2.1. ДИАГРАМА НА ПРЕХОДИТЕ

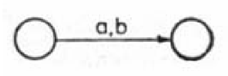
1. Определение за краен автомат като идентификатор на низове

Краен автомат A над $\Sigma = \{a_1, a_2, ..., a_n,\}$ е краен ориентиран граф, от всеки връх на който излизат n насочени дъги, всяка обозначена с етикет a_i. Всяка дума $w \in \Sigma^*$ се дефинира с w - път от връх i до връх j в A и представлява конкатенацията на етикетите на насочените дъги, през които преминава. Думата $w \in \Sigma^*$ се разпознава от крайния автомат A, ако w пътят от началния връх води до крайния връх. Празната дума Λ се разпознава от A точно тогава, когато началният връх е и краен връх. Множеството от думи, разпознавани от крайния автомат A е крайно и се бележи с $\widetilde{A}$. Едно множество е регулярно над Σ точно тогава, когато се разпознава от краен автомат над Σ.

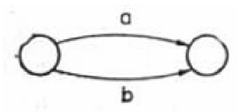
Пример:

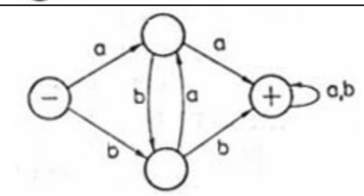
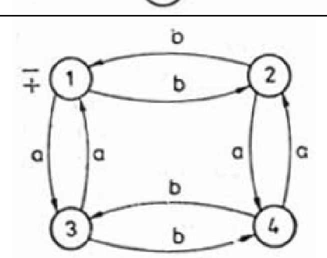
За краен автомат над $\Sigma = \{a, b\}$ са в сила следните означения.

Дъга от вида



се означава с



No	A - автомат	$\widetilde{A}$ - множество от думи
1		$\varnothing$
2		Σ^*
3		Всички думи над Σ, съдържащи две последователни a или две последователни b.
4		Всички думи над Σ, съдържащи четен брой a и четен брой b.

С всеки връх i се свързва множеството S_i от всички думи w, за които имат w-път от началния връх води до i. На множество S_i се съпоставя множество S_j на онези върхове j, от които се стига до i по една единствена дъга. Обединението на всички множества S_i за крайните върхове е $\widetilde{A}$

За автомата (4) се получава:

S_1 : Всички думи с четен брой a и четен брой b;
S_2 : Всички думи с четен брой a и нечетен брой b;
S_3 : Всички думи с нечетен брой a и четен брой b;
S_4 : Всички думи с нечетен брой a и нечетен брой b.

Диаграма на преходите T над Σ е краен ориентиран граф, всяка насочена дъга на който е белязана с дума $w \in \Sigma^*$ (евентуално празната дума Λ), наричана етикет. Има поне един връх, белязан с "-" и множество (евентуално празно) от върхове, белязани с "+" (крайни върхове). Даден връх може да бъде едновременно начален и краен.

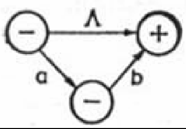
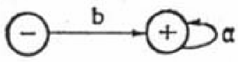
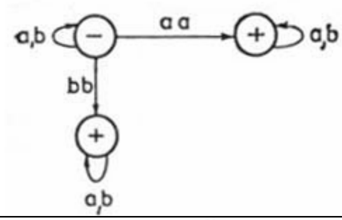
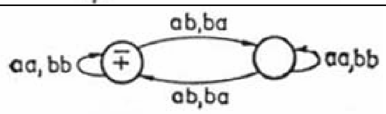
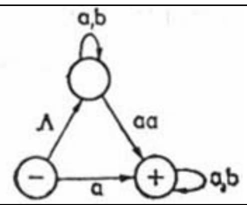
Ако думата $w \in \Sigma^*$, съществува w - краен път от връх i до връх j, конкатенацията на етикетите на насочените дъги на който представлява думата w (празните думи Λ се пренебрегват). Думата $w \in \Sigma^*$ *се разпознава* от *T*, ако съществува път w от начален до краен връх. Празната дума се разпознава от *T*, ако съществува връх в Т, който е едновременно начален и краен или има Λ - път, който води от начален до краен връх. Множеството от думи, разпознавани от *T*, се означават с $\widetilde{T}$.

Диаграмата на преходите е обобщено понятие на краен автомат. Макар класът на крайните автомати е същински подклас на диаграмите на преходите, всяко регулярно множество, което е разпознава от диаграмите на преходите, се разпознава и от някакъв краен автомат. Всеки краен автомат е диаграма на преходите. Обратното не винаги е вярно. Крайният автомат е детерминиран, т.е. $\forall\, w \in \Sigma^*$ и връх i $\exists$ единствен w-път с начало i. Диаграмата на преходите е недетерминирана. Тя съдържа повече от един (или нито един) w-път с начало i.

Теорема на Клини: 1) За всяка диаграма на преходите *T* над Σ съществува регулярен израз *R* над Σ, за който $\widetilde{R} = \widetilde{T}$;2) За всеки регулярен израз *R* над Σ съществува краен автомат *A* над Σ, за който $\widetilde{A} = \widetilde{R}$. Следствие: 1) едно множество е регулярно над Σ точно тогава, когато се разпознава от някакъв краен автомат над Σ; 2) едно множество е регулярно над Σ точно тогава, когато се разпознава от някаква диаграма на преходите над Σ.

Диаграмата на преходите *T* и множеството от думи $\widetilde{T}$ над $\Sigma = \{a, b\}$, разпознавани от нея са представени в таблица -1.

