(x) = \bar{t}(\sigma(x))\}, \\ (\text{resp. } \sigma^{-1}(\mathcal{R}) &:= \bigcup_{\bar{t} \in \mathcal{R}} \sigma^{-1}(\bar{t})) . \end{aligned} \quad (2.1)$$

The upturn of  $\mathcal{R}$  along an inclusion  $\mathcal{X} \supseteq \mathcal{Y}$  is called a *cylindrification of  $\mathcal{R}$  up to  $\mathcal{X}$*  and is written  $\mathcal{X}(\mathcal{R})$ . The upturn along an inclusion  $\mathcal{X} \subseteq \mathcal{Y}$  (resp.  $\mathcal{Y} \setminus \mathcal{X} \subseteq \mathcal{Y}$ ) is a projection of  $\mathcal{R}$  on (resp. out of)  $\mathcal{X}$  and is written  $\pi_{\mathcal{X}}(\mathcal{R})$  (resp.  $\pi_{\exists \mathcal{X}}(\mathcal{R})$ ).

Let  $\square \notin \Gamma$  be an additional symbol called *blank* and set  $\Gamma_{\square} := \Gamma \cup \{\square\}$ . A *convolution*  $\otimes \bar{w}$  of an  $\mathcal{X}$ -tuple of words  $\bar{w} \in (\Gamma^*)^{\mathcal{X}}$  is a word  $u : [l] \rightarrow \Gamma_{\square}^{\mathcal{X}}$ , with  $l = \max_{x \in \mathcal{X}} |\bar{w}(x)|$ , on the alphabet  $\Gamma_{\square}^{\mathcal{X}}$  of  $\mathcal{X}$ -tuples of letters of  $\Gamma_{\square}$ , with  $\otimes \bar{w} := u$  defined by

$$u(j)(x) = \begin{cases} \bar{w}(x)(j), & \text{if } 1 \leq j \leq |\bar{w}(x)|, \\ \square, & \text{if } |\bar{w}(x)| < j \leq l. \end{cases}$$

A *hole* in word  $t \in (\Gamma_{\square}^{\mathcal{X}})^*$  is a position  $i$  such that  $t(i)(x) = \square$  for some component  $x \in \mathcal{X}$  and there exists  $j > i$  such that  $t(j)(x) \neq \square$ . A word is *hole-free* if it has no hole and a language on  $(\Gamma_{\square}^{\mathcal{X}})^*$  is *hole-free* if it has only hole-free words. Note that the result of a convolution is always *hole-free*. Operation  $\dot{\otimes}$  is extended to  $\mathcal{X}$ -relations:  $\otimes R := \{\otimes \bar{w} : \bar{w} \in R\}$ , for every  $R \subseteq (\Gamma^*)^{\mathcal{X}}$ . The inverse of the convolution  $\otimes^{-1}$  is defined for hole-free words over  $\Gamma_{\square}^{\mathcal{X}}$  in the usual way.

A (*finite*)  $\mathcal{X}$ -automaton [18, 17, 16] on  $\Gamma$  is a (finite word) automaton on  $\Gamma_{\square}^{\mathcal{X}}$ . An  $\mathcal{X}$ -automaton  $\mathcal{A}$  is *hole-free* if its language  $\mathcal{L}(\mathcal{A})$  is hole-free. As the set of hole-free words on  $\Gamma_{\square}^{\mathcal{X}}$  is a regular language, any  $\mathcal{X}$ -automaton may be restricted in such a way that it accepts only hole-free words. In the sequel, all  $\mathcal{X}$ -automata are assumed to be hole-free. Therefore, every (hole-free)  $\mathcal{X}$ -automaton  $\mathcal{A}$  induces the relation  $[\mathcal{A}] \subseteq (\Gamma^*)^{\mathcal{X}}$  defined by  $[\mathcal{A}] := \otimes^{-1}(\mathcal{L}(\mathcal{A}))$ . When  $\mathcal{X}$  is clear from the context or irrelevant, we speak of tuple automata or, more specifically, of pair

(when  $|\mathcal{X}| = 2$ ) or triple (when  $|\mathcal{X}| = 3$ ) automata. The arity  $\text{ar}(\mathcal{A})$  of an  $\mathcal{X}$ -automaton is the arity of the induced relation:  $\text{ar}(\mathcal{A}) := \text{ar}[\llbracket \mathcal{A} \rrbracket]$ . We assume that the reader is familiar with standard constructions on finite (word) automata notably Boolean operations and concatenation and we only give one construction specific to tuple automata. For  $\mathcal{X} \in \mathcal{P}_n(\mathcal{X})$  and  $\mathcal{Y} \in \mathcal{P}_m(\mathcal{X})$  let  $\sigma: \mathcal{X} \rightarrow \mathcal{Y}$  be a (possibly partial) function and let  $\mathcal{A}$  be an  $\mathcal{Y}$ -automaton,  $\mathcal{A} = (\mathbb{Q}, \Delta, \mathbb{1}, \mathbb{F})$ . An *upturn of  $\mathcal{A}$  along  $\sigma$*  is an  $\mathcal{X}$ -automaton  $\sigma^{-1}(\mathcal{A}) = (\mathbb{Q}, \sigma^{-1}(\Delta), \mathbb{1}, \mathbb{F}_{\square})$ , where  $\mathbb{F}_{\square}$  and  $\sigma^{-1}(\Delta)$  are defined by:

$$\begin{aligned} \mathbb{F}_{\square} &:= \{q \in \mathbb{Q} : \exists f \in \mathbb{F} \ q \overset{(\square^x)^*}{\dashrightarrow} f\}, \\ \sigma^{-1}(\Delta) &:= \{p \overset{\bar{a}}{\rightarrow} q : \exists \bar{b} \ (p \overset{\bar{b}}{\rightarrow} q \in \Delta \wedge \bar{a} \in \sigma^{-1}(\bar{b}))\} . \end{aligned}$$

Specific variants of the upturn for  $\mathcal{Y}$ -automata are written as for  $\mathcal{Y}$ -relations:  $\mathcal{X}(\mathcal{A})$  is a cylindrification up to  $\mathcal{X}$ ,  $\pi_{\mathcal{X}}(\mathcal{A})$  (resp.  $\pi_{\exists \mathcal{X}}(\mathcal{A})$ ) is a projection on (resp. out of)  $\mathcal{X}$ .

**Lemma 2.1.** *For every  $\mathcal{Y}$ -automaton and every (possibly partial) function  $\sigma: \mathcal{X} \rightarrow \mathcal{Y}$ , one has  $\llbracket \sigma^{-1}(\mathcal{A}) \rrbracket = \sigma^{-1}(\llbracket \mathcal{A} \rrbracket)$ .*

### 3 Logic and structures

A *relational signature* is a set  $\Pi$  of predicate symbols, each symbol  $\rho \in \Pi$  possessing its arity  $\text{ar}(\rho) \in \mathbb{N}$ . A  $\Pi$ -*relational structure* on a set  $D$  is a map  $\mathfrak{D}: \Pi \rightarrow \bigcup_{i \in \text{ar}(\Pi)} \mathcal{P}(D^i)$  such that every predicate symbol  $\rho \in \Pi$  is mapped into a relation  $\mathfrak{D}(\rho) \subseteq D^{\text{ar}(\rho)}$ .

*First-order formulae* on  $\Pi$  are built, from atomic formulae with predicate symbols of  $\Pi$ , using variables in  $\mathcal{X}$ , connectives  $\neg, \wedge, \vee, \supset$  and quantifiers  $\forall, \exists$ . Without loss of generality, each quantified variable is assumed occurring free in the subformula. The set they form is written  $\text{FO}(\Pi)$ . We allow grouping of any sequence of universal or existential quantifiers  $Q x_1 \dots, Q x_k$  with  $Q \in \{\forall, \exists\}$  into *quantifier block*  $Q\mathcal{X}$  where  $\mathcal{X} = \{x_1 \dots x_k\}$ . Every atomic formula is considered as being of the form  $\rho(\sigma)$  where  $\sigma: [\text{ar}(\rho)] \rightarrow \mathcal{X}$  and  $\mathcal{X} \in \mathcal{P}_m(\mathcal{X})$  with  $m \leq \text{ar}(\rho)$ . In other words, a tuple of variables of an atomic formula is seen as an explicit mapping. For instance, assuming  $x, y \in \mathcal{X}$ , the traditional writing  $r(y, x, y)$  where  $r \in \Pi$  is of arity 3 is seen as  $r(\sigma)$  where  $\sigma: [3] \rightarrow \{x, y\}$  with  $\sigma(1) = y, \sigma(2) = x$  and  $\sigma(3) = y$ . When  $\mathcal{X}$  is the set of free variables of a formula  $\varphi$ , the latter is often written  $\varphi(\mathcal{X})$ . As every  $\text{FO}(\Pi)$ -formula can be transformed into an equivalent one without  $\wedge, \supset$  nor  $\forall$ , we often give definitions or constructions for  $\text{FO}(\Pi)$ -formulae using no other connectives than  $\neg, \vee$  and no other quantifiers than  $\exists$ .

Every formula  $\varphi(\mathcal{X}) \in \text{FO}(\Pi)$  induces on a structure  $\mathfrak{D}: \Pi \rightarrow \bigcup_{i \in \text{ar}(\Pi)} \mathcal{P}(D^i)$  an  $\mathcal{X}$ -relation  $\mathfrak{D}(\varphi(\mathcal{X})) \subseteq D^{\mathcal{X}}$  defined inductively as follows

$$\begin{aligned} \mathfrak{D}(\rho(\sigma)) &:= \sigma^{-1} \circ \mathfrak{D}(\rho), \quad \text{where } \rho \in \Pi, \sigma: [\text{ar}(\rho)] \rightarrow \mathcal{X} \text{ and } \sigma^{-1} \text{ is given by (2.1),} \\ \mathfrak{D}(\neg \varphi(\mathcal{X})) &:= D^{\mathcal{X}} \setminus \mathfrak{D}(\varphi(\mathcal{X})) \\ \mathfrak{D}(\exists \mathcal{Z} \varphi(\mathcal{X})) &:= \pi_{\exists \mathcal{Z}}(\mathfrak{D}(\varphi(\mathcal{X}))) \\ \mathfrak{D}(\varphi(\mathcal{X}) \vee \psi(\mathcal{Y})) &:= (\mathcal{X} \cup \mathcal{Y})(\mathfrak{D}(\varphi(\mathcal{X}))) \cup (\mathcal{X} \cup \mathcal{Y})(\mathfrak{D}(\psi(\mathcal{Y}))) \end{aligned}$$

By an analogy between a structure  $\mathfrak{D}$  and a database, a formula  $\varphi(\mathcal{X})$  may be considered as a (*first-order*) *query* with corresponding relation  $\mathfrak{D}(\varphi(\mathcal{X}))$  induced on  $\mathfrak{D}$  being the result of the query.

#### Automatic structures

An *automatic presentation* of a relational structure  $\mathfrak{D}: \Pi \rightarrow \bigcup_{i \in \text{ar}(\Pi)} \mathcal{P}(D^i)$  is given by a tuple of automata on  $\Gamma$ ,  $\mathcal{A}(\mathfrak{D}) = (\mathcal{A}_{\delta}, \mathcal{A}_{\text{eq}}, (\mathcal{A}(\rho))_{\rho \in \Pi})$  where  $\mathcal{A}_{\delta}$  is a usual finite automaton,  $\mathcal{A}_{\text{eq}}$  is

a pair automaton, and,  $\text{ar}(\mathcal{A}(\varrho)) = \text{ar}(\varrho)$ , for each  $\varrho \in \Pi$ , together with a surjective mapping  $g: \mathcal{L}(\mathcal{A}_\delta) \rightarrow D$  such that

- $(u, v) \in \llbracket \mathcal{A}_{\text{eq}} \rrbracket$  if, and only if  $g(u) = g(v)$ , for all  $u, v \in \mathcal{L}(\mathcal{A}_\delta)$ ,
- $(u_1, \dots, u_{\text{ar}(\varrho)}) \in \llbracket \mathcal{A}(\varrho) \rrbracket$  if, and only if  $(g(u_1), \dots, g(u_{\text{ar}(\varrho)})) \in \mathfrak{D}(\varrho)$ , for all  $u_1, \dots, u_{\text{ar}(\varrho)} \in \mathcal{L}(\mathcal{A}_\delta)$ .

$\mathcal{A}(\mathfrak{D})$  is also said to be *an automatic presentation of  $\mathfrak{D}$  via  $g$* . Such a presentation is said *injective* when  $g$  is so. Then  $\mathcal{A}_{\text{eq}} = \mathcal{A}(=)$  is an automaton realising the identity on  $\mathcal{L}(\mathcal{A}_\delta)$ . A structure possessing an automatic presentation is called an *automatic structure*. Using an appropriate intersection of automata, every  $\mathcal{A}(\varrho)$  can be arranged so that  $\llbracket \mathcal{A}(\varrho) \rrbracket \subseteq \llbracket \mathcal{A}_\delta \rrbracket^{\text{ar}(\varrho)}$ . This is assumed for the sequel. Automatic structures have been introduced in 1976 by Hodgson [18, 17] and independently by Koussainov and Nerode in 1994 [23] (see also survey [13]).

## Construction of query automaton

Perhaps the most prominent property of automatic structures is that the result of every first-order query is encoded by a regular language. In other words, every relation that is FO-definable on an automatic structure may be encoded as a relation induced by a tuple automaton.

Let  $\mathcal{A}(\mathfrak{D}) = (\mathcal{A}_\delta, \mathcal{A}_{\text{eq}}, (\mathcal{A}(\varrho))_{\varrho \in \Pi})$  be an automatic presentation of a relational structure  $\mathfrak{D}: \Pi \rightarrow \bigcup_{i \in \text{ar}(\Pi)} \mathcal{P}(D^i)$ . To every formula  $\varphi(\mathcal{X}) \in \text{FO}(\Pi)$ , an automaton  $\mathcal{H}(\varphi(\mathcal{X}))$ , called a *query automaton of  $\varphi(\mathcal{X})$* , is associated inductively as follows

$$\begin{aligned} \mathcal{H}(\varrho(\sigma)) &:= \sigma^{-1}(\mathcal{A}(\varrho)) && \text{where } \varrho \in \Pi \text{ and } \sigma: [\text{ar}(\varrho)] \rightarrow \mathcal{X} \\ \mathcal{H}(\neg\varphi(\mathcal{X})) &:= \mathcal{A}_\delta^{\mathcal{X}} \setminus \mathcal{H}(\varphi(\mathcal{X})), \\ \mathcal{H}(\varphi(\mathcal{X}) \vee \psi(\mathcal{Y})) &:= (\mathcal{X} \cup \mathcal{Y})(\mathcal{H}(\varphi(\mathcal{X}))) \cup (\mathcal{X} \cup \mathcal{Y})(\mathcal{H}(\psi(\mathcal{Y}))), \\ \mathcal{H}(\exists \mathcal{Z} \varphi(\mathcal{X})) &:= \pi_{\exists \mathcal{Z}}(\mathcal{H}(\varphi(\mathcal{X}))) \end{aligned}$$

where  $\mathcal{A}_\delta^{\mathcal{X}}$  is an  $\mathcal{X}$ -automaton such that  $\llbracket \mathcal{A}_\delta^{\mathcal{X}} \rrbracket = \llbracket \mathcal{A}_\delta \rrbracket^{\mathcal{X}}$ , viz.,  $\mathcal{X}$ -indexed free product of  $|\mathcal{X}|$  copies of  $\mathcal{A}_\delta$ . The fundamental property of  $\mathcal{H}$  is that  $g(\llbracket \mathcal{H}(\varphi(\mathcal{X})) \rrbracket) = \mathfrak{D}(\varphi(\mathcal{X}))$ , for every  $\varphi(\mathcal{X}) \in \text{FO}(\Pi)$ . In particular, for a sentence  $\psi \in \text{FO}(\Pi)$ , one has  $\mathfrak{D} \models \psi \Leftrightarrow \mathcal{L}(\mathcal{H}(\psi)) \neq \emptyset$ .<sup>3</sup> This leads to the well-known result that the FO-model checking (resp. query evaluation) for automatic structures is decidable (resp. computable), even in the case when  $\text{FO}(\Pi)$  is extended with counting quantifiers saying that “there are  $k \bmod m$  many elements” or “there are infinitely many elements”. Indeed, the above inductive construction of the query automaton can be extended with two corresponding cases [4]. The results of our paper remain valid in that extension.

## Remarks on complexity

The construction of query automaton, also known as query evaluation problem, has a reputation of non-elementary complexity [13]. However, for quantifier-free formula, this construction can be done in time  $\mathcal{O}(m^n)$  where  $m$  is the number of transitions of the largest automaton of  $\mathcal{A}(\mathfrak{D})$  and  $n$  is the number of atoms and their negations in the input formula [4]. In the general case, input formula can be put in prenex form in linear time. Then the time complexity is an exponential tower of height  $k$  with  $m^n$  on top of it, where  $k$  is the number of quantifier alternations. Note that this worst-case complexity should not be considered prohibitive for real life applications. Properties that need four quantifier alternations or more are very uncommon. Also uncommon is the exponential blow-up of the determinisation. As is observed in [12], it occurs only in cases similar to the usual example.

<sup>3</sup>Note that, for a sentence  $\psi \in \text{FO}(\Pi)$ ,  $\mathcal{H}(\psi)$  is an automaton over single letter alphabet  $\{()\}$ .

## 4 Lindenmayer graph languages

The grammatical formalism developed in this section subsumes those of graph grammars discussed in [21]. A 0L graph grammar introduced in the sequel generates (concrete) graphs as incidence structures of the form  $(V, E, \ell, \mathfrak{I})$  where

- $V \subseteq \Theta^+$  are the vertices given as words over an alphabet  $\Theta$  of *vertex names*,
- $E \subseteq \Theta^* \Omega^+$  are the edges given as words over an alphabet  $\Omega$  of *edge names*,  $\Theta \cap \Omega = \emptyset$ , possibly with a prefix in  $\Theta^*$ ,
- $\ell: V \rightarrow \Gamma$  is the labelling of vertices and
- $\mathfrak{I} \subseteq \Sigma \times V \times E \times V$  is the incidence relation given as a set of *labelled incidences* of the form  $a(u, w, v)$  with  $u, v \in V$ ,  $w \in E$  and  $a \in \Sigma$ , such that  $\{a(u, w, v) \mapsto w : a(u, w, v) \in \mathfrak{I}\}$  is a bijection between  $\mathfrak{I}$  and  $E$ .

The set of (concrete) graphs over vertex and edge alphabets  $\Theta$  and  $\Omega$  with  $\Gamma$ -labelled vertices and  $\Sigma$ -labelled edges is written  $\mathcal{G}(\Gamma, \Sigma, \Theta, \Omega)$ . The set of those graphs of  $\mathcal{G}(\Gamma, \Sigma, \Theta, \Omega)$  that have their edges and vertices all of length  $l \in \mathbb{N}$  is written  $\mathcal{G}_l(\Gamma, \Sigma, \Theta, \Omega)$ . This formalism is general enough to handle unlabelled or undirected graphs too. Undirected incidences  $a(\{u, v\}, w)$  (resp. with labels at both extremities  $(\{(a, u), (b, v)\}, w)$ ) often drawn as  $u \overset{c}{-} v$  (resp.  $u \overset{a}{-} \overset{b}{-} v$ ) are represented by two opposite directed incidences (resp. with a pair of labels):  $a(u, w, v)$  and  $a(v, w, u)$  (resp.  $(a, b)(u, w, v)$  and  $(b, a)(v, w, u)$ ).

A *0L graph grammar* over a (finite) set  $\Gamma$  of vertex labels, a (finite) set  $\Sigma$  of edge labels, a (finite) set  $\Theta$  of vertex names and a (finite) set  $\Omega$  of edge names is a triple  $\mathcal{G} = (\mathbf{V}_{\mathcal{G}}, \mathbf{E}_{\mathcal{G}}, \mathbf{A}_{\mathcal{G}})$  where

- $\mathbf{A}_{\mathcal{G}}$  is the axiom of the grammar representing a graph with no edge and single vertex encoded by the empty word labelled  $\mathbf{A}_{\mathcal{G}}$  which can but is not necessarily in  $\Gamma$ ,
- $\mathbf{V}_{\mathcal{G}}$  is a finite set of *vertex productions* of the form  $A \rightarrow G$ , with  $A \in \Gamma \cup \{\mathbf{A}_{\mathcal{G}}\}$  and  $G \subseteq \mathcal{G}_1(\Gamma, \Sigma, \Theta, \Omega)$ , such that the right hand sides of  $\mathbf{V}_{\mathcal{G}}$  are pairwise disjoint, respectively to their vertex and edge sets,<sup>4</sup>
- $\mathbf{E}_{\mathcal{G}}$  is a finite set of *edge productions* of the form  $(A_1 \rightarrow G_1) \xrightarrow{a} (A_2 \rightarrow G_2) \rightarrow \sqsupset$ , with  $\sqsupset$  using two disjoint sets of edge labels  $\overrightarrow{\Sigma}$  and  $\overleftarrow{\Sigma}$ , each one in bijection with  $\Sigma$ , where

$$\begin{aligned} a &\in \Sigma, \\ A_1 \rightarrow G_1, A_2 \rightarrow G_2 &\in \mathbf{V}_{\mathcal{G}}, \text{ with } G_1 = (V_1, \_, \_, \_) \text{ and } G_2 = (V_2, \_, \_, \_), \\ \text{and } \sqsupset &\subseteq (\Sigma \times V_1 \times \Omega \times V_2) \cup (\Sigma \times V_2 \times \Omega \times V_1) \text{ are production's incidences} \end{aligned}$$

such that every edge name of  $\Omega$  appears at most once in the right hand sides of grammar's productions  $\mathbf{V}_{\mathcal{G}} \cup \mathbf{E}_{\mathcal{G}}$ .<sup>4</sup>The set of such grammars is written  $0L(\Gamma, \Sigma, \Theta, \Omega)$ . The edge label from  $\overrightarrow{\Sigma}$  (resp.  $\overleftarrow{\Sigma}$ ) corresponding to  $a \in \Sigma$  is written  $\overrightarrow{a}$  (resp.  $\overleftarrow{a}$ ). For every production  $P \in \mathbf{V}_{\mathcal{G}} \cup \mathbf{E}_{\mathcal{G}}$  let  $\text{lhs}(P)$  (resp.  $\text{rhs}(P)$ ) stands for its left (resp. right) hand side, viz., for vertex production  $P_1 = A \rightarrow G$ ,  $\text{lhs}(P_1) = A$  (resp.  $\text{rhs}(P_1) = G$ ) and for edge production  $P = P_1 \xrightarrow{a} P_2 \rightarrow \sqsupset$ ,  $\text{lhs}(P) = P_1 \xrightarrow{a} P_2$  (resp.  $\text{rhs}(P) = \sqsupset$ ). As every symbol  $i \in \Theta \cup \Omega$  appears at most once in  $\text{rhs}(\mathbf{V}_{\mathcal{G}} \cup \mathbf{E}_{\mathcal{G}})$ , we denote by  $\text{prd}(i)$  the production  $P \in \mathbf{V}_{\mathcal{G}} \cup \mathbf{E}_{\mathcal{G}}$  where  $i$  appears in  $\text{rhs}(P)$ . When  $\mathbf{A}_{\mathcal{G}} \notin \Gamma$ , a vertex production with the axiom on its left hand side is called *an axiom production*. A vertex (resp. edge) production of the form  $A \rightarrow (\emptyset, \_, \_, \_)$  (resp.  $P \xrightarrow{a} P' \rightarrow \emptyset$ ) is called *erasing*. A grammar with no erasing productions is called *non-erasing*.

