

Скінченні автомати з виходом. Скінченні автомати без виходу. Подання мов.

Запровадження ЕОМ у сферу інтелектуальної діяльності людини покликане до життя нову комунікативну систему «людина – машина – людина», в межах якої функціонування природної мови відрізняється від функціонування її в безпосередньому людському спілкуванні. Дослідження й опис природної мови в нових комунікативних системах вимагає й нових методів та підходів. Для розв’язування поставлених проблем прикладна лінгвістика повинна, використовуючи власне лінгвістичні дані, звертатися до багатьох інших дисциплін – кібернетики, математики, психології, фізики, медицини. Цим вона сприяє розширенню контактів мовознавчої науки з іншими науками і збагаченню лінгвістики новими точними методами дослідження мови.

Скінченний автомат з виходом

Скінченний автомат $M = (S, I, O, f, g, s_0)$ складається з:

- скінченної множини станів S ;
- скінченного вхідного алфавіту I ;
- скінченного вихідного алфавіту O ;
- функції переходів $f: S \times I \rightarrow S$;
- функції виходів $g: S \times I \rightarrow O$;
- початкового стану s_0 .

Скінченний автомат може бути заданий двома способами:

За допомогою таблиці станів, яка містить значення функції переходів f та функції виходів g для всіх пар (s, i) , де $s \in S$, $i \in I$.

За допомогою діаграми станів, яка є орієнтованим графом з поміченими дугами. Кожна вершина графа відповідає стану. Дуга позначається вхідним і вихідним сигналами, які відповідають заданому цією дугою переходу з одного стану в інший.

Приклад 1. Задання автомата таблицею (табл. 1) і діаграмою (рис. 1).

Задання автомата таблично

Табл. 1

Стан	f		g	
	вхід		вхід	
	0	1	0	1
s_0	s_1	s_0	1	0
s_1	s_3	s_0	1	1
s_2	s_1	s_2	0	1
s_3	s_2	s_1	0	0

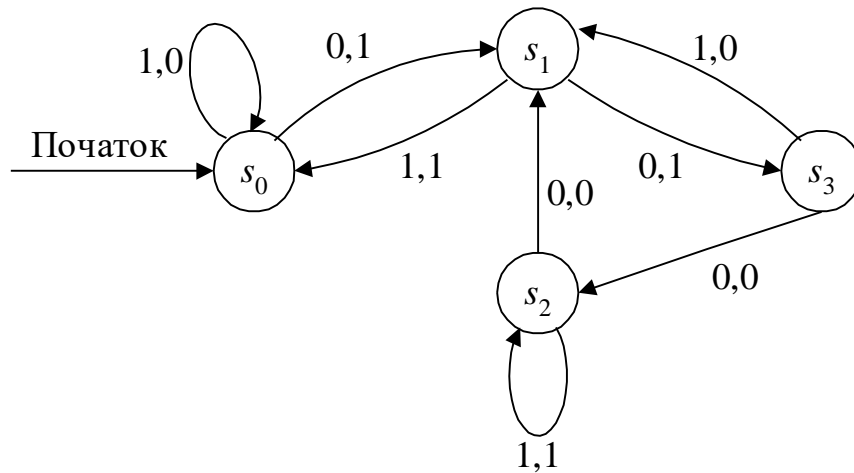


Рис. 1. Задання автомата діаграмою станів.

Приклад 2. Побудувати скінченний автомат для додавання двох цілих додатних чисел у двійковій системі.

Вхідний алфавіт складається з чотирьох символів: $I = \{00, 01, 10, 11\}$. Це необхідно для зображення можливих значень x_i та y_i – значень i -го розряду обох доданків. Вихідний алфавіт: $O = \{0, 1\}$. Множина станів $S = \{s_0, s_1\}$. Стан s_0 відповідає відсутності переносу з попереднього розряду, цей же стан початковий. Стан s_1 відповідає наявності 1 переносу з попереднього розряду. Розв’язок подано у табл. 2 і на рис. 2.

