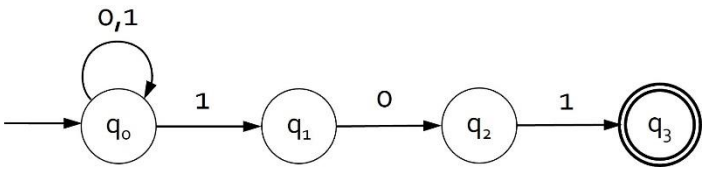
Ćwiczenie 11

# Transformacja NFA na DFA synteza strukturalna

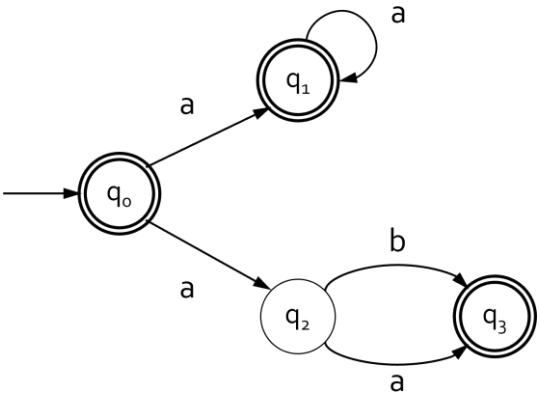
## 1. Wprowadzenie

Niedeterministyczny automat skończony NFA (*Nondeterministic Finite Automaton*) jest modelem matematycznym działania układu fizycznego akceptującego słowa języka formalnego generowanego przez gramatykę bezkontekstową regularną. W automacie niedeterministycznym z każdego stanu może prowadzić zero, jedno bądź wiele przejść skojarzonych z określonym symbolem wejściowym (rysunek 1, 2). Automat niedeterministyczny akceptuje słowo wejściowe oznacza to, że istnieje co najmniej jedna ścieżka (ciąg przejść) prowadzących ze stanu początkowego do stanu końcowego (akceptującego). Automat niedeterministyczny nie akceptuje słowa wejściowego tylko wtedy, gdy żadna ze ścieżek (rozgałęzień) nie prowadzi do stanu akceptującego – wszystkie ścieżki prowadzą do stanów nieakceptujących (rysunek 3). Jeżeli w NFA istnieje stan z więcej niż jednym przejściem skojarzonym z określonym symbolem wejściowym mówimy o rozgałęzieniu (rysunek 3). Jeżeli wchodzimy do stanu, w którym nie ma przejścia, wtedy rozgałęzienie zanika (umiera).
Przykład automatu niedeterministycznego akceptującego wszystkie słowa binarne kończące się 101:




Rysunek 1.

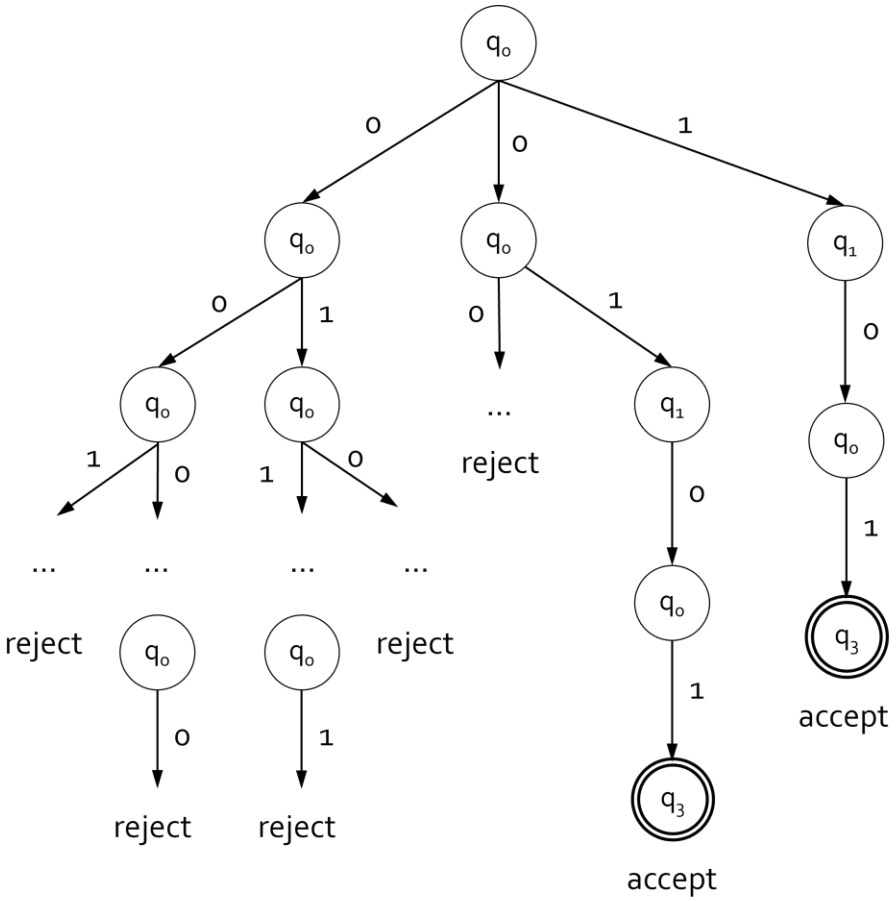
Przykład automatu deterministycznego dla wyrażenia regularnego $a^* + (ab)^*$:




Rysunek 2.

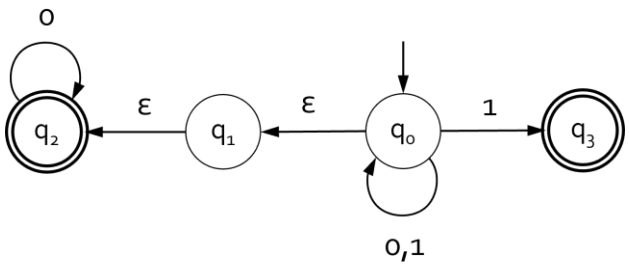
Automaty niedeterministyczne pozwalają modelować działanie złożonych procesów, których modelowanie przy użyciu automatów deterministycznych byłoby pracochłonne i nieefektywne ze względu na konieczność spełnienia ograniczeń automatów deterministycznych.

O niedeterminizmie, w przypadku działania NFA można myśleć jak o realizacji procesów równoległych. Graficznie, niedeterminizm można przedstawić jako drzewo z rozgałęzieniami reprezentującymi przejścia ze stanów pod wpływem jednakowych symboli wejściowych.



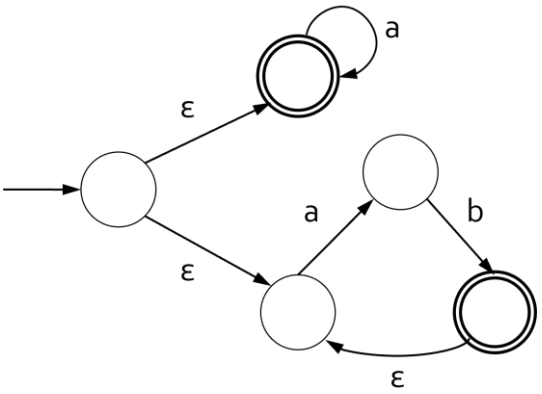

Rysunek 3.

