Quantum Finite State Automata

Sicheol Sung

October 2023

Introduction

We will see...

1. The definition of 2-way quantum FA (2QFA).
2. 2QFAs are strictly more powerful than 2-way probabilistic FAs, under bounded error and polynomial time constraints.
3. 1QFAs are strictly less powerful than DFAs.

Deterministic Finite State Automata

Definition (DFA)

A *deterministic finite state automaton (DFA)* is a 5-tuple $(Q, \Sigma, \delta, q_0, F)$, where

1. Q is a finite set of *states*;
2. Σ is a finite *input alphabet*;
3. $\delta : Q \times \Sigma \to Q$ is a *transition function*;
4. $q_0 \in Q$ is the *initial state*; and
5. $F \subseteq Q$ is a set of *(final) accepting states*.

2-Way Deterministic Finite State Automata

Definition

Definition (2DFA)

A *2-way deterministic finite state automaton (2DFA)* is a 6-tuple $(Q, \Sigma, \delta, q_0, Q_{\mathrm{acc}}, Q_{\mathrm{rej}})$, where

1. $\delta : Q \times \Gamma \to Q \times \{-1, 0, 1\}$ is a *transition function*; and
2. $Q_{\mathrm{acc}} \subseteq Q$ and $Q_{\mathrm{rej}} \subseteq Q$ are the sets of *accepting states* and *rejecting states*, respectively.

2-Way Deterministic Finite State Automata

Details

Details:

1. $Q_{\text{non}} := Q \setminus (Q_{\text{acc}} \cup Q_{\text{rej}})$;
2. $q_0 \in Q_{\text{non}}$;
3. $Q_{\text{acc}} \cap Q_{\text{rej}} = \emptyset$;
4. ¢ $\notin \Sigma$ and \$ $\notin \Sigma$ are the *start of string* and *end of string* symbols, respectively; and
5. The *tape alphabet* $\Gamma := \Sigma \cup \{¢, \$\}$.

2-Way Deterministic Finite State Automata

Tapes

A *tape* is a mapping $x : \mathbb{Z}_n \to \Gamma$, where $n =: |x|$ is the length of the tape. (At this point, we assume a circular tape.)

For a string $w = w_1 \cdots w_{|w|} \in \Sigma^*$, we define the tape x_w of w as

1. $x_w(0) := ¢$,
2. $x_w(i) := w_i$ for $1 \leq i \leq |w|$, and
3. $x_w(|w| + 1) := \$$.

Fix a 2DFA $M := (Q, \Sigma, \delta, q_0, Q_{\mathrm{acc}}, Q_{\mathrm{rej}})$ and a tape x with length n. $C_n := Q \times \mathbb{Z}_n$ is the set of *configurations* of M.

The *time-evolution operator* $U_\delta^x : C_n \to C_n$ of M on tape x is defined as:

$$U_\delta^x(q, k) := (p, k + d),$$

where $\delta(q, x(k)) =: (p, d) \in Q \times \{-1, 0, 1\}$.

For each time step t, let $(q_t, \_) := (U_\delta^x)^t(q_0, 0)$.
If $q_t \in Q_{\mathrm{acc}}$, then M *accepts* a string w at a time step t.

Definition (2PFA)

A *2-way probabilistic finite state automaton (2PFA)* is a 6-tuple $(Q, \Sigma, \delta, q_0, Q_{\mathrm{acc}}, Q_{\mathrm{rej}})$, where

$$\delta : \underline{(Q \times \Gamma)} \times \underline{(Q \times \{-1, 0, 1\})} \to \mathbb{R}.$$

A *distribution* of M on x is a probabilistic distribution $D : C_n \to \mathbb{R}$.

1. For each $c \in C_n$, $[\![c]\!]$ is denotes the distribution $c \mapsto 1$.
2. We can denote a distribution D by $\sum_{c \in C_n} p_c \cdot [\![c]\!]$, where $p_c := D(c)$.

2-Way Probabilistic Finite State Automata

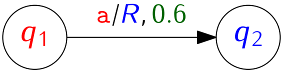
Operator

$$\delta : \underline{(Q \times \Gamma)} \times \underline{(Q \times \{-1, 0, 1\})} \to \mathbb{R}$$

The operator $U_\delta^x : D \mapsto U_\delta^x D$ is defined as:

$$U_\delta^x [\![q, k]\!] := \sum_{q', d} \delta(\underline{q, x(k)}, \underline{q', d}) \cdot [\![q', k + d]\!],$$

and is extended to all distributions of M on x by linearity.



Note

δ is restricted to U_δ^x be valid. What should the restriction be?

Known Results

Known results:

1. DFA and 2DFA have the same power of expression (regular).
2. Under constant error bound and *exponential* expected time constraints, 2PFA can express the non-regular language $\{\mathtt{a}^n\mathtt{b}^n \mid n > 0\}$.
3. Under constant error bound and *polynomial* expected time constraints, 2PFA cannot express non-regular languages.

