The McNaughton Theorem

McNaughton Theorem

Theorem 1 *Let Σ be an alphabet. Any recognizable subset of Σ^ω can be recognized by a Rabin automaton.*

Determinisation algorithm by S. Safra (1989) uses a special subset construction to obtain a Rabin automaton equivalent to a given Büchi automaton. The Safra algorithm is optimal $2^{O(n \log n)}$.

This proves that recognizable ω-languages are closed under complement (Büchi Theorem).

Oriented Trees

Let Σ be an alphabet of labels.

An oriented tree is a pair of partial functions $t = \langle l, s \rangle$:

- $l : \mathbb{N} \mapsto \Sigma$ denotes the labels of the nodes
- $s : \mathbb{N} \mapsto \mathbb{N}^*$ gives the ordered list of children of each node

$$dom(l) = dom(s) \stackrel{def}{=} dom(t)$$

$\leq$ denotes the successor, and $\preceq$ the lexicographical ordering on tree positions

Safra Trees

Let $A = \langle S, I, T, F \rangle$ be a Büchi automaton.

A *Safra tree* is a pair $\langle t, m \rangle$, where t is a finite oriented tree labeled with non-empty subsets of S, and $m \subseteq dom(t)$ is the set of *marked positions*, such that:

- each marked position is a leaf
- for each $p \in dom(t)$, the union of labels of its children is a strict subset of $t(p)$
- for each $p, q \in dom(t)$, if $p \not\leq q$ and $q \not\leq p$ then $t(p) \cap t(q) = \emptyset$

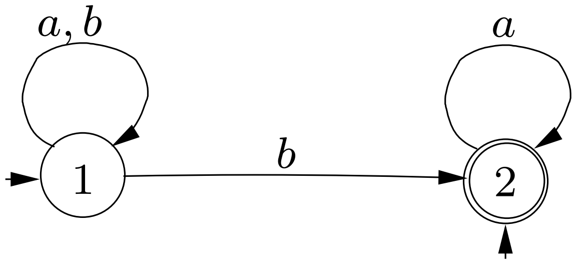
Proposition 1 *A Safra tree has at most $\|S\|$ nodes.*

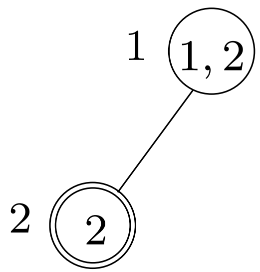
$$
\begin{aligned}
r(p) &= t(p) \setminus \bigcup_{q<p} t(q) \\
\|dom(t)\| &= \sum_{p \in dom(t)} 1 \leq \sum_{p \in dom(t)} \|r(p)\| \leq \|S\|
\end{aligned}
$$

Initial State

We build a Rabin automaton $B = \langle S_B, i_B, T_B, \Omega_B \rangle$, where:

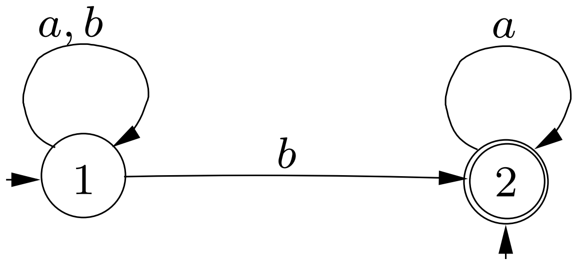
- S_B is the set of all Safra trees $\langle t, m \rangle$ labeled with subsets of S
- $i_B = \langle t, m \rangle$ is the Safra tree defined as either:
  - $dom(t) = \{\epsilon\}, t(\epsilon) = I$ and $m = \emptyset$ if $I \cap F = \emptyset$
  - $dom(t) = \{\epsilon\}, t(\epsilon) = I$ and $m = \{\epsilon\}$ if $I \subseteq F$
  - $dom(t) = \{\epsilon, 0\}, t(\epsilon) = I, t(0) = I \cap F$ and $m = \{0\}$ if $I \cap F \neq \emptyset$

