

Алгоритми і скінченні автомати

Про теорію автоматів

Теорія автоматів — логіко-математична теорія, об'єктом дослідження якої є так звані абстрактні автомати — покрокові перетворювачі інформації. *Інформація* — абстрактне судження, що має різні значення залежно від контексту. Спроб формального визначення інформації є багато, і вони відмінні в різних областях людської діяльності. На інтуїтивному рівні інформація означає зміст того, про що отримувач довідався.

Варто розрізняти два терміни: *дані* та *інформація*. Від даних інформація відрізняється доступністю отримувачу. Наприклад, запис

[18, 22, 21, 25, 19, 24, 24]

є сам по собі є лише даними. А слово "інформація" має декілька значень: роз'яснення; виклад фактів, подій; витлумачення; представлення, поняття. Отже, інформацією для наведеного приклада може бути:

"денна температура повітря за тиждень": 18->22->21->25->19->24->24

Однією з найважливіших властивостей інформації є *достовірність* — здатність інформації об'єктивно відбивати процеси та явища, що відбуваються в навколишньому світі. Як правило, такою вважається насамперед інформація, яка несе у собі безпомилкові та істинні дані.

Автомат визначають двома сутностями: "стан автомата", "перехід з одного стану в інший". Поведінку автомата можна пояснити на простому прикладі. Задано 3 цілі числа a, b, c , надруковані в рядок через кому. Для простоти вважати всі числа попарно різними, двох однакових чисел немає. Знайти найбільше шляхом елементарних бінарних операцій.

Позначимо початковий стан автомата як S . Інші стани для прикладу можна позначати літерами. Якщо $a > b$, то перейти в стан нехай R , а при $a < b$ перейти в стан T . Зі стану R (було $a > b$) перейти в стан M при $a > c$, чи перейти в стан K при $a < c$. Стан M є кінцевим, в ньому маємо найбільше число рівне a . Стан K є кінцевим, в ньому маємо найбільше число рівне c . Зі стану T (було $a < b$) перейти в стан H при $b > c$, а при $b < c$ перейти в стан D . Стан H дає кінцеву відповідь про найбільше b , а стан D дає кінцеву відповідь про найбільше c .

Графічно можна уявити автомат як діаграму, на якій колами позначені стани автомата, а переходи з одного стану в інший позначені відповідними дугами. Іншими словами, автомат — це граф з вершинами і ребрами.

Скінченні автомати (СА)

Скінченний автомат (англ. *finite-state machine*, машина зі скінченною кількістю станів) — особливий різновид автомата — абстракції, що використовується для опису шляху зміни стану об'єкта в залежності від поточного стану та інформації отриманої ззовні. Його особливістю є скінченність множини станів автомата.

Існує дві різних групи автоматів: *акцептори/розпізнавачі* і *перетворювачі (трансдуктори)*. Кожну групу можна будувати за типом "автомат Мура" чи "автомат Мілі".

Акцептори і розпізнавачі

Акцептори і розпізнавачі мають виконати дослідження деякого об'єкта і побудувати двійковий вихід "так" або "ні" щодо правильності об'єкта. Отже, треба встановити лише факт правильності об'єкта. Повернемося до задачі про три числа: чи правильно задані вхідні дані? За умовою, числа мають бути надруковані через кому одним рядком, все разом є об'єктом. Приклад такого СА показаний на рис.1.

