Linguagens Formais e Autômatos

Aula 03 - Minimização de DFAs

Prof. Dr. Daniel Lucrédio
Departamento de Computação / UFSCar
Última revisão: ago/2015

Referências bibliográficas

- **Introdução à teoria dos autômatos, linguagens e computação / John E. Hopcroft, Rajeev Motwani, Jeffrey D. Ullman** ; tradução da 2.ed. original de Vandenberg D. de Souza. - Rio de Janeiro : Elsevier, 2002 (Tradução de: Introduction to automata theory, languages, and computation - ISBN 85-352-1072-5)
  - Capítulo 4 - Seção 4.4

Minimização de DFAs

- Existe um procedimento que minimiza um DFA
  - Ou seja, dado um DFA, ele permite encontrar um DFA equivalente que tenha o número mínimo de estados.
- De fato, esse DFA é mínimo:
  - Teorema: Se A é um DFA e M é o DFA construído a partir de A pelo algoritmo descrito a seguir, então M tem tão poucos estados quanto qualquer DFA equivalente a A
  - Em outras palavras, podemos testar a equivalência entre DFAs
    - Minimizando os dois e verificando se são iguais (com exceção, possivelmente, dos nomes dos estados)

Minimização de DFAs

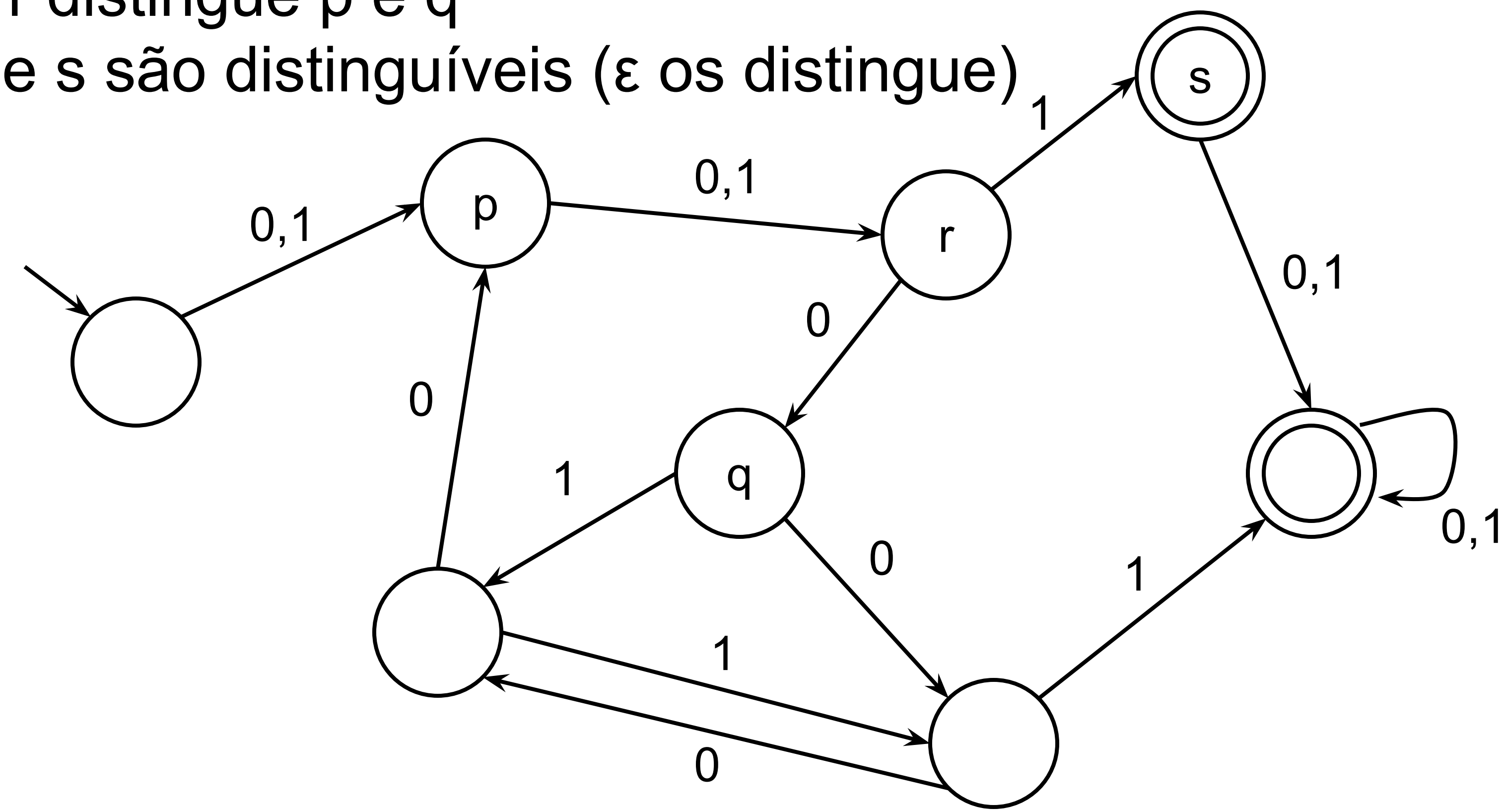
- Conceito de estados equivalentes
  - Objetivo: entender quando dois estados distintos p e q podem ser substituídos por um único estado que se comporte como p e q
- Formalmente:
  - Dois estados p e q são equivalentes se:
    - Para todas as cadeias de entrada w, $\delta^\wedge(p, w)$ é um estado de aceitação se e somente se $\delta^\wedge(q, w)$ é um estado de aceitação

Minimização de DFAs

- Menos formalmente:
  - Existe uma cadeia w que leva p à aceitação e w à não-aceitação (ou vice-versa)?
  - Se existir pelo menos uma cadeia assim, os estados são distinguíveis
    - Caso contrário, são equivalentes!

Minimização de DFAs

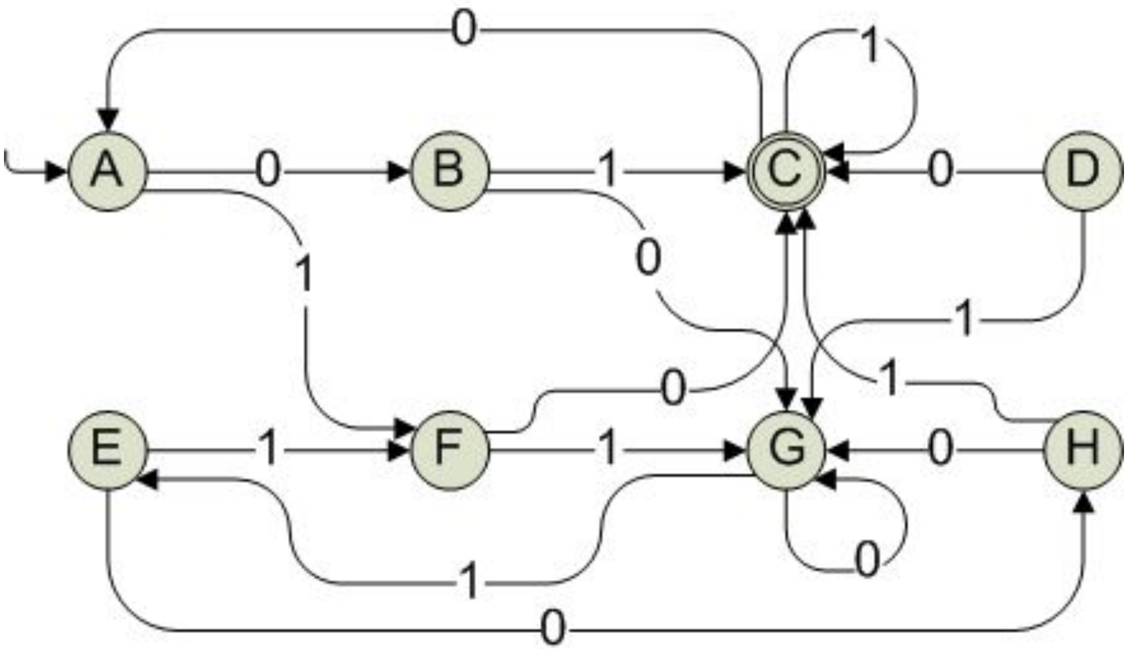
- Ilustrando:
  - 0,1, 010, 111 não distingue p e q
  - 11 distingue p e q
  - r e s são distinguíveis (ε os distingue)



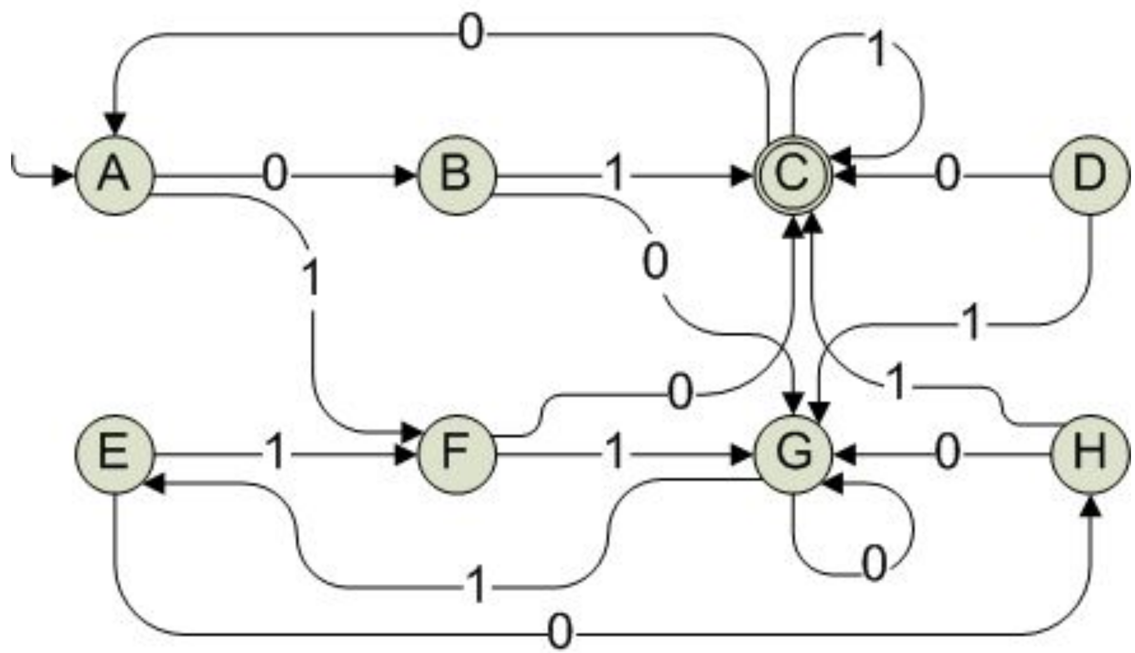
Minimização de DFAs

- Difícil encontrar estados equivalentes apenas "olhando" para o DFA
  - Muitas combinações, fácil se perder
- Estratégia sistemática: encontrar todos os pares de estados que sejam distinguíveis
  - Se fizermos o melhor possível
    - Qualquer par de estados que não considerarmos distinguíveis serão equivalentes
- Algoritmo de preenchimento de tabela
  - Descoberta recursiva de pares distinguíveis
  - Cada célula da tabela marca um par distinguível
  - Células em branco marcam pares equivalentes

