

۱) برای زبان های زیر که روی الفبای $\Sigma = \{0, 1\}$ تعریف شده اند، عبارت منظم بنویسید.

الف) مجموعه تمام رشته هایی که در آنها دو کاراکتر پشت هم مانند هم نیستند.

$$((1+\epsilon)(01)^*(0+\epsilon))+((0+\epsilon)(10)^*(1+\epsilon))$$

ب) مجموعه تمام رشته هایی که یک عدد توان ۲ را به صورت باینری نشان می دهند. داشتن 0 در آخر عدد مشکلی ندارد
یعنی به این شکل: 001000

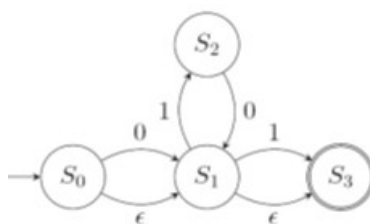
$$0^*10^*$$

پ) تمام رشته هایی که شامل حداقل یکی از زیررشته های 1011 و 1110 باشند.

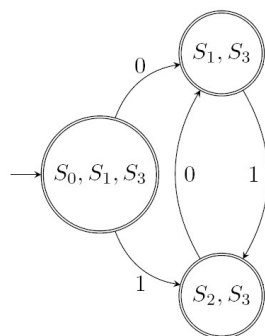
$$(1+0)^*(1011+1110)(1+0)^*$$

۲- برای هر یک از عبارات منظم سوال ۱ یک DFA رسم کنید.

الف)

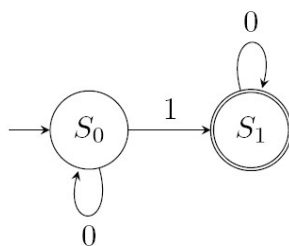


NFA

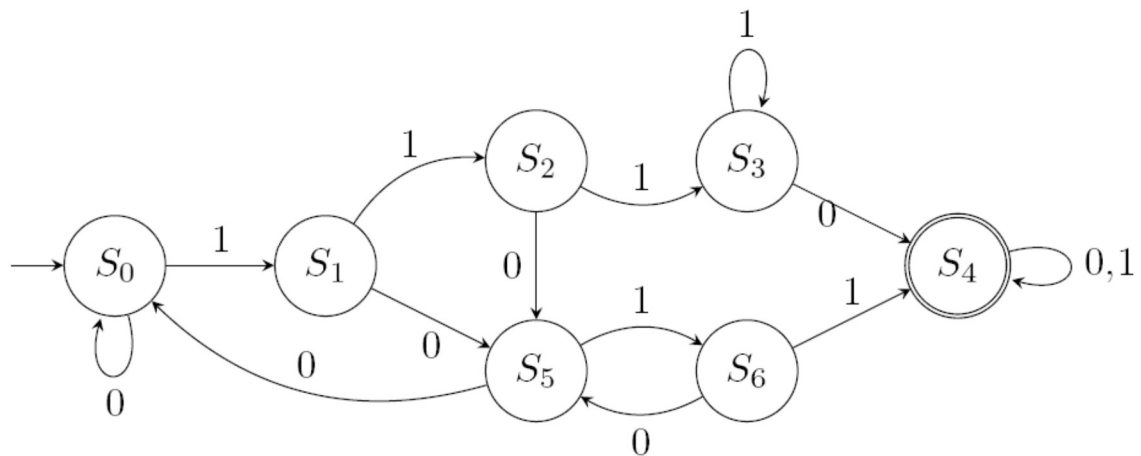


DFA

ب)

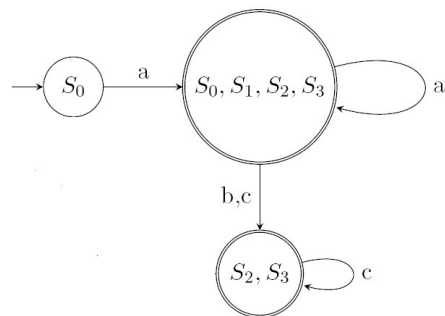


(پ)

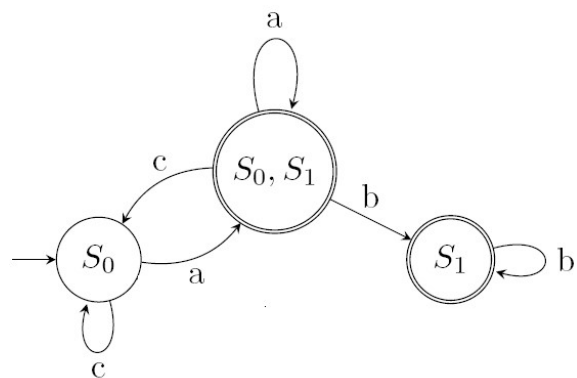


۳- با استفاده از الگوریتمی که در کلاس آموزش داده شد، NFA هایی که در ادامه آمده است را به DFA تبدیل کنید.

