Infinite Automata 2025/26

Lecture Notes 6

Henry Sinclair-Banks

Claim 5.6. Let V be a VASS and let $(p, \mathbf{u}), (q, \mathbf{v})$ be a pair of configurations. We define $R((p, \mathbf{u})) \coloneqq \{(p', \mathbf{u}') : (p, \mathbf{u}) \xrightarrow{*}_V (p', \mathbf{u}')\}$. The following two statements are equivalent.

(1) There exists $(q, \mathbf{v}') \in R((p, \mathbf{u}))$ such that $\mathbf{v}' \geq \mathbf{v}$.

(2) There is a node α in $\mathcal{T}$ such that $\text{state}(\alpha) = q$ and $\text{vector}(\alpha) \geq \mathbf{v}$.

Proof. For (1) $\implies$ (2), consider the run $(p, \mathbf{u}) \xrightarrow{\pi}_V (q, \mathbf{v}')$ for some $\mathbf{v}' \geq \mathbf{v}$. We shall repeatedly apply the following procedure to π. From left to right, scan through the configurations in the run and search for either of the following two properties.

(i) If there is a configuration $(r, \mathbf{y})$ that repeats in π, then delete the second occurrence of $(r, \mathbf{y})$ and all subsequent configurations in $(p, \mathbf{u}) \xrightarrow{\pi}_V (q, \mathbf{v}')$.

(ii) If there are two configurations $(r, \mathbf{y})$ and $(r, \mathbf{y}')$ such that $\mathbf{y}' \geq \mathbf{y}$, there is an $i \in \{1, \ldots, d\}$ such that $\mathbf{y}' > \mathbf{y}$, and $(r, \mathbf{y})$ occurs before $(r, \mathbf{y}')$, then in every configuration $(r', \mathbf{z})$ *after* $(r, \mathbf{y}')$, the component $\mathbf{z}[i]$ is replaced with ω.

This procedure must terminate because property (i) can only be observed finitely many times (the length of $(p, \mathbf{u}) \xrightarrow{\pi}_V (q, \mathbf{v}')$ is finite) and (ii) can only be observed finitely many times (again because the dimension of V is finite and the length of $(p, \mathbf{u}) \xrightarrow{\pi}_V (q, \mathbf{v}')$ is finite). Let $(p, \mathbf{u}) \xrightarrow{\pi'} (q, \mathbf{w})$ be the run that is obtained; note that this is now a run over "pseudo-configurations", so $(q, \mathbf{w}) \in Q \times \mathbb{N}_\omega^d$. This run is *exactly* contains the pseudo-configurations that one would witness in the path down the coverability tree from the root node labelled with $(p, \mathbf{u})$ to the node labelled with $(q, \mathbf{w})$. It is true that $\mathbf{w} \geq \mathbf{v}'$ because $\mathbf{w}$ is obtained by replacing some components of $\mathbf{v}'$ with 'ω's. Since $\mathbf{v}' \geq \mathbf{v}$, it follows that $\mathbf{w} \geq \mathbf{v}$.

For (2) $\implies$ (1), we will prove the following claim by induction. Let α be a node in $\mathcal{T}$ and let Ω_α be the set of indices i for which $\text{vector}(\alpha)[i] = \omega$. The claim is that for every node α in $\mathcal{T}$ and for every $N \in \mathbb{N}$, there exist $N_i \in \mathbb{N}$ for every $i \in \Omega_\alpha$ such that $N_i \geq N$ and the configuration $(\text{state}(\alpha), \mathbf{v}) \in R((p, \mathbf{u}))$ is reachable from $(p, \mathbf{u})$ where $\mathbf{v}$ is defined by

$$\mathbf{v}[i] \coloneqq \begin{cases} N_i & \text{if } i \in \Omega_\alpha, \\ \text{vector}(\alpha)[i] & \text{if } i \notin \Omega_\alpha. \end{cases}$$

We will prove this claim by induction on the depth of the nodes in the tree.