W automatach niederministycznych możliwe są również przejścia (powodujące zmianę stanu) bez wprowadzania symbolu wejściowego. Są to tak zwane $\varepsilon$-przejścia (rysunek 4).




Rysunek 4. NFA akceptujący słowa binarne kończące się 1 lub zawierające dowolną liczbę 0.

Przykład NFA dla wyrażenia regularnego: $a^* + (ab)^*$:




Rysunek 5.

W ujęciu formalnym automat skończony to piątka < $Z$, $Q$, $\Psi$, $q_0$, $F$ >, gdzie:

$Z$ – alfabet wejściowy (skończony zbiór symboli wejściowych),

$Z = \{ z_0, z_1, z_2, ..., z_{n-1} \}$,

$Q$ – skończony zbiór stanów wewnętrznych automatu

$Q = \{ q_0, q_1, q_2, ..., q_{m-1} \}$,

$\Psi$ – funkcja przejść automatu

$\Psi( q(t), z(t) ) = \{ q_{i0}(t+1), q_{i1}(t+1), ..., q_{in-1}(t+1) \}$,

$q_0$ – stan początkowy, $q_0 \in Q$,

$F$ – zbiór stanów końcowych automatu, $F \subseteq Q$.

O ile modelowanie działania układów (przebiegu procesów) przy użyciu automatów niedeterministycznych jest prostsze, a działanie układu jest łatwiejsze do analizy, o tyle synteza układów realizujących określone zadania na podstawie opisu w postaci NFA (z lub bez $\varepsilon$-przejść), jest niemożliwe. Należy pamiętać, że rozumienie działania układu (modelowanego NFA) jako równoległej realizacji podzadań prowadzących (co najmniej w jednym przypadku) do wykonania zadania, jest nierealizowalna w praktyce ze względu na brak możliwości tworzenia (rozgałęzienie) i redukcji kopii fizycznego układu w trakcie jego działania.

Aby uzyskać możliwość syntezy fizycznego układu należy wykonać konwersję automatu niedeterministycznego na automat deterministyczny (NFA → DFA), bazując na twierdzeniu:

> Dla każdego niedeterministycznego automatu skończonego można wskazać ekwiwalentny deterministyczny automat skończony. (1)

Dowód twierdzenia (1) można znaleźć w [1].

## 2. Algorytm transformacji NFA → DFA

Bazując na dowodzie twierdzenia (1) możemy określić sposób transformacji NFA → DFA. Możemy go zapisać w postaci krótkiego algorytmu.

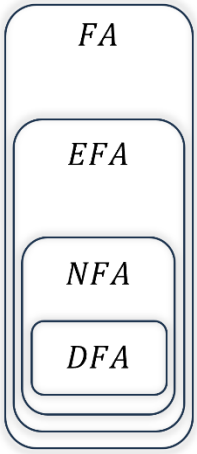
1. Na podstawie liczby stanów NFA określ liczbę możliwych stanów DFA.

Należy pamiętać, że każdy stan DFA reprezentuje określony podzbiór stanów NFA. Stąd dla $n$-stanowego NFA, maksymalna liczba stanów DFA to $2^n$.

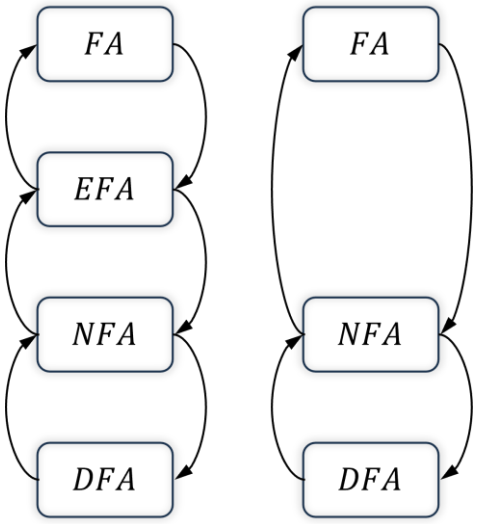
2. Określ stan początkowy i stany końcowe DFA, uwzględniając $\varepsilon$-przejścia.

   Stan początkowy może być zbiorem jedno- bądź wieloelementowym. Zależnie od przyjętej terminologii podziału automatów skończonych i sposobu konwersji. Według terminologii Stoughtona (rysunek 6), stan początkowy jest zbiorem jednoelementowym ze względu na określony sposób konwersji (rysunek 7). Według terminologii proponowanej przez Sipsera, gdzie pomijana jest konwersja z $NFA$ z ε-przejściami (*nondeterministic finite automaton with* ε-*moves*) na $NFA$ bez ε-przejść, a realizowana jest konwersja $FA \rightarrow NFA$ z uwzględnieniem ε-przejść, stan początkowy może być zbiorem stanów osiąganych z wejścia z uwzględnieniem ε-przejść. W ćwiczeniu przyjmujemy terminologię Sipsera.

3. W NFA określ przejścia z każdego ze stanów dla wszystkich symboli alfabetu uwzględniając przejścia ε.
4. Zminimalizuj liczbę stanów DFA, czyli usuń stany bez wejść.




Rysunek 6. Podklasy automatów skończonych wg Stoughtona

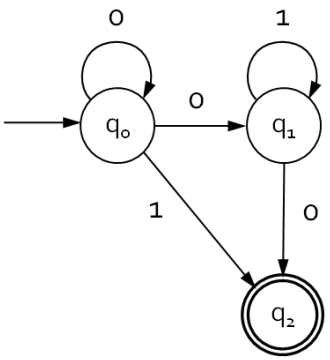



Rysunek 7. Konwersja automatów skończonych wg Stoughtona i Sipsera

EFA (*Empty-string finite Automaton*) to automat skończenie stanowy (FA), w którym etykieta każdego przejścia ma długość 0 lub 1.
NFA – automat, w którym etykieta każdego przejścia ma długość 1.
DFA – automat, w którym etykieta każdego przejścia ma długość 1 i każdy symbol alfabetu opisuje jedno przejście ze każdego stanu.

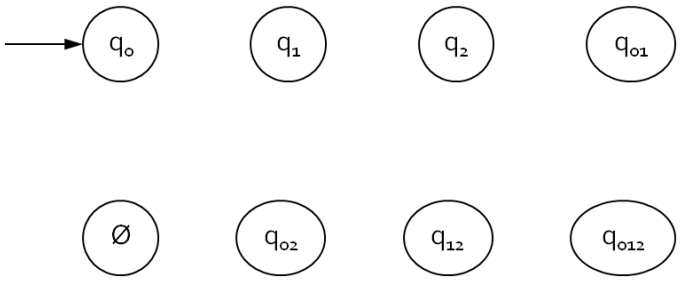
### 3. Przebieg transformacji NFA → DFA

Przykład 1
NFA zadane w postaci grafu bez ε-przejść.

