グラフマイナー定理に基づく
線形時間アルゴリズムの自動生成

小川 瑞史（NTT）

近年 Robertson-Seymour により証明されたグラフマイナー定理（Wagnber の予想）により有限グラフ上の minor-closed な性質の検出は多項式時間アルゴリズムが存在することが知られている。しかし、その証明は非構成的であり一般に実際のアルゴリズムを構成するのは難しかった。本稿ではグラフマイナー定理の簡単な場合である Higman の補題の構成的証明を用い、未解決問題であった時間概念を含むデータベース上の選言的単項質問処理の線形時間アルゴリズムを自動生成により与えた。

Polynomial-time algorithm generation based on graph minors (draft)

Mizuhito Ogawa (mizuhito@ntt-20.ntt.jp)
NTT Basic Research Laboratories
3-9-11 Midori-cho Musashino-shi Tokyo 180 Japan

In 1988, Robertson-Seymour have proved graph-minor theorem (Wagner's conjecture). Fellows shows that the detection of a minor-closed property on finite graphs has a polynomial-time algorithm. However, the proof of graph-minor theorem is highly non-constructive and in general it is very difficult to construct an actual polynomial-time algoritm. In this report, using a constructive proof of Higman's lemma (which is a restricted version of graph-minor theorem), we generate a liner-time algorithm for a disjunctive monadic query processing on indefinite database, which has been an open problem.

Polynomial-time algorithm generation based on graph minors (draft)

Mizuhito Ogawa (mizuhito@ntt-20.ntt.jp)
NTT Basic Research Laboratories
3-9-11 Midori-cho Musashino-shi Tokyo 180 Japan

1 Introduction

Well-quasi-order (wqo, for short) is frequently used concept in theoretical computer science. A quasi-order $\leq$ (i.e., reflexive transitive relation) is wqo if for any infinite sequence $x_1, x_2, x_3, \cdots$ there exist $i < j$ s.t. $x_i \leq x_j$. A wqo (or its variations) $\leq$ on some simple set is extended to a wqo $\sqsubseteq$ on more complex structures - such as, direct product, finite power set, finite words [7], finite trees [8, 14], finite graphs [18], etc. infinite words [15], infinite trees [16], etc. (But, not for infinite graphs [20].) The extension is constructed by a *homeomorphical embedding* - that is, if there exists an one-to-one *structure-preserving* map ϕ from x into y s.t. $x_i \leq f(x_i)$ for all elements x_i in x, then $x \sqsubseteq y$.

One application of wqo is to show termination, such as simple termination in term rewriting systems [2]. Another application is to show the existence of a polynomial-time algorithm to detect a *minor-closed* property [3]. A property P is minor-closed if $P(x)$ and $x \sqsubseteq y$ shows $P(y)$. Graph minor theorem (or, Wagner's conjecture) is recently proved by Robertson and Seymour, which states that an ordering $\sqsubseteq$ based on a homeomorphical embedding is a wqo on finite graphs [18]. Then, the minimal elements which hold P must be finite - otherwise, it contradicts to wqo property of $\sqsubseteq$. These minimal elements are called *forbidden-minors.* The problem whether $P(x)$ holds is reduced to the simple test whether there exists a forbidden-minor y s.t. $y \sqsubseteq x$. This is computed in $O(n^3)$-time where n is a number of edges in x. However, this does not suggest any actual algorithm - a set of all forbidden-minors is usually difficult to detect. For instance, the non-planarity of finite graphs requires Kuratowski's theorem. It shows K_5 and $K_{3,3}$ are required forbidden-minors.

This paper is the first step to an automatic polynomial-time algorithm generation for a minor-closed property. The basic idea is to construct an algorithm which detects forbidden-minors. This construction is based on a constructive proofs of Higman's lemma [13]. This is applied to a linear-time algorithm generation for a fixed disjunctive query processing on indefinite database, which is an open problem in [9].

2 Well quasi order

A *quasi order* on a set Σ is a reflexive and transitive binary relation on Σ, and denoted as $x \leq y$. A *decidable quasi order* is a quasi order for which the condition $x \leq y$ is decidable.

Definition 2.1 A quasi order on a set Σ is a *well-quasi-order* (*wqo*, for short) if for any infinite sequence $x_1, x_2, x_3, \cdots$ in Σ there exist $i < j$ s.t. $x_i \leq x_j$. We also say $(\Sigma, \leq)$ or simply Σ as a *well-quasi-ordered set* (*wqo* set, for short).

In 1952, Higman shows *wqo* on a set Σ can be extended to a set Σ^* of finite words on Σ [7].