Every 0L graph grammar  $\mathcal{G} = (\mathbf{V}_{\mathcal{G}}, \mathbf{E}_{\mathcal{G}}, \mathbf{A}_{\mathcal{G}})$  can be seen as a graph with vertices  $\mathbf{V}_{\mathcal{G}}$  labelled in  $\Gamma$  and edges  $\text{lhs}(\mathbf{E}_{\mathcal{G}})$  labelled in  $\Sigma$  together with an additional information in  $\text{rhs}(\mathbf{E}_{\mathcal{G}})$ . Every vertex production  $A \rightarrow H$  is an  $A$ -labelled vertex of graph  $\mathcal{G}$ . Every edge production  $(A \rightarrow H) \xrightarrow{a} (A' \rightarrow H') \rightarrow \sqsupset$  defines an  $a$ -labelled edge  $(A \rightarrow H) \xrightarrow{a} (A' \rightarrow H')$  of graph  $\mathcal{G}$ . With this view of  $\mathcal{G}$ , from a (concrete) graph  $G = (V, E, \ell, \mathfrak{I}) \in \mathcal{G}_l(\Gamma, \Sigma, \Theta, \Omega)$  with vertices and edges of length  $l$ , one

<sup>4</sup>This requirement is not necessarily respected in examples in order to keep them simple.

*derives* in a single step a (concrete) graph  $G' = (V', E', \ell', \mathfrak{I}') \in \mathcal{G}_{l+1}(\Gamma, \Sigma, \Theta, \Omega)$  upon selecting a graph homomorphism  $h: G \rightarrow \mathcal{G}$  satisfying  $h(v) = \ell(v) \rightarrow \_$  for every vertex  $v \in V$  and, for every edge  $w \in E$  such that  $a(u, w, v) \in \mathfrak{I}$ , satisfying  $h(w) = h(u) \xrightarrow{a} h(v) \rightarrow \_$ . Resulting graph  $G'$  is said to be *produced by  $h$  from  $G$* , written  $G \xrightarrow[\mathcal{G}]{h} G'$ , or *derived (in a single step) from  $G$* , written  $G \xrightarrow[\mathcal{G}]{} G'$  and is obtained as follows:

$$\begin{aligned} \ell' &:= \bigcup_{v \in V} \{vi \mapsto \ell_v(i) : h(v) = \_ \rightarrow (V_v, \_, \ell_v, \_) \wedge i \in V_v\}, \\ V' &:= \bigcup_{A \in \Gamma} (\ell')^{-1}(A), \\ \mathfrak{I}' &:= \{b(ui, wj, vk) : (h(w) = h(u) \xrightarrow{a} h(v) \rightarrow \sqsupset \wedge \vec{b}(i, j, k) \in \sqsupset)\} \cup \\ &\quad \{b(vk, wj, ui) : (h(w) = h(u) \xrightarrow{a} h(v) \rightarrow \sqsupset \wedge \overleftarrow{b}(i, j, k) \in \sqsupset)\} \cup \\ &\quad \{b(ui, uj, uk) : h(u) = \_ \rightarrow (\_, \_, \_, \mathfrak{I}_u) \wedge b(i, j, k) \in \mathfrak{I}_u\}, \\ E' &:= \{w' : b(\_, w', \_) \in \mathfrak{I}'\}. \end{aligned}$$

In the above definition of  $\mathfrak{I}'$ , incidences and corresponding edges in the 1st or 2nd (resp. in the 3rd) member of the union are called *inherited* (resp. *innate*). Every  $ui \in V' \cup E'$  *derives (in a single step) from  $u \in V \cup E$  via  $P$* , written  $v \xrightarrow[\mathcal{G}]{P} vi$ , where  $P \in \mathbf{V}_{\mathcal{G}} \cup \mathbf{E}_{\mathcal{G}}$  is a production such that  $h(u) = P$ . If  $u_0 \xrightarrow[\mathcal{G}]{P_1} \dots \xrightarrow[\mathcal{G}]{P_n} u_n$ , we write  $u_0 \xrightarrow[\mathcal{G}]{P_1 \dots P_n} u_n$  and say that  $u_n$  *derives from  $u_0$  via (sequence of productions)  $P_1 \dots P_n$* . The transitive closure of the single-step derivation relation over graphs is the *derivation relation*, written “ $\xrightarrow[\mathcal{G}]{}^+$ ”, and the *language of  $\mathcal{G}$* , viz., the *set of graphs derived or generated by  $\mathcal{G}$*  is defined as usual:  $\mathcal{L}(\mathcal{G}) := \{G : A_{\mathcal{G}} \xrightarrow[\mathcal{G}]{}^+ G\}$ .

A slight generalisation of  $\mathcal{L}(\mathcal{G})$ , written  $\mathcal{L}(\mathcal{G})_N$ , arises when one distinguishes a subset of *non-terminal* labels  $N \subseteq \Gamma \cup \Sigma$  not to appear in the final result:  $\mathcal{L}(\mathcal{G})_N := \mathcal{L}(\mathcal{G}) \cap \mathcal{G}(\Gamma \setminus N, \Sigma \setminus N, \Theta, \Omega)$ . A label in  $\Gamma \cup \Sigma$  (resp. production of  $\mathcal{G}$ ) is *accessible* in (resp. is *useful* for)  $\mathcal{G}$ , if it appears in (resp. is used in a derivation of) a graph of  $\mathcal{L}(\mathcal{G})$ . A graph grammar is *trimmed* if it has only useful productions. A 0L graph grammar is *edge-deterministic* (resp. *vertex-deterministic*) and called a *eD0L* (resp. *vD0L*) *graph grammar* if, for every triple  $(A_1, a, A_2) \in \Gamma \times \Sigma \times \Gamma$  (resp. vertex label  $A \in \Gamma$ ), it has at most one edge  $(A_1 \rightarrow \_) \xrightarrow{a} (A_2 \rightarrow \_) \rightarrow \_ \in \mathbf{E}_{\mathcal{G}}$  (resp. one vertex  $A \rightarrow \_ \in \mathbf{V}_{\mathcal{G}}$ ). A 0L graph grammar is *deterministic* and is called *D0L graph grammar* if it is both vD0L and eD0L. The subset of 0L( $\Gamma, \Sigma, \Theta, \Omega$ ) that consists of D0L graph grammars is denoted by  $\text{D0L}(\Gamma, \Sigma, \Theta, \Omega)$ . A 0L graph grammar is *unix* if for every  $l \in \mathbb{N}$  it has at most one derivation of length  $l$ . Obviously, every D0L graph grammar is unix.

We say that  $\mathcal{G}$  is *complete*, if it satisfies the following two conditions:

- (1) for every edge label  $a \in \Sigma$  and all vertices  $A_1 \rightarrow H_1, A_2 \rightarrow H_2 \in \mathbf{V}_{\mathcal{G}}$  such that  $A_{\mathcal{G}} \notin \{A_1, A_2\}$  and an incidence  $a(u, \_, v)$  with  $u$  labelled  $A_1$  and  $v$  labelled  $A_2$  appear a right hand side of a production, there exists an edge  $(A_1 \rightarrow H_1) \xrightarrow{a} (A_2 \rightarrow H_2) \rightarrow \_ \in \mathbf{E}_{\mathcal{G}}$ ,
- (2) for every isolated vertex with label  $A \in \Gamma$  appearing in a right hand side of a production, there exists a vertex  $A \rightarrow \_ \in \mathbf{V}_{\mathcal{G}}$ .

Derivations according to a complete 0L graph grammar are “easy”, in the sense that every choice of vertex productions to apply on some graph  $G$ , for deriving in a single step a “child”, can be completed with a compatible choice of edge productions. This is formalised in the following lemma.

**Lemma 4.1.** *Let  $\mathcal{G} = (A_{\mathcal{G}}, \mathbf{V}_{\mathcal{G}}, \mathbf{E}_{\mathcal{G}})$  be a complete graph grammar in  $\text{0L}(\Gamma, \Sigma, \_, \_)$  and let  $G = (V, E, \ell, \mathfrak{I})$  be a graph in  $\mathcal{L}(\mathcal{G})$ . Every labelling preserving map  $h: V \rightarrow \mathbf{V}_{\mathcal{G}}$  extends into a homomorphism from  $G$  into  $\mathcal{G}$ .*

The following proposition asserts that the completeness can be bypassed with the help of one non-terminal vertex label and one non-terminal edge label.



**Proposition 4.2.** *For every 0L graph grammar  $\mathcal{G} = (\mathbb{V}_{\mathcal{G}}, \_, \_)$  one can construct in time  $\mathcal{O}(|\mathbb{V}_{\mathcal{G}}|^2)$  using additional non-terminal vertex label “ $\mathfrak{i}$ ” and non-terminal edge label “ $\mathfrak{!}$ ” complete 0L graph grammar  $\mathcal{G}^{\mathfrak{i}!}$  such that  $\mathcal{L}(\mathcal{G}) = \mathcal{L}(\mathcal{G}^{\mathfrak{i}!})_{\{\mathfrak{i}, \mathfrak{!}\}}$ . If  $\mathcal{G}$  is vD0L or eD0L, then  $\mathcal{G}^{\mathfrak{i}!}$  is so.*

Note that, in general, the construction does not preserve the property of being unix, even for unix eD0L graph grammars.

The following graphical formalism is used in subsequent examples. Every edge production  $(A_1 \rightarrow H_1) \xrightarrow{a} (A_2 \rightarrow H_2) \rightarrow \sqsupset$  in  $\mathbb{E}_{\mathcal{G}}$  is drawn separately using colours as an (abstract) edge  $A_1 \xrightarrow{a} A_2$  on the left hand side and  $H_1 \cup H_2$  augmented with  $\sqsupset$  on the right hand side. Vertex productions  $A_1 \rightarrow H_1$  and  $A_2 \rightarrow H_2$  are implicit form colour codes and are not listed unless they cannot be inferred, notably when the axiom uses an extra label not in  $\Gamma$ . Incidences  $\sqsupset$  are drawn as black edges. In addition, vertex or edge names from  $\Theta \cup \Omega$  can be mentioned in parentheses. During a rewriting step of a graphical representation of a graph  $(V, E, \ell, \mathfrak{l})$ , every incidence in  $\mathfrak{l}$  together with corresponding vertex labels is matched to a left hand side of a graphical representation of an edge production so that the matching agrees on common vertices. In other words, for each pair of incident edges drawn as e.g.  $A \xrightarrow{a} B \xrightarrow{b} C$ , the vertex production used at  $B$  must be common to both edge productions applied on  $A \xrightarrow{a} B$  and  $B \xrightarrow{b} C$  (and similarly for other orientations of edges like  $A \xleftarrow{a} B \xrightarrow{b} C$  or  $A \xrightarrow{a} B \xleftarrow{b} C$ ). This is a graphical counterpart of the homomorphism condition in the definition of derivation.

Note that arrows of  $\overrightarrow{\Sigma}$  and  $\overleftarrow{\Sigma}$  are not needed in a graphical representation. Moreover, even a formal statement of  $\mathcal{G}$  can be simplified by removing arrows of  $\overrightarrow{\Sigma}$  and  $\overleftarrow{\Sigma}$  in every edge production, the left hand side of which is not a loop. If however it is a loop, like in e.g.  $(A \rightarrow H) \xrightarrow{a} (A \rightarrow H) \rightarrow \sqsupset$ , when homomorphism  $h: G \rightarrow G'$  producing  $G'$  from  $G$  maps to it an edge between two distinct vertices of  $G$ , then  $\overrightarrow{\Sigma}$  and  $\overleftarrow{\Sigma}$  are necessary to avoid a confusion about the direction of edges created by  $\sqsupset$  in  $G'$  between the two copies of  $H$ . Now, in the corresponding graphical representation of  $\mathcal{G}$  we draw an edge  $A \xrightarrow{a} A$  at the left hand side and there is no ambiguity about the direction of edges added on the right hand side between the two copies  $H$  and  $H$ . Nevertheless, in such a case, one needs to keep in mind that, in a graphical representation of the above edge production,  $A \xrightarrow{a} A$  corresponds to a loop of  $\mathcal{G}$  which can be matched by loop of  $G$  over one of its vertices, say  $v$ . Then, in produced  $G'$ , there is single copy of  $H$  corresponding to  $v$  (instead of two copies of  $H$ , in case the matched edge of  $G$  is not a loop) and edges created by  $\sqsupset$  for that copy are loops.

Another difference between a formal statements and a graphical formalism concerns the meaning of “right hand side”. For edge production  $P = (A_1 \rightarrow H_1) \xrightarrow{a} (A_2 \rightarrow H_2) \rightarrow \sqsupset$ ,  $\text{rhs}(P)$  appears only as a portion of the right hand side of the graphical representation of  $P$ . Indeed, the latter includes drawings of  $H_1$  and  $H_2$  whereas the former is  $\text{rhs}(P) = \sqsupset$ . The following example illustrates the differences between the graphical representation and the formal statement of a 0L graph grammar.

### Example 4.3

Consider  $\Gamma := \{A\}$ ,  $\Sigma := \{a\}$ ,  $\Theta := \{0, 1, 2, 3, 4\}$ ,  $\Omega := \{\alpha, \beta, \gamma, \delta\}$  and  $(\text{Ax}, \mathbb{V}, \mathbb{E}) \in \text{0L}(\Gamma, \Sigma, \Theta, \Omega)$  where  $\mathbb{V} := \{\text{Ax} \rightarrow G_0, A \rightarrow G_1, A \rightarrow G_2\}$ ,

$$\mathbb{E} := \{(A \rightarrow G_1) \xrightarrow{a} (A \rightarrow G_1) \rightarrow \sqsupset_1, (A \rightarrow G_2) \xrightarrow{a} (A \rightarrow G_2) \rightarrow \sqsupset_2\}$$

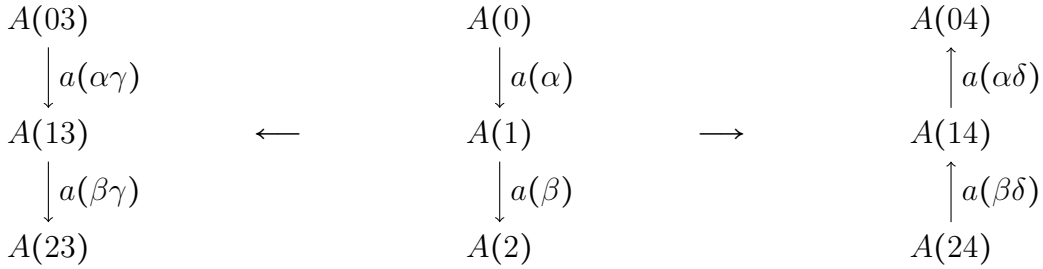
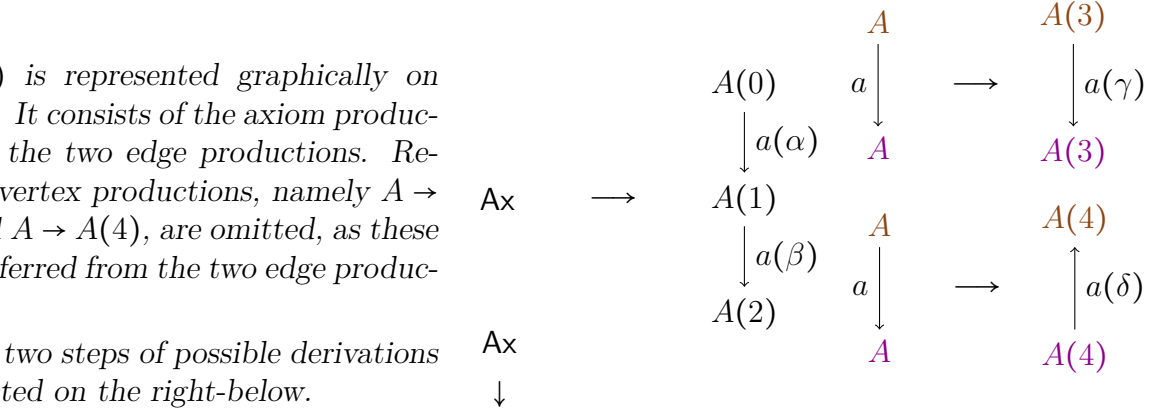
with  $G_0 := (\{0, 1, 2\}, \{\alpha, \beta\}, \{0 \mapsto A, 1 \mapsto A, 2 \mapsto A\}, \{a(0, \alpha, 1), a(1, \beta, 2)\})$ ,

$$G_1 := (\{3\}, \emptyset, \{3 \mapsto A\}, \emptyset), \quad \sqsupset_1 := \{\overrightarrow{a}(3, \gamma, 3)\},$$

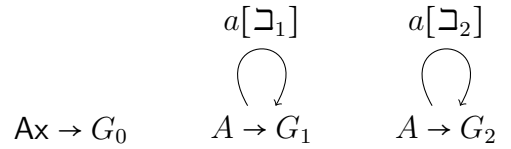
$$G_2 := (\{4\}, \emptyset, \{4 \mapsto A\}, \emptyset), \quad \sqsupset_2 := \{\overleftarrow{a}(4, \delta, 4)\}.$$

$(Ax, V, E)$  is represented graphically on the right. It consists of the axiom production and the two edge productions. Remaining vertex productions, namely  $A \rightarrow A(3)$  and  $A \rightarrow A(4)$ , are omitted, as these can be inferred from the two edge productions.

The first two steps of possible derivations are depicted on the right-below.



Seen as a graph, this 0L graph grammar consists of three isolated vertices forming  $V$ , two of which are equipped with a loop, as depicted on the right. The additional information about  $E$  mentioned in the definition of 0L graph grammar is put in brackets.



This grammar is not complete. It lacks edges between the two above  $A$ -labelled vertices (in each direction). It is neither vD0L nor eD0L.

**Example 4.4**

By extending Example 4.3 with an additional edge production  $(A \rightarrow G_1) \xrightarrow{a} (A \rightarrow G_2) \rightarrow \sqsupset_3$  where  $\sqsupset_3 := \{\vec{a}(3, \zeta, 4)\}$  which can be depicted as  $A \xrightarrow{a} A \longrightarrow A(3) \xrightarrow{a(\zeta)} A(4)$ , two more graphs derive from  $G_0$  in a single step:

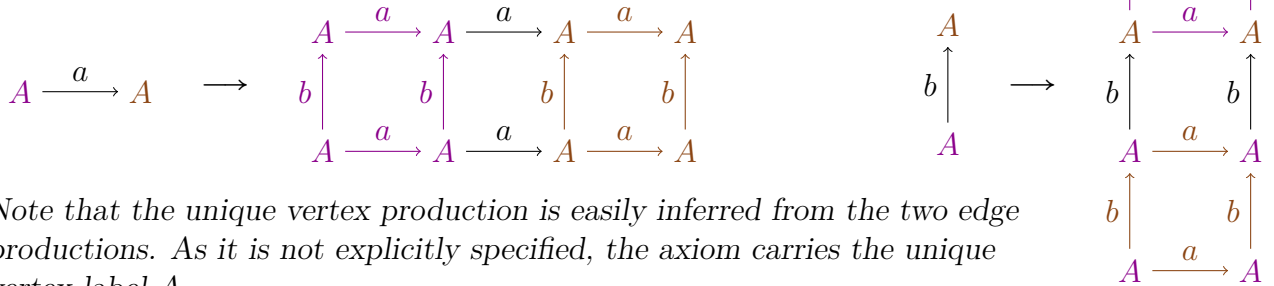
$$A(03) \xrightarrow{a(\alpha\zeta)} A(14) \xleftarrow{a(\beta\delta)} A(24)$$

$$A(03) \xrightarrow{a(\alpha\gamma)} A(13) \xrightarrow{a(\beta\zeta)} A(24) .$$

Observe that these graph share their vertices with graphs derived in Example 4.3. Also, each new graph shares one edge with a graph from the latter example.

**Example 4.5**

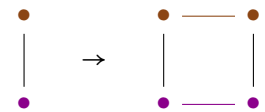
D0L grammar generating a language of  $2^n \times 2^n$  grids, for  $n \geq 1$ .

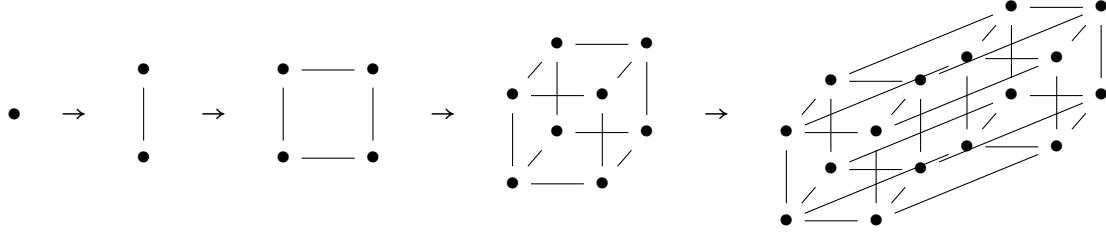


Note that the unique vertex production is easily inferred from the two edge productions. As it is not explicitly specified, the axiom carries the unique vertex label  $A$ .

**Example 4.6**

D0L grammar for undirected unlabelled graphs (formally  $|\Gamma| = |\Sigma| = 1$ ) generating all hypercubes via its unique derivation depicted below.





Observe that by definition, a 0L graph grammar cannot be ambiguous. Indeed, every graph in  $\mathcal{L}(\mathcal{G})$  has a unique derivation. If, two distinct derivations lead to isomorphic graphs, say  $(V_1, E_1, \_, \_)$  and  $(V_2, E_2, \_, \_)$ , then  $V_1 \cup E_1 \neq V_2 \cup E_2$  (see Example 4.3). However, two distinct graphs need not to be disjoint as pointed out in Example 4.4.