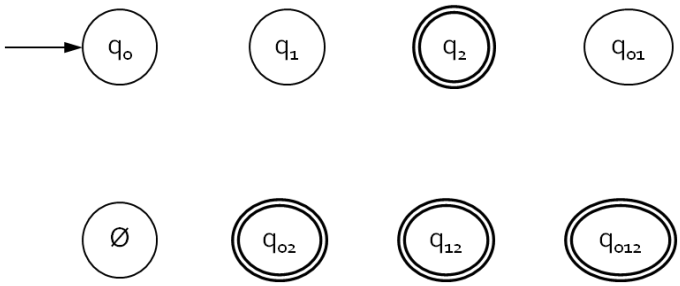



Postępujemy zgodnie z procesem transformacji

1. Określamy liczbę możliwych stanów DFA
NFA posiada trzy stany, $N = \{q_0, q_1, q_2\}$. DFA może mieć co najwyżej osiem stanów. Tworzymy osiem stanów, po jednym dla każdego podzbioru z $N$. Stany DFA oznaczamy (numerujemy) zgodnie porządkiem podzbiorów $D$. Powstaje zbiór stanów $D = \{Ø, q_0, q_1, q_2, q_{01}, q_{02}, q_{12}, q_{012}\}$.

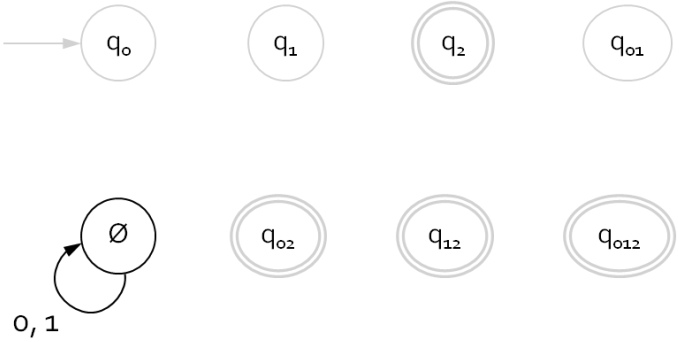


2. Określamy stany początkowe i końcowe DFA. Stan początkowy $E\{q_0\}$ to zbiór stanów osiągalnych z $q_0$ uwzględniając przejścia ε oraz stan $q_0$. Zatem $E\{q_0\} = \{q_0\}$, ponieważ nie występują w NFA przejścia ε. Stany końcowe, to wszystkie stany zawierające stan końcowy NFA. Stanami końcowymi są: $q_2$, $q_{02}$ ,$q_{12}$ i $q_{012}$.

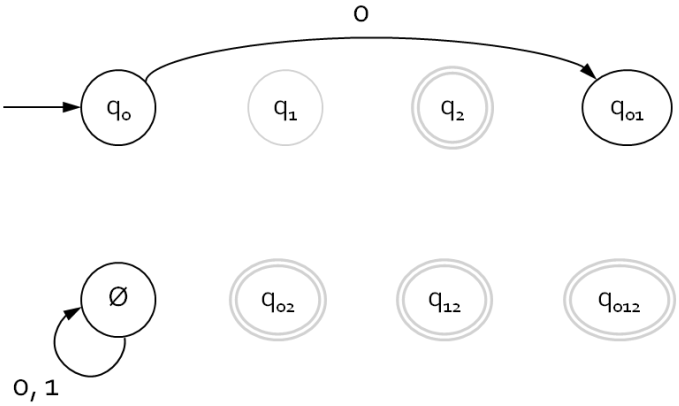



3. W NFA określamy wszystkie przejścia z każdego ze stanów dla wszystkich symboli alfabetu wejściowego. Tworzymy funkcję przejść DFA.
Kroki tego procesu zostały omówione, a ich efekty zilustrowane odpowiednimi rysunkami.

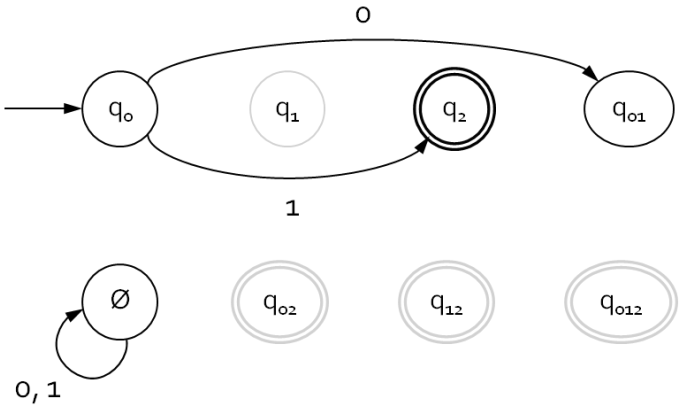
krok 1: Rozpoczynamy od stanu Ø, do którego trafiają wszystkie nieokreślone w NFA przejścia.



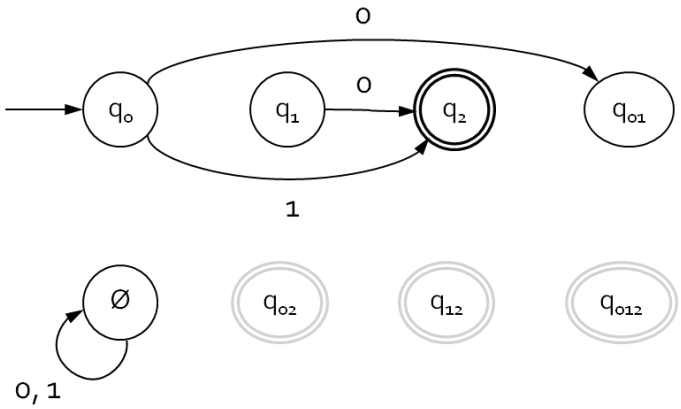

krok 2: Ze stanu $q_0$ przechodzimy do stanu $q_{01}$ pod wpływem symbolu 0, ponieważ w NFA ze stanu $q_0$ następuje przejście do stanów $q_0$ i $q_1$ oraz dalej, z żadnego ze stanów $q_0$, $q_1$ nie ma przejść ε. Ponieważ w przykładzie nie występują przejścia ε, fakt braku wykonania takiego przejścia w opisach kolejnych kroków będzie pomijany.



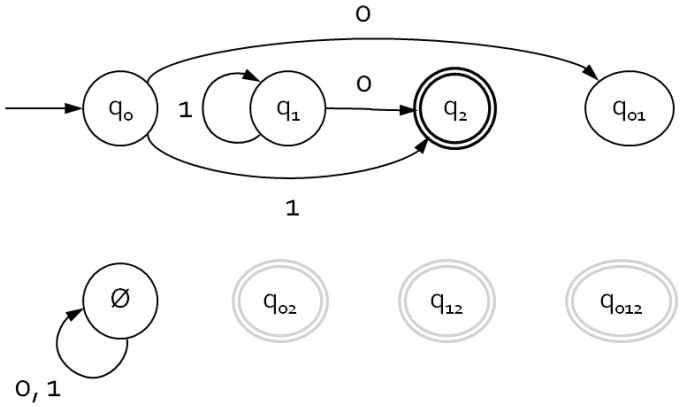
krok 3: Ze stanu $q_0$ przechodzimy do stanu $q_2$ pod wpływem symbolu 1, ponieważ w NFA ze stanu $q_0$ istnieje jedynie przejście do stanu $q_2$.



