Infinite Automata 2025/26

Lecture Notes 5

Henry Sinclair-Banks

Definition 5.1. Universality of letter-labelled VASS (problem).

Input. A letter-labelled VASS V over the alphabet Σ.

Question. Is Σ^* the language that is recognised by V?

Theorem 5.2. Universality of letter-labelled VASS is decidable.

Before we begin the proof of Theorem 5.2, we will introduce and work with downward closed sets of $\mathbb{N}^d$. Precisely, let $\mathbb{D} \coloneqq \{X \subseteq \mathbb{N}^d : X = X\!\downarrow\}$. Here

$$X\!\downarrow \coloneqq \{\mathbf{y} \in \mathbb{N}^d : \text{ there exists } \mathbf{x} \in X \text{ such that } \mathbf{x} \geq \mathbf{y}\}.$$

In Exercise 5.3, we proved that $(\mathbb{D}, \subseteq)$ is a WQO. Fix $k \in \mathbb{N}$; $\mathbb{D}^k$ is the collection of k-tuples of sets in $\mathbb{D}$. Moreover, we defined the component-wise inclusion relation $\subseteq^k$ as follows. Let $X, Y \in \mathbb{D}^k$, $X \subseteq^k Y$ if and only if, for every $i \in \{1, \ldots, k\}$, $X[i] \subseteq Y[i]$. The following claim follows from Exercise 5.3 and Dickson's Lemma (Lemma 4.5).

Claim 5.3. For every $k \in \mathbb{N}$, $(\mathbb{D}, \subseteq^k)$ is a WQO.

Proof of Theorem 5.2. Let $V = (\Sigma, Q, T, q_0, F)$ be a letter-labelled d-VASS. First observe that upon reading a word $w \in \Sigma^*$, it is only possible to reach a finite (exponential) number of different configurations. We need not maintain *all* possible configurations that can be reached by reading w, we only need to maintain the maximal configurations reached. For example, suppose that $(p, \mathbf{x})$ and $(p, \mathbf{y})$ can be reached after reading w_1 and $\mathbf{x} \leq \mathbf{y}$. Now suppose that from $(p, \mathbf{x})$ the word w_2 can be read. Since $\mathbf{x} \leq \mathbf{y}$, we know that w_2 can also be read starting from $(p, \mathbf{y})$ as well. Thus, when verifying whether the word $w_1 w_2$ is accepted by V or not, we only needed to consider the greater configuration $(p, \mathbf{y})$.

For a word $w \in \Sigma^*$, let $R_w \subseteq Q \times \mathbb{N}^d$ be the set of configurations that can be reached by reading w from the initial configuration $(q_0, \mathbf{0})$. We attribute, to w, the $|Q|$-tuple of downward closed sets $X_w \in \mathbb{D}^{|Q|}$ such that, for every state $p \in Q$, $X_w[p] = \{\mathbf{v} : (p, \mathbf{v}) \in R_w\}\!\downarrow$. In other words, $X_w[p]$ is the downward closure of counter value vectors that can be reached by reading w and ending at state p. If it is not possible to reach the state p after reading w, then $X_w[p] = \emptyset$.

Since $X_w[p]$ is the downward closure of a finite set in $\mathbb{N}^d$, then $X_w[p]$ is also finite. Moreover $X_w[p]$ can be represented as the downward closure of finitely many maximal elements. It is also true that the representation of $X_w[p]$ as the downward closure of its maximal elements can be computed in finite time.

We say that X_w is an *accepting* tuple if there exists $q \in F$ such that $X_w[q] \neq \emptyset$. It immediately follows that w is accepted by V if and only if X_w is an accepting tuple.

Now, we will construct a tree $\mathcal{T}$ whose edges are labelled by letters in Σ and whose nodes are the tuples X_w. In a similar fashion to the proof of Theorem 4.6, consider generating this tree (in a BFS or DFS style). Since $(\mathbb{D}^{|Q|}, \subseteq^{|Q|})$ is a WQO (Claim 5.3), we know that in every path down the tree, there will eventually exist a node X_w which has an ancestor $X_{w'}$ (i.e. w' is a prefix of w) such that $X'_w \subseteq^{|Q|} X_w$. We shall terminate the exploration of a branch in the tree once we observe such a node X_w (that is greater than an ancestor). By König's Lemma (Lemma 4.4) and Claim 5.3, $\mathcal{T}$ is finite.