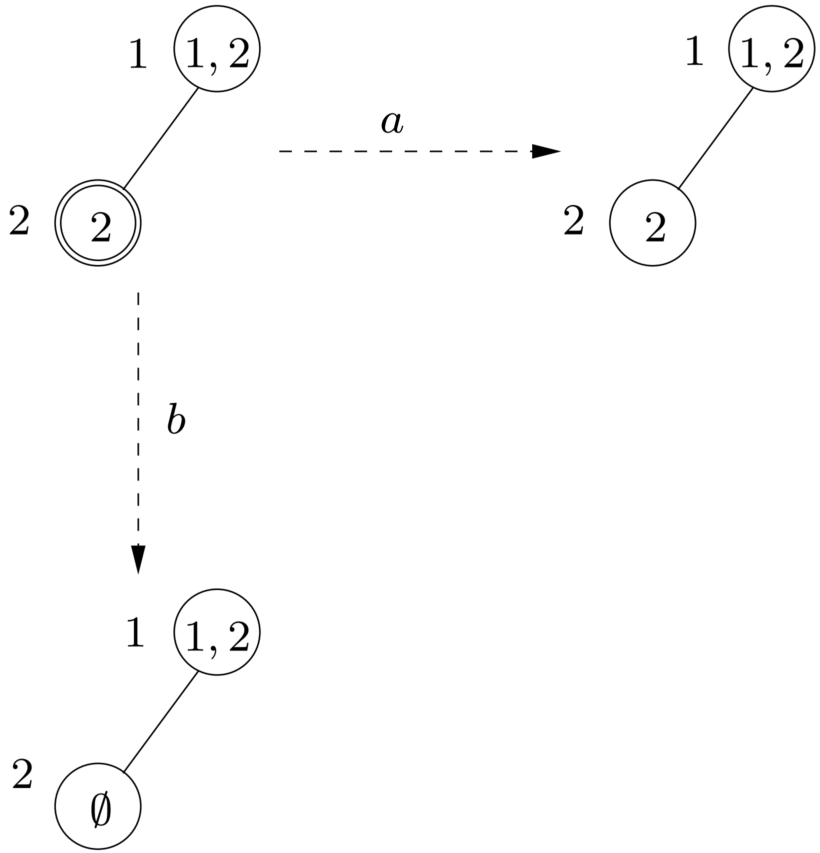




Classical Subset Move

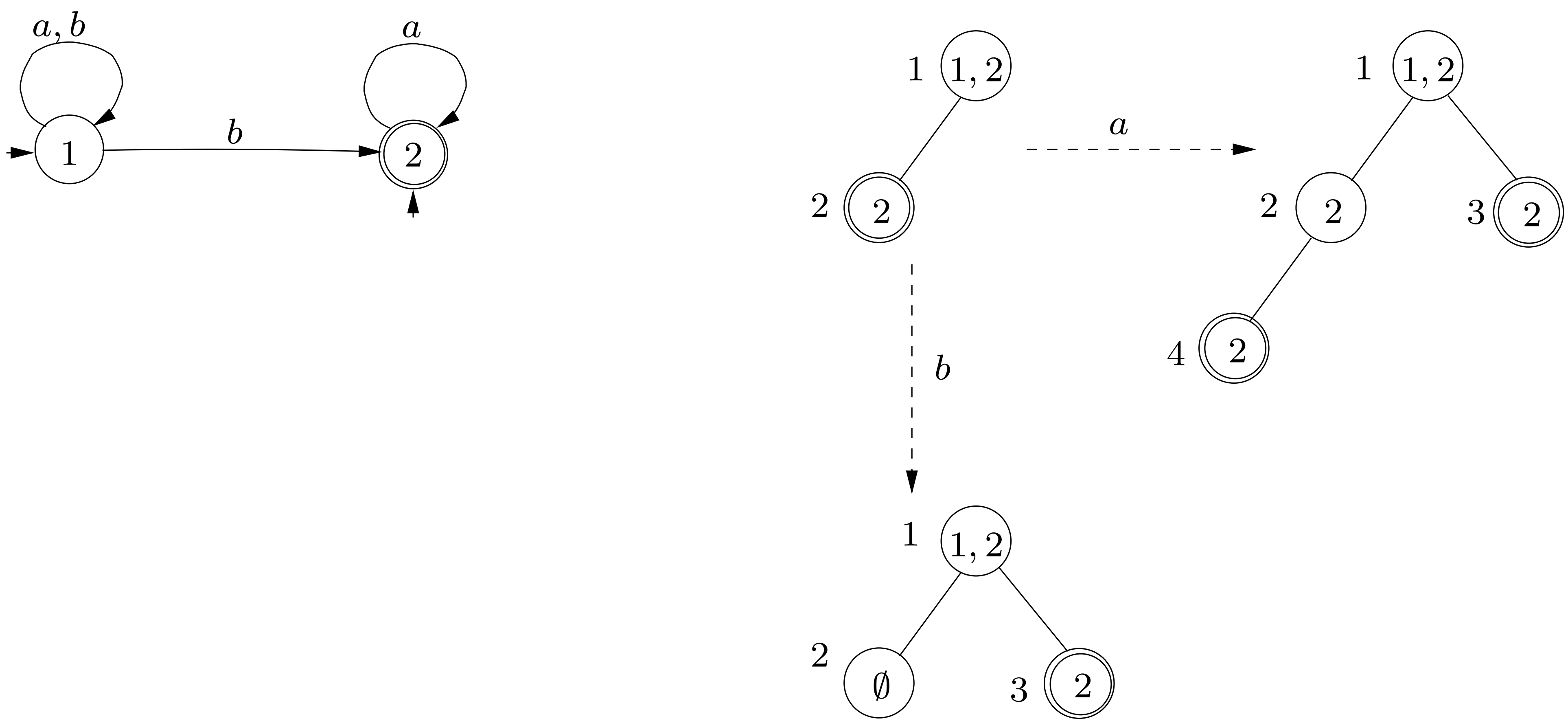
[Step 1] $\langle t_1, m_1 \rangle$ is the tree with $dom(t_1) = dom(t), m_1 = \emptyset$, and
$t_1(p) = \{s' \mid s \xrightarrow{\alpha} s',\ s \in t(p)\}$, for all $p \in dom(t)$





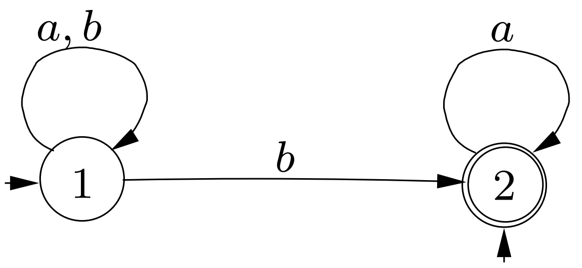
Spawn New Children

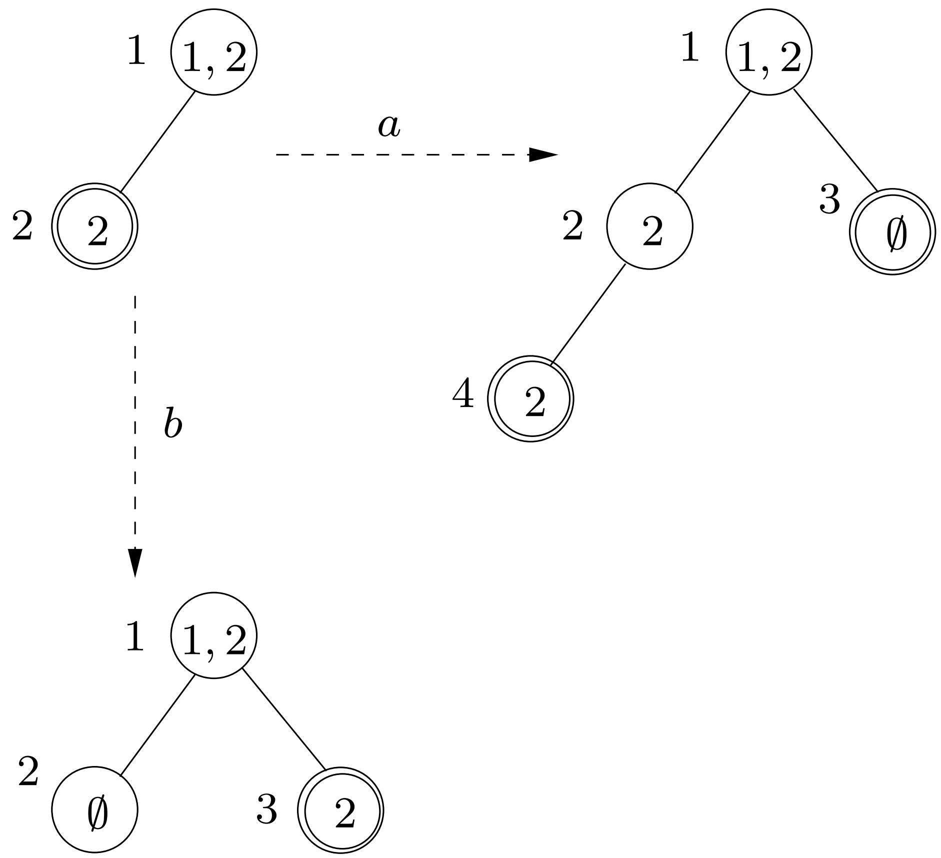
[Step 2] $\langle t_2, m_2 \rangle$ is the tree such that, for each $p \in dom(t_1)$, if $t_1(p) \cap F \neq \emptyset$ we add a new child to the right, identified by the first available id, and labeled $t_1(p) \cap F$, and m_2 is the set of all such children



Horizontal Merge

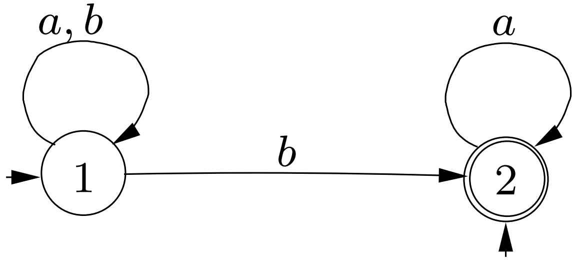
[Step 3] $\langle t_3, m_3 \rangle$ is the tree with $dom(t_3) = dom(t_2)$, $m_3 = m_2$, such that, for all $p \in dom(t_3)$, $t_3(p) = t_2(p) \setminus \bigcup_{q \prec p} t_2(q)$

