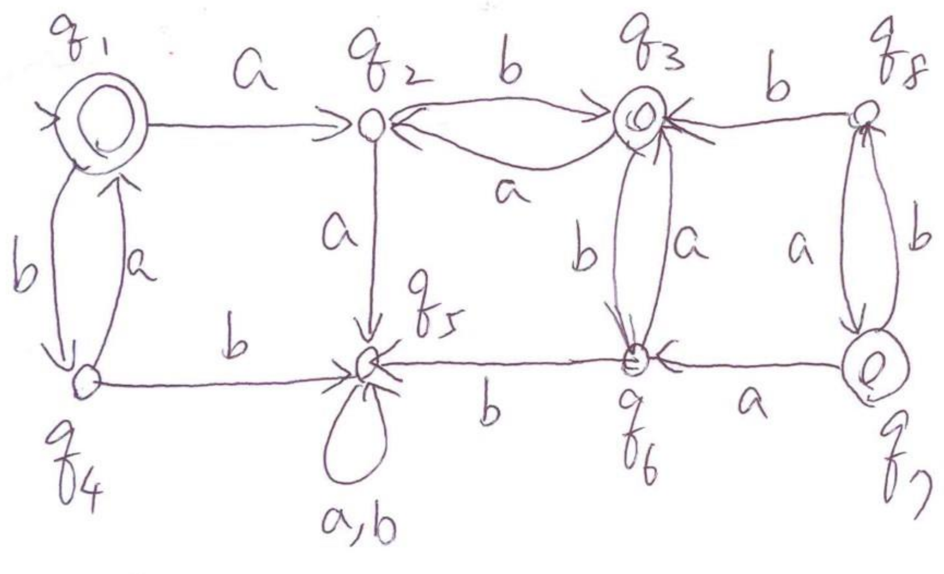
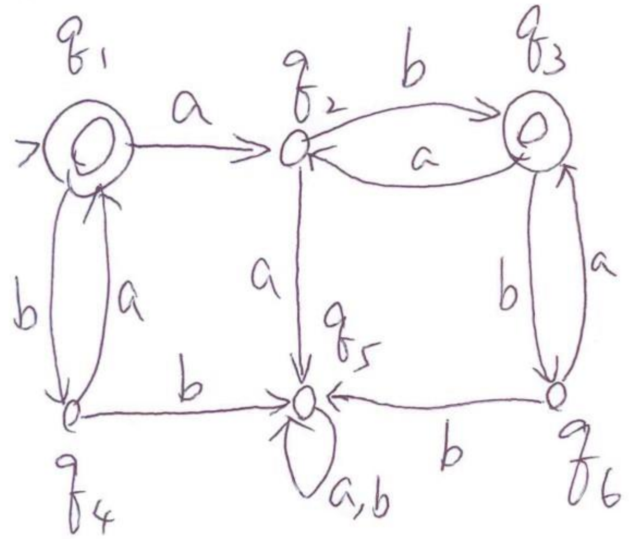
✡ State Minimization

Kun-Mao Chao



q_7 and q_8 : unreachable.