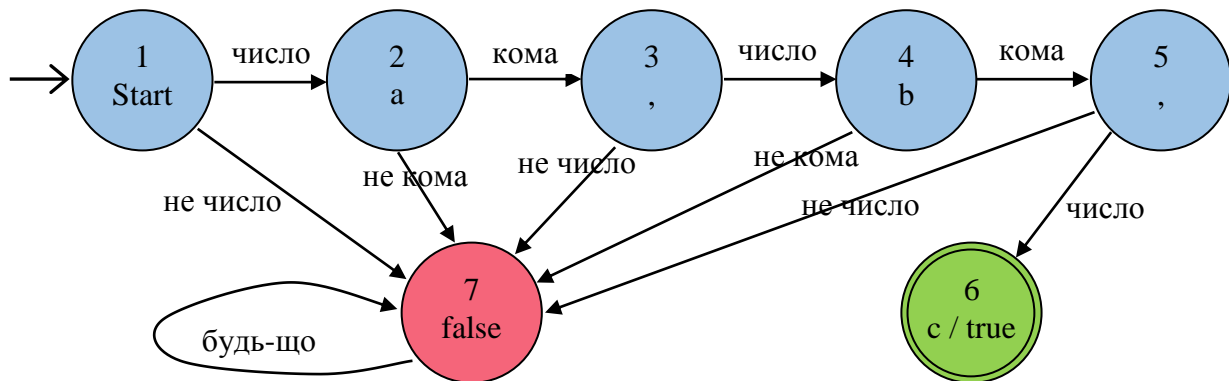


Рис.1

На рис.1 стани (кола) позначені номерами і умовними підписами.

Початковий стан зазвичай показують зі стрілкою «звідкись». Таким станом є 1/Start.

Допустимий стан (або *кінцевий*) – досягаючи якого, автомат дає ствердну відповідь про правильність об'єкта. Зазвичай позначають двома колами. Таким станом є 6.

Дуги (ребра) позначають вхідними даними з врахуванням порядку поступлення.

Правила роботи автомата-розпізнавача: 1) найпершим станом вважати Start; 2) розпочати від найпершого елемента об'єкта і повторювати крок 3 доти, поки не досягнемо останнього елемента об'єкта; 3) читати і аналізувати наступний елемент об'єкта, і перейти ребром, позначеним цим елементом, до наступного стану; 4) якщо досягаємо останнього елемента об'єкта, тоді цілий об'єкт є правильним тоді і тільки тоді, коли останнім станом автомата є допустимий.

Перетворювачі (трансдуктори)

Такі СА будують послідовність операцій для розв'язання задачі. На кожному кроці показують операції для поточної групи вхідних даних, переходять до наступного стану, і будують далі операції для наступної групи даних. Вирізняють два види перетворювачів, подібно до розпізнавачів: "автомат Мура"; "автомат Мілі".

Автомат Мура

Роботу такого автомата можна визначити парами кроків: 1) перейти в деякий стан; 2) в позиції стану виконати операцію або інші дії. Ребра графа залишаємо без позначень, а біля вершин записуємо результат операції, чи саму операцію, чи вибір даних для операцій. Схема: "перейти-виконати". Отже, перехід (функція виходів автомата Мура m) до наступного стану $next$ залежить лише від попереднього стану $prev$: $y(next)=m(g(prev))$.

На рис.2 показаний автомат Мура для задачі пошуку найбільшого числа з трьох.

