Algorithmic Aspects of WQO (Well-Quasi-Ordering) Theory Part II: Algorithmic Applications of WQOs

Philippe Schnoebelen

LSV, CNRS & ENS Cachan

Chennai Mathematical Institute, Jan. 2017

Based on joint work with Sylvain Schmitz, Prateek Karandikar, K. Narayan Kumar, Alain Finkel, ..

Lecture notes & exercises available via www.lsv.ens-cachan.fr/~phs

IF YOU MISSED PART I

 (X, \leqslant) is a well-quasi-ordering (a wqo) if any <u>infinite</u> sequence $x_0, x_1, x_2...$ over X contains an increasing pair $x_i \leqslant x_j$ (for some i < j)

Examples.

- 1. $(\mathbb{N}^k, \leq_{\times})$ is a wqo (Dickson's Lemma) where, e.g., $(3,2,1) \leq_{\times} (5,2,2)$ but $(1,2,3) \leq_{\times} (5,2,2)$
- (Σ*,≤*) is a wqo (Higman's Lemma) where, e.g., abc ≤* bacbc but cba ≤* bacbc

Objectives for today's course:

- See algorithms that rely on wqos: verification of WSTS's
- Reduce complexity analysis to bounds on bad sequences

IF YOU MISSED PART I

 (X, \leqslant) is a well-quasi-ordering (a wqo) if any <u>infinite</u> sequence $x_0, x_1, x_2...$ over X contains an increasing pair $x_i \leqslant x_j$ (for some i < j)

Examples.

- 1. $(\mathbb{N}^k, \leq_{\times})$ is a wqo (Dickson's Lemma) where, e.g., $(3,2,1) \leq_{\times} (5,2,2)$ but $(1,2,3) \leq_{\times} (5,2,2)$
- (Σ*,≤*) is a wqo (Higman's Lemma) where, e.g., abc ≤* bacbc but cba ≤* bacbc

Objectives for today's course:

- See algorithms that rely on wqos: verification of WSTS's
- Reduce complexity analysis to bounds on bad sequences

OUTLINE FOR PART II

- Well-structured transition systems (WSTS's)
- Deciding Termination
- Deciding Coverability
- (in lecture notes only:) other wqo-based algorithms: other termination proofs, relevance logic, Karp-Miller trees, ...

All of these are actual examples of algorithms that terminate thanks to wqo-theoretical arguments

Question for Part III. terminate in how many steps exactly?

OUTLINE FOR PART II

- Well-structured transition systems (WSTS's)
- Deciding Termination
- Deciding Coverability
- (in lecture notes only:) other wqo-based algorithms: other termination proofs, relevance logic, Karp-Miller trees, ..

All of these are actual examples of algorithms that terminate thanks to wqo-theoretical arguments

Question for Part III. terminate in how many steps exactly?

In program verification, wqo's appear prominently under the guise of WSTS.

Def. A WSTS is a system $(S, \rightarrow, \leqslant)$ where

- 1. (S, \rightarrow) with $\rightarrow \subseteq S \times S$ is a transition system
- 2. the set of states (S, \leq) is wqo, and
- 3. the transition relation is compatible with the ordering (also called "monotonic"): $s \to t$ and $s \leqslant s'$ imply $s' \to t'$ for some $t' \ge t$



A run of M: $(\ell_0, 0, 1, 4) \rightarrow (\ell_1, 1, 1, 4) \rightarrow (\ell_2, 1, 0, 4) \rightarrow (\ell_3, 1, 0, 0)$

Ordering states: $(\ell_1, 0, 0, 0) \leq (\ell_1, 0, 1, 2)$ but $(\ell_1, 0, 0, 0) \leq (\ell_2, 0, 1, 2)$. This is wqo as a product of wqo's: $(Loc, =) \times (\mathbb{N}^3, \leq_{\times})$

Compatibility: easily checked when guards are upward-closed and assignments are monotonic functions of the variables.

NB. Other updates can be considered as long as they are monotonic. Extending guards require using a finer ordering.