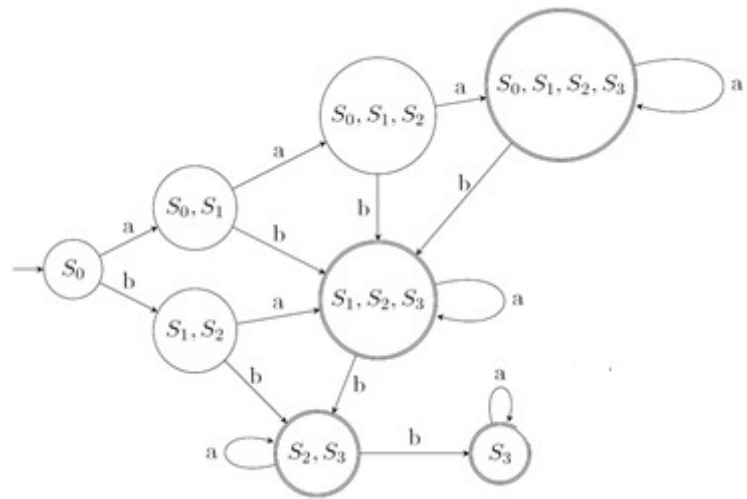
(الف)



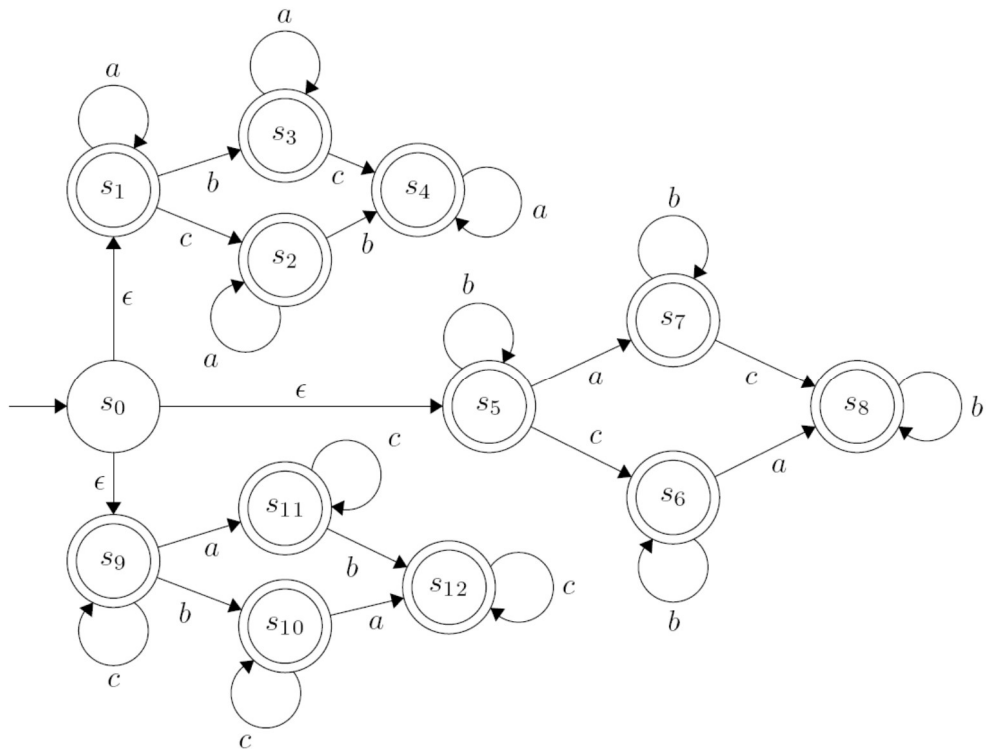
(ب)