Lemma 2.1 (Higman's lemma) Let $x = (x_1, \cdots, x_m)$, $y = (y_1, \cdots, y_n)$ be words (finite sequences) on a wqo set $(\Sigma, \leq)$. A set of words on Σ is noted as Σ^*. A relation $x \lll y$ is defined that x is termwise dominated by a subword in y. (i.e., there exists a strongly increasing map ψ from $\{1, \cdots, m\}$ to $\{1, \cdots, n\}$ s.t. $x_i \leq y_{\psi(i)}$ for $\forall i \leq m$.) Then $(\Sigma^*, \lll)$ is a wqo set.

In early 60's, this lemma is extended to more complex structures, such as finite trees, infinite trees, transfinite sequences, etc. by Nash-Williams[14, 15, 16]. The most well-known form would be Kruskal's theorem[8, 14]

Definition 2.2 Let T be a tree. For a node x in T, *position*(x) is inductively defined as

$$\begin{cases} x = \epsilon & \text{if } x \text{ is a root.} \\ x = p \cdot i & \text{if } x \text{ is an } i\text{-th branch of } x' \text{ and } position(x') = p. \end{cases}$$

Two positions p, q are $p \prec q$ if $\exists r$ s.t. $q = p \cdot r$. We denote $p \wedge q$ as a greatest lower bound of positions p and q.

Definition 2.3 Let T, T' be finite trees. T is *homeomorphically embedded* in T' if there exists an injection ψ from T to T' s.t. $\psi(x \wedge y) = \psi(x) \wedge \psi(y)$ for any nodes x, y in T. ψ is said to be a *homomorphic embedding.*

Theorem 2.1 (Kruskal's theorem) Let T, T' be finite trees, and $T \sqsubseteq T'$ be a relation s.t. T is homeomorphically embeddable to T'. Then, $\sqsubseteq$ is a wqo on a set of finite trees.

Theorem 2.2 Let T, T' be finite terms on a wqo set W. We define $T \sqsubseteq T'$ if there exists an injection map ψ from T to T' s.t. ψ is a homomorphic embedding (regarding a term as a tree) and $x \leq \psi(x)$ for each node in T. Then, $\sqsubseteq$ is a wqo on a set of finite terms on W.

These results are surveyed in [11]. Wagner conjectured the extension of Kruskal's theorem for graphs. Wagner's conjecture (graph-minor theorem) has been open for more than 20 years. In 1988, it is positively solved by Robertson and Seymour for finite graphs [18]. (For infinite graphs, Wagner's conjecture is refuted [20].)

The proofs are based on highly non-constructive reasoning called *minimal bad sequence* - this is far beyond Peano Arithmetic (**PA**). The proofs are non-constructive from two reasons: (1) the proofs proceed by contradiction, (2) the arguments on minimal-bad-sequence require heavily impredicative, such as universal quantification over all bad sequences. Next, Kruskal's theorem can prove a transfinite induction of an ordinal Γ_0. This shows Kruskal's theorem can be proved neither in **PA** nor relatively strong logical systems such as $\mathbf{ACA}_0$, $\mathbf{ATR}_0$, etc. These proof-theoretical view on Kruskal's theorem is surveyed in [4].

Principally, the proof by contradiction for Π_2^0 sentences can be transferred to constructive (impredicative) proof by *A-translation* [6]. Thus, these theorems would have (imperative) constructive proofs. However, the rest of facts above shows difficulty to find predicative constructive proofs for graph-minor and related theorem.

Murthy-Russell and Richman-Stolzenberg have independently given an explicit constructive proof of Higman's lemma [13, 17]. Similar idea to [17] is also found in [19]. Gupta extended Murthy-Russell's regular expression techniques to finite trees and have given an explicit constructive proof of Kruskal's theorem [5]. Murthy-Russell's proof takes a finite sequence and maps it to a finite set of regular expressions of elements incomparable to those in the sequence. They define a well-founded ordering on these sequences which decreases as the sequence is extended with incomparable elements. This ordering can be translated into a function which maps finite sequences to ordinals. Murthy-Russell are implementing an algorithm extracted from their constructive proof using the NuPRL proof development system. We will discuss their detailed proof techniques in the next section.

3 Constructive proof of Higman's lemma

Briefly speaking, Murthy-Russell's proof (1) takes a finite sequence and maps it to some sort of regular expressions of the elements incomparable to those in the sequence. (2) They define a well-founded ordering on these sequences which decreases as the sequence is extended with incomparable elements. This is shown by an induction on sets of sets of sets over a fixed carrier set. We will obey notations and definitions to [13]. For detailed proofs please refer [13]. Let $(\Sigma, \leq)$ be a wqo set, Σ^* be a set of finite words on Σ, and $\lll$ be a quasi-order defined in theorem 2.1.

Assumption

1. Let A and B be non-increasing sequences of Σ and let $A \sqsubset_{seq} B$ if and only if A is a proper extension of B. $\sqsubset_{seq}$ is well-founded and equipped with a well-founded induction scheme.