A run of M: $(\ell_0, 0, 1, 4) \rightarrow (\ell_1, 1, 1, 4) \rightarrow (\ell_2, 1, 0, 4) \rightarrow (\ell_3, 1, 0, 0)$

Ordering states: $(\ell_1, 0, 0, 0) \leq (\ell_1, 0, 1, 2)$ but $(\ell_1, 0, 0, 0) \leq (\ell_2, 0, 1, 2)$. This is wqo as a product of wqo's: $(Loc, =) \times (\mathbb{N}^3, \leq_{\times})$

Compatibility: easily checked when guards are upward-closed and assignments are monotonic functions of the variables.

NB. Other updates can be considered as long as they are monotonic. Extending guards require using a finer ordering.



A run of M: $(\ell_0, 0, 1, 4) \rightarrow (\ell_1, 1, 1, 4) \rightarrow (\ell_2, 1, 0, 4) \rightarrow (\ell_3, 1, 0, 0)$

Ordering states: $(\ell_1, 0, 0, 0) \leq (\ell_1, 0, 1, 2)$ but $(\ell_1, 0, 0, 0) \leq (\ell_2, 0, 1, 2)$. This is wqo as a product of wqo's: $(Loc, =) \times (\mathbb{N}^3, \leq_{\times})$

Compatibility: easily checked when guards are upward-closed and assignments are monotonic functions of the variables.

NB. Other updates can be considered as long as they are monotonic. Extending guards require using a finer ordering.



A run of M: $(\ell_0, 0, 1, 4) \rightarrow (\ell_1, 1, 1, 4) \rightarrow (\ell_2, 1, 0, 4) \rightarrow (\ell_3, 1, 0, 0)$

Ordering states: $(\ell_1, 0, 0, 0) \leq (\ell_1, 0, 1, 2)$ but $(\ell_1, 0, 0, 0) \leq (\ell_2, 0, 1, 2)$. This is wqo as a product of wqo's: $(Loc, =) \times (\mathbb{N}^3, \leq_{\times})$

Compatibility: easily checked when guards are upward-closed and assignments are monotonic functions of the variables.

NB. Other updates can be considered as long as they are monotonic. Extending guards require using a finer ordering.



A run of M: $(\ell_0, 0, 1, 4) \rightarrow (\ell_1, 1, 1, 4) \rightarrow (\ell_2, 1, 0, 4) \rightarrow (\ell_3, 1, 0, 0)$

Ordering states: $(\ell_1, 0, 0, 0) \leq (\ell_1, 0, 1, 2)$ but $(\ell_1, 0, 0, 0) \leq (\ell_2, 0, 1, 2)$. This is wqo as a product of wqo's: $(Loc, =) \times (\mathbb{N}^3, \leq_{\times})$

Compatibility: easily checked when guards are upward-closed and assignments are monotonic functions of the variables.

NB. Other updates can be considered as long as they are monotonic. Extending guards require using a finer ordering.

SOME WSTS'S: RELATIONAL AUTOMATA



One does not use \leqslant_{\times} to compare states!! Rather

 $\begin{aligned} (a_1, \dots, a_k) \leqslant_{\text{sparse}} (b_1, \dots, b_k) \\ & \Leftrightarrow \forall i, j = 1, \dots, k : (a_i \leqslant a_j \text{ iff } b_i \leqslant b_j) \land (|a_i - a_j| \leqslant |b_i - b_j|). \end{aligned}$

Fact. $(\mathbb{Z}^k, \leq_{\text{sparse}})$ is wqo

$$\begin{split} (\ell,a_1,\ldots,a_k) \leqslant (\ell',b_1,\ldots,b_k) \stackrel{\text{def}}{\Leftrightarrow} \\ \textbf{Compatibility: We use} \\ \ell = \ell' \wedge (a_1,\ldots,a_k,-1,10) \leqslant_{\text{sparse}} (b_1,\ldots,b_k,-1,10) \,. \end{split}$$

SOME WSTS'S: RELATIONAL AUTOMATA



One does not use \leq_{\times} to compare states!! Rather

$$\begin{aligned} &(a_1, \dots, a_k) \leqslant_{\text{sparse}} (b_1, \dots, b_k) \\ &\Leftrightarrow \forall i, j = 1, \dots, k : \left(a_i \leqslant a_j \text{ iff } b_i \leqslant b_j \right) \land \left(|a_i - a_j| \leqslant |b_i - b_j| \right). \end{aligned}$$