Theorem (Dwork89)

For any 2PFA recognizing a non-regular language with a constant error bound, the 2PFA must take exponential expected time with respect to the length of the input.

Definition of 2-Way Quantum Finite State Automata

Definition (2QFA)

A *2-way quantum finite state automaton (2QFA)* is a 6-tuple $(Q, \Sigma, \delta, q_0, Q_{\text{acc}}, Q_{\text{rej}})$, where

$$\delta : Q \times \Gamma \times Q \times \{-1, 0, 1\} \to \mathbb{C}.$$

A *superposition* of M on x is a $|C_n|$-dimensional quantum state.

1. $\mathcal{H}_n$ denotes the set of all superpositions.
2. For each $c \in C_n$, $|c\rangle$ denotes the unit vector with value 1 at c.
3. For $|\psi\rangle = \sum_{c \in C_n} \alpha_c |c\rangle$, $\alpha_c \in \mathbb{C}$ is the *amplitude* of c in $|\psi\rangle$.

Transitions of 2QFA

$$\delta : \underline{(Q \times \Gamma)} \times \underline{(Q \times \{-1, 0, 1\})} \to \mathbb{C}.$$

For a tape x, the *time-evolution operator* $U_\delta^x : \mathcal{H}_n \to \mathcal{H}_n$ of M on tape x is defined as:

$$U_\delta^x |q, k\rangle := \sum \delta(\underline{q, x(k)}, \underline{q', d}) \cdot |q', k + d\rangle,$$

and is extended to all $|\psi\rangle \in \mathcal{H}_n$ by linearity.

Note

- δ is restricted to U_δ^x be valid, that is, U_δ^x must be *unitary*.
- The configuration of M is in "superposition", we must "carefully observe" whether the configuration is accepting.

Observables

An *observable* $\mathcal{O}$ is a decomposition $\{E_1, \dots, E_k\}$ of the Hilbert space $\mathcal{H}_n$ into subspaces, where

- $\mathcal{H}_n = E_1 \oplus E_2 \cdots \oplus E_k$; and
- E_j are pairwise orthogonal.

Consider that we observe $|\psi\rangle \in \mathcal{H}_n$ with an observable $\mathcal{O} = \{E_1, \dots, E_k\}$.

Let $|\psi_j\rangle$ be the projection of $|\psi\rangle$ onto E_j.
Then, after the observation,

1. We observe each outcome j with probability $\||\psi_j\rangle\|^2$.
2. The machine "collapse" to $\frac{1}{\||\psi_j\rangle\|}|\psi_j\rangle$.

Note

It is similar to *conditional probabilities*.

Observables (Examples)

Let $|\psi\rangle = \frac{1}{2}|00\rangle + \frac{1}{2}|01\rangle + \frac{1}{2}|10\rangle + \frac{1}{2}|11\rangle$.

1. Using observable $\{\langle|00\rangle\rangle, \langle|10\rangle\rangle, \langle|01\rangle\rangle, \langle|11\rangle\rangle\}$, with the same probability 0.25,
   - $|\psi\rangle$ collapses to $|00\rangle, |10\rangle, |01\rangle$, or $|11\rangle$.

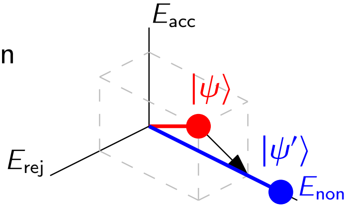
2. Using observable $\{\langle|00\rangle, |10\rangle\rangle, \langle|01\rangle, |11\rangle\rangle\}$, with the same probability 0.5,
   - $|\psi\rangle$ collapses to $\frac{\sqrt{2}}{2}(|00\rangle + |10\rangle)$; or
   - $|\psi\rangle$ collapses to $\frac{\sqrt{2}}{2}(|01\rangle + |11\rangle)$.

Observable of 2QFA

For a 2QFA M and an input x,
we use an observable $\mathcal{O} := \{E_{\mathrm{acc}}, E_{\mathrm{rej}}, E_{\mathrm{non}}\}$, where

- $E_{\mathrm{acc}} := \langle C_{\mathrm{acc}} \rangle$, $C_{\mathrm{acc}} := Q_{\mathrm{acc}} \times \mathbb{Z}_n$
  (C_{acc} is the set of all accepting configurations);
- $E_{\mathrm{rej}} := \langle C_{\mathrm{rej}} \rangle$, $C_{\mathrm{rej}} := Q_{\mathrm{rej}} \times \mathbb{Z}_n$; and
- $E_{\mathrm{non}} := \langle C_{\mathrm{non}} \rangle$, $C_{\mathrm{non}} := Q_{\mathrm{non}} \times \mathbb{Z}_n$.