As  $\mathcal{L}(\mathcal{G})$  is in one-to-one correspondence with the set of derivations of  $\mathcal{G}$ , one can stratify  $\mathcal{L}(\mathcal{G})$  according to derivation length. For  $l \in \mathbb{N}$ , an  $l$ -layer of  $\mathcal{L}(\mathcal{G})$  is the set, written  $\mathcal{L}_l(\mathcal{G})$ , of graphs having derivations of length  $l$  according to  $\mathcal{G}$ . With that respect, the following two lemmas are as obvious as important.

**Lemma 4.7.**  $\mathcal{L}_l(\mathcal{G}) \subseteq \mathcal{G}_l(\Gamma, \Sigma, \Theta, \Omega)$ , for every  $l \in \mathbb{N}$  and every  $\mathcal{G} \in \text{0L}(\Gamma, \Sigma, \Theta, \Omega)$ .

**Lemma 4.8.** For every unix 0L graph grammar  $\mathcal{G}$  and every  $l \in \mathbb{N}$

- (1)  $|\mathcal{L}_l(\mathcal{G})| \leq 1$  and
- (2) if  $|\mathcal{L}(\mathcal{G})| = \omega$ , then  $|\mathcal{L}_l(\mathcal{G})| = 1$ .

We now turn to a few typical language-theoretic decision problems.

**Theorem 4.9.** The following problems

- Instance: 0L graph grammar  $\mathcal{G}$ , label  $\lambda$  of  $\mathcal{G}$  and production  $P$  of  $\mathcal{G}$ .  
Finiteness: Is  $\mathcal{L}(\mathcal{G})$  finite?  
Emptiness:  $\mathcal{L}(\mathcal{G}) = \emptyset$ ?  
Accessibility: Is  $\lambda$  accessible in  $\mathcal{G}$ ?  
Usefulness: Is  $P$  useful for  $\mathcal{G}$ ?

are

- (i) decidable for complete 0L graph grammars without non-terminals and for D0L graph grammars even with non-terminals,
- (ii) undecidable for unix eD0L graph grammars, with non-terminals in the case of emptiness.

For every 0L graph grammar  $\mathcal{G}$ , by enumerating  $\mathcal{L}_l(\mathcal{G})$ , one can decide if a concrete graph  $G \in \mathcal{G}_l(\Gamma, \Sigma, \Theta, \Omega)$  is in  $\mathcal{L}(\mathcal{G})$ . The membership up to isomorphism considered in the next statement is a bit more challenging.

**Proposition 4.10.** The (abstract) membership problem is decidable for D0L graph grammars even with non-terminals. It is undecidable for unix eD0L graph grammars.

We let  $\mathcal{V}(\mathcal{G}) := \bigcup_{(V, \_, \_, \_) \in \mathcal{L}(\mathcal{G})} V$  (resp.  $\mathcal{E}(\mathcal{G}) := \bigcup_{(\_, E, \_, \_) \in \mathcal{L}(\mathcal{G})} E$ ,  $\mathcal{J}(\mathcal{G}) := \bigcup_{(\_, \_, \_, \_) \in \mathcal{L}(\mathcal{G})} \mathcal{J}$ ) denote the sets of all vertices (resp. edges, incidences) generated by  $\mathcal{G}$ . We write  $\mathcal{V}(\mathcal{G})_N$ ,  $\mathcal{E}(\mathcal{G})_N$  and  $\mathcal{J}(\mathcal{G})_N$ , when generating is restricted to  $\mathcal{L}(\mathcal{G})_N$  for some set  $N$  of non-terminals. The sets so defined are stratified, with layers  $\mathcal{V}_l(\mathcal{G}) := \{v \in \mathcal{V}(\mathcal{G}) : |v| = l\}$ ,  $\mathcal{E}_l(\mathcal{G}) := \{e \in \mathcal{E}(\mathcal{G}) : |e| = l\}$  and  $\mathcal{J}_l(\mathcal{G}) := \{a(u, w, v) \in \mathcal{J}(\mathcal{G}) : |u| = |v| = |w| = l\}$  for  $l \in \mathbb{N}$ . Similarly for  $\mathcal{V}_l(\mathcal{G})_N$ ,  $\mathcal{E}_l(\mathcal{G})_N$  and  $\mathcal{J}_l(\mathcal{G})_N$ .

The *incompatibility relation* (of  $\mathcal{G}$ ), written “ $\chi$ ”, is defined on  $(\Theta \cup \Omega)^*$  as

$$\chi := \chi_0 \cup \chi_1 \cup \chi_1^{-1} \cup \chi_2 \cup \chi_2^{-1}$$

where

$$\begin{aligned}
\chi_0 &:= \{(uis, ujt) \in (\mathcal{V}_l(\mathcal{G}) \cup \mathcal{E}_l(\mathcal{G}))^2 : l \in \mathbb{N}_+ \wedge \text{lhs}(\text{prd}(i)) = \text{lhs}(\text{prd}(j)) \wedge \text{prd}(i) \neq \text{prd}(j)\} \\
\chi_1 &:= \{(us, wt) \in (\mathcal{V}_l(\mathcal{G}) \cup \mathcal{E}_l(\mathcal{G}))^2 : l \in \mathbb{N}_+ \wedge \exists a \in \Sigma \exists u_1, u_2 \in \mathcal{V}(\mathcal{G}) (a(u_1, w, u_2) \in \mathcal{J}(\mathcal{G}) \wedge \\
&\quad (u \chi_0 u_1 \vee u \chi_0 u_2))\} \\
\chi_2 &:= \{(ws, w't) \in \mathcal{E}_l(\mathcal{G})^2 : l \in \mathbb{N}_+ \wedge \exists a \in \Sigma \exists u_1, u_2 \in \mathcal{V}(\mathcal{G}) (a(u_1, w, u_2) \in \mathcal{J}(\mathcal{G}) \wedge \\
&\quad (u_1 \chi_1 w' \vee u_2 \chi_1 w'))\}
\end{aligned}$$

The incompatibility relation is intended to capture the impossibility of two elements, vertices or edges, of the same layer, belonging to the same graph. Indeed,  $\chi_0$  says that applying two different productions on the same vertex or edge results in two sets (of vertices or edges) that cannot belong to the same graph. A bit more subtle is the impossibility related to inherited edges as defined by  $\chi_1$  and  $\chi_2$ . The former says that an inherited edge is incompatible with every vertex or edge which is incompatible with its source or its target. The latter says that two inherited edges are incompatible if one is incident to a vertex incompatible with the other. Note that  $\chi$  is irreflexive and, as  $\chi_0 = \chi_0^{-1}$ , it is also symmetric. The layerwise complement of  $\chi$  is the *compatibility relation*, written  $\circ$  and defined by:

$$\circ := \bigcup_{l \in \mathbb{N}} ((\mathcal{V}_l(\mathcal{G}) \cup \mathcal{E}_l(\mathcal{G}))^2 \setminus \chi) .$$

A set  $S \subseteq \mathcal{V}_l(\mathcal{G}) \cup \mathcal{E}(\mathcal{G})$  is *compatible* (resp. *maximal compatible*) if its elements are pairwise compatible, viz.,  $S^2 \subseteq \circ$  (resp. and is  $\subseteq$ -maximal among compatible sets). Note that  $\circ$  is reflexive and symmetric but not transitive. The relevance of this notion is highlighted by the following lemma closing this section.

**Lemma 4.11.** *If  $\mathcal{G}$  is a complete 0L graph grammar, then*

$$\{S \subseteq \mathcal{V}_l(\mathcal{G}) \cup \mathcal{E}_l(\mathcal{G}) : S \text{ is maximal compatible}\} = \{V \cup E : (V, E, \_, \_) \in \mathcal{L}_l(\mathcal{G})\} .$$

## 5 Expanders

In this section we focus on families of efficiently connected graphs. Such families called expanders (see [19] and [26] for more details) have been studied mostly in the case of undirected graphs of uniform degree, also called  $d$ -regular, where all members of a family have its vertices of the same degree  $d$ . This section is written assuming such setting.

Intuitively, an expander is a large graph which is sparse but efficiently connected relatively to its degree. This paradigm can be formalised using the Cheeger constant also known as the isoperimetric number, written  $h(G)$  for graph  $G = (V, E)$  and defined by

$$h(G) := \min \left\{ \frac{|\partial S|}{|S|} : S \subset V \wedge 0 < |S| \leq \frac{1}{2}|V| \right\} \text{ where } \partial S := \{ \{u, v\} \in E : u \in S \wedge v \in V \setminus S \} .$$

An infinite family  $\{G_i : i \in I\}$  is called *an expander family* if all its members have the same degree and there exists  $\epsilon \in \mathbb{R}_+$  such that  $h(G_i) > \epsilon$  for all  $i \in I$ . Although among graphs with  $n$  vertices the maximum value of the Cheeger constant is obtained for clique  $h(K_n) = n/2$ , because the degree of cliques grows with their size,  $\{K_n : n \in \mathbb{N}\}$  is not an expander family. As the problem of computing the Cheeger constant is NP-hard, one uses the isoperimetric inequalities of [1]

$$\frac{d - \lambda_2}{2} \leq h(G) \leq \sqrt{2d(d - \lambda_2)}$$

where  $d$  is the degree of  $G$  and  $\lambda_2$  stands for the second largest eigenvalue of the normalised, by factor  $1/d$ , adjacency matrix of  $G$ . Because  $\lambda_2 > 0$  one looks for graph families where all members have the second eigenvalue as small as possible. The smallest values are hit by *Ramanujan graphs* where  $\lambda_2 \leq 2\sqrt{d-1}$ . Efficient construction of expanders is a challenging research topic. In this section we review a few graph operations used in the most elementary

combinatorial constructions of expanders and show how these can be implemented within 0L graph grammar framework. Algebraic constructions (extensively covered in survey [26]) where expanders occur as Cayley graphs are beyond the scope of this paper.

### 5.1 2-lifts



As established in [29] and [30], starting from an arbitrary bipartite Ramanujan graph, the above very simple two rules let generate an infinite family of bipartite Ramanujan graphs. These two edge productions together with vertex production  $\bullet \rightarrow \bullet \bullet$  and another one, which from the axiom yields the starting bipartite Ramanujan graph, say  $B$ , form a complete vD0L graph grammar, say  $\mathcal{G}_B$ . However, only a strict subset of  $\mathcal{L}(\mathcal{G}_B)$  is Ramanujan. Indeed, it is shown in [29] that there exists a derivation along which encountered graphs form the desired family.<sup>5</sup> As it requires a nondeterministic choice among the above two vertex productions, the  $n$ -th graph of the family is obtained in time  $o(2^n)$ . However, an independently discovered essential improvement of this method leads to a polynomial time construction [6]. A family obtained is not limited to  $d$ -regular graphs. Indeed, the construction lets derive a family of  $(c, d)$ -biregular Ramanujan graphs starting from one such graph. A bipartite graph is  $(c, d)$ -biregular if all vertices in one part have degree  $c$  and those in the remaining part have degree  $d$ .

### 5.2 Shift 4-lifts



The construction using 2-lifts is further improved in [5] by using shift 4-lifts (or shift 3-lifts defined similarly). This operation can be written as the above four edge productions and one vertex production  $\bullet \rightarrow \bullet \bullet \bullet \bullet$ . Together with an axiom, we get a complete vD0L graph grammar. As in the example of 2-lifts, by selecting an appropriate derivation, one obtains an infinite family of bipartite Ramanujan graphs, provided that from the axiom one starts with its smallest member.

### 5.3 Graph operations using rotation maps

In this subsection we review graph operations introduced in [34] together with rotation maps. The latter equip an undirected graph with an additional structure: every vertex has its numbering of incident edges. Thus, every edge is equipped with two numbers. Such an enriched edge between vertices  $u$  and  $v$  can be written  $\{(u, i), (v, j)\}$ , where  $i, j \in [d]$  for a  $d$ -regular graph, to mean that  $v$  is the  $i$ -th neighbour of  $u$  and  $u$  is the  $j$ -th neighbour of  $v$ . In [34] it is denoted by  $\text{Rot}(u, i) = (v, j)$ , or, equivalently, by  $\text{Rot}(v, j) = (u, i)$ . To make it sound, one requires that for all enriched edges  $\{(u, i), (v, l)\}$  and  $\{(u, j), (w, k)\}$ , if  $i = j$  then  $v = w$  and  $l = k$ . Note that with this notation multiple edges between a pair of vertices arise naturally without a need of declaring an explicit set of edges. Rotation maps can be drawn using undirected edges with labels at each extremity,  $u \overset{i}{\text{---}} v$ , and implemented in a 0L graph grammar by two opposite directed incidences with a pair of labels drawn as  $u \overset{(i,j)}{\rightarrow} v$  and  $v \overset{(j,i)}{\rightarrow} u$ . One might suggest simplifying the numbering of the neighbours of each vertex so that  $i = j$  for every enriched edge  $\{(u, i), (v, j)\}$  and represent it as  $u \overset{i}{\text{---}} v$  but this is not possible in general as it can be quickly checked on  $K_3$ .

For the remainder of this section  $\mathcal{G}(V, d, \lambda)$  stands for the set of  $d$ -regular graphs with vertices  $V$  and the second largest eigenvalue of the normalised adjacency matrix at most  $\lambda$ . The set of enriched edges of graph  $G$  is written  $\mathcal{E}(G)$ .

<sup>5</sup>The original statement of [29] does not use graph grammars terminology.

## Replacement product

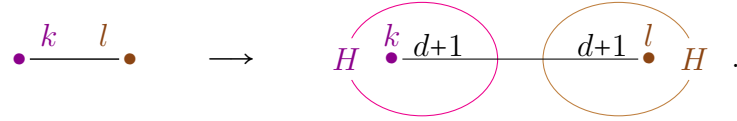
The replacement product takes two graphs  $G_1 \in \mathcal{G}([n_1], d_1, \lambda_1)$  and  $G_2 \in \mathcal{G}([d_1], d_2, \lambda_2)$  and outputs  $G_1 \oplus G_2 \in \mathcal{G}([n_1] \times [d_1], d_2 + 1, \lambda)$  with the following bound on  $\lambda$  as established in [34]:

$$\lambda \leq \left( \frac{d_2^3}{d_2^3 + 1} + \left(1 - \frac{d_2^3}{d_2^3 + 1}\right) \left( \frac{1}{2}(1 - \lambda_2^2)\lambda_1 + \frac{1}{2}\sqrt{(1 - \lambda_2^2)\lambda_1^2 + 4\lambda_2^2} \right) \right)^{\frac{1}{3}}.$$

Enriched edges defining  $G_1 \oplus G_2$  are

$$\begin{aligned} \mathcal{E}(G_1 \oplus G_2) &:= \{ \{((u, k), i), ((u, l), j)\} : u \in [n_1] \wedge \{(k, i), (l, j)\} \in \mathcal{E}(G_2)\} \\ &\cup \{ \{((u, k), d_2 + 1), ((v, l), d_2 + 1)\} : \{(u, k), (v, l)\} \in \mathcal{E}(G_1)\} \}. \end{aligned}$$

For all graphs  $G_0 \in \mathcal{G}(\_, d + 1, \_)$  and  $H \in \mathcal{G}([d + 1], d, \_)$  we can write D0L graph grammar  $\mathcal{G}_{\oplus H, G_0}$  with axiom production  $A \rightarrow G_0$ , another vertex production  $\bullet \rightarrow H$  and the set of edge productions  $\{P_{k,l} : k, l \in [d + 1]\}$  where each  $P_{k,l}$  has the following form



Note that  $k$  and  $l$  label the edge of  $\text{lhs}(P_{k,l})$  whereas in either copy of  $H$  in  $P_{k,l}$  they name its vertices. This grammar rewrites any graph  $G \in \mathcal{G}(\_, d + 1, \_)$  into  $G \oplus H \in \mathcal{G}(\_, d + 1, \_)$ . When  $G_0$  is a good expander and  $H$  is an odd cycle,  $\mathcal{L}(\mathcal{G}_{\oplus H, G_0})$  is an expander family of degree three [34]. However, in general, the expansion due to replacement product is not always satisfactory.

## Balanced replacement product

A slight improvement of expansion is obtained with the balanced replacement product. It takes two graphs  $G_1 \in \mathcal{G}([n_1], d_1, \lambda_1)$  and  $G_2 \in \mathcal{G}([d_1], d_2, \lambda_2)$  and outputs  $G_1 \odot G_2 \in \mathcal{G}([n_1] \times [d_1], 2d_2, \lambda)$  with the following bound on  $\lambda$  as established in [34]:

$$\lambda \leq \left( \frac{7}{8} + \frac{1}{16}(1 - \lambda_2^2)\lambda_1 + \frac{1}{16}\sqrt{(1 - \lambda_2^2)\lambda_1^2 + 4\lambda_2^2} \right)^{\frac{1}{3}}.$$

Enriched edges defining  $G_1 \odot G_2$  are

$$\begin{aligned} \mathcal{E}(G_1 \odot G_2) &:= \{ \{((u, k), i), ((u, l), j)\} : u \in [n_1] \wedge \{(k, i), (l, j)\} \in \mathcal{E}(G_2)\} \\ &\cup \{ \{((u, k), i), ((v, l), i)\} : \{(u, k), (v, l)\} \in \mathcal{E}(G_1) \wedge d_2 < i \leq 2d_2\} \}. \end{aligned}$$

Observe that the balanced replacement product differs from the former replacement product only by the multiplicity of edges between two copies of  $G_2$  which is 1 in the former and  $d_2$  in the present product. Thus, a vertex has equal number of neighbours within and outside each copy of  $G_2$ . The construction of D0L graph grammars is similar to those from the latter subsection except for multiple edges. For all graphs  $G_0 \in \mathcal{G}(\_, 2d, \_)$  and  $H \in \mathcal{G}([2d], d, \_)$ , one can write D0L graph grammar  $\mathcal{G}_{\odot H, G_0}$ . Expander families are obtained as for the former replacement product. In particular, the degree is kept constant because any graph  $G \in \mathcal{G}(\_, 2d, \_)$  rewrites in a single step into  $G \odot H \in \mathcal{G}(\_, 2d, \_)$ .

## Zig-zag product

An even better expansion is obtained with the zig-zag product. It takes two graphs  $G_1 \in \mathcal{G}([n_1], d_1, \lambda_1)$  and  $G_2 \in \mathcal{G}([d_1], d_2, \lambda_2)$  and outputs  $G_1 \otimes G_2 \in \mathcal{G}([n_1] \times [d_1], d_2^2, \lambda)$  with the following bound on  $\lambda$  as established in [34]:

$$\lambda \leq \left( \frac{1}{2}(1 - \lambda_2^2)\lambda_1 + \frac{1}{2}\sqrt{(1 - \lambda_2^2)\lambda_1^2 + 4\lambda_2^2} \right)^{\frac{1}{3}}.$$

The “number” of a neighbour in  $G_1 \otimes G_2$  is a pair in  $[d_2]^2$  according to the original definition. Enriched edges defining  $G_1 \otimes G_2$  are

$$\mathcal{E}(G_1 \otimes G_2) := \left\{ \left\{ ((u, k), (i, j)), ((v, l), (j', i')) \right\} : \right. \\ \left. \exists k', l' \in [d_1] \quad \left( \{(k, i), (k', i')\}, \{(l, j'), (l', j)\} \in \mathcal{E}(G_2) \wedge \right. \right. \\ \left. \left. \{(u, k'), (v, l')\} \in \mathcal{E}(G_1) \right) \right\} .$$

An enriched edge  $u \xrightarrow{k' l'} v$  of  $G_1$  gives rise to  $d_2^2$  such new enriched edges between copies  $\{u\} \times V_2$  and  $\{v\} \times V_2$  of vertices  $V_2$  of  $G_2$ . Label  $k'$  (resp.  $l'$ ) determines the vertex that is “the entry point” of copy  $\{u\} \times V_2$  (resp.  $\{v\} \times V_2$ ). There is an edge between vertex  $(u, k)$  in copy  $\{u\} \times V_2$  and vertex  $(v, l)$  in copy  $\{v\} \times V_2$  if one can take an edge  $k \xrightarrow{i i'} k'$  in  $G_2$  from  $k$  to entry point  $k'$  and an edge  $l \xrightarrow{j' j} l'$  in  $G_2$  from  $l$  to entry point  $l'$ . Because of edge  $u \xrightarrow{k' l'} v$  in  $G_1$ , these entry points allow to “jump” between copies  $\{u\} \times V_2$  and  $\{v\} \times V_2$ . The label on  $(u, k)$ 's side of created enriched edge  $(u, k) \xrightarrow{(i, j) \quad (j', i')} (v, l)$  describes the corresponding “route with jump” from  $(u, k)$  to  $(v, l)$  as follows. *From  $k$  take direction  $i$  to get to the entry point of current copy. Jump and you land at the entry point of  $v$ 's copy. From there, take direction  $j$  and you are there.* Label  $(j', i')$  on  $(v, l)$ 's side describes the same route in the reverse direction.

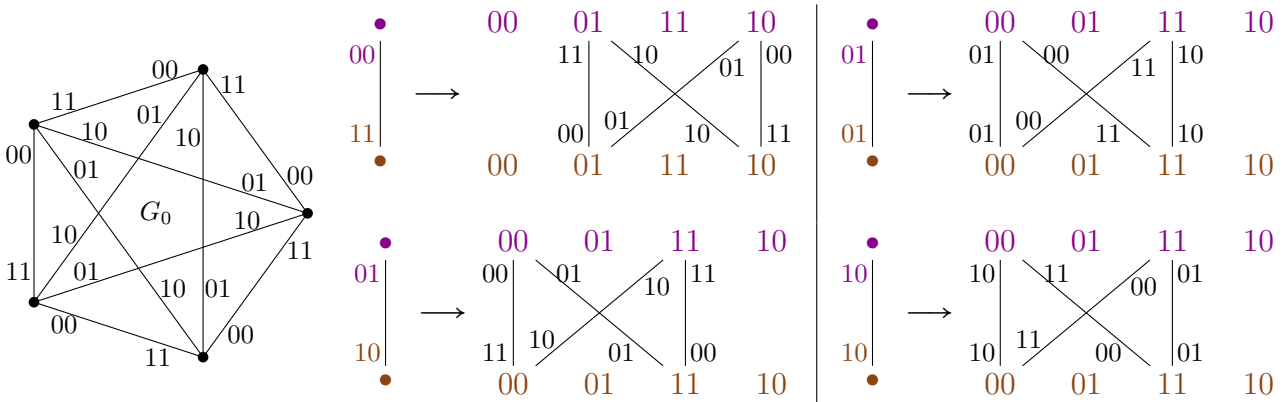
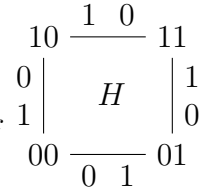
In order to implement the zig-zag product in DOL graph grammar so that it can be iterated, the set of labels used for neighbours numbering must be kept constant at every derivation step. The choice we made is to take, as the right hand side of the axiom production,  $G_0 \in \mathcal{G}(\_, d^2, \_)$ , where each edge has at either extremity a pair in  $[d]^2$  as label. With  $H \in \mathcal{G}([d]^2, d, \_)$ , one can complete the construction of DOL graph grammar  $\mathcal{G}_{\otimes H, G_0}$  suitable for such iteration. Besides the axiom,  $\mathcal{G}_{\otimes H, G_0}$  has one vertex production  $\bullet \rightarrow V$  where  $V$  is the edgeless version of  $H$ . The set of edge productions  $\{P_{k', l'} : k', l' \in [d]^2\}$  is defined so that each lhs( $P_{k', l'}$ ) has  $\bullet \xrightarrow{k' l'} \bullet$  and the corresponding rhs( $P_{k', l'}$ ) creates edges between each pair  $\{k, l\}$  of neighbours of  $k'$  and  $l'$  in their respective copies of  $V$ . Every such edge of rhs( $P_{k', l'}$ ) can be drawn as

$$\begin{array}{c} \bullet \xrightarrow{k' l'} \bullet \\ \vdots \quad \vdots \\ \bullet \xrightarrow{i k(i, j) \quad (j', i') l j'} \bullet \end{array}$$

without dashed edges which serve here only as explanation. Remember that  $k, l, k', l'$  are pairs in  $[d]^2$  whereas  $i, j, i', j'$  are numbers in  $[d]$ .