Fact. $(\mathbb{Z}^k, \leq_{sparse})$ is wqo

$$\begin{split} (\ell,a_1,\ldots,a_k) \leqslant (\ell',b_1,\ldots,b_k) \stackrel{\text{def}}{\Leftrightarrow} \\ \textbf{Compatibility:} \ \text{We use} \\ \ell = \ell' \wedge (a_1,\ldots,a_k,-1,10) \leqslant_{\text{sparse}} (b_1,\ldots,b_k,-1,10) \,. \end{split}$$

SOME WSTS'S: RELATIONAL AUTOMATA



One does not use \leq_{\times} to compare states!! Rather

$$\begin{aligned} &(a_1, \dots, a_k) \leqslant_{\text{sparse}} (b_1, \dots, b_k) \\ &\stackrel{\text{def}}{\Leftrightarrow} \forall i, j = 1, \dots, k : \left(a_i \leqslant a_j \text{ iff } b_i \leqslant b_j \right) \land \left(|a_i - a_j| \leqslant |b_i - b_j| \right). \end{aligned}$$

Fact. $(\mathbb{Z}^k, \leq_{sparse})$ is wqo

$$\begin{split} (\ell, a_1, \dots, a_k) \leqslant (\ell', b_1, \dots, b_k) & \stackrel{\text{def}}{\Leftrightarrow} \\ & \textbf{Compatibility: We use} \\ \ell = \ell' \wedge (a_1, \dots, a_k, -1, 10) \leqslant_{\text{sparse}} (b_1, \dots, b_k, -1, 10) \,. \end{split}$$

SOME WSTS'S: RELATIONAL AUTOMATA



One does not use \leq_{\times} to compare states!! Rather

$$\begin{aligned} &(a_1, \dots, a_k) \leqslant_{\text{sparse}} (b_1, \dots, b_k) \\ &\stackrel{\text{def}}{\Leftrightarrow} \forall i, j = 1, \dots, k : \left(a_i \leqslant a_j \text{ iff } b_i \leqslant b_j \right) \land \left(|a_i - a_j| \leqslant |b_i - b_j| \right). \end{aligned}$$

Fact. $(\mathbb{Z}^k, \leq_{sparse})$ is wqo

$$\begin{split} (\ell,a_1,\ldots,a_k) \leqslant (\ell',b_1,\ldots,b_k) \stackrel{\text{def}}{\Leftrightarrow} \\ \textbf{Compatibility: We use} \\ \ell = \ell' \wedge (a_1,\ldots,a_k,-1,10) \leqslant_{\text{sparse}} (b_1,\ldots,b_k,-1,10) \,. \end{split}$$

SOME WSTS'S: LCS / LOSSY CHANNEL SYSTEMS



A configuration $\sigma = (\ell_1, \ell_2, w_1, w_2)$ with $w_i \in \Sigma^*$. E.g., $w_1 = \text{hup.ack.ack}$.

Reliable steps: $\sigma \rightarrow_{\mathsf{rel}} \rho$ read in front of channels, write at end (FIFO)

Lossy steps: messages may be lost nondeterministically $\sigma \rightarrow \sigma' \stackrel{\text{def}}{\Leftrightarrow} \sigma \sqsupseteq \rho \rightarrow_{\text{rel}} \rho' \sqsupseteq \sigma' \text{ for some } \rho, \rho'$ where (S, \sqsubseteq) is the wqo $(Loc_1, =) \times (Loc_2, =) \times (\Sigma_{c_1}^*, \leqslant_*) \times (\Sigma_{c_2}^*, \leqslant_*)$

A model useful for concurrent protocols but also timed automata, metric temporal logic, products of modal logics, ...

SOME WSTS'S: LCS / LOSSY CHANNEL SYSTEMS



A configuration $\sigma = (\ell_1, \ell_2, w_1, w_2)$ with $w_i \in \Sigma^*$. E.g., $w_1 = \text{hup.ack.ack}$.