Observation



Expressiveness of 2QFAs

We will show that the 2QFA is more powerful than 2PFA by the following theorem.

Theorem (Kondacs97, Proposition 2)

For any error bound $\epsilon > 0$*, there is a 2QFA* M *which recognizes* $\{a^m b^m \mid m \geq 1\}$ *with the one-sided error bound* ϵ*, in linear time with respect to the length of the input.*

2QFA M_N for $\{a^m b^m \mid m \geq 1\}$

Definition

$$V_{¢}|q_0\rangle = |q_0\rangle,\quad V_{¢}|q_1\rangle = |q_3\rangle,\quad V_{¢}|r_{j,0}\rangle = \frac{1}{\sqrt{N}}\sum_{l=1}^{N}\exp\left(\frac{2\pi i}{N} j\, l\right)|s_l\rangle,\ 1 \leq j \leq N,$$

$$V_{\$}|q_0\rangle = |q_3\rangle,\quad V_{\$}|q_2\rangle = \frac{1}{\sqrt{N}}\sum_{j=1}^{N}|r_{j,0}\rangle,$$

$$V_a|q_0\rangle = |q_0\rangle,\quad V_a|q_1\rangle = |q_2\rangle,\quad V_a|q_2\rangle = |q_3\rangle,$$
$$V_a|r_{j,0}\rangle = |r_{j,j}\rangle,\ 1 \leq j \leq N,\quad V_a|r_{j,k}\rangle = |r_{j,k-1}\rangle,\ 1 \leq k \leq j,\ 1 \leq j \leq N,$$

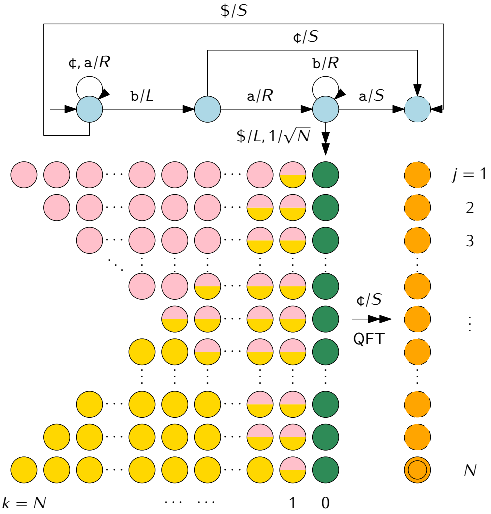
$$V_b|q_0\rangle = |q_1\rangle,\quad V_b|q_2\rangle = |q_2\rangle,$$
$$V_b|r_{j,0}\rangle = |r_{j,N-j+1}\rangle,\ 1 \leq j \leq N,\quad V_b|r_{j,k}\rangle = |r_{j,k-1}\rangle,\ 1 \leq k \leq N-j+1,\ 1 \leq j \leq N,$$

$$D(q_0) = +1,\quad D(q_1) = -1,\quad D(q_2) = +1,\quad D(q_3) = 0,$$
$$D(r_{j,0}) = -1,\ 1 \leq j \leq N,\quad D(r_{j,k}) = 0,\ 1 \leq j \leq N,\ k \neq 0,\quad D(s_j) = 0,\ 1 \leq j \leq N.$$

Figure 1: Specification of the transition function of M_N.