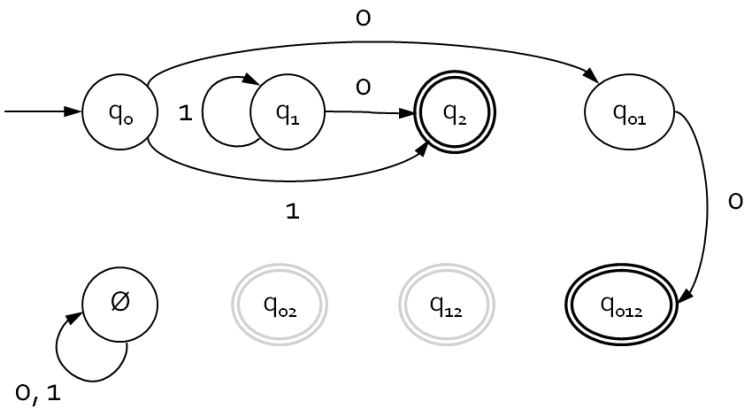

krok 4: Ze stanu $q_1$ przechodzimy do stanu $q_2$ pod wpływem symbolu 0, ponieważ w NFA ze stanu $q_1$ istnieje jedynie przejście do stanu $q_2$.



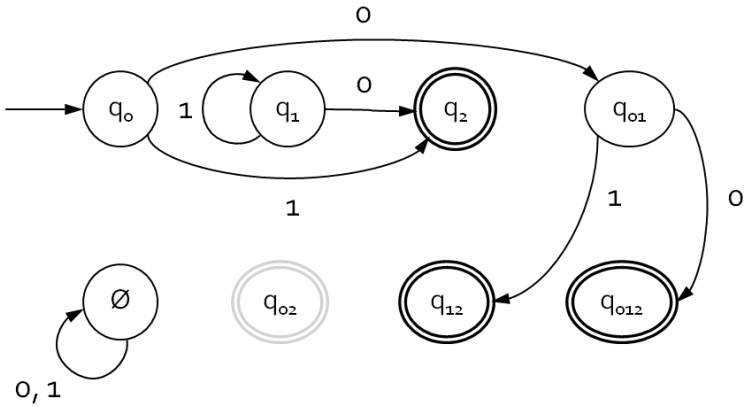

krok 5: Ze stanu $q_1$ przechodzimy do stanu $q_1$ pod wpływem symbolu 1, ponieważ w NFA ze stanu $q_1$ istnieje jedynie przejście do stanu $q_1$.



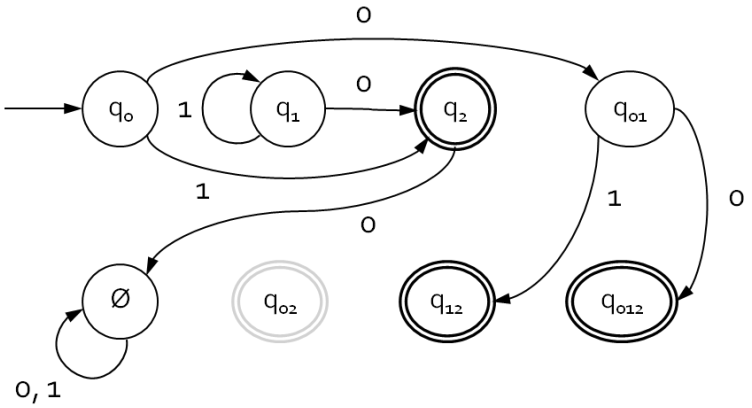
krok 6: Stan $q_{012}$ reprezentuje połączone stany $q_0$, $q_1$ i $q_2$. W związku z tym, ponieważ ze stanu $q_0$ przez 0 przechodzimy do $q_0$ i jednocześnie do stanu $q_1$ oraz ze stanu $q_1$ przez 0 przechodzimy do stanu $q_2$, to ze stanu $q_{01}$ przechodzimy do stanu $q_{012}$ pod wpływem symbolu 0.



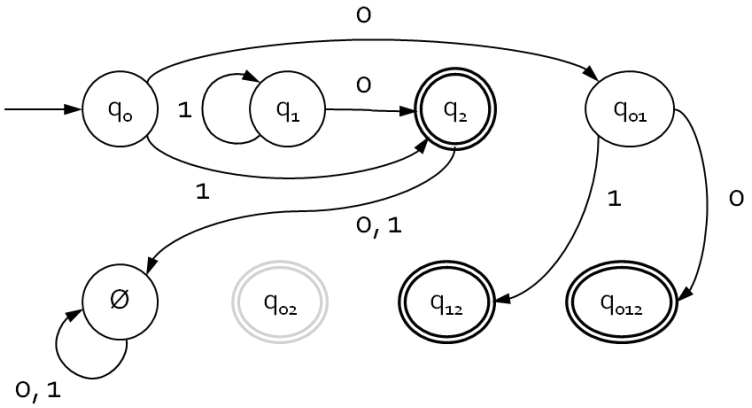

krok 7: Podobnie rzecz ma się przy przejściu ze stanu $q_{01}$ do stanu $q_{12}$ pod wpływem symbolu 1.



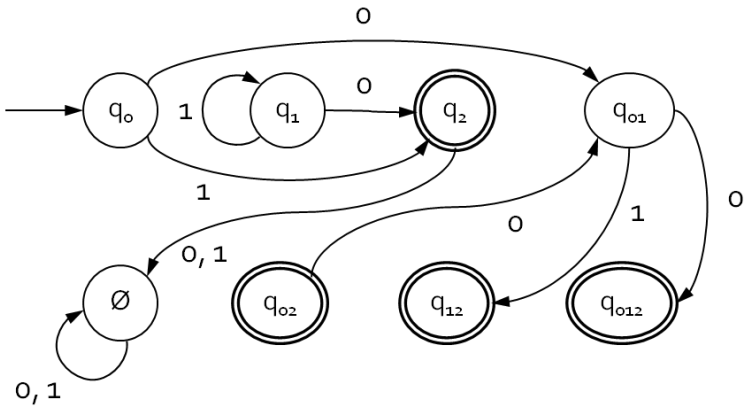

krok 8: Ze stanu $q_2$ w NFA nie występuje przejście pod wpływem 0, co powoduje w DFA przejście do stanu Ø.