2. The wqo $\leq$ on Σ is decidable.

Classically, the first assumption is obvious from the wqo property of $\leq$. However, constructively it is not. The wqo which satisfies assumptions above is called *constructive well-quasi-order* (*cwqo*, for short) [19].

Step 1: Sequential r.e.'s and their reductions

Definition 3.1 A *sequential r.e.* (on Σ) is a (possibly empty) concatenation of either *constant expressions* or *starred expressions* (on Σ) defined below. Let $b \in \Sigma$ and $A = a_1, a_2, \cdots, a_k$ be a finite non-increasing sequence of Σ of length k.

$$\begin{cases} \textit{constant expression} & (b-A) & = & \{x \mid b \leq x \text{ and } a_i \not\leq x \text{ for } \forall i \leq k\} \\ \textit{starred expression} & (\Sigma - A)^* & = & \{w = w_1 w_2 \cdots w_n \in \Sigma^* \mid \forall i (\leq n), j(\leq k) \text{ s.t. } w_i \not\leq a_j\} \end{cases}$$

Let s_1, s_2, s_3, $\cdots$ be a non-increasing sequence of elements of Σ^*. We assign a set of sequential r.e.'s Θ_i to each stage s_i, and analyze how Θ_i is reduced to Θ_{i+1} at the next stage $s_{i+1} \in \sigma \in \Theta_i$. The intuition behind is that $\cup_{\sigma \in \Theta_i} \sigma$ is a set of finite words not in the upward closure of $\{s_1, s_2, \cdots, s_{i-1}\}$. The basic idea is that for a word not to be a superword of w, it can contain only a proper subword of w. So what we do is write down sequential r.e.'s which each accept classes of words containing different subwords of w. The following lemmas mean that if we remove the sequential r.e. σ from E and replace it with the set $\Theta(\sigma, s)$, then the resulting set of sequential r.e.'s includes all the finite words in $\cup_{\forall \sigma \in E}\, \sigma$ not containing s.

Definition 3.2 Let $s \in \Sigma^*$. We define $s^\circ = \{x \in \Sigma^* \mid s \lll x\}$.

Lemma 3.1 Let $\sigma \subseteq \Sigma^*$ be a sequential r.e., and a finite word $s \in \sigma$. Then we can construct a finite set of sequential r.e.'s $\Theta(\sigma, s)$ s.t. if $x \in \sigma$ and $s \not\lll x$ then there is a sequential r.e. $\theta \in \Theta(\sigma, s)$ with $x \in \theta$. (See [13] for an actual construction of $\Theta(\sigma, s)$.)

Lemma 3.2 Let $S \subseteq \Sigma^*$ and a finite set of sequential r.e.'s E s.t. for $\forall s \in S$ there is $\sigma \in E$ with $s \in \sigma$. Then for any s, σ s.t. $s \in \sigma \in E$ we have $\forall s' \in S - s^\circ$ $\exists \sigma' \in (E - \{\sigma\}) \cup \Theta(\sigma, s)$ s.t. $s' \in \sigma'$.

Step 2: Well-founded ordering on a finite set of sequential r.e.'s

Definition 3.3 We define orderings on starred and constant r.e.'s as

$$\begin{cases} \text{starred r.e.} & (\Sigma - A)^* \sqsubset_* (\Sigma - B)^* & \Leftrightarrow & A \sqsubset_{seq} B \\ \text{constant r.e.} & (a - A) \sqsubset_{const} (b - B) & \Leftrightarrow & a = b \wedge A \sqsubset_{seq} B \end{cases}$$

Let $\sqsubset_{exp} = \sqsubset_{seq} \cup \sqsubset_*$ and let a multiset extension [1] of $\sqsubset_{exp}$ be $\sqsubset_{setexp}$. We define an ordering $\sqsubset_{re}$ on sequential r.e.'s as $\alpha \sqsubset_{re} \beta \Leftrightarrow \uplus_{i=1}^{k}\{a_i\} \sqsubset_{setexp} \uplus_{i=1}^{l}\{b_i\}$, where sequential r.e.'s $\alpha = a_1 \cdots a_k$ and $\beta = b_1 \cdots b_l$. We also denote a multiset extension of $\sqsubset_{re}$ as $\sqsubset_{setre}$.

Lemma 3.3 (1) Let σ be a sequential r.e. and $w \in \sigma$. Then, $\theta \sqsubset_{re} \sigma$ for $\forall \theta \in \Theta(\sigma, w)$.
(2) Let E be a set of sequential r.e.'s and $w \in \sigma \in E$. Then, $(E - \{\sigma\}) \cup \Theta(\sigma, w) \sqsubset_{setre} E$.