Розв’язок для **прикладу 2**

Таб.
2

С тан	f				g			
	вхід				вхід			
	0	1	0	1	0	1	0	1
s_0								
s_1								

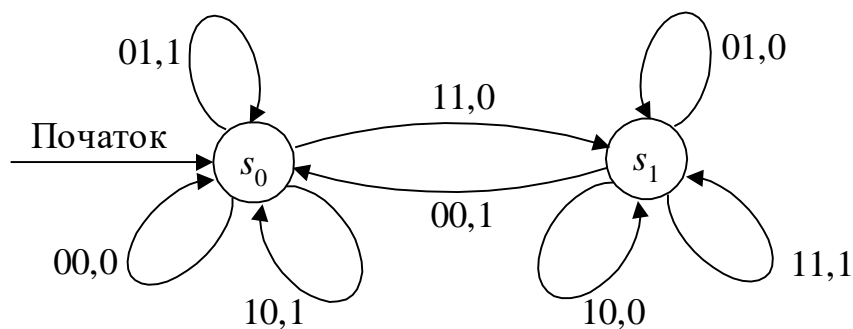


Рис. 2. Розв’язок для Прикладу 14

Приклад 3. Побудувати скінченний автомат, який видає на виході 1 тоді і тільки тоді, коли на вході останніми трьома символами були 1.

Розв'язок: $S = \{s_0, s_1, s_2\}$, $I = \{0, 1\}$, $O = \{0, 1\}$, діаграма зображена на рис. 3.

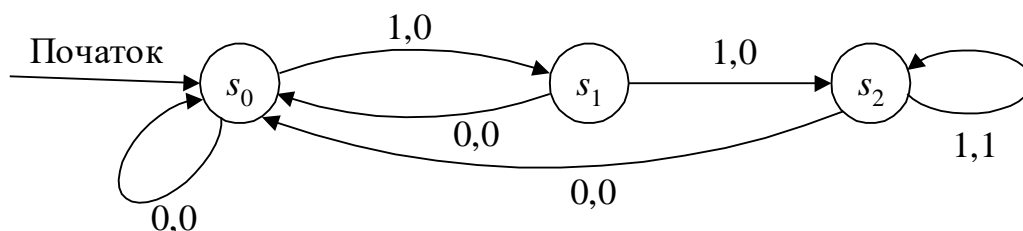


Рис. 3. Розв'язок для Прикладу 3

У вигляді вправи пропонуємо задати цей же автомат таблицею станів.

Автомат з останнього прикладу розпізнає мову, оскільки він продукує на виході 1 тоді і тільки тоді, коли вхідний ланцюжок (слово) має спеціальні властивості.

Автомати, які ми розглянули, називаються автоматами Мілі (G.H. Mealy). Вперше введені у 1955 році.

Існує також інший тип автоматів з виходом, так звані автомати Мура (E.F. Moore), які запроваджені у 1956 році. У цих автоматах вихід визначається лише станом, тобто не залежить від вхідного сигналу.

У прикладі 15 показано, як автомат Мілі може бути використаний для розпізнавання мови. Проте з цією метою звичайно використовують інший тип автоматів – скінченні автомати без виходу. Такі автомати мають множину кінцевих станів і розпізнають ланцюжок тоді і тільки тоді, коли цей ланцюжок переводить автомат без виходу із початкового стану у кінцевий стан.

Скінченний автомат без виходу

Скінченний автомат без виходу – це п'ятірка $M = (S, I, f, s_0, F)$, яка містить:

- скінченну множину S станів;
- скінченний вхідний алфавіт I ;
- функцію переходів $f : S \times I \rightarrow S$;
- початковий стан s_0 ;
- підмножину $F \subset S$, елементи F називаються заключними станами.

Скінченний автомат без виходу може бути заданий:

- таблицею станів;
- діаграмою станів, заключні стани на діаграмі позначаються подвійними кружечками.

Приклад 4. Зобразити діаграму станів для автомата $M = (S, I, f, s_0, F)$, де $S = \{s_0, s_1, s_2, s_3\}$, $I = \{0, 1\}$, $F = \{s_0, s_3\}$. Функція переходів задана таблицею 3. Діаграма станів зображена на рис. 4.

Поняття функції переходів f можна розширити і визначити її для всіх пар станів і ланцюжків. У такому разі, нехай $x = x_1x_2...x_k$ – ланцюжок з I^* . Тоді $f(s_1, x)$ – стан, що обчислений з використанням послідовних символів з x зліва направо, як вхідних символів, починаючи зі стану s_1 . Процес іде так: $s_2 = f(s_1, x_1)$; $s_3 = f(s_2, x_2)$... Покладаємо

$f(s_1, x) = f(s_k, x_k)$. Ланцюжок x сприймається, або розпізнається скінченним автоматом $M = (S, I, O, f, s_0, F)$, якщо він переводить початковий стан s_0 у кінцевий стан – це означає, що стан $f(s_0, x)$ є елементом множини F .

Табл. 3.