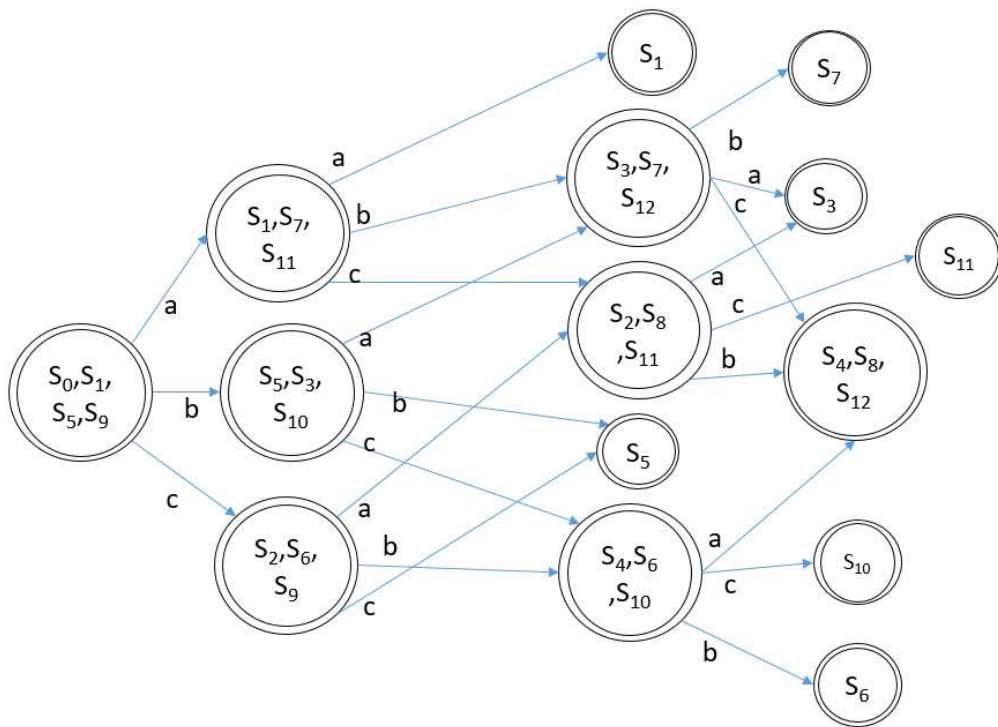
(پ)



-۴



NFA



DFA

Stateهایی از DFA که شامل یک حالت از NFA هستند، به همان صورت که در NFA هستند، ادامه پیدا می کنند.

۵- تصویر کنید که توکن های زیر به ترتیب اولویت در یک scanner تعریف شده باشند. چه ورودی ای باید به این scanner داده شود تا خروجی مطابق $HI^3(BUY^{11}GOOD^2)$ باشد.

$((0011)^{11}(1011)^2011001$

۶- به خاطر دارید که برای رفع ابهام در scanner، همیشه بزرگترین زیررشته ممکن را در نظر می گرفتیم. به این معنا که اگر توکن های زیر را در یک scanner داشته باشیم :

T_DO : DO

T_ID : [A-Za-z0-9]*

و در رشته ورودی "dot" را ببینیم، برای این رشته توکن T_ID تشخیص داده می شود. ولی ممکن است عبارت های منظمی داشته باشیم که بتوانیم یک رشته خاص را مطابق آنها به توکن ها تقسیم کنیم ولی با استفاده از قانون بیشتر زیررشته ممکن، تقسیم رشته به توکن ها با خطا مواجه شود. مثالی از تعدادی عبارت منظم (به عنوان تعریف کننده توکن ها) و یک رشته ورودی بزنید که رشته ورودی را بتوان با استفاده از عبارت های منظم تعریف شده به توکن ها تقسیم کرد ولی در صورتیکه قانون بزرگترین زیررشته در نظر گرفته شود، تقسیم رشته به توکن ها با خطا مواجه شود.

A : a

B : aba

C : bab

در صورتیکه رشته ورودی "**abab**" باشد، می توان آن را به دو توکن **AC** تقسیم کرد. ولی با استفاده از قانون بزرگترین زیررشته ممکن در ابتدا توکن **B** تشخیص داده می شود و سپس از آنجاییکه **b** با هیچ توکنی **match** نمی شود، خطا داده می شود.