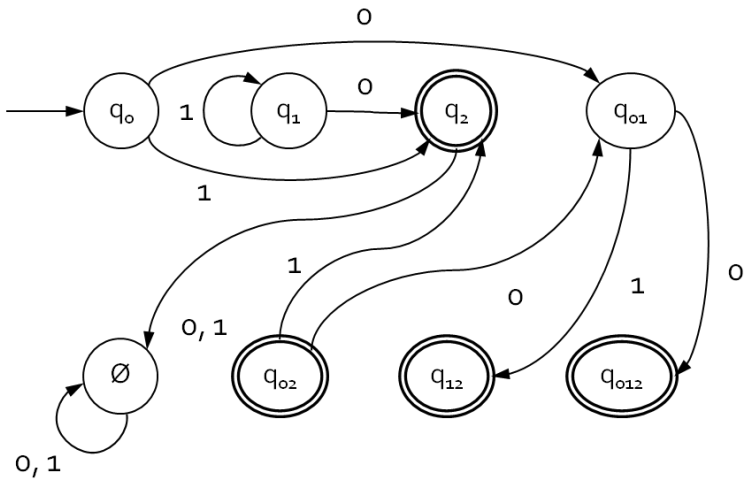
krok 9: Ze stanu $q_2$ w NFA nie występuje również przejście pod wpływem 1, co powoduje w DFA przejście do stanu Ø.



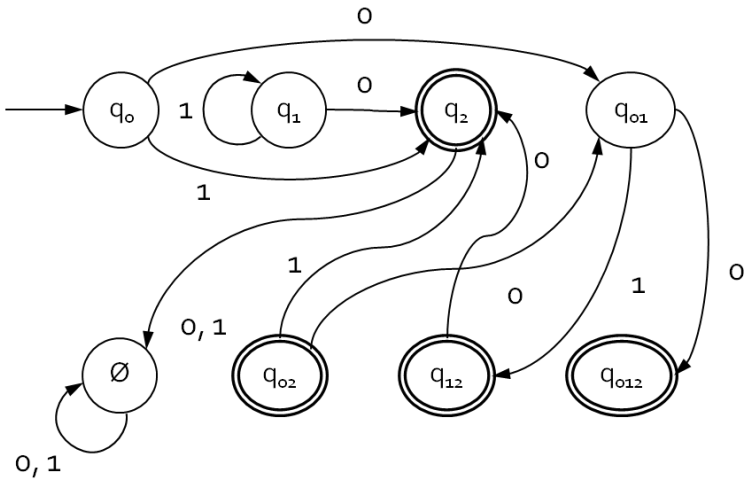

krok 10: Przejście ze stanu $q_{02}$ do stanu $q_{01}$ następuje z podobnych powodów co w krokach 6 i 7.



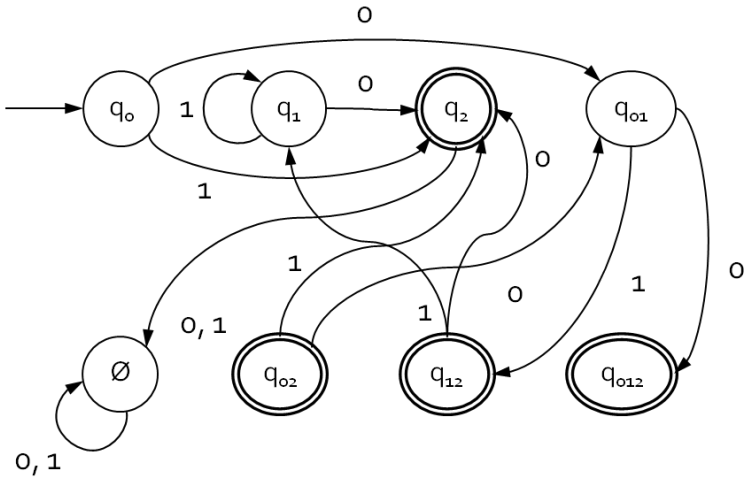

krok 11: Przejście ze stanu $q_{02}$ do stanu $q_2$ pod wpływem 1.



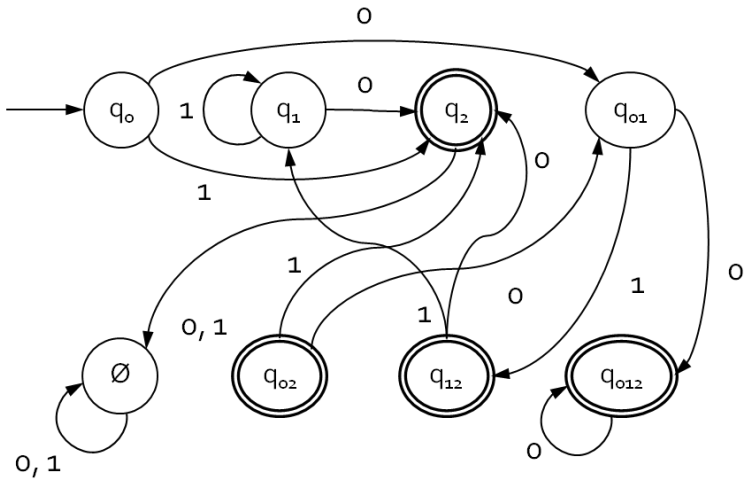
krok 12: Przejście ze stanu $q_{12}$ do stanu $q_2$ pod wpływem 0.



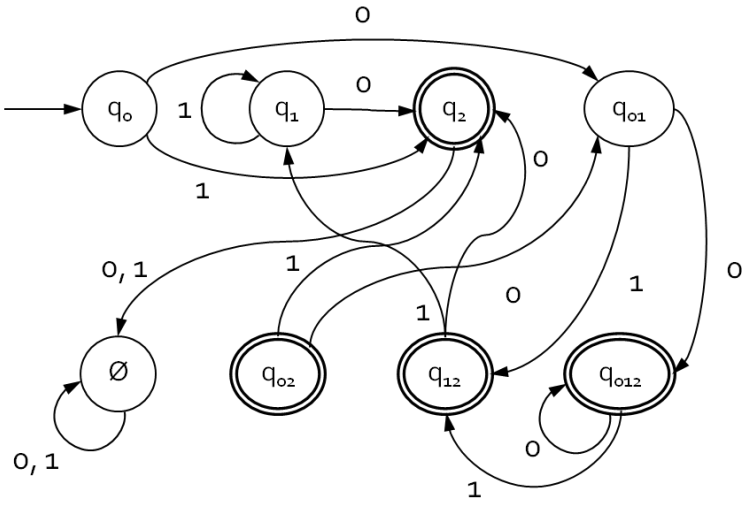

krok 13: Przejście ze stanu $q_{12}$ do stanu $q_1$ pod wpływem 1.




krok 14: Przejście ze stanu $q_{012}$ do stanu $q_{012}$ pod wpływem 0. Sytuacja podobna do tej z kroku 6.