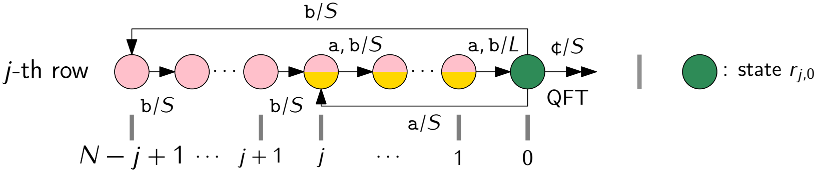
The 2QFA M_N

Overview 1



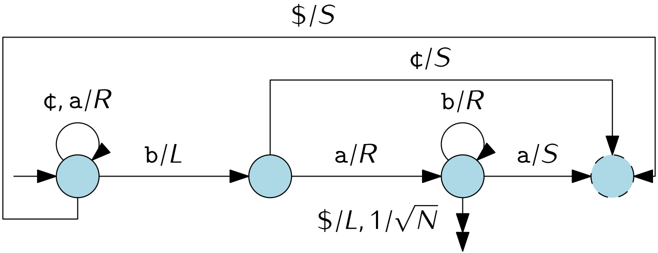
The 2QFA M_N

Overview 2

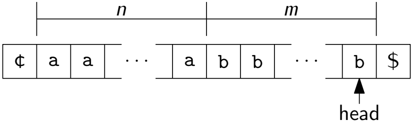


The 2QFA M_N

Phase 1



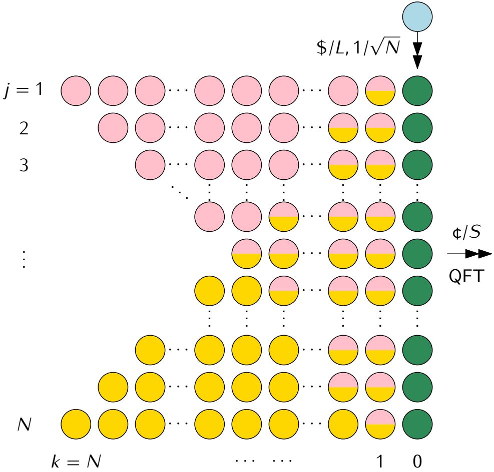
- First, if the input is not in $\{\mathtt{a}^n\mathtt{b}^m \mid n, m > 0\}$, then rejects.



- At the start of the next phase, the head is as above.

The 2QFA M_N

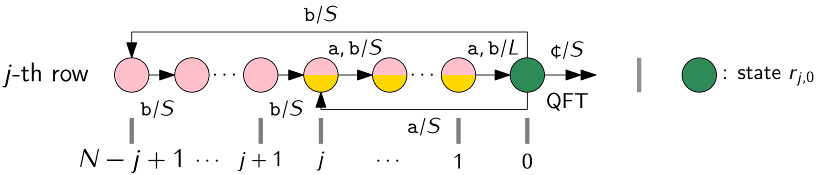
Phase 2 (1)

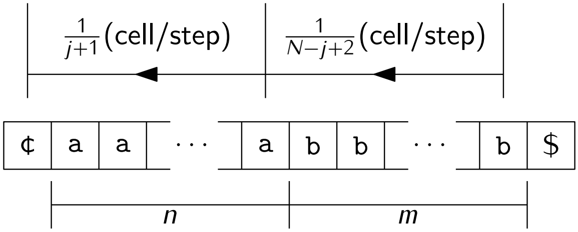


- The superposition consists of 0-column whose amplitude is $1/\sqrt{N}$ each at the beginning of phase 2.

The 2QFA M_N

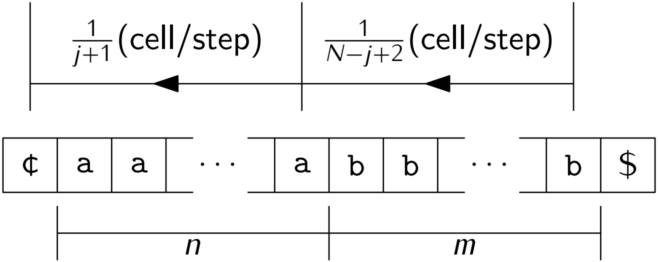
Phase 2 (2)





The 2QFA M_N

Phase 2 (3)



- For j-th row, reaching the beginning of the tape takes $(j+1) \cdot n + (N-j+2) \cdot m$ steps.
- At the same time, the amplitude of $|r_{j,0}, 0\rangle$ is $1/\sqrt{N}$.

The 2QFA M_N

Phase 2 (4)

- The amplitude of $|r_{j,0}, 0\rangle$ is $\frac{1}{\sqrt{N}}$, after $(j+1)n + (N-j+2)m$ steps.

Suppose that j-th and j'-th rows reach the beginning at the same time step. That is,

$$
\begin{aligned}
& (j+1)n + (N-j+2)m = (j'+1)n + (N-j'+2)m \\
\iff & (j-j')(m-n) = 0
\end{aligned}
$$

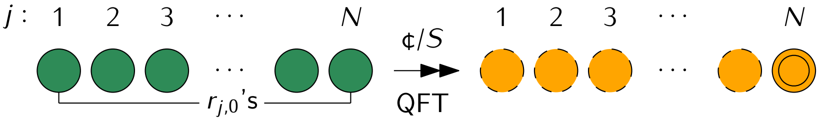
---

1. If $m = n$, the amp. of every $|r_{j,0}, 0\rangle$ is $\frac{1}{\sqrt{N}}$ at some time step t' in $O(N(m+n))$.
2. Otherwise, at most one amp. of $|r_{j,0}, 0\rangle$ is $\frac{1}{\sqrt{N}}$ at a time.

The 2QFA M_N

Phase 3 (1)

1. If $m = n$, the amp. of every $|r_{j,0}, 0\rangle$ is $\frac{1}{\sqrt{N}}$ at some t'.
2. Otherwise, at most one amp. of $|r_{j,0}, 0\rangle$ is $\frac{1}{\sqrt{N}}$ at a time.



Finally, for a string $w = \mathtt{a}^n\mathtt{b}^n$, M_N accepts w at $t' + 1$.

The 2QFA M_N

Phase 3 (2)

If $m \neq n$, consider each observation.
Let j' be the first row that reaches the beginning of the tape.