For depth $n = 0$, the only node is ε and its label is $(p, \mathbf{u}) \in Q \times \mathbb{N}^d$. This base case is trivial because $\mathbf{u} \in \mathbb{N}^d$ has no ω-components ($\Omega_\varepsilon = \emptyset$). Thus $(p, \mathbf{u}) \in R((p, \mathbf{u}))$ immediately satisfies the property.

Now, for the inductive step, assume that the claim holds for very node in the tree with depth n. Let γ be a node in $\mathcal{T}$ at depth $n+1$. Let β be the parent of γ in $\mathcal{T}$ at depth n. Let $t = (\text{state}(\beta), \mathbf{x}, \text{state}(\beta)) \in T$ be the transition taken from β to γ. There are two cases to consider.

The first case is when $\Omega_\gamma = \Omega_\beta$ ($\text{vector}(\gamma)$ and $\text{vector}(\beta)$ have the same ω-components). Recall that we wish to reach a configuration $(\text{state}(\gamma), \mathbf{v})$ such that $\mathbf{v}[i] \geq N$ if i is an omega component of γ. By the inductive assumption with $N' = N + \|\mathbf{x}\|_\infty$, we know that there is a configuration $(\text{state}(\beta), \mathbf{v}')$ that is reachable such that $\mathbf{v}'[i] \geq N' = N + \|\mathbf{x}\|_\infty$ for all $i \in \Omega_\beta$, and $\mathbf{v}'[i] = \text{vector}(\beta)[i]$ for all $i \notin \Omega_\beta$. Thus, from $(\text{state}(\beta), \mathbf{v}')$, when the transition t is taken, a configuration $(\text{state}(\gamma), \mathbf{v})$ is reached where $\mathbf{v} = \mathbf{v}' + \mathbf{x}$. Therefore, for all $i \in \Omega_\gamma = \Omega_\beta$, $\mathbf{v}[i] = \mathbf{v}'[i] + \mathbf{x}[i] \geq N + \|\mathbf{x}\|_\infty - \mathbf{x}[i] \geq N$, as required. So there is a run

$$(p, \mathbf{u}) \xrightarrow{*}_V (\text{state}(\beta), \mathbf{v}') \xrightarrow{t}_V (\text{state}(\gamma), \mathbf{v})$$

and $\mathbf{v}$ is at least N on its ω-components.

The second case is when $\Omega_\gamma \neq \Omega_\beta$. The only possible scenario is that $\Omega_\gamma \supset \Omega_\beta$ because of the definition of the coverability tree (Definition 5.4), ω-components are persistent. Let $p = |\Omega_\gamma \setminus \Omega_\beta|$ be the number of new ω-components in vector(γ) that were not present in vector(β). This case is similar to the first case, except it could be true that in order to get a new ω-component to a value that is at least N, a cycle is used that is positive on the this new ω-component but negative on all components $i \in \Omega_\beta$ that were already ω.

For every $j \in \Omega_\gamma \setminus \Omega_\beta$, let C_j be the cycle that allowed vector$(\gamma)[j] = \omega$ in $\mathcal{T}$. Specifically, for each $j \in \Omega_\gamma \setminus \Omega_\beta$ there exists some node α_j that is an ancestor of β such that for every $i \notin \Omega_\beta$, the cycle C_j is nonnegative on the i-th component, and for component j, it is true that C_j is strictly positive. Let k be the maximum of all lengths of cycles C_j; we know that $k < n$ because β is at depth n in the tree. Let M be the absolute value of the greatest counter update on any transition in the VASS.

