minimization of dfa’s

1 ’Historical Background’

I found 2 pages on this topic in my old lecture notes.

- great algorithm, great example, too good to leave out!
- arguments missing
- I have never seen this
- somebody else must have given my lecture
- using material I did not know

1 'Historical Background'

I found 2 pages on this topic in my old lecture notes.

- great algorithm, great example, too good to leave out!
- arguments missing
- I have never seen this
- somebody else must have given my lecture
- using material I did not know

the good news: I found the material - with the same example and all arguments at as skript from 2005 of Jan Peleska

http://www.informatik.uni-bremen.de/agbs/lehre/ss05/pi2/hintergrund/minimize_dfa.pdf

notation in that document

$$B = (Q, \Sigma, \delta, q_0, F)$$

uploaded to class materials

1 ’Historical Background’

I found 2 pages on this topic in my old lecture notes.

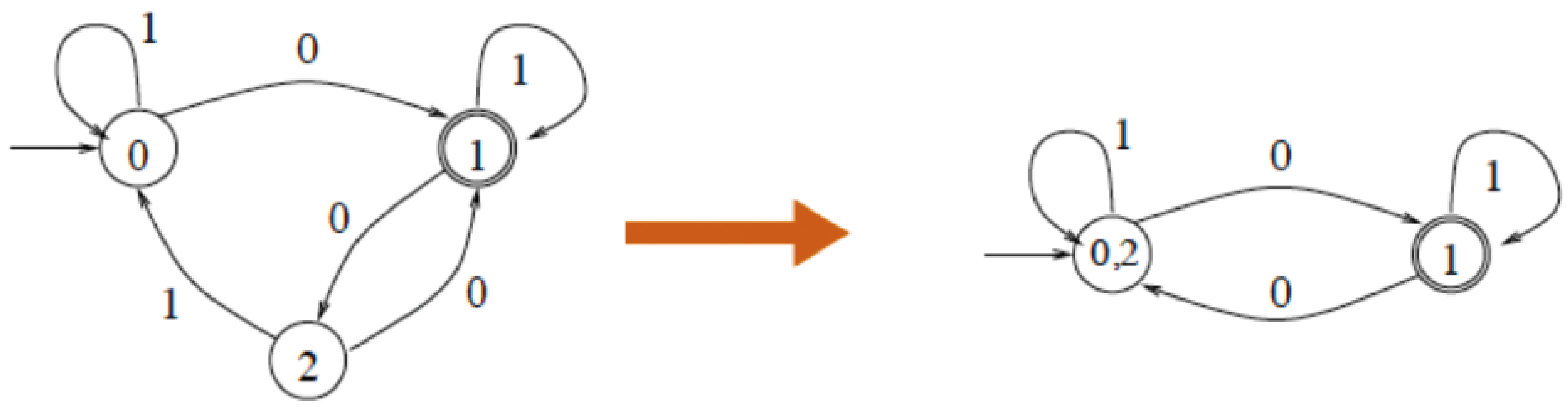
- great algorithm, great example, too good to leave out!
- arguments missing
- I have never seen this
- somebody else must have given my lecture
- using material I did not know

2 The problem

Given dfa M: find an automaton M' with

$$L(M) = L(M')$$

and a minimal number of states.



the good news: I found the material - with the same example and all arguments at as skript from 2005 of Jan Peleska

http://www.informatik.uni-bremen.de/agbs/lehre/ss05/pi2/hintergrund/minimize_dfa.pdf

notation in that document

$$B = (Q, \Sigma, \delta, q_0, F)$$

uploaded to class materials

3 Preliminaries

reachable states: a state $p \in Q$ of a fa is *reachable* if there is a path from the start state q_0 to p.

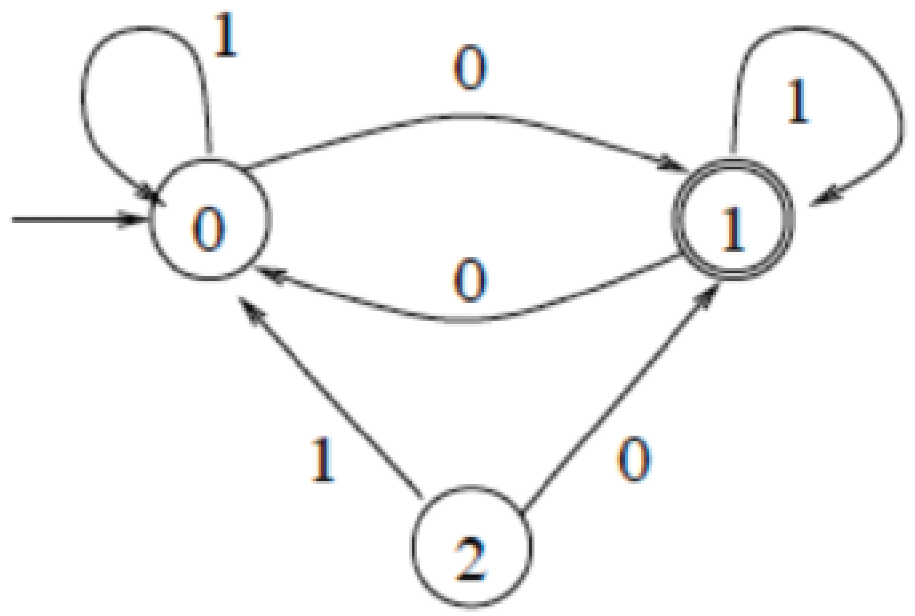
- find reachable states by a graph search (DFS or BFS)

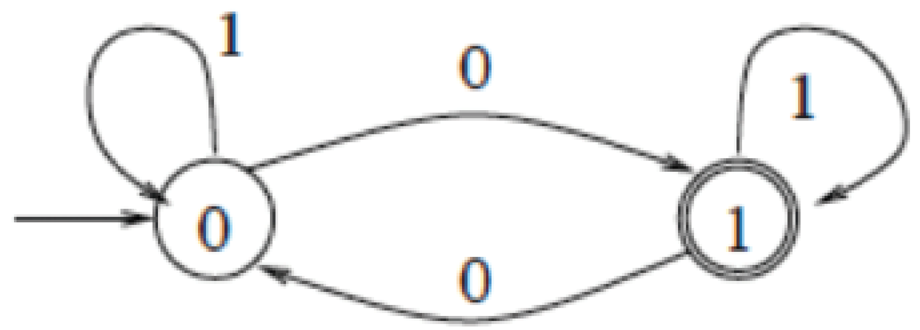
- eliminate non reachable states

3 Preliminaries

reachable states: a state $p \in Q$ of a fa is *reachable* if there is a path from the start state q_0 to p.

- find reachable states by a graph search (DFS or BFS)
- eliminate non reachable states





iterated transition function: transition function

$$\delta : Q \times \Sigma \to Q$$

gives iterated transition function

$$\delta^* : Q \times A^* \to Q$$

For $p \in Q, a \in \Sigma, w \in \Sigma^*$ define

$$\begin{aligned} \delta^*(p,\varepsilon) &= p \\ \delta^*(p,aw) &= \delta^*(\delta(p,a),w) \end{aligned}$$

making steps until w is consumed

4 Distiguishable states

States $p,q \in Q$ are *distinguishable* if there is $w \in \Sigma^*$ such that

$$\delta^*(p,w) \in F \quad , \quad \delta^*(q,w) \notin F$$

or

$$\delta^*(q,w) \in F \quad , \quad \delta^*(p,w) \notin F$$

We say: w *distinguishes* between p and q.

4 Distiguishable states

States $p,q \in Q$ are *distinguishable* if there is $w \in \Sigma^*$ such that

$$\delta^*(p,w) \in F \quad , \quad \delta^*(q,w) \notin F$$

or

$$\delta^*(q,w) \in F \quad , \quad \delta^*(p,w) \notin F$$

We say: w *distinguishes* between p and q.

Lemma 1. *Let B be a dfa in which all states are reachable. Then B is minimal iff all states are distinguishable.*

Proof: ($\Rightarrow$) This implication is easy, for if B has two undistinguishable states, one of them can be eliminated, and the transitions into this state can be changed to go to the other. This will not affect the accepted language. (Formal proof omitted.)

4 Distiguishable states

States $p, q \in Q$ are *distinguishable* if there is $w \in \Sigma^*$ such that

$$\delta^*(p,w) \in F \quad , \quad \delta^*(q,w) \notin F$$

or

$$\delta^*(q,w) \in F \quad , \quad \delta^*(p,w) \notin F$$

We say: w *distinguishes* between p and q.

Lemma 1. *Let B be a dfa in which all states are reachable. Then B is minimal iff all states are distinguishable.*

- $\rightarrow$: trival. Merge non distiguishable states.
- $\leftarrow$: Let $|Q| = k$ and let B' be automaton with $|Q'| = \ell < k$. Claim:

$$L(B) \neq L(B')$$

For each $q \in Q$ let

$$\delta^*(q_0, w_q) = q$$

exists because q reachable.

pidgeon hole argument: $\ell < k \rightarrow$:

$$\exists p, q. \quad \delta'^*(q_0', w_p) = \delta'^*(q_0', w_q) = r$$

4 Distiguishable states

States $p, q \in Q$ are *distinguishable* if there is $w \in \Sigma^*$ such that

$$\delta^*(p,w) \in F \quad , \quad \delta^*(q,w) \notin F$$

or

$$\delta^*(q,w) \in F \quad , \quad \delta^*(p,w) \notin F$$

We say: w *distinguishes* between p and q.