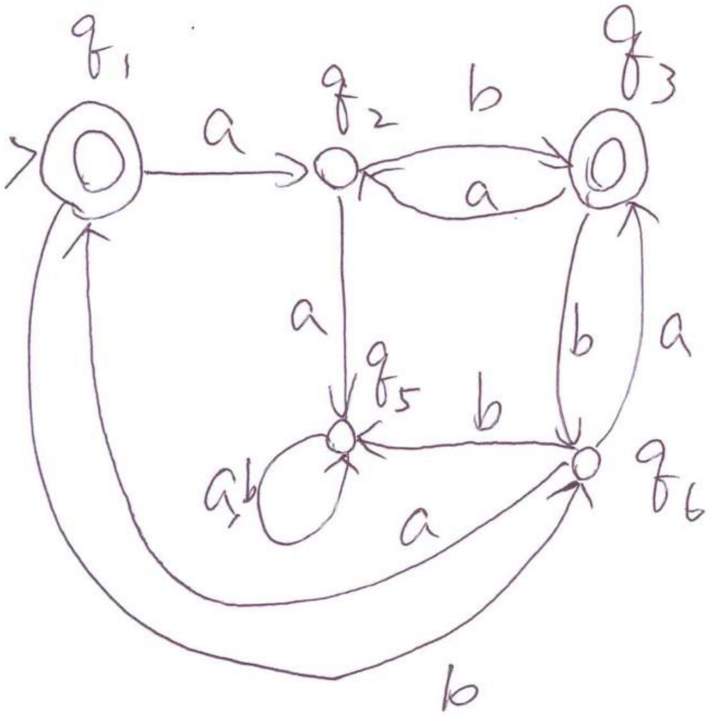
$(ab \cup ba)^*$

$q_4 \xrightarrow{a(ba \cup ab)^*} f \in F$

$q_6 \xrightarrow{a(ba \cup ab)^*} f' \in F$

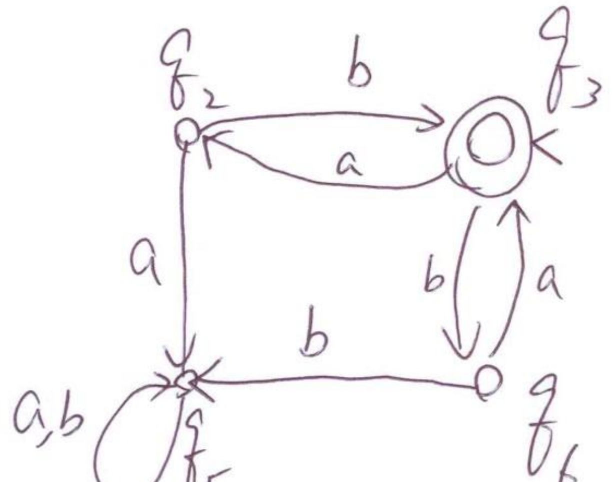
equivalent.

Kun-Mao Chao



nondeterministic.

If $(q_1, x) \vdash_M^* (f, e)$, where $f \in F$, then
$(q_3, x) \vdash_M^* (f', e)$, where $f' \in F$.



Def. Let $L \subseteq \Sigma^*$ be a language, and let $x, y \in \Sigma^*$. We say $x \approx_L y$ if for all $z \in \Sigma^*$, $xz \in L$ iff $yz \in L$. ($\approx_L$ is an equivalence relation.)

Kun-Mao Chao

$[x]$: the equivalence class with respect to L to which x belongs.

$$L = (ab \cup ba)^*$$

Four equivalence classes:

$[e] = L$ $\{e, ab, ba, abab, abba, \dots\}$

$[a] = La$ $\{a, aba, baa, ababa, abbaa, \dots\}$

$[b] = Lb$ $\{b, abb, bab, ababb, abbab, \dots\}$

$[aa] = L(aa \cup bb)\Sigma^*$ $\{aa, bb, abaa, abbb, \dots\}$

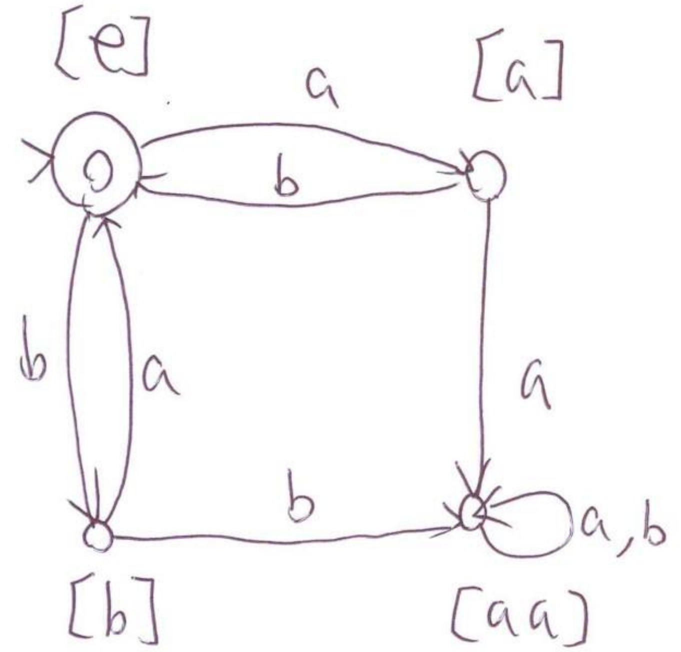
For any string $x \in [e]$, we have $xa \in [a]$ and $xb \in [b]$.