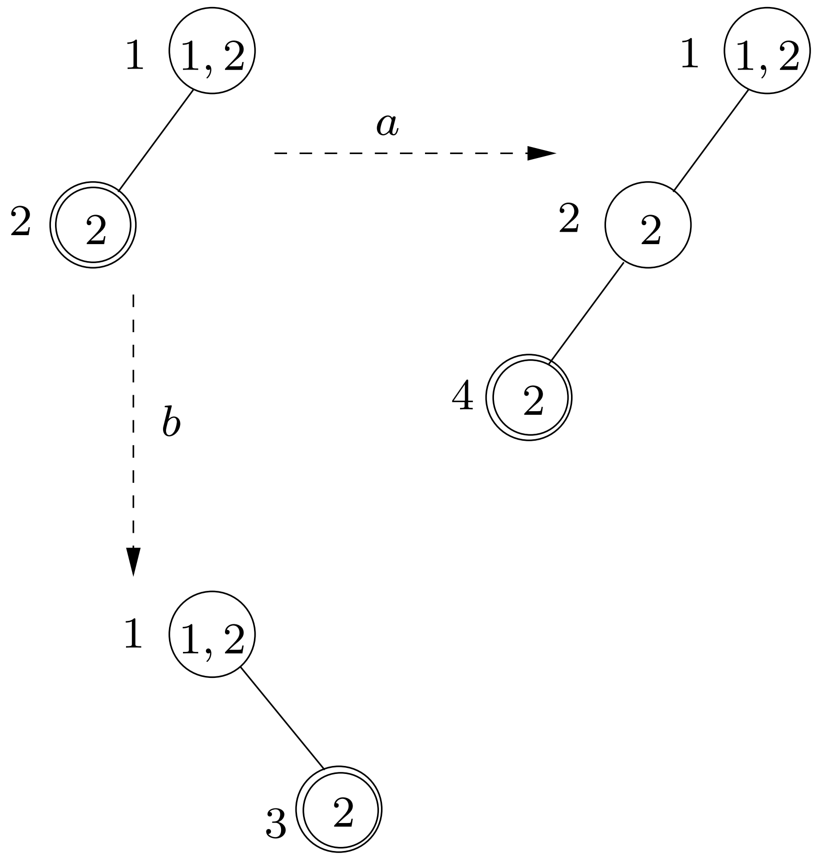




Delete Empty Nodes

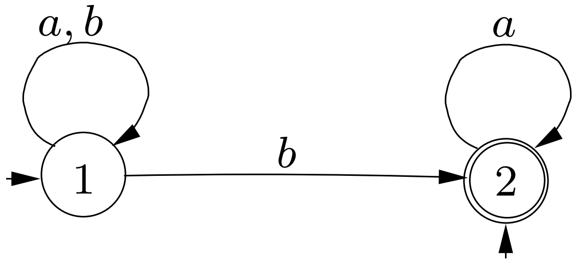
[Step 4] $\langle t_4, m_4 \rangle$ is the tree such that $dom(t_4) = dom(t_3) \setminus \{p \mid t_3(p) = \emptyset\}$ and $m_4 = m_3 \setminus \{p \mid t_3(p) = \emptyset\}$

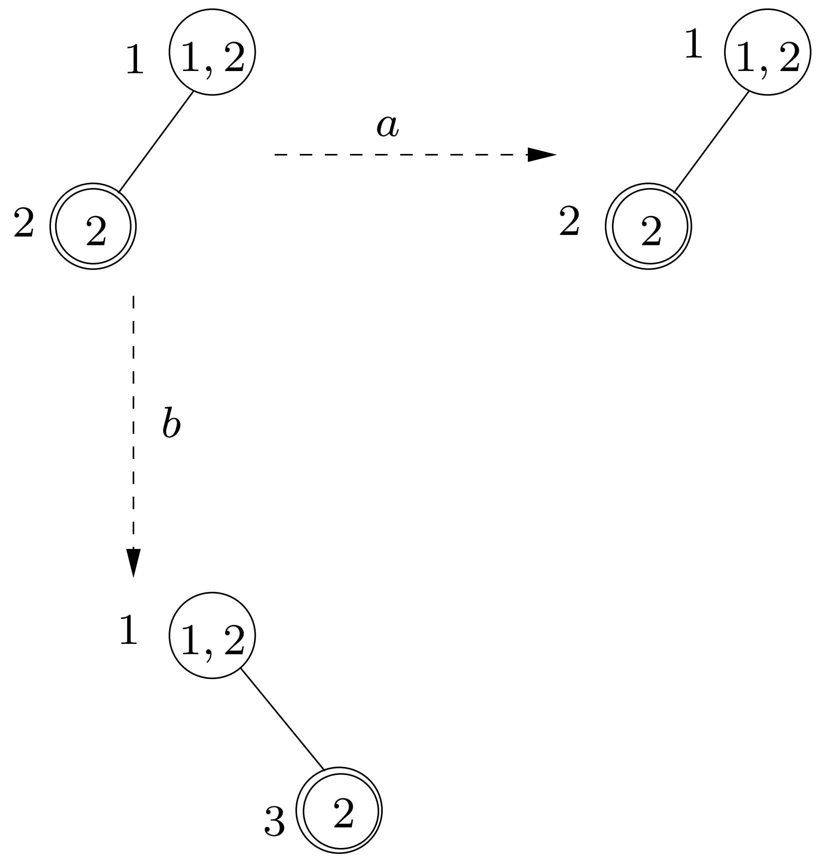




Vertical Merge

[Step 5] $\langle t_5, m_5 \rangle$ is $m_5 = m_4 \cup V$, $dom(t_5) = dom(t_4) \setminus \{q \mid p \in V,\ p < q\}$,
$V = \{p \in dom(t_4) \mid t_4(p) = \bigcup_{p<q} t_4(q)\}$

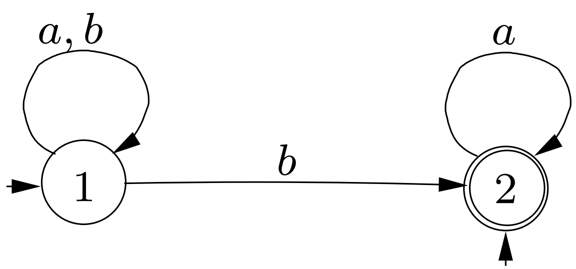


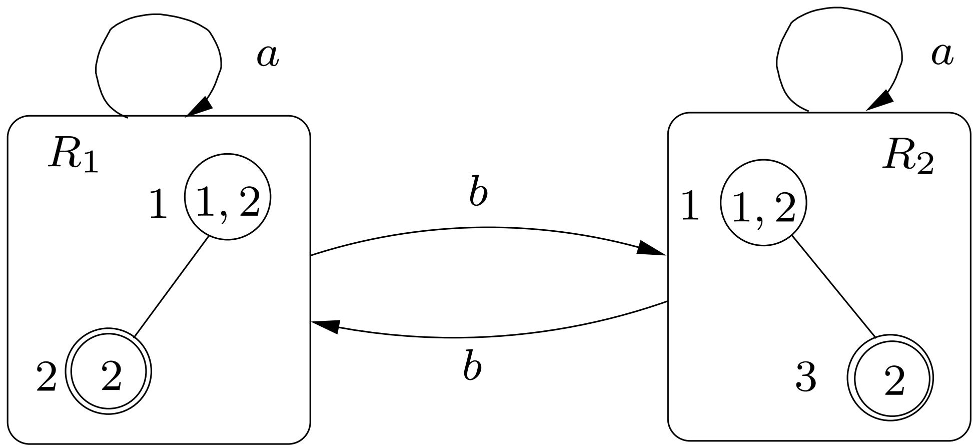


Accepting Condition

The Rabin accepting condition is defined as
$\Omega_B = \{(N_q, P_q) \mid q \in \bigcup_{\langle t,m\rangle \in S_B} dom(t)\}$, where:

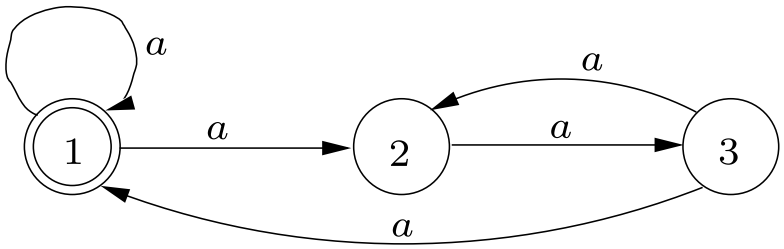
- $N_q = \{\langle t, m\rangle \in S_B \mid q \notin dom(t)\}$
- $P_q = \{\langle t, m\rangle \in S_B \mid q \in m\}$

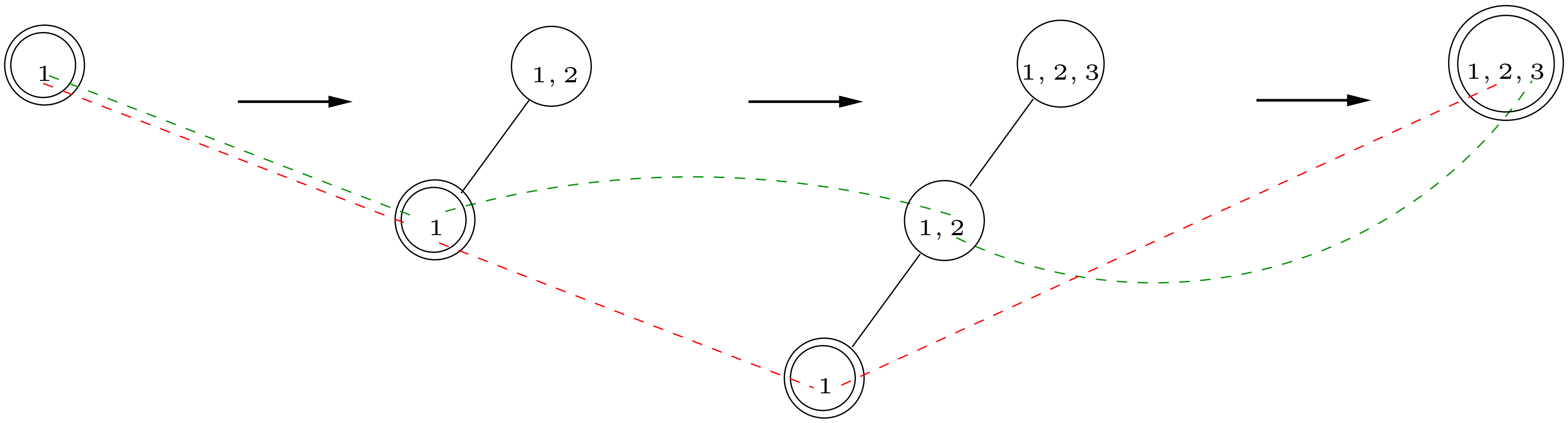




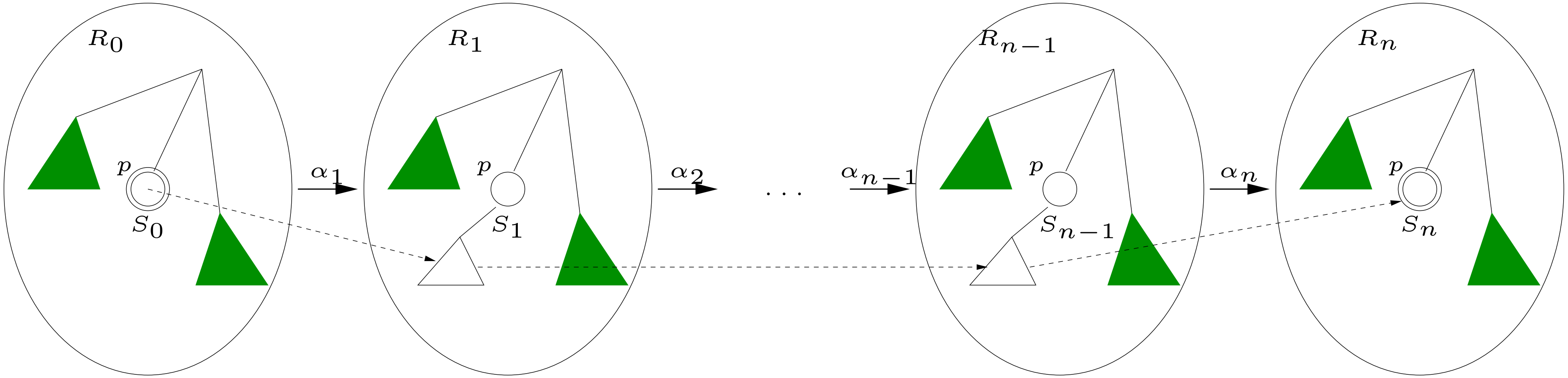
$$\Omega_B = \{(\{R_1\}, \{R_2\}), (\{R_2\}, \{R_1\})\}$$

Example





Correctness of Safra Construction



Lemma 1 *For* $0 \leq i \leq n-1$, $S_{i+1} \subseteq T(S_i, \alpha_{i+1})$*. Moreover, for every* $q \in S_n$*, there is a path in* A *starting in some* $q_0 \in S_0$*, ending in* q *and visiting at least one final state after its origin.*

An infinite accepting path in B corresponds to an infinite accepting path in A (König's Lemma)

Correctness of Safra Construction

Conversely, an infinite accepting path of A over $u = \alpha_0\alpha_1\alpha_2\ldots$

$$\pi \ : \ q_0 \xrightarrow{\alpha_0} q_1 \xrightarrow{\alpha_1} q_2 \ldots$$

corresponds to a unique infinite path of B:

$$i_B = R_0 \xrightarrow{\alpha_0} R_1 \xrightarrow{\alpha_1} R_2 \ldots$$

where each q_i belongs to the root of R_i

If the root is marked infinitely often, then u is accepted. Otherwise, let n_0 be the largest number such that the root is marked in R_{n_0}. Let $m > n_0$ be the smallest number such that q_m is repeated infinitely often in π.

Since $q_m \in F$ it appears in a child of the root. If it appears always on the same position p_m, then the path is accepting. Otherwise it appears to the left of p_m from some n_1 on (step 3). This left switch can only occur a finite number of times.

Complexity of the Safra Construction

Given a Büchi automaton with n states, how many states we need for an equivalent Rabin automaton?