### Example 5.1

In this example  $d = 2$ . We use  $\{0, 1\}$  for neighbours numbering in  $H$  (depicted on the right) and  $\{0, 1\}^2$  for naming its vertices and for neighbours numbering of  $G_0 := K_5$ . We have the following edge productions that apply on  $K_5$  below.



For more clarity, the names of vertices are written in the right hand sides of edge productions. Although the number of possible edge labellings is  $|\{\{ab, cd\} : a, b, c, d \in \{0, 1\}\}| = 10$ , but only four appear in  $G_0$  and these form the left hand sides of the above edge productions. Moreover, their right hand sides only introduce three types of enriched edges among those of  $G_0$ . Thus,

this DOL graph grammar is like  $\mathcal{G}_{\otimes H, G_0}$  except that it has no useless productions. Consequently the language generated here is as expected  $\{G_n : n \in \mathbb{N}\}$  where  $G_{i+1} = G_i \otimes H$ .

To close this section, we need to mention that the most efficient construction of expanders of [34] with the zig-zag product requires an additional operation of squaring. It consists in replacing an incident pair of edges  $u-v-w$  by a single edge  $u-w$ . Within Lindenmayer graph grammars, squaring can be handled in 1L (instead of 0L) framework, unfortunately, at the expense of most decidability results.

## 6 0L graph languages as models

We start this section by fixing first order language  $\text{FO}(\Gamma \cup \Sigma)$  and its extension  $\text{FO}(\Upsilon)$  for expressing properties or queries concerning 0L graph languages. Before focusing on  $\text{FO}(\Gamma \cup \Sigma)$ , questions one would like to ask are conveniently qualified in a more general scope than graph languages, namely that of arbitrary family of structures  $\mathcal{S}$  together with adequate logic  $\mathcal{L}$ . We consider the

LANGUAGE  $\mathcal{L}$ -CHECKING PROBLEM FOR FAMILY  $\mathcal{S}$  OF STRUCTURES

Instance: a countable set of countable structures  $\mathcal{S} \in \mathcal{S}$  and an  $\mathcal{L}$ -sentence  $\varphi$

Question:  $\exists S \in \mathcal{S} S \models \varphi$ ?

Replacing the existential quantifier in the question by the universal one leads to an equivalent checking problem. Consequently, any finite Boolean combination of existential and universal questions leads to an equivalent checking problem too. The reader should note the difference of the above problems with the well-known

MODEL  $\mathcal{L}$ -CHECKING PROBLEM FOR CLASS  $\mathcal{S}$  OF STRUCTURES

Instance: a countable structure  $S \in \mathcal{S}$  and an  $\mathcal{L}$ -sentence  $\varphi$  (resp.  $\mathcal{L}$ -formula  $\varphi(\mathcal{X})$ ),

Question:  $S \models \varphi$ ?

Whereas the instances in the latter problem range over individual structures, they range over sets of structures in the former. Both can be confused when the set of sentences is, up to logical equivalence, in one-to-one correspondence with family  $\mathcal{S}$ , like for instance in the case of MSO and the family of regular sets.

In the present paper, we are interested in the following case of the former generic problems: *the language  $\text{FO}(\Gamma \cup \Sigma)$ -checking (resp.  $\text{FO}(\Gamma \cup \Sigma)$ -querying) problem for the family of 0L graph languages.* Given a 0L grammar  $\mathcal{G}$  and an  $\text{FO}(\Gamma \cup \Sigma)$  sentence  $\varphi$ , we ask whether a graph in  $\mathcal{L}(\mathcal{G})$  satisfies  $\varphi$ . Sentence  $\varphi$  can for instance express the property that a graph has a clique of a given fixed size as induced subgraph. To establish the decidability of the above problem, we proceed by reduction of the language  $\text{FO}(\Gamma \cup \Sigma)$ -checking problem for the family of 0L graph languages to the model  $\text{FO}(\Upsilon)$ -checking problem for a class of infinite structures associated with those languages, where the sentences are written using relational signature  $\Upsilon$  extending  $\Gamma \cup \Sigma$ . By representing  $\mathcal{L}(\mathcal{G})$  as an adequate single structure, say  $\mathcal{G}^\dagger$ , given 0L graph grammar  $\mathcal{G}$  and sentence  $\varphi \in \text{FO}(\Gamma \cup \Sigma)$ , under certain assumptions, one can reduce a question “ $\exists G \in \mathcal{L}(\mathcal{G}) G \models \varphi$ ?” to a question “ $\mathcal{G}^\dagger \models \psi$ ?” with  $\psi \in \text{FO}(\Upsilon)$ . In doing so, one needs to separate somehow individual graphs within  $\mathcal{G}^\dagger$  using an appropriate sentence  $\psi$  associated to  $\varphi$ . Indeed, as pointed out in Example 4.4. two graphs in  $\mathcal{L}(\mathcal{G})$  need not be disjoint.

Variables from a countable set  $\mathcal{X}$  are used for writing  $\text{FO}(\Gamma \cup \Sigma)$  formulae expressing properties of a graph in  $\mathcal{G}(\Gamma, \Sigma, \_, \_)$ . Atomic formulae of  $\text{FO}(\Gamma \cup \Sigma)$  are

- “ $x$  is an  $A$ -labelled vertex”:  $A(x)$ , for  $A \in \Gamma$  and  $x \in \mathcal{X}$ ,
- “ $y$  is an  $a$ -labelled edge from vertex  $x_1$  to vertex  $x_2$ ”:  $a(x_1, y, x_2)$ , for  $a \in \Sigma$ ,  $x_1, y, x_2 \in \mathcal{X}$ ,
- equalities  $x = y$ , for  $x, y \in \mathcal{X}$ ,



- and those for discriminating vertices and edges:  $\text{vrt}(x)$ ,  $\text{edg}(x)$ , for  $x \in \mathcal{X}$ .

The closure of the atomic formulae under connectives  $\neg$ ,  $\supset$ ,  $\vee$ , and quantifiers  $\exists$ ,  $\forall$  yields the set of  $\text{FO}(\Gamma \cup \Sigma)$  formulae.

Additional symbols in  $\text{FO}(\Upsilon) \setminus (\Gamma \cup \Sigma)$  are used for writing the following atomic formulae:

- “vertex (resp. edge)  $y$  is in a graph which derives from vertex (resp. or edge)  $x$ ”:  $x \sqsubseteq y$  for  $x, y \in \mathcal{X}$ ,
- “vertices or edges  $x$  and  $y$  are compatible”:  $x \circ y$  for  $x, y \in \mathcal{X}$ ,
- “vertices or edges  $x$  and  $y$  belong to graphs of the same layer”  $x \hat{=} y$  for  $x, y \in \mathcal{X}$ .

The structure, written  $\mathcal{G}^\dagger$ , underlying the language  $\mathcal{L}(\mathcal{G})$  of 0L graph grammar  $\mathcal{G}$  is a relational structure over signature  $\Upsilon$ .  $\mathcal{G}^\dagger$  consists of the universe of vertices and edges  $\mathcal{V}(\mathcal{G}) \cup \mathcal{E}(\mathcal{G})$  where predicate symbols of  $\Upsilon$  are interpreted in the following way:

- $A \in \Gamma$  as unary relation

$$\mathcal{G}^\dagger(A) = \{v \in \mathcal{V}(\mathcal{G}) : \exists (V, \_, \ell, \_) \in \mathcal{L}(\mathcal{G}) \ell(v) = A\},$$

- $a \in \Sigma$ , as ternary relation

$$\mathcal{G}^\dagger(a) = \{(u, w, v) : \exists (\_, \_, \_, \_] \in \mathcal{L}(\mathcal{G}) \ a(u, w, v) \in \]\},$$

- “ $\circ$ ”, as the compatibility relation defined in Sect. 4,
- “ $\sqsubseteq$ ”, as prefix ordering on  $\mathcal{V}(\mathcal{G}) \cup \mathcal{E}(\mathcal{G})$ ,
- $\mathcal{G}^\dagger(\hat{=}) = \bigcup_{l \in \mathbb{N}} (\mathcal{V}_l(\mathcal{G}) \cup \mathcal{E}_l(\mathcal{G}))^2$ , and,
- $\mathcal{G}^\dagger(\text{vrt}) = \mathcal{V}(\mathcal{G})$  and  $\mathcal{G}^\dagger(\text{edg}) = \mathcal{E}(\mathcal{G})$ .

For a set  $N$  of non-terminals,  $\mathcal{G}^{\dagger \setminus N}$  is a restriction of  $\mathcal{G}^\dagger$  onto  $\mathcal{V}(\mathcal{G})_N \cup \mathcal{E}(\mathcal{G})_N$ .

In Sect. 7 we give a construction that yields an automatic presentation of  $\mathcal{G}^\dagger$  for every complete 0L graph grammar  $\mathcal{G}$ . Our construction brings the following result.

**Theorem 6.1.**  $\mathcal{G}^\dagger$  (resp.  $\mathcal{G}^{\dagger \setminus N}$ ) has a decidable  $\text{FO}(\Upsilon)$  theory for every 0L graph grammar  $\mathcal{G}$  which is complete (resp. deterministic with non-terminals  $N$ ).

The following theorem is essential for  $\text{FO}(\Gamma \cup \Sigma)$ -language checking.

**Theorem 6.2.** There is a translation  $\tau: \text{FO}(\Gamma \cup \Sigma) \rightarrow \text{FO}(\Gamma \cup \Sigma \cup \{\hat{=}\})$  in time  $\mathcal{O}(nk + m)$ , where  $m$  is the length of the input formula,  $n$  is its alternation rank and  $k$  is the number of its variables, such that for every unix 0L graph grammar  $\mathcal{G} \in \text{0L}(\Gamma, \Sigma, \_, \_)$  and every formula  $\varphi(\mathcal{X}) \in \text{FO}(\Gamma \cup \Sigma)$ , one has

$$\mathcal{G}^\dagger(\tau(\varphi(\mathcal{X}))) = \bigcup_{G \in \mathcal{L}(\mathcal{G})} G(\varphi(\mathcal{X})) .$$

In particular, when  $\mathcal{X} = \emptyset$ , there exists  $G \in \mathcal{L}(\mathcal{G})$  such that  $G \models \varphi$ , if, and only if,  $\mathcal{G}^\dagger \models \tau(\varphi)$ .

Additional translation steps can be done using the following proposition.

**Proposition 6.3.** For every set of non-terminals  $N \subseteq \Gamma \cup \Sigma$ , there is a translation  $\tau$  of  $\text{FO}(\Gamma \cup \Sigma)$  in time  $\mathcal{O}(|N|)$ , such that for every graph grammar  $\mathcal{G} \in \text{0L}(\Gamma, \Sigma, \_, \_)$  and every formula  $\varphi(\mathcal{X}) \in \text{FO}(\Gamma \cup \Sigma)$ , one has  $G(\tau(\varphi(\mathcal{X}))) = G(\varphi(\mathcal{X}))$  for every  $G \in \mathcal{L}(\mathcal{G})_N$  and  $G(\tau(\varphi(\mathcal{X}))) = \emptyset$  for every  $G \in \mathcal{L}(\mathcal{G}) \setminus \mathcal{L}(\mathcal{G})_N$ .

In particular, when  $\mathcal{X} = \emptyset$ , there exists  $G_N \in \mathcal{L}(\mathcal{G})_N$  such that  $G_N \models \varphi$ , if, and only if, there exists  $G \in \mathcal{L}(\mathcal{G})$  such that  $G \models \tau(\varphi)$ .

*Proof.* The translation is obvious:  $\tau(\varphi(\mathcal{X})) = \varphi(\mathcal{X}) \wedge \neg \exists x \bigvee_{A \in N \cap \Gamma} A(x) \wedge \neg \exists xyz \bigvee_{a \in N \cap \Sigma} a(x, y, z)$ .  $\square$

Applied to Prop. 4.2, the above proposition allows for extending the translation of Thm. 6.2. Then, using Thm. 6.1, we conclude this section with the following result.

**Corollary 6.4.** *FO( $\Gamma \cup \Sigma$ )-language checking problem is decidable DOL graph grammars even with non-terminals.*

It is important to note that the corollary does not extend to unix graph grammars. The essential step of the proof uses Prop. 4.2 for transforming the input grammar into a complete grammar so as to benefit from Thm. 6.1. However, the grammar resulting from that step is not unix in general.

## 7 Automatic presentation of Lindenmayer languages

In this section we build an injective automatic presentation of  $\mathcal{G}^\dagger$ , where the corresponding injection is in fact an identity map. The reader is supposed to be familiar with standard constructions on finite automata such as Boolean operations, concatenation or projection. The following additional constructions are used in the sequel.

- For a tuple automaton  $\mathcal{B}$  we denote by  $\mathcal{B}^=$  the tuple automaton which differs from  $\mathcal{B}$  by labelling of transitions: each transition  $p \xrightarrow{\bar{a}} q$  of  $\mathcal{B}$  is replaced in  $\mathcal{B}^=$  by  $p \xrightarrow{\bar{a}\bar{a}} q$ . Thus  $\llbracket \mathcal{B}^= \rrbracket = \{(\bar{u}, \bar{u}) : \bar{u} \in \llbracket \mathcal{B} \rrbracket\}$  with  $\text{ar}(\mathcal{B}^=) = 2\text{ar}(\mathcal{B})$ .
- The *loose product* of tuple automata  $\mathcal{A}_1 = (\mathbb{Q}_1, \Delta_1, \mathbb{l}_1, \mathbb{F}_1)$  and  $\mathcal{A}_2 = (\mathbb{Q}_2, \Delta_2, \mathbb{l}_2, \mathbb{F}_2)$  is  $\mathcal{A}_1 \times \mathcal{A}_2 = (\mathbb{Q}_1 \times \mathbb{Q}_2, \Delta_1 \times \Delta_2, \mathbb{l}_1 \times \mathbb{l}_2, \mathbb{F}_1 \times \mathbb{F}_2)$  where

$$\Delta_1 \times \Delta_2 := \{(p, q) \xrightarrow{\bar{a}\bar{b}} (p', q') : p \xrightarrow{\bar{a}} p' \in \Delta_1 \wedge q \xrightarrow{\bar{b}} q' \in \Delta_2\}.$$

Thus  $\mathcal{L}(\mathcal{A}_1 \times \mathcal{A}_2) = \bigcup_{l \in \mathbb{N}} \mathcal{L}_l(\mathcal{A}_1) \times \mathcal{L}_l(\mathcal{A}_2)$ .

- For  $\mathcal{A}_1$  and  $\mathcal{A}_2$  as above and  $\Delta \subseteq \mathbb{Q}_1 \times \Pi \times \mathbb{Q}_2$ , where  $\Pi$  is an alphabet, the *concatenation of  $\mathcal{A}_1$  and  $\mathcal{A}_2$  via  $\Delta$* , written  $\mathcal{A}_1 \xrightarrow{\Delta} \mathcal{A}_2$  is the automaton  $(\mathbb{Q}_1 \cup \mathbb{Q}_2, \Delta_1 \cup \Delta \cup \Delta_2, \mathbb{l}_1, \mathbb{F}_2)$ .

Each predicate symbol  $\varrho$  of the signature of  $\mathcal{G}^\dagger$  has the corresponding automaton  $\mathcal{A}(\varrho)$  of arity  $\text{ar}(\mathcal{A}(\varrho)) = \text{ar}(\varrho)$  given as quadruple  $\mathcal{A}(\varrho) = (\mathbb{Q}_\varrho, \Delta_\varrho, \iota_\varrho, \mathbb{F}_\varrho)$ , where  $\mathbb{Q}_\varrho$  is a finite set of states,  $\Delta_\varrho \subseteq \mathbb{Q}_\varrho \times (\Theta \cup \Omega)^{\text{ar}(\varrho)} \times \mathbb{Q}_\varrho$  is a transition relation,  $\iota_\varrho$  is the initial state of  $\mathcal{A}(\varrho)$  and  $\mathbb{F}_\varrho \subseteq \mathbb{Q}_\varrho$  is its set of final states. The relevant predicate symbols are

- the equality,
- unary **vrt** and **edg**,
- binary  $\circ$ ,  $\hat{=}$  and  $\sqsubseteq$ .
- $A$ , with  $\text{ar}(A) = 1$ , for  $A \in \Gamma$ ,
- $a$ , with  $\text{ar}(a) = 3$ , for  $a \in \Sigma$ ,

Automata for  $\hat{=}$ ,  $\sqsubseteq$  and for the equality are obvious. Besides direct constructions, we also consider automata for FO-definable predicates without providing their construction. Indeed, such automata can be constructed in standard way from the corresponding formula as query automata (see Subsect. 3).

Here are the details of main automata, starting with vertex related ones.

- $\mathcal{A}(\text{vrt}) := (\mathbb{Q}_{\text{vrt}}, \Delta_{\text{vrt}}, \iota_{\text{vrt}}, \mathbb{F}_{\text{vrt}})$  is a deterministic automaton on  $\Theta$  where

$\mathbb{Q}_{\text{vrt}} := \Theta \cup \{\iota_{\text{vrt}}\}$  is its set states,

$\iota_{\text{vrt}} \notin \Theta$  is its initial state,

$\mathbb{F}_{\text{vrt}} := \Theta$  is its set of final states,

$\Delta_{\text{vrt}} \subseteq \mathbb{Q}_{\text{vrt}} \times \Theta \times \mathbb{Q}_{\text{vrt}}$  is its set of transitions defined by

$$\Delta_{\text{vrt}} := \{i \xrightarrow{j} j : \text{prd}(j) = \ell(i) \rightarrow \_ \} \cup \{\iota_{\text{vrt}} \xrightarrow{j} j : \text{prd}(j) = A_G \rightarrow \_ \} .$$

As  $\mathcal{A}(\text{vrt})$  is deterministic,  $\Delta_{\text{vrt}}$  is implemented as a table with rows indexed by  $\mathbb{Q}_{\text{vrt}}$  and columns indexed by input symbols  $\Theta$ . Filling this table is done in time  $\mathcal{O}(|\Theta|^2)$ .

**Lemma 7.1.** *If  $\mathcal{G}$  is complete, then  $\llbracket \mathcal{A}(\text{vrt}) \rrbracket = \mathcal{G}^\dagger(\text{vrt})$ .*

- For every  $A \in \Gamma$  automaton  $\mathcal{A}(A)$  over  $\Theta$  is like  $\text{vrt}$  except for its final states, namely  $\mathcal{A}(A) := (\mathbb{Q}_{\text{vrt}}, \Delta_{\text{vrt}}, \iota_{\text{vrt}}, \mathbb{F}_A)$  where  $\mathbb{F}_A := \{i \in \mathbb{F}_{\text{vrt}} : \ell(i) = A\}$ . It is built on-the-fly when constructing  $\mathcal{A}(\text{vrt})$ . It shares its transition table with  $\mathcal{A}(\text{vrt})$  and needs only different marking of final states. Such marking requires linear time with respect to the number of states.

**Corollary 7.2.** *If  $\mathcal{G}$  is complete, then  $\llbracket \mathcal{A}(A) \rrbracket = \mathcal{G}^\uparrow(A)$  for every  $A \in \Gamma$ .*

- Let  $\Sigma_\diamond := \overrightarrow{\Sigma} \cup \overleftarrow{\Sigma} \cup \{\diamond\}$  with  $\diamond \notin \Sigma$  and  $\Omega_\square := \Omega \cup \{\square\}$  with  $\square \notin \Omega$ . We define  $\mathcal{A}_j := (\mathbb{Q}_j, \Delta_j, \iota_j, \mathbb{F}_j)$ , running over incidences, where
 
$$\begin{aligned} \mathbb{Q}_j &:= \Sigma_\diamond \times \mathbb{Q}_{\text{vrt}} \times \Omega_\square \times \mathbb{Q}_{\text{vrt}}, \\ \iota_j &:= (\diamond, \iota_{\text{vrt}}, \square, \iota_{\text{vrt}}), \\ \mathbb{F}_j &:= (\overrightarrow{\Sigma} \cup \overleftarrow{\Sigma}) \times \mathbb{F}_{\text{vrt}} \times \Omega \times \mathbb{F}_{\text{vrt}}, \end{aligned}$$

and  $\Delta_j \subseteq \mathbb{Q}_j \times \Theta \times (\Theta \cup \Omega) \times \Theta \times \mathbb{Q}_j$  is defined assuming the following shorthands:

- $\text{prd}(i) \xrightarrow{\overleftarrow{b}} \text{prd}(k) \rightarrow \sqsupset$  stands for  $\text{prd}(i) \xrightarrow{b} \text{prd}(k) \rightarrow \sqsupset$  when  $\overleftarrow{b} = \overrightarrow{b}$  and for  $\text{prd}(k) \xrightarrow{b} \text{prd}(i) \rightarrow \sqsupset$  when  $\overleftarrow{b} = \overleftarrow{b}$ ,
- $[\overrightarrow{a}, i, j, k]_j$  means  $(\overrightarrow{a}, i, j, k)$  and one has  $\{a(i, j, k), \overrightarrow{a}(i, j, k), \overleftarrow{a}(k, j, i)\} \cap \sqsupset \neq \emptyset$ , or it means  $(\overleftarrow{a}, i, j, k)$  and one has  $\{a(k, j, i), \overrightarrow{a}(k, j, i), \overleftarrow{a}(i, j, k)\} \cap \sqsupset \neq \emptyset$ .