Lemma 1. *Let B be a dfa in which all states are reachable. Then B is minimal iff all states are distinguishable.*

- $\rightarrow$: trival. Merge non distiguishable states.
- $\leftarrow$: Let $|Q| = k$ and let B' be automaton with $|Q'| = \ell < k$. Claim:

  $$L(B) \neq L(B')$$

  For each $q \in Q$ let

  $$\delta^*(q_0, w_q) = q$$

  exists because q reachable.

  pidgeon hole argument: $\ell < k \rightarrow$:

  $$\exists p, q. \quad \delta'^*(q_0', w_p) = \delta'^*(q_0', w_q) = r$$

p, q in B distinguished by v. w.l.o.g:

$$\delta^*(p,v) \in F \quad , \quad \delta^*(q,v) \notin F$$

$$w_p v \in L(B) \quad , \quad w_q v \notin L(B)$$

4 Distiguishable states

States $p, q \in Q$ are *distinguishable* if there is $w \in \Sigma^*$ such that

$$\delta^*(p,w) \in F \quad , \quad \delta^*(q,w) \notin F$$

or

$$\delta^*(q,w) \in F \quad , \quad \delta^*(p,w) \notin F$$

We say: w *distinguishes* between p and q.

Lemma 1. *Let B be a dfa in which all states are reachable. Then B is minimal iff all states are distinguishable.*

- $\rightarrow$: trival. Merge non distiguishable states.
- $\leftarrow$: Let $|Q| = k$ and let B' be automaton with $|Q'| = \ell < k$. Claim:

$$L(B) \neq L(B')$$

For each $q \in Q$ let

$$\delta^*(q_0, w_q) = q$$

exists because q reachable.

pidgeon hole argument: $\ell < k \rightarrow$:

$$\exists p, q. \quad \delta'^*(q_0', w_p) = \delta'^*(q_0', w_q) = r$$

p, q in B distinguished by v. w.l.o.g:

$$\delta^*(p,v) \in F \quad , \quad \delta^*(q,v) \notin F$$

$$w_p v \in L(B) \quad , \quad w_q v \notin L(B)$$

in B':

$$w_p v, w_q v \in L(B') \leftrightarrow \delta'^*(r,v) \in F'$$

4 Distiguishable states

States $p, q \in Q$ are *distinguishable* if there is $w \in \Sigma^*$ such that

$$\delta^*(p,w) \in F \quad , \quad \delta^*(q,w) \notin F$$

or

$$\delta^*(q,w) \in F \quad , \quad \delta^*(p,w) \notin F$$

We say: w *distinguishes* between p and q.

Lemma 1. *Let B be a dfa in which all states are reachable. Then B is minimal iff all states are distinguishable.*

- $\rightarrow$: trival. Merge non distiguishable states.

Lemma 2. *indistinguishability is an equivalence relation*

Proof. exercise

4 Distiguishable states

States $p, q \in Q$ are *distinguishable* if there is $w \in \Sigma^*$ such that

$$\delta^*(p,w) \in F \quad , \quad \delta^*(q,w) \notin F$$

or

$$\delta^*(q,w) \in F \quad , \quad \delta^*(p,w) \notin F$$

We say: w *distinguishes* between p and q.

Lemma 1. *Let B be a dfa in which all states are reachable. Then B is minimal iff all states are distinguishable.*

- $\rightarrow$: trival. Merge non distiguishable states.

Lemma 2. *indistinguishability is an equivalence relation*

Proof. exercise

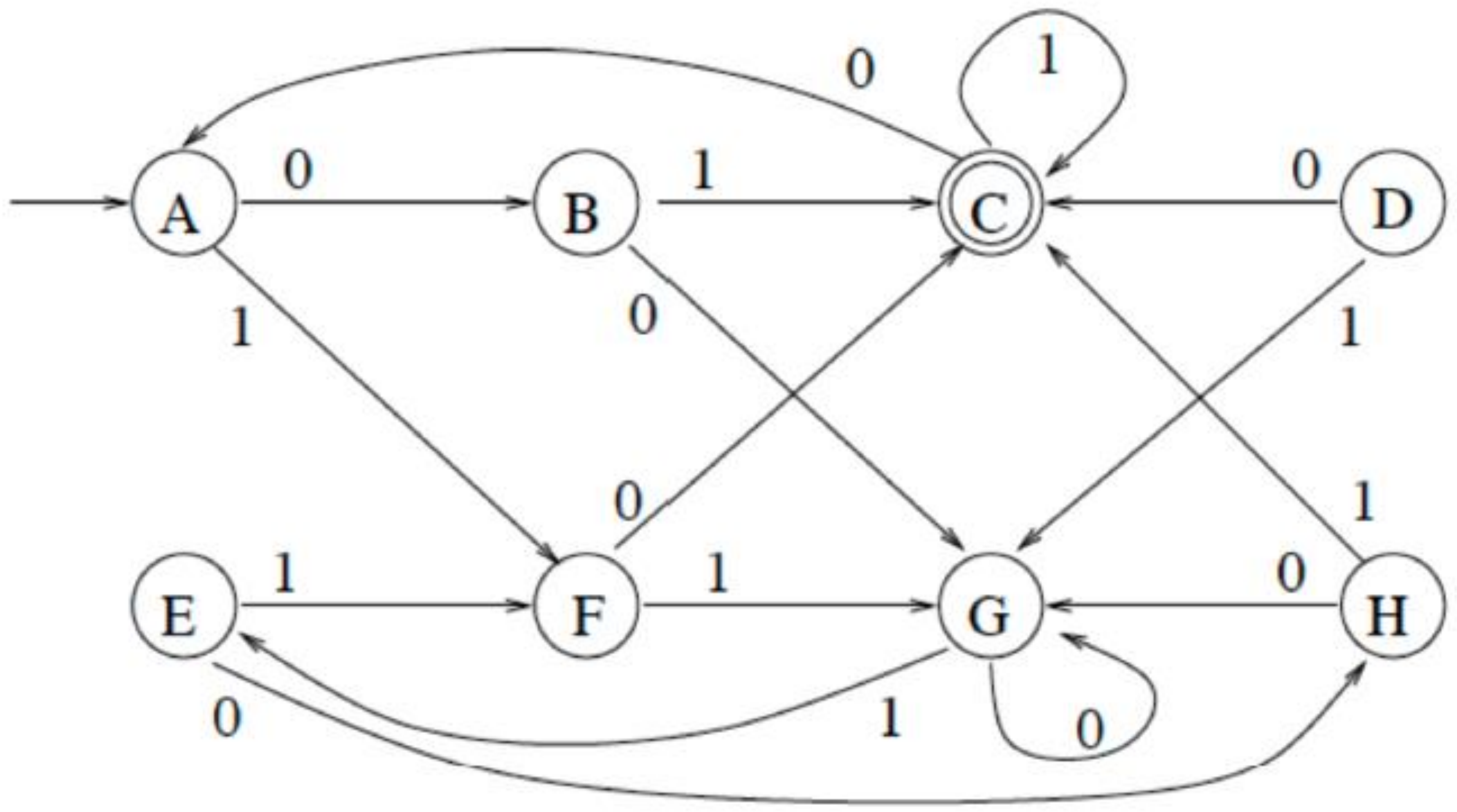
Lemma 3. *Let $\delta(p,a) = p'$ and $\delta(q,a) = q'$. Then:*

$$p', q' \textit{ distinguishable } \rightarrow p, q \textit{ distinguishable}$$

Proof. If w distinguishes p', q', then aw distinguishes p, q.

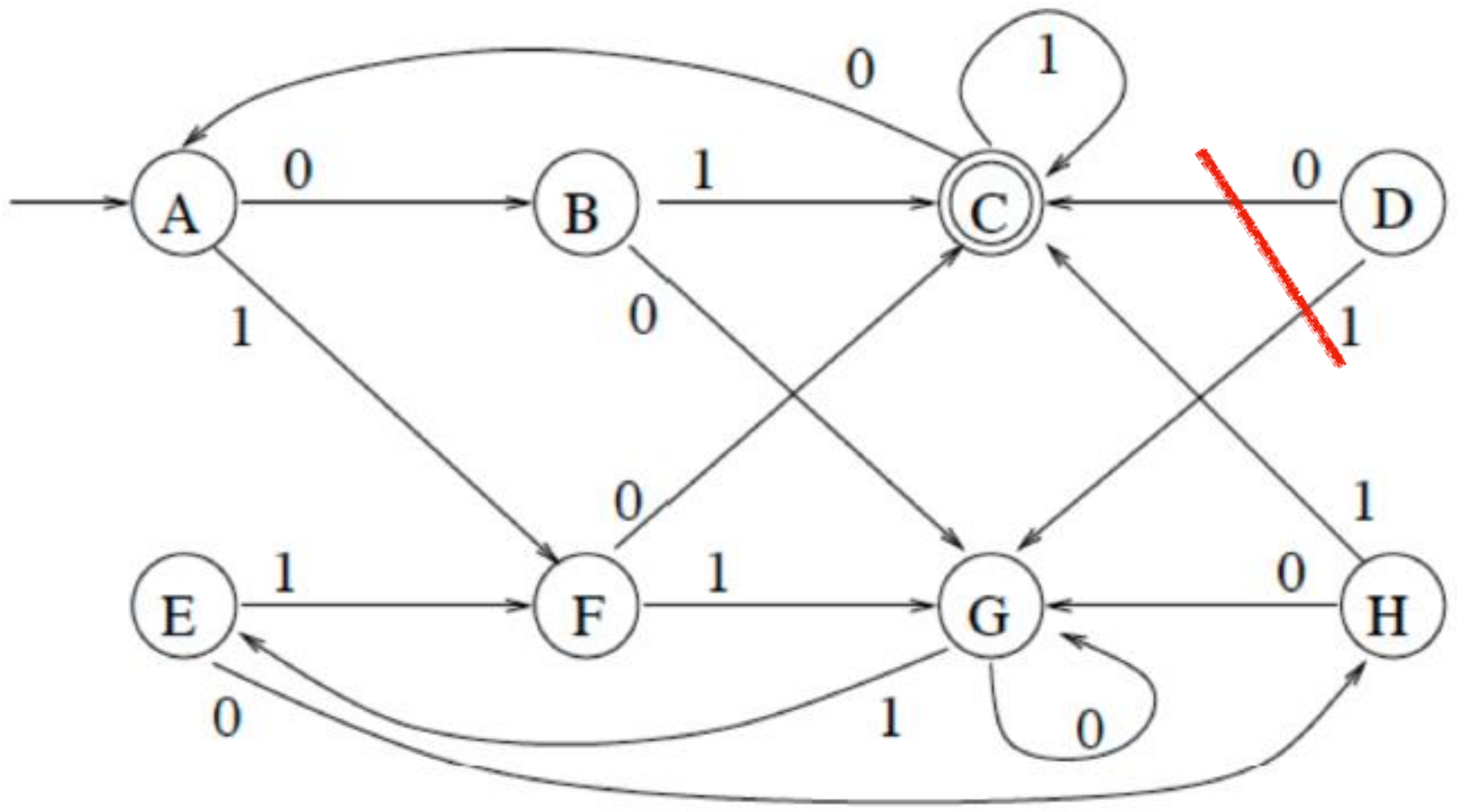
5 Minimization algorithm

running example :