- The upper bound is $2^{\mathcal{O}(n \log n)}$ states
- The lower bound is of at least $n!$ states

Maximum Number of Safra Trees

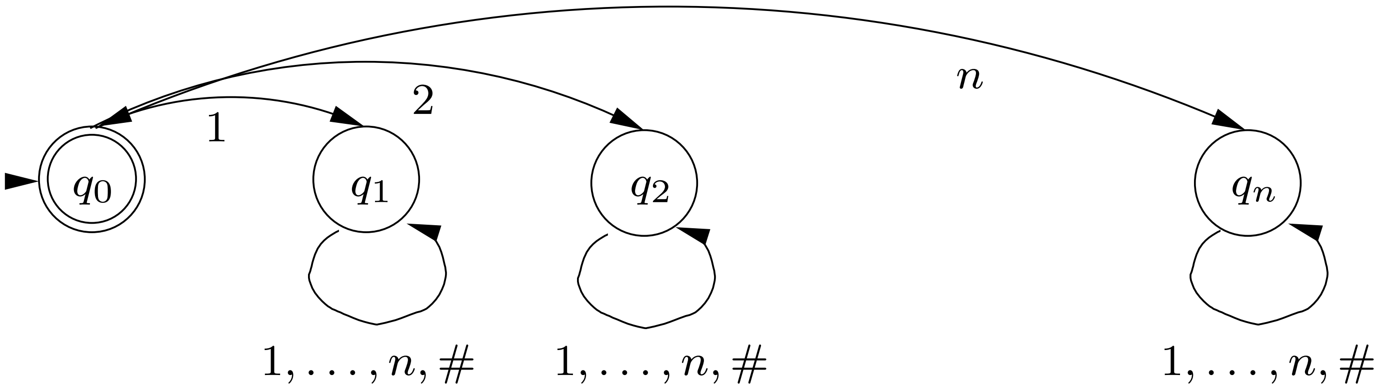
Each Safra tree has at most n nodes.

A Safra tree $\langle t, m \rangle$ can be uniquely described by the functions:

- $S \to \{0, \ldots, n\}$ gives for each $s \in S$ the characteristic position $p \in dom(t)$ such that $s \in t(p)$, and s does not appear below p
- $\{1, \ldots, n\} \to \{0, 1\}$ is the marking function
- $\{1, \ldots, n\} \to \{0, \ldots, n\}$ is the parent function
- $\{1, \ldots, n\} \to \{0, \ldots, n\}$ is the older brother function

Altogether we have at most $(n+1)^n \cdot 2^n \cdot (n+1)^n \cdot (n+1)^n \leq (n+1)^{4n}$ Safra trees, hence the upper bound is $2^{\mathcal{O}(n \log n)}$.

The Language L_n



$\alpha \in L_n$ if there exist $i_1, \ldots, i_n \in \{1, \ldots, n\}$ such that

- $\alpha_k = i_1$ is the first occurrence of i_1 in α and $q_0 \xrightarrow{\alpha_0 \ldots \alpha_k} q_{i_1}$
- the pairs $i_1 i_2, i_2 i_3, \ldots, i_n i_1$ appear infinitely often in α.

Example 1

$(3\#32\#21\#1)^\omega \in L_3$

$(312\#)^\omega \notin L_3$

The Language L_n

Lemma 2 *(Permutation) For each permutation* $i_1, i_2, \ldots, i_n$ *of* $1, 2, \ldots, n$*, the infinite word* $(i_1 i_2 \ldots i_n \#)^\omega \notin L_n$.

Lemma 3 *(Union) Let* $A = (S, i, T, \Omega)$ *be a Rabin automaton with* $\Omega = \{\langle N_1, P_1 \rangle, \ldots, \langle N_k, P_k \rangle\}$ *and* ρ_1, ρ_2, ρ *be runs of* A *such that*

$$\inf(\rho_1) \cup \inf(\rho_2) = \inf(\rho)$$

If ρ_1 *and* ρ_2 *are not successful, then* ρ *is not successful either.*

Proving the $n!$ Lower Bound

Suppose that A recognizes L_n. We need to show that A has $\geq n!$ states.

Let $\alpha = i_1, i_2, \ldots, i_n$ and $\beta = j_1, j_2, \ldots, j_n$ be two permutations of $1, 2, \ldots, n$. Then the words $(i_1 i_2 \ldots i_n \#)^\omega$ and $(j_1 j_2 \ldots j_n \#)^\omega$ are not accepted.

Let ρ_α, ρ_β be the non-accepting runs of A over α and β, respectivelly.

Claim 1 $\inf(\rho_\alpha) \cap \inf(\rho_\beta) = \emptyset$

Then A must have $\geq n!$ states, since there are $n!$ permutations.

Proving the $n!$ Lower Bound

By contradiction, assume $q \in \inf(\rho_\alpha) \cap \inf(\rho_\beta)$. Then we can build a run ρ such that $\inf(\rho) = \inf(\rho_1) \cup \inf(\rho_2)$ and α, β appear infinitely often. By the union lemma, ρ is not accepting.

$$\begin{array}{ccccccccc} i_1 & \dots & i_{k-1} & i_k & i_{k+1} & \dots & i_{l-1} & i_l & \dots & i_n \\ = & & = & \neq & & & & & & \\ j_1 & \dots & j_{k-1} & j_k & j_{k+1} & \dots & j_{r-1} & j_r & \dots & j_n \end{array}$$

$$i_k \quad i_{k+1}, \quad \dots \quad i_l = j_k \quad j_{k+1}, \quad \dots \quad j_{r-1} \quad j_r = i_k$$

The new word is accepted since the pairs $i_k i_{k+1}, \dots, j_k j_{k+1}, \dots, j_{r-1} i_k$ occur infinitely often. Contradiction with the fact that ρ is not accepting.

Büchi Complementation Theorem

Büchi Complementation Theorem

Theorem 2 *For every Büchi automaton A there exists a Büchi automaton B such that $\mathcal{L}(A) = \overline{\mathcal{L}(B)}$.*

Already a consequence of McNaughton Theorem, since from A we can build a Rabin automaton R, complement it to $\overline{R}$, and build B from $\overline{R}$.

Next we present a direct proof.

Congruences

Definition 1 *An equivalence relation* $R \subseteq \Sigma^* \times \Sigma^*$ *is said to be a* left-congruence *iff for all* $u, v, w \in \Sigma^*$ *we have* $u \cong v \Rightarrow wu \cong wv$.

Definition 2 *An equivalence relation* $R \subseteq \Sigma^* \times \Sigma^*$ *is said to be a* right-congruence *iff for all* $u, v, w \in \Sigma^*$ *we have* $u \; R \; v \Rightarrow uw \; R \; vw$.

Definition 3 *An equivalence relation* $R \subseteq \Sigma^* \times \Sigma^*$ *is said to be a* congruence *iff it is both a left- and a right-congruence.*

Ex: the Myhill-Nerode equivalence $\sim_L$ is a right-congruence.

Congruences

Let $A = \langle S, I, T, F \rangle$ be a Büchi automaton and $s, s' \in S$.

$$W_{s,s'} = \{w \in \Sigma^* \mid s \xrightarrow{w} s'\}$$

For $s, s' \in S$ and $w \in \Sigma^*$, we denote $s \rightarrow_w^F s'$ iff $s \xrightarrow{w} s'$ visiting a state from F.

$$W_{s,s'}^F = \{w \in \Sigma^* \mid s \rightarrow_w^F s'\}$$

For any two words $u, v \in \Sigma^*$ we have $u \cong v$ iff for all $s, s' \in S$ we have:

- $s \xrightarrow{u} s' \iff s \xrightarrow{v} s'$, and
- $s \rightarrow_u^F s' \iff s \rightarrow_v^F s'$.