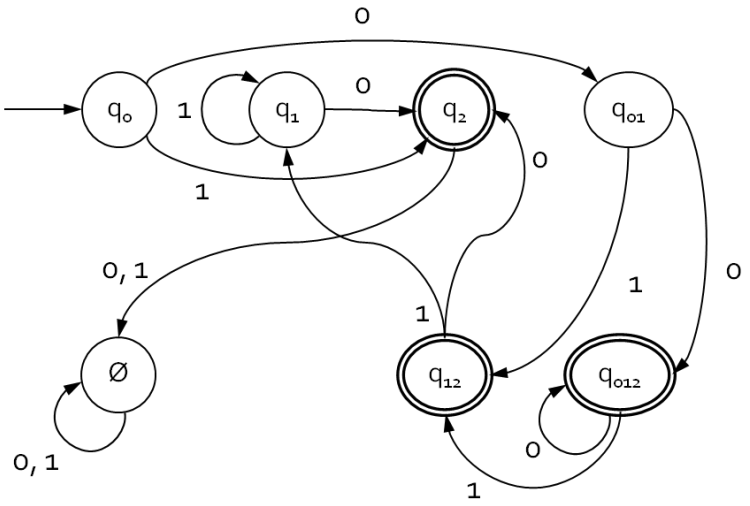


krok 15: Przejście ze stanu $q_{012}$ do stanu $q_{12}$ pod wpływem 1.

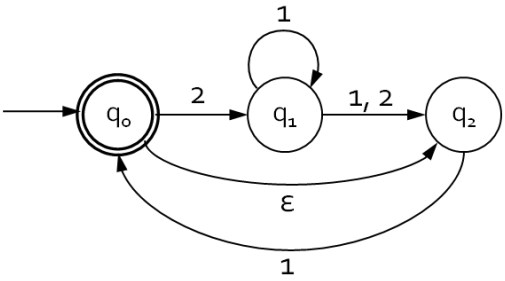



4. Minimalizujemy liczbę stanów DFA, czyli usuwamy stany nieosiągalne.

Jedynym nieosiągalnym stanem w DFA jest stan $q_{02}$.
Ostateczna postać DFA wygląda następująco.

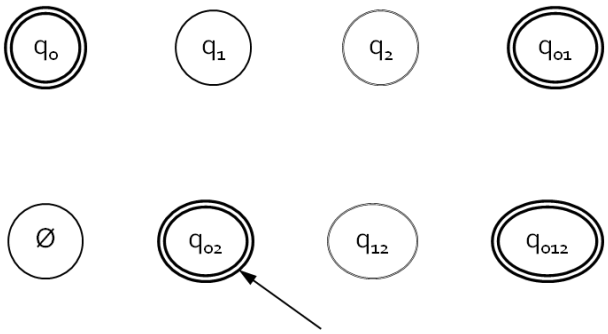


Przykład 2
NFA zadane w postaci grafu z przejściami ε.



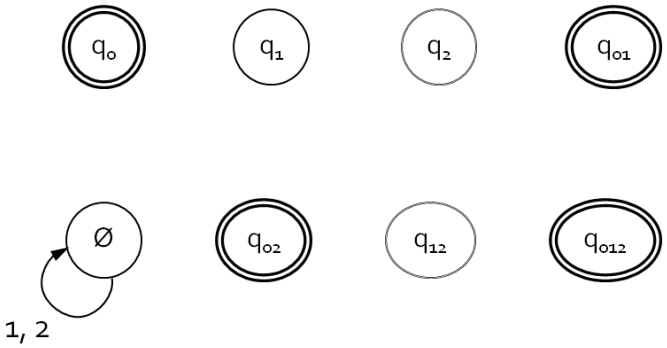

1. Określamy liczbę możliwych stanów DFA
NFA posiada trzy stany, $N = \{q_0, q_1, q_2\}$. DFA może mieć co najwyżej osiem stanów. Tworzymy osiem stanów, po jednym dla każdego podzbioru z $N$. Stany DFA oznaczamy (numerujemy) zgodnie z porządkiem podzbiorów $D$. Powstaje zbiór stanów $D = \{Ø, q_0, q_1, q_2, q_{01}, q_{02}, q_{12}, q_{012}\}$.

2. Określamy stany początkowe i końcowe DFA. Stan początkowy $E\{q_0\}$ to zbiór stanów osiągalnych z $q_0$ uwzględniając przejścia ε oraz stan $q_0$. Zatem $E\{q_0\} = \{q_0, q_2\}$, ponieważ występuje przejście ε z $q_0$ do $q_2$. Stany końcowe, to wszystkie stany zawierające stan końcowy NFA. Stanami końcowymi są: $q_0$, $q_{01}$ ,$q_{02}$ i $q_{012}$. Stanem początkowym w DFA będzie stan $q_{02}$.

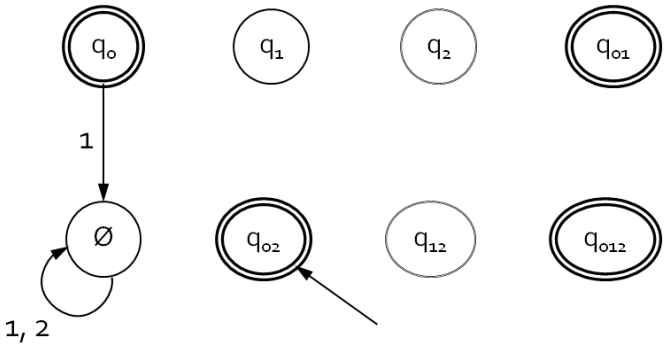



3. W NFA określamy wszystkie przejścia z każdego ze stanów dla wszystkich symboli alfabetu wejściowego. Tworzymy funkcję przejść DFA.
Kroki tego procesu zostały omówione, a ich efekty zilustrowane odpowiednimi rysunkami.

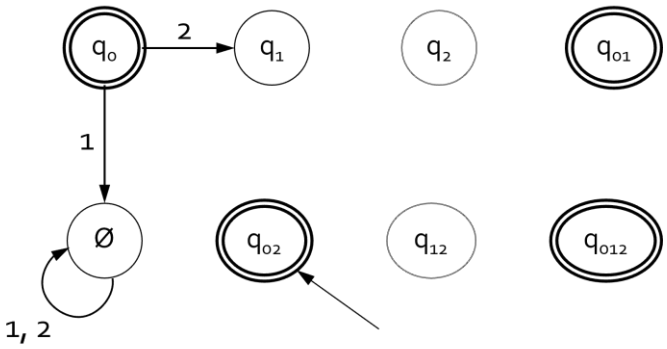
krok 1: Rozpoczynamy od stanu Ø, do którego trafiają wszystkie nieokreślone w NFA przejścia.



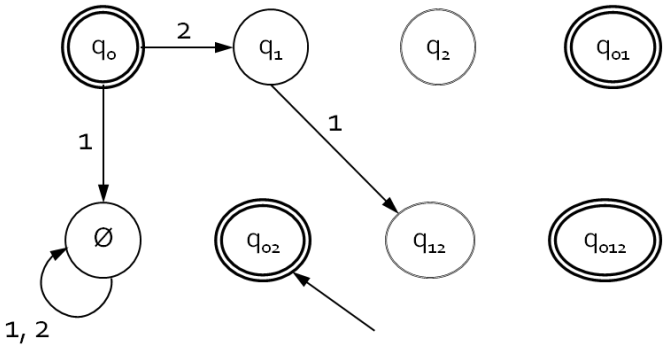
krok 2: Nie istnieje przejście ze stanu $q_0$ pod wpływem symbolu 1. Zatem idziemy do stanu Ø.