Стан	f	
	вхід	
	0	1
s_0	s_0	s_1
s_1	s_0	s_2
s_2	s_0	s_0
s_3	s_2	s_1

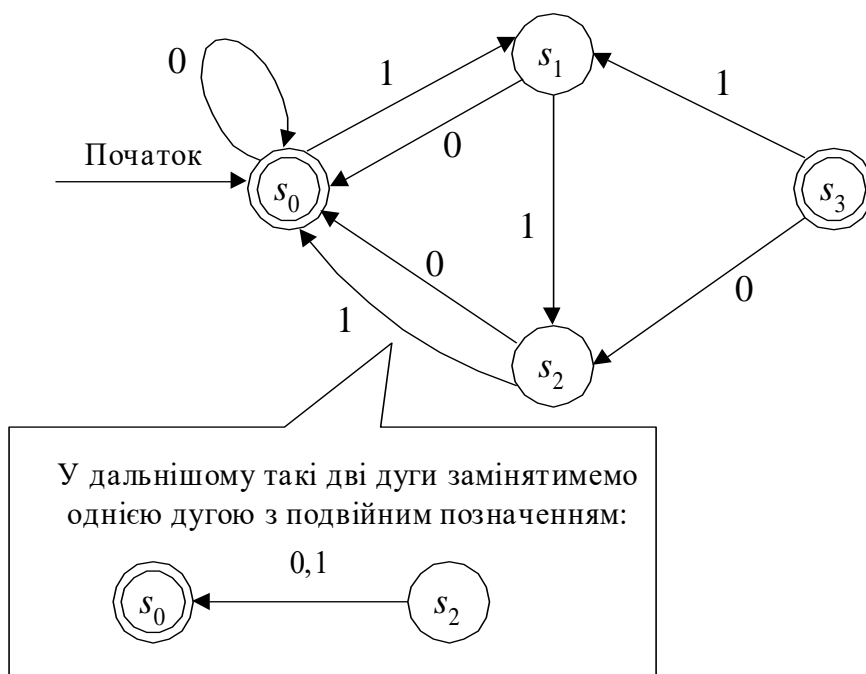


Рис. 4.

Подання мов

Мова, що розпізнається, або сприймається автоматом M , позначається через $L(M)$ – це множина всіх ланцюжків, які розпізнаються автоматом M . Два автомати називаються еквівалентними, якщо вони розпізнають одну і ту саму мову.

Приклад 5. Визначити мову, що розпізнається автоматом M з діаграмою на рис. 8.

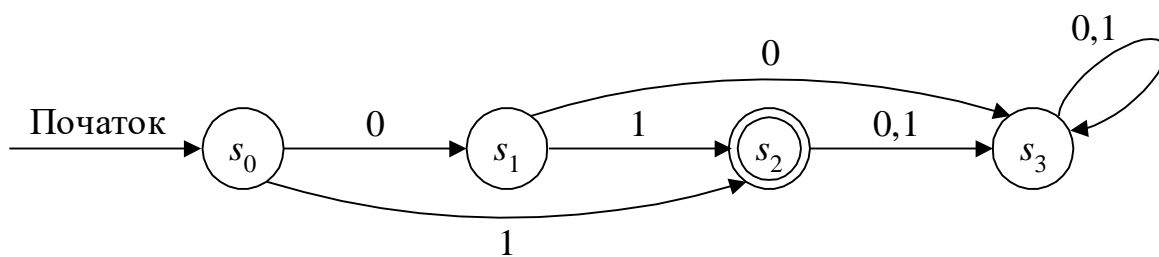


Рис. 5.

Відповідь: $L(M) = \{1, 01\}$.

Приклад 6. Визначити мову, що розпізнається автоматом M з діаграмою на рис. 6.

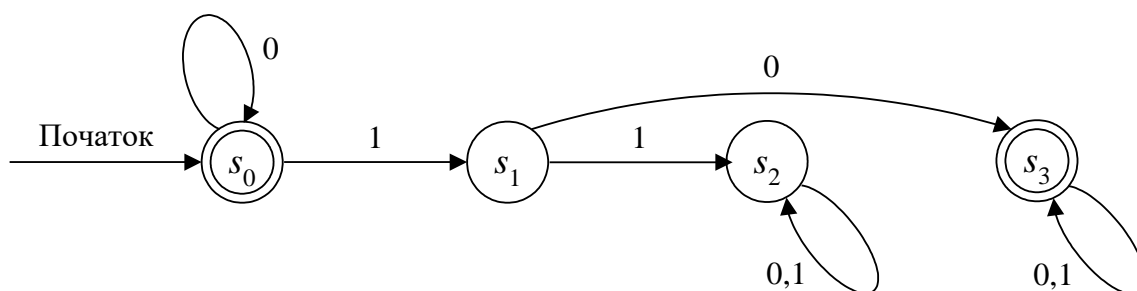


Рис. 6.