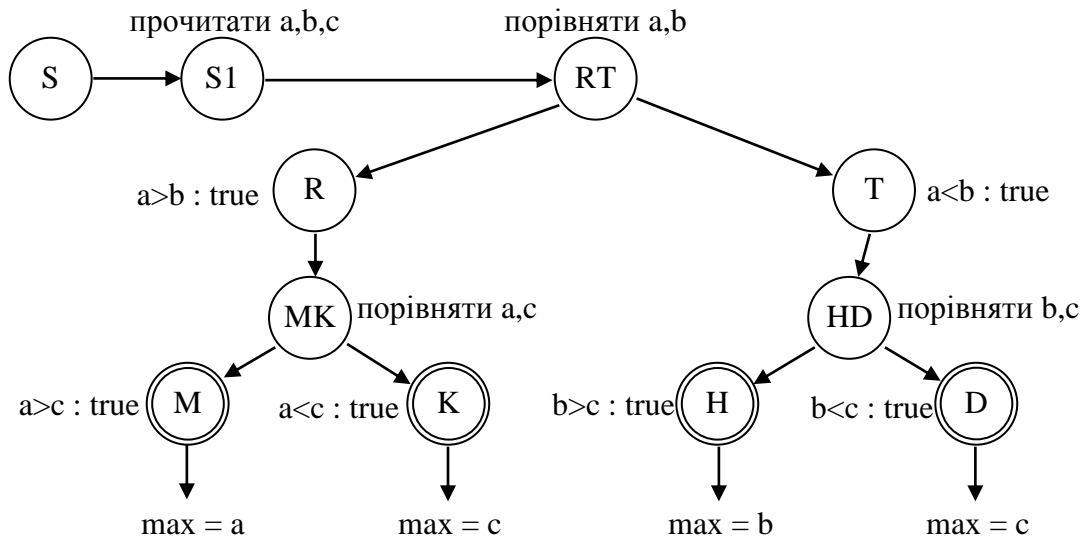


Рис.2

Автомат Мілі

Роботу автомата Мілі можна визначити такими парами кроків: 1) на ребрі позначаємо деяку умову, чи вхідні і вихідні символи даних, чи реалізуємо операцію; 2) виконуємо перехід в деякий стан. Перехід до наступного стану залежить від попереднього стану і від умови чи даних. Схема: "виконати-перейти" (обернена до автомата Мура).

На рис.3 показаний автомат Мілі для тієї самої задачі пошуку найбільшого числа з трьох заданих.

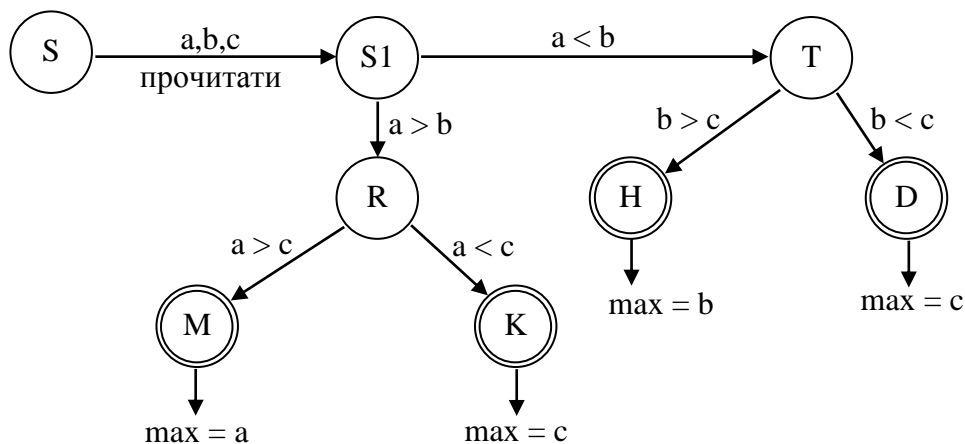


Рис.3

Поведінка автомата Мілі визначається функцією переходів (функцією відображення) $\Omega(\Sigma, \alpha) = \Delta$, яка означає, що з стану Σ при умові чи вхідному символі α відбувається перехід в стан Δ . Наприклад, для показаного автомата маємо: $\Omega(R, "a > c") = M$, $\Omega(R, "a < c") = K$.

Приклади задач розпізнавання

Задача 1. Задано слово як послідовність знаків. Визначити, чи є слово ідентифікатором алгоритмічної мови. Вважати правила будови ідентифікатора такими: найпершим знаком має бути буква; наступними знаками можуть бути букви і цифри; кількість знаків ідентифікатора не обмежена.

Автомат Мура цієї задачі показаний на рис.4.

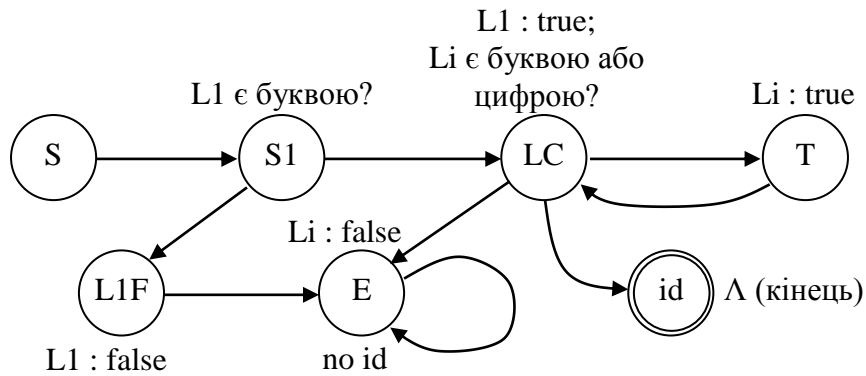


Рис.4

Автомат Мілі цієї ж задачі показаний на рис.5.