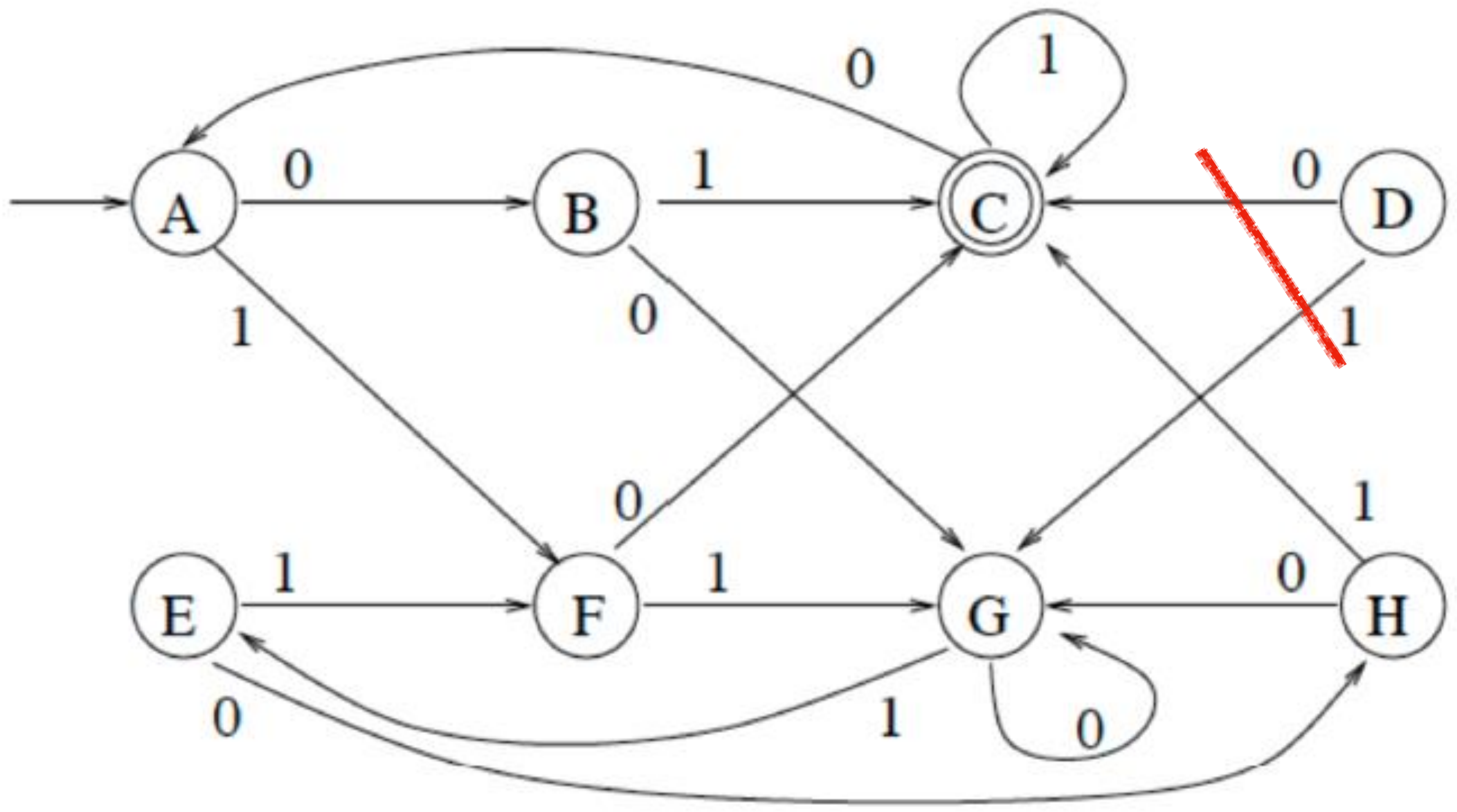
5 Minimization algorithm

running example :



step 1: eliminate unreachable states

running example :



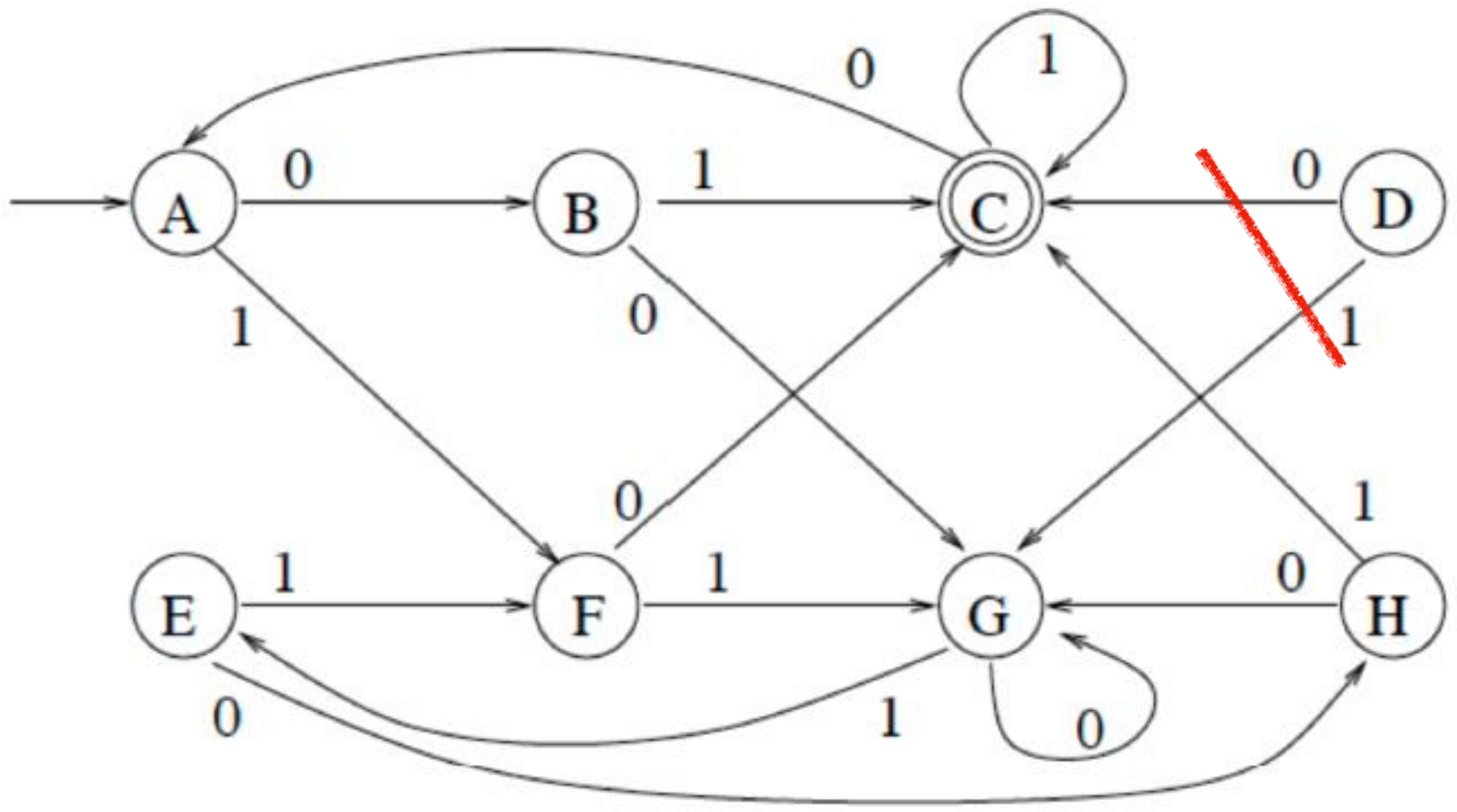
step 1: eliminate unreachable states

step 2: Mark the distinguishable pairs of states. To achieve this task, we first mark all pairs p, q, where $p \in F$ and $q \notin F$ as distinguishable.

1. initiate:

B						
C	X	X				
E			X			
F			X			
G			X			
H			X			
	A	B	C	E	F	G

running example :



step 1: eliminate unreachable states

step 2: Mark the distinguishable pairs of states. To achieve this task, we first mark all pairs p, q, where $p \in F$ and $q \notin F$ as distinguishable.