Note that the amplitude of $|j', 0\rangle$ is $\sqrt{\frac{1}{N}}$.
Then, at the next observation, one of the following happens.

We observe

1. "acc" and "rej" with probability $\frac{1}{N} \cdot \frac{1}{N}$ and $\frac{1}{N} \cdot \frac{N-1}{N}$, resp..
2. "non" with probability $1 - \frac{1}{N} = \frac{N-1}{N}$.
   (Then, amp. of each config.s $|c\rangle \in C_{\mathrm{acc}}$ are multiplied by $\sqrt{\frac{N}{N-1}}$)

For each i-th reaching, the above $\frac{1}{N}$'s are replaced by $\frac{1}{N-i+1}$'s.
As a result, M *wrongly* accepts w with probability $\frac{1}{N}$.

Some Languages of 2QFAs Are Not in Context-Free

Similarly, we can construct QFAs M_N for the non-context-free language $\{\mathtt{a}^n\mathtt{b}^n\mathtt{c}^n \mid n > 0\}$.

2-Way Reversible Finite Automata

Recall the following definitions of 2QFAs.

- $\delta : (Q \times \Gamma) \times (Q \times \{-1, 0, 1\}) \to \mathbb{C}$.
- $U_\delta^x |q, k\rangle := \sum \delta(q, x(k), q', d) \cdot |q', k + d\rangle$.
- U_δ^x is unitary.

Definition (2RFA)

A *2-way reversible finite state automata (2RFA)* is a 2QFA whose each transition amplitude is 0 or 1.

Note

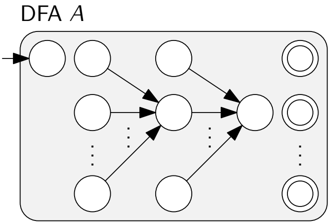
For each configuration of 2RFA,
the configuration has one unique "previous" configuration.

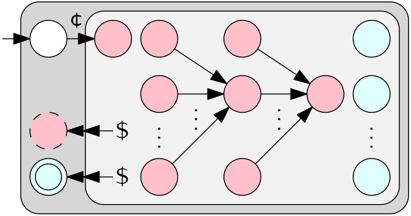
Reversible Simulation of 1DFAs

Theorem (Kondacs97, Proposition 4)

For any 1DFA A, there exists a 2RFA M which exactly recognize $L(A)$ in linear time.

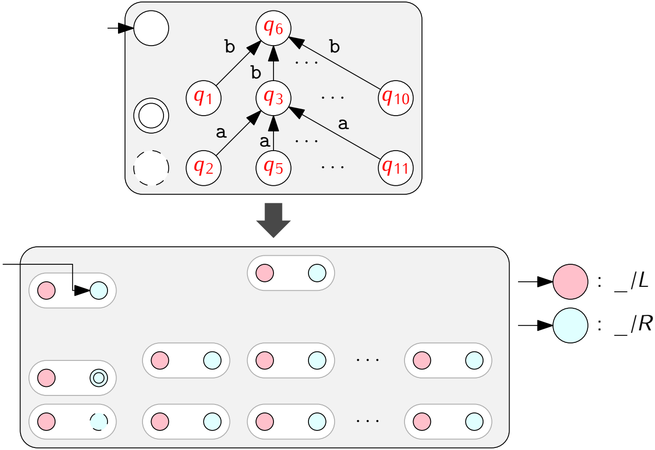
Construction of the 2RFA (1)





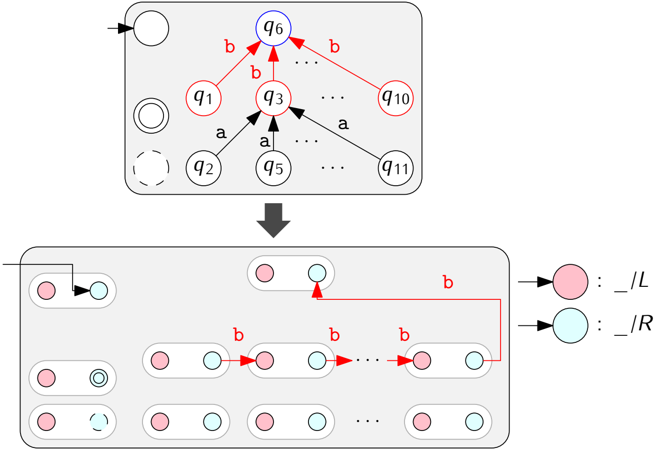
First, add some states for ¢ and $.