The relation $\cong$ is a congruence of finite index on Σ^*

Congruences

Let $[w]_{\cong}$ denote the equivalence class of $w \in \Sigma^*$ w.r.t. $\cong$.

Lemma 4 *For any* $w \in \Sigma^*$, $[w]_{\cong}$ *is the intersection of all sets of the form* $W_{s,s'}, W^F_{s,s'}, \overline{W_{s,s'}}, \overline{W^F_{s,s'}}$, *containing* w.

$$T_w = \bigcap_{w \in W_{s,s'}} W_{s,s'} \cap \bigcap_{w \in W^F_{s,s'}} W^F_{s,s'} \cap \bigcap_{w \in \overline{W_{s,s'}}} \overline{W_{s,s'}} \cap \bigcap_{w \in \overline{W^F_{s,s'}}} \overline{W^F_{s,s'}}$$

We show that $[w]_{\cong} = T_w$.

"$\subseteq$" If $u \cong w$ then clearly $u \in T_w$.

Congruences

"$\supseteq$" Let $u \in T_w$

- if $s \xrightarrow{w} s'$, then $w \in W_{s,s'}$, hence $u \in W_{s,s'}$, then $s \xrightarrow{u} s'$ as well.
- if $s \not\xrightarrow{w} s'$, then $w \in \overline{W_{s,s'}}$, hence $u \in \overline{W_{s,s'}}$, then $s \not\xrightarrow{u} s'$.

Also,

- if $s \to_w^F s'$, then $w \in W^F_{s,s'}$, hence $u \in W^F_{s,s'}$, then $s \to_u^F s'$ as well.
- if $s \not\to_w^F s'$, then $w \in \overline{W^F_{s,s'}}$, hence $u \in \overline{W^F_{s,s'}}$, then $s \not\to_u^F s'$.

Then $u \cong w$.

This lemma gives us a way to compute the $\cong$-equivalence classes.

Outline of the proof

We prove that:

$$\mathcal{L}(A) = \bigcup_{VW^\omega \cap \mathcal{L}(A) \neq \emptyset} VW^\omega$$

where V, W are $\cong$-equivalence classes

Then we have

$$\Sigma^\omega \setminus \mathcal{L}(A) = \bigcup_{VW^\omega \cap \mathcal{L}(A) = \emptyset} VW^\omega$$

Finally we obtain an algorithm for complementation of Büchi automata

Saturation

Definition 4 *A congruence relation* $R \subseteq \Sigma^* \times \Sigma^*$ saturates *an* ω*-language* L *iff for all* R*-equivalence classes* V *and* W*, if* $VW^\omega \cap L \neq \emptyset$ *then* $VW^\omega \subseteq L$*.*

Lemma 5 *The congruence relation* $\cong$ *saturates* $\mathcal{L}(A)$*.*

Every word belongs to some VW^ω

Let $\alpha \in \Sigma^\omega$ be an infinite word.

Since $\cong$ is an equivalence relation, there exists a mapping $\varphi : \Sigma^+ \to \Sigma^+_{/\cong}$ such that $\varphi(u) = [u]_{/\cong}$, for all $u \in \Sigma^+$.

Then there exists a Ramseyan factorization of $\alpha = uv_0v_1v_2\ldots$ such that $\varphi(v_i) = [v]_{/\cong}$ for some $v \in \Sigma^+$ and for all $i \geq 0$.

Together with the saturation lemma, this proves

$$\mathcal{L}(A) = \bigcup_{VW^\omega \cap \mathcal{L}(A) \neq \emptyset} VW^\omega$$

Complementation of Büchi Automata

Theorem 3 *For any Büchi automaton A there exists a Büchi automaton $\overline{A}$ such that $\mathcal{L}(\overline{A}) = \Sigma^\omega \setminus \mathcal{L}(A)$.*

$$\mathcal{L}(A) = \bigcup_{VW^\omega \cap \mathcal{L}(A) \neq \emptyset} VW^\omega$$

where V, W are $\cong$-equivalence classes

$$\Sigma^\omega \setminus \mathcal{L}(A) = \bigcup_{VW^\omega \cap \mathcal{L}(A) = \emptyset} VW^\omega$$

Ramseyan Factorizations

Ramsey Theorem

Theorem 4 *Let X be some countably infinite set and colour the subsets of X of size n in c different colours. Then there exists some infinite subset M of X such that the size n subsets of M all have the same colour.*

A Particular Case of Ramsey Theorem

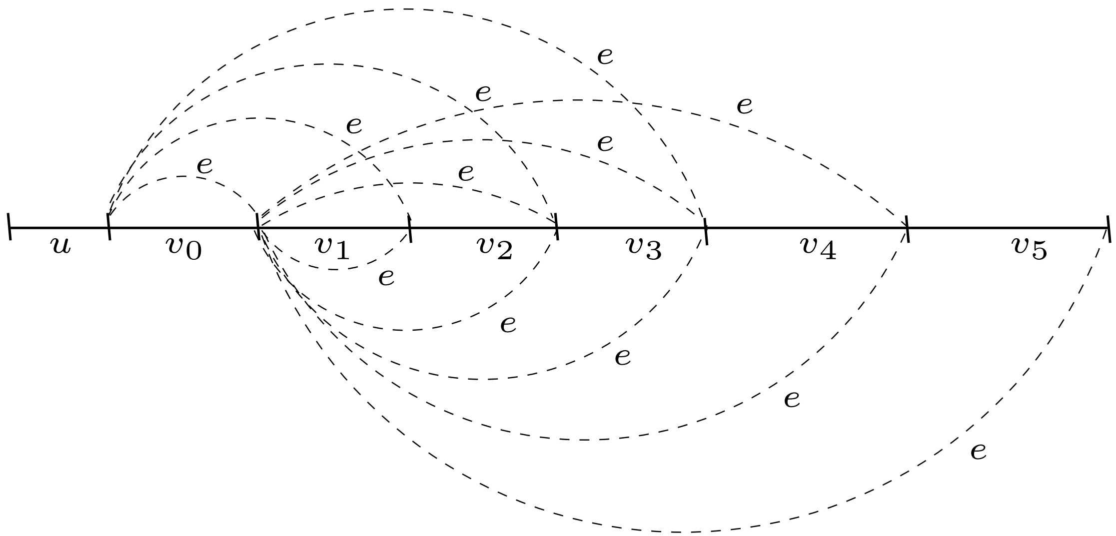
Let $\alpha \in \Sigma^\omega$ be an infinite word.

A *factorization* of α is an infinite sequence $\{\alpha_i\}_{i=0}^{\infty}$ of finite words such that $\alpha = \alpha_0\alpha_1\ldots$

Let E be a finite set of colors and $\varphi : \Sigma^+ \to E$. A factorization $\alpha = uv_0v_1v_2\ldots$ is said to be *Ramseyan for φ* if there exists $e \in E$ such that

$$\varphi(v_iv_{i+1}\ldots v_{i+j}) = e$$

for all $i \leq j$.