Then, we proceed as follows:

repeat
for all non-marked pairs p, q do
for each letter a do
if the pair $\delta(p, a), \delta(q, a)$ is marked then mark p, q
until no new pairs are marked

1. initiate:

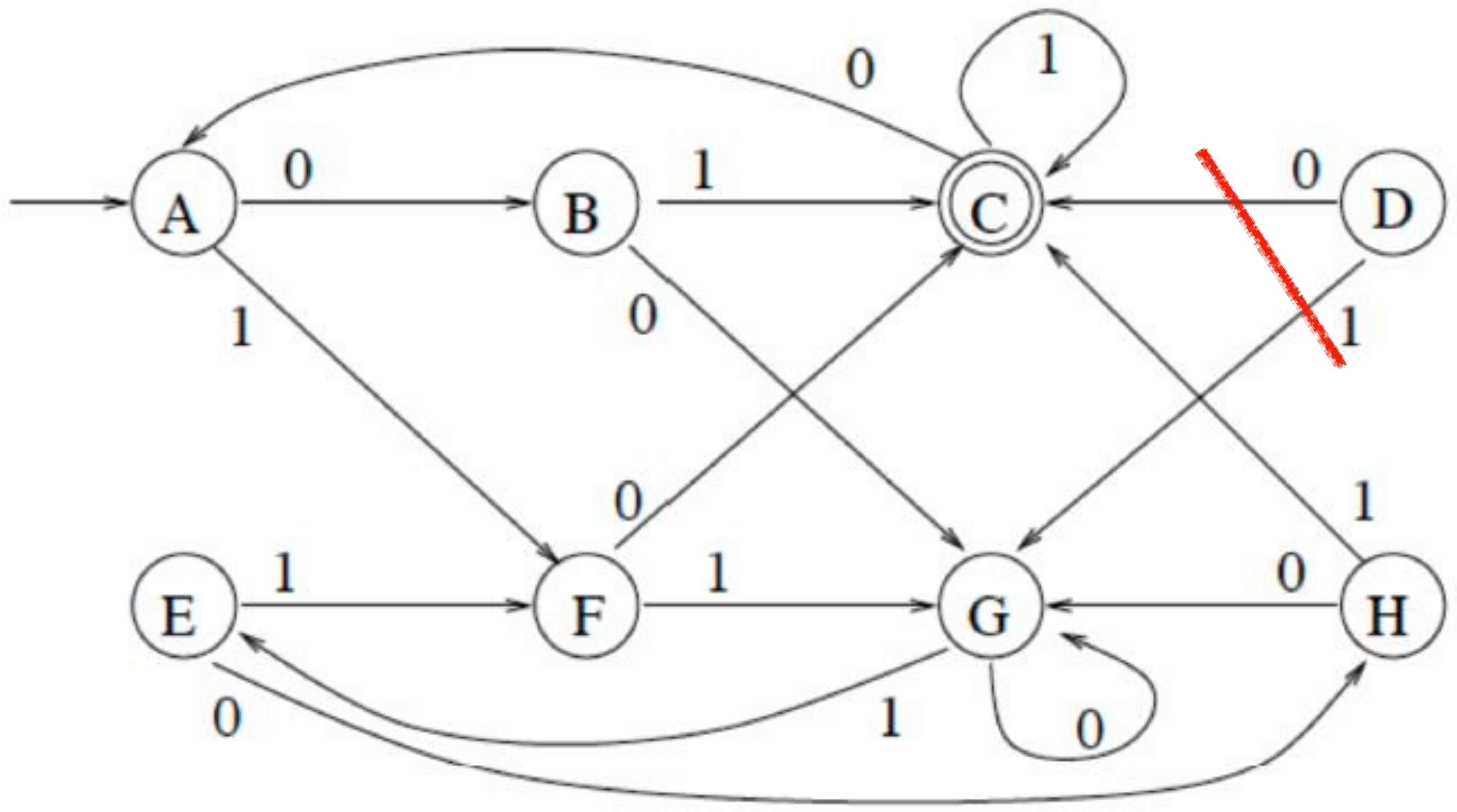
B						
C	X	X				
E			X			
F			X			
G			X			
H			X			
	A	B	C	E	F	G

2. e.g. $\delta(A,1) = F, \delta(B,1) = C$, mark A, B

B	X					
C	X	X				
E		X	X			
F	X	X	X	X		
G		X	X		X	
H	X		X	X	X	X
	A	B	C	E	F	G

Figure 5: after 1 iteration

running example :



step 1: eliminate unreachable states

step 2: Mark the distinguishable pairs of states. To achieve this task, we first mark all pairs p,q, where $p \in F$ and $q \notin F$ as distinguishable.

Then, we proceed as follows:

repeat
for all non-marked pairs p,q do
for each letter a do
if the pair $\delta(p,a), \delta(q,a)$ is marked then mark p,q
until no new pairs are marked

2. e.g. $\delta(A,1) = F, \delta(B,1) = C$, mark A,B

B	X					
C	X	X				
E		X	X			
F	X	X	X	X		
G		X	X		X	
H	X		X	X	X	X
	A	B	C	E	F	G

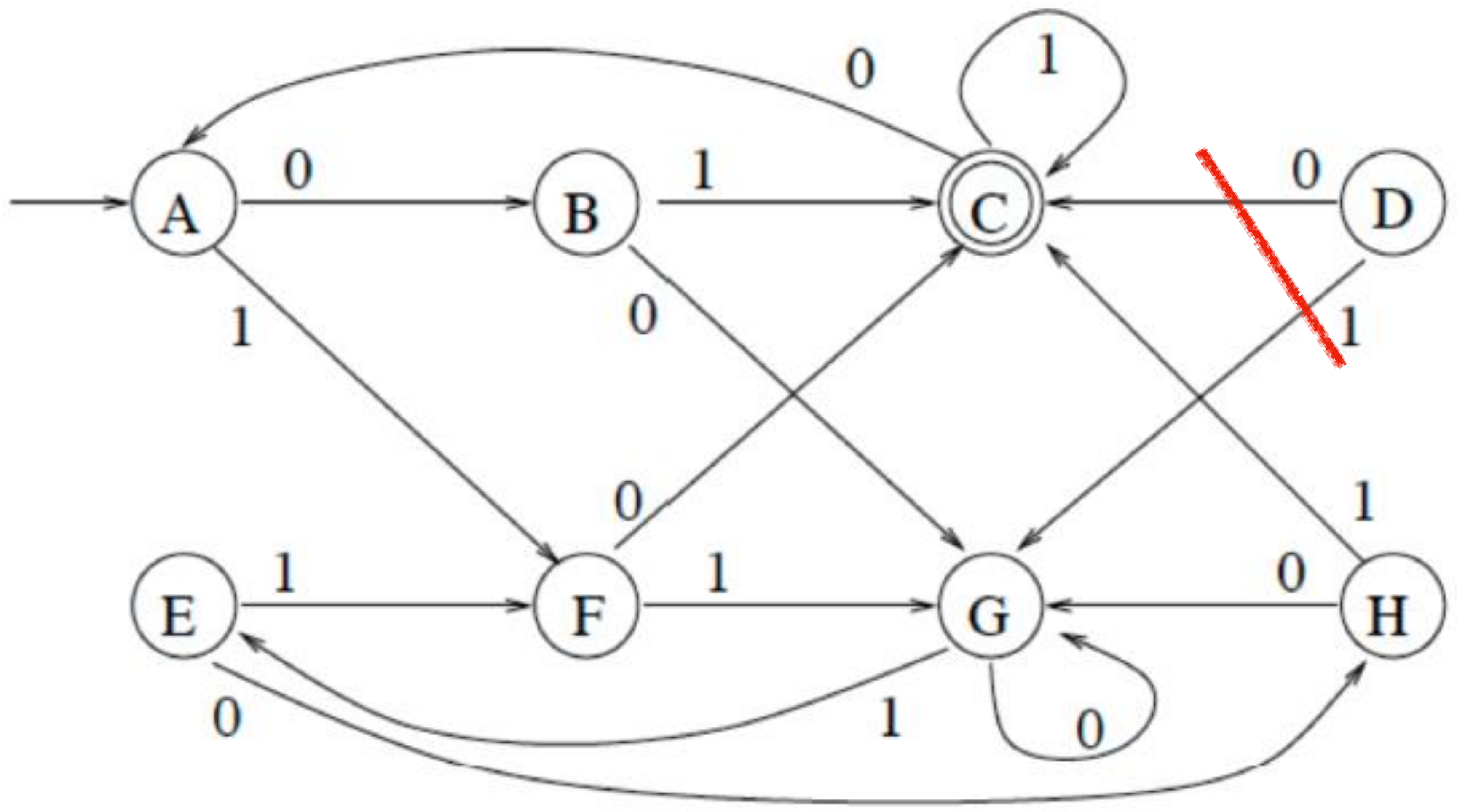
Figure 5: after 1 iteration

3. e.g. B,G marked, $\delta(A,0) = B, \delta(G,0) = G$, mark A,G

B	X					
C	X	X				
E		X	X			
F	X	X	X	X		
G	X	X	X	X	X	
H	X		X	X	X	X
	A	B	C	E	F	G

Figure 6: after 2 iterations

running example :



step 1: eliminate unreachable states

step 2: Mark the distinguishable pairs of states. To achieve this task, we first mark all pairs p,q, where $p \in F$ and $q \notin F$ as distinguishable.