Theorem 3.1 Let $s = s_1, s_2, \cdots$ be a sequence of finite words in Σ^*, and let E be a set of sequential r.e.'s. Then, for any integer $k \geq 1$,

$$\Sigma^* - (s_1^\circ \cup s_2^\circ \cup \cdots \cup s_{k-1}^\circ) \subseteq \cup_{\sigma \in E}\, \sigma \quad \Rightarrow \quad \exists j \geq k\ \exists i < j \text{ s.t. } s_i \lll s_j$$

Note that $\sqsubset_*$, $\sqsubset_{const}$, $\sqsubset_{exp}$, $\sqsubset_{setexp}$, $\sqsubset_{re}$, and $\sqsubset_{setre}$ are well-founded. The proof of the theorem is due to an induction wrt $\sqsubset_{setre}$ on E. Then, Higman's lemma is directly proved as a corollary of theorem 3.1 by setting $E = \{\Sigma^*\}$ and $k = 1$. (Furthermore, by repeating similar techniques, $(\Sigma^*, \lll)$ is shown to be a cwqo set [13].)

Miscellaneous: Finite sets on a wqo set

Next, we consider $\mathcal{FP}(\Sigma)$ which is a set of all finite sets of Σ. Assume Σ satisfy cwqo assumptions same as Higman's lemma. We define $s_1 \leq_m s_2$ for $s_1, s_2 \in \mathcal{FP}(\Sigma)$ when for each $x \in s_1$ there exists $y \in s_2$ s.t. $x \leq y$. Let us take the maximal elements as a representative of an element in $\mathcal{FP}(\Sigma)/\equiv$ for $\equiv = \leq_m \cup \geq_m$. The similar and easier discussion to Higman's lemma is repeated. We will also denote $s^\circ = \{x \in \mathcal{FP}(\Sigma) \mid s \leq_m x\}$.

Definition 3.4 Let $A = a_1, a_2, \cdots, a_k$ be a finite non-increasing sequence of elements of Σ. A *base r.e.* (on Σ) is $\Sigma \ominus A = \mathcal{FP}(\{x \in \Sigma \mid a_i \not\leq x \text{ for } i \leq k\})$. We define $\Sigma \ominus A \sqsubset_{base} \Sigma \ominus B$ if and only if $A \sqsubset_{seq} B$.

Let $s_1, s_2, s_3, \cdots$ be a non-increasing sequence of elements of $\mathcal{FP}(\Sigma)$. The intuition behind is that $\cup_{\sigma \in \Phi_i} \sigma$ is a set of finite sets not in the upward closure of $\{s_1, s_2, \cdots, s_{i-1}\}$. The basic idea is that for a finite set not to be a superset of s (which is a representative of $\mathcal{FP}(\Sigma)/\equiv$), it cannot contain one of the elements in s. So what we do is write down base r.e.'s which each accept classes of finite sets not containing different elements of s. The following lemmas mean that if we remove the r.e. σ from E and replace it with the set $\Phi(\sigma, s)$ (similar to $\Theta(\sigma, s)$ in step 2), then the resulting set of r.e.'s includes all the finite sets in $\cup_{\forall \sigma \in E}\, \sigma$ not containing s.

Lemma 3.4 Let σ be a base r.e. and let $s \in \sigma$. Then we can construct a finite set of base r.e.'s $\Phi(\sigma, s)$ s.t. if $x \in \sigma$ and $s \not\leq_m x$ then there is a base r.e. $\theta \in \Phi(\sigma, s)$ with $x \in \theta$.

Proof Let $\sigma = \Sigma \ominus A$ for a non-increasing sequence $A = a_1, a_2, \cdots, a_k$ of elements of Σ and $s_A = \{c \in s \mid a_i \not\leq c \text{ for } \forall i \leq k\}$. Since $s \in \sigma$, $s_A \neq \phi$. Then $\Phi(\sigma, s) = \{\Sigma \ominus A|c \mid c \in s_A\}$. ■

Lemma 3.5 Let $S \subseteq \mathcal{FP}(\Sigma)$ and let E be a finite set of base r.e.'s s.t. for $\forall s \in S$ there is $\sigma \in E$ with $s \in \sigma$. Then for any s, σ s.t. $s \in \sigma \in E$ we have $\forall s' \in S - s^\circ\ \exists \sigma' \in (E - \{\sigma\}) \cup \Phi(\sigma, s)$ s.t. $s' \in \sigma'$.

Since $\sqsubset_{seq}$ is well-founded, $\sqsubset_{base}$ and its multiset extension are well-founded, too [1]. By the same discussion for the constructive proof of Higman's lemma, we obtain a constructive proof of the wqo-property of $(\mathcal{FP}(\Sigma), \leq_m)$.

Lemma 3.6 If $(\Sigma, \leq)$ is a cwqo set, $(\mathcal{FP}(\Sigma), \leq_m)$ is a wqo set.