Таблица- 1

No	T -диаграма на преходите	$\widetilde{T}$ - множество от думи
1		$\varnothing$
2		$\{\Lambda\}$
3		Σ
4		$\{\Lambda, a, b\}$
5		Всички думи над Σ, започващи с b, последвано само от a – та.
6		Всички думи над Σ, съдържащи две последователни a - та и две последователни b – та.
7		Всички думи над Σ, съдържащи четен брой a - та и четен брой b – та.
8		Всички думи над Σ, които започват с a или съдържат aa.

Обобщената диаграма на преходите представлява диаграма на преходите, чиито насочени дъги са белязани с регулярни изрази

2. Метод на подмножествата за конструиране на краен автомат

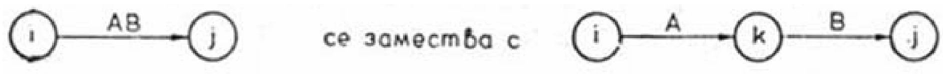
Да се конструира краен автомат A по зададен регулярен израз R над Σ, така, че $\widetilde{A} = \widetilde{R}$.

Алгоритъм:

1. Конструира се диаграма на преходите T, за която $\widetilde{T} = \widetilde{R}$. Започва се от обобщената диаграма на преходите, съставена от начален връх x и краен връх y, с дъга между тях обозначена с регулярен израз R.

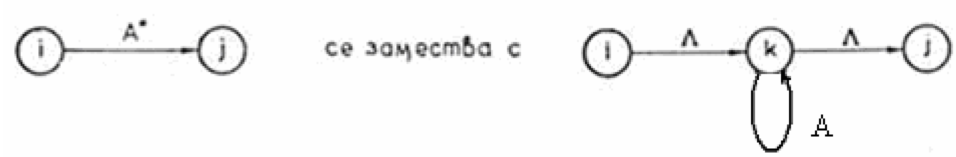
2. Последователно се разклонява R, като се прибавя нови върхове и дъги, докато всяка дъга се бележи само с буква от Σ или Λ, като се прилагат следните правила:

а. Ако между два върха i и j съществува регулярен израз, съставен от конкатенацията на два регулярни израза A и B, то се въвежда нов връх k между i и j и насочена дъга от i към k, отбелязана с етикет A и дъга от k към j, отбелязана с етикет B.



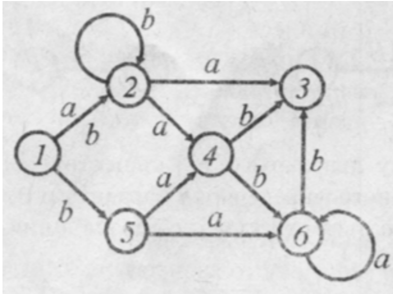
б. Ако между два върха i и j съществува регулярен израз, описващ обединението на два регулярни израза A и B, между върховете i и j се добавят две дъги от i и j, обозначени съответно A и B.

в. Ако между два върха i и j има регулярен израз от вида A^*, добавя се нов връх k между i и j и насочена дъга от i към k, отбелязана с етикет Λ и дъга от k към j, отбелязана с етикет Λ, и дъга от k към k, обозначена с регулярен израз A.



3. **Построяване на таблица на преходите**.

Нека T е диаграмата на преходите, получена след стъпка 2. Приема се, че $\Sigma = \{a, b\}$ е азбука и M е произволно подмножество на върхове на T. За всяка дума $w \in \Sigma^*$ се дефинира подмножество M_w от всички върхове на T, до които съществува w-път, водещ от някакъв връх на подмножеството M. Например, подмножеството M_{ab} се състои от всички върхове на T, до които има ab-път от някакъв връх на M или все едно всички върхове на T, до които има b-път от някакъв връх на M_a.



Пример:

1) M = {1} M_{ab} = {2} M_a = {2}
2) M = {1, 5} M_{ab} = {2, 3, 6} M_b = {5}
3) M = {1, 4, 6} M_{ab} = {2, 3} M_a = {2, 6}

Таблицата на преходите се състои от три стълба. Елементите от таблицата са подмножества от върховете на T (евентуално празното множество). Подмножеството от първи ред, първи стълб е $\{x\}_\Lambda$, т.е. подмножество, състоящо се от началния връх x и всички върхове на T, до които има Λ - път от x. За всеки ред на таблицата се прави следното: ако подмножеството M е в първи стълб, във втори стълб се прибавя множество M_a - множество от върхове на T, до които има a – път от някакъв връх на M, в трети стълб се прибавя M_b - множество от върхове на T, до които има b – път от някакъв връх на M. Ако M_a не се среща преди това в първи стълб, то се поставя в на следващия ред в първи стълб и процесът се повтаря. Аналогично се процедира с M_b. Процесът приключва, когато няма нови подмножества във втори и трети стълб на таблицата.

M	M_a	M_b
M$\{x\}_\Lambda$	-	-
-	-	-

4. **Построяване на крайния автомат**.

От таблицата на преходите се построява крайният автомат. Той се състои от толкова върха, колкото са редовете в таблицата. На всяко подмножество M в първи стълб от таблицата съответства връх $\overline{M}$ в A. Върхът $\{\overline{x}\}_\Lambda$, съответстващ на подмножеството от първи ред и първи стълб, е единственият начален връх на A. Насочените дъги на крайния автомат съответстват на наредените двойки $\langle M, M_a \rangle$ (за дъга с етикет a) и $\langle M, M_b \rangle$ (за дъга с етикет b). Даден връх $\overline{M}$ от A е краен точно тогава, когато M съдържа краен връх на T. Крайните върхове на крайния автомат са тези върхове, чиито множества в таблицата съдържат краен връх на диаграмата, от която е построена таблицата.