For any string $x \in [a]$, we have $xb \in [e]$ and $xa \in [aa]$.

For any string $x \in [b]$, we have $xa \in [e]$ and $xb \in [aa]$.

For any string $x \in [aa]$, we have $xa \in [aa]$ and $xb [aa]$.

Kun-Mao Chao



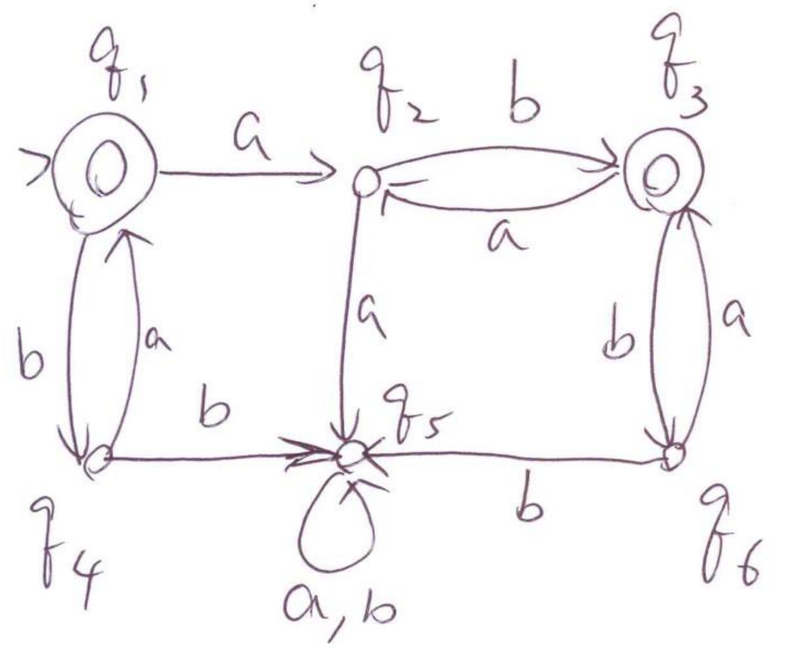
Thm. Let L be a regular language. There is a DFA with as many states as there are equivalence classes in $\approx_L$ that accepts L.

Def. Let M be a DFA. Two strings $x, y \in \Sigma^*$ are equivalent with respect to M, denoted $x \sim_M y$, if they both drive M from the initial state to the same state. That is, $x \sim_M y$ if $\exists$ a state q such that $(s, x) \vdash_M^* (q, e)$ and

$$(s, y) \vdash_M^* (q, e).$$

$L = (ab \cup ba)^*$

Kun-Mao Chao



$E_{q_1} = (ba)^*$ $\subseteq [e]$

$E_{q_2} = (ba)^* a (bLa \cup \phi^*)$ $\subseteq [a]$

$E_{q_3} = (ba)^* ab L$ $\subseteq [e]$

$E_{q_4} = b(ab)^*$ $\subseteq [b]$

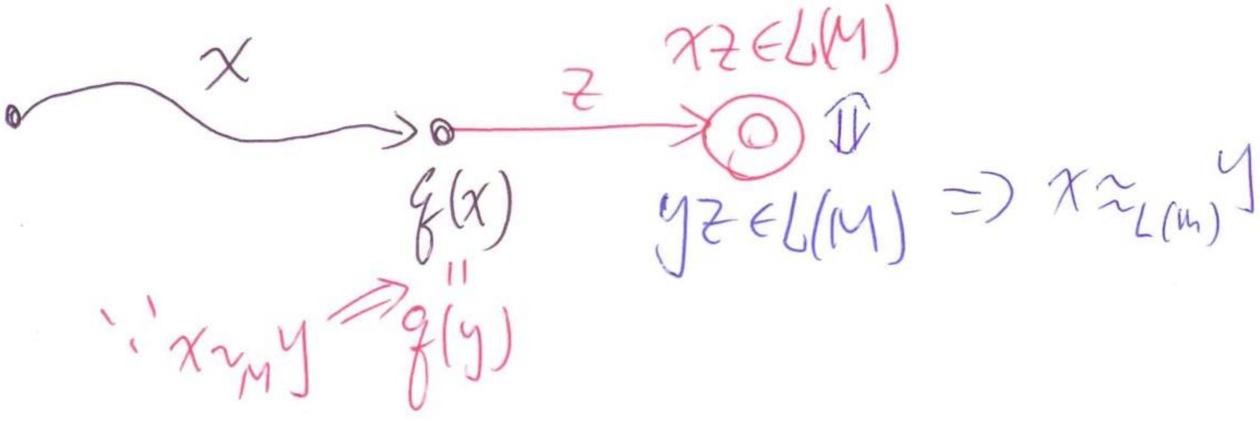
$E_{q_5} = L(aa \cup bb)\Sigma^*$ $\subseteq [aa]$