Construction of the 2RFA (2)



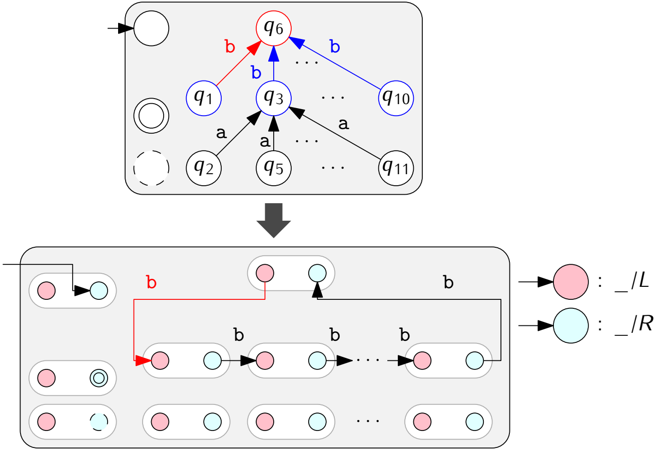
- Fix arbitrary ordering for states.
- Split each state into two states.

Construction of the 2RFA (3)



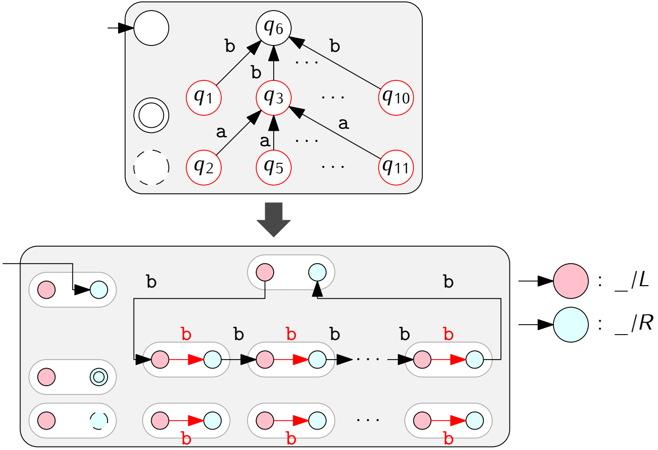
- Add b-transitions of states with the same target state.

Construction of the 2RFA (4)



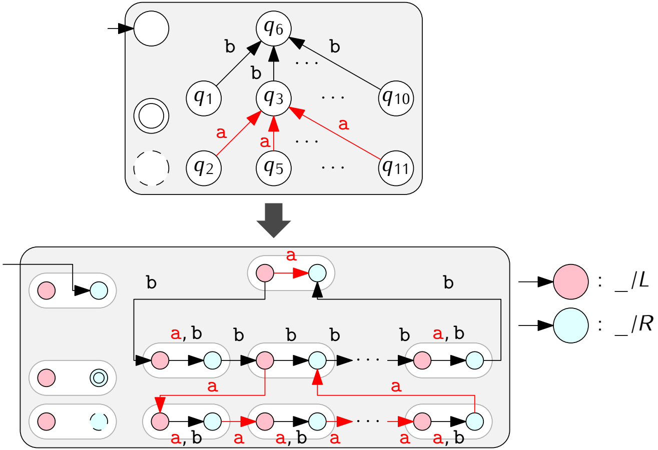
- Add "reverse" b-transitions of the target states.

Construction of the 2RFA (5)



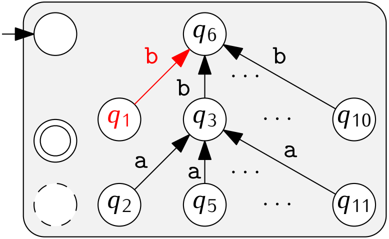
- Add b-transitions for states without b-labeled in-transition.

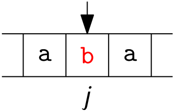
Construction of the 2RFA (6)



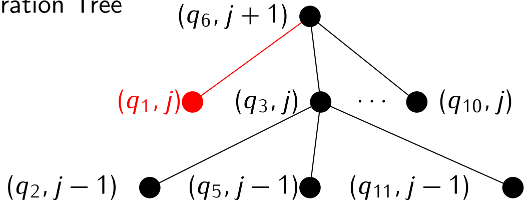
- Add a-transitions in a similar way.

Irreversibility of the DFA (1)



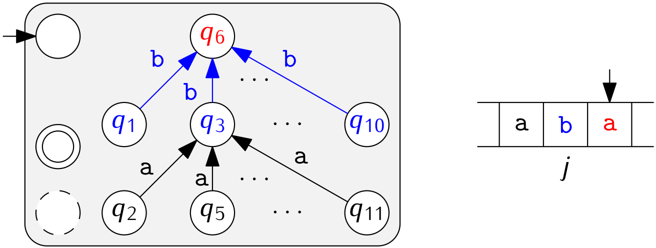


Configuration Tree

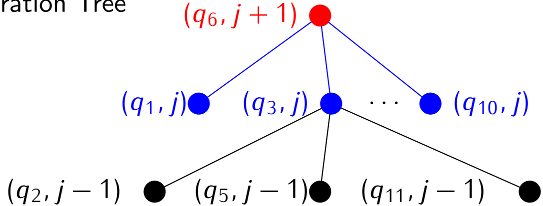


- Suppose we reached the configuration $(q_6, j+1)$ from (q_1, j).