Then, we proceed as follows:

repeat
for all non-marked pairs p,q do
for each letter a do
if the pair $\delta(p,a),\delta(q,a)$ is marked then mark p,q
until no new pairs are marked

3. e.g. B,G marked, $\delta(A,0) = B, \delta(G,0) = G$, mark A,G

B	X					
C	X	X				
E		X	X			
F	X	X	X	X		
G	X	X	X	X	X	
H	X		X	X	X	X
	A	B	C	E	F	G

Figure 6: after 2 iterations

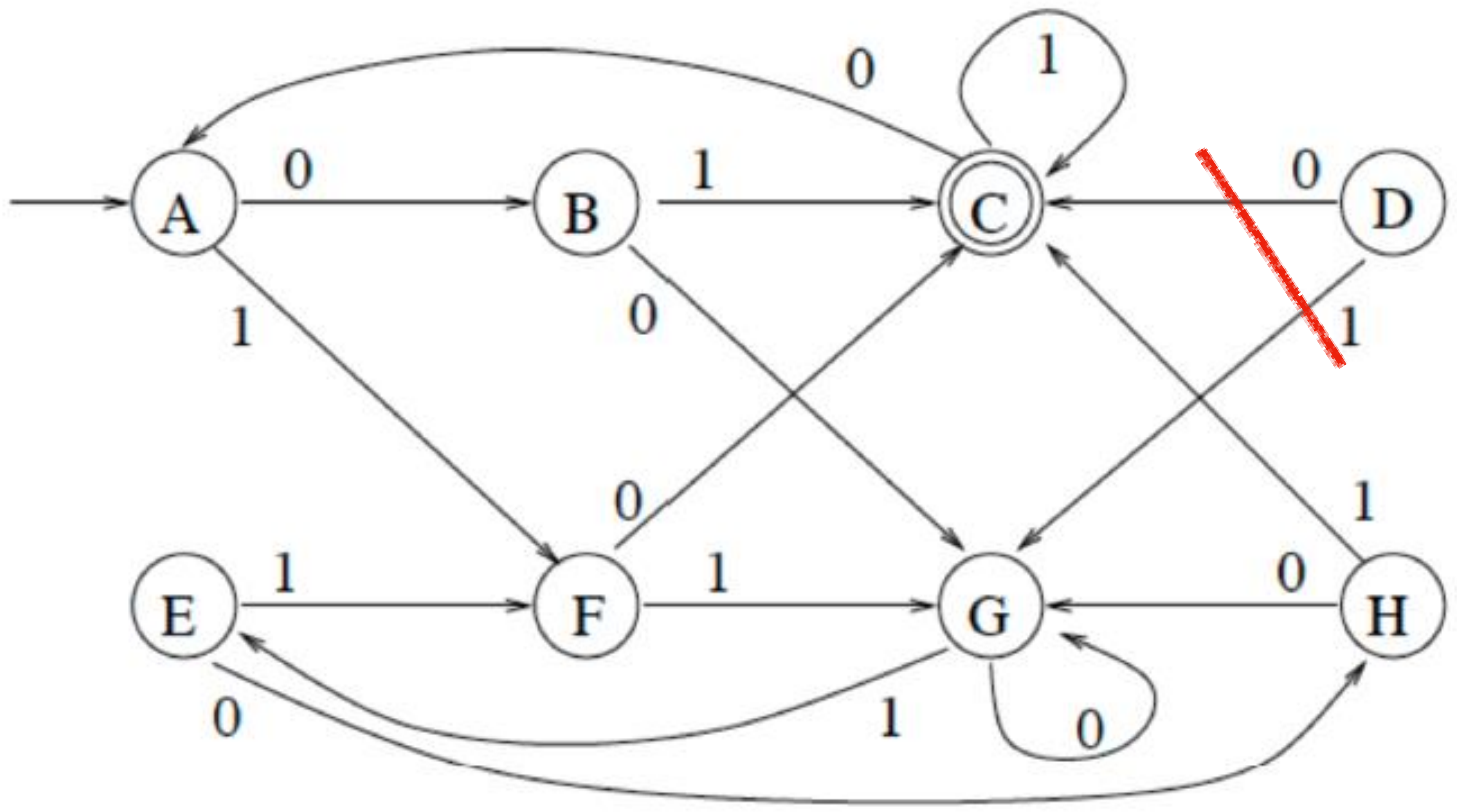
4. then nothing new

Lemma 4. *for all $p,q \in Q$:*

$$p,q \text{ marked} \leftrightarrow p,q \text{ distinguishable}$$

Proof. Induction on the length of the shortest string that distinguishes p,q, using Lemma 3 . Exercise. $\square$

running example :



step 1: eliminate unreachable states

step 2: Mark the distinguishable pairs of states. To achieve this task, we first mark all pairs p,q, where $p \in F$ and $q \notin F$ as distinguishable.

Then, we proceed as follows:

repeat
for all non-marked pairs p,q do
for each letter a do
if the pair $\delta(p,a), \delta(q,a)$ is marked then mark p,q
until no new pairs are marked

3. e.g. B,G marked, $\delta(A,0) = B, \delta(G,0) = G$, mark A,G

B	X					
C	X	X				
E		X	X			
F	X	X	X	X		
G	X	X	X	X	X	
H	X		X	X	X	X
	A	B	C	E	F	G

Figure 6: after 2 iterations

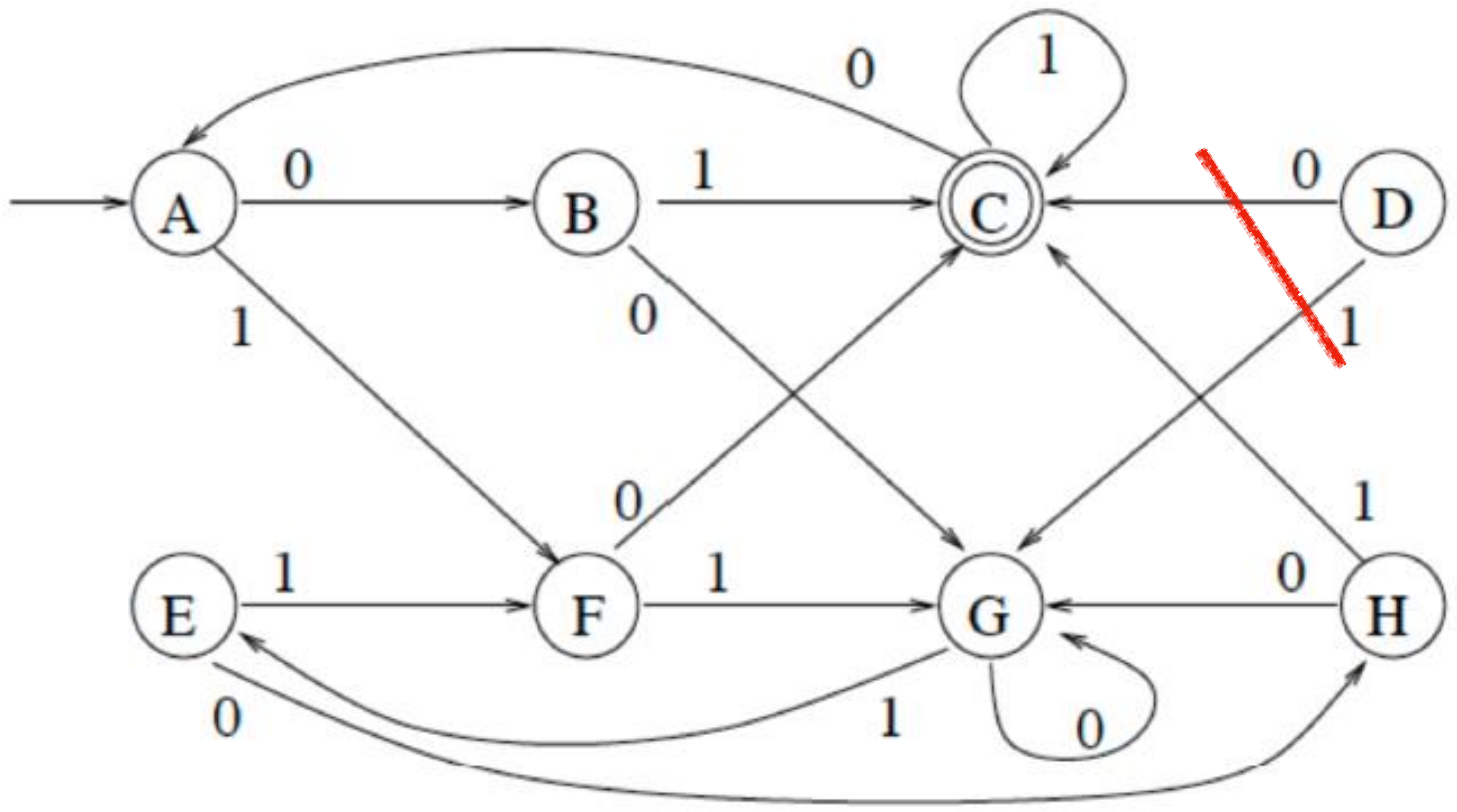
$$\hat{Q} = \{\{A,E\},\{B,H\},\{C\},\{F\},\{G\}\}$$

step 3: Construct the reduced automaton $\hat{B}$.