	A	B	C	D	E	F	G
B							
C							
D							
E							
F							
G							
H							

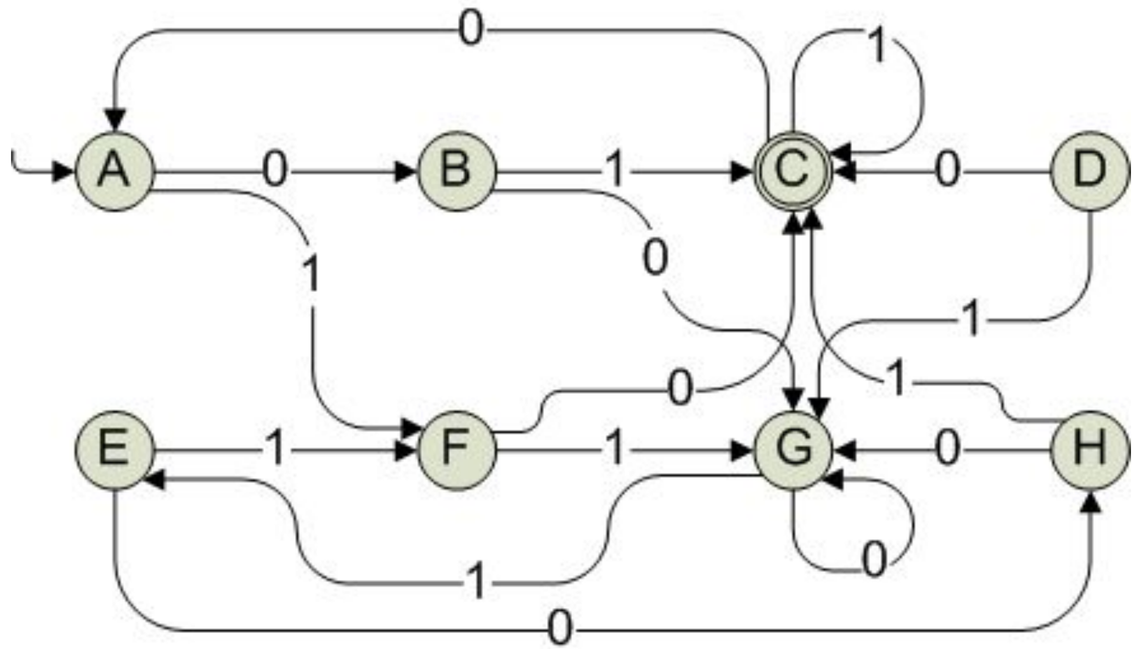


	A	B	C	D	E	F	G
B							
C	x	x					
D			x				
E			x				
F			x				
G			x				
H			x				

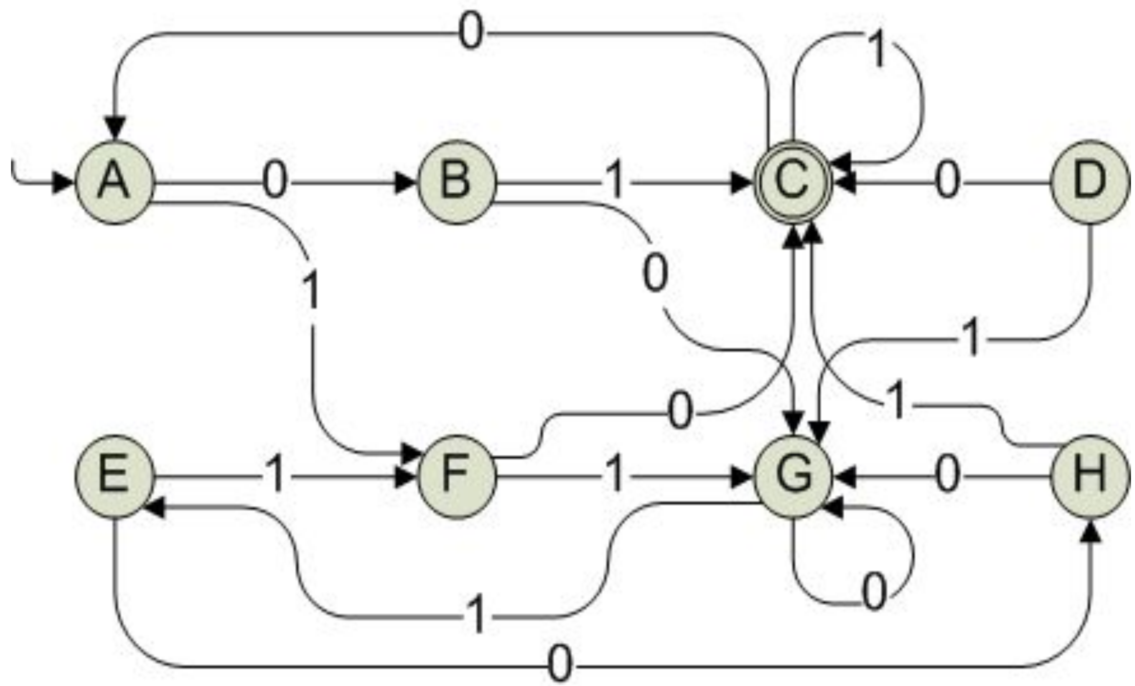


Começamos pelos estados de aceitação/não-aceitação. São obviamente pares distinguíveis pela cadeia vazia

	A	B	C	D	E	F	G
B							
C	x	x					
D			x				
E			x				
F			x				
G			x				
H			x				



Agora tentamos encontrar outros estados que "chegam" em um par conhecido, dada uma mesma entrada.
A técnica é seguir, para cada par distinguível, as setas pelo lado inverso, com um mesmo rótulo
Fica mais fácil se marcar as células já analisadas

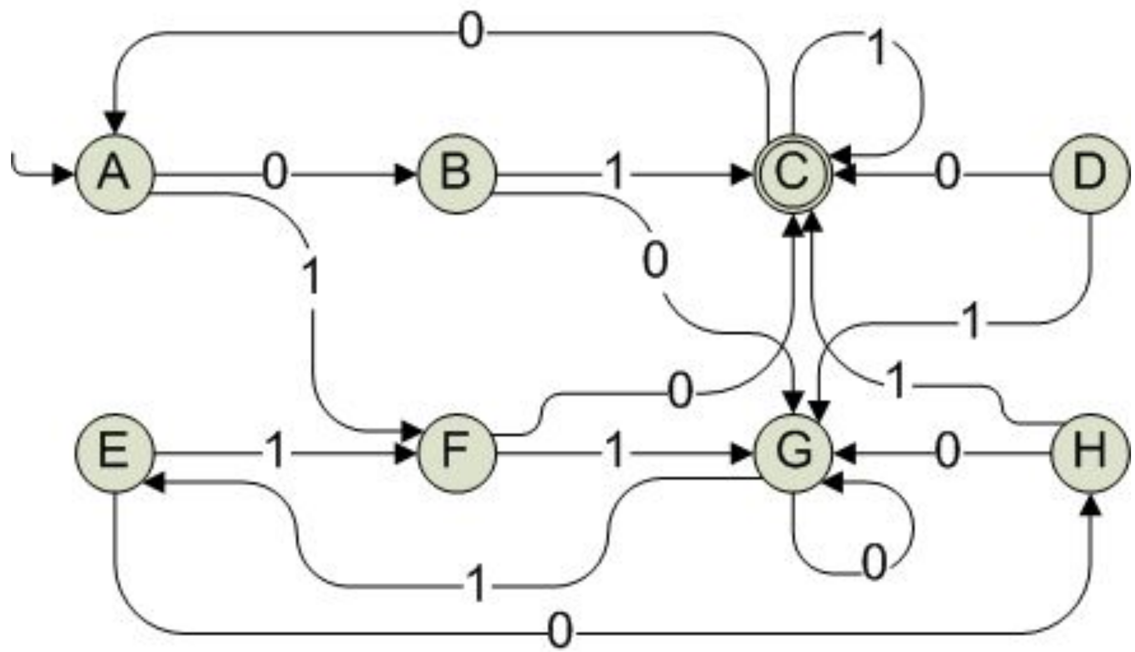


	A	B	C	D	E	F	G
B							
C	x	x					
D			x				
E			x				
F			x				
G			x				
H			x				

Exs:

(a) Seguindo as setas que “chegam” em A e C (um par distinguível), mediante entrada 0, temos:

- SetasA_0:{C}, SetasC_0:{D,F}
- Novos pares = SetasA_0 x SetasC_0 = {(C,D), (C,F)}
- Estes pares já estão marcados na tabela, com um x
- Analisando para entrada 1, temos:
- SetasA_1: {}, SetasC_1: {B,C,H}
- Novos pares = SetasA_1 X SetasC_1 = {} (nenhum novo par)
- Uma vez que já analisamos as entradas 0 e 1, a célula (A,C) foi analisada e é marcada

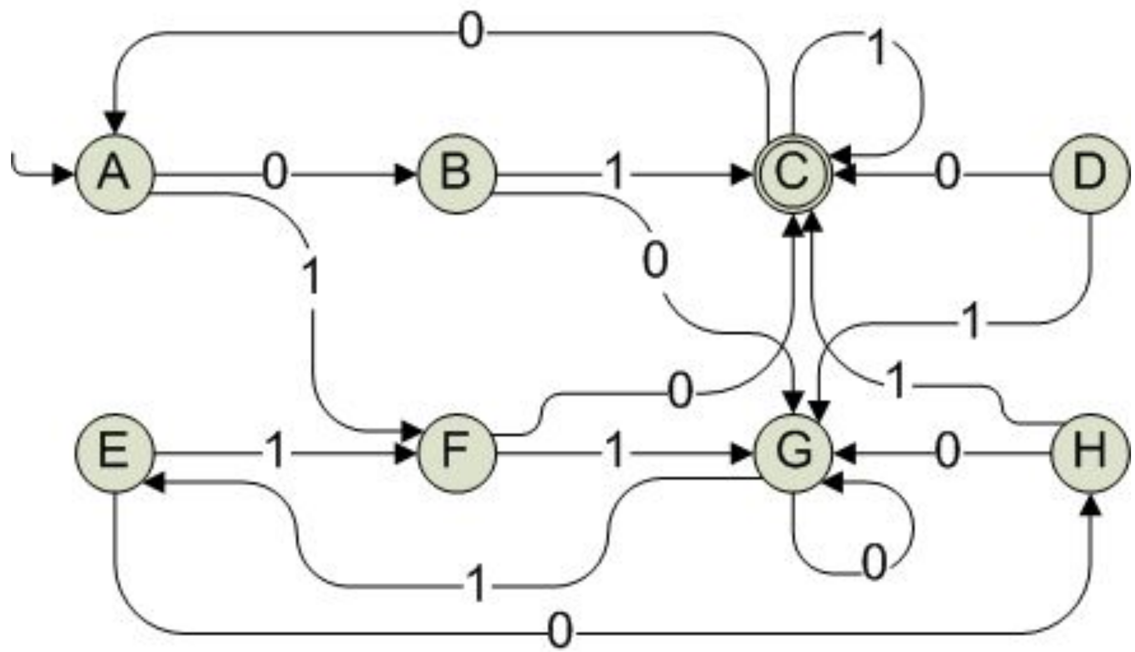


	A	B	C	D	E	F	G
B							
C	x	x					
D	x		x				
E			x				
F	x		x				
G			x				
H			x				

Exs:

(b) Continuando para par (B,C):