We will now argue that if there exists $w \in \Sigma^*$ such that w is *not* accepted by V, then there exists w' in $\mathcal{T}$ such that w' is not accepted by V. Suppose that there exists a word that is not accepted by V; in fact let w be the shortest word that is not accepted by V. Assume, for sake of contradiction, that w is not in $\mathcal{T}$. Consider the decomposition of w into subwords x, y, and z such that (i) $w = xyz$, (ii) X_{xy} is a leaf node of $\mathcal{T}$, and (iii) X_x is the ancestor tuple such that $X_x \subseteq^{|Q|} X_{xy}$. Note that $|y| > 0$ because X_x is an ancestor of X_{xy}. Since $X_x \subseteq^{|Q|} X_{xy}$, it is true that $X_{xz} \subseteq^{|Q|} X_{xyz}$. Moreover, since w is not accepted by V, we know that $X_{xyz} = X_w$ is not an accepting tuple. This means that for all $q \in F$, $X_w[q] = \emptyset$. Given that $X_{xz} \subseteq^{|Q|} X_{xyz}$, it follows that $X_{xyz}[q] = \emptyset$ implies that $X_{xz}[q] = \emptyset$. Thus X_{xz} is not an accepting tuple, and we can therefore conclude that $xz \in \Sigma^*$ is not accepted by V. Since $|y| > 0$, it is true that $|xz| < |xyz| = |w|$. This contradicts the minimality of the length of w.

Now, to conclude this proof, observe that in order to decide whether V accepts all words $w \in \Sigma^*$, it suffices to check that w is accepted by V for all words that have nodes in the tree $\mathcal{T}$. This algorithm therefore generates $\mathcal{T}$, at each step computes X_w and checks whether X_w is an accepting tuple. If at any point, a tuple if found that is not accepting, then the algorithm outputs "no" (the language of V is not universal). Otherwise, at the end of generating the tree, if all nodes are accepting, then the algorithm outputs "yes" (the language of V is universal). □

The following definition differs only slightly from the original definition presented by Karp and Miller in 1969; they used VAS and we will use VASS. We define ω to be the symbol such that, for every $z \in \mathbb{Z}$, $\omega + z = \omega$ and $z < \omega$. We also define $\mathbb{N}_\omega \coloneqq \mathbb{N} \cup \{\omega\}$.

Definition 5.4. Let $V = (Q, T)$ be a d-VASS and let $(p, \mathbf{u}) \in Q \times \mathbb{N}^d$ be a configuration. The *coverability tree* $\mathcal{T}$of V rooted with $(p, \mathbf{u})$ is defined as follows. Nodes in the tree will are elements in T^* (i.e. runs in V). Each node is labelled by a "pseudo-configuration" $(q, \mathbf{v}) \in Q \times \mathbb{N}_\omega^d$. For convenience, we define the functions state(α) and vector(α) as one would expect: if α is labelled with $(q, \mathbf{v})$ then state$(\alpha) = q \in Q$ and vector$(\alpha) = \mathbf{v} \in \mathbb{N}_\omega^d$.

The root of $\mathcal{T}$ is $\varepsilon \in T^*$ and it is labelled with state$(\varepsilon) = p$ and vector$(\varepsilon) = \mathbf{u}$. Let β be an arbitrary node in $\mathcal{T}$. If there exists an ancestor α of β such that the labels are the same (state(α) = state(β) and vector(α) = vector(β)) then β is a leaf node (it is the end of its branch of the tree and has no children). Assuming that is not the case (β is not a leaf), there will be one child γ of β for every transition such that (state$(\beta), \mathbf{x}, q) \in T$ such that vector$(\beta) + \mathbf{x} \geq 0$. The state-label of γ will be state$(\gamma) = q$ and the vector-label of γ will be defined as follows. First, we say that α is an *i-pumpable* ancester or β if α is an ancestor of β, state(α) = state(β), vector$(\alpha) \leq$ vector(β), and vector$(\alpha)[i] <$ vector$(\beta)[i]$. Now, for $i \in \{1, \ldots, d\}$,

$$\text{vector}(\gamma)[i] = \begin{cases} \omega & \text{if } \beta \text{ has an } i\text{-pumpable ancestor,} \\ \text{vector}(\beta)[i] + \mathbf{x}[i] & \text{otherwise.} \end{cases}$$

