1 Infinite Computations

Theorem 1. *NBA-recognizable languages are closed under complement.*

Proof. Let $\mathcal{A} = (Q, \Sigma, q_0, \Delta, F)$ be an NBA. We define the transition profile of a word $w \in \Sigma^*$ as a finite directed graph $\mathrm{t}(w) = (Q, E, E_F)$ with $E = \{(p, q) \in Q \times Q \mid \mathcal{A} : p \xrightarrow{w} q\}$ and $E_F = \{(p, q) \in Q \times Q \mid \mathcal{A} : p \underset{F}{\xrightarrow{w}} q\}$. Then we define $\mathrm{TP} = \{\mathrm{t}(w) \mid w \in \Sigma^*\}$ and the transition profile automaton $\mathrm{TPA}_F = (\mathrm{TP}, \Sigma, \mathrm{t}(\varepsilon), \delta_{\mathrm{TPA}}, F)$ with $\delta_{\mathrm{TPA}}(\mathrm{t}(u), a) = \mathrm{t}(ua)$.

For $t \in \mathrm{TP}$, let $U_t = \{u \in \Sigma^+ \mid \mathrm{t}(u) = t\}$. These sets are regular, as they are accepted by the NFA $\mathrm{TPA}_{\{t\}}$.

Let $\bar{t} = t_0 t_1 \cdots \in \mathrm{TP}^\omega$ be an infinite sequence of transition profiles. We call $\bar{t}$ accepting if there are $q_1 \cdots \in Q^\omega$ such that every t_i has an edge (q_i, q_{i+1}) and infinitely many of these edges are labeled by F. Let $\mathrm{NTP} = \{(t_0, t_1) \in \mathrm{TP} \times \mathrm{TP} \mid t_0 t_1^\omega$ is non-accepting$\}$.

Claim : $L(\mathcal{A})^\complement = \bigcup_{(t_0,t_1) \in \mathrm{NTP}} U_{t_0} U_{t_1}^\omega$. Then one can construct an NBA for $L(\mathcal{A})^\complement$.

For every $\alpha \in \Sigma^\omega$, let $\alpha = u_0 u_1 \cdots \in (\Sigma^*)^\omega$ be a factorization of α into finite words and let $\bar{t} = \mathrm{t}(u_0)\mathrm{t}(u_1) \ldots$. Then $\bar{t}$ is accepting iff $\alpha \in L(\mathcal{A})$.

In particular, this is true for $\mathrm{t}(u_1) = \mathrm{t}(u_2) = \ldots$, so it remains to be shown that every word α has such a factorization.

Ramsey's Theorem: Let C be a finite set, X with $|X| = \aleph_0$, $E = \{(x, y) \in X \times X \mid x \neq y\}$, and $f : E \to C$. Then there is an infinite $Y \subseteq X$ such that $|f(E \cap (Y \times Y))| = 1$.

For a $\alpha \in \Sigma^\omega$, let $X = \mathbb{N}$ and $C = \mathrm{TP}$. For $i < j$ we set $f(j, i) = f(i, j) = \mathrm{t}(\alpha[i, j])$. Using Ramsey's theorem, there is an infinite $Y \subseteq \mathbb{N}$ such that $f(E \cap (Y \times Y)) = \{t\}$. Let $Y = \{i_j \mid j > 0\}$ such that $i_j < i_{j+1}$ and $i_0 := 0$. We define $u_j = \alpha[i_j, i_{j+1}]$. Then $\alpha = u_0 u_1 u_2 \ldots$ and $\mathrm{t}(u_1) = \mathrm{t}(u_2) = \ldots$. ☐

Theorem 2.

Proof. ☐

Theorem 3 (Landweber)**.** *Let $\mathcal{A} = (Q, \Sigma, q_0, \delta, \mathcal{F})$ be a DMA.*

1. *$L(\mathcal{A})$ is DBA-recognizable iff $\mathcal{F}$ is closed under super loops.*
2. *$L(\mathcal{A})$ is E-recognizable iff $\mathcal{F}$ is closed under reachable loops.*

Proof. ☐

Theorem 4. *NBAs, S1S, S1S$_0$, and ∃S1S have the same expressive power.*

Proof. **NBA ⇒ ∃S1S** Let $\mathcal{A} = (Q, \mathbb{B}^n, 1, \Delta, F)$ be the NBA with $Q = \{1, \ldots, m\}$. We encode a run of $\mathcal{A}$ on a word with a formula.