- SetasB_0:{A}, SetasC_0:{D,F}
- Novos pares = SetasB_0 x SetasC_0 = {(A,D), (A,F)}
- Esses pares ainda não foram marcados, e portanto a tabela precisa ser atualizada
- SetasB_1: {}, SetasC_1: {B,C,H}
- Novos pares = SetasB_1 X SetasC_1 = {} (nenhum novo par)
- Uma vez que já analisamos as entradas 0 e 1, a célula (B,C) foi analisada e é marcada

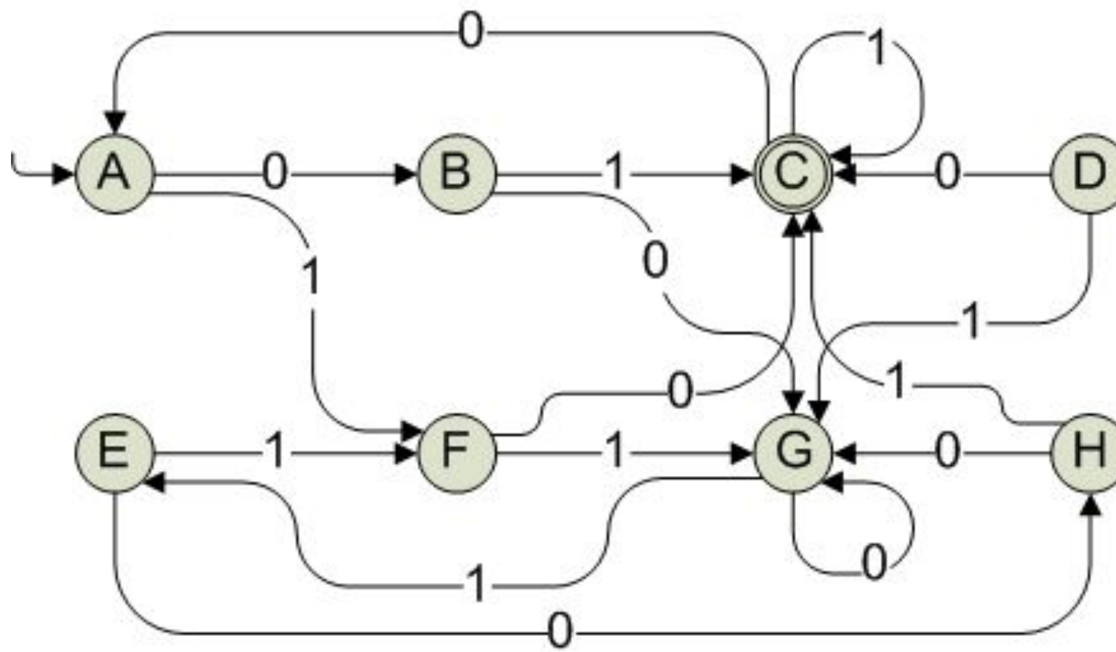


	A	B	C	D	E	F	G
B							
C	x	x					
D	x		x				
E			x				
F	x		x				
G			x				
H			x				

Exs:

(c) Continuando para par (C,D):

- SetasC_0:{D,F}, SetasD_0:{}
- Novos pares = {}
- SetasC_1:{B,C,H}, SetasD_1: {}
- Novos pares = {}
- Nenhum novo par

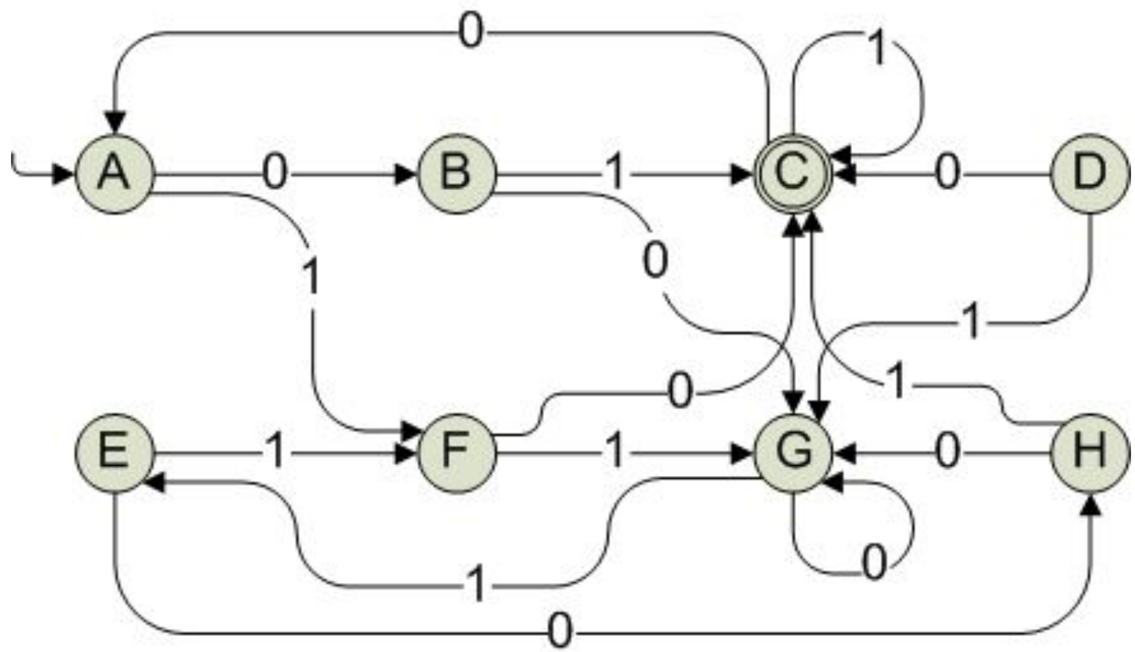


	A	B	C	D	E	F	G
B							
C	x	x					
D	x		x				
E			x				
F	x		x				
G		x	x				
H			x				x

Exs:

(d) Continuando para par (C,E):

- SetasC_0:{D,F}, SetasE_0:{}
- Novos pares = {}
- SetasC_1:{B,C,H}, SetasE_1: {G}
- Novos pares = {(B,G),(C,G),(H,G)}
- Novos pares e a célula (C,E) são marcados

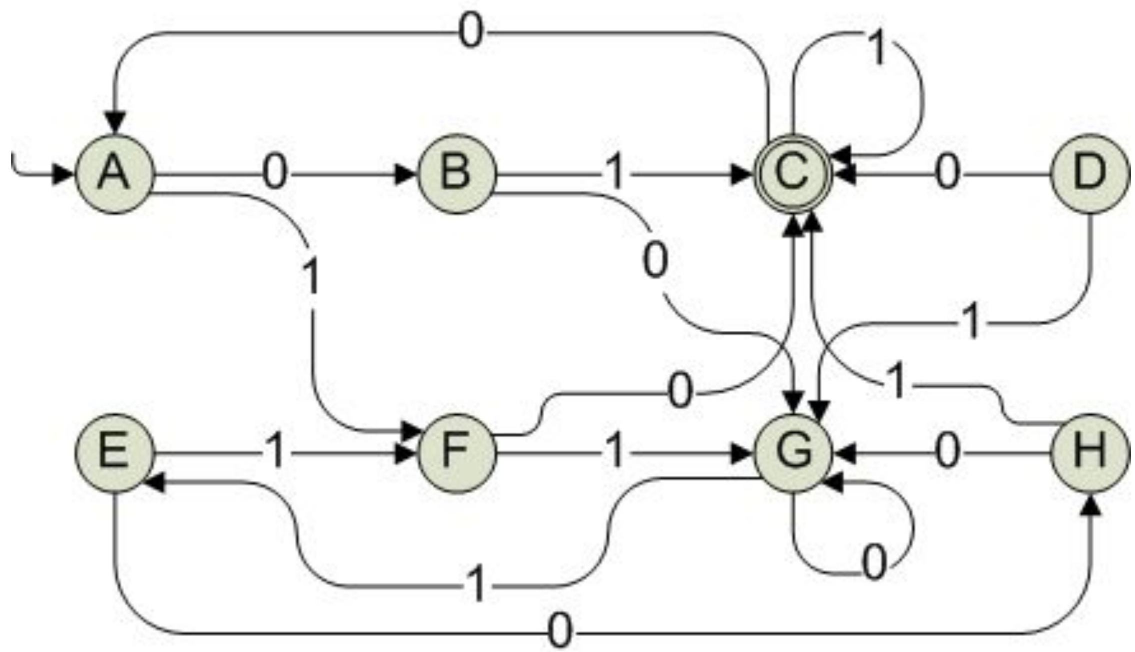


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x		x				
E		x	x				
F	x		x				
G		x	x				
H	x		x		x		x

Exs:

(e) Continuando para par (C,F):

- SetasC_0:{D,F}, SetasF_0:{}
- Novos pares = {}
- SetasC_1:{B,C,H}, SetasF_1: {A,E}
- Novos pares = {(B,A),(B,E),(C,A),(C,E),(H,A),(H,E)}
- Novos pares e a célula (C,F) são marcados

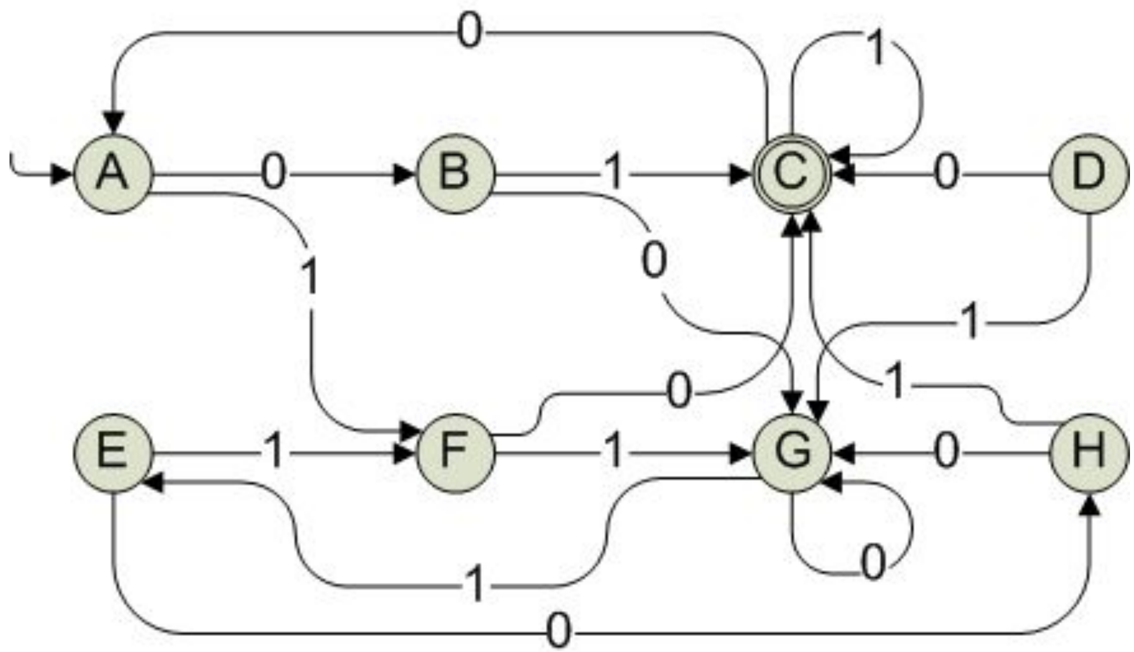


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x				
F	x	x	x				
G		x	x	x		x	
H	x		x	x	x	x	x

Exs:

(f) Continuando para par (C,G):