$E_{q_6} = (ba)^* ab L b$ $\subseteq [b]$

Kun-Mao Chao

Thm.

$$x \sim_M y \Rightarrow x \approx_{L(M)} y$$



The number of states of a DFA accepting L is no less than the number of equivalence classes under $\approx_L$.

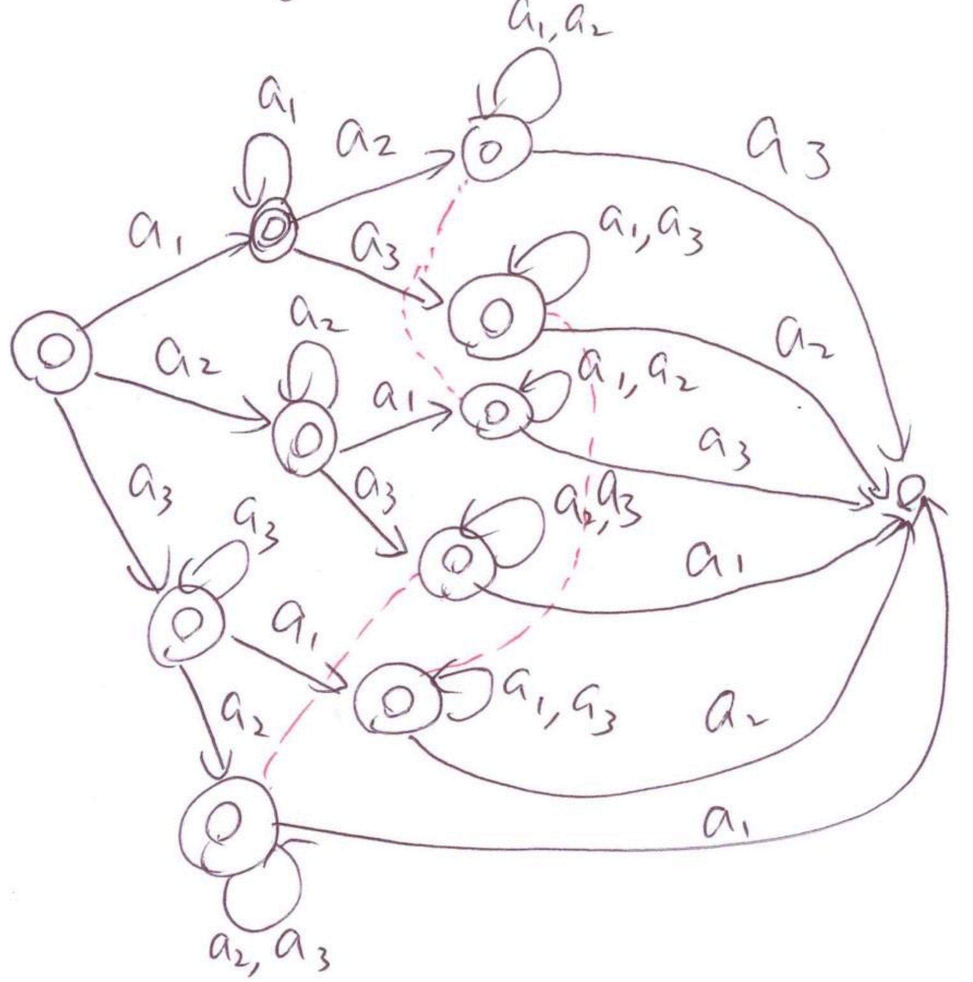
Corollary: A language L is regular iff $\approx_L$ has finitely many equivalence classes.