4 Disjunctive monadic query on indefinite database

The problem indicated in this section is given by R. van der Meyden [9]. His open problem is: *Let fix a disjunctive monadic query. There exists an algorithm answering to the query which is linear to a size of indefinite database. Then, what is an actual algorithm ?* He solved this problem by a manual construction[10]. We will observe the same problem as an example of an algorithm generation based on graph-minors. We obey definitions and notations to [9]. For detailed descriptions and proofs, please refer [9].

Proper atoms are of the form $P(a)$ where P is a predicate symbol and a is a tuple of constants or variables. *Order atoms* are of the form $u < v$, where u and v are order constants or variables.

Indefinite database D is a set of ground atoms, where atoms are either proper atoms or order atoms. Indefinite database D is a collection of facts on a linearly ordered domain, such as time.

A *query* is a positive existential first-order clause constructed from proper atoms and order atoms using only $\exists$, $\wedge$, and $\vee$. A *conjunctive query* is a first-order clause constructed from proper atoms and order atoms using only $\exists$ and $\wedge$. For simplicity, we assume queries are in disjunctive normal forms.

Definition 4.1 A conjunctive query is *sequential* if its form is

$$\exists t_1 t_2 \cdots t_n \; [t_1 < t_2 < \cdots < t_n \;\wedge\; \Psi(t_1, t_2, \cdots, t_n)]$$

where Ψ contains no quantification over order variables $t_1, t_2, \cdots, t_n$.

We concentrate on *monadic queries* - queries in which each proper atom allows only monadic predicate symbols. A predicate symbol is *monadic* if its arity is at most one. A class of monadic queries is restrictive, but contains non-trivial problems, such as comparing two gene alignments (regarding C, G, A, T as monadic predicates) [9].

Let *Pred* be a set of monadic predicates, and let $\Sigma = \mathcal{P}(Pred)$ be the power set of *Pred*. The set Σ^* is the set of all finite words of symbols in Σ. Without loss of generality, we can assume that a monadic query does not contain constants and indefinite database is monadic (i.e., each proper atom in database is monadic). Then, up to variable-renaming sequential monadic queries have an one-to-one correspondence to words in Σ^* by regarding a set of monadic predicate symbols which hold at t_i as an order variable t_i in an order atom. For instance, $\exists t_1 t_2 t_3 \; [P(t_1) \wedge Q(t_1) \wedge P(t_2) \wedge R(t_3) \wedge \; t_1 < t_2 < t_3]$ corresponds to $\{P, Q\}\{P\}\{R\}$. If Ψ is a conjunctive monadic query a *path* in Ψ is a maximal sequential subquery of Ψ. In other words, a path is the maximal words in Ψ in the above representation. We write $Paths(\Psi)$ for the subset of Σ^* corresponding to paths of Ψ.

Lemma 4.1 Let D be a monadic database and Ψ be a conjunctive monadic query. Then, $D \models \Psi$ if and only if $D \models p$ for every path $p \in Paths(\Psi)$.

Let $P_1, P_2, \cdots, P_n$ be either proper or order atoms. Regarding a monadic database $D = \{P_1, P_2, \cdots, P_n\}$ as a conjunctive monadic formula $P_1 \wedge P_2 \wedge \cdots \wedge P_n$, paths of database are similarly defined. We denote a set of paths of D as $Paths(D)$. Note that detecting a path in monadic database is a kind of *sorting*, thus its complexity is linear to the size of database.

Lemma 4.2 Let p be a sequential query, and $p \lll q$ is a subword relation constructed from a set inclusion on Σ. Then $D \models p$ if and only if there exists a path $q \in Paths(D)$ s.t. $p \lll q$.

Since Σ is finite, the set-inclusion $\subseteq$ in $\Sigma = \mathcal{P}(Pred)$ is a wqo. $(\Sigma^*, \lll)$ is also a cwqo set by Higman's lemma and $(\mathcal{FP}(\Sigma^*), \leq_m)$ is also a cwqo set by lemma 3.6 Two monadic database D_1 and D_2 are $D_1 \sqsubseteq D_2$ if and only if $Paths(D_1) \leq_m Paths(D_2)$.

Theorem 4.1 For any disjunctive monadic query Ψ, if $D_1 \models \Psi$ and $D_1 \sqsubseteq D_2$ then $D_2 \models \Psi$.

This theorem is deduced from lemmas above and consideration on models of indefinite database. Higman's lemma shows an existence of an algorithm for a fixed disjunctive monadic query processing on indefinite database which is linear to the size on database. Note that a disjunctive query $p \vee q$ may be $D \models p \vee q$ although neither $D \models p$ nor $D \models q$ (where p and q are conjunctive queries) [9], though for each model α of D, either $\alpha \models p$ or $\alpha \models q$. For instance, let $D = \{p(a), q(b), a < b, q(c), r(d), c < d, r(e), p(f), e < f\}$, $P = \exists xyz[p(x) \wedge q(y) \wedge r(z) \wedge\ x < y < z]$, $Q = \exists xyz[q(x) \wedge r(y) \wedge p(z) \wedge\ x < y < z]$, and $R = \exists xyz[r(x) \wedge p(y) \wedge q(z) \wedge\ x < y < z]$. Then $D \models P \vee Q \vee R$ but neither $D \models P$ nor $D \models Q$ nor $D \models R$.