- SetasC_0:{D,F}, SetasG_0:{B,G,H}
- Novos pares = {(D,B),(D,G),(D,H),(F,B),(F,G),(F,H)}
- SetasC_1:{B,C,H}, SetasG_1: {D,F}
- Novos pares = {(B,D),(B,F),(C,D),(C,F),(H,D),(H,F)}
- Novos pares e a célula (C,G) são marcados

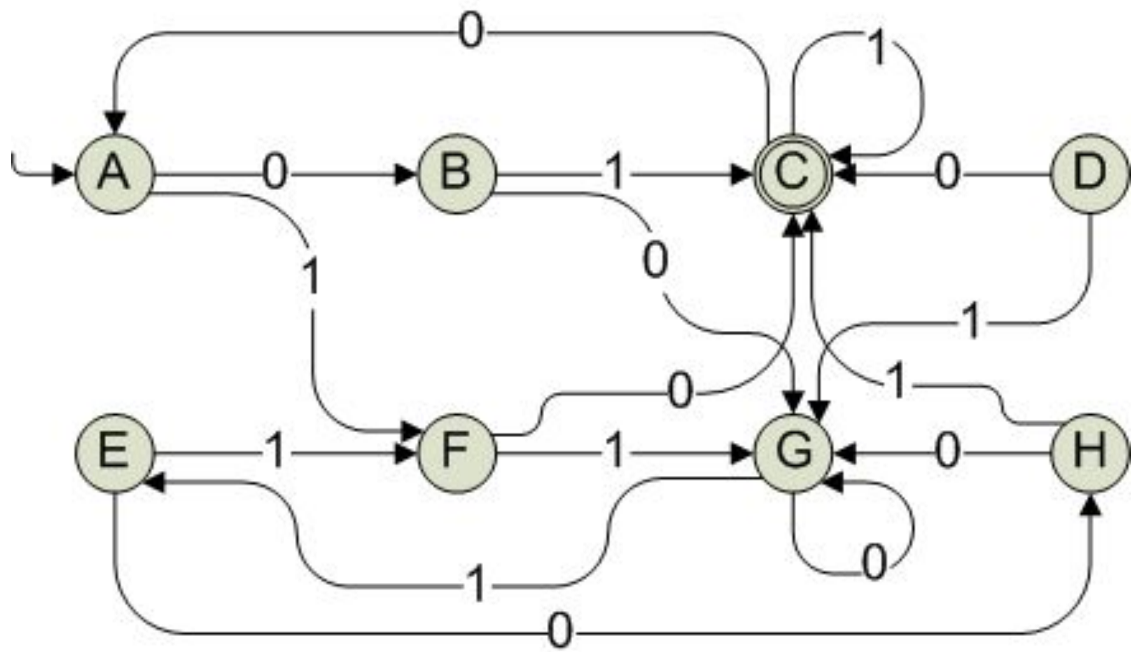


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G		x	x	x		x	
H	x		x	x	x	x	x

Exs:

(g) Continuando para par (C,H):

- SetasC_0:{D,F}, SetasH_0:{E}
- Novos pares = {(D,E),(F,E)}
- SetasC_1:{B,C,H}, SetasH_1: {}
- Novos pares = {}
- Novos pares e a célula (C,H) são marcados

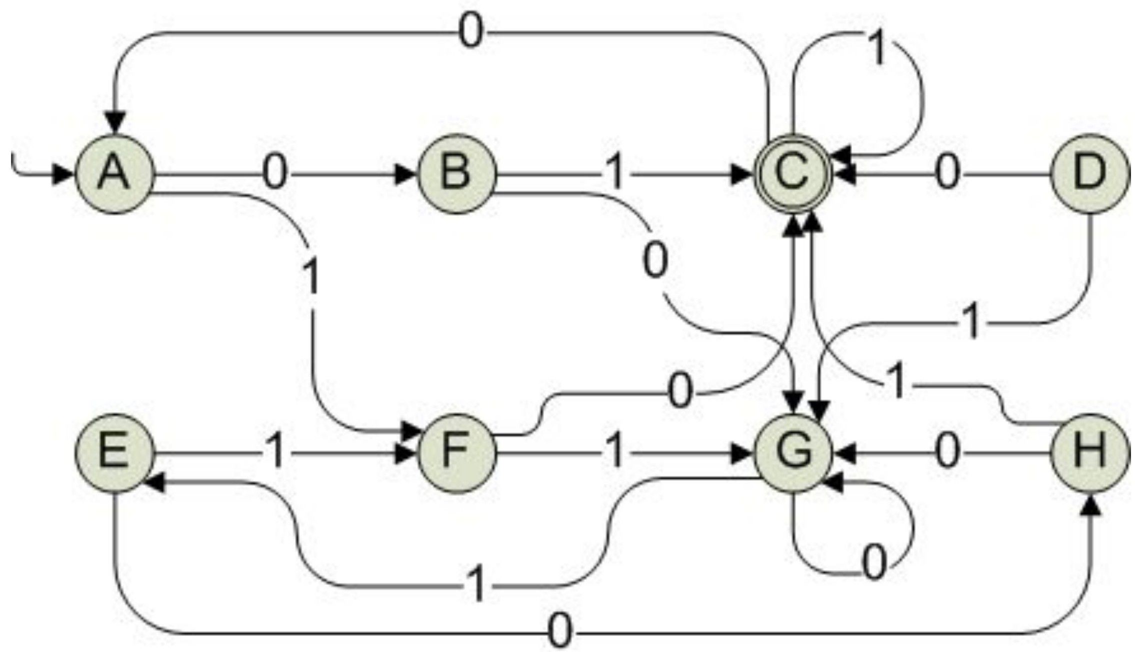


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G		x	x	x		x	
H	x		x	x	x	x	x

Exs:

(h) Continuando para par (A,B):

- SetasA_0:{C}, SetasB_0:{A}
- Novos pares = {(A,C)}
- SetasA_1:{}, SetasB_1: {}
- Novos pares = {}
- Novos pares e a célula (A,B) são marcados

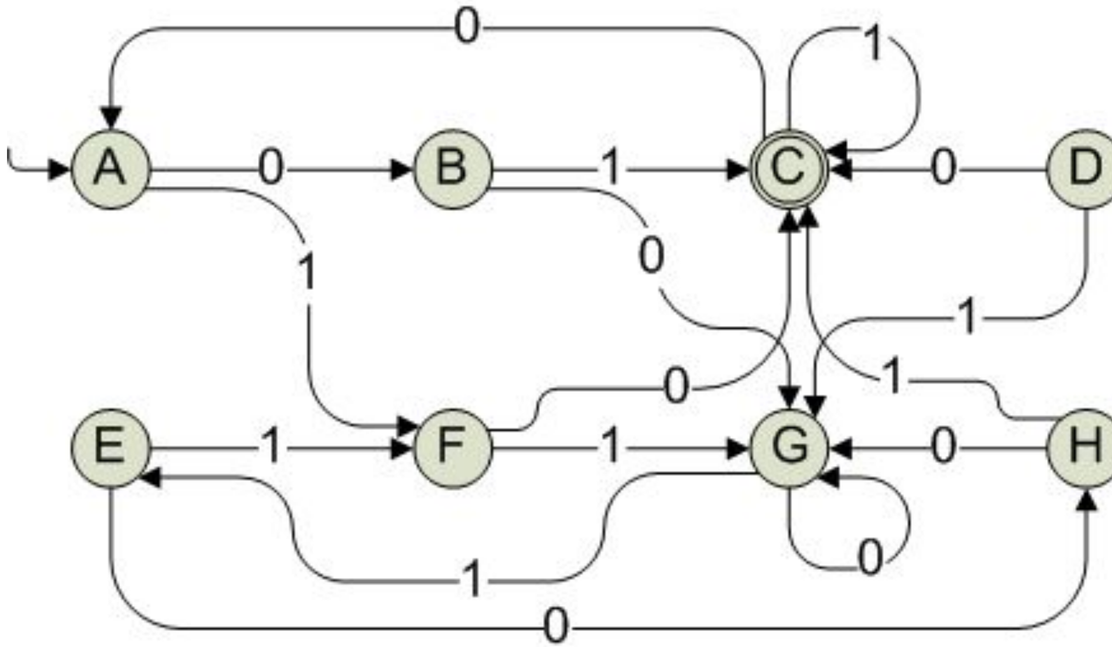


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G		x	x	x		x	
H	x		x	x	x	x	x

Exs:

(h) Continuando para par (A,D):

- SetasA_0:{C}, SetasD_0:{}
- Novos pares = {(A,C)}
- SetasA_1:{}, SetasD_1: {}
- Novos pares = {}
- Célula (A,D) é marcada

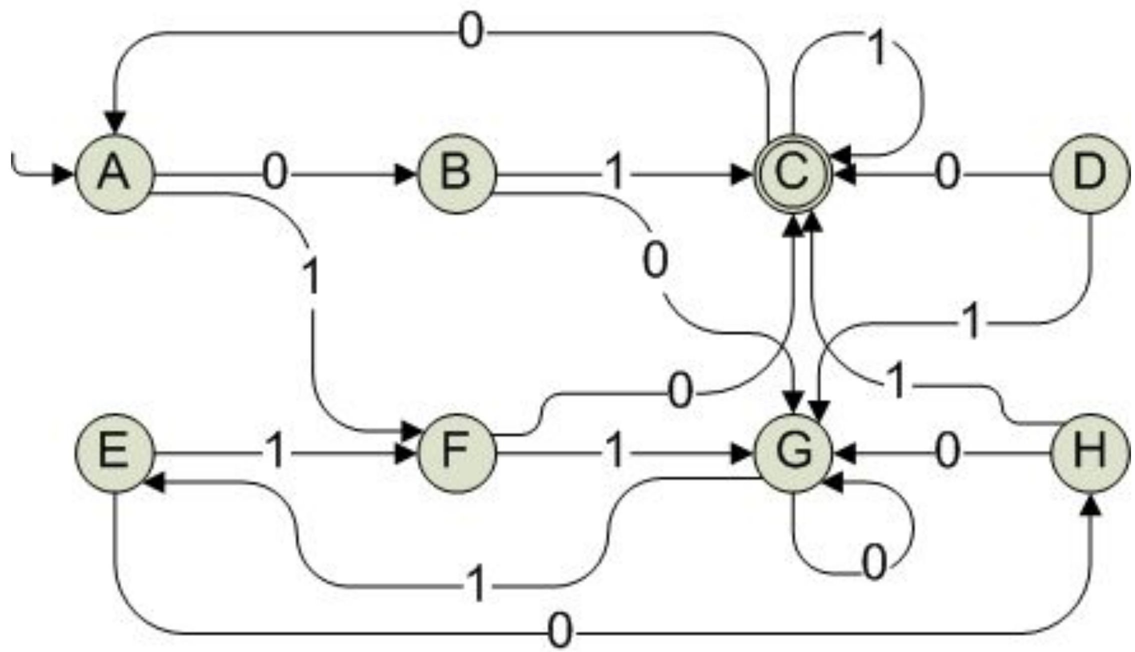


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G		x	x	x		x	
H	x		x	x	x	x	x

Exs:

(h) Continuando para par (A,F):

- SetasA_0:{C}, SetasF_0:{}
- Novos pares = {(A,C)}
- SetasA_1:{}, SetasF_1: {A,E}
- Novos pares = {}
- Célula (A,F) é marcada