$$\varphi_{\mathcal{A}}(X_1, \ldots, X_n) = \exists Y_1 \ldots \exists Y_m \mathrm{Part}(\overline{Y}) \wedge Y_1(0) \wedge \mathrm{Trans}(\overline{X}, \overline{Y}) \wedge \mathrm{Fin}(\overline{Y})$$

$$\mathrm{Part}(Y_1, \ldots, Y_m) = \forall x \left(\bigvee_{i=1}^{m} Y_i x \wedge \bigwedge_{i=1}^{m} \bigwedge_{j \neq i} \neg Y_i x \vee \neg Y_j x \right)$$

$$\mathrm{Trans}(\overline{X}, \overline{Y}) = \forall x \bigvee_{\tau \in \Delta} \psi_\tau(x, \overline{X}, \overline{Y})$$

$$\psi_{(p,a,q)}(x, X_1, \ldots, X_n, Y_1, \ldots, Y_m) = Y_p x \wedge Y q(x+1) \wedge X_a x$$

$$\mathrm{Fin}(Y_1, \ldots, Y_m) = \forall x \exists y \left(x < y \wedge \bigvee_{q \in F} Y_q y \right)$$

∃S1S ⇒ S1S trivial

S1S ⇒ S1S$_0$ Let $\varphi \in$ S1S. We define an equivalent $\varphi' \in$ S1S$_0$ inductively.

- Boolean operators stay the same.
- $\exists x \psi(x) \mapsto \exists X_x(\mathrm{Sing}(X_x) \wedge \psi'(X_x))$
- $Xx \mapsto X_x \subseteq X$
- $x + 1 = y \mapsto \mathrm{Succ}(X_x, X_y)$

S1S$_0$ ⇒ NBA $\exists$, $\vee$, $\neg$ correspond to projection, union, and complement of automata respectively. Therefore it suffices to give NBAs for the three base formulas.

- $X_1 \subseteq X_2$: $Q = F = \{q_0\}$, $\Delta = \{(q_0, a, q_0) \mid a \neq (1\ 0)\}$.
- $\mathrm{Sing}(X)$: $Q = \{q_0, q_f\}$, $F = \{q_f\}$, $\Delta = \{(q_0, 0, q_0), (q_f, 0, q_f), (q_0, 1, q_f)\}$.
- $\mathrm{Succ}(X_1, X_2)$: $Q = \{q_0, q_1, q_2\}$, $F = \{q_2\}$,
  $\Delta = \{(q_0, (0\ 0), q_0), (q_2, (0\ 0), q_2), (q_0, (1\ 0), q_1), (q_1, (0\ 1), q_2)\}$.

The complexity of this construction is $2 \uparrow\uparrow |\varphi|$.

□

Theorem 5 (McNaughton)**.** *A language is NBA-recognizable iff it is DMA-recognizable.*

Proof. ⇐ NBA with $L(\mathcal{A}) = \bigcup_{F \in \mathcal{F}} \left(\bigcap_{q \in F} L(\mathcal{A}_q) \cap \bigcap_{q \notin F} \overline{L(\mathcal{A}_q)} \right)$ where $\mathcal{A}_q$ is $\mathcal{A}$ starting in q.

⇒ □

Theorem 6. *For each Muller automaton, one can construct an equivalent parity automaton. The construction preserves determinism.*

Proof. □

Theorem 7. *For each ABA, one can construct an equivalent NBA with states at most $3^{|Q|}$.*

Proof. Let $\mathcal{A} = (Q, \Sigma, q_0, \delta, F)$ be an ABA. We define $\mathcal{A}' = (\{0,1,2\}^Q, \Sigma, f_0, \Delta, \{0,2\}^Q)$ with $f_0(q) = \begin{cases} 1 & \text{if } q = q_0 \\ 0 & \text{else} \end{cases}$ and Δ as described below.

A state is a function $f : Q \to \{0,1,2\}$. Consider a run-tree of $\mathcal{A}$ on some word and all states that are "active" on one level in the tree. $f(q)$ is 0 for all states which are not active and 1 or 2 for all others. A state is assigned to 1 by default, which just means that it is active in some path. It is assigned 2 if "recently" on all paths it is active on, a final state in F was seen. Recently means that it occurred since the last time that an entire level in the run tree was made up of final states.