Recall that we wish to reach a configuration (state(γ), $\mathbf{v}$) where $\mathbf{v}[i] \geq N$ on all components $i \in \Omega_\gamma$ for which vector$(\gamma)[i] = \omega$. By the inductive assumption with $N' = N + \|\mathbf{x}\|_\infty + p \cdot N \cdot kM$, we know that there is a configuration (state(β), $\mathbf{v}'$) that is reachable such that

$$\begin{aligned} \mathbf{v}'[i] &\geq N' = N + \|\mathbf{x}\|_\infty + p \cdot N \cdot kM && \text{for all } i \in \Omega_\beta, \text{ and} \\ \mathbf{v}'[i] &= \text{vector}(\beta)[i] && \text{for all } i \notin \Omega_\beta. \end{aligned}$$

So, from (state(β), $\mathbf{v}'$), we will take N iterations of each cycle C_j in order to get the new ω-components $i \in \Omega_\gamma \setminus \Omega_\beta$ to take a value that is at least N. Then the transition t is taken in order to go from state(β) to state(γ):

$$(\text{state}(\beta), \mathbf{w}) \xrightarrow{\cdots C_j^N \cdots t} (\text{state}(\gamma), \mathbf{w}').$$

Accordingly, a configuration (state(γ), $\mathbf{v}$) is reached where $\mathbf{v} = \mathbf{v}' + \mathbf{x} + \sum_{j \in \Omega_\gamma \setminus \Omega_\beta} N \cdot \text{effect}(C_j)$. On the components $i \in \Omega_\beta$, we know that

$$\begin{aligned} \mathbf{v}[i] &= \mathbf{v}'[i] + \mathbf{x}[i] + \sum_{j \in \Omega_\gamma \setminus \Omega_\beta} N \cdot \text{effect}(C_j)[i] \\ &\geq N + \|\mathbf{x}\|_\infty + p \cdot N \cdot kM + (-\|\mathbf{x}\|_\infty) + p \cdot N \cdot (-kM) \\ &\geq N. \end{aligned}$$

On the new ω-components $i \in \Omega_\gamma \setminus \Omega_\beta$, we know that

$$\begin{aligned} \mathbf{v}[i] &= \mathbf{v}'[i] + \mathbf{x}[i] + \sum_{j \in \Omega_\gamma \setminus \Omega_\beta} N \cdot \text{effect}(C_j)[i] \\ &\geq \mathbf{v}'[i] + \mathbf{x}[i] + N \cdot \text{effect}(C_i)[i] \\ &\geq \mathbf{v}'[i] + \mathbf{x}[i] + N \\ &\geq N. \end{aligned}$$

Lastly, on the components $i \notin \Omega_\gamma$, since $i \notin \Omega_\beta$ as well, we know effect$(C_j)[i] \geq 0$ (for all j) and so simply $\mathbf{v}[i] = \mathbf{v}'[i] + \mathbf{x}[i] = \text{vector}(\gamma)[i]$. We therefore obtain a run

$$(p, \mathbf{u}) \xrightarrow{*}_V (\text{state}(\beta), \mathbf{v}') \xrightarrow{\cdots C_j^N \cdots t} (\text{state}(\gamma), \mathbf{v})$$

where $\mathbf{v}$ is at least N on all ω components. This concludes the proof of the claim.

Finally, to conclude the proof of Claim 5.6, we simply use the previous claim as follows. Given that there is a node α in $\mathcal{T}$ such that state$(\alpha) = q$ and vector$(\alpha) \geq \mathbf{v}$, we know that we can find a reachable configuration $(q, \mathbf{v}') \in R((p, \mathbf{u}))$ such that $\mathbf{v}'[i] \geq N = \|\mathbf{v}\|_\infty$ for all $i \in \Omega_\alpha$. □

Corollary 6.1. The decision problems coverability, unboundedness, and simultaneous unboundedness are decidable in VASS.

Proof idea. Compute the coverability tree for a given VASS with a given initial configuration and simply look for the property in question in the labels of the nodes in the tree. This can be completed in finite time thanks to Lemma 5.5. □

We will now move onto stating and proving Parikh's theorem. For this we must define semilinear sets and Parikh's function.