5 Algorithm generation based on graph-minors

Let $(\Sigma, \leq)$ be a wqo set. A property P on $(\Sigma, \leq)$ is *minor-closed* if $P(x)$ implies $P(y)$ for all $y \in \Sigma$ s.t. $x \leq y$. If P is minor-closed then P has the set of minimal elements in Σ which satisfy P. Since Σ is a wqo set, such minimal elements are finite - if there are infinite minimal elements, regarding them as an infinite sequence this contradicts to the definition of wqo and their minimality. These elements are called *forbidden minors.*

For simplicity, we consider an automatic linear-time algorithm generation based on Higman's lemma. Murthy-Russel's constructive proof [13] gives insight to an actual algorithm detecting forbidden-minors under Curry-Howard isomorphism. Let $(\Sigma, \leq)$ be a cwqo set. Then a set of finite words $(\Sigma^*, \lll)$ on Σ is a wqo set. Consider a minor-closed property P on $(\Sigma^*, \lll)$. We assume Σ is a recursively enumerable set. (Then Σ^* is so. i.e., there exists a recursive function $f : \mathrm{N} \to \Sigma^*$ s.t. $\forall x \in \Sigma\ \exists n\ [f(n) = x]$). Since the wqo property of $\lll$ implies that forbidden-minors for P is finite, if we have *subroutines*

- `P-test(w)`: For $\forall w \in \Sigma^*$, a judgment whether $P(w)$ holds.
- `Min-test(w)`: For $\forall w \in \Sigma^*$ s.t. $P(w)$ holds, a judgment whether w is the minimal element wrt $\lll$. (i.e., $p(w)$ does not hold for $\forall x \lll w$).
- `Fin-test(FM)`: For all finite subset FM of forbidden-minors, a judgment whether $\exists x \in \Sigma - \cup_{\forall m \in FM} m^\circ$ s.t. $P(x)$.

the algorithm below to detect forbidden minors with an enumerating function $f : \mathrm{N} \to \Sigma^*$ always terminates.

```
FM:={}, n=1
  do until Fin-test(FM)
     if P-test(f(n)) and Min-test(f(n)) then add f(n) to FM
     n:=n+1
  od
```

Proposition If the following conditions are satisfied, an algorithm to detect forbidden-minors for the minor-closed property P on Σ^* is constructed.

1. Σ is finite. (Thus, a quasi order $\leq$ is a cwqo.)
2. For $\forall w \in \Sigma^*$, a judgment whether $P(w)$ holds is decidable.
3. For all sequential r.e. σ, a judgment whether $\exists w \in \sigma$ s.t. $P(w)$ is decidable.

Assumption 1 implies the constructions of `f` and `Min-test`. Assumption 2 gives `P-test`. Lemma 3.5 shows that for any FM there exists a set of sequential r.e.'s E s.t. $\Sigma - \bigcup_{\forall m \in FM} m^\circ = \bigcup_{\forall \sigma \in E} \sigma$ (regarding FM as a non-increasing finite sequence). Thus `Fin-test` is reduced to $\bigvee_{\forall \sigma \in E} [\exists x \in \sigma\ P(x)]$ which is a collection of assumption 3.

Example: a disjunctive monadic query on indefinite database

We fix a disjunctive monadic query $Q = Q_1 \vee Q_2 \vee \cdots \vee Q_n$ where Q_i's for $\forall i (\leq n)$ are conjunctive components (i.e., conjunctive monadic queries.) Let $Pred$ be a set of monadic predicate symbols appears in Q and let us set $\Sigma = \mathcal{P}(Pred)$. We use a symbol D for indefinite database as default and let $\mathcal{M}_Q$ be a set of forbidden-minors for the property $D \models Q$.

The idea is basically same to those in the proposition except that the wqo property of $\leq_m$ on $\mathcal{FP}(\Sigma^*)$ is obtained by two steps : the first by Higman's lemma and the second by lemma 3.6. Let $\{m_1, m_2, \cdots, m_k\}$ be a finite subset of $\mathcal{M}_Q$. Lemma 3.5 shows that there exists a finite set E of base r.e.'s s.t. $\mathcal{FP}(\Sigma^*) - m_1^\circ \cup m_2^\circ \cup \cdots \cup m_k^\circ = \bigcup_{\forall \sigma \in E} \mathcal{FP}(\sigma)$. Lemma 3.2 show that for all base r.e. σ there exists a finite set F_σ of sequential r.e.'s s.t. $\sigma = \bigcup_{\forall \theta \in F_\sigma} \theta$. Thus $\mathcal{FP}(\Sigma^*) - m_1^\circ \cup m_2^\circ \cup \cdots \cup m_k^\circ = \bigcup_{\forall \sigma \in E} \mathcal{FP}(\sigma) = \bigcup_{\forall \sigma \in E} \mathcal{FP}(\bigcup_{\forall \theta \in F_\sigma} \theta)$.