Reliable steps: $\sigma \rightarrow_{\mathsf{rel}} \rho$ read in front of channels, write at end (FIFO)

Lossy steps: messages may be lost nondeterministically $\sigma \rightarrow \sigma' \stackrel{\text{def}}{\Leftrightarrow} \sigma \sqsupseteq \rho \rightarrow_{\text{rel}} \rho' \sqsupseteq \sigma' \text{ for some } \rho, \rho'$ where (S, \sqsubseteq) is the wqo $(Loc_1, =) \times (Loc_2, =) \times (\Sigma_{c_1}^*, \leqslant_*) \times (\Sigma_{c_2}^*, \leqslant_*)$

A model useful for concurrent protocols but also timed automata, metric temporal logic, products of modal logics, ...

SOME WSTS'S: LCS / LOSSY CHANNEL SYSTEMS



A configuration $\sigma = (\ell_1, \ell_2, w_1, w_2)$ with $w_i \in \Sigma^*$. E.g., $w_1 = \text{hup.ack.ack}$.

Reliable steps: $\sigma \rightarrow_{\mathsf{rel}} \rho$ read in front of channels, write at end (FIFO)

Lossy steps: messages may be lost nondeterministically $\sigma \rightarrow \sigma' \stackrel{\text{def}}{\Leftrightarrow} \sigma \sqsupseteq \rho \rightarrow_{\text{rel}} \rho' \sqsupseteq \sigma' \text{ for some } \rho, \rho'$ where (S, \sqsubseteq) is the wqo $(Loc_1, =) \times (Loc_2, =) \times (\Sigma_{c_1}^*, \leqslant_*) \times (\Sigma_{c_2}^*, \leqslant_*)$

A model useful for concurrent protocols but also timed automata, metric temporal logic, products of modal logics, ...

Termination is the question, given a TS $(S, \rightarrow, ...)$ and a state $s_{init} \in S$, whether there are no infinite runs starting from s_{init}

Lem. [Finite Witnesses for Infinite Runs] A WSTS (S, \rightarrow, \leq) has an infinite run from s_{init} iff it has a finite run from s_{init} that is a good sequence.

Recall: $s_0, s_1, s_2, \dots, s_n$ is good $\stackrel{\text{def}}{\Leftrightarrow}$ there exist i < j s.t. $s_i \leq s_j$

Coro. One can decide Termination for a WSTS by enumerating all finite runs from s_{init} and reject when/if a good sequence is found.

NB: This requires some minimal effectiveness assumptions on the WSTS, e.g., that the ordering is decidable

Termination is the question, given a TS $(S, \rightarrow, ...)$ and a state $s_{init} \in S$, whether there are no infinite runs starting from s_{init}

Lem. [Finite Witnesses for Infinite Runs] A WSTS (S, \rightarrow, \leq) has an infinite run from s_{init} iff it has a finite run from s_{init} that is a good sequence.

```
Recall: s_0, s_1, s_2, \dots, s_n is good \stackrel{\text{def}}{\Leftrightarrow} there exist i < j s.t. s_i \leq s_j
```

Coro. One can decide Termination for a WSTS by enumerating all finite runs from *s*_{init} and reject when/if a good sequence is found.

NB: This requires some minimal effectiveness assumptions on the WSTS, e.g., that the ordering is decidable

Termination is the question, given a TS $(S, \rightarrow, ...)$ and a state $s_{init} \in S$, whether there are no infinite runs starting from s_{init}

Lem. [Finite Witnesses for Infinite Runs]

A WSTS (S, \rightarrow, \leq) has an infinite run from s_{init} iff it has a finite run from s_{init} that is a good sequence.

Recall: $s_0, s_1, s_2, \dots, s_n$ is good $\stackrel{\text{def}}{\Leftrightarrow}$ there exist i < j s.t. $s_i \leq s_j$

 $\begin{array}{l} \textbf{Proof.} \Rightarrow: \text{the infinite run contains an increasing pair} \\ \Leftarrow: \text{good finite run } s_0 \xrightarrow{*} s_i \xrightarrow{+} s_j \text{ can be extended by simulating } s_i \xrightarrow{+} s_j \\ \text{from above: } s_j \xrightarrow{+} s_{2j-i}, \text{ then } s_{2j-i} \xrightarrow{+} s_{3j-2i}, \text{ etc.} \end{array}$

Coro. One can decide Termination for a WSTS by enumerating all finite runs from *s*_{init} and reject when/if a good sequence is found.