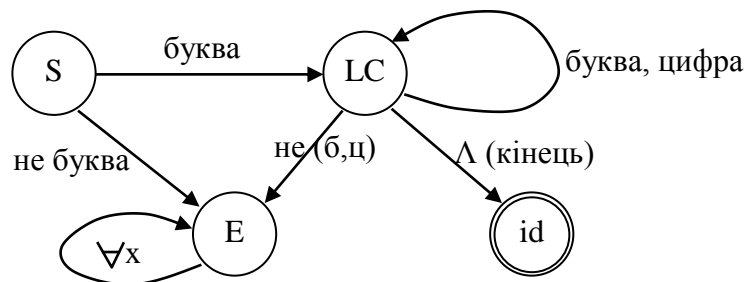


Рис.5

Задача 2. Побудувати автомат Мілі типу розпізнавача для аналізу чисел у всіх форматах: ціле; фіксована крапка; експоненціальний формат.

Граматичні правила для чисел:

$\langle \text{число} \rangle ::= \langle \text{ціле} \rangle \mid \langle \text{фіксована крапка} \rangle \mid \langle \text{експоненціальний формат} \rangle$
 $\langle \text{ціле} \rangle ::= \text{цифра} \mid \langle \text{ціле} \rangle \text{ цифра}$
 $\langle \text{фіксована крапка} \rangle ::= \langle \text{ціле} \rangle . \langle \text{ціле} \rangle$
 $\langle \text{експоненціальний формат} \rangle ::= \langle \text{мантиса} \rangle \langle \text{порядок} \rangle$
 $\langle \text{мантиса} \rangle ::= \langle \text{ціле} \rangle \mid \langle \text{фіксована крапка} \rangle$
 $\langle \text{порядок} \rangle ::= E \langle \text{знак} \rangle \langle \text{ціле} \rangle \mid E \langle \text{ціле} \rangle$
 $\langle \text{знак} \rangle ::= + \mid -$

На рис.6 показаний вигляд автомата Мілі цієї задачі. Тут допустимим станом вважаємо "out (число)". Знаком Λ (пусто) позначаємо відсутність умови. Якщо з деякого стану виходять декілька ребер, то перехід за ребром Λ виконують лише тоді, коли не виконуються умови переходу жодного іншого ребра, тобто – "у всіх інших випадках".

