MPRI 2-9-1: Well-Quasi-Orders for Algorithms

Final Exam

To be returned no later than Friday Dec. 4th 2020, 09:00 AM.

All questions below are subjectively ranked from (*) to (***) stars depending on whether they are more or less difficult, hence require more careful and detailed answers.

This exam consists of 2 independent exercises. When answering the questions, be rigorous in form and complete in reasoning. Rigour does not mean length: you can omit trivial justifications (but do not omit corner cases!).

I. An ordering on strings.

Let \( A = \{a_1, \ldots, a_k\} \) be a finite alphabet where letters are linearly ordered by \( a_1 <_A a_2 <_A \cdots <_A a_k \).

For two words \( x, y \in A^* \), we say that \( x \) can be simplified into \( y \), written \( x \rightarrow y \), when \( x \) is some \( x' \cdot a \cdot a' \cdot x'' \) and \( y \) is \( x' \cdot a'' \cdot x'' \) for \( a'' = \max(a, a') \). We write \( x \overset{*}{\rightarrow} y \) when there is a sequence \( x = x_0 \rightarrow x_1 \rightarrow x_2 \cdots \rightarrow x_m = y \) of simplifications from \( x \) to \( y \) (note that \( m \) is possible).

For example, if \( A \) is the classic 26-letter Latin alphabet ordered with \( a < b < \cdots < z \), then \( baba \rightarrow bba \) and \( baba \rightarrow bab \) but \( baba \not\rightarrow baa \). Observe that \( baba \overset{*}{\rightarrow} b \) while \( baba \not\rightarrow \epsilon \), where as usual \( \epsilon \) denotes the empty word.

Finally we define a binary relation \( \leq \) over \( A^* \) by \( x \leq y \iff y \overset{*}{\rightarrow} x \).

(*) 1. Prove that \((A^*, \leq)\) is a partial ordering.

**Solution:**
By definition \( \rightarrow \) is reflexive and transitive. It is antisymmetric since \( x \rightarrow y \) implies that \( y \)

(***) 2. Prove that it is a wqo.

**Solution:**
We prove that \((A^*, \leq)\) is wqo by induction on the size of \( A \). The result is clear when \( |A| = 0 \) since then \( A^* = \{\epsilon\} \) is finite.

For the inductive step, we write \( z \) for the maximal letter in \( A \) and let \( B \overset{\text{def}}{=} A \setminus \{z\} \). By induction hypothesis, \((B^*, \leq)\) is a wqo. Let us now consider an arbitrary infinite sequence \((x_i)_{i \in \mathbb{N}}\) over \( A^* \): our task is to prove that it contains an increasing pair. If \((x_i)_{i \in \mathbb{N}}\) contains infinitely many words from \( B^* \), then it contains an increasing pair by the induction hypothesis.

So we may assume that a suffix \((x_i)_{i \geq N}\) contains only words with at least one \( z \) and we factor any such \( x_i \) under the form \( x_i = x_i^0 \cdot z \cdot x_i^2 \cdot \cdots \cdot z \cdot x_i^{\ell_i} \) with \( \ell_i > 0 \) and \( x_i^j \in B^* \) for all \( j = 0, \ldots, \ell_i \).

Since \( B^* \) is a wqo, we can extract an infinite subsequence \((x_i)_{i \in I}\) with \( x_i^0 \leq x_j^0 \) and \( x_i^{\ell_i} \leq x_j^{\ell_j} \) for all \( i < j \).

Furthermore, \(((B^*)^*, \leq_s)\) is wqo too by Higman’s Lemma, so we can further extract a pair \( x_i, x_{i'} \) with \( i < i' \) and

\[
(x_i^1, x_i^2, \ldots, x_i^{\ell_i-1}) \leq_s (x_{i'}^1, x_{i'}^2, \ldots, x_{i'}^{\ell_{i'}-1})
\]

Note that this entails \( \ell_i \leq \ell_{i'} \) and the existence of an increasing map \( f : \{1, \ldots, \ell_i\} \rightarrow \{1, \ldots, \ell_{i'}\} \) such that \( x_i^j \leq x_{i'}^{f(j)} \) when \( 1 \leq j < \ell_i \).