Formally, we have $(f, a, g) \in \Delta$ for all $g : Q \to \{0,1,2\}$ which satisfy the following:

- The active states need to be passed on, i.e. for all $q \in Q$: if $f(q) \in \{1,2\}$ then there must be an $X_q \subseteq Q$ with $X_q \models \delta(q,a)$ such that $g(X_q) \subseteq \{1,2\}$.
- A state is assigned 2 if it is final, i.e. $g(q) = 2$ if $q \in F$.
- Otherwise, a 1 overwrites a 2 for succession, i.e. for all $p \in P$, if there is a q with $f(q) = 1$ and $p \in X_q$, then also $g(p) = 1$.
- If all states are marked with a 2, $\mathcal{A}'$ reached a final state. We reset the values to $g(p) = 2$ iff $p \in F$.

□

2 Tree Automata

Theorem 8. *Let $\mathcal{A} = (Q, \Sigma, q_0, \Delta, F)$ be a TWA. There is an NTA $\mathcal{A}'$ of size exponential in $|Q|$ that recognizes $T(\mathcal{A})$.*

Proof. Let $\sim_{T(\mathcal{A})}$ be the usual equivalence relation, i.e. $t_1 \sim_{T(\mathcal{A})} t_2$ iff $\forall s \in S_\Sigma : s \cdot t_1 \in T(\mathcal{A}) \leftrightarrow s \cdot t_2 \in T(\mathcal{A})$. We define a relation $\sim \subseteq T_\Sigma \times T_\Sigma$ such that $\text{index}(\sim_{T(\mathcal{A})}) \leq \text{index}(\sim) \leq 2^{|Q|^2 \cdot m+1}$, where m is the maximal rank in Σ.

Let $t_0 \in T_\Sigma$ and $a_m \in \Sigma_m$ be arbitrary. For every $t \in T_\Sigma$ and $1 \leq i \leq m$, we define $t^{(i)} = a_m(t_0, t_0, \ldots, t, \ldots, t_0)$, meaning the i-th subtree below the root is t. Further, we define a relation $B_t^i \subseteq Q \times Q$ with $(p, q) \in B_t^i$ iff there is a run segment ρ of $\mathcal{A}$ on $t^{(i)}$, such that the run begins at the root of t, never leaves that subtree until the end. Meaning, $\rho = (p, i)(q_1, iu_1) \ldots (q_n, iu_n)(q, \varepsilon)$.

Finally, let $t_1 \sim t_2$ iff $t_1 \in T(\mathcal{A}) \leftrightarrow t_2 \in T(\mathcal{A})$ and $\forall i : B_{t_1}^i = B_{t_2}^i$.

Idea: $(p, q) \in B_t^i$ if $\mathcal{A}$ can enter t as i-th child with state p and after some while leaves it again with state q.

Claim : Let $t_1 \sim t_2$. Then $t_1 \sim_{T(\mathcal{A})} t_2$.

Let $s \in S_\Sigma$. Due to the symmetric definition of $\sim$, it suffices to show that $t_1 \in T(\mathcal{A})$ implies $t_2 \in T(\mathcal{A})$, so let $t_1 \in T(\mathcal{A})$. If $s = \circ$, then $s \cdot t_1 = t_1 \in T(\mathcal{A})$. By definition of $\sim$, this implies $s \cdot t_2 = t_2 \in T(\mathcal{A})$.

Otherwise $s \neq \circ$. Let $\rho_1\rho_2\rho_3$ be an accepting run of $\mathcal{A}$ on $s \cdot t_1$ such that ρ_1 only stays outside of t_1 and ρ_2 only stays inside of t_1. Since $B_{t_1}^i = B_{t_2}^i$, there is a run segment of $\mathcal{A}$ on t_2 which enters and exits the tree with the same states as ρ_2 does, meaning it can replace ρ_2 in the accepting run. Repeating this procedure gives an accepting run of $\mathcal{A}$ on $s \cdot t_2$, so $t_2 \in T(\mathcal{A})$.

Notes on the construction: each state in the NTA corresponds to a list of Q-states that $\mathcal{A}$ had when visiting this node, together with the direction from which it was coming. That can be used to check the correctness of a run. □

Theorem 9. *A language of finite trees $T \subseteq T_\Sigma$ can be recognized by an NTA iff it can be described by a regular tree expression.*

Proof. ⇒ Let $\mathcal{A} = (Q, \Sigma, \Delta, F)$ be an NTA for T. For $R, I \subseteq Q, q \in Q$, we define $T(R, I, q) \subseteq T_{\Sigma \cup C_R}$ (where $C_R = \{c_p \mid p \in R\}$) as the set of all trees on which $\mathcal{A}$ has a run that only uses states in I and ends in q. We can inductively define regular expressions for $T(R, I, q)$. Then $T(\mathcal{A}) = \bigcup\limits_{q \in F} T(\emptyset, Q, q)$.