NB: This requires some minimal effectiveness assumptions on the WSTS, e.g., that the ordering is decidable

Termination is the question, given a TS $(S, \rightarrow, ...)$ and a state $s_{init} \in S$, whether there are no infinite runs starting from s_{init}

Lem. [Finite Witnesses for Infinite Runs] A WSTS (S, \rightarrow, \leq) has an infinite run from s_{init} iff it has a finite run from s_{init} that is a good sequence.

Recall: $s_0, s_1, s_2, \dots, s_n$ is good $\stackrel{\text{def}}{\Leftrightarrow}$ there exist i < j s.t. $s_i \leqslant s_j$

Proof. \Rightarrow : the infinite run contains an increasing pair \Leftarrow : good finite run $s_0 \stackrel{*}{\rightarrow} s_i \stackrel{+}{\rightarrow} s_j$ can be extended by simulating $s_i \stackrel{+}{\rightarrow} s_j$ from above: $s_j \stackrel{+}{\rightarrow} s_{2j-i}$, then $s_{2j-i} \stackrel{+}{\rightarrow} s_{3j-2i}$, etc.

Coro. Termination is co-r.e.

Since it is also r.e. (for finitely branching systems), it is decidable.

Coro. One can decide Termination for a WSTS by enumerating all finite runs from s_{init} and reject when/if a good sequence is found.

NB: This requires some minimal effectiveness assumptions on the WSTS, e.g., that the ordering is decidable

Termination is the question, given a TS $(S, \rightarrow, ...)$ and a state $s_{init} \in S$, whether there are no infinite runs starting from s_{init}

Lem. [Finite Witnesses for Infinite Runs] A WSTS (S, \rightarrow, \leq) has an infinite run from s_{init} iff it has a finite run from s_{init} that is a good sequence.

Recall: $s_0, s_1, s_2, \dots, s_n$ is good $\stackrel{\text{def}}{\Leftrightarrow}$ there exist i < j s.t. $s_i \leq s_j$

Coro. One can decide Termination for a WSTS by enumerating all finite runs from s_{init} and reject when/if a good sequence is found.

NB: This requires some minimal effectiveness assumptions on the WSTS, e.g., that the ordering is decidable

Coverability is the question, given $(S, \rightarrow, ...)$, a state s_{init} and a target state t, whether there is a run $s_{init} \rightarrow s_1 \rightarrow s_2 \cdots \rightarrow s_n$ with $s_n \ge t$.

This is equivalent to having a pseudo-run $s_{init}, s_1, ..., s_n$ with $s_n \ge t$, where a pseudo-run is a sequence of pseudo-steps $s_{i-1} \rightarrow s'_i \ge s_i$.



Lem. [Finite Witnesses for Covering] There is a pseudo-run $s_{\text{init}}, \ldots, s_n$ covering t **iff** there is a minimal pseudo-run $s_0 \rightarrow \ge \cdots \rightarrow \ge s_{n'} = t$ from some $s_0 \le s_{\text{init}}$ to t such that $s_{n'}, s_{n'-1}, \ldots, s_0$ is a bad sequence.

NB. a pseudo-run $s_0, \ldots, s_{n'}$ is minimal $\stackrel{\text{def}}{\Leftrightarrow}$ for all $0 \leq i < n'$, s_i is a minimal (pseudo) predecessor of s_{i+1} .

Coverability is the question, given $(S, \rightarrow,...)$, a state s_{init} and a target state t, whether there is a run $s_{init} \rightarrow s_1 \rightarrow s_2 \cdots \rightarrow s_n$ with $s_n \ge t$.

This is equivalent to having a pseudo-run $s_{init}, s_1, ..., s_n$ with $s_n \ge t$, where a pseudo-run is a sequence of pseudo-steps $s_{i-1} \rightarrow s'_i \ge s_i$.



Lem. [Finite Witnesses for Covering] There is a pseudo-run $s_{\text{init}}, \ldots, s_n$ covering t **iff** there is a minimal pseudo-run $s_0 \rightarrow \ge \cdots \rightarrow \ge s_{n'} = t$ from some $s_0 \le s_{\text{init}}$ to t such that $s_{n'}, s_{n'-1}, \ldots, s_0$ is a bad sequence.