Using these shorthands,  $\Delta_j$  is defined as follows:

$$\begin{aligned} \Delta_j &:= \left\{ (\overleftarrow{b}, i', j', k') \xrightarrow{(i, j, k)} [\overrightarrow{a}, i, j, k]_j : \text{prd}(j) = \text{prd}(i) \xrightarrow{\overleftarrow{b}} \text{prd}(k) \rightarrow \sqsupset \in \mathbb{E}_G \wedge \right. \\ &\quad \left. \text{prd}(i) = \ell(i') \rightarrow \_ \in \mathbb{V}_G \wedge \text{prd}(k) = \ell(k') \rightarrow \_ \in \mathbb{V}_G \right\} \cup \\ &\quad \left\{ (\diamond, i', \square, i') \xrightarrow{(i, j, k)} (\overrightarrow{a}, i, j, k) : \ell(i') \rightarrow (\_, \_, \_, \mathbb{1}) \in \mathbb{V}_G \wedge a(i, j, k) \in \mathbb{1} \right\} \cup \\ &\quad \left\{ (\diamond, i', \square, i') \xrightarrow{(i, i, i)} (\diamond, i, \square, i) : i' \xrightarrow{i} i \in \Delta_{\text{vrt}} \right\} \cup \\ &\quad \left\{ \iota_j \xrightarrow{(i, i, i)} (\diamond, i, \square, i) : \iota_{\text{vrt}} \xrightarrow{i} i \in \Delta_{\text{vrt}} \right\} \cup \\ &\quad \left\{ \iota_j \xrightarrow{(i, j, k)} (\overrightarrow{a}, i, j, k) : \mathbb{A}_G \rightarrow (\_, \_, \_, \mathbb{1}) \wedge a(i, j, k) \in \mathbb{1} \right\}. \end{aligned}$$

As  $\mathcal{A}_j$  is deterministic and every accessible state  $(\overleftarrow{c}, \_, \_, \_ ) \in \mathbb{Q}_j$  is in  $\mathbb{F}_j$ , the following notation makes sense  $\llbracket \mathcal{A}_j \rrbracket := \{ \overleftarrow{c}(u, w, v) : \iota_j \xrightarrow{\otimes(u, w, v)} (\overleftarrow{c}, \_, \_, \_) \}$ , with dashed arrow representing a path of  $\mathcal{A}_j$ . By considering that every  $\overleftarrow{c} \in \overrightarrow{\Sigma} \cup \overleftarrow{\Sigma}$  represents an operation mapping triples to incidences,  $\overleftarrow{c} : (u, w, v) \mapsto c(u, w, v)$  and  $\overleftarrow{c} : (u, w, v) \mapsto c(v, w, u)$ , the soundness of the construction of  $\mathcal{A}_j$  is stated as follows.

**Lemma 7.3.** *If  $\mathcal{G}$  is complete, then  $\llbracket \mathcal{A}_j \rrbracket = \mathcal{J}(\mathcal{G})$ , where  $\mathcal{J}(\mathcal{G})$  is the set of incidences of  $\mathcal{G}^\uparrow$ .*

For understanding the time complexity of the construction of  $\mathcal{A}_j$ , remember that  $\mathcal{E}(\mathcal{G})$  is in bijection with  $\mathcal{J}(\mathcal{G})$ . Although  $\mathbb{Q}_j = \Sigma_\diamond \times \mathbb{Q}_{\text{vrt}} \times \Omega_\square \times \mathbb{Q}_{\text{vrt}}$ , for every  $j \in \Omega$ , there is exactly one state  $(\_, \_, j, \_) \in \mathbb{Q}_j$ . Indeed, for subset  $\Sigma_\diamond \times \mathbb{Q}_{\text{vrt}} \times \Omega \times \mathbb{Q}_{\text{vrt}}$  of  $\mathbb{Q}_j$ , its components other than  $\Omega$  aim only at keeping their track for making the definition and the proof simple. Thus, the transition table storing  $\Delta_j$  for this subset of  $\mathbb{Q}_j$ , can be thought as having its rows indexed by  $\Omega$ . For subset  $\Sigma_\diamond \times \mathbb{Q}_{\text{vrt}} \times \{\square\} \times \mathbb{Q}_{\text{vrt}}$  of  $\mathbb{Q}_j$ ,  $\Delta_j$  is isomorphic to  $\Delta_{\text{vrt}}$ . We do not need to construct this part of  $\Delta_j$  explicitly. We can plug  $\Delta_{\text{vrt}}$  into  $\Delta_j$  via transitions corresponding to the second member of the union in the above definition of  $\Delta_j$ . When so plugging  $\mathcal{A}(\text{vrt})$  into the optimised version of  $\mathcal{A}_j$ , we only need to “triple” the input word of  $\mathcal{A}(\text{vrt})$ , say  $u$ . We consider that  $\otimes(u, u, u)$  instead  $u$  is being read when switching from  $\mathcal{A}(\text{vrt})$  to the optimised version of  $\mathcal{A}_j$ . Thus, when constructing this version, we only need to build from scratch the portion of  $\Delta_j$  with rows indexed by  $\Omega$ . Although the columns for this portion are indexed by triples in  $\Theta \times \Omega \times \Theta$ , only those triples that correspond to incidences in  $\text{rhs}(\mathbb{V}_G)$  are relevant. The number of such incidences is bound by  $|\Omega|$ . Consequently, the time complexity of this construction is in  $\mathcal{O}(|\Omega|^2)$ .

Edge related automata are obtained from auxiliary automaton  $\mathcal{A}_j$ . For every  $\vec{a} \in \vec{\Sigma} \cup \overleftarrow{\Sigma}$ , we define  $\mathcal{A}(\vec{a})$  which is like  $\mathcal{A}_j$ , except for its final states  $\mathbb{F}_{\vec{a}} := \{\vec{a}\} \times \mathbb{F}_{\text{vrt}} \times \Omega \times \mathbb{F}_{\text{vrt}}$ . Then

- for every  $a \in \Sigma$ , we define  $\mathcal{A}(a) := \mathcal{A}(\vec{a}) \cup \text{swap}_{1 \leftrightarrow 3}(\mathcal{A}(\overleftarrow{a}))$  where  $\text{swap}_{1 \leftrightarrow 3}$  only modifies transitions' labels by exchanging their 1st and 3rd components:

$$\text{swap}_{1 \leftrightarrow 3}(\Delta_{\vec{a}}) := \{p \xrightarrow{(i,j,k)} q : p \xrightarrow{(k,j,i)} q \in \Delta_{\vec{a}}\}.$$

- $\mathcal{A}(\text{edg})$  is obtained by projecting  $\mathcal{A}_j$  on the second component:  $\mathcal{A}(\text{edg}) = \pi_2(\mathcal{A}_j)$ . Note that this projection preserves the determinism, again, thanks to the bijection between edges and incidences.

All these automata share their transition table with  $\mathcal{A}_j$  and are derived from the latter in linear time.

**Corollary 7.4.** *If  $\mathcal{G}$  is complete, then  $\llbracket \mathcal{A}(a) \rrbracket = \mathcal{G}^\uparrow(a)$  for every  $a \in \Gamma$  and  $\llbracket \mathcal{A}(\text{edg}) \rrbracket = \mathcal{E}(\mathcal{G})$ .*

Since the compatibility relation comes as a bonus that is not essential for the other results, before focusing on it, we summarise the complexity of the basic construction.

**Theorem 7.5.** *For every complete graph grammar  $\mathcal{G} \in \text{OL}(\_, \_, \Theta, \Omega)$ , one can construct in time  $\mathcal{O}(|\Theta|^2 + |\Omega|^2)$  an automatic presentation of  $\mathcal{G}^\uparrow$  without the compatibility relation.*

It remains to define  $\mathcal{A}(\circ)$ . Although its direct construction has a better complexity, to keep this presentation simple, we get  $\mathcal{A}(\circ)$  as the complement  $\mathcal{A}(\chi)$  which in turn is obtained as an adequate query automaton. We start with  $\mathcal{A}(\chi_0)$ . Observe first that, as  $\mathcal{A}_j$  mimics  $\mathcal{A}(\text{vrt})$  before reaching its accepting states, if we modify its projection  $\mathcal{A}(\text{edg})$  by making accepting all its states but  $\iota_j$ , we get precisely  $\mathcal{A}_\delta$ :  $\llbracket \mathcal{A}_\delta \rrbracket = \mathcal{V}(\mathcal{G}) \cup \mathcal{E}(\mathcal{G})$ . Let then  $\mathcal{B}_1 := \mathcal{A}_\delta^-$  and  $\mathcal{B}_2 := \mathcal{A}_\delta^2$  be the loose product of  $\mathcal{A}_\delta$  by itself. Observe that  $\llbracket \mathcal{B}_1 \rrbracket$  is the equality on  $\mathcal{V}(\mathcal{G}) \cup \mathcal{E}(\mathcal{G})$  and  $\llbracket \mathcal{B}_2 \rrbracket = \bigcup_{l \in \mathbb{N}} (\mathcal{V}_l(\mathcal{G}) \cup \mathcal{E}_l(\mathcal{G}))^2$ . Note also that  $\mathcal{B}_1$  and  $\mathcal{B}_2$  have disjoint sets of states. Then, we define  $\mathcal{A}(\chi_0)$  as their concatenation  $\mathcal{A}(\chi_0) := \mathcal{B}_1 \xrightarrow{\Delta} \mathcal{B}_2$  where  $\Delta$  switches the control from  $\mathcal{B}_1$  to  $\mathcal{B}_2$  when two distinct productions apply on a same state. Assuming that  $\Delta_1$  is the set of transitions of  $\mathcal{B}_1$ , the switch is defined as follows:

$$\Delta := \{p \xrightarrow{(i,j)} (q,r) : \{p \xrightarrow{i} q, p \xrightarrow{j} r\} \subseteq \Delta_1 \wedge \text{lhs}(\text{prd}(i)) = \text{lhs}(\text{prd}(j)) \wedge \text{prd}(i) \neq \text{prd}(j)\}.$$

The following lemma attests the soundness of this construction.

**Lemma 7.6.** *If  $\mathcal{G}$  is complete, then  $\llbracket \mathcal{A}(\chi_0) \rrbracket = \chi_0$ .*

Observe that,  $\mathcal{B}_1$  is derived from  $\mathcal{A}_j$  in linear time. However  $\mathcal{B}_2$  is quadratic with respect to  $\mathcal{A}_j$ . Therefore,  $\mathcal{A}(\chi_0)$  is built in time  $\mathcal{O}(|\Theta \cup \Omega|^4)$ , because the time of making  $\Delta$  is in  $\mathcal{O}(|\Theta \cup \Omega|^2)$ .

Finally, observe that  $\chi$  is FO definable using relations so far defined by automata:

$$x \chi y \quad :\Leftrightarrow \quad x \chi_0 y \vee \exists x_1 x_2 x_3 y_1 y_2 y_3 \left( (x_1, x_2, x_3) \in \mathcal{J}(\mathcal{G}) \wedge (y_1, y_2, y_3) \in \mathcal{J}(\mathcal{G}) \wedge \left( \bigvee_{i,j \in \{1,3\}} x_i \chi_0 y_j \right) \wedge \right. \\ \left. \left( \left( \bigvee_{i \in [3]} x_i \sqsubseteq x \right) \wedge y_2 \sqsubseteq y \right) \vee \left( \left( \bigvee_{i \in [3]} y_i \sqsubseteq y \right) \wedge x_2 \sqsubseteq x \right) \right),$$

where  $(x, y, z) \in \mathcal{J}(\mathcal{G})$  is a shorthand for  $\bigvee_{a \in \Sigma} a(x, y, z)$ . The corresponding query automaton  $\mathcal{A}(\chi)$  is built from  $\mathcal{A}(\chi_0)$ ,  $\mathcal{A}_j$  and  $\mathcal{A}(\sqsubseteq)$ . Then  $\mathcal{A}(\circ)$  is obtained as the complement of  $\mathcal{A}(\chi)$ .

The overall soundness of the automatic presentation of  $\mathcal{G}^\uparrow$  developed in this section is stated in the following theorem.

**Theorem 7.7.** *For every complete OL graph grammar  $\mathcal{G}$ , the result of every  $\text{FO}(\Upsilon)$  query on  $\mathcal{G}^\uparrow$  is effectively given by a finite automaton.*

Thm. 6.1 is a direct consequence of the above construction. Corollary 6.4 derives from the latter and from Thm. 6.2.

## 8 Undecidability of language FO-checking for 0L graph languages

Unfortunately, in the unrestricted case, most of decision problems for 0L graph languages considered so far turn out to be undecidable. In this section, we construct a 0L graph grammar simulating a deterministic Turing machine. The resulting grammar is edge-deterministic and unix. All undecidability results of this paper follow either directly from the construction or from its variants.

A transition rule  $pA \rightarrow \vec{q}B$  (resp.  $pA \rightarrow \overleftarrow{q}B$ ) of a Turing machine means “in state  $p$  when the symbol read is  $A$ , overwrite it with  $B$ , move the head to the next (resp. previous) cell and change the state from  $p$  into  $q$ ”.

Consider a deterministic Turing machine  $\mathcal{T} = (\mathbb{Q}, \Delta, \iota, f)$  on some alphabet  $\Xi$  and let  $\Xi_{\square} := \Xi \cup \{\square\}$ . To  $\mathcal{T}$ , we associate eD0L graph grammar  $\mathcal{G}_{\mathcal{T}}$  with vertex labels  $\Gamma := \Xi_{\square} \cup \{\triangleright, \triangleleft\} \cup \{pA : p \in \mathbb{Q} \wedge A \in \Xi_{\square}\}$  and unlabelled edges in such a way that

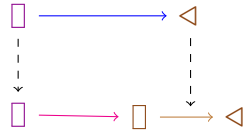
$$\varepsilon \in \mathcal{L}(\mathcal{T}) \iff \exists G \in \mathcal{L}(\mathcal{G}_{\mathcal{T}}) \quad G \models \exists x f(x) . \quad (8.1)$$

We list the elements of  $\Xi_{\square}$  (resp.  $\Xi_{\square, \triangleright} := \Xi_{\square} \cup \{\triangleright\}$ ) as  $\Xi_{\square} = \{C_1, \dots, C_n\}$  (resp.  $\Xi_{\square, \triangleright} = \{D_1, \dots, D_{n+1}\}$ ). Grammar  $\mathcal{G}_{\mathcal{T}}$  has the following vertex productions

$$\begin{aligned} \mathcal{V}_{\mathcal{G}_{\mathcal{T}}} := & \{ \triangleleft \rightarrow (\square \rightarrow \triangleleft) \} \cup \{ C \rightarrow C : C \in \Xi_{\square} \} \cup \{ pA \rightarrow A : pA \rightarrow \_ \in \Delta \} \cup \\ & \{ C \rightarrow qC : C \in \Xi_{\square} \wedge \_ \rightarrow \vec{q} \_ \in \Delta \} \cup \{ A_{\mathcal{G}_{\mathcal{T}}} \rightarrow (\triangleright \rightarrow \iota \square \rightarrow \square \rightarrow \triangleleft) \} \end{aligned}$$

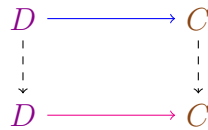
with axiom  $A_{\mathcal{G}_{\mathcal{T}}}$ . Note that  $\mathcal{G}_{\mathcal{T}}$  is not vD0L. Instead of an exhaustive listing of edge productions in  $\mathcal{E}_{\mathcal{G}_{\mathcal{T}}}$ ,  $\mathcal{G}_{\mathcal{T}}$  is given using the graphical formalism introduced in Sect. 4, except that, for better readability, vertex production are explicit and use dashed arrows in this representation. Moreover, every matching edge is drawn in blue and the inherited edge is drawn in magenta.

- The following edge production “extends” the tape of  $\mathcal{T}$  to the right

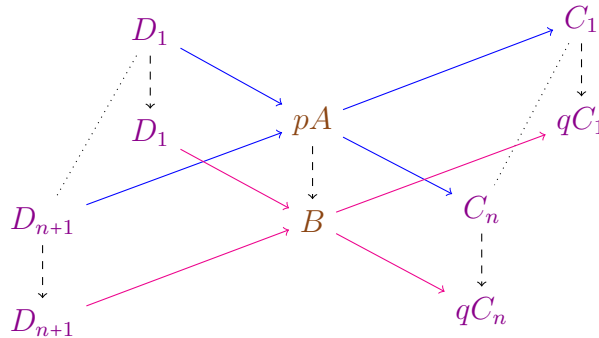


so that the simulated head of  $\mathcal{T}$  cannot reach  $\triangleleft$ .

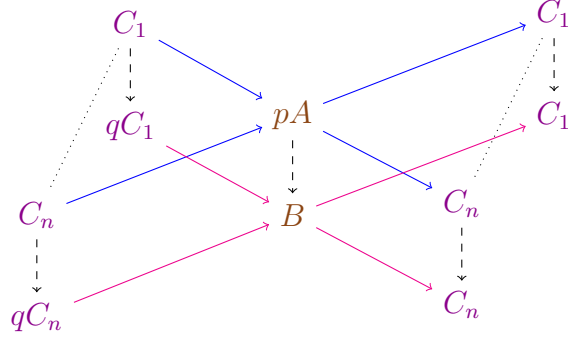
- For all  $D \in \Xi_{\square, \triangleright}$  and  $C \in \Xi_{\square}$ , we have edge productions



- For each rule  $pA \rightarrow \vec{q}B \in \Delta$ , we have edge productions



- For each rule  $pA \rightarrow \overleftarrow{q}B \in \Delta$ , we have edge productions



Observe that every  $pA$  in the left hand side of  $\Delta$  uniquely determines a vertex production of  $V_{\mathcal{G}_{\mathcal{T}}}$  because  $\mathcal{T}$  is deterministic. Then, that vertex production falls into one and only one of the two latter diagrams of edge productions. We get eD0L graph grammar  $\mathcal{G}_{\mathcal{T}}$  where every derived graph is a word with the corresponding successor represented by unlabelled edges. Moreover, each such word represents a reachable configuration of  $\mathcal{T}$  possibly padded with some blanks on the right and ended by  $\triangleleft$ . Indeed, by rewriting the graph given by the axiom, there is always only one vertex with a label of the form  $pA$  and all other vertices have a label in  $\Xi_{\square}$  except the leftmost one labelled  $\triangleright$  and rightmost one labelled  $\triangleleft$ . Therefore, (8.1) holds. Note that  $\mathcal{G}_{\mathcal{T}}$  is not complete. However it is unix. Consequently, we have the following:

**Proposition 8.1.**

- (1) *The model  $\text{FO}(\Gamma \cup \Sigma)$ -checking problem is undecidable for the class of structures  $\{\mathcal{G}^{\dagger} : \mathcal{G} \text{ is eD0L and unix}\}$ .*
- (2) *The language  $\text{FO}(\Gamma \cup \Sigma)$ -checking problem is undecidable for graph languages generated by unix eD0L graph grammars.*

Many other undecidability results derive from this construction. Here is an example.

**Proposition 8.2.** *The following problems*

*Instance:* graph grammar  $\mathcal{G}$ .

*Connectedness:* Are all graphs in  $\mathcal{L}(\mathcal{G})$  (strongly) connected?

*Hamiltonicity:* Do all graphs in  $\mathcal{L}(\mathcal{G})$  have a Hamiltonian path?

*are undecidable for unix eD0L graph grammars.*

*Proof.* We extend the construction of  $\mathcal{G}_{\mathcal{T}}$  by adding erasing productions for vertex labels involving the final state of  $\mathcal{T}$ . Then both connectedness and Hamiltonicity are equivalent to  $\varepsilon \notin \mathcal{L}(\mathcal{T})$ .  $\square$

A slightly different consequence of the construction of  $\mathcal{G}_{\mathcal{T}}$  is stated in the following.

**Proposition 8.3.** *The language  $\text{FO}(\Gamma \cup \Sigma)$ -checking problem is undecidable for graph languages generated by complete eD0L graph grammars*

*Proof.* Let  $\mathcal{G}_{\mathcal{T}}^{\dagger}$  be the completion of  $\mathcal{G}_{\mathcal{T}}$  stipulated by Prop. 4.2. Observe that translation  $\tau_{\dagger}$  from the proof of Corollary 6.4, allows reducing the question “ $\varepsilon \in \mathcal{L}(\mathcal{T})$ ?” via the construction of  $\mathcal{G}_{\mathcal{T}}$  to the question whether there exists  $G \in \mathcal{L}(\mathcal{G}_{\mathcal{T}}^{\dagger})_{\{\dagger, !\}}$  such that  $G \models \tau_{\dagger}(\exists x f(x))$ .  $\square$

From this proposition and Thm. 6.1, we conclude that the decidability of  $\text{FO}(\Gamma \cup \Sigma)$ -theory of complete 0L grammars is not sufficient for the decidability of the  $\text{FO}(\Gamma \cup \Sigma)$ -language checking problem for complete eD0L graph grammars. In the presence of vertex nondeterminism, it is not possible in general to “separate” graphs of the same layer sharing some vertices.

## 9 Conclusion

Although this work only addresses first-order logic, its application potential beyond mathematics and theoretical computer science would benefit from extending our approach to quantitative logics. Such extensions also seem desirable for automatic structures. To our knowledge, examples of application of automatic structures outside of mathematics are still awaited. We hope that the connection of this field of scientific achievements with expanders and Lindenmayer graph grammars proposed in this paper can stimulate even more relevant theoretical developments that would enable such applications.