For all $R \subseteq Q, q \in Q$, $T(R, \emptyset, q)$ contains only trees of height 0 or 1, so it is a finite union of singular languages.

$$T(R, I \cup \{i\}, q) = T(R, I, q) + T(R \cup \{i\}, I, q) \cdot^{c_i} (T(R \cup \{i\}, I, i))^{*_{c_i}} \cdot^{c_i} T(R, I, i)$$

⇐ Show by induction that a regular expression r can be transformed to a NTA $\mathcal{A}_r$.

- If $r = t \in T_{\Sigma \cup C}$, then there is an automaton $\mathcal{A}_t$ with $T(\mathcal{A}_t) = \{t\}$.
- If $r = s + t$, then $\mathcal{A}_r$ is the union automaton of $\mathcal{A}_s$ and $\mathcal{A}_{\sqcup}$.
- If $r = s \cdot^c t$, then $\mathcal{A}_r = (Q_s \cup Q_t, \Sigma \cup C, \Delta_r, F_s)$, where Δ_r simulates $\mathcal{A}_s$ and $\mathcal{A}_t$ but the final transitions in $\mathcal{A}_t$ are replaced by initial transitions to $\mathcal{A}_s$.
- If $r = s^{+_c}$, let $\mathcal{A}_r = (Q_s, \Sigma \cup C, \Delta_r, F_s)$, where Δ_r simulates $\mathcal{A}_s$ and allows "restarts" when a final state could be reached.

$\square$

Theorem 10. *Let D be an EDTD. We can construct a NUTA $\mathcal{A}$ with $T(D)^{\complement} = T(\mathcal{A})$ in polynomial time in $|D|$.*

Proof. Let $D = (\Sigma', P, S^{(1)})$. For $a^{(i)} \in \Sigma'$, let $a^{(i)} \to r_{a,i}$ be the according rule in P, where $r_{a,i}$ is a regular expression that can be transformed to a DFA in polynomial time. If $b \in \Sigma$ occurs in $r_{a,i}$ then it must have a unique type j (since D is single-typed). We call $b^{(j)}$ the b-successor of $a^{(i)}$. If b does not occur in $r_{a,i}$, we say that the b-successor is $b^{\perp}$. Furthermore, we assume $r_{a,\perp} = \varepsilon$ for all a.

We define a typing function $f : \mathrm{dom}_t \to \{\perp, 1, \ldots, k\}$. We assign $f(\varepsilon) = \begin{cases} 1 & \text{if } \mathrm{val}_t(\varepsilon) = S \\ \perp & \text{else} \end{cases}$. For the other nodes, let u be a node with parent v. We call $\mathrm{val}_t(v) = a$ and $\mathrm{val}_t(u) = b$. Then there is a unique $i \in \{\perp, 1, \ldots, k\}$ such that $b^{(i)}$ is the b-successor of $a^{f(v)}$. We set $f(u) := i$.

Claim: $t \in T(D)$ iff $\forall v \in \mathrm{dom}_t : f(v) \neq \perp$ and $a_1^{f(v1)} \ldots a_m^{f(vm)} \in L(r_{a,f(v)})$, where $a = \mathrm{val}_t(v)$ and $a_j = \mathrm{val}_t(vj)$. (without proof)

Using this claim, we can provide an automaton $\mathcal{A} = (Q, \Sigma, \Delta, F)$ that accepts a tree iff it violates the constraint.

- $Q = \Sigma \times \{\perp, 1, \ldots, k\} \times \{0, 1\}$, where the last component denotes whether a violation was found.
- $\Delta = \{(L_{a,i,x}, a, (a, i, x)) \mid a \in \Sigma, i \in \{\perp, 1, \ldots, k\}, x \in \{0, 1\}\}$
  Let $\alpha = (a_1, i_1, x_1) \ldots (a_m, i_m, x_m) \in Q^m$ and $w := a_1^{(i_1)} \ldots a_m^{(i_m)} \in (\Sigma')^m$.
  - succ $:\Leftrightarrow$ for all $1 \leq j \leq m$, $a_j^{(i_j)}$ is the a_j-successor of $a^{(i)}$.
  - $\mathrm{sat}_0 :\Leftrightarrow w \in L(r_{a,i})$ and for all $1 \leq j \leq m$, $x_j = 0$ and $i_j \neq \perp$.
  - $\mathrm{sat}_1 :\Leftrightarrow w \notin L(r_{a,i})$ or there is a $1 \leq j \leq m$ such that $x_j = 1$ or $i_j = \perp$.