Definition 6.2. A *linear set* $L \subseteq \mathbb{N}^d$ is a set that can be defined as follows for a *base* $\mathbf{b} \in \mathbb{N}^d$ and *periods* $P \subseteq \mathbb{N}^d$:

$$L(\mathbf{b}, P) \coloneqq \{\mathbf{x} + x_1\mathbf{p_1} + \ldots + x_k\mathbf{p_k} : \mathbf{p_1}, \ldots, \mathbf{p_k} \in P \text{ and } x_1, \ldots, x_k \in \mathbb{N}\}.$$

A *semilinear set* is a finite union of linear sets.

Proposition 6.3. Semilinear sets are closed under union, intersection, complement, and projection.

Definition 6.4. Fix an alphabet $\Sigma = \{\sigma_1, \ldots, \sigma_k\}$. The *Parikh function* $\psi : \Sigma^* \to \mathbb{N}^k$ is defined for a word $w \in \Sigma^*$ as

$$\psi(w) \coloneqq (|w|_{\sigma_1}, \ldots, |w|_{\sigma_k}),$$

where $|w|_{\sigma_i}$ is the number of times the letter σ_i occurs in w. We overload this notation to take Parikh images of languages; for $S \subseteq \Sigma^*$, $\psi(S) \coloneqq \{\psi(w) : w \in S\}$.

Theorem 6.5. *Parikh's Theorem.* The Parikh image of any context free language is semilinear.

Before proving Theorem 6.5, we will present a slightly stronger form of the pumping lemma for context free languages.

Lemma 6.6. *Stronger pumping lemma for CFLs.* If L is generated by a Chomsky normal form grammar, then there exists $N \geq 1$ such that the following holds. For every $k \in \mathbb{N}$, and for every $w \in L$ such that $|w| \geq N^k$, there exists a nonterminal symbol A such that $S \overset{*}{\Rightarrow} u\,A\,v$ (where S is the starting symbol) and w can be decomposed into $w = u\,x_1 \cdots x_k\,z\,y_k \cdots y_1\,v$ such that

(i) $|x_i\,y_i| > 1$, for every $i \in \{1, \ldots, k\}$;

(ii) $|x_1 \cdots x_k\,z\,y_k \cdots y_1| \leq N^k$;

(iii) $A \overset{*}{\Rightarrow} x_i\,A\,y_i$, for every $i \in \{1, \ldots, k\}$; and

(iv) $A \overset{*}{\Rightarrow} z$.

Proof idea. One can prove Lemma 6.6 in the same way that one would prove the classical pumping lemma for context free language, instead one extends the length of the words considered to N^k to allow for pigeonhole principle arguments to facilitate the existence of a greater number of "pumping locations" $A \overset{*}{\Rightarrow} x_i\,A\,y_i$. □

Before we prove Theorem 6.5, we shall state a corollary to at least partially motivate Parikh's theorem and its proof.

Corollary 6.7. Pushdown automata over a singleton alphabet recognise regular languages.

Proof. First, observe that all unary languages $S \subseteq \{a\}^*$ satisfy: $w \in S$ if and only if $|w| \in \psi(S)$. This means that, roughly speaking, there is no difference between a unary language and its Parikh image.

Let P be a pushdown automata over a singleton alphabet $\Sigma = \{a\}$. Convert P into a (language-equivalent) context free grammar G. By Theorem 6.5, we know that the Parikh language $\psi(L(G))$ is semilinear. It is also straightforward to see that, for a given semilinear set S, one can construct a unary NFA N such that $\psi(N) = S$. For a linear set, one can a linear chain of states (where the length of the chain is equal to the base value in the linear set) leading to a collection of cycles where there is one cycle for each period in the linear set (where the length of each cycle corresponds to each of the period values in the linear set). Since semilinear sets are unions of linear sets, we can just use ε branches in the NFA to point to each linear set in the union of the given semilinear set. $\square$

Proof of Theorem 6.5. Let L be the given context free language. First, construct a Chomsky normal form grammar for L. For each finite nonempty subset of nonterminals X, define L_X to be the subset of words in L such that there exists a derivation that uses *every* nonterminal in X and *no* nonterminals not in X. It is clear that $L = \bigcup_X L_X$. Given that semilinear sets are closed under union (Proposition 6.3), it suffices to show that $\psi(L_X)$ is semilinear for an arbitrary subset of nonterminals X.