Intuitively speaking there are two reasons that in which vector$(\gamma)[i]$ could be set to or equal to ω. The first is given by "i-pumpable ancestors". This means that there is some ancestor which has the same state and which the vector-label of β is at least the vector-label of α. This means that the cycle $\beta - \alpha \in T^*$ can be repeated arbitrarily many times. Moreover, since we insist that vector$(\beta)[i] >$ vector$(\alpha)[i]$, it is therefore possible to attain arbitrary high counter values on the i-th coordinate; hence we set vector$(\gamma)[i] = \omega$. The other way that vector$(\gamma)[i] = \omega$ is if vector$(\beta)[i] = \omega$ already, in this case, regardless of $\mathbf{x}$, it will be true that vector$(\gamma)[i]$ = vector$(\beta)[i] + \mathbf{x}[i] = \omega$.

Lemma 5.5. Let V be a VASS and let $(p, \mathbf{u})$ be a configuration. The coverability tree $\mathcal{T}$ of V rooted with $(p, \mathbf{u})$ is finite.

Proof. Suppose, for sake of contradiction, that $\mathcal{T}$ is infinite. Since $\mathcal{T}$ is finitely branching, by König's Lemma (Lemma 4.4), there must exist an infinite path down the tree. Since $(\mathbb{N}_\omega^d, \leq)$ is a WQO and by pigeonhole principle over the set of states Q, there exists an infinite sequence $(\alpha_i)_{i=1}^\infty$ of nodes in $\mathcal{T}$ such that state(α_1) = state$(\alpha_2) = \ldots$ and vector$(\alpha_1) \leq$ vector$(\alpha_2) \leq \ldots$. By definition of $\mathcal{T}$, we know that for every i, vector$(\alpha_i) \neq$ vector(α_{i+1}). Again, by definition of the tree, that means that vector(α_{i+1}) must contain at least one more ω than vector(α_i). However, since there are only finitely many coordinates and the sequence of α_i is infinite, this cannot be the case. We therefore conclude that $\mathcal{T}$ must be finite. □

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 sketch. For (1) $\implies$ (2): consider the run π from $(p, \mathbf{u})$ to $(q, \mathbf{v}')$ for some $\mathbf{v}' \geq \mathbf{v}$. Repeatedly apply the following procedure to π: (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; (ii) if there is a configuration $(r, \mathbf{y})$ such that there is a prior configuration $(r, \mathbf{x})$ such that $\mathbf{x} \leq \mathbf{y}$ and $\mathbf{x}[i] < \mathbf{y}[i]$, then replace the i-th entry of $\mathbf{y}$ with ω. This procedure must terminate because there are only finitely many (i) operations that are possible and between every (i) operation, there can only be finitely many (ii) operations. Let $(q, \mathbf{w}) \in Q \times \mathbb{N}_\omega^d$ be the final "pseudo-configuration" in the modified sequence. Observe that the final sequence obtained is exactly a route through $\mathcal{T}$ to $(q, \mathbf{w}) \geq (q, \mathbf{v}') \geq (q, \mathbf{v})$.

For (2) $\implies$ (1): consider the path down $\mathcal{T}$ from ε to α such that $\text{state}(\alpha) = q$ and $\text{vector}(\alpha) \geq \mathbf{v}$. Assume WLOG that the first $0 \leq h \leq d$ components of $\text{vector}(\alpha)$ are ω. Moreover, assume WLOG that the ω components are introduced in order via the path down the tree. Every time a new ω value is added to the tree, there must have been an i-pumpable ancestor. This can be used to create a cycle that is non-negative on all non-ω places and strictly positive on the ω place. One can obtain a cycle for each ω places and can create a valid run in the original VASS that follows the path down the tree, but whenever a new ω is introduced its corresponding i-pumpable cycle is taken a (very) large number of times — so much so that later cycles that may be negative on the ω components still do not bring the counters below zero. Intuitively speaking, it is possible, for every $N \in \mathbb{N}$ to find a configuration in $R((p, \mathbf{u}))$ with counter values $(N_1, \ldots, N_h, v_{h+1}, \ldots, v_d)$ for some $N_1, \ldots, N_h \geq N$. □