Відповідь: $L(M) = \{0^n, 0^n 10x \mid n=0,1,2,\dots \text{ та } x \text{ є довільним ланцюжком}\}$.

Розглянуті скінченні автомати без виходу називаються *детермінованими*, оскільки для кожної пари *стан – вхідний символ* існує єдиний наступний стан, який задається функцією переходів. Існує інший тип автоматів без виходу – це *недетерміновані* автомати. У таких автоматах може бути *декілька можливих наступних станів* для кожної пари *стан – вхідний символ*.

Недетермінований скінченний автомат без виходу – це п'ятірка $M = (S, I, f, s_0, F)$, де:

- S – скінченна множина станів;
- I – скінченний вхідний алфавіт;
- f – функція переходів, яка кожній парі стан – вхідний символ ставить у відповідність множину станів;
- s_0 – початковий стан;
- $F \subset S$, де F – множина кінцевих станів.

Приклад 7. На рис. 7 і у табл. 4 зображено діаграму і таблицю станів деякого недетермінованого автомата.

Табл. 2

**Табличне задання
недетермінованого автомата**

Стан	f	
	ВХІД	
	0	1
s_0	s_0, s_2	s_1
s_1	s_3	s_4
s_2	—	s_4
s_3	s_3	—
s_4	s_3	s_3

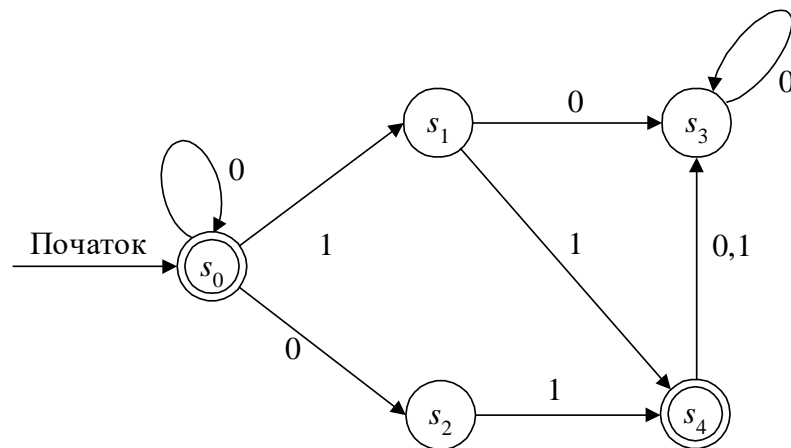


Рис. 7. Задання недетермінованого автомата множиною станів

Що означає, що недетермінований автомат розпізнає ланцюжок $x = x_1x_2...x_k$?

Щоразу розглядається *множина станів*, яка визначається функцією f . Автомат розпізнає, або сприймає ланцюжок x , якщо є заключний стан у тій множині станів, яка отримана з початкового стану s_0 та ланцюжка x . Мова розпізнається недетермінованим автоматом, якщо множина всіх ланцюжків (слів) цієї мови розпізнається цим автоматом.

Приклад 8. Знайти мову, що розпізнається автоматом M з прикладу 7.

Неважко переконатись, що

$$L(M) = \{0^n, 0^n10, 0^n11 \mid n = 0,1,2,\dots\}.$$

Теорема 1. Якщо мова L розпізнається недетермінованим скінченним автоматом M_0 , то існує також детермінований скінченний автомат M_1 , який розпізнає цю мову.

Теорема 2. Регулярна мова і тільки вона породжується регулярною граматикою.

Теорема 3. Для того, щоб мова була регулярною, необхідно і достатньо, щоб існував скінченний автомат, який її розпізнає.

Проілюструємо кілька часткових випадків у розпізнаванні мов.

1. Множину \emptyset допускає недетермінований автомат без заключних станів, зображений на рис. 8.

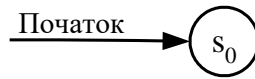


Рис.8.

2. Недетермінований автомат, який допускає $\{\lambda\}$, має початковим і заключним станами s_0 (рис. 16).

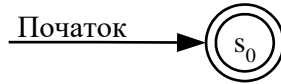


Рис.9.