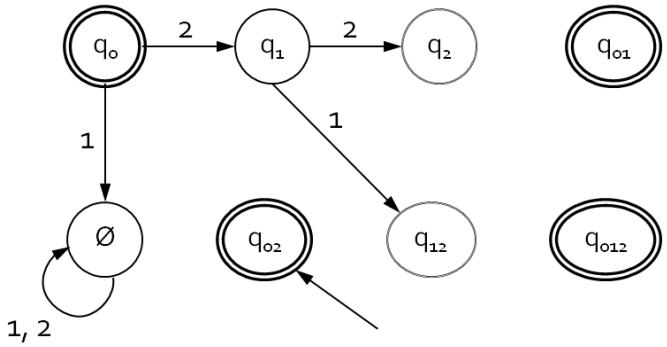

krok 3: Ze stanu $q_0$ pod wpływem symbolu 2 przechodzimy do stanu $q_1$. A ponieważ nie istnieje ścieżka prowadząca przez krawędzie oznaczone ε, $q_1$ jest jedynym stanem, do którego możemy w ten sposób przejść. Należy pamiętać, że podążanie pustymi przejściami ε występuje tylko po wprowadzeniu symbolu wejściowego. Tu taka sytuacja nie występuje.



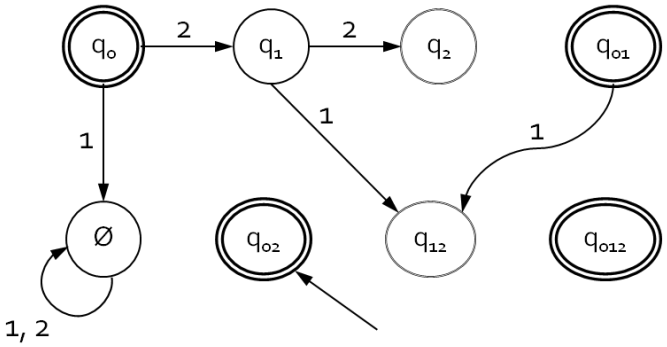

krok 4: Ze stanu $q_1$ pod wpływem symbolu 1 przechodzimy do stanu $q_{12}$ ponieważ w NFA ze stanu $q_1$ pod wpływem 1 przechodzimy do obu stanów: $q_1$ i $q_2$. Na tym kończymy przejścia ponieważ nie istnieje ścieżka prowadząca przez krawędzie oznaczone ε.



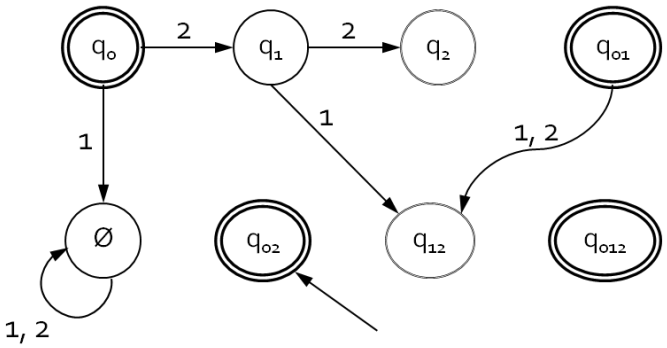
krok 5: Ze stanu $q_1$ pod wpływem 2 przechodzimy do stanu $q_2$. Brak przejść ε.



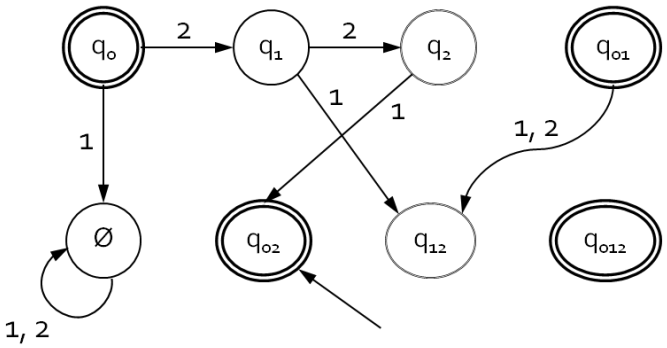

krok 6: Ze stanu $q_{01}$ pod wpływem 1 przechodzimy do stanu $q_{12}$. Brak przejść ε.



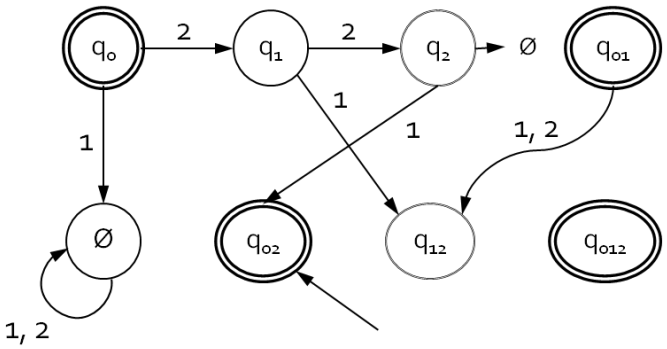

krok 7: Ze stanu $q_{01}$ pod wpływem 2 również przechodzimy do stanu $q_{12}$. Brak przejść ε.



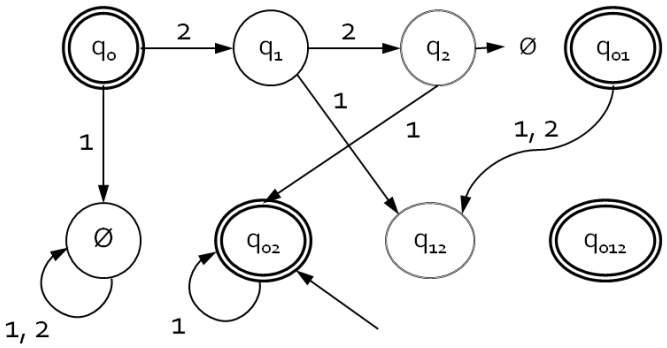
krok 8: Ze stanu $q_2$ pod wpływem 1 idziemy do stanu $q_{02}$, ponieważ 1 prowadzi nas do stanu $q_0$ oraz powoduje powrót do stanu $q_2$ przez krawędź ε. Istnieje przejście ε.



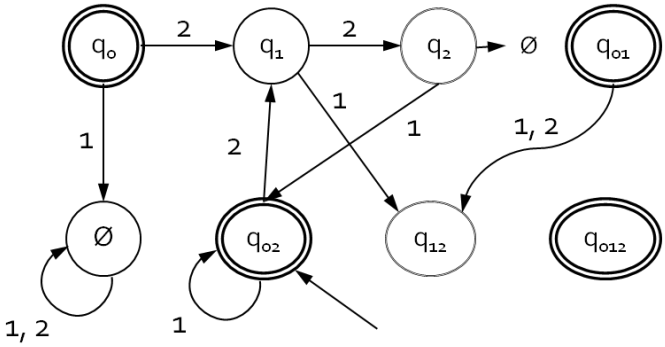

krok 9: Ze stanu $q_2$ pod wpływem 2 idziemy do stanu Ø, ponieważ nie ma w NFA przejścia pod wpływem 2 do innego stanu.