NB. a pseudo-run $s_0, \ldots, s_{n'}$ is minimal $\stackrel{\text{def}}{\Leftrightarrow}$ for all $0 \leq i < n'$, s_i is a minimal (pseudo) predecessor of s_{i+1} .

Coverability is the question, given $(S, \rightarrow,...)$, a state s_{init} and a target state t, whether there is a run $s_{init} \rightarrow s_1 \rightarrow s_2 \cdots \rightarrow s_n$ with $s_n \ge t$.

This is equivalent to having a pseudo-run $s_{init}, s_1, ..., s_n$ with $s_n \ge t$, where a pseudo-run is a sequence of pseudo-steps $s_{i-1} \rightarrow s'_i \ge s_i$.



Lem. [Finite Witnesses for Covering] There is a pseudo-run $s_{\text{init}}, \ldots, s_n$ covering t iff there is a minimal pseudo-run $s_0 \rightarrow \geqslant \cdots \rightarrow \geqslant s_{n'} = t$ from some $s_0 \leqslant s_{\text{init}}$ to t such that $s_{n'}, s_{n'-1}, \ldots, s_0$ is a bad sequence.

NB. a pseudo-run $s_0, \ldots, s_{n'}$ is minimal $\stackrel{\text{def}}{\Leftrightarrow}$ for all $0 \leq i < n'$, s_i is a minimal (pseudo) predecessor of s_{i+1} .

Coverability is the question, given $(S, \rightarrow, ...)$, a state s_{init} and a target state t, whether there is a run $s_{init} \rightarrow s_1 \rightarrow s_2 \cdots \rightarrow s_n$ with $s_n \ge t$.

This is equivalent to having a pseudo-run $s_{init}, s_1, ..., s_n$ with $s_n \ge t$, where a pseudo-run is a sequence of pseudo-steps $s_{i-1} \rightarrow s'_i \ge s_i$.



Lem. [Finite Witnesses for Covering] There is a pseudo-run $s_{\text{init}}, \ldots, s_n$ covering t **iff** there is a minimal pseudo-run $s_0 \rightarrow \ge \cdots \rightarrow \ge s_{n'} = t$ from some $s_0 \le s_{\text{init}}$ to t such that $s_{n'}, s_{n'-1}, \ldots, s_0$ is a bad sequence.

NB. a pseudo-run $s_0, \ldots, s_{n'}$ is minimal $\stackrel{\text{def}}{\Leftrightarrow}$ for all $0 \leq i < n'$, s_i is a minimal (pseudo) predecessor of s_{i+1} .

The two algorithms we have seen guess a finite sequence s_0, s_1, \ldots, s_ℓ that is bad (for Coverability) or almost bad (for non-Termination) and check that they are indeed correct witnesses.

We can give a complexity upper bound in (CO)NTIME(f(n)) or (CO)NSPACE(f(n)) if we can bound the size of the sequence —in practice: bound its length ℓ — as a function of the input $(S, \rightarrow, \leqslant), s_{init}, t, ...$

This is the topic for next course ...

The two algorithms we have seen guess a finite sequence s_0, s_1, \ldots, s_ℓ that is bad (for Coverability) or almost bad (for non-Termination) and check that they are indeed correct witnesses.

We can give a complexity upper bound in (CO)NTIME(f(n)) or (CO)NSPACE(f(n)) if we can bound the size of the sequence —in practice: bound its length ℓ — as a function of the input $(S, \rightarrow, \leqslant), s_{init}, t, ...$

This is the topic for next course

The two algorithms we have seen guess a finite sequence s_0, s_1, \ldots, s_ℓ that is bad (for Coverability) or almost bad (for non-Termination) and check that they are indeed correct witnesses.

We can give a complexity upper bound in (CO)NTIME(f(n)) or (CO)NSPACE(f(n)) if we can bound the size of the sequence —in practice: bound its length ℓ — as a function of the input $(S, \rightarrow, \leqslant), s_{init}, t, ...$

This is the topic for next course