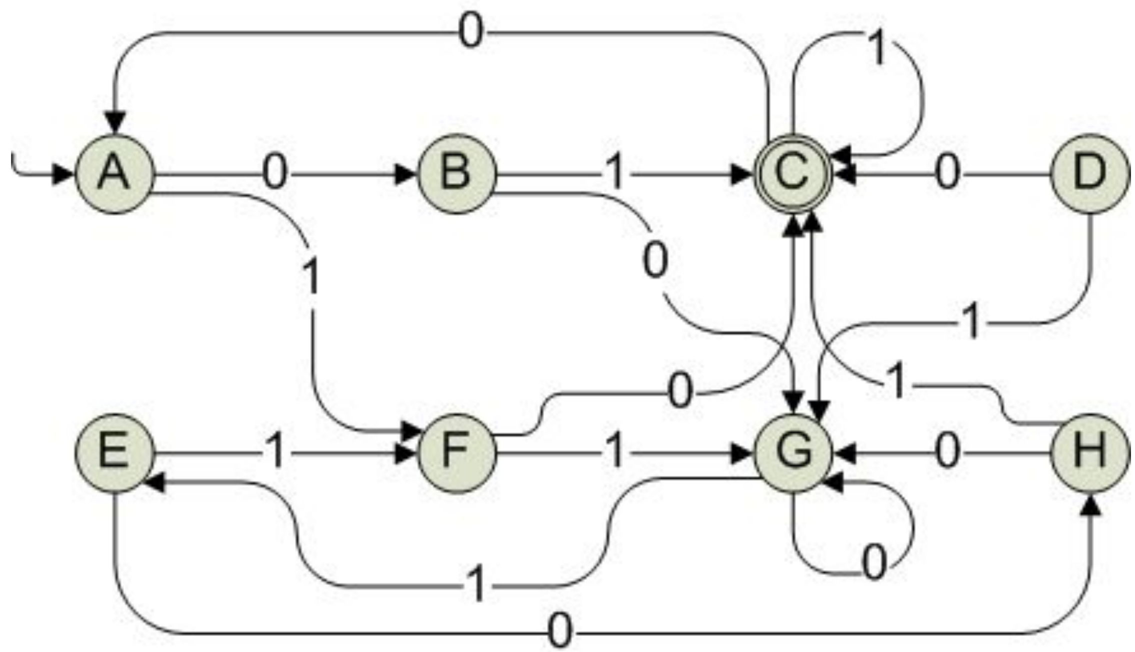


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G		x	x	x		x	
H	x		x	x	x	x	x

Exs:

(h) Continuando para par (A,H):

- SetasA_0:{C}, SetasH_0:{E}
- Novos pares = {(C,E)}
- SetasA_1:{}, SetasH_1: {}
- Novos pares = {}
- Célula (A,H) é marcada

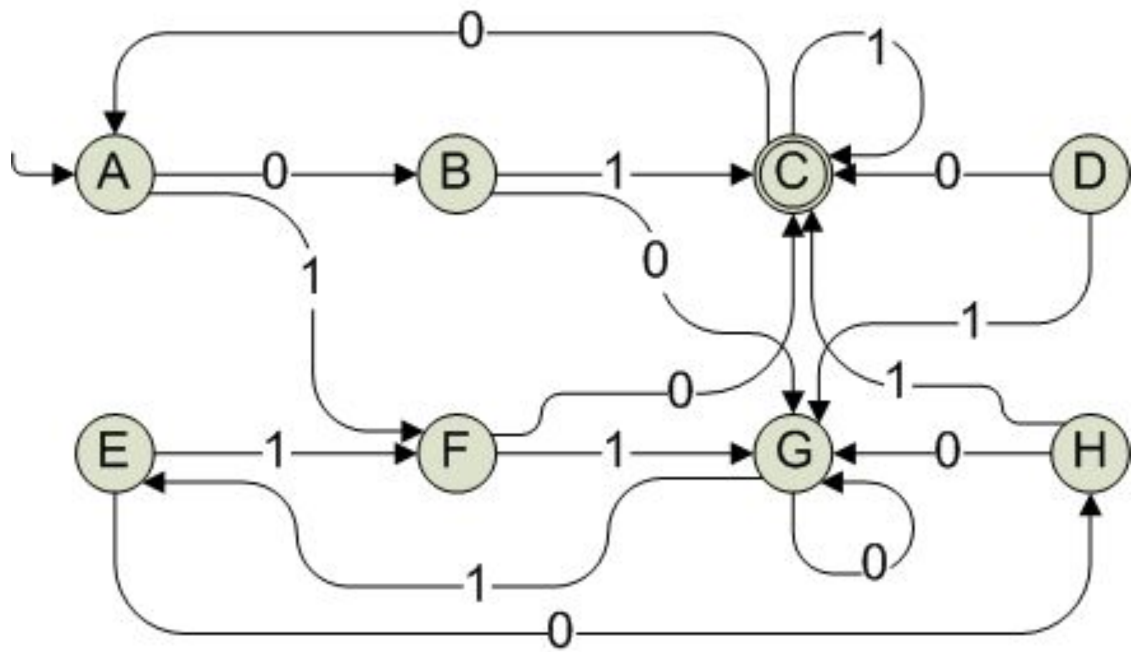


B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G		x	x	x		x	
H	x		x	x	x	x	x
	A	B	C	D	E	F	G

Exs:

(i) Continuando para par (B,D):

- SetasB_0:{A}, SetasD_0:{}
- Novos pares = {}
- SetasB_1:{}, SetasD_1: {}
- Novos pares = {}
- Célula (B,D) é marcada

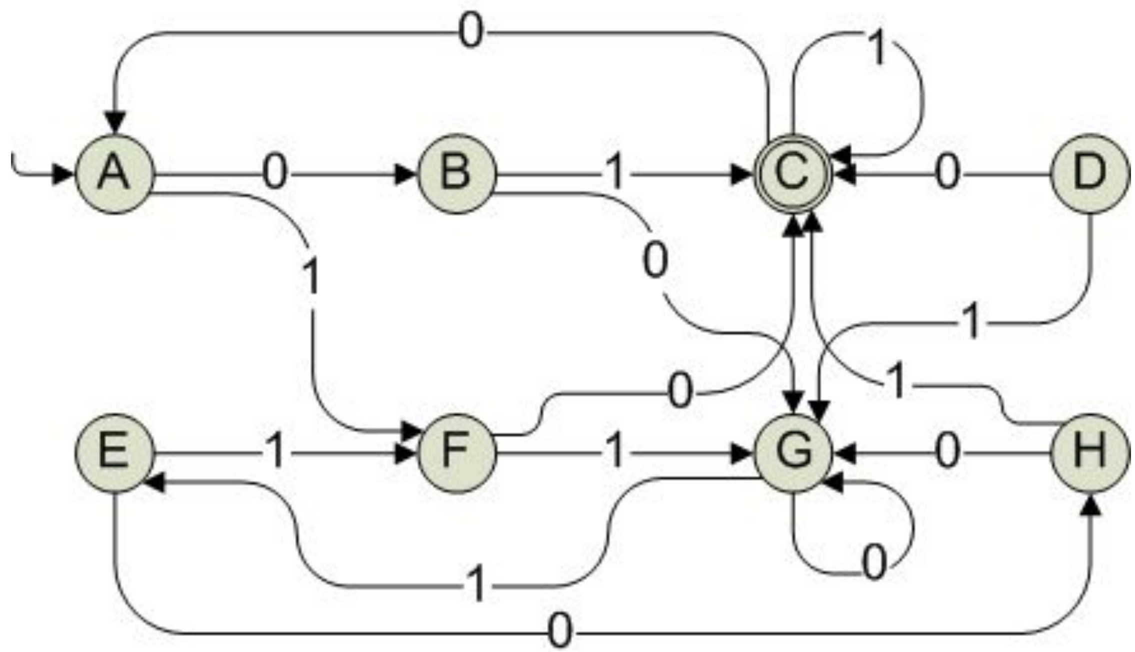


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G		x	x	x		x	
H	x		x	x	x	x	x

Exs:

(i) Continuando para par (B,E):

- SetasB_0:{A}, SetasE_0:{}
- Novos pares = {}
- SetasB_1:{}, SetasE_1: {G}
- Novos pares = {}
- Célula (B,E) é marcada

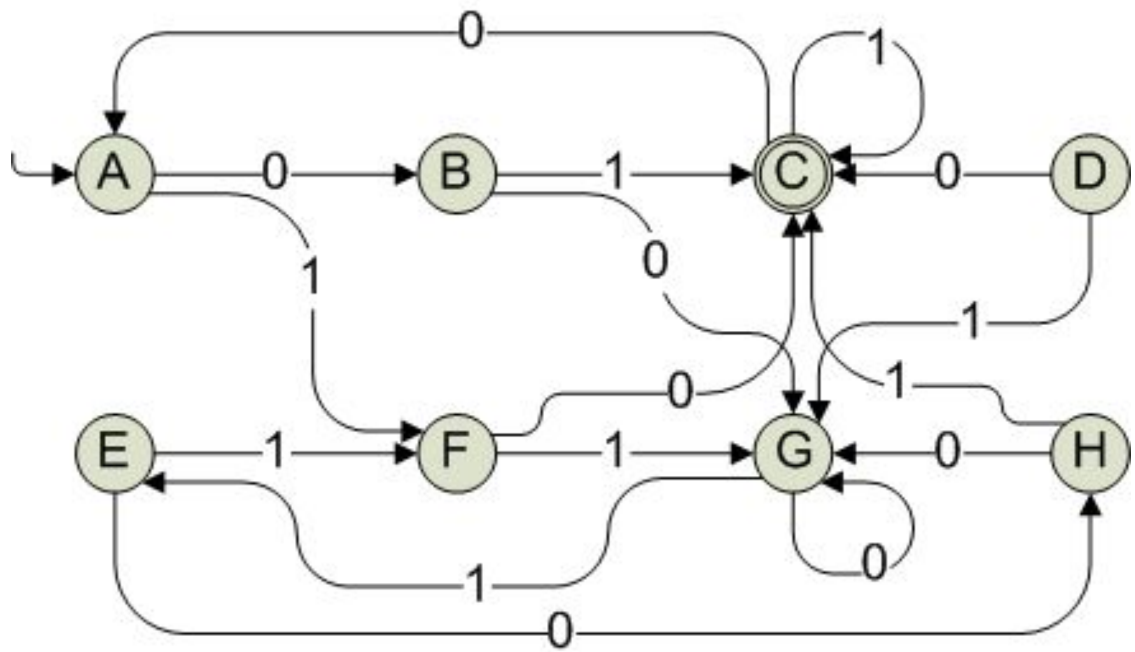


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G		x	x	x		x	
H	x		x	x	x	x	x

Exs:

(j) Continuando para par (B,F):

- SetasB_0:{A}, SetasF_0:{}
- Novos pares = {}
- SetasB_1:{}, SetasF_1: {A,E}
- Novos pares = {}
- Célula (B,F) é marcada