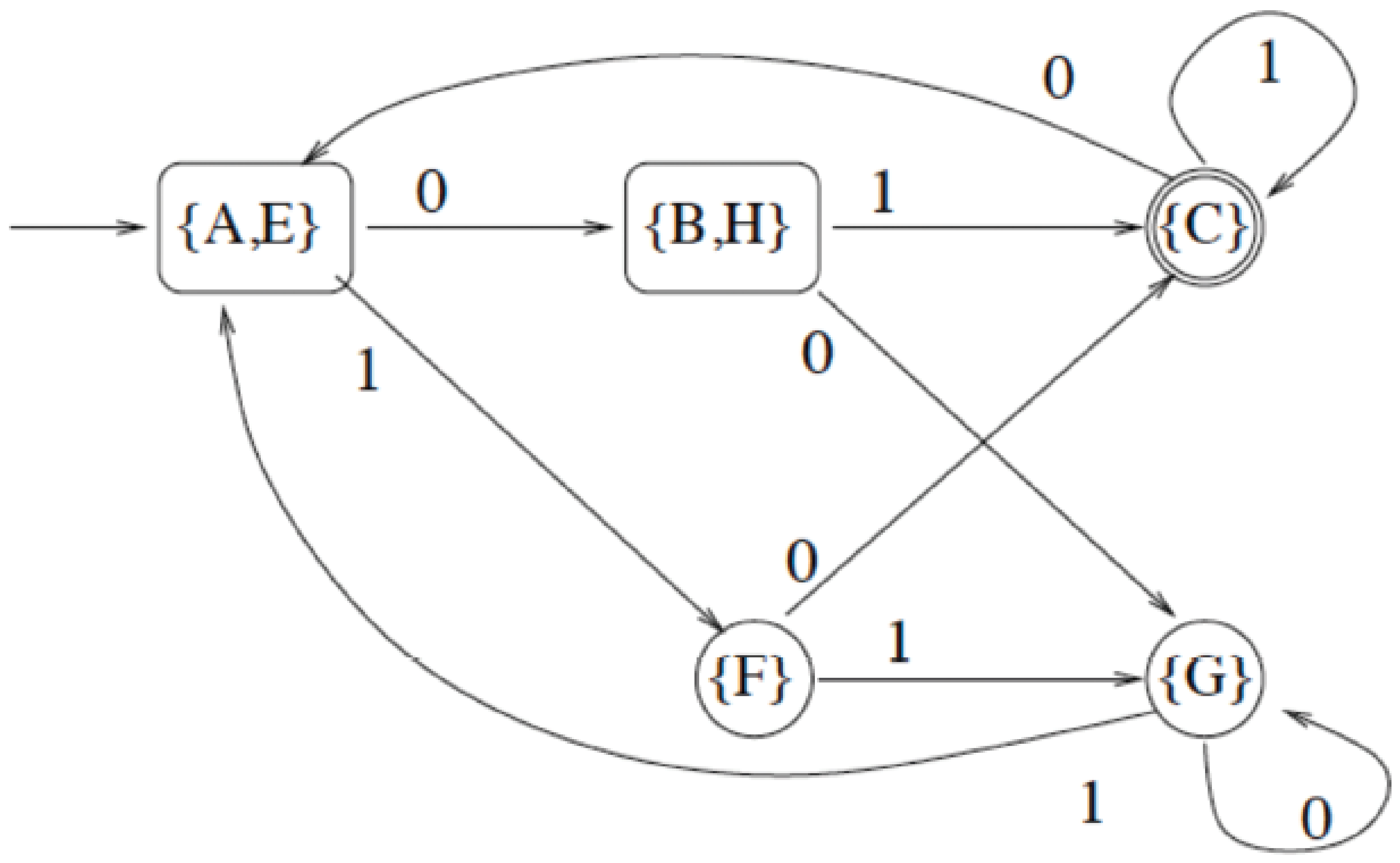
- determine the equivalence classes of the indistinguishability relation. For $q \in Q$

$$[q] = \{a \in Q \; : \; a,q \text{ not marked}\}$$

running example :



$\hat{Q} = \{\{A,E\},\{B,H\},\{C\},\{F\},\{G\}\}$



step 3: Construct the reduced automaton $\hat{B}$.

- determine the equivalence classes of the indistinguishability relation. For $q \in Q$

$$[q] = \{a \in Q \; : \; a,q \text{ not marked}\}$$

- states

$$\hat{Q} = \{[q] \; : \; q \in Q\}$$

- initial state

$$\hat{q_0} = [q_0]$$

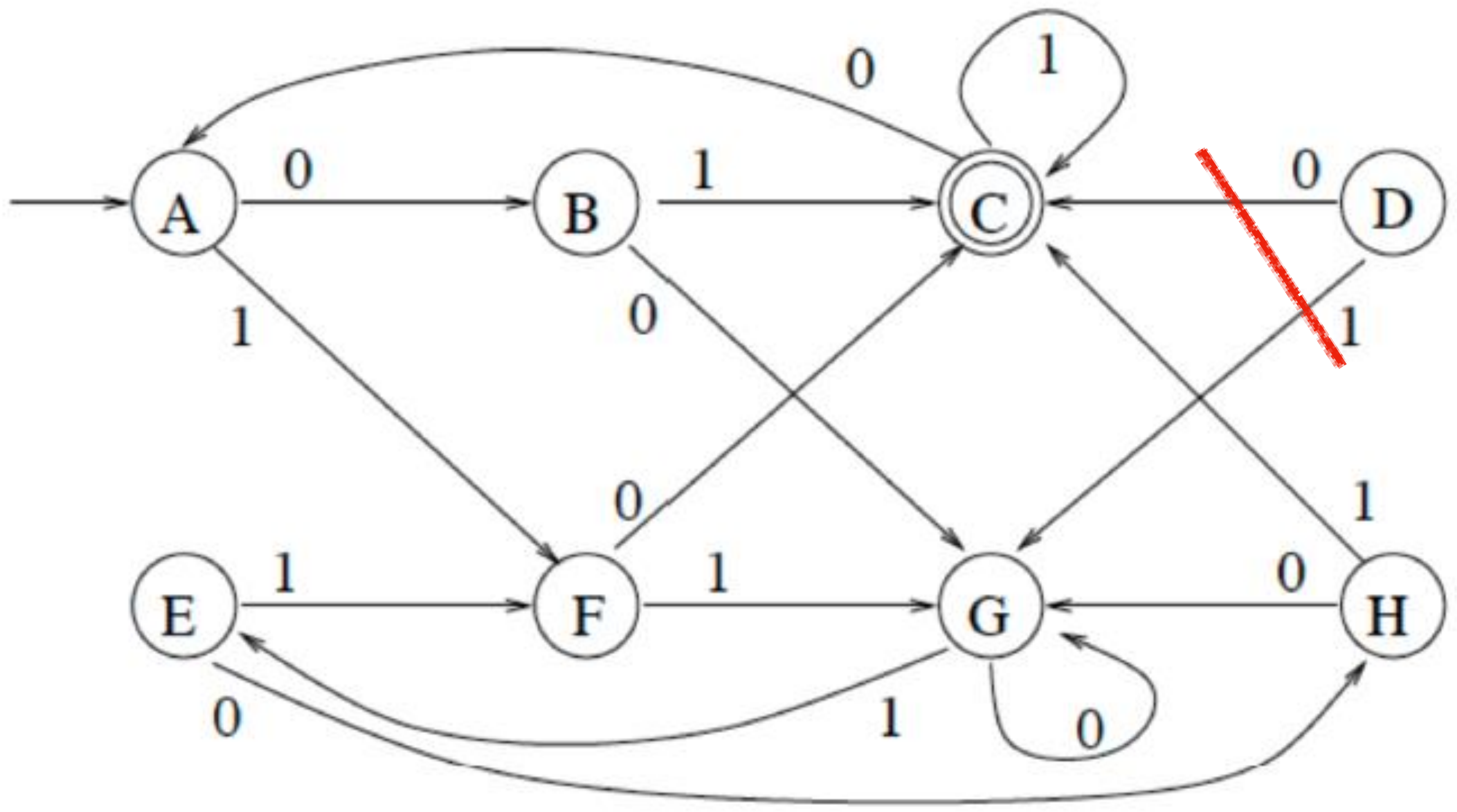
- final states

$$\hat{F} = \{[q] \; : \; q \in F\}$$

- transition function

$$\hat{\delta}([q],a) = [\delta(q,a)]$$

running example :



step 3: Construct the reduced automaton $\hat{B}$.

- determine the equivalence classes of the indistinguishability relation. For $q \in Q$