## References

- [1] Noga Alon and V. D. Milman. “ $\lambda_1$ , Isoperimetric inequalities for graphs, and superconcentrators”. In: *J. Comb. Theory, Ser. B* 38.1 (1985), pp. 73–88.
- [2] Yanis M. Barzdin. “On the Realization of Networks in Three-Dimensional Space”. In: *Selected Works of A. N. Kolmogorov: Volume III: Information Theory and the Theory of Algorithms*. Ed. by A. N. Shiriyayev. Springer Netherlands, 1993, pp. 194–202. ISBN: 978-94-017-2973-4.
- [3] Brent R. Bielefeldt et al. “Development and validation of a genetic L-System programming framework for topology optimization of multifunctional structures”. In: *Computers & Structures* 218 (2019), pp. 152–169.
- [4] Achim Blumensath and Erich Grädel. “Finite Presentations of Infinite Structures: Automata and Interpretations”. In: *Theory of Computing Systems* 37.6 (2004), pp. 641–674.
- [5] Karthekeyan Chandrasekaran and Ameya Velingker. “Shift lifts preserving Ramanujan property”. In: *Linear Algebra and its Applications* 529 (2017), pp. 199–214.
- [6] Michael Benjamin Cohen. “Ramanujan Graphs in Polynomial Time”. In: *FOCS 2016*. Ed. by Irit Dinur. 2016, pp. 276–281.
- [7] Bruno Courcelle and Joost Engelfriet. *Graph structure and Monadic Second-Order Logic, a Language Theoretic Approach*. Cambridge University Press, 2012.
- [8] Irit Dinur. “The PCP theorem by gap amplification”. In: *J. ACM* 54.3 (2007), p. 12.
- [9] Hartmut Ehrig and Hans-Jörg Kreowski. “Parallelism of manipulations in multidimensional information structures”. In: *Mathematical Foundations of Computer Science 1976*. Ed. by Antoni Mazurkiewicz. LNCS 45. 1976, pp. 284–293.
- [10] Hartmut Ehrig et al., eds. *Handbook of graph grammars and computing by graph transformation: vol. 3: concurrency, parallelism, and distribution*. World Scientific Publishing, 1999. ISBN: 981024021X.
- [11] Alexandre Fernandez, Luidnel Maignan, and Antoine Spicher. “Non-Determinism in Lindenmayer Systems and Global Transformations”. In: *MFCS 2022*. Vol. 241. LIPIcs, 49:1–49:13.
- [12] John P. Gallagher, Mai Ajspur, and Bishoksan Kafle. “Optimised determinisation and completion of finite tree automata”. In: *J. Log. Algebraic Methods Program.* 95 (2018), pp. 1–16.
- [13] Erich Grädel. “Automatic Structures: Twenty Years Later”. In: *LICS '20*. Ed. by Holger Hermanns et al. 2020, pp. 21–34.

- [14] Erich Grädel and Yuri Gurevich. “Metafinite Model Theory”. In: *Inf. Comput.* 140.1 (1998), pp. 26–81.
- [15] Erich Grädel et al. *Finite Model Theory and Its Applications*. Springer, 2007.
- [16] Bernard R. Hodgson. “Décidabilité par automate fini”. In: *Annales des sciences mathématiques du Québec* 7.3 (1983), pp. 39–57.
- [17] Bernard R. Hodgson. “On direct automaton product decidable theories”. In: *Theoretical Computer Science* 19 (1982), pp. 331–335.
- [18] Bernard R. Hodgson. “Théories décidables par automate fini”. PhD thesis. Université de Montreal, 1976.
- [19] Shlomo Hoory, Nathan Linial, and Avi Wigderson. “Expander graphs and their applications”. In: *Bulletin of the AMS* 43.4 (2001).
- [20] Karel Culik II and Aristid Lindenmayer. “Parallel graph generating and graph recurrence systems for multicellular development”. In: *International Journal of General Systems* 3.1 (1976), pp. 53–66.
- [21] Dirk Janssens, Grzegorz Rozenberg, and Raymond Verraedt. “On Sequential and Parallel Node-Rewriting Graph Grammars”. In: *Computer Graphics and Image Processing* 18 (1982), pp. 279–304.
- [22] Paris C. Kanellakis. “Elements of Relational Database Theory”. In: *Formal Models and Semantics*. Ed. by Jan van Leeuwen. Vol. B. Handbook of Theoretical Computer Science. Elsevier, 1990, pp. 1073–1156.
- [23] Bakhadyr Khoussainov and Anil Nerode. “Automatic Presentations of Structures”. In: *Logic and Computational Complexity*. Ed. by Daniel Leivant. LNCS 960. 1994, pp. 367–392.
- [24] Leonid Libkin. *Elements of Finite Model Theory*. Springer, 2004.
- [25] Aristid Lindenmayer. “Mathematical models for cellular interaction in development”. In: *Journal of Theoretical Biology* 18 (1968), pp. 300–315.
- [26] Alexander Lubotzky. “Expander graphs in pure and applied mathematics”. In: *Bulletin of the AMS* 49.1 (2012).
- [27] Jacqueline Lück, Peter W. Barlow, and Hermann B. Lück. “Deterministic Patterns of Cellular Growth and Division Within a Meristem”. In: *Annals of Biology* 73.1 (1994), pp. 1–11.
- [28] Sebastian von Mammen et al. “A Graph-Based Developmental Swarm Representation and Algorithm”. In: *Swarm Intelligence - 7th International Conference, ANTS*. Vol. 6234. LNCS. 2010, pp. 1–12.
- [29] Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava. “Interlacing families I: Bipartite Ramanujan graphs of all degrees”. In: *Annals of Mathematics* 182.1 (2015), pp. 307–325.
- [30] Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava. “Interlacing families IV: Bipartite Ramanujan graphs of all sizes”. In: *SIAM Journal on computing* 47.6 (2018), pp. 2488–2509.
- [31] Yongzhi Ong and Winfried Kurth. “A graph model and grammar for multi-scale modelling using XL”. In: *2012 IEEE International Conference on Bioinformatics and Biomedicine Workshops*. 2012, pp. 1–8.



- [32] M. Pinsky. “On the complexity of a concentrator”. In: *Proceeding of the 7th International Teletraffic Congress*. Stockholm, June 1973, pp. 318/1–318/4.
- [33] Przemysław Prusinkiewicz and Aristid Lindenmayer. *The Algorithmic Beauty of Plants*. Przemysław Prusinkiewicz, 2004. URL: <http://algorithmicbotany.org/papers/abop/abop.pdf>.
- [34] Omer Reingold, Salil P. Vadhan, and Avi Wigderson. “Entropy Waves, the Zig-Zag Graph Product, and New Constant-Degree Expanders and Extractors”. In: *Electron. Colloquium Comput. Complex.* TR01-018 (2001).
- [35] Grzegorz Rozenberg and Arto Salomaa, eds. *Lindenmayer Systems. Impacts on Theoretical Computer Science, Computer Graphics, and Developmental Biology*. Springer-Verlag, 1992. ISBN: 978-3-642-58117-5.
- [36] Thierry Boy de la Tour. “Subsumptions of Algebraic Rewrite Rules”. In: *EPTCS* 397 (2023), pp. 20–38.
- [37] Thierry Boy de la Tour and Rachid Echahed. “Combining Parallel Graph Rewriting and Quotient Graphs”. In: *Rewriting Logic and Its Applications - 13th International Workshop*. Vol. 12328. LNCS. Revised selected papers. 2020, pp. 1–18.

## Appendices

The statements without proof in the main part appear with their original numbers in the following appendices.

### A Fundamental facts on complete and incomplete 0L graph grammars

In this appendix, the reader will find complete proofs of statements of Sect. 4.

**Lemma 4.1.** *Let  $\mathcal{G} = (\mathbf{A}_{\mathcal{G}}, \mathbf{V}_{\mathcal{G}}, \mathbf{E}_{\mathcal{G}})$  be a complete graph grammar in  $0L(\Gamma, \Sigma, \_ , \_)$  and let  $G = (V, E, \ell, \mathfrak{L})$  be a graph in  $\mathcal{L}(\mathcal{G})$ . Every labelling preserving map  $h: V \rightarrow \mathbf{V}_{\mathcal{G}}$  extends into a homomorphism from  $G$  into  $\mathcal{G}$ .*

*Proof.* Let  $h: V \rightarrow \mathbf{V}_{\mathcal{G}}$  be such that  $h(v) = \ell(v) \rightarrow \_$  for every  $v \in V$ . In order to extend  $h$  into a homomorphism from  $G$  to  $\mathcal{G}$ , consider  $w \in E$  and the corresponding  $a(u, w, v) \in \mathfrak{L}$ . Set  $h(w) := P$  for some  $P = h(u) \xrightarrow{a} h(v) \rightarrow \_$ . Observe that  $P \in \mathbf{E}_{\mathcal{G}}$  because  $\mathcal{G}$  is complete. Clearly,  $h$  so extended onto edges is a homomorphism from  $G$  to  $\mathcal{G}$ .  $\square$

**Proposition 4.2.** *For every 0L graph grammar  $\mathcal{G} = (\mathbf{V}_{\mathcal{G}}, \_ , \_)$  one can construct in time  $\mathcal{O}(|\mathbf{V}_{\mathcal{G}}|^2)$  using additional non-terminal vertex label “j” and non-terminal edge label “!” complete 0L graph grammar  $\mathcal{G}^!$  such that  $\mathcal{L}(\mathcal{G}) = \mathcal{L}(\mathcal{G}^!)_{\{i, !\}}$ . If  $\mathcal{G}$  is vD0L or eD0L, then  $\mathcal{G}^!$  is so.*

*Proof.* Let  $\mathcal{G} = (\mathbf{V}_{\mathcal{G}}, \mathbf{E}_{\mathcal{G}}, \mathbf{A}_{\mathcal{G}})$  be a 0L graph grammar in  $0L(\Gamma, \Sigma, \Theta, \Omega)$ . Consider bijection  $g_1: \Theta' \rightarrow \Gamma \cup \{j\}$  where  $\Theta'$  is a new set of vertex names  $\Theta' \cap \Theta = \emptyset$ ,  $j \notin \Gamma$  is a new vertex label and  $\Gamma'$  consists of labels missing in  $\text{lhs}(\mathbf{V}_{\mathcal{G}})$  but labelling isolated vertices in  $\text{rhs}(\mathbf{V}_{\mathcal{G}})$ . Let then  $\mathcal{G}^! := (\mathbf{V}_{\mathcal{G}} \cup \mathbf{V}_i, \mathbf{E}_{\mathcal{G}}, \mathbf{A}_{\mathcal{G}})$  where  $\mathbf{V}_i := \{g_1(i) \rightarrow (\{i\}, \emptyset, \{i \mapsto j\}, \emptyset) : i \in \Theta'\}$ . Every right hand side of  $\mathbf{V}_i$  is a graph with a single j-labelled vertex with no loop. Observe that a possible incompleteness

of  $\mathcal{G}^i$  can only be due to edge productions but no more to vertex productions. Observe also that if  $\mathcal{G}$  is vD0L or eD0L, then  $\mathcal{G}^i$  is so.

Next, we transform  $\mathcal{G}^i$  into the target grammar  $\mathcal{G}^{\dagger}$  by completing  $E_{\mathcal{G}}$ . Let  $M$  be the set of all triples in  $V_{\mathcal{G}} \times \Sigma \times V_{\mathcal{G}}$  that are not among left hand sides of  $E_{\mathcal{G}}$  but which could be matched by incidences appearing in right hand sides of productions of  $\mathcal{G}$ . In symbols

$$M := \{P_1 \xrightarrow{a} P_2 \in V_{\mathcal{G}} \times \Sigma \times V_{\mathcal{G}} \setminus \text{lhs}(E_{\mathcal{G}}) : \exists a(i, k, j) \in \mathfrak{C}_{\mathcal{G}} \ell_{\mathcal{G}}(i) = \text{lhs}(P_1) \wedge \ell_{\mathcal{G}}(j) = (\text{lhs}(P_2))\}$$

where  $\mathfrak{C}_{\mathcal{G}} := \bigcup \pi_4(\text{rhs}(V_{\mathcal{G}})) \cup \bigcup \text{rhs}(E_{\mathcal{G}})$  and  $\ell_{\mathcal{G}} := \bigcup \pi_3(\text{rhs}(V_{\mathcal{G}}))$ . Consider bijection  $g_2: \Omega' \rightarrow M$  for some set of new edge names  $\Omega'$  such that  $\Omega' \cap (\Theta \cup \Theta' \cup \Omega) = \emptyset$  and injective mapping  $f: V_{\mathcal{G}} \rightarrow \Theta$  such that  $f(A \rightarrow (V, \_, \_, \_)) \in V$  for every  $A \rightarrow (V, \_, \_, \_) \in V_{\mathcal{G}}$ . Set

$$E := \{P_1 \xrightarrow{a} P_2 \rightarrow \{!(f(P_1), i, f(P_2))\} : i \in \Omega' \wedge g_2(i) = P_1 \xrightarrow{a} P_2\} .$$

As every triple in  $M$  witnesses the incompleteness of  $\mathcal{G}^i$ ,  $E$  form a partial completion of  $\mathcal{G}^i$ . Indeed, all incidences of  $\mathfrak{C}_{\mathcal{G}}$  have a match in  $E_{\mathcal{G}} \cup E$ . However, adding  $E$  alone to  $E_{\mathcal{G}}$  is not enough because now we have new incidences with label  $!$  that have no matching edge productions in  $E_{\mathcal{G}} \cup E$ . Consider therefore bijection  $g_2^!: \Omega^! \rightarrow V_{\mathcal{G}} \times V_{\mathcal{G}}$  for some yet newer set of edge names  $\Omega^!$  such that  $\Omega^! \cap (\Theta \cup \Theta' \cup \Omega \cup \Omega') = \emptyset$  and set

$$E^! := \{P_1 \xrightarrow{!} P_2 \rightarrow \{!(f(P_1), i, f(P_2))\} : i \in \Omega^! \wedge g_2^!(i) = (P_1, P_2)\} .$$

Let then

$$\mathcal{G}^{\dagger} := (V_{\mathcal{G}} \cup V^i, E_{\mathcal{G}} \cup E \cup E^!, A_{\mathcal{G}}) .$$

Observe that  $\mathcal{G}^{\dagger} \in \text{OL}(\Gamma \cup \{i\}, \Sigma \cup \{!\}, \Theta \cup \Theta', \Omega \cup \Omega' \cup \Omega^!)$  is complete and that  $\mathcal{L}(\mathcal{G}) = \mathcal{L}(\mathcal{G}^{\dagger})_{\{i, !\}}$ . Note also that  $\mathcal{G}^{\dagger}$  has the same vertex production as  $\mathcal{G}^i$ . Thus, if  $\mathcal{G}^i$  is vD0L,  $\mathcal{G}^{\dagger}$  is so. To see that the transformation from  $\mathcal{G}^i$  to  $\mathcal{G}^{\dagger}$  preserves the property of being eD0L, observe first that  $\text{lhs}(E \cup E^!) \cap \text{lhs}(E_{\mathcal{G}}) = \emptyset$ . Obviously also  $\text{lhs}(E) \cap \text{lhs}(E^!) = \emptyset$ . Moreover, two distinct productions of  $E \cup E^!$  never share their left hand sides as both  $g_2$  and  $g_2^!$  are bijections. Consequently,  $\mathcal{G}^{\dagger}$  is eD0L when  $\mathcal{G}_1$  is so. Hence, the transformation from  $\mathcal{G}$  is both vD0L and eD0L preserving.  $\square$

**Theorem 4.9.** *The following problems*

Instance: 0L graph grammar  $\mathcal{G}$ , label  $\lambda$  of  $\mathcal{G}$  and production  $P$  of  $\mathcal{G}$ .

Finiteness: Is  $\mathcal{L}(\mathcal{G})$  finite?

Emptiness:  $\mathcal{L}(\mathcal{G}) = \emptyset$ ?

Accessibility: Is  $\lambda$  accessible in  $\mathcal{G}$ ?

Usefulness: Is  $P$  useful for  $\mathcal{G}$ ?

are

- (i) decidable for complete 0L graph grammars without non-terminals and for D0L graph grammars even with non-terminals,
- (ii) undecidable for unix eD0L graph grammars, with non-terminals in the case of emptiness.

*Proof.*

- (i) Except for the emptiness, the decidability of each problem results from Thm. 6.1 by reduction to the model checking of  $\mathcal{G}^{\dagger}$  or  $\mathcal{G}^{\dagger \setminus N}$ .

**Accessibility** For complete 0L graph grammar  $\mathcal{G}$  and  $\lambda \in \Gamma$  (resp.  $\lambda \in \Sigma$ ) we check if  $\mathcal{G}^{\dagger}$  models  $\exists x \lambda(x)$  (resp.  $\exists x y z \lambda(x, y, z)$ ). For D0L graph grammar  $\mathcal{G}$  with non-terminals  $N$  we similarly check  $\mathcal{G}^{\dagger \setminus N}$ .

**Usefulness** We modify  $P$  by introducing an extra label to  $\text{rhs}(P)$  and ask if the new label is accessible.

**Finiteness** Depending on whether  $\mathcal{G}$  is complete or deterministic with non-terminals  $N$ , we check  $\mathcal{G}^{\dagger}$  or  $\mathcal{G}^{\dagger \setminus N}$  against the negation of the following formula

$$\forall x \exists y (x \hat{=} y \Rightarrow \exists z (y \neq z \wedge y \sqsubseteq z)) .$$

The formula says that every layer has an element which is a strict prefix of another element. Hence, the formula requires an infinite number of layers to hold. This

is equivalent to having  $\mathcal{L}(\mathcal{G})$  (resp.  $\mathcal{L}(\mathcal{G})_N$ ) infinite when  $\mathcal{G}$  is without non-terminals (resp. with non-terminals  $N$ ).

**Emptiness** For 0L graph grammar  $\mathcal{G} = (V_{\mathcal{G}}, E_{\mathcal{G}}, A_{\mathcal{G}})$  without non-terminals, we have  $\mathcal{L}(\mathcal{G}) = \emptyset$  if and only if  $A_{\mathcal{G}} \notin \text{lhs}(V_{\mathcal{G}})$ . If  $\mathcal{G}$  is D0L with non-terminals  $N$ , the emptiness is equivalent to language checking  $\mathcal{L}(\mathcal{G})_N$  against a tautology, for instance  $\forall x x=x$ . The decidability follows from Corollary 6.4.

(ii) The undecidability of each problem results from the construction, given in Sect. 8, of unix eD0L grammar  $\mathcal{G}_{\mathcal{T}}$  associated to arbitrary deterministic Turing machine  $\mathcal{T}$  running on an empty tape, viz., with input word  $\varepsilon$ .

**Accessibility** We ask, for each vertex label  $fC$  formed with accepting state  $f$  of  $\mathcal{T}$ , if it is accessible in  $\mathcal{G}_{\mathcal{T}}$ . The disjunction of all answers is equivalent to  $\varepsilon \in \mathcal{L}(\mathcal{T})$ .

**Usefulness** We ask, for each vertex production introducing  $f$ , if it is useful. The disjunction of all answers is equivalent to  $\varepsilon \in \mathcal{L}(\mathcal{T})$ .

**Finiteness** As  $\mathcal{G}_{\mathcal{T}}$  has a unique derivation, the halting of  $\mathcal{T}$  is equivalent to the finiteness of  $\mathcal{L}(\mathcal{G}_{\mathcal{T}})$ .

**Emptiness** This problem is trivially decidable for 0L graph grammars without non-terminals (see the decidability of the emptiness above). To establish the undecidability of the emptiness for unix eD0L graph grammars with non-terminals, the construction of Sect. 8 is adapted as follows. We consider all vertex and edge labels of  $\mathcal{G}_{\mathcal{T}}$  as non-terminals, say  $N$ , and we add one terminal vertex label  $\blacklozenge$  together with vertex productions  $\{fC \rightarrow \blacklozenge : C \in \Xi_{\square}\} \cup \{C \rightarrow \blacklozenge : C \in \Xi_{\triangleright, \square, \triangleleft}\}$ . Moreover

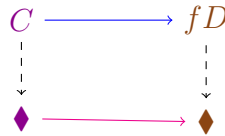
- for all  $D \in \Xi_{\square, \triangleright}$  and  $C \in \Xi_{\square, \triangleleft}$ , we add edge productions



- for all  $C \in \Xi_{\square}$  and  $D \in \Xi_{\square, \triangleleft}$ , we add edge productions



- and for all  $C \in \Xi_{\square, \triangleright}$  and  $D \in \Xi_{\square}$ , we add edge productions

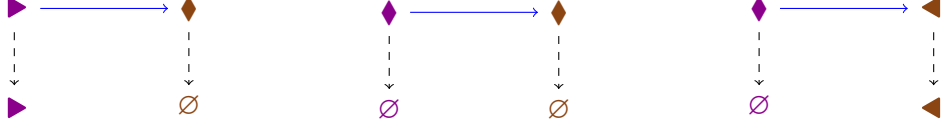


Let  $\mathcal{G}'_{\mathcal{T}}$  be the grammar so extended. Observe that (see Sect. 8) reaching  $f$  by  $\mathcal{T}$  is reflected on  $\mathcal{G}_{\mathcal{T}}$  by the unique derivation ending with a graph representing an accepting configuration with exactly one vertex labelled  $fC$  for some  $C \in \Xi_{\square}$ . For  $\mathcal{G}'_{\mathcal{T}}$ , the presence of such vertex triggers the simultaneous application of new productions. The resulting graph  $\blacklozenge \rightarrow \dots \rightarrow \blacklozenge$  has the terminal label everywhere and is the unique graph in  $\mathcal{L}(\mathcal{G}'_{\mathcal{T}})_N$ . If, however,  $\mathcal{T}$  cannot reach an accepting configuration, then  $\mathcal{L}(\mathcal{G}'_{\mathcal{T}})_N = \emptyset$ . □

**Proposition 4.10.** *The (abstract) membership problem is decidable for D0L graph grammars even with non-terminals. It is undecidable for unix eD0L graph grammars.*

*Proof.* For the undecidability claim we modify the adaptation of the construction of Sect. 8 considered in the proof of Thm. 4.9 for establishing the undecidability of the emptiness problem.

The starting step of the present modification is that all labels have now the same status without distinguishing non-terminals. We take two additional vertex labels  $\blacktriangleright$  and  $\blacktriangleleft$  and replace vertex productions  $\triangleright \rightarrow \blacklozenge$  and  $\triangleleft \rightarrow \blacklozenge$  by  $\triangleright \rightarrow \blacktriangleright$  and  $\triangleleft \rightarrow \blacktriangleleft$  in the above adaptation. Edge productions are modified accordingly, so that we can derive graph  $\blacktriangleright \rightarrow \blacklozenge \rightarrow \dots \rightarrow \blacklozenge \rightarrow \blacktriangleleft$  (instead of  $\blacklozenge \rightarrow \dots \rightarrow \blacklozenge$  of the above adaptation), if, and only if,  $\varepsilon \in \mathcal{L}(\mathcal{T})$ . We further extend the construction by adding two vertex productions  $\blacktriangleright \rightarrow \blacktriangleright$ ,  $\blacktriangleleft \rightarrow \blacktriangleleft$ , one erasing vertex production  $\blacklozenge \rightarrow \emptyset$  and three erasing edge productions



These let derive in one step from  $\blacktriangleright \rightarrow \blacklozenge \rightarrow \dots \rightarrow \blacklozenge \rightarrow \blacktriangleleft$  the graph which consists of two isolated vertices only:  $\blacktriangleright \blacktriangleleft$ . The membership of this graph in the language of the modified graph grammar, which still remains unix and eD0L, is equivalent to the membership of  $\varepsilon$  in  $\mathcal{L}(\mathcal{T})$ .