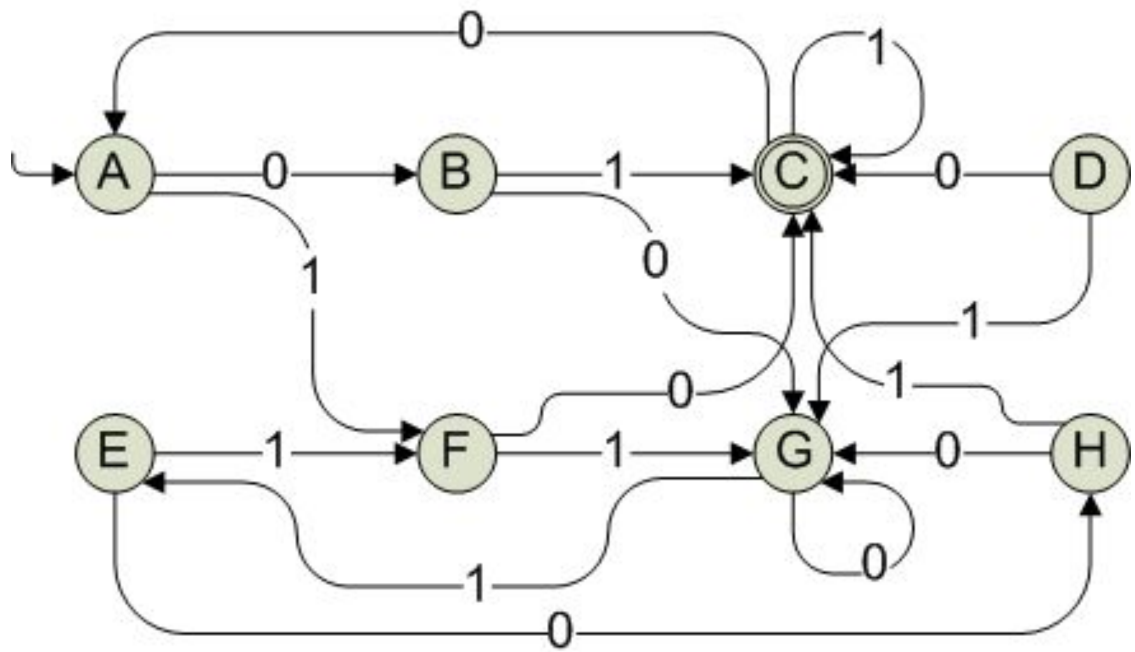


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G	x	x	x	x		x	
H	x		x	x	x	x	x

Exs:

(k) Continuando para par (B,G):

- SetasB_0:{A}, SetasG_0:{B,G,H}
- Novos pares = {(A,B),(A,G),(A,H)}
- SetasB_1:{}, SetasG_1: {D,F}
- Novos pares = {}
- Novo par e célula (B,G) são marcados

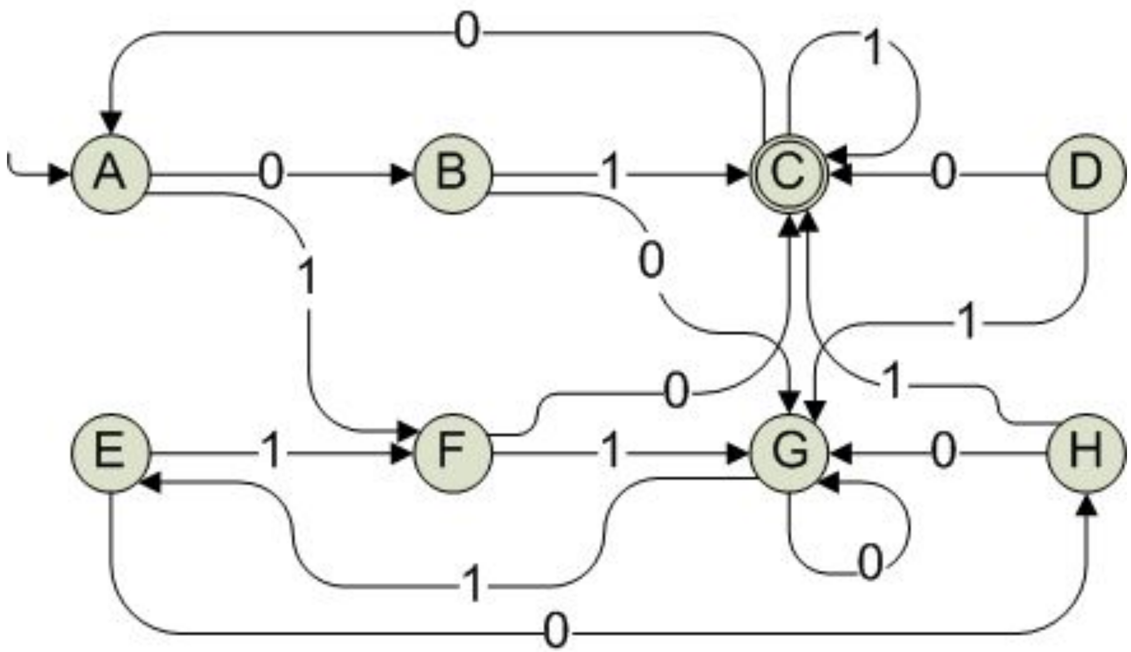


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G	x	x	x	x		x	
H	x		x	x	x	x	x

Exs:

(I) Continuando para par (A,G):

- SetasA_0:{C}, SetasG_0:{B,G,H}
- Novos pares = {(C,B),(C,G),(C,H)}
- SetasA_1:{}, SetasG_1: {D,F}
- Novos pares = {}
- Célula (A,G) é marcada

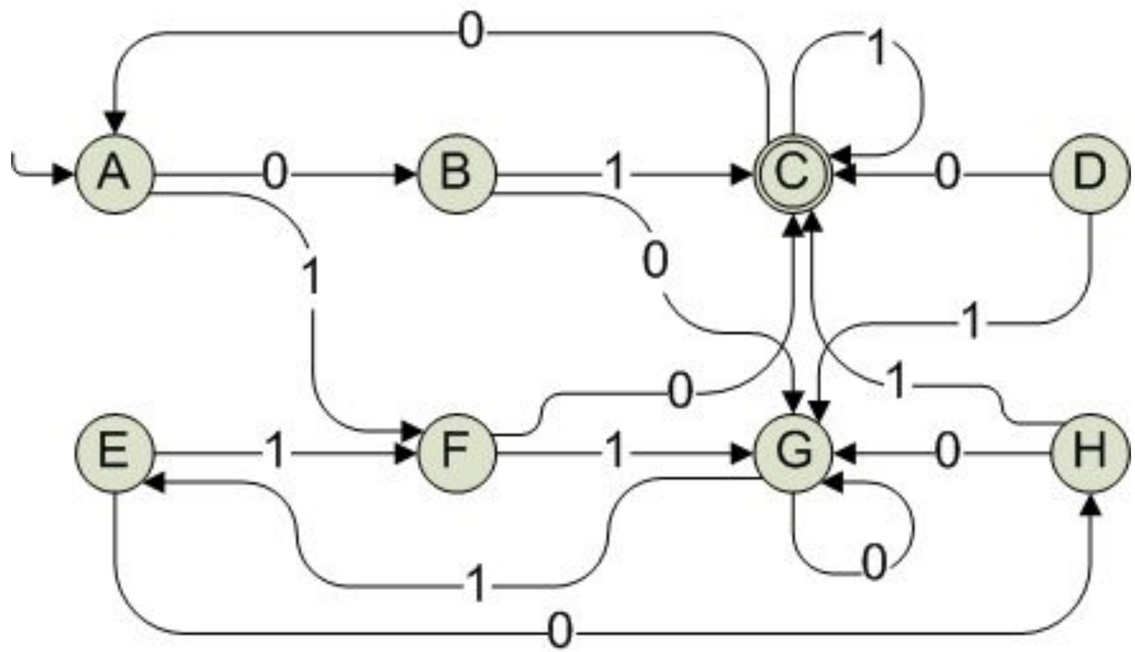


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G	x	x	x	x		x	
H	x		x	x	x	x	x

Exs:

(m) Continuando para par (D,E):

- SetasD_0:{}, SetasE_0:{}
- Novos pares = {}
- SetasD_1:{}, SetasE_1: {G}
- Novos pares = {}
- Células (D,E), (D,G) e (D,H) são marcadas (pois SetasD_0 e SetasD_1 são vazios)

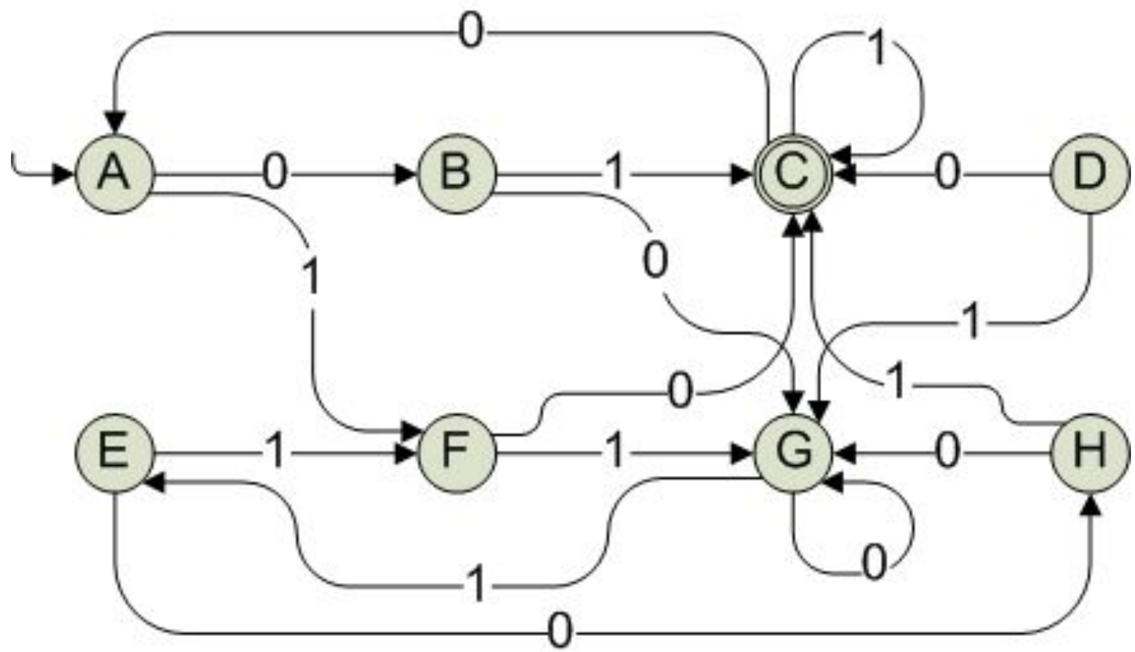


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G	x	x	x	x		x	
H	x		x	x	x	x	x

Exs:

(n) Continuando para par (E,F):

- SetasE_0:{}, SetasF_0:{}
- Novos pares = {}
- SetasE_1:{G}, SetasF_1: {A,G}
- Novos pares = {(A,G)}
- Célula (E,F) é marcada