$$[q] = \{a \in Q \ : \ a, q \text{ not marked}\}$$

- states

$$\hat{Q} = \{[q] \ : \ q \in Q\}$$

- initial state

$$\hat{q_0} = [q_0]$$

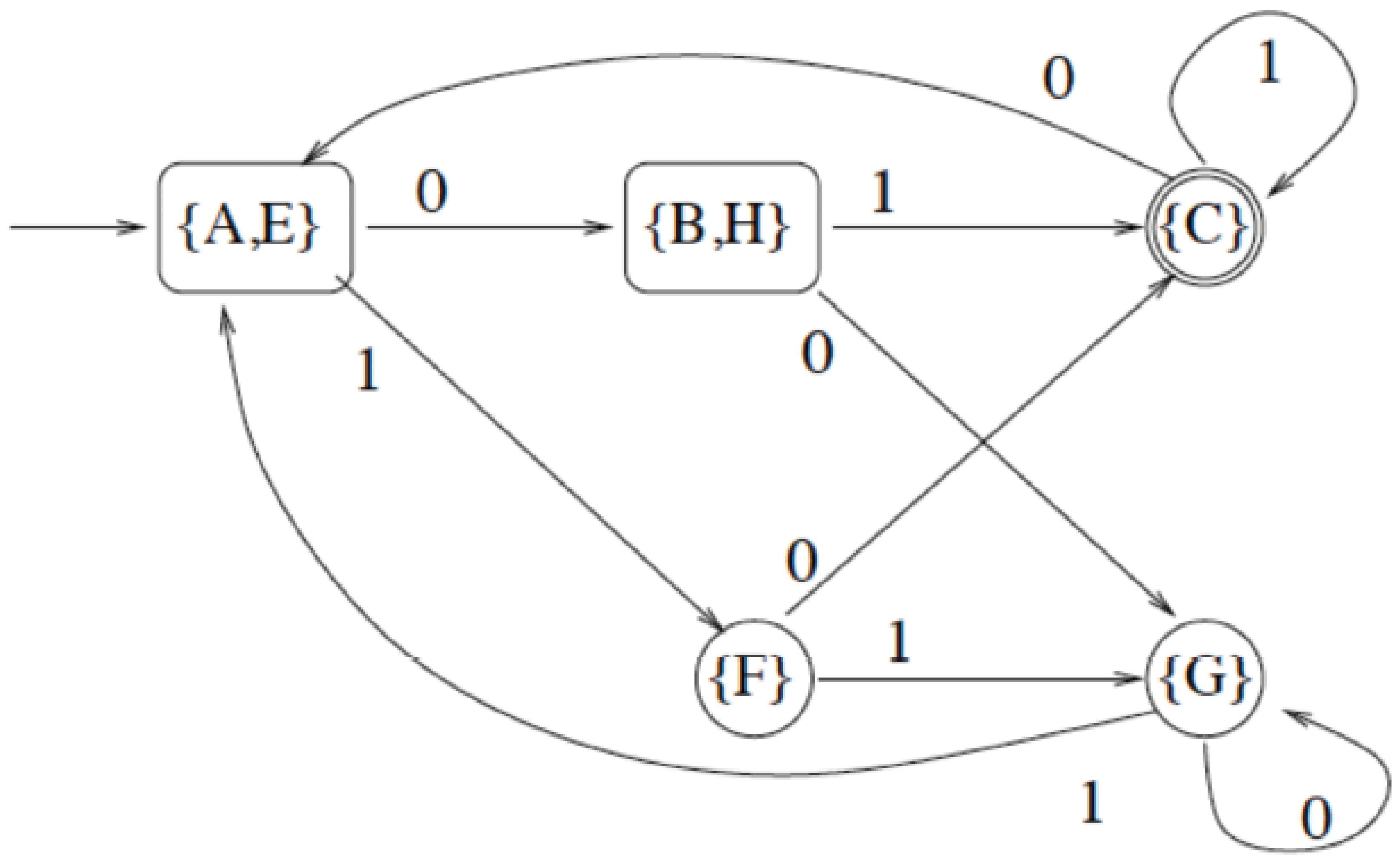
- final states

$$\hat{F} = \{[q] \ : \ q \in F\}$$

- transition function

$$\hat{\delta}([q], a) = [\delta(q, a)]$$

$$\hat{Q} = \{\{A,E\},\{B,H\},\{C\},\{F\},\{G\}\}$$



Lemma 5. *$\hat{B}$ is well defined*

Claim: definitions of $\hat{F}$ and $\hat{\delta}$ are independent of choice of $q \in [q]$

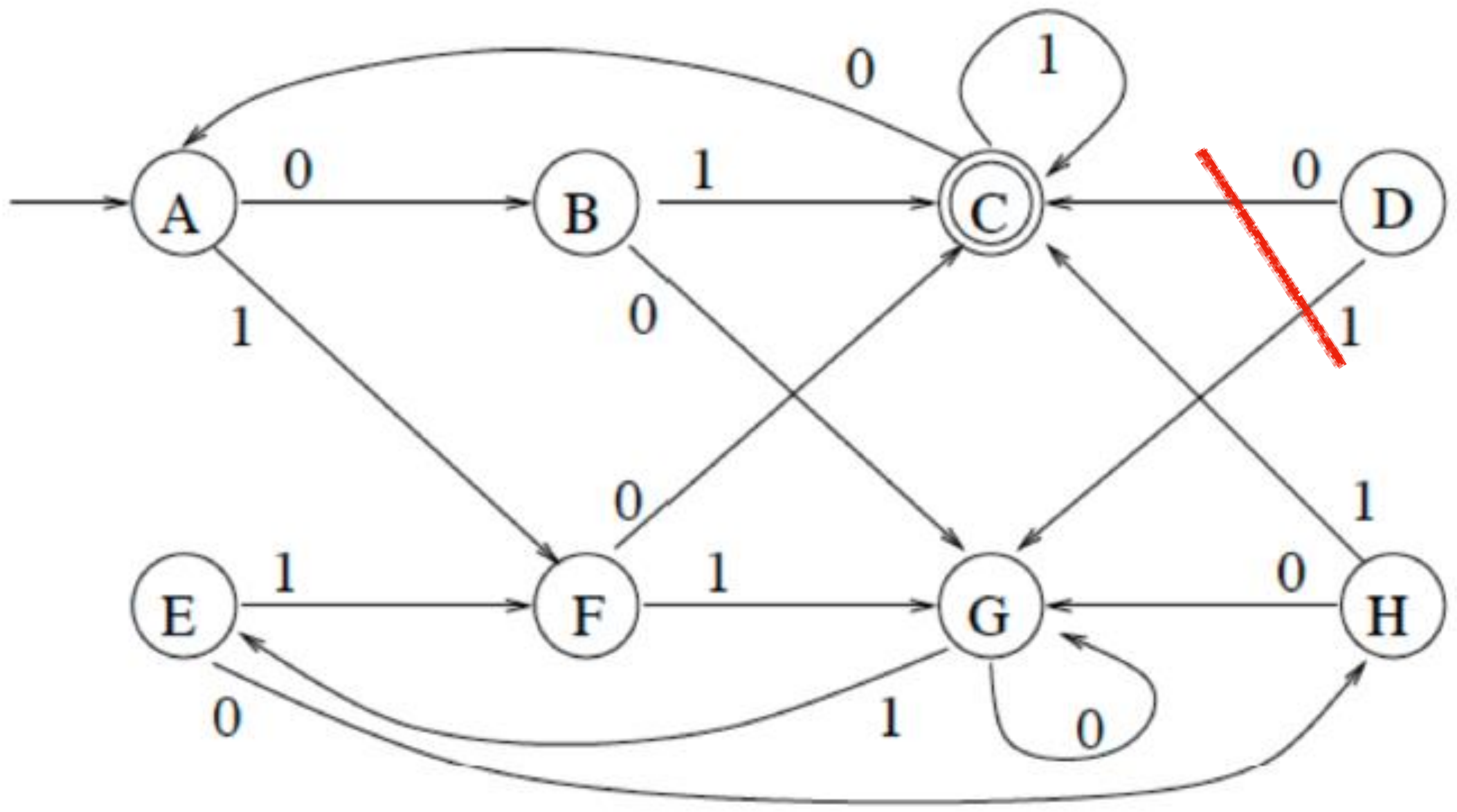
- $\hat{F}$: Let $q \in F$ and $p \notin F$. Then p, q marked in initial phase and stay marked. Hence

$$p \notin [q]$$

- $\hat{\delta}$ Let $p \in [q]$. As a does not distinguish p and q:

$$\delta(p,a) = \delta(q,a) \quad , \quad [\delta(p,a)] = [\delta(q,a)]$$

running example :



step 3: Construct the reduced automaton $\hat{B}$.

- determine the equivalence classes of the indistinguishability relation. For $q \in Q$

$$[q] = \{a \in Q \ : \ a, q \text{ not marked}\}$$

- states

$$\hat{Q} = \{[q] \ : \ q \in Q\}$$

- initial state

$$\hat{q}_0 = [q_0]$$

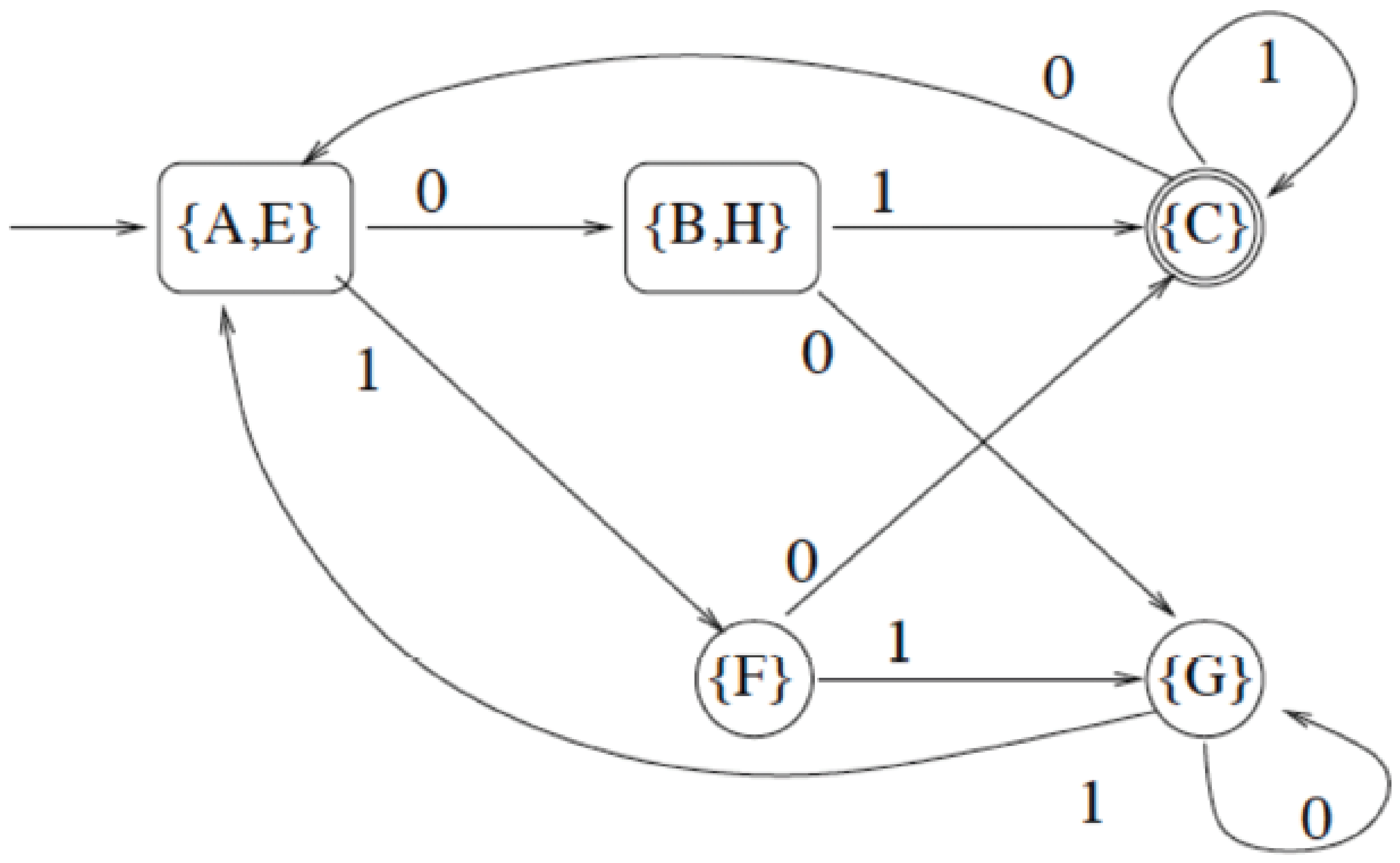
- final states

$$\hat{F} = \{[q] \ : \ q \in F\}$$

- transition function

$$\hat{\delta}([q], a) = [\delta(q, a)]$$

$$\hat{Q} = \{\{A, E\}, \{B, H\}, \{C\}, \{F\}, \{G\}\}$$



Lemma 6.