The decidability claim follows from the fact that, for every finite graph  $H$ , we can write a sentence  $\varphi_H$  such that  $\{G : G \models \varphi_H\}$  is the isomorphism class of  $H$ . Consequently,  $H$  belongs up to isomorphism to  $\mathcal{L}(\mathcal{G})_N$ , if language checking of  $\mathcal{L}(\mathcal{G})_N$  against  $\varphi_H$  answers yes. This is decidable by Corollary 6.4.  $\square$

## B When graphs' elements form maximal compatible subsets of $\mathcal{V}(\mathcal{G}) \cup \mathcal{E}(\mathcal{G})$

This appendix is devoted to the proof of the the following lemma.

**Lemma 4.11.** *If  $\mathcal{G}$  is a complete 0L graph grammar, then*

$$\{S \subseteq \mathcal{V}_l(\mathcal{G}) \cup \mathcal{E}_l(\mathcal{G}) : S \text{ is maximal compatible}\} = \{V \cup E : (V, E, \_, \_) \in \mathcal{L}_l(\mathcal{G})\}.$$

With the help of subsequent lemmas, the proof goes as follows.

*Proof.* Let  $\mathcal{G}$  be a complete 0L graph grammar.

( $\subseteq$ ) Let  $S \subseteq \mathcal{V}_l(\mathcal{G}) \cup \mathcal{E}_l(\mathcal{G})$  be a maximal compatible set. By Lemma B.2, there exists  $(V, E, \_, \_) \in \mathcal{L}_l(\mathcal{G})$  such that  $S \subseteq V \cup E$ . By definition of derivation,  $(V \cup E)^2 \subseteq \circ$ . In other words, it is a compatible set, say  $S' = V \cup E$ . But as  $S$  is maximal compatible, whereas  $S \subseteq S'$ , we have  $S = S' = V \cup E$ .

( $\supseteq$ ) Let  $(V, E, \ell, \mathfrak{I}) \in \mathcal{L}_l(\mathcal{G})$ . As observed above,  $V \cup E$  is a compatible set. To see that it is maximal compatible, consider a compatible set  $S \subseteq \mathcal{V}_l(\mathcal{G}) \cup \mathcal{E}_l(\mathcal{G})$  such that  $V \cup E \subseteq S$ . By Lemma B.2, there exists  $(V', E', \ell', \mathfrak{I}') \in \mathcal{L}_l(\mathcal{G})$  such that  $S \subseteq V' \cup E'$ . Consequently  $V \cup E \subseteq V' \cup E'$ . As  $\Theta \cap \Omega = \emptyset$ , the latter inclusion splits into  $V \subseteq V'$  and  $E \subseteq E'$ . Observe also that for a vertex (resp. edge), say  $si$  in  $\mathcal{V}_l(\mathcal{G})$  (resp. in  $\mathcal{E}_l(\mathcal{G})$ ), where  $i \in \Theta$  (resp.  $i \in \Omega$ ), its label is uniquely determined within the corresponding production  $\text{prd}(i)$ . Moreover, when  $si$  is an edge,  $\text{prd}(i)$  also uniquely determines the incidence of  $si$ . Consequently, from inclusions  $V \subseteq V'$  and  $E \subseteq E'$ , we obtain  $\ell \subseteq \ell'$  and  $\mathfrak{I} \subseteq \mathfrak{I}'$ . As  $(V, E, \ell, \mathfrak{I}) \subseteq (V', E', \ell', \mathfrak{I}')$ , by Lemma B.1,  $(V, E, \ell, \mathfrak{I}) = (V', E', \ell', \mathfrak{I}')$ . Hence,  $S \subseteq (V, E, \_, \_)$  and we conclude that  $V \cup E$  is maximal compatible.  $\square$

**Lemma B.1.** *For every 0L graph grammar  $\mathcal{G}$  and all  $G, G' \in \mathcal{L}_l(\mathcal{G})$ , if  $G \subseteq G'$  then  $G = G'$ .*

*Proof.* By induction on  $l \in \mathbb{N}_{>0}$ . Let  $\mathcal{G} = (\mathbf{A}_{\mathcal{G}}, \mathbf{V}_{\mathcal{G}}, \mathbf{E}_{\mathcal{G}})$  be a 0L graph grammar and let  $G_1 = (V_1, E_1, \_, \_) \in \mathcal{L}_l(\mathcal{G})$  and  $G_2 = (V_2, E_2, \_, \_) \in \mathcal{L}_l(\mathcal{G})$  be such that  $V_1 \subseteq V_2$  and  $E_1 \subseteq E_2$ .

( $\mathbf{l} = \mathbf{0}$ ) Assume by contradiction that  $G_1 \neq G_2$ . Then  $\mathcal{G}$  has productions  $A_{\mathcal{G}} \rightarrow G_1$  and  $A_{\mathcal{G}} \rightarrow G_2$  but the definition of 0L graph grammar requires that  $(V_1 \cup E_1) \cap (V_2 \cup E_2) = \emptyset$  which contradicts  $V_1 \subseteq V_2$  and  $E_1 \subseteq E_2$ .

( $\mathbf{l} > \mathbf{0}$ ) By definition of derivation, there exist  $G'_1, G'_2 \in \mathcal{L}_{l-1}(\mathcal{G})$  and homomorphisms  $h_1: G'_1 \rightarrow \mathcal{G}$ ,  $h_2: G'_2 \rightarrow \mathcal{G}$  such that  $G'_1 \xrightarrow{h_1} G_1$  and  $G'_2 \xrightarrow{h_2} G_2$ . As  $G_1 \subseteq G_2$ , also  $G'_1 \subseteq G'_2$ . By induction hypothesis  $G'_1 = (V', E', \_, \_ ) = G'_2$  for some  $V' \subseteq \mathcal{V}_{l-1}(\mathcal{G})$  and  $E' \subseteq \mathcal{E}_{l-1}(\mathcal{G})$ . To show that  $h_1 = h_2$ , assume by contradiction that there exists  $w \in V' \cup E'$  such that  $h_1(w) = P_1 \neq P_2 = h_2(w)$  for some productions  $P_1, P_2 \in \mathbf{V}_{\mathcal{G}} \cup \mathbf{E}_{\mathcal{G}}$ . By definition of 0L graph grammar, the set of vertex or edge names appearing in  $\text{rhs}(P_1)$  is disjoint from that appearing in  $\text{rhs}(P_2)$ . Then also  $\{s \in V_1 \cup E_1 : w \xrightarrow{P_1} s\} \cap \{s \in V_2 \cup E_2 : w \xrightarrow{P_2} s\} = \emptyset$  which contradicts  $G_1 \subseteq G_2$ .  $\square$

**Lemma B.2.** *Let  $\mathcal{G} \in \text{0L}(\Gamma, \Sigma, \Theta, \Omega)$  be complete and  $S \subseteq \mathcal{V}_l(\mathcal{G}) \cup \mathcal{E}_l(\mathcal{G})$  be such that  $S^2 \subseteq \circ$ . Then  $S \subseteq V \cup E$  for some  $(V, E, \_, \_) \in \mathcal{L}_l(\mathcal{G})$ .*

*Proof.* By induction on product order  $(\mathbb{N}, \leq)^2$  over pairs  $(l, m)$  with  $m = |S|$ . Let  $\mathcal{G} = (\mathbf{V}_{\mathcal{G}}, \mathbf{E}_{\mathcal{G}}, \mathbf{A}_{\mathcal{G}})$ .

( $\mathbf{l} = \mathbf{1}$ ) Then  $S \subseteq V \cup E$  for some production  $A_{\mathcal{G}} \rightarrow (V, E, \_, \_) \in \mathbf{V}_{\mathcal{G}}$ .

( $\mathbf{m} = \mathbf{0}$ ) Trivial case.

( $\mathbf{l} > \mathbf{1}$ )  $S$  can be written as  $S = \{s_j i_j : j \in [m]\}$  with  $S' = \{s_j : j \in [m]\} \subseteq \mathcal{V}_{l-1}(\mathcal{G}) \cup \mathcal{E}_{l-1}(\mathcal{G})$  and  $\{i_j : j \in [m]\} \subseteq \Theta \cup \Omega$ . Let  $si \in S$  with  $s \in S'$  and let  $S_1 := S \setminus \{si\}$ . As  $si \in \mathcal{V}_l(\mathcal{G}) \cup \mathcal{E}_l(\mathcal{G})$ , there exists a graph, say  $G_2 = (V_2, E_2, \_, \_) \in \mathcal{L}_l(\mathcal{G})$ , such that  $si \in V_2 \cup E_2$ . By definition of  $\circ$  we have  $S_1^2 \subseteq \circ$  and as  $|S_1| < m$ , by induction hypothesis, there exists  $G_1 = (V_1, E_1, \_, \mathfrak{I}_1) \in \mathcal{L}_l(\mathcal{G})$  such that  $S_1 \subseteq V_1 \cup E_1$ . Observe that  $(S')^2 \subseteq \circ$ , by definition of  $\circ$ . As also  $S' \subseteq \mathcal{V}_{l-1}(\mathcal{G}) \cup \mathcal{E}_{l-1}(\mathcal{G})$ , by induction hypothesis, there exists  $G' = (V', E', \_, \mathfrak{I}') \in \mathcal{L}_{l-1}(\mathcal{G})$  such that  $S' \subseteq V' \cup E'$ . By definition of derivation, there exist homomorphisms  $h_1, h_2: G' \rightarrow \mathcal{G}$  such that  $G' \xrightarrow{h_1} G_1$  and  $G' \xrightarrow{h_2} G_2$ . If  $h_2(s) = h_1(s)$ , then  $si \in V_1 \cup E_1$  and we are done because  $S \subseteq V_1 \cup E_1$ . Assume therefore that  $h_2(s) \neq h_1(s)$ . For any  $t \in V'$ , we denote by  $E'_t$  the set of edges of  $E'$  incident to  $t$ :

$$E'_t := \{w \in E' : \{ \_ (t, w, \_ ), \_ (\_ , w, t) \} \cap \mathfrak{I}' \neq \emptyset \} .$$

Using this notation, we construct homomorphism  $h: G' \rightarrow \mathcal{G}$  setting  $h(s) := h_2(s)$  and such that  $G' \xrightarrow{h} G$  for some  $G = (V, E, \_, \_) \in \mathcal{L}_l(\mathcal{G})$  satisfying  $S \subseteq V \cup E$ . The construction has two cases.

( $\mathbf{s} \in \mathbf{V}'$ ) Observe that  $E'_s \cap S' = \emptyset$ . If instead  $s_k \in E'_s$  for some  $k \in [m]$ , then we would have  $\{ \_ (s, s_k, \_ ), \_ (\_ , s_k, s) \} \cap \mathfrak{I}' \neq \emptyset$  and  $\{ \_ (si', s_k i_k, \_ ), \_ (\_ , s_k i_k, si') \} \cap \mathfrak{I}_1 \neq \emptyset$  for some vertex name  $i'$  such that  $\text{prd}(i') = h_1(s)$ . We would then conclude that  $si' \times_0 si$  because  $h_1(s) \neq h_2(s)$  whereas  $\text{lhs}(h_1(s)) = \text{lhs}(h_2(s))$ , and consequently we would get  $si' \times s_k i_k$  since  $si'$  is an extremity of  $s_k i_k$  in  $G_1$ . As both  $si$  and  $s_k i_k$  are in  $S$ , the latter would contradict  $S^2 \subseteq \circ$ . Now, as  $E'_s \cap S' = \emptyset$ , for every loop  $w' \in E'_s$  with corresponding incidence  $a(s, w', s)$ , we pick  $P' \in \mathbf{E}_{\mathcal{G}}$  such that  $\text{lhs}(P') = h_2(s) \xrightarrow{a} h_2(s)$  and we set  $h(w') := P'$ . For every edge  $w \in E'_s$ , with corresponding incidence  $a(s, w, v) \in \mathfrak{I}'$  (resp.  $a(v, w, s) \in \mathfrak{I}'$ ) and  $s \neq v$ , we pick  $P \in \mathbf{E}_{\mathcal{G}}$  such that  $\text{lhs}(P) = h_2(s) \xrightarrow{a} h_1(v)$  (resp.  $\text{lhs}(P) = h_1(v) \xrightarrow{a} h_2(s)$ ) and we set  $h(w) := P$ . The existence of picks  $P$  and  $P'$  results from the completeness of  $\mathcal{G}$ . Finally, for every  $t \in (V' \setminus \{s\}) \cup (E' \setminus E'_s)$ , we set  $h(t) := h_1(t)$ . Hence  $G' \xrightarrow{h} G$  for some  $G = (V, E, \_, \_) \in \mathcal{L}_l(\mathcal{G})$  such that  $S \subseteq V \cup E$ .

( $\mathbf{s} \in \mathbf{E}'$ ) Let  $a(u, s, v) \in \mathfrak{I}'$  be the incidence of  $s$ . Observe that if there were  $s_k \in (\{u, v\} \cup E'_u \cup E'_v) \cap (S' \setminus \{s\})$  for some  $k \in [m]$  then we would have  $si' \times s_k i_k$  (by a reasoning similar to the above case of  $s \in V'$ ) contradicting  $S \subseteq \circ$ . As  $(\{u, v\} \cup E'_u \cup E'_v) \cap S' = \{s\}$ , for

every edge  $w \in E'_u \setminus \{s\}$  with corresponding incidence  $b(u, w, t) \in \mathfrak{I}'$  (resp.  $b(t, w, u) \in \mathfrak{I}'$ ), we pick  $P \in \mathbf{E}_{\mathcal{G}}$  with  $\text{lhs}(P) = h_2(u) \xrightarrow{b} h_1(t)$  (resp.  $\text{lhs}(P) = h_1(t) \xrightarrow{b} h_2(u)$ ) and we set  $h(w) := P$ . Similarly, for every edge  $w \in E'_v \setminus \{s\}$  with corresponding incidence  $c(v, w, t) \in \mathfrak{I}'$  (resp.  $c(t, w, v) \in \mathfrak{I}'$ ), we pick  $P \in \mathbf{E}_{\mathcal{G}}$  with  $\text{lhs}(P) = h_2(v) \xrightarrow{c} h_1(t)$  (resp.  $\text{lhs}(P) = h_1(t) \xrightarrow{c} h_2(v)$ ) and we set  $h(w) := P$ . Thanks to the completeness of  $\mathcal{G}$ , all those picks are possible. To close the construction, we set  $h(u) := h_2(u)$ ,  $h(v) := h_2(v)$  and, for every  $t \in (V' \setminus \{u, v\}) \cup (E' \setminus (E'_u \cup E'_v))$ , we set  $h(t) := h_1(t)$ .  $\square$

## C Theorems of Section 6

**Theorem 6.1.**  $\mathcal{G}^\dagger$  (resp.  $\mathcal{G}^{\dagger \setminus N}$ ) has a decidable  $\text{FO}(\Upsilon)$  theory for every 0L graph grammar  $\mathcal{G}$  which is complete (resp. deterministic with non-terminals  $N$ ).

*Proof.* For a complete 0L graph grammar, the result follows directly from Thm. 7.7. Let then  $\mathcal{G}$  be a D0L graph grammar with non-terminals  $N$  and consider its completion  $\mathcal{G}^{i!}$  according to Prop. 4.2. It is easy to see that  $\mathcal{G}^{\dagger \setminus N}$  can be interpreted within  $(\mathcal{G}^{i!})^\dagger$  by means of an  $\text{FO}(\Upsilon)$ -interpretation. Then, the result is obtained using the standard backwards translation associated to the interpretation. The interpretation leaves unchanged atomic predicates other than  $\text{vrt}$  and  $\text{edg}$ . Formula  $\theta_{\text{vrt}}(x)$  selecting  $\mathcal{V}(\mathcal{G})_N$  within  $(\mathcal{G}^{i!})^\dagger$  brings together all vertices from layers where no labels from  $\Gamma' := \{j\} \cup (N \cap \Gamma)$  nor  $\Sigma' := \{!\} \cup (N \cap \Sigma)$  appear

$$\theta_{\text{vrt}}(x) \quad :\Leftrightarrow \quad \text{vrt}(x) \wedge \forall y \left( x \hat{=} y \Rightarrow \neg \left( \bigvee_{\lambda \in \Gamma'} \lambda(y) \vee \exists xz \bigvee_{\lambda \in \Sigma'} \lambda(x, y, z) \right) \right) .$$

Formula  $\theta_{\text{edg}}(x)$  selecting  $\mathcal{E}(\mathcal{G})_N$  within  $(\mathcal{G}^{i!})^\dagger$  is similar. Finally, formula  $\delta(x)$  selecting the interpretation domain is obvious:  $\delta(x) :\Leftrightarrow \theta_{\text{vrt}}(x) \vee \theta_{\text{edg}}(x)$ .  $\square$

**Theorem 6.2.** There is a translation  $\tau: \text{FO}(\Gamma \cup \Sigma) \rightarrow \text{FO}(\Gamma \cup \Sigma \cup \{\hat{=}\})$  in time  $\mathcal{O}(nk + m)$ , where  $m$  is the length of the input formula,  $n$  is its alternation rank and  $k$  is the number of its variables, such that for every unix 0L graph grammar  $\mathcal{G} \in \text{0L}(\Gamma, \Sigma, \_, \_)$  and every formula  $\varphi(\mathcal{X}) \in \text{FO}(\Gamma \cup \Sigma)$ , one has

$$\mathcal{G}^\dagger(\tau(\varphi(\mathcal{X}))) = \bigcup_{G \in \mathcal{L}(\mathcal{G})} G(\varphi(\mathcal{X})) .$$

In particular, when  $\mathcal{X} = \emptyset$ , there exists  $G \in \mathcal{L}(\mathcal{G})$  such that  $G \models \varphi$ , if, and only if,  $\mathcal{G}^\dagger \models \tau(\varphi)$ .

*Proof.* The input formula is first put into prenex form:  $Q_1 \mathcal{X}_1 \dots Q_n \mathcal{X}_n \psi(\mathcal{Y})$  where  $\psi(\mathcal{Y})$  is quantifier-free and every  $Q_i \mathcal{X}_i$  with  $Q_i \in \{\exists, \forall\}$ ,  $\mathcal{X}_i = \{x_{i1} \dots x_{ik_i}\}$  and  $Q_i \neq Q_{i+1}$  is a quantifier block replacing  $Q_i x_{i1} \dots, Q_i x_{ik_i}$ . Next, every universal block  $\forall \mathcal{X}_j$  is replaced with  $\neg \exists \mathcal{X}_j \neg$ . For this transformation, we need first to parse the input formula. This is done in linear time as the set of formulae is deterministic context-free. Once parsed, putting it into prenex form is done in linear time too.

For any  $\mathcal{Z} = \{z_1, \dots, z_{|\mathcal{Z}|}\}$ , we set  $\mathcal{Z}^{\hat{=}} := z_1 \hat{=} z_2 \hat{=} \dots \hat{=} z_{|\mathcal{Z}|}$ . With this notation  $\tau$  is defined inductively as follows for every formula  $\varphi(\mathcal{X}) \in \text{FO}(\Gamma \cup \Sigma)$  in prenex form:

$$\begin{aligned} \tau(\varphi(\mathcal{X})) &= \mathcal{X}^{\hat{=}} \wedge \varphi(\mathcal{X}), & \text{for any quantifier-free } \varphi(\mathcal{X}), \\ \tau(\exists \mathcal{Y} \varphi(\mathcal{X})) &= \exists \mathcal{Y} \tau(\varphi(\mathcal{X})), \\ \tau(\neg \varphi(\mathcal{X})) &= \mathcal{X}^{\hat{=}} \wedge \neg \tau(\varphi(\mathcal{X})) . \end{aligned}$$

Observe that, as the translation stops after the prefix with quantifiers, there is no need to deal with binary connectives. For the quantitative claim, observe that  $nk$  term in the time complexity  $\mathcal{O}(nk + m)$  of the translation comes from the alternation of quantifiers. Every such alternation brings a linear size formula  $\mathcal{Z}^{\hat{=}}$  where  $\mathcal{Z}$  is the set of free variables in the subformula under the quantifier block. The qualitative claim is established as follows by structural induction.

$$\begin{aligned}
(\varphi(\mathcal{X}) \text{ is quantifier-free}) \text{ Using Lemma 4.8, we have } & \mathcal{G}^\dagger(\tau(\varphi(\mathcal{X})) = \mathcal{G}^\dagger(\mathcal{X}^\pm \wedge \varphi(\mathcal{X})) = \\
& = \mathcal{G}^\dagger(\mathcal{X}^\pm) \cap \mathcal{G}^\dagger(\varphi(\mathcal{X})) = \bigcup_{\substack{G \in \mathcal{L}(\mathcal{G}) \\ G=(V,E,-,-)}} (V \cup E)^x \cap \mathcal{G}^\dagger(\varphi(\mathcal{X})) = \bigcup_{G \in \mathcal{L}(\mathcal{G})} G(\varphi(\mathcal{X})).
\end{aligned}$$

$$\begin{aligned}
(\exists \mathcal{X} \varphi(\mathcal{X})) \text{ We have } & \mathcal{G}^\dagger(\tau(\exists \mathcal{Y} \varphi(\mathcal{X})) = \mathcal{G}^\dagger(\exists \mathcal{Y} \tau(\varphi(\mathcal{X}))) = \\
& = \pi_{\exists \mathcal{Y}}(\mathcal{G}^\dagger(\tau(\varphi(\mathcal{X})))) \stackrel{\text{i.h.}}{=} \pi_{\exists \mathcal{Y}}\left(\bigcup_{G \in \mathcal{L}(\mathcal{G})} G(\varphi(\mathcal{X}))\right) = \bigcup_{G \in \mathcal{L}(\mathcal{G})} \pi_{\exists \mathcal{Y}}(G(\varphi(\mathcal{X}))) = \bigcup_{G \in \mathcal{L}(\mathcal{G})} G(\exists \mathcal{Y} \varphi(\mathcal{X})).
\end{aligned}$$

$$\begin{aligned}
(\neg \varphi(\mathcal{X})) \text{ Using Lemma 4.8 at } (*) \text{, we have } & \mathcal{G}^\dagger(\tau(\neg \varphi(\mathcal{X})) = \\
& = \mathcal{G}^\dagger(\mathcal{X}^\pm \wedge \neg \tau(\varphi(\mathcal{X}))) \\
& = \mathcal{G}^\dagger(\mathcal{X}^\pm) \cap \mathcal{G}^\dagger(\neg \tau(\varphi(\mathcal{X}))) \\
& = \mathcal{G}^\dagger(\mathcal{X}^\pm) \cap ((\mathcal{V}(\mathcal{G}) \cup \mathcal{E}(\mathcal{G}))^x \setminus \mathcal{G}^\dagger(\tau(\varphi(\mathcal{X})))) \\
& = (\mathcal{G}^\dagger(\mathcal{X}^\pm) \cap (\mathcal{V}(\mathcal{G}) \cup \mathcal{E}(\mathcal{G}))^x) \setminus (\mathcal{G}^\dagger(\mathcal{X}^\pm) \cap \mathcal{G}^\dagger(\tau(\varphi(\mathcal{X})))) \\
& = \mathcal{G}^\dagger(\mathcal{X}^\pm) \setminus (\mathcal{G}^\dagger(\mathcal{X}^\pm) \cap \mathcal{G}^\dagger(\tau(\varphi(\mathcal{X})))) \\
& = \mathcal{G}^\dagger(\mathcal{X}^\pm) \setminus \mathcal{G}^\dagger(\tau(\varphi(\mathcal{X}))) \\
& \stackrel{\text{i.h.}}{=} \mathcal{G}^\dagger(\mathcal{X}^\pm) \setminus \bigcup_{G \in \mathcal{L}(\mathcal{G})} G(\varphi(\mathcal{X})) \\
& \stackrel{(*)}{=} \bigcup_{\substack{G \in \mathcal{L}(\mathcal{G}) \\ G=(V,E,-,-)}} (V \cup E)^x \setminus \bigcup_{G \in \mathcal{L}(\mathcal{G})} G(\varphi(\mathcal{X})) \\
& \stackrel{(*)}{=} \bigcup_{\substack{G \in \mathcal{L}(\mathcal{G}) \\ G=(V,E,-,-)}} ((V \cup E)^x \setminus G(\varphi(\mathcal{X}))) \\
& = \bigcup_{G \in \mathcal{L}(\mathcal{G})} G(\neg \varphi(\mathcal{X})) .
\end{aligned}$$