Since Σ is finite, an enumeration function `f` and subroutines `Min-test` are constructed. `P-test` is also decidable. The difficult one is `Fin-test`. `Fin-test` is reduced to a collection of simpler subproblems. That is, the judgment $(\mathcal{FP}(\Sigma^*) - m_1^\circ \cup m_2^\circ \cup \cdots \cup m_k^\circ) \cap \mathcal{M}_Q \neq \phi$ is reduced to the judgment $\exists \sigma \in E$ s.t. $\mathcal{FP}(\sigma) \cap \mathcal{M}_Q \neq \phi$, and further reduced to the judgment in terms of $\theta \in F_\sigma$. An actual form of `Fin-test` is given in theorem 5.1. We first prepare several abbreviate notations.

Definition 5.1 Let $c, c_1, c_2, \cdots, c_l \in \Sigma$ and let $C_1, C_2, \cdots, C_l \in \mathcal{FP}(\Sigma)$. Then

$$\begin{aligned} [\![C_1 \cdots C_l]\!] &= \bigwedge_{\forall i \forall c_i \in C_i} [c_1 \cdots c_l] \\ [c_1 \cdots c_l] &= \exists x_1 \cdots x_l\ [p_1(x_1) \wedge \cdots \wedge p_l(x_l) \wedge\ x_1 < \cdots < x_l] \end{aligned}$$

where $p_i(x) = \bigwedge_{p \in c_i} p(x_i)$ for $i \leq l$. We also denote

$$\mathcal{PS}(c \wedge \neg c_1 \wedge \cdots \wedge \neg c_l) = \{x \in \mathcal{FP}(\Sigma) \mid c \subseteq x\ \wedge\ c_1 \not\subseteq x\ \wedge\ \cdots\ \wedge\ c_l \not\subseteq x\}.$$

Definition 5.2 Let $A = a_1, a_2, \cdots, a_k$ be a finite non-increasing sequence of Σ and let $b \in \Sigma$. Let $(b - A)$ be a constant expression, let $(\Sigma - A)^*$ be a starred expression, and let $\theta = \theta_1 \theta_2 \cdots \theta_l$ be a sequential r.e. where θ_i's are either constant or stared r.e.'s. We set $\Delta(Q)$ as the sum of lengths of all paths of all conjunctive components of Q. We define a function Ψ as

$$\begin{cases} \psi((b - A)) &= \mathcal{PS}(b \wedge \neg a_1 \wedge \neg a_2 \wedge \cdots \wedge \neg a_k) \\ \psi((\Sigma - A)^*) &= \mathcal{PS}(\neg a_1 \wedge \neg a_2 \wedge \cdots \wedge \neg a_k) \\ \Psi(\theta, n) &= [\![\psi(\theta_1)^{\alpha_1} \cdots \psi(\theta_l)^{\alpha_l}]\!] \end{cases}$$

where for $\forall i \leq l$ $\alpha_i = 1$ if θ_i is a constant expression and $\alpha_i = n$ if θ_i is a starred expression.

Lemma 5.1 (1) Let θ be a sequential r.e. on Σ. Then $\forall w \in \theta\ \exists n$ s.t. $w \lll w'$ for some path $w' \in Paths(\Psi(\theta, n))$.
(2) Let θ be a sequential r.e. on Σ. Then $\forall S \subseteq \mathcal{FP}(\theta)\ \exists n$ s.t. $S \leq_m Paths(\Psi(\theta, n))$.
(3) Let σ be a base r.e. on Σ^*. Then $\forall T \in \mathcal{FP}(\sigma)\ \exists n$ s.t. $T \leq_m Paths(\bigwedge_{\forall \theta \in F_\sigma} \Psi(\theta, n))$.

Lemma 5.2 Let Q be a disjunctive monadic query and let F be a finite set of sequential r.e.'s. Then for $\forall n \geq \Delta(Q)$, $\bigwedge_{\theta \in F} \Psi(\theta, \Delta(Q)) \models Q$ if and only if $\bigwedge_{\theta \in F} \Psi(\theta, n) \models Q$.