$$L(B) = L(\hat{B})$$

- Claim: for all w

$$\delta^*(q_0, w) \in \hat{\delta}^*(\hat{q}_0, w)$$

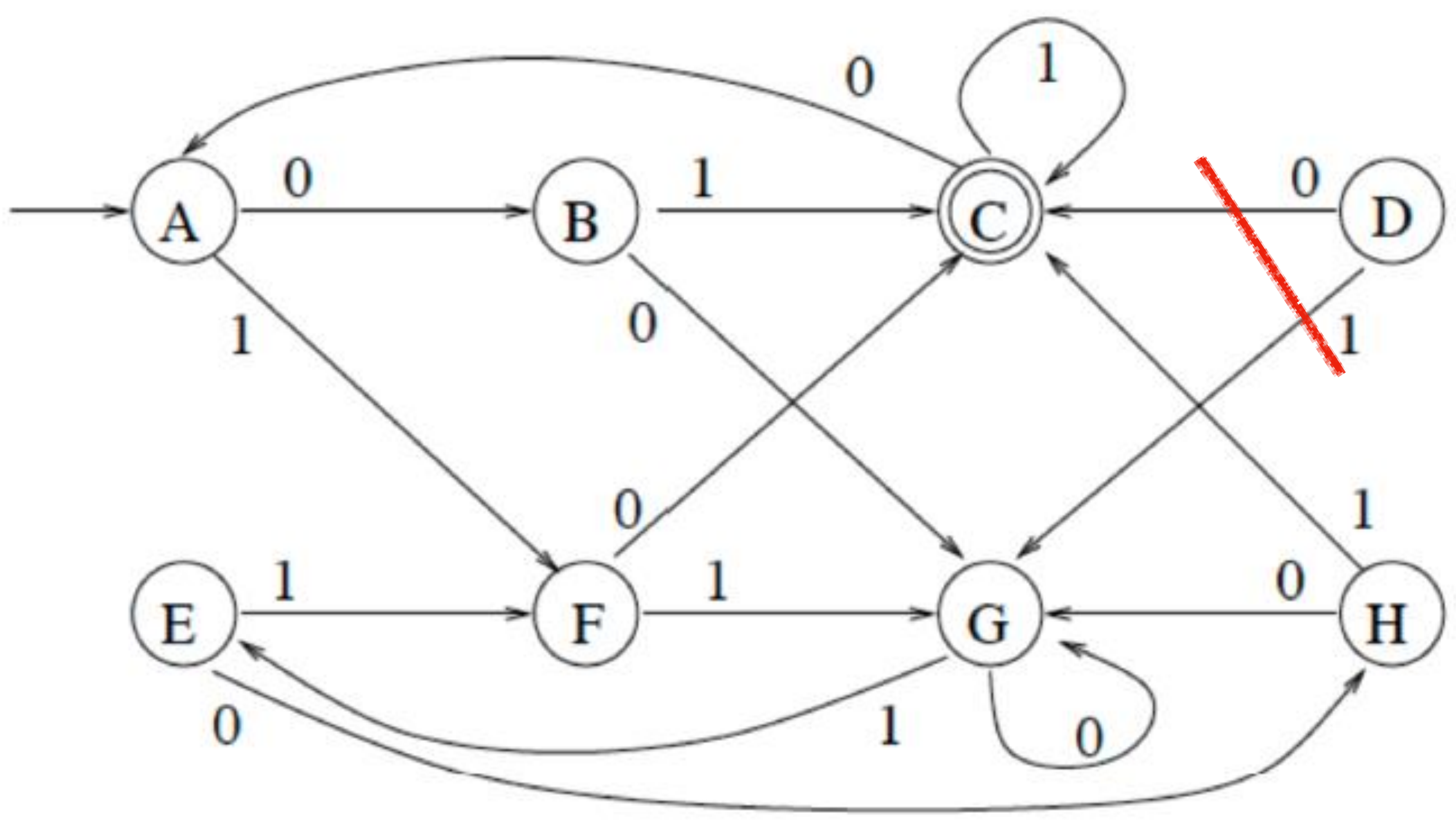
  Proof: induction on$|w|$.

-

$$\delta^*(q_0, w) \in F \ \rightarrow \hat{\delta}^*(\hat{q}_0, w) = [\delta^*(q_0, w)] \in \hat{F}$$

$$\delta^*(q_0, w) \notin F \ \rightarrow \hat{\delta}^*(\hat{q}_0, w) = [\delta^*(q_0, w)] \notin \hat{F}$$

running example :



step 3: Construct the reduced automaton $\hat{B}$.

- determine the equivalence classes of the indistinguishability relation. For $q \in Q$

$$[q] = \{a \in Q \ : \ a, q \text{ not marked}\}$$

- states

$$\hat{Q} = \{[q] \ : \ q \in Q\}$$

- initial state

$$\hat{q_0} = [q_0]$$

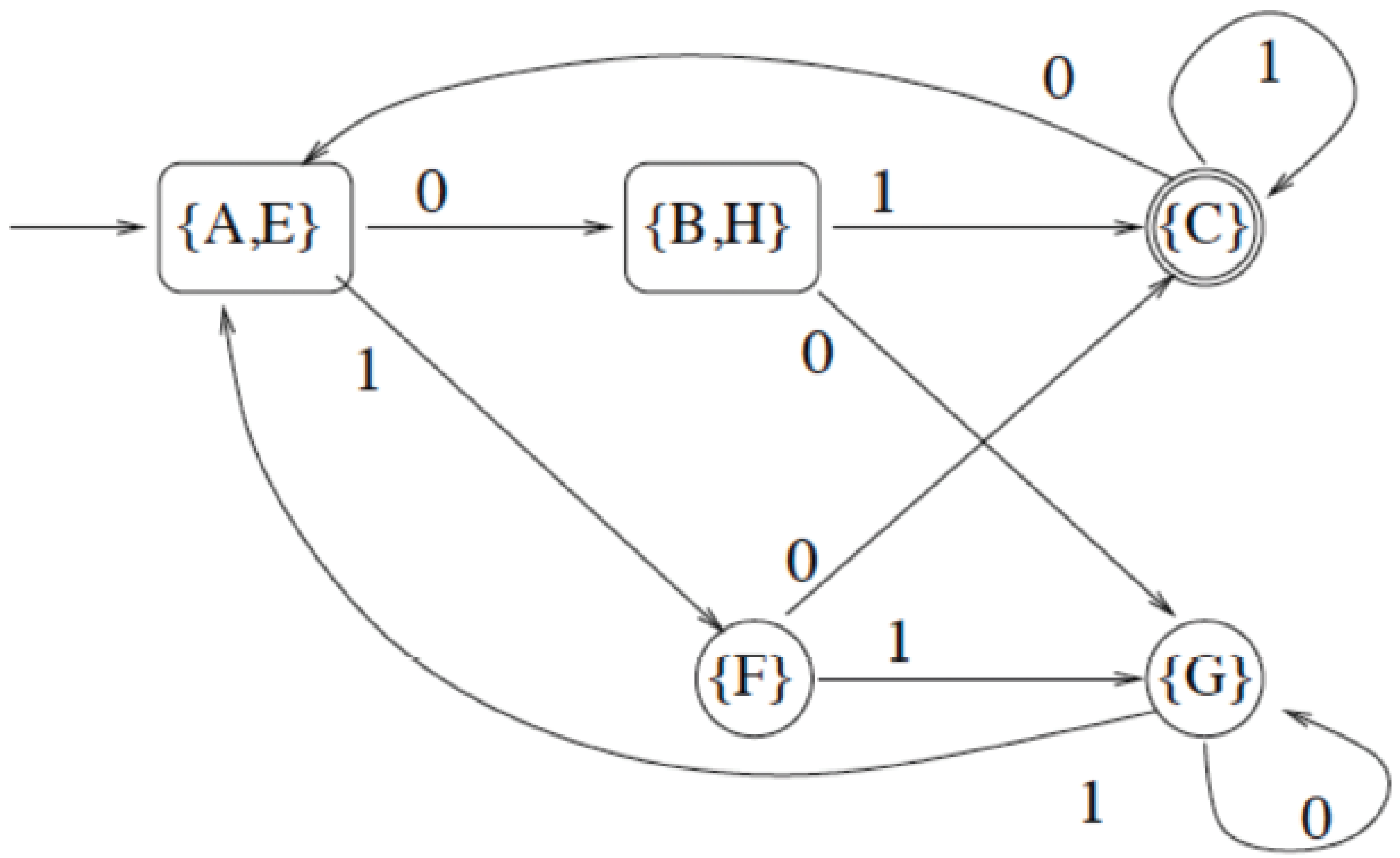
- final states

$$\hat{F} = \{[q] \ : \ q \in F\}$$

- transition function

$$\hat{\delta}([q], a) = [\delta(q, a)]$$

$$\hat{Q} = \{\{A,E\},\{B,H\},\{C\},\{F\},\{G\}\}$$



Lemma 7. *$\hat{B}$ is minimal.*

- states of $\hat{B}$ are distinguishable.
- apply lemma 1