Irreversibility of the DFA (2)

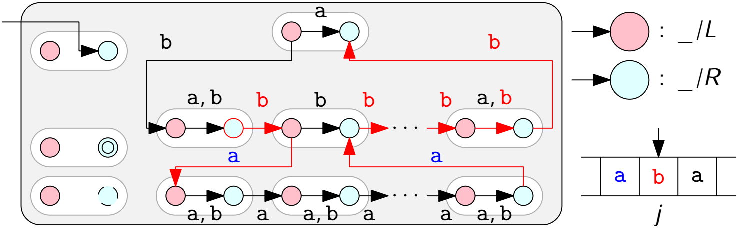


Configuration Tree

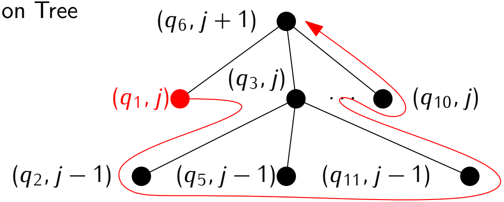


- We cannot determine a previous config. of the configuration.

Reversible Simulation using the 2RFA (1)

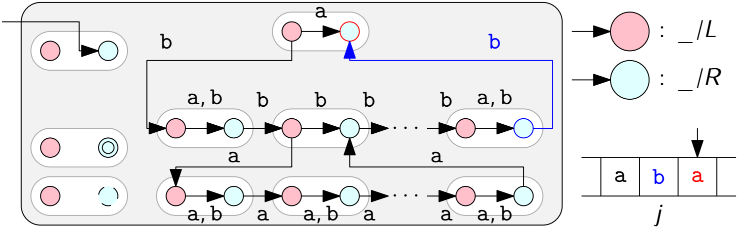


Configuration Tree

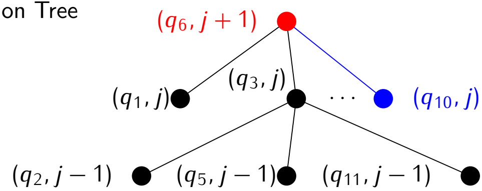


- In the 2RFA M, every valid configuration sequence reaching $(q_6, j+1)$ must passes (q_{10}, j).

Reversible Simulation using the 2RFA (2)



Configuration Tree



- A previous config. is *uniquely* determined for each config..
- The linear time comes from the num. of config.s.
  According to the pigeon hole principle, the number of steps more than configurations results in infinity loop.

1-Way Quantum Finite State Automata

Definition (1QFA)

A (measure-many) 1-way quantum finite state automata (1QFA) is a 2QFA $M = (Q, \Sigma, \delta, q_0, Q_{\mathrm{acc}}, Q_{\mathrm{rej}})$ that, for each $\sigma \in \Gamma$, there exists an unitary matrix $V_\sigma : Q \times Q \to \mathbb{C}$ satisfying

1. $\delta(q, \sigma, q', 1) = \langle q' | V_\sigma | q \rangle$, and
2. $\delta(q, \sigma, q', 0) = \delta(q, \sigma, q', -1) = 0$.

Total-States and Computation of 1QFA

A *total-state* of an 1QFA M is $(\psi, p_{\mathrm{acc}}, p_{\mathrm{rej}}) \in \mathcal{V} := \ell_2(Q) \times \mathbb{R} \times \mathbb{R}$. Intuitively,

1. ψ denotes *unnormalized* superposition $|\psi\rangle$,
2. p_{acc} is the (accumulated) accepting probability, and
3. p_{rej} is the rejecting probability.

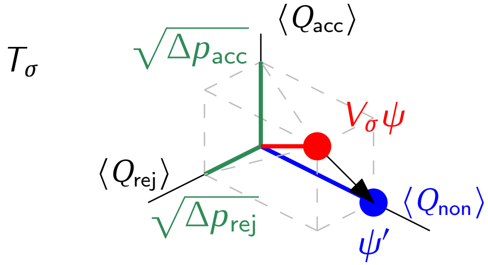
The intuition become clearer with the following operator T_σ.

$$T_\sigma : (\psi, p_{\mathrm{acc}}, p_{\mathrm{rej}}) \mapsto (P_{\mathrm{non}} V_\sigma \psi, \|P_{\mathrm{acc}} V_\sigma \psi\|^2 + p_{\mathrm{acc}}, \|P_{\mathrm{rej}} V_\sigma \psi\|^2 + p_{\mathrm{rej}}),$$

where P_{acc}, P_{rej} and P_{non} are the projection matrices onto $\langle Q_{\mathrm{acc}} \rangle$, $\langle Q_{\mathrm{rej}} \rangle$ and $\langle Q_{\mathrm{non}} \rangle$.