We now claim that \( x_i \overset{*}{\rightarrow} x_{i'} \) for this we rely on the fact that \( \leq_s \) is a precongruence, i.e.,

\[
x \overset{*}{\rightarrow} y \quad \text{entails} \quad u x v \overset{*}{\rightarrow} u y v.
\]

We know that \( x_i^0 \overset{*}{\rightarrow} x_{i'}^0, x_i^j \overset{*}{\rightarrow} x_{i'}^{f(j)}, \) and \( x_i^{(j)} \overset{*}{\rightarrow} x_{i'}^{(j)} \) when
1 \leq j < \ell_i. Furthermore, if \( j \) is not in the image of \( f \), we can use \( z x_i^j z \to z \) to make the \( x_i^j \) factor disappear. Gluing these rewrite steps together, we easily show \( x_i \to x_i \), i.e., \( x_i \leq x_i^j \).

(**) 3. We consider the special case where \( A = \{a, b, c\} \) with \( a < b < c \), and define some subsets \( L_1, \ldots \) of \( A^* \) via standard regular expressions:

\[
L_1 = a^* c^* a^*, \quad L_2 = a^+ c^+ b^*, \quad L_3 = a^+ b^* c, \quad L_4 = a^* c^* b^* c^+, \quad L_5 = c^+ a^* c^+.
\]

Which of the \( L_i \)'s are ideals of \( (A^*, \leq) \)? Justify your answers.

Solution:
None of the \( L_i \)'s are ideals:

\( L_1 \) is not directed since it contains \( \epsilon \) and \( a \), two words with no common upper bound.

\( L_2 \) and \( L_3 \) are not downward-closed since they contain \( ac \) but not \( c \).

\( L_4 \) is not directed since it contains \( bc \) and \( bcb \), two words with no common upper bound.

\( L_5 \) is not downward-closed since it contains \( cc \) but not \( c \).

II. Computing \( L_{g,A}(n) \).
For this exercise, we reuse the notations and definitions from Chapter 2 of the lecture notes.

Let \( (A, \leq, |\.|_A) \) be a normed wqo, \( g : \mathbb{N} \to \mathbb{N} \) a control function, and \( n \in \mathbb{N} \) an initial norm. Let \( a_0, a_1, \ldots, a_{\ell-1} \) be a \((g, n)\)-controlled bad sequence of maximal length, i.e., such that \( \ell = L_{g,A}(n) \).

To simplify notations, we use \( A_i \) to denote \((A/a_0/\cdots/a_{i-1}) \) and \( A_i^g \) to denote \((A_i)_{\leq g(n)} \). In order to simplify proofs, we further assume that \( \leq \) is antisymmetric.

(**) 4. Prove that, for all \( 0 \leq i < \ell, a_i \) is a maximal element of \( A_i^g \).

Solution:
Assume, by way of contradiction, that the claim is false and let \( i \) be an index such that \( a_i \) is not maximal in \( A_i^g \). This means that \( A_i^g \) contains some \( a > a_i \).

We now claim that the sequence \( a_0, a_1, \ldots, a_{i-1}, a, a_i, \ldots, a_{\ell-1} \) obtained by inserting \( a \) before \( a_i \) is a \((g, n)\)-controlled bad sequence. Since it has length \( \ell + 1 \), we reach a contradiction so our assumption that \( a_i \) was not maximal is false.

To prove the claim we first observe that the new sequence is \((g, n)\)-controlled since \(|a| \leq g' (n) \) and since the other elements never appear at an earlier position compared to the original sequence.

To see that the new sequence is bad let assume that it has an increasing pair. This pair must involve the new element \( a \) otherwise it would already be an increasing pair in the original sequence, which is impossible. So it is some \( a_j \leq a \) for some \( j < i \), or \( a \leq a_i \), or some \( a \leq a_j \) for some \( j > i \). But the first case is impossible since \( a \) belongs to \( A/a_0/a_i/\cdots/a_{i-1} \), hence to \( A/a_j \upharpoonright A \not\supseteq a_j a_i \leq a_j \). The second case is impossible since \( a > a_i \). And the third case is impossible since \( a \leq a_j \) would entail \( a_i \leq a_j \), contradicting the fact that the original sequence is bad.