Тепер визначимо, як недетермінований скінченний автомат допускає ланцюжок $\alpha = x_1 x_2 \dots x_k$. Перший вхідний символ x_1 переводить стан s_0 у множину S_1 , яка може містити більше одного стану. Наступний вхідний символ x_2 переводить кожний зі станів S_1 у деяку множину станів, і нехай S_2 буде об'єднанням цих множин. Ми продовжуємо цей процес, вибираючи на кожній стадії всі стани, отримані з використанням поточного вхідного символу і всіх станів, отриманих на попередній стадії. Недетермінований скінченний автомат *допускає* ланцюжок α , якщо у множині станів, отриманій з початкового стану s_0 під дією ланцюжка α , є заключний стан. Мова, яка *допускається* недетермінованим скінченим автоматом, – це множина всіх ланцюжків, які допускаються цим автоматом.

Приклад 9. Знайдемо мову, яка допускається недетермінованим скінченим автоматом з рис. 7.

Оскільки s_0 є заключним станом, і вхід 0 переводить s_0 у себе, то автомат допускає ланцюжки $\lambda, 0, 00, 000, 0000, \dots$ Стан s_4 також заключний, і нехай s_4 є у множині станів, що досягаються зі стану s_0 з ланцюжком α на вході. Тоді ланцюжок α допускається. Такими ланцюжками є $0^n 01$ та $0^n 11$, де $n = 0, 1, 2, \dots$ Оскільки інших заключних станів немає, то мова, яка допускається цим недетермінованим автоматом, є такою:

$$\{0^n, 0^n 01, 0^n 11 \mid n = 0, 1, 2, \dots\}.$$

Важливо зазначити: якщо мова допускається недетермінованим автоматом, то вона також допускається детермінованим автоматом.

З теореми 1 випливає, що недетерміновані скінченні автомати допускають ті самі мови (множини ланцюжків), що і детерміновані скінченні автомати. Проте є причини розглядати недетерміновані автомати. Вони часто компактніші, і їх легше побудувати, ніж детерміновані. Крім того, хоча недетермінований автомат завжди можна перетворити у детермінований, детермінований може мати експоненціально більше станів. На щастя, такі випадки досить рідкісні.

Приклад 10. Знайдемо детермінований скінченний автомат, який допускає ту саму мову, що й недетермінований автомат з рис. 7.

Детермінований скінченний автомат, зображений на рис. 10, побудований з недетермінованого автомата з прикладу 7.

Стани цього детермінованого автомата є підмножинами станів недетермінованого автомата з рис. 13. Їх отримують так, як описано у доведенні теореми 1. Зокрема, вхід 0 переводить стан $\{s_0\}$ у стан $\{s_0, s_2\}$ в детермінованому автоматі, оскільки s_0 у разі входу 0 переходить у самого себе і в s_2 у недетермінованому автоматі. Аналогічно, у разі входу 1 множина $\{s_0, s_2\}$ переходить у $\{s_1, s_4\}$, оскільки s_0 переходить в s_1 , а s_2 – в s_4 у недетермінованому автоматі у

разі входу 1. Нарешті, вхід 0 переводить множину $\{s_1, s_4\}$ у множину $\{s_3\}$, оскільки в недетермінованому автоматі вхід 0 переводить обидва стани s_1 та s_4 у стан s_3 . Усі підмножини, які отримують таким способом, є станами детермінованого скінченного автомата з рис. 10. Наголосимо, що одним із станів цього автомата є порожня множина: вона містить усі стани, в які вхід 1 переводить $\{s_3\}$. Початковий стан – $\{s_0\}$, а заключними є всі ті стани, які містять s_0 або s_4 .

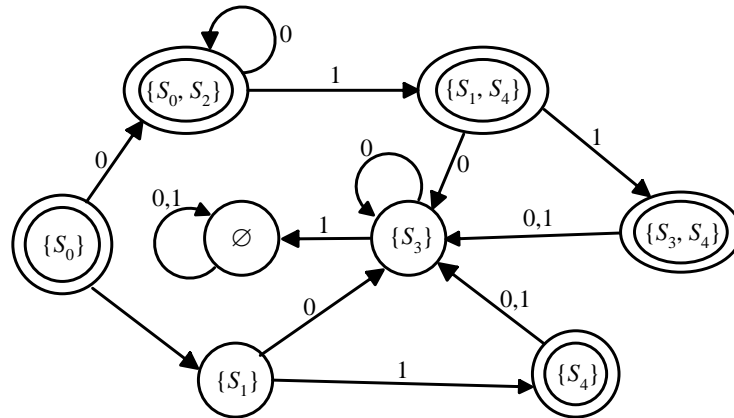


Рис.10.