Note that equality  $(*)$  is due to inclusions  $G(\varphi(\mathcal{X})) \subseteq (V \cup E)^x$ , for every  $G = (V, E, -, -)$  in  $\mathcal{L}(\mathcal{G})$ , whereas the uniqueness of derivation guarantees that that every pair of distinct graphs  $G_1 \neq G_2$  of  $\mathcal{L}(\mathcal{G})$  is disjoint, viz.,  $(V_1 \cup E_1) \cap (V_2 \cup E_2) = \emptyset$  for  $G_1 = (V_1, E_1, -, -)$  and  $G_2 = (V_2, E_2, -, -)$ .  $\square$

**Corollary 6.4.** *FO( $\Gamma \cup \Sigma$ )-language checking problem is decidable D0L graph grammars even with non-terminals.*

*Proof.* Let  $\varphi$  be a sentence in FO( $\Gamma \cup \Sigma$ ) and  $\mathcal{G}$  be a D0L graph grammar possibly using non-terminals labels in  $M$ . Consider  $\mathcal{G}^!$  obtained by completing  $\mathcal{G}$  according to Prop. 4.2 and let  $\tau_{!i}$  be the translation of Prop. 6.3 for  $N = \{i, !\}$ , namely  $\tau_{!i}(\varphi) := \varphi \wedge \neg \exists x i(x) \wedge \neg \exists x y z !(x, y, z)$ , where  $i$  is a vertex label and  $!$  is an edge label. Let  $\tau_M$  be the translation of Prop. 6.3 for  $N = M$ . Let  $\tau_0$  be the translation of Thm. 6.2. Then

$$\begin{aligned}
& \text{there exists } G \in \mathcal{L}(\mathcal{G})_M \text{ such that } G \models \varphi \\
& \text{iff there exists } G \in \mathcal{L}(\mathcal{G}) \text{ such that } G \models \tau_M(\varphi) && \text{by Prop. 6.3} \\
& \text{iff there exists } G \in \mathcal{L}(\mathcal{G}^!)_{\{i, !\}} \text{ such that } G \models \tau_M(\varphi) && \text{as } \mathcal{L}(\mathcal{G}) = \mathcal{L}(\mathcal{G}^!)_{\{i, !\}} \text{ by Prop. 4.2} \\
& \text{iff there exists } G \in \mathcal{L}(\mathcal{G}^!) \text{ such that } G \models \tau_{!i}(\tau_M(\varphi)) && \text{by Prop. 6.3} \\
& \text{iff } (\mathcal{G}^!)^\dagger \models \tau_0(\tau_{!i}(\tau_M(\varphi))) && \text{by Thm. 6.2.}
\end{aligned}$$

$\square$

## D Automatic presentation of $\mathcal{G}^\dagger$ in Section 7 is correct for complete grammars

This appendix brings together missing proofs of Sect. 7.

**Lemma 7.1.** *If  $\mathcal{G}$  is complete, then  $\llbracket \mathcal{A}(\text{vrt}) \rrbracket = \mathcal{G}^\dagger(\text{vrt})$ .*

For facilitating the proof of this lemma, we need one additional lemma. We also recall the following notation:  $w(\leq i) = a_1 a_2 \dots a_i$  when  $w = a_1 a_2 \dots a_n$  is a word over an alphabet, say  $\Xi$ , and  $a_i \in \Xi$  for  $i \in [n]$ . Note that we omit subscript  $\mathcal{G}$  in derivation related notations and that the dashed arrow represents the existence of a derivation of specified length.

**Lemma D.1.** *For all  $(V', E', \_, \_) \in \mathcal{L}_l(\mathcal{G})$  and  $n \in [l-1]$*

- (1) *there exists a unique  $(V, E, \_, \_) \in \mathcal{L}_{l-n}(\mathcal{G})$  such that  $A_{\mathcal{G}} \xrightarrow{l-n} (V, E, \_, \_) \xrightarrow{n} (V', E', \_, \_)$ ,*
- (2)  *$u'(\leq l-n) \in V \cup E$  for every  $u' \in V' \cup E'$  and  $u'(\leq l-n) \xrightarrow{\text{prd}(u'(l-n+1)) \dots \text{prd}(u'(l))} u'$ ,*
- (3) *in particular,  $\varepsilon \xrightarrow{\text{prd}(u'(1)) \dots \text{prd}(u'(l))} u'$  where  $\text{prd}(u'(1)) = A_{\mathcal{G}} \rightarrow \_$ ,*
- (4) *for every  $u \in V \cup E$  there exists  $v \in (\Theta \cup \Omega)^n$  such that  $uv \in V' \cup E'$ .*

*Proof.* Follows directly from the definition of the derivation according to  $\mathcal{G}$ .  $\square$

With the help of the above lemma, the proof of Lemma 7.1 goes as follows.

*Proof of Lemma 7.1.* Let  $\mathcal{G} \in 0L(\Gamma, \Sigma, \Theta, \Omega)$  be a complete graph grammar. Because  $\text{ar}(\text{vrt}) = 1$  and  $\square$  is not used in  $\Delta_{\text{vrt}}$ , one has  $\llbracket \mathcal{A}(\text{vrt}) \rrbracket = \mathcal{L}(\mathcal{A}(\text{vrt}))$ .

( $\subseteq$ ) Let  $v \in \llbracket \mathcal{A}(\text{vrt}) \rrbracket$ . We show by induction on  $l = |v|$  that there exists  $(V_l, \_, \_, \_) \in \mathcal{L}_l(\mathcal{G})$  such that  $v \in V_l$ . The inclusion follows then from  $V_l \subseteq \mathcal{V}(\mathcal{G}) = \mathcal{G}^\uparrow(\text{vrt})$ .

( $l = 1$ ) Then  $\iota_{\text{vrt}} \xrightarrow{v} v \in \Delta_{\text{vrt}}$  for some  $A_{\mathcal{G}} \rightarrow (V_1, \_, \_, \_) \in \mathcal{V}_{\mathcal{G}}$  such that  $v \in V_1$ .

( $l > 1$ ) Let  $uij = v$  for some  $i, j \in \Theta$ . By construction of  $\mathcal{A}(\text{vrt})$ , there is a path  $\iota_{\text{vrt}} \xrightarrow{ui} i$  in  $\mathcal{A}(\text{vrt})$ . Then  $ui \in \mathcal{L}(\mathcal{A}(\text{vrt}))$  because  $i \in F_{\text{vrt}}$ . By induction hypothesis there exists  $G_{l-1} = (V_{l-1}, E_{l-1}, \ell_{l-1}, \mathfrak{I}_{l-1}) \in \mathcal{L}_{l-1}(\mathcal{G})$  such that  $ui \in V_{l-1}$ . As  $uij \in \mathcal{L}(\mathcal{A}(\text{vrt}))$ , there is a transition  $i \xrightarrow{j} j \in \Delta_{\text{vrt}}$  such that  $\text{prd}(j) = P = \ell(i) \rightarrow \_$ . Then, one can consider label preserving map  $h: V_{l-1} \rightarrow \mathcal{V}_{\mathcal{G}}$  such that  $h(v) = P$ . According to Lemma 4.1, this map extends into homomorphism  $h: G_{l-1} \rightarrow \mathcal{G}$ . The latter produces from  $G_{l-1}$  a graph in  $\mathcal{L}_l(\mathcal{G})$ , say  $(V_l, \_, \_, \_)$ , such that  $v = uij \in V_l$ .

( $\supseteq$ ) Let  $v \in \mathcal{V}_l(\mathcal{G}) \subseteq \mathcal{G}^\uparrow(\text{vrt})$ . Thus, there exists  $(V_l, \_, \_, \_) \in \mathcal{L}_l(\mathcal{G})$  such that  $v \in V_l$ . By (2) and (3) of Lemma D.1, we have  $\varepsilon \xrightarrow{\text{prd}(v(1)) \dots \text{prd}(v(l))} v$  where  $\text{prd}(v(1)) = A_{\mathcal{G}} \rightarrow \_$  and, for every  $k \in [l-1]$ , we have  $v(\leq k) \xrightarrow{\text{prd}(v(k+1))} v(\leq k+1)$  where  $\text{prd}(v(k+1)) = \ell(v(k)) \rightarrow \_$ . Then, the construction of  $\mathcal{A}(\text{vrt})$  allows us to affirm that  $\iota_{\text{vrt}} \xrightarrow{v(1)} \dots \xrightarrow{v(l)} v(l)$  is an accepting path of  $\mathcal{A}(\text{vrt})$ . Hence,  $v \in \mathcal{L}(\mathcal{A}(\text{vrt}))$ .  $\square$

**Lemma 7.3.** *If  $\mathcal{G}$  is complete, then  $\llbracket \mathcal{A}_j \rrbracket = \mathcal{J}(\mathcal{G})$ , where  $\mathcal{J}(\mathcal{G})$  is the set of incidences of  $\mathcal{G}^\uparrow$ .*

Remember that  $\mathcal{J}(\mathcal{G})$  is defined in Sect. 4 as  $\mathcal{J}(\mathcal{G}) := \bigcup_{(\_, \_, \_, \mathfrak{I}) \in \mathcal{L}(\mathcal{G})} \mathfrak{I}$ . Notation  $\llbracket \mathcal{A}_j \rrbracket$  is introduced in Sect. 7 just above Lemma 7.3 together with a view of  $\overrightarrow{\Sigma} \cup \overleftarrow{\Sigma}$  as a family of mappings. To keep the proof simple, all orientations of  $\overrightarrow{a}$  and  $\overleftarrow{b}$  are taken as in the construction of  $\mathcal{A}_j$  without being explicitly mentioned. If instead orientations were set explicitly, there would be four similar cases to consider in the proof.

*Proof.* Let  $\mathcal{G} \in 0L(\Gamma, \Sigma, \Theta, \Omega)$  be a complete graph grammar. Observe that  $\square$  is not used in  $\Delta_j$ . Consequently every path of length  $n$  of  $\mathcal{A}_j$  is labelled by some  $\otimes(s_1, s_2, s_3)$  such that  $s_1, s_2, s_3 \in (\Theta \cup \Omega)^n$ .

( $\subseteq$ ) Let  $\overleftarrow{a}(u, w, v) \in \llbracket \mathcal{A}_j \rrbracket$ . We show by induction on  $l = |u| = |v| = |w|$  that there exists  $(\_, \_, \_, \mathfrak{I}_l) \in \mathcal{L}_l(\mathcal{G})$  such that  $\overleftarrow{a}(u, w, v) \in \mathfrak{I}_l$ . The inclusion follows then from  $\mathfrak{I}_l \subseteq \mathcal{J}(\mathcal{G})$ .

( $l = 1$ ) Then  $\iota_j \xrightarrow{(u, w, v)} (\overleftarrow{a}, u, w, v) \in \Delta_j$  for some  $A_{\mathcal{G}} \rightarrow (\_, \_, \_, \mathfrak{I}_1) \in \mathcal{V}_{\mathcal{G}}$  such that  $a(u, w, v) \in \mathfrak{I}_1$ .



( $l > 1$ ) Let  $(u'i'i, w'j'j, v'k'k) = (u, w, v)$  for some  $i, i', k, k' \in \Theta$ ,  $j \in \Omega$  and  $j' \in \Theta \cup \Omega$ . By construction of  $\mathcal{A}_j$ , there is a path  $\iota_j \xrightarrow{\otimes(u'i', w'j', v'k')} (\beta, i', q, k')$  in  $\mathcal{A}_j$  where  $q = \square$  and  $\beta = \diamond$  or,  $q = j'$  and  $\beta = \overleftarrow{b}$ , viz.,  $\beta \in \{\overleftarrow{b}, \overrightarrow{b}\}$ , for some  $b \in \Sigma$ .

- If  $q = \square$  then the above path can be more precisely written as

$$\iota_j \xrightarrow{(r_1, r_1, r_1)} (\diamond, r_1, \square, r_1) \xrightarrow{(r_2, r_2, r_2)} \dots \xrightarrow{(r_{l-1}, r_{l-1}, r_{l-1})} (\diamond, r_{l-1}, \square, r_{l-1}) \quad (*)$$

where  $i' = j' = k' = r_{l-1}$  and  $u' = v' = w' = r_1 \dots r_{l-2}$  with  $\{r_1, \dots, r_{l-1}\} \subseteq \Theta$ . This path mimics path  $\iota_{\text{vrt}} \xrightarrow{r_1} r_1 \xrightarrow{r_2} \dots \xrightarrow{r_{l-1}} r_{l-1}$  of  $\mathcal{A}(\text{vrt})$ . By Lemma 7.1, this is an accepting path. Thus  $u'i' = w'j' = v'k' = \bar{r}$ , where  $\bar{r} := r_1 \dots r_{l-1}$ , is in  $\mathcal{V}_{l-1}(\mathcal{G})$ . As  $\mathcal{A}_j$  is deterministic and  $\overleftarrow{a}(u, w, v) \in \llbracket \mathcal{A}_j \rrbracket$ , path  $(*)$  extends in  $\mathcal{A}_j$  into the path accepting  $\otimes(u, w, v)$ . Consequently, there is  $(\diamond, r_{l-1}, \square, r_{l-1}) \xrightarrow{(i, j, k)} (\overrightarrow{a}, i, j, k) \in \Delta_j$  such that  $\ell(r_{l-1}) \rightarrow (\_, \_, \_, \mathfrak{I}) \in \mathcal{V}_{\mathcal{G}}$  and  $a(i, j, k) \in \mathfrak{I}$ . By definition of the derivation, there exists  $G_l = (V_l, E_l, \ell_l, \mathfrak{I}_l) \in \mathcal{L}_l(\mathcal{G})$  such that  $a(\bar{r}i, \bar{r}j, \bar{r}k) \in \mathfrak{I}_l$ . Hence,  $a(u, w, v) = (\bar{r}i, \bar{r}j, \bar{r}k) \in \mathfrak{I}_l(\mathcal{G})$ .

- If  $q = j'$ , then  $\iota_j \xrightarrow{\otimes(u'i', w'j', v'k')} (\overleftarrow{b}, i', j', k')$  is an accepting path of  $\mathcal{A}_j$ , viz.,  $\overleftarrow{b}(u'i', w'j', v'k') \in \llbracket \mathcal{A}_j \rrbracket$ . By induction hypothesis there exists  $G_{l-1} = (\_, \_, \ell_{l-1}, \mathfrak{I}_{l-1}) \in \mathcal{L}_{l-1}(\mathcal{G})$  such that  $\overleftarrow{b}(v'k', w'j', u'i') \in \mathfrak{I}_{l-1}$ . As  $\otimes(u, w, v) \in \mathcal{L}(\mathcal{A}_j)$ , there is a transition  $(\overleftarrow{b}, i', j', k') \xrightarrow{(i, j, k)} (\overrightarrow{a}, i, j, k) \in \Delta_j$ . By construction of  $\mathcal{A}_j$ , there exist  $\text{prd}(i) = \ell_{l-1}(i') \rightarrow \_$ ,  $\text{prd}(k) = \ell_{l-1}(k') \rightarrow \_ \in \mathcal{V}_{\mathcal{G}}$  and  $\text{prd}(j) = \text{prd}(i) \xrightarrow{\overleftarrow{b}} \text{prd}(k) \rightarrow \sqsupset \in \mathcal{E}_{\mathcal{G}}$  such that  $\overrightarrow{a}(i, j, k) \in \sqsupset$ . Then, one can consider label preserving map  $h: V_{l-1} \rightarrow \mathcal{V}_{\mathcal{G}}$  such that  $h(u'i') = \text{prd}(i)$ ,  $h(v'k') = \text{prd}(k)$ . Observe that this homomorphism can be chosen so that  $h(w'j') = \text{prd}(j)$ . Consequently,  $h$  produces from  $G_{l-1}$  a graph in  $\mathcal{L}_l(\mathcal{G})$ , say  $(\_, \_, \_, \mathfrak{I}_l)$ , such that  $\overrightarrow{a}(u, w, v) \in \mathfrak{I}_l$ .

( $\supseteq$ ) Let  $\overrightarrow{a}(u, w, v) \in \mathfrak{I}_l(\mathcal{G})$ . Thus, there exists  $(\_, \_, \_, \mathfrak{I}_l) \in \mathcal{L}_l(\mathcal{G})$  such that  $\overrightarrow{a}(u, w, v) \in \mathfrak{I}_l$ . We show by induction on  $l \in \mathbb{N}_{>0}$  that  $\iota_j \xrightarrow{\otimes(u, w, v)} (\overrightarrow{a}, \_, \_, \_)$  is a path in  $\mathcal{A}_j$ . Note that such a path is accepting.

( $l = 1$ ) As  $\overrightarrow{a}(u, w, v) \in \mathfrak{I}_1$  for some  $G_1 = (\_, \_, \_, \mathfrak{I}_1) \in \mathcal{L}_1(\mathcal{G})$ , there is production  $A_{\mathcal{G}} \rightarrow G_1 \in \mathcal{V}_{\mathcal{G}}$ . Then more precisely,  $\overrightarrow{a}(u, w, v) = a(u, w, v) \in \mathfrak{I}_1$ . Consequently,  $\iota_j \xrightarrow{(u, w, v)} (\overrightarrow{a}, u, w, v) \in \Delta_j$ .

( $l > 1$ ) As  $\overrightarrow{a}(u, w, v) \in \mathfrak{I}_l$  for some  $G_l = (\_, E_l, \_, \mathfrak{I}_l) \in \mathcal{L}_l(\mathcal{G})$ , by Lemma D.1(1), there exists  $G_{l-1} = (V_{l-1}, \_, \ell_{l-1}, \mathfrak{I}_{l-1}) \in \mathcal{L}_{l-1}(\mathcal{G})$  such that  $G_{l-1} \xrightarrow{h} G_l$  for some homomorphism  $h: G_{l-1} \rightarrow \mathcal{G}$ . As  $w \in E_l$  is either innate or inherited, we have two cases.

- $(u, w, v) = (si, sj, sk)$  for some  $i, k \in V_0$ ,  $j \in E_0$  and  $s \in V_{l-1}$  such that  $G_0 = (V_0, E_0, \_, \mathfrak{I}_0)$  is the right hand side of vertex production  $h(s) = \ell_{l-1}(s) \rightarrow G_0$ . Moreover  $a(i, j, k) \in \mathfrak{I}_0$ . By Lemma 7.1,  $\iota_{\text{vrt}} \xrightarrow{s} i' = s(l-1)$  is an accepting path in  $\mathcal{A}(\text{vrt})$ . It is mimicked in  $\mathcal{A}_j$  by  $\iota_j \xrightarrow{\otimes(s, s, s)} (\diamond, i', \square, i')$ . From there one more step can be done as, by construction,  $\{(\diamond, i', \square, i') \xrightarrow{(i, j, k)} (\overrightarrow{a}, i, j, k) \in \Delta_j$ . Consequently,  $\overrightarrow{a}(u, w, v) = \overrightarrow{a}(si, sj, sk) \in \llbracket \mathcal{A}_j \rrbracket$ .
- $(u, w, v) = (u'i, w'j, v'k)$  where  $\overleftarrow{b}(u', w', v') \in \mathfrak{I}_{l-1}$  and  $\overrightarrow{a}(i, j, k) \in \sqsupset$  for some edge production  $h(w') = \text{prd}(i) \xrightarrow{\overleftarrow{b}} \text{prd}(k) \rightarrow \sqsupset$ . By induction hypothesis,  $\overleftarrow{b}(u', w', v') \in \llbracket \mathcal{A}_j \rrbracket$ . Consequently,  $\iota_j \xrightarrow{\otimes(u', w', v')} (\overleftarrow{b}, i', j', k')$  where  $i' = u'(l-1)$ ,  $j' = w'(l-1)$  and  $k' = v'(l-1)$ . By construction,  $(\overleftarrow{b}, i', j', k') \xrightarrow{(i, j, k)} (\overrightarrow{a}, i, j, k) \in \Delta_j$ . Therefore, the above path extends into  $\iota_j \xrightarrow{\otimes(u'i, w'j, v'k)} (\overrightarrow{a}, i, j, k)$ . Hence,  $\overrightarrow{a}(u, w, v) = \overrightarrow{a}(u'i, w'j, v'k) \in \llbracket \mathcal{A}_j \rrbracket$ .

□

**Lemma 7.6.** *If  $\mathcal{G}$  is complete, then  $\llbracket \mathcal{A}(\chi_0) \rrbracket = \chi_0$ .*

*Proof.* Since  $\mathcal{B}_1$  accepts all elements of  $\mathcal{V}(\mathcal{G}) \cup \mathcal{E}(\mathcal{G})$  duplicated, after switching via  $\Delta$  into  $\mathcal{B}_2$  every pair  $(w, w) \in \llbracket \mathcal{B}_1 \rrbracket$  extends into pair  $(wu, wv)$  such that  $(u, v) \in \llbracket \mathcal{B}_2 \rrbracket$ , where  $(u, v)$  is any pair in  $(\mathcal{V}_l(\mathcal{G}) \cup \mathcal{E}_l(\mathcal{G}))^2$ , for  $l \in \mathbb{N}$ . However, the switch is activated by any pair of distinct productions sharing the same left hand side. This behaviour matches precisely the definition of  $\chi_0$ . □