(**) 5. Assume that \( a \) is an element of \( A_i^g \) that is maximal in \( A_i \) (not just in \( A_i^g \)).

Prove that there exists a \((g, n)\)-controlled bad sequence of maximal length that starts with \( a_0, a_1, \ldots, a_{i-1}, a \).

Solution:
(1) We first claim that \( a \) necessarily appears in the sequence as some \( a_j \) with \( j \geq i \).

To see this, recall that the sequence of residuals satisfies \( A = A_0 \supseteq A_1 \supseteq A_2 \cdots \supseteq A_\ell = \emptyset \). We can thus define \( j \) as the first index with \( a \not\in A_j \), entailing \( a \geq a_j \). Necessarily \( j \geq i \).
since $a \in A_i$. Now assume $a > a_j$; this means that $a_j$ is not maximal in $A_i^g$, contradicting what we proved in the previous question. Therefore $a = a_j$ and indeed $a$ appears in the sequence.

(2) Then we claim that the sequence $a_0, a_1, \ldots, a_{i-1}, a, a_i, a_{i+1}, \ldots, a_{j-1}, a_j, \ldots$ obtained by moving $a$ at position $i$ fulfils the claim. It is clearly $(g, n)$-controlled. To see that it is bad we reason that any increasing pair in the new sequence that would not appear in the original sequence must be some $a \leq a_k$ with $i \leq k \neq j$. But this would contradict the maximality of $a$ in $A_i$.

(**) 6. Using the results proven in the previous two questions, compute $L_{g, A}(2)$ in the following situation:

- $A = \mathbb{N}^2$ with $|(a, b)|_A \overset{\text{def}}{=} \max(a, b)$,
- $g$ is the successor function $x \mapsto x + 1$.

**Solution:**

We build a controlled bad sequence $s = (a_0, a_1, \ldots)$ of maximal length, relying on Questions 4 and 5 to simplify the exhaustive search. $s$ must start with $(2, 2)$, since this is the only maximal element in $A_0^g = (\mathbb{N}^2)_{\leq 2}$.

Then the residual $A_1^g$ has two maximal elements, $(1, 3)$ and $(3, 1)$, so $s$ must continue with one of these. Since the two options are completely symmetrical, we can assume $a_1 = (1, 3)$ without loss of generality.

The residual $A_2^g$ now has three maximal elements: $(0, 4), (1, 2)$ and $(4, 1)$. Since $(1, 2)$ is maximal not just in $A_2^g$ but more generally in $A_2$, we can safely pick $(1, 2)$ and extend our sequence.

Now we have to choose between the two maximal elements of $A_3^g$: $(0, 5)$ and $(5, 1)$. If we choose $(0, 5)$ the rest of the sequence is completely determined by our principles and we end up with

$$s_1 = (2, 2), (1, 3), (1, 2), (0, 5), (0, 4), (0, 3), (0, 2), (9, 1), (8, 1), \ldots, (0, 1), (18, 0), (17, 0), \ldots, (0, 0).$$

If we choose $(5, 1)$, the sequence necessarily continues

$$s_2 = (2, 2), (1, 3), (1, 2), (5, 1), (4, 1), (3, 1), (2, 1), (1, 1),$$

and we must now choose between $(0, 10)$ and $(10, 0)$, the two maximal elements of $A_4^g$. Since these choices are completely symmetrical, we can continue with $(0, 10)$ without any loss of generality, ending up with

$$s_2 = (2, 2), (1, 3), (1, 2), (5, 1), (4, 1), (3, 1), (2, 1), (1, 1), (0, 10), (0, 9), \ldots, (0, 1), (20, 0), (19, 0), \ldots, (0, 0).$$

The second option leads to a longer sequence. Since we have exhausted all possibilities, we can conclude with $L_{\text{Succ}, \mathbb{N}^2}(2) = 39$. 

Page 3