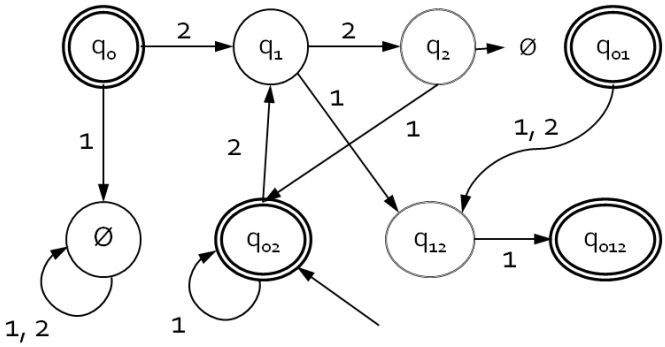

krok 10: Kolejne przejście przez ε. Ze stanu $q_{02}$ pod wpływem 1 idziemy do stanu $q_{02}$ pomimo braku bezpośredniego przejścia z $q_0$ do $q_2$. Ze stanu $q_2$ istnieje przejście (przez 1) do $q_0$, natomiast ze stanu $q_0$ mamy przejście przez ε, co powoduje przejście do stanu $q_2$.



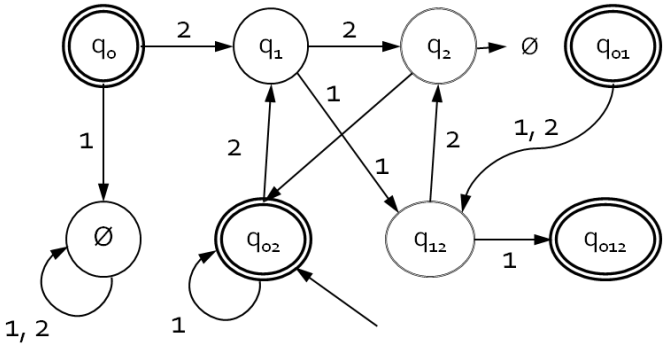
krok 11: Ze stanu $q_{02}$ pod wpływem 2 idziemy do stanu $q_1$.



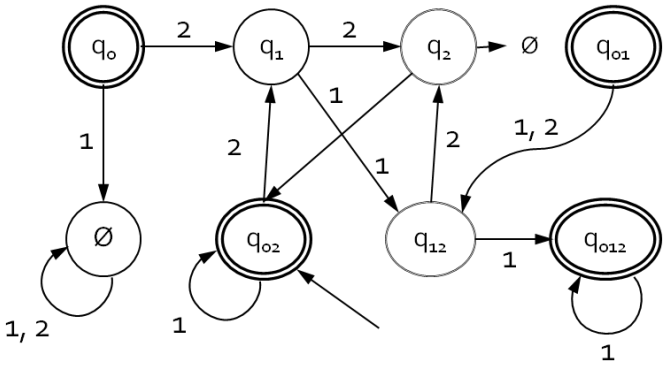

krok 12: W stanie $q_{12}$, przez 1 idziemy z $q_1$ do $q_1$, z $q_1$ do $q_2$, z $q_2$ do $q_0$ oraz powracamy przez ε do $q_2$. Stąd ze stanu $q_{12}$ pod wpływem 1 idziemy do stanu $q_{012}$.



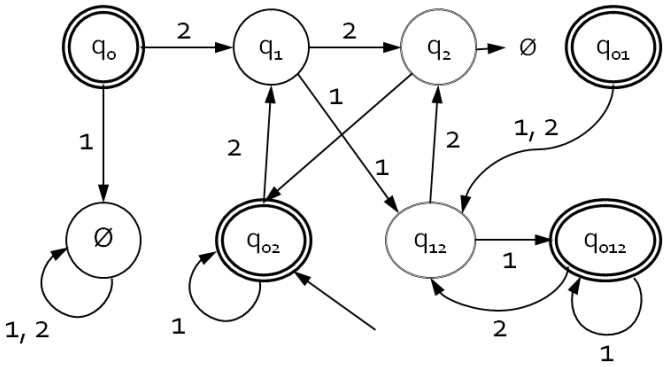

krok 13: Z $q_1$ do $q_2$ jest przejścia pod wpływem 2, natomiast a z $q_2$ nie ma przejścia oznaczonego 2. Zatem pod wpływem 2 ze stanu $q_{12}$, przechodzimy do $q_2$.



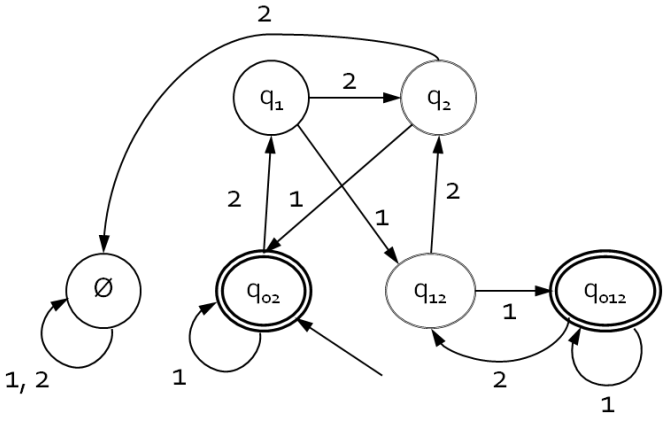
krok 14: Ze stanu $q_{012}$ pod wpływem 1 osiągalny jest każdy stan. Stąd, z $q_{012}$ do $q_{012}$.




krok 15: Ze stanu $q_{012}$ pod wpływem 2 do $q_{12}$, ponieważ nie istnieje możliwość osiągnięcia stanu $q_0$. Nie ma zdefiniowanego przejścia przez 2 ani też nie prowadzi do $q_0$ żadne puste przejście ε.




4. Minimalizujemy liczbę stanów DFA, czyli usuwamy stany nieosiągalne.
Stanami nieosiągalnymi w DFA są $q_0$ i $q_{01}$. Ostatecznie DFA ma następującą postać:




Synteza strukturalna realizowana jest zgodnie z procedurą opisaną w instrukcjach 205 i 207.

**Przebieg ćwiczenia**

Na podstawie zadania przedstawionego przez prowadzącego należy wykonać transformację automatu z postaci niedeterministycznej na postać deterministyczną (NFA $\rightarrow$ DFA). Zadanie może zostać sformułowane w postaci opisu działania układu, wyrażenia regularnego bądź grafu automatu.
Należy:
- przedstawić graf NFA (jeśli nie został podany),
- wykonać transformację na DFA, sposobem podanym w instrukcji lub uzgodnionym z prowadzącym,
- wykonać syntezę abstrakcyjną i strukturalną automatu,
- przedstawić schemat układu w postaci uwzględniającej dostępne w laboratorium układy scalone,
- zgodnie z poleceniem prowadzącego, zrealizować układ na zestawie Unilog lub wykonać symulację jego działania.

Literatura
[1] M. Sipser, Wprowadzenie do teorii obliczeń