Proof Since each model of Q relates to at most $\Delta(S)$ points on a linearly ordered domain, for detecting the validity of Q, the n-times repetition of a starred r.e. in each $\theta \in E$ is equivalent to the $\Delta(Q)$-times repetition. ■

Theorem 5.1 Let E be a finite set of base r.e.'s and let F_σ be a finite set of sequential r.e. for $\sigma \in E$ s.t. $\mathcal{FP}(\Sigma^*) - m_1^\circ \cup m_2^\circ \cup \cdots \cup m_k^\circ = \cup_{\forall \sigma \in E} \mathcal{FP}(\sigma) = \cup_{\forall \sigma \in E} \mathcal{FP}(\cup_{\forall \theta \in F_\sigma} \theta)$. Then, $\exists D \in \mathcal{FP}(\Sigma^*) - m_1^\circ \cup m_2^\circ \cup \cdots \cup m_k^\circ\ [D \models Q]$ if and only if $\exists \sigma \in E\ [\wedge_{\theta \in F_\sigma} \Psi(\theta, \Delta(Q)) \models Q]$.

6 Conclusion

This paper discussed an automatic generation of a linear-time algorithm based on Murthy-Russel's constructive proof of Higman's lemma. The example is a monadic disjunctive query processing on indefinite database (indicated in [9]). Its forbidden-minors is shown to be automatically detected. This is only the first step of research on this field - there remain many problems. For instance,

- Current relation between a constructive proof and automatic generation is not so clear. A suitable logical system and more direct generator extraction method are required.
- The current automatic generation technique is very inefficient. A smarter method and an optimization technique are required from practical viewpoint.
- Currently we apply only Higman's lemma, etc. A generalization to a more complex structure is desired. For instance, a constructive proof of Kruskal's theorem has been given [5].

We hope Wagner's conjecture to have a constructive proof, and hope that it enables us to develop an algorithm generation on general finite graphs.

Acknowledgments

The author would like to thank Dr. Ronald van der Meyden (NTT Basic Research Labo.) for private lectures on indefinite database theory.

References

[1] N.Dershowitz and Z.Manna, "Proving termination with multiset orderings", CACM, 22, 8, pp.465-476 (1979)

[2] N.Dershowitz, "Ordering for term-rewriting systems", Proc. 20th IEEE FOCS, pp.123-131 (1979)

[3] M.R.Fellows and M.A.Langston, "Nonconstructive tools for proving polynomial-time decidability", JACM, 35, 3, pp.727-739 (1988)

[4] J.H.Gallier, "What's so special about Kruskal's theorem and the ordinal Γ_0 ? A survey of some results in proof theory", Annals of Pure and Applied Logic 53, pp.199-260 (1991)

[5] A.Gupta, "A constructive proof that tree are well-quasi-ordered under Minors", 2nd Int. sympo. Tver, logical foundations of computer science, *LNCS 620*, pp.174-185 (1992)

[6] H.Friedman, "Classically and intuitionistically provably recursive functions", Higher Set Theory (G.H.Müller and D.S.Scott eds.), *LNM699*, pp.21-28 (1978)

[7] G.Higman, "Ordering by divisibility in abstract algebras", Proc. London Math. Soc., 3, 2, pp.326-336 (1952)

[8] J.B.Kruskal, "Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture", Trans. American Math. Soc., 95, pp.210-225 (1960)

[9] R.van der Meyden, "The complexity of querying indefinite data about linearly ordered domains (preliminary version)", Proc. 11th ACM PODS, pp.331-345 (1992)

[10] R.van der Meyden, *Private communications*, February 1993

[11] E.C.Milner, "Basic wqo- and bqo- theory", Graphs and Orders (I.Rival ed.), pp.487-592, D.Redidel Publishing Germany (1985)

[12] R.Motwani, A.Raghunathan, and H.Saran, "Constructive results from Graph Minors: Linkless embeddings", Proc. 29th IEEE FOCS, pp.398-409 (1988)

[13] C.R.Murthy and J.R.Russell, "A constructive proof of Higman's lemma", Proc. 5th IEEE LICS, pp.257-267 (1990)

[14] C.ST.J.A.Nash-Williams, "On well-quasi-ordering finite trees", Proc. Cambridge Phil. Soc., 59, pp.833-835 (1963)

[15] C.ST.J.A.Nash-Williams, "On well-quasi-ordering transfinite sequences", Proc. Cambridge Phil. Soc., 61, pp.33-39 (1965)

[16] C.ST.J.A.Nash-Williams, "On well-quasi-ordering infinite trees", Proc. Cambridge Phil. Soc., 61, pp.697-720 (1965)

[17] F.Richman and G.Stolzenberg, "Well quasi-ordered sets", Advances in Mathematics, 97, pp.145-153 (1993)

[18] N.Robertson and P.D.Seymour, "Graph Minors XV. Wagner's Conjecture", *preprint* (1988)

[19] S.G.Simpson, "Ordinal numbers and the Hilbert basis theorem", JSL, 53, 3, pp.961-974 (1988)

[20] R.Thomas, "A counter-example to 'Wagner's conjecture' for infinite graphs", Proc. Cambridge Phil. Soc., 103, pp.55-57 (1988)