A Particular Case of Ramsey Theorem

Theorem 5 *Let $\varphi : \Sigma^+ \to E$ be a map from Σ^+ into a finite set E. Then every infinite word of Σ^ω admits a Ramseyan factorization for φ.*

Let $\{U_i\}_{i=0}^{\infty}$ be an infinite sequence of infinite subsets of $\mathbb{N}$ defined as:

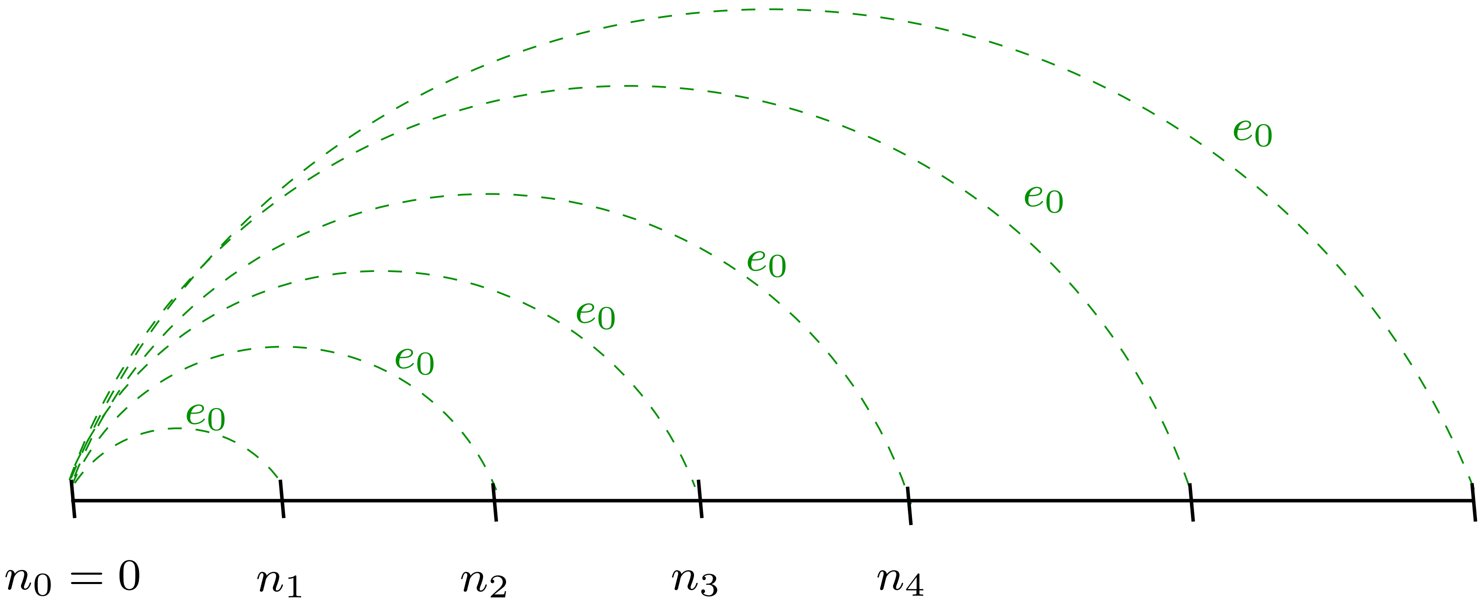
$$
\begin{aligned}
U_0 &= \mathbb{N} \\
U_{i+1} &= \{n \in U_i \mid \varphi(\alpha(\min U_i, n)) = e_i\}
\end{aligned}
$$

where $e_i \in E$ is chosen such that the set U_{i+1} is infinite (show the existence of e_i)

Since E is finite, there exists an infinite subsequence of integers $i_0, i_1, \ldots$ such that $e_{i_0} = e_{i_1} = \ldots = e$.

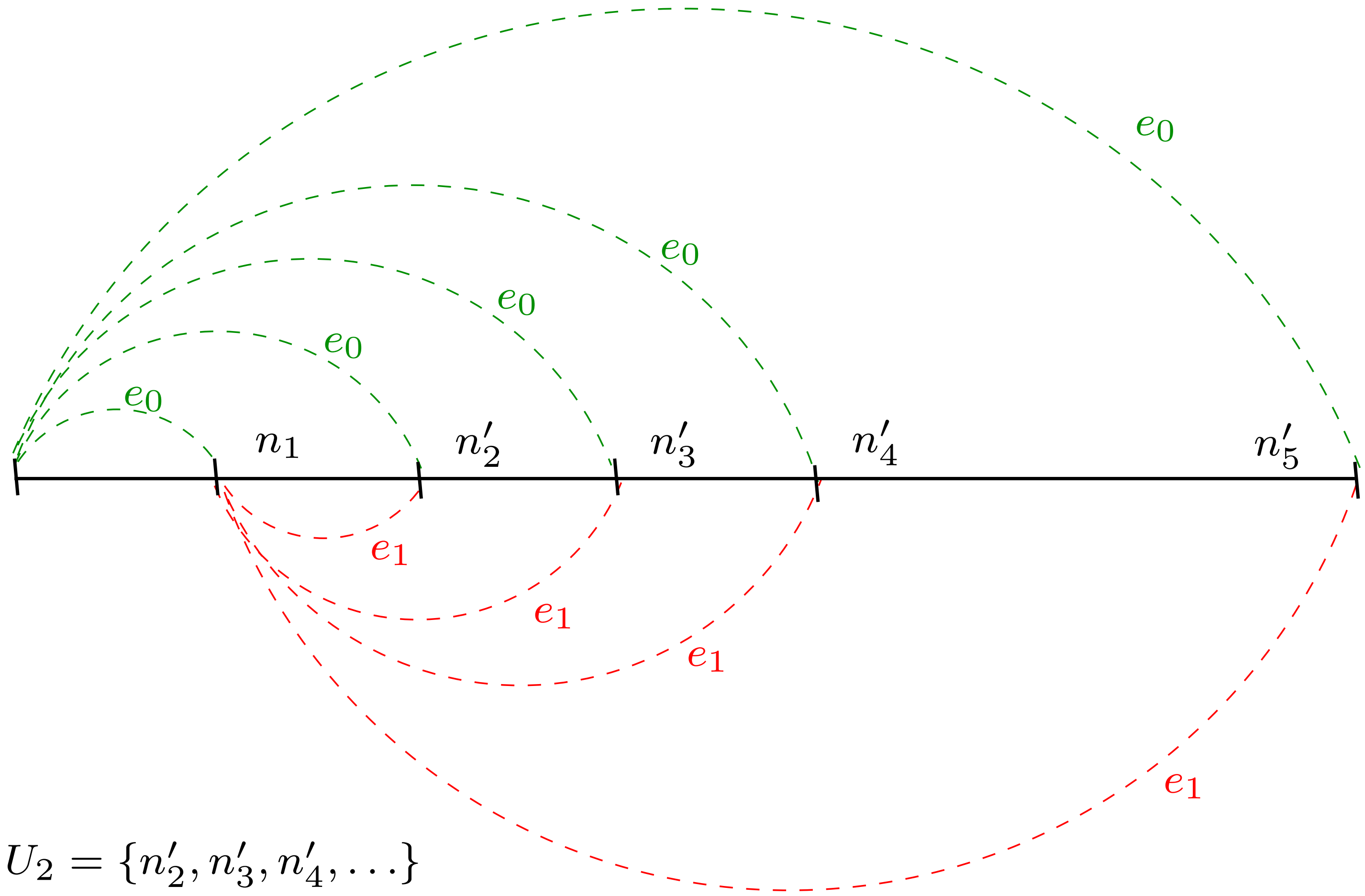
Then $v_j = \alpha(n_{i_j}, n_{i_{j+1}})$ is the required factorization.