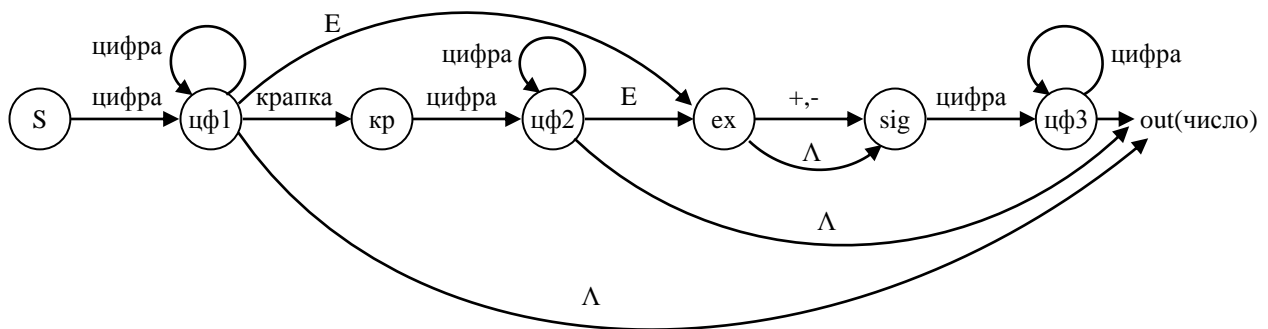


Рис.6

Порівняння автоматів Мура і Мілі

- ❖ Для будь-якого автомата Мілі існує еквівалентний йому автомат Мура. Кожна машина Мура m еквівалентна машині Мілі з тими ж станами та переходами.
- ❖ У машині Мура кожен вузол (стан) позначений вихідним значенням.
- ❖ У машині Мілі кожне ребро (перехід) позначене вихідним значенням.
- ❖ Використання автомата Мілі часто призводить до зменшення кількості станів.
- ❖ Автомат Мура може мати деяку неоднозначність поведінки при зовнішньому аналізі. Наприклад, для задачі перевірки слова на ідентифікатор (рис.4 копія) з стану S1 можна перейти в стан LC або в стан L1F, в яких фіксуємо відповідь про перевірку найпершого знака ідентифікатора. Але перехід треба виконати до фіксування відповіді. З діаграми автомата не впливає спосіб визначення переходу.

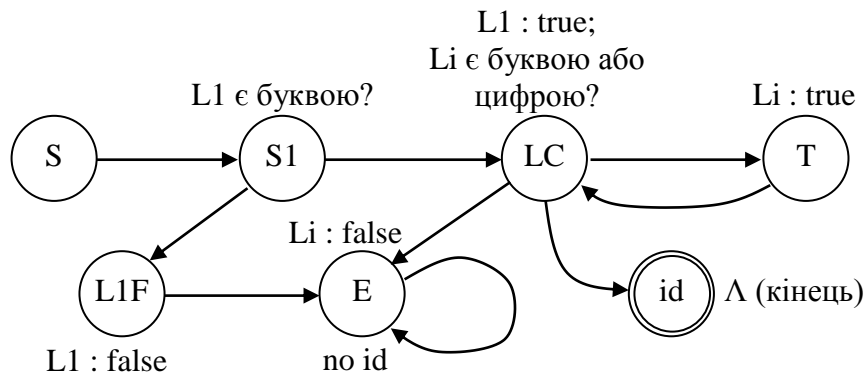


Рис.4 (копія)

- ❖ Автомат Мура краще будувати для задач технічних пристроїв, коли перехід в стан означає виконання деяких дій, наприклад: "увімкнути двигун", "вимкнути двигун", "повернути ліворуч", "повернути праворуч", "читати покази датчика температури". Приклади задач можна знайти за посиланнями на матеріали теми.
- ❖ Автомат Мілі більш придатний для зображення алгоритмів розв'язування обчислювальних задач, коли в результаті одного кроку обчислень переходимо в наступний стан для наступного кроку обчислень.
- ❖ На практиці часто використовують суміщення моделей автоматів Мура і Мілі.

Суміщення автоматів Мура і Мілі

Суміщення автоматів Мура і Мілі виконаємо за таким правилом. На ребрах позначаємо умову, а біля вершин записуємо операції.

Задача. В масиві чисел X порахувати середні арифметичні значення p для додатніх і m для від'ємних.

Масив складається з окремих чисел $[x_1, x_2, \dots, x_k]$. Числа читаємо по черзі окремо кожне. Для кожного числа будуємо відповідну функцію переходу.

Позначення величин: x – чергове число; sp – сума додатніх; sm – сума від'ємних; np – кількість додатніх; nm – кількість від'ємних. Знаком Λ (пусто) позначаємо відсутність умови чи відсутність дії. (Рис.7)