  Then $\alpha \in L_{a,i,x}$ iff succ and sat_x hold.
- $F = \{(S, 1, 1)\} \cup \{(a, i, x) \mid a \neq S\}$, meaning either a violation was found or the starting symbol was not S.

$\square$

Theorem 11. *The class of DTWA-recognizable languages is closed under complement.*

Proof. Let $\mathcal{A} = (Q, \Sigma, q_0, \delta, \{q_f\})$ be a DTWA that only moves to q_f at the root. The **backwards configuration graph** $\mathrm{BCG}(\mathcal{A}, t)$ is defined as a tree over $Q \times \mathrm{dom}_t$ with root (q_f, ε). For a node (q, u), the children are all (p, v) such that $(p, v) \to_{\mathcal{A}} (q, u)$. We define $\overline{\mathcal{A}}$ in a way that it performs DFS on the BCG of the input tree and accepts iff the node (q_0, ε) is found.

For that, let $\prec \subseteq (Q \times \mathrm{Dir})^2$ be an arbitrary linear order on $Q \times \mathrm{Dir}$. We set $\overline{\mathcal{A}} = (\overline{Q}, \Sigma, q_0, \overline{\delta}, q_f)$ with $\overline{Q} = \{q_0, q_f\} \cup \{\langle p, (q, d)\rangle \mid p, q \in Q, d \in \mathrm{Dir}\}$. The behavior of $\overline{\delta}$ is described below. Let $\langle q, (q', d)\rangle$ be a state.

Case 1 : In the ordering $\prec$, $(\hat{q}, \hat{d})$ is the next largest element after (q', d). (for q_0 we also consider this case with the $\prec$-minimal pair.)

Case 2 : (q', d) □

Theorem 12. *Let $T \subseteq T_\Sigma$. T is regular iff fcns(T) is regular.*

Proof. $\Rightarrow$ Let $\mathcal{A} = (Q, \Sigma, \Delta, F)$ be a NUTA with $T(\mathcal{A}) = T$. Wlog we assume that $\mathcal{A}$ is normalized. For every transition $\tau = (L_{a,q}, a, q) \in \Delta$, let $\mathcal{B}_{a,q} = (P_{a,q}, Q, p^0_{a,q}, \Delta_{a,q}, F_{a,q})$ be a NFA with $L(\mathcal{B}_{a,q}) = L_{a,q}$. We define $\mathcal{A}_{\mathrm{fcns}} = (Q_{\mathrm{fcns}}, \Gamma, \Delta_{\mathrm{fcns}}, F_{\mathrm{fcns}})$ so that $T(\mathcal{A}_{\mathrm{fcns}}) = \mathrm{fcns}(T)$.

- $Q_{\mathrm{fcns}} = \{q_f, q_\#\} \cup \bigcup_{a \in \Sigma, q \in Q} P_{a,q}$
- $F_{\mathrm{fcns}} = \{q_f\}$
- Δ_{fcns}:
  - $(\#, q_\#)$
  - For all $p \in \bigcup_{a \in \Sigma, q \in Q} F_{a,q}$: $(\#, p)$
  - For all $a \in \Sigma, q \in F$: $(p^0_{a,q}, q_\#, a, q_f)$
  - For all $a \in \Sigma, p \in P, p' \in P, q \in Q$ with $(p, q', p') \in \bigcup_{b \in \Sigma, q' \in Q} \Delta_{b,q'}$: $(p^0_{a,q}, p', b, p)$

Via induction on $t_1 \dots t_n$, one can show that $p \in \Delta^*_{\mathrm{fcns}}(\mathrm{fcns}(t_1 \dots t_n))$ iff there are $q_1, \dots q_n \in Q$ such that $\forall i : q_i \in \Delta^*(t_i)$ and $q_1 \dots q_n \in L(\mathcal{B}_{a,q}$ init $p)$.

$\Leftarrow$ Let $\mathcal{A} = (Q, \Sigma, \Delta, F)$ be an NTA with $T(\mathcal{A}) = \mathrm{fcns}(T)$. We define □