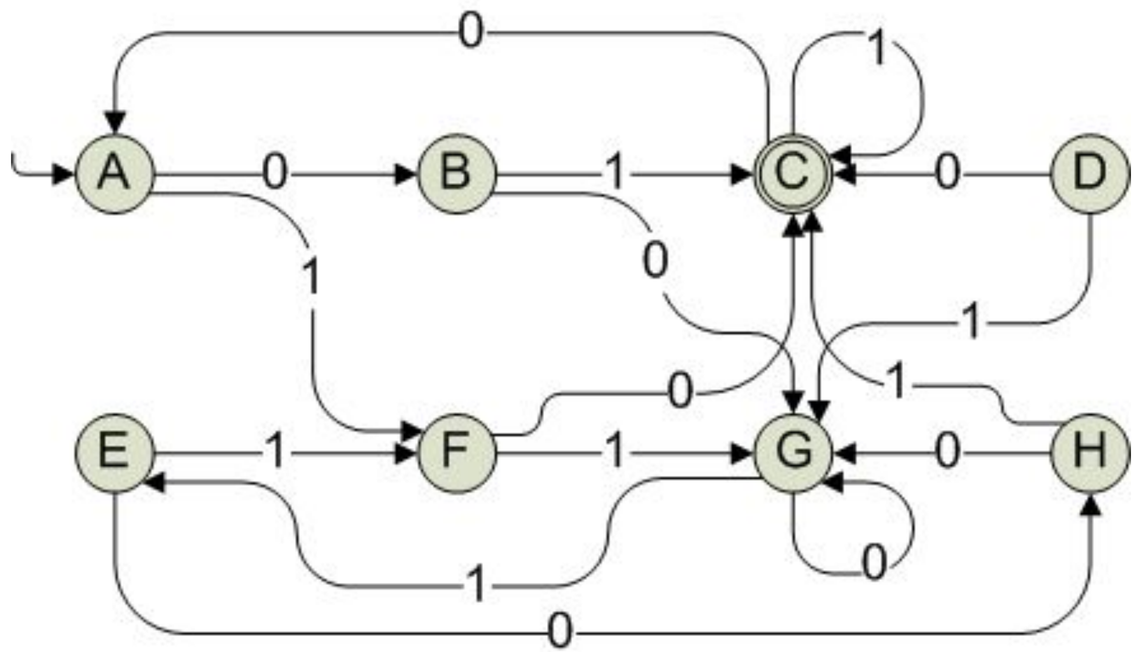


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G	x	x	x	x		x	
H	x		x	x	x	x	x

Exs:

(o) Continuando para par (E,H):

- SetasE_0:{}, SetasH_0:{E}
- Novos pares = {}
- SetasE_1:{G}, SetasH_1: {}
- Novos pares = {}
- Célula (E,H) é marcada

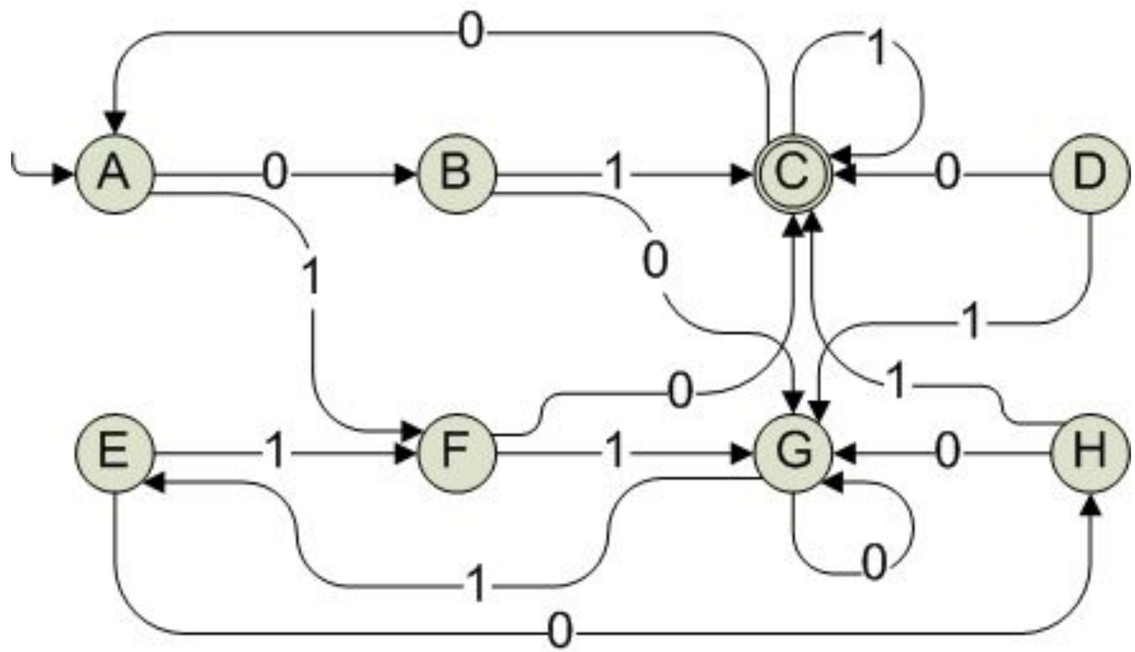


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G	x	x	x	x		x	
H	x		x	x	x	x	x

Exs:

(p) Continuando para par (F,G):

- SetasF_0:{}, SetasG_0:{B,G,H}
- Novos pares = {}
- SetasF_1:{A,E}, SetasG_1: {D,F}
- Novos pares = {(A,D),(A,F),(E,D),(E,F)}
- Célula (F,G) é marcada

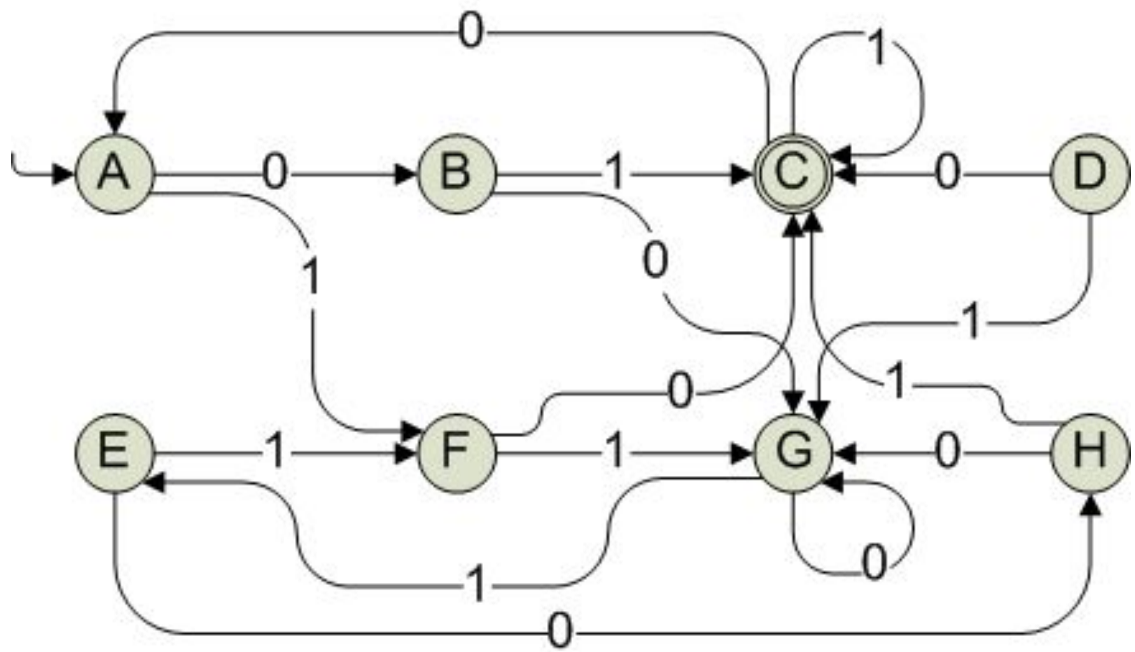


Exs:

(q) Continuando para par (F,H):

- SetasF_0:{}, SetasH_0:{E}
- Novos pares = {}
- SetasF_1:{A,E}, SetasH_1: {}
- Novos pares = {}
- Célula (F,H) é marcada

	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G	x	x	x	x		x	
H	x		x	x	x	x	x

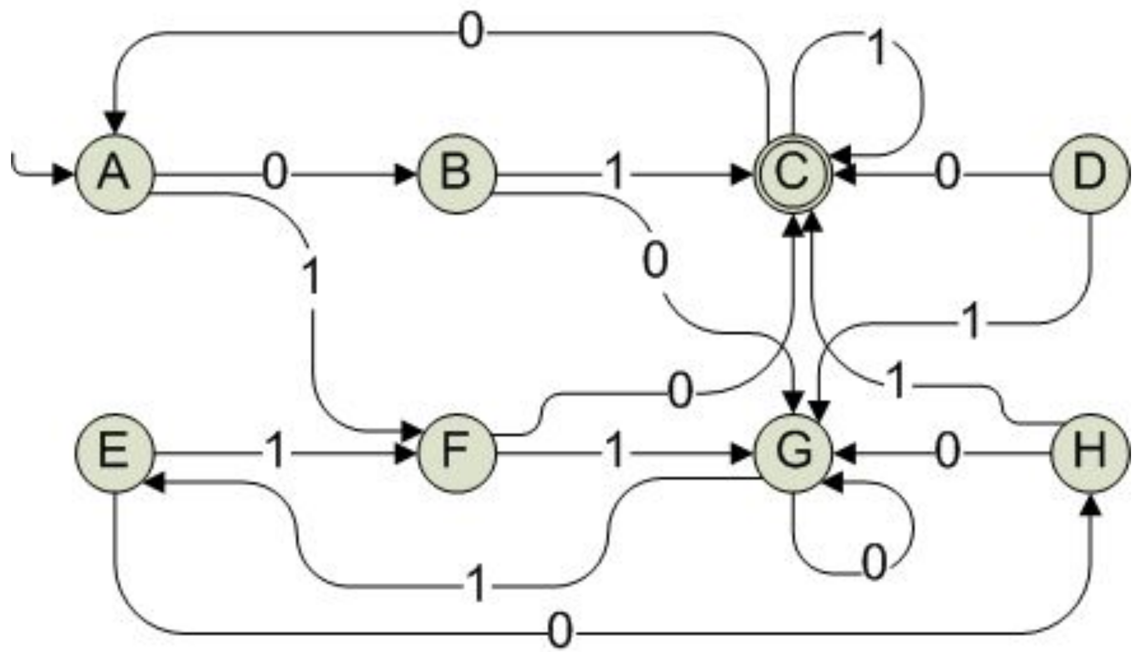


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G	x	x	x	x	x	x	
H	x		x	x	x	x	x

Exs:

(q) Continuando para par (G,H):

- SetasG_0:{B,G,H}, SetasH_0:{E}
- Novos pares = {(B,E),(G,E),(H,E)}
- SetasG_1:{D,F}, SetasH_1: {}
- Novos pares = {}
- Novo par e célula (G,H) são marcados

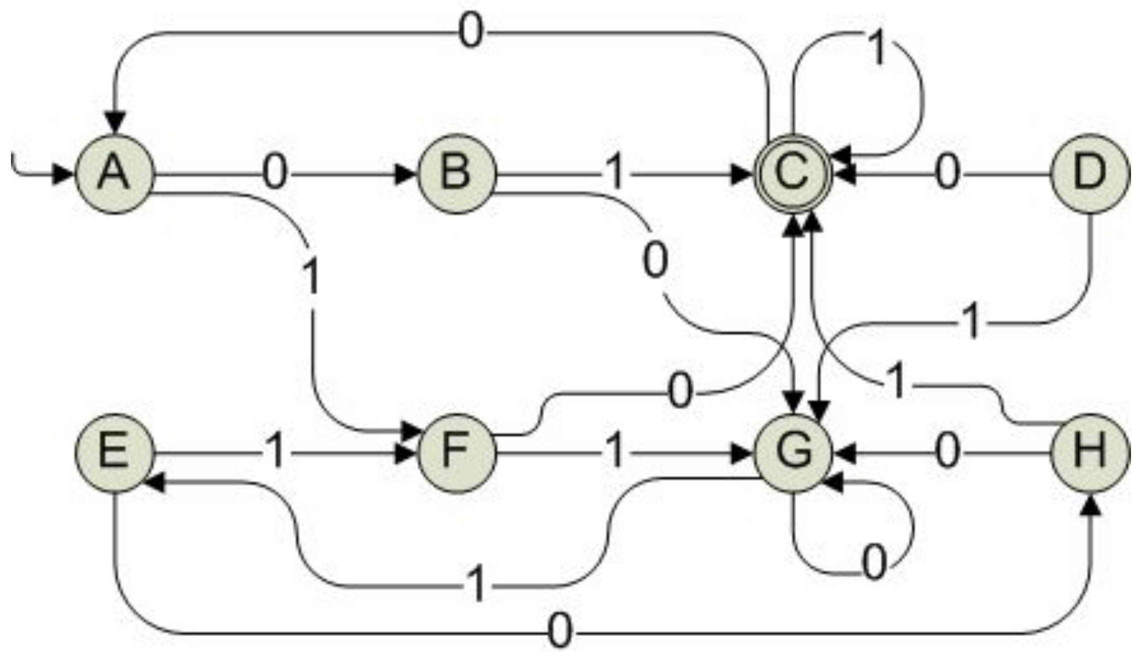


	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G	x	x	x	x	x	x	
H	x		x	x	x	x	x