Fix a subset of nonterminals X and let $k = |X|$. We will construct two finite sets F and G such that $\psi(L_X) = \psi(F \cdot G^*)$. Given that F and G are finite, $\psi(F \cdot G^*) = \bigcup_{f \in F} \psi(\{f\} \cdot G^*)$ is clearly semilinear. We define

$$\begin{aligned} F &\coloneqq \{w \in L_X : |w| < N^k\}, \text{ and} \\ G &\coloneqq \{x\,y : 1 \leq |x\,y| \leq N^k \text{ and there exists } A \in X : A \overset{*}{\underset{X}{\Rightarrow}} x\,A\,y\}. \end{aligned}$$

Here, $N \in \mathbb{N}$ is the value given by Lemma 6.6 (the stronger pumping lemma for CFLs) applied to the context free language L_X.

First, we will prove that $\psi(L_X) \subseteq \psi(F \cdot G^*)$ by induction on the length of a word $w \in L_X$. If $|w| < N^k$, then $w \in F$, so $\psi(w) \in \psi(F \cdot G^*)$. Otherwise, since $w \in L_X$, there is a derivation of w that uses all of the non-terminals in U. By the Lemma 6.6, this derivation can be rewritten to be of the form:

$$\begin{aligned} S \overset{*}{\underset{d_0}{\Longrightarrow}} u\,A\,v \overset{*}{\underset{d_1}{\Longrightarrow}} u\,x_1\,A\,y_1\,v \overset{*}{\underset{d_2}{\Longrightarrow}} u\,x_1\,x_2\,A\,y_2\,y_1\,v \overset{*}{\underset{d_3}{\Longrightarrow}} \cdots \\ \cdots \overset{*}{\underset{d_k}{\Longrightarrow}} u\,x_1\,x_2 \cdots x_k\,A\,y_k \cdots y_2\,y_1\,v \overset{*}{\underset{d_{k+1}}{\Longrightarrow}} u\,x_1\,x_2 \cdots x_k\,z\,y_k \cdots y_2\,y_1\,v. \end{aligned}$$

Since there are only $k-1$ nonterminals in $X \setminus \{A\}$ and there are k many subderivations $d_1, \ldots, d_k$, then there must exist a subderivation d_i which uses non-terminals $X' \subseteq X \setminus \{A\}$ which are *all* used by other subderivations $d_1, \ldots, d_{i-1}, d_{i+1}, \ldots, d_k$. We can therefore delete this subderivation d_i to obtain a shorter word w' that still uses all of the nonterminals in X; $w' \in L_X$. Precisely,

$$w' = u\,x_1 \cdots x_{i-1}\,x_{i+1} \cdots x_k\,z\,y_k \cdots y_{i+1}\,y_{i-1} \cdots y_1 v.$$

By Lemma 6.6, we know that $1 \leq |x_i\,y_i| \leq N^k$, so both $|w'| \leq |w|$ and $x_i\,y_i \in G$. Moreover, since $|w'| \leq |w|$, we know that $\psi(w') \in \psi(F \cdot G^*)$ by the inductive assumption. As $\psi(w) = \psi(w') + \psi(x_i\,y_i)$, we conclude that $\psi(w) \in \psi(F \cdot G^*)$ as well.

End of lecture, $\psi(F \cdot G^*) \subseteq \psi(L_X)$ will be proved next lecture.