A Particular Case of Ramsey Theorem



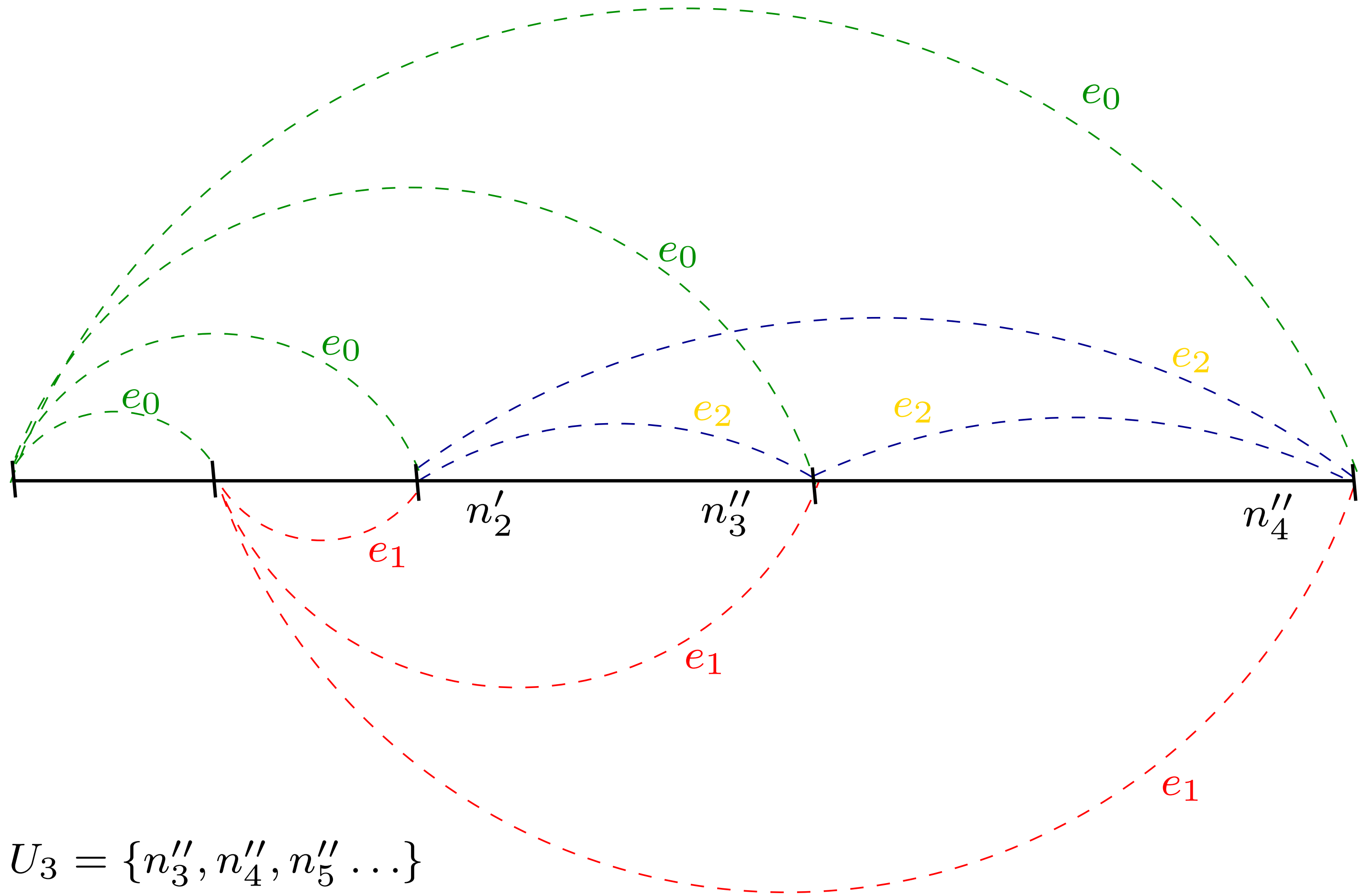
$U_1 = \{n_1, n_2, n_3, n_4, \ldots\}$

A Particular Case of Ramsey Theorem



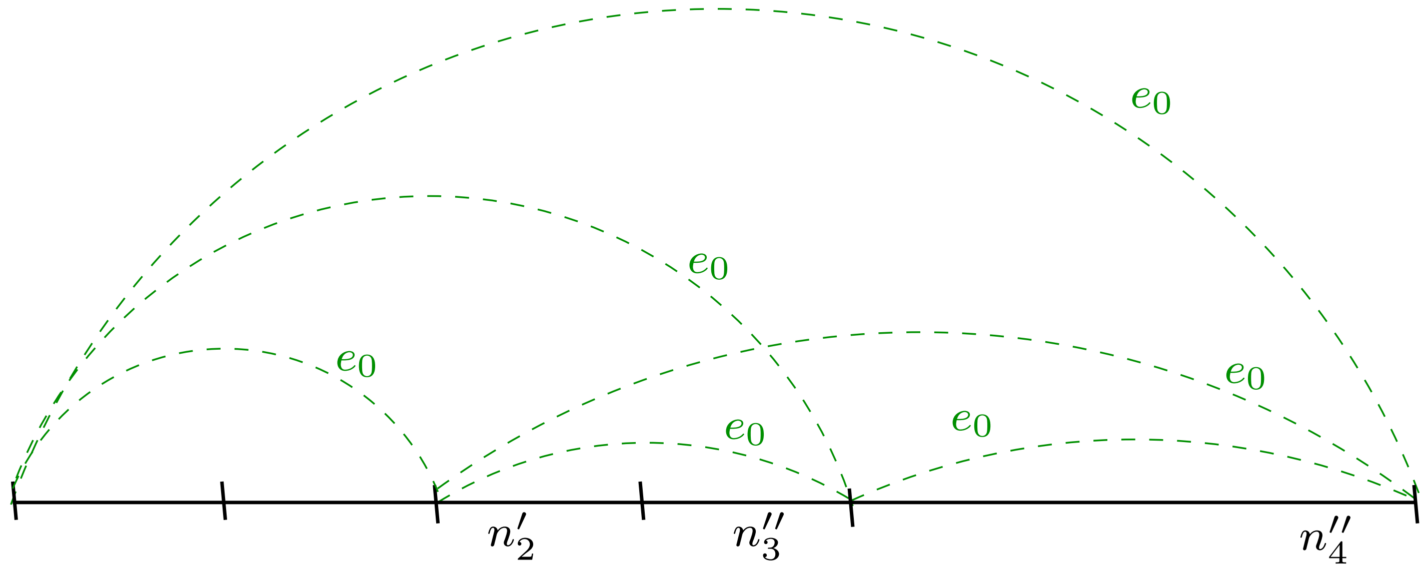
$U_2 = \{n'_2, n'_3, n'_4, \ldots\}$

A Particular Case of Ramsey Theorem



$U_3 = \{n''_3, n''_4, n''_5 \ldots\}$

A Particular Case of Ramsey Theorem



$U_3 = \{n''_3, n''_4, \ldots\}$