Exs:

(r) Continuando para par (G,E):

- SetasG_0:{B,G,H}, SetasE_0:{}
- Novos pares = {}
- SetasG_1:{D,F}, SetasE_1: {G}
- Novos pares = {(D,G),(F,G)}
- Célula (G,E) é marcada



	A	B	C	D	E	F	G
B	x						
C	x	x					
D	x	x	x				
E		x	x	x			
F	x	x	x		x		
G	x	x	x	x	x	x	
H	x		x	x	x	x	x

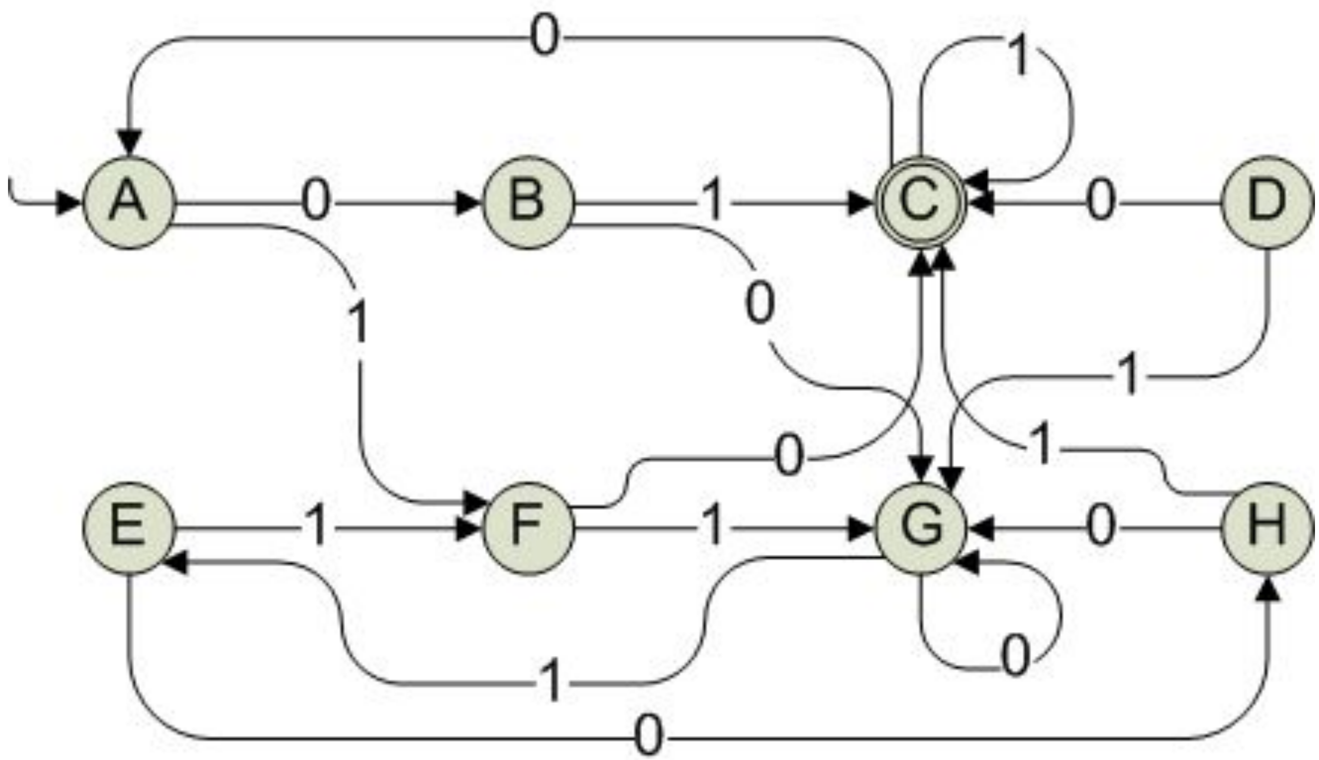
Resultado:

- São pares equivalentes: (A,E), (B,H) e (D,F)

Minimização de DFAs

- Algoritmo em duas etapas:
  - a. Eliminar estados inalcançáveis
    - Reduz o trabalho do algoritmo de preenchimento de tabela
  - b. Particionar os estados restantes em blocos de estados equivalentes
    - Primeiro deve-se identificar os pares equivalentes
    - Depois formar os grupos de estados equivalentes

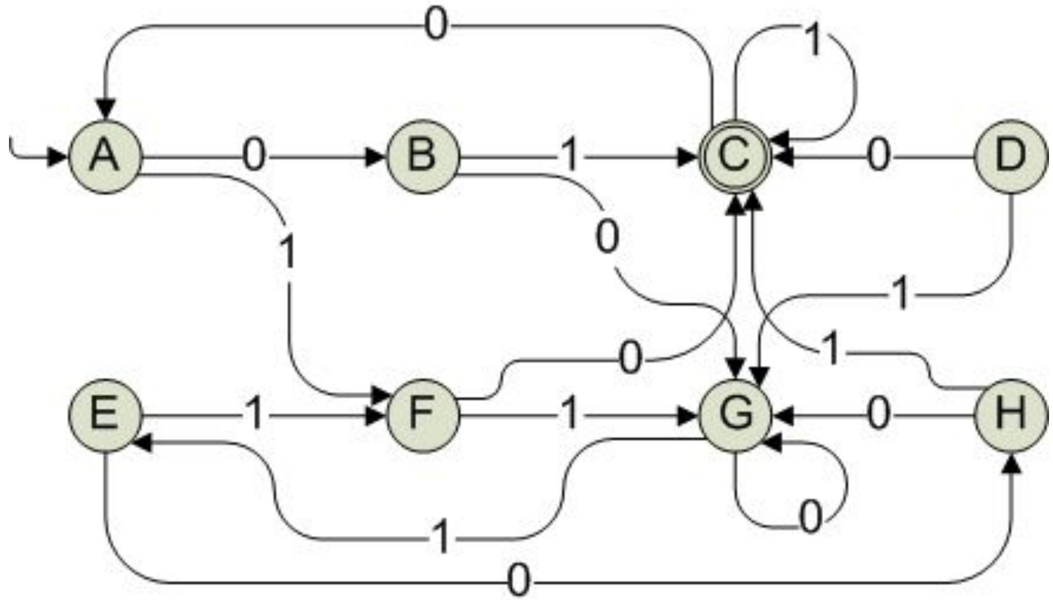
Estados inalcançáveis



Estados alcançáveis devem ter um caminho a partir do estado inicial

Neste exemplo, o estado D é inalcançável

Particionamento em grupos de estados equivalentes



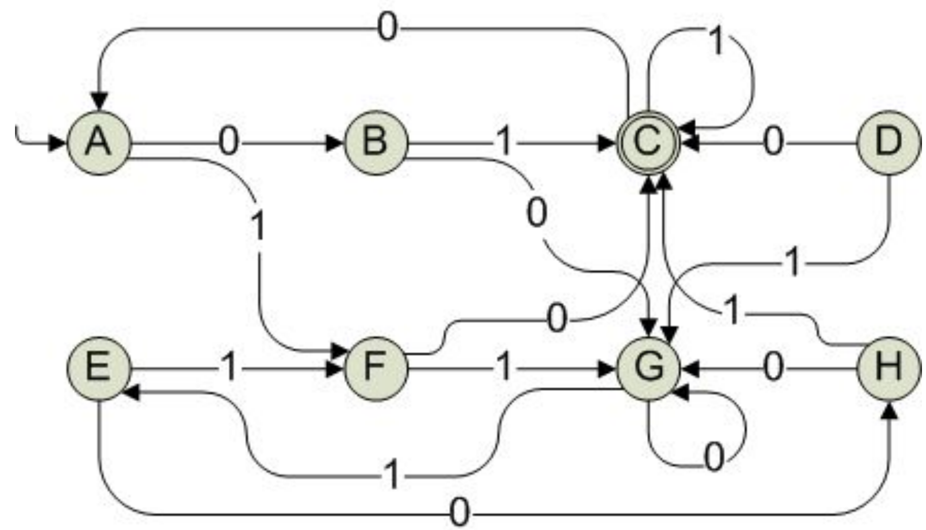
Neste exemplo, são pares equivalentes: (A,E), (B,H) e (D,F)
- Partição: {A,E},{B,H},{D,F},{C},{G}

Importante: deve-se considerar o caráter transitivo da equivalência. Por exemplo, se os pares equivalentes fossem: (A,E), (E,H), (D,F)
- A partição seria: {A,E,H},{D,F},{B},{C},{G}

(Ou seja, A é equivalente a E, E é equivalente a H, portanto A é equivalente a H, e os três formam um único grupo)

Particionamento em grupos de estados equivalentes

Para concluir a minimização, basta definir a nova função de transição



	0	1
{A,E}	{B,H}	{F}
{B,H}	{G}	{C}
{D,F}	{C}	{G}
{C}	{A}	{C}
{G}	{G}	{E}

Para isso, monta-se uma tabela vazia, onde cada estado é um grupo da partição

As transições são definidas como a união das transições no autômato original

Particionamento em grupos de estados equivalentes

	0	1
{A,E}	{B,H}	{F}
{B,H}	{G}	{C}
{D,F}	{C}	{G}
{C}	{A}	{C}
{G}	{G}	{E}

	0	1
{A,E}	{B,H}	{D,F}
{B,H}	{G}	{C}
{D,F}	{C}	{G}
{C}	{A,E}	{C}
{G}	{G}	{A,E}

De {F} para {D,F}

De {A} para {A,E}

De {E} para {A,E}

Agora, basta substituir os valores das células por grupos que representam estados válidos (a primeira coluna da tabela)

Nunca haverá conflito, devido ao algoritmo de preenchimento da tabela

Particionamento em grupos de estados equivalentes

Estados iniciais e de aceitação são os grupos que contém os estados iniciais e de aceitação do DFA original

	0	1
→ {A,E}	{B,H}	{D,F}
{B,H}	{G}	{C}
{D,F}	{C}	{G}
* {C}	{A,E}	{C}
{G}	{G}	{A,E}

Para concluir, renomeie os estados para ficar mais legível

	0	1
→ Q1	Q2	Q3
Q2	Q5	Q4
Q3	Q4	Q5
* Q4	Q1	Q4
Q5	Q5	Q1

Fim

Aula 03 - Minimização de DFAs