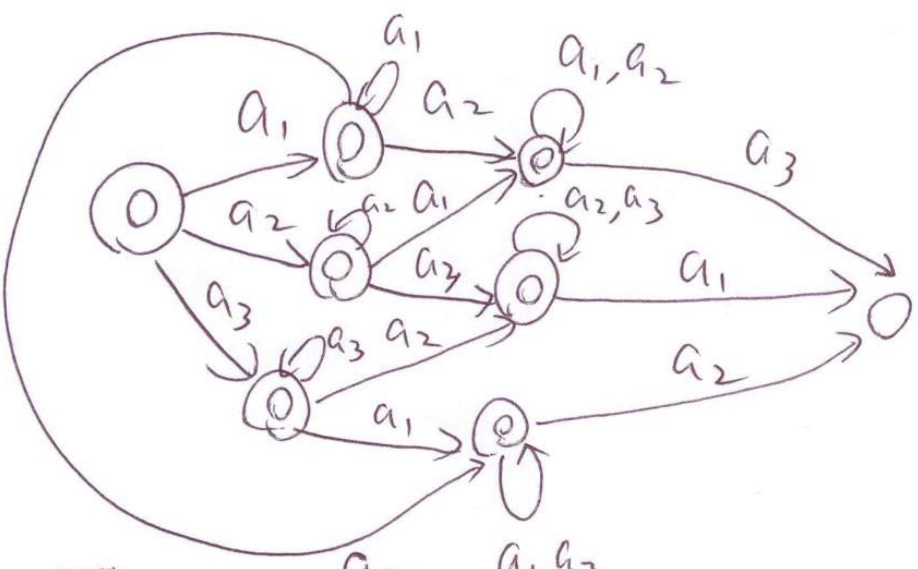
* $L = \{a^i b^i : i \geq 0\}$ is not regular because $\approx_L$ has infinitely many equivalence classes

$[e], [a], [aa], [aaa], [aaaa], \ldots$

Kun-Mao Chao

Ex. $L = \{w \in \{a_1, a_2, a_3\}^* : w$ does not contain occurrences of all three symbols$\}$





$[e] = \{e\}$

$[a_1] = \{a_1, a_1a_1, \dots\}$

$[a_2] = \{a_2, a_2a_2, \dots\}$

$[a_3] = \{a_3, a_3a_3, \dots\}$

$[a_1a_2] = \{a_1a_2, a_1a_1a_2, \dots\}$

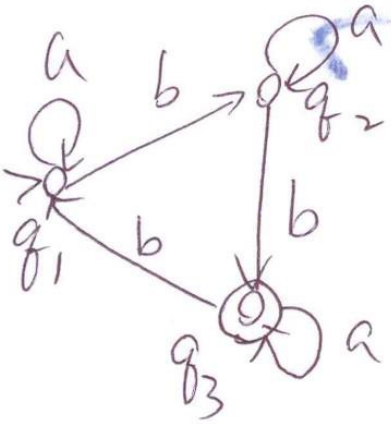
$[a_1a_3] = \{a_1a_3, \dots\}$

$[a_2a_3] = \{a_2a_3, a_2a_2a_3, \dots\}$

$[a_1a_2a_3] = \{a_1a_2a_3, \dots\}$

These states are all necessary because $\approx_L$ has 8 equivalence classes.

Kun-Mao Chao



q_1: $\sigma \leftarrow$ get-next-symbol;
if σ is EOF then Reject;
else if $\sigma = a$ then goto q_1;
else if $\sigma = b$ then goto q_2;

q_2: $\sigma \leftarrow$ get-next-symbol;
if σ is EOF then Reject;
else if $\sigma = a$ then goto q_2;
else if $\sigma = b$ then goto q_3;

q_3: $\sigma \leftarrow$ get-next symbol;
if σ is EOF then Accept;
else if $\sigma = a$ then goto q_3;
else if $\sigma = b$ then goto q_1;