Total-States and Computation of 1QFA

Figure



$$T_\sigma : (\psi, p_{\mathrm{acc}}, p_{\mathrm{rej}}) \mapsto (P_{\mathrm{non}} V_\sigma \psi, \|P_{\mathrm{acc}} V_\sigma \psi\|^2 + p_{\mathrm{acc}}, \|P_{\mathrm{rej}} V_\sigma \psi\|^2 + p_{\mathrm{rej}}),$$

We define T_x for $x = \sigma_1 \cdots \sigma_n \in \sigma^*$ as $T_x = T_{\sigma_n} \cdots T_{\sigma_1}$.

Distance of Total-States

$$T_\sigma : \mathcal{V} : (\psi, p_{\text{acc}}, p_{\text{rej}})$$
$$\mapsto (P_{\text{non}} V_\sigma \psi, \|P_{\text{acc}} V_\sigma \psi\|^2 + p_{\text{acc}}, \|P_{\text{rej}} V_\sigma \psi\|^2 + p_{\text{rej}}),$$

For two total-states $v = (\psi, p_{\text{acc}}, p_{\text{rej}})$ and $v' = (\psi', p'_{\text{acc}}, p'_{\text{rej}})$, we define a *norm* of v as:

$$\|v\| := \frac{1}{2}(\|\psi\| + |p_{\text{acc}}| + |p_{\text{rej}}|).$$

Then a *distance* between total-states v and v' is

$$d(v, v') := \|v - v'\|.$$

Reachable Total-States

If $v = T_{¢w}|q_0, 0, 0\rangle$ for some $w \in \Sigma^*$, we call v is *reachable* by w.

Let $\mathcal{B} := \{v \in \mathcal{V} \mid \|v\| \leq 1\}$.
Clearly, any valid total-state v must be in $\mathcal{B}$.

Note that

1. T_x increases the distance at most linearly:
   $d(T_\sigma v, T_\sigma v') \leq c \cdot d(v, v')$,
2. $A \subset \mathcal{B}$ and $\exists \epsilon > 0, \forall v, v' \in A, d(v, v') > \epsilon$ implies A is finite.

Construction of the DFA

1. $d(T_x v, T_x v') \leq c \cdot d(v, v')$, and
2. $A \subset \mathcal{B}$ and $\exists \epsilon > 0, \forall v, v' \in A, d(v, v') > \epsilon$ implies A is finite.

Fix an (two-sided) error bound $\epsilon > 0$. Total-states v and v' are *distinguishable* if there exists $y \in \Sigma^*$ such that

1. accepting probability of $T_{y\$} v$ is greater than $\frac{1}{2} + \epsilon$, and
2. accepting probability of $T_{y\$} v'$ is less than $\frac{1}{2} - \epsilon$,

or vice versa.

1. Note that $2\epsilon < d(T_{y\$} v, T_{y\$} v')$ by the definition.
2. From the bound of $T_{y\$}$ (in 1. above), $\frac{2\epsilon}{c} < d(v, v')$.
3. Thus, by the finiteness (in 2. above)
   a set of distinguishable-and-reachable total-states is finite.

Limitation of 1QFA

1QFA cannot recognize $L = \{a, b\}^* a$ with a bounded error.

Let $\psi_x := (P_{\text{non}} V_{\sigma_n})(P_{\text{non}} V_{\sigma_{n-1}}) \cdots (P_{\text{non}} V_{\sigma_1})|q_0\rangle$.

1. Let $\mu = \inf\{\|\psi_{¢w}\| \mid w \in \{a, b\}^*\}$. If $\mu = 0$, M cannot recognize L with a bounded error. So we may assume $\mu > 0$.
2. Let $\xi > 0$, and choose w such that $\|\psi_{¢w}\| < \mu + \xi$.
3. Then, $\mu \leq \|\psi_{¢wy}\| < \mu + \xi$ for every $y \in \{a, b\}^*$.
4. Specifically, for any $j \geq 0$,

$$\mu \leq \psi_{¢wab^j} = \|(P_{\text{non}} V_b)^j \psi_{¢wa}\| < \mu + \xi.$$

5. Fix j, k satisfying $\|\psi_{¢wab^j} - \psi_{¢wab^{k+j}}\| < \xi$.
6. Then,

$$\begin{aligned} &\|\psi_{¢wa} - \psi_{¢wab^k}\| < c' \cdot \xi^{1/4} \\ &\implies d(T_{¢wa\$}(|q_0\rangle, 0, 0), T_{¢wab^k\$}(|q_0\rangle, 0, 0)) < c'' \cdot \xi^{1/4}. \end{aligned}$$

Thank you for your attention!

Any Questions?