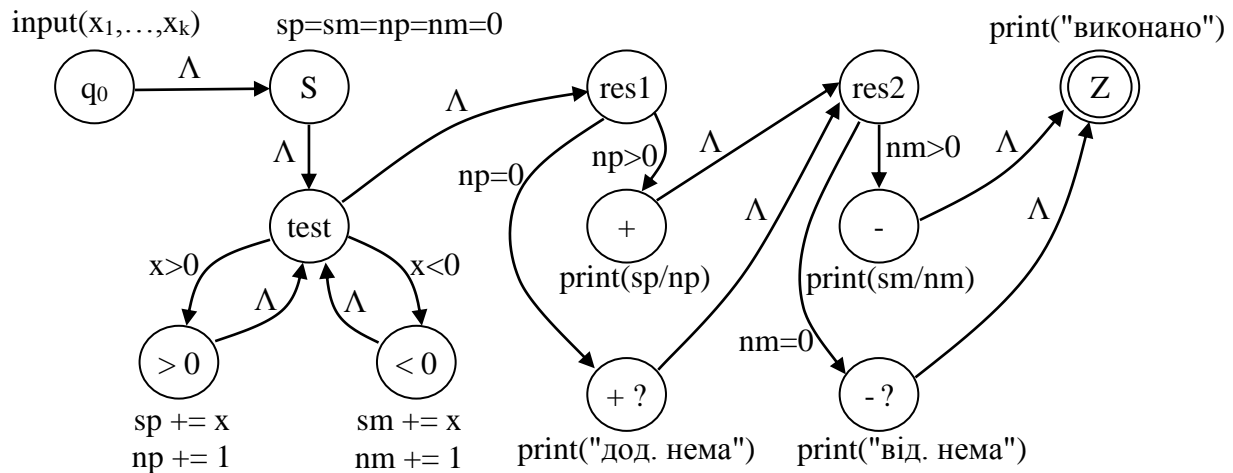


Рис.7

Звернемо увагу, що суміщення автоматів Мура і Мілі за логікою будови близьке до зображення алгоритма блок-схемою.

Детермінований скінченний автомат

Детермінований скінченний автомат (ДСА) – скінченний автомат *групи розпізнавачів* або *групи перетворювачів*, який приймає скінченний рядок символів, тобто рядок є об'єктом дослідження. Це важливий частковий випадок скінченних автоматів. ДСА розпізнає набір регулярних мов, що є корисним для проведення лексичного аналізу і зіставлення із взірцем. ДСА можна використати або в режимі приймача для перевірки належності вхідного рядка до мови, або в режимі генерування для створення списку всіх рядків у мові.

Для кожного стану існує ребро переходу в наступний стан для кожного символу. Після зчитування символу ДСА переходить детерміновано з одного стану в інший за відповідним ребром. *Детермінованість* означає наявність лише одного результату (тобто переходу в наступний стан для кожного символу ($S_0 \rightarrow S_i$) або повернення в той самий стан ($S_0 \rightarrow S_0$)). ДСА має початковий стан (позначений графічно стрілкою нізвідки), де починаються обчислення, і набір допустимих станів (позначених графічно двійними колами), які визначають успішність обчислень.

ДСА є окремим підкласом автоматів Мілі для задач розпізнавання чи перетворення. Приклади задач показані вище: 1) чи є слово ідентифікатором (рис.4); 2) правильність запису числа (рис.6).

Формальне визначення. Детермінований скінченний автомат M – це п'ятірка $(Q, \Sigma, \delta, q_0, Z)$, де:

Q – скінченний набір станів $\{q_i\}$;

Σ – скінченний набір вхідних символів $\{a_i\}$ (абетка), літери, які можуть зустрітись у вхідному рядку;

δ – функція переходу (відображення) декартового добутку множин $Q \times \Sigma \rightarrow Q$; якщо $\delta(q_1, a_i) = q_2$, то це означає, що з стану q_1 при вхідній літері a_i відбувається перехід в стан q_2 ;

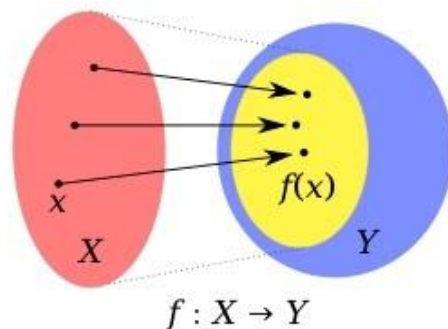
$q_0 \in Q$ – початковий стан;

$Z \subseteq Q$ – непуста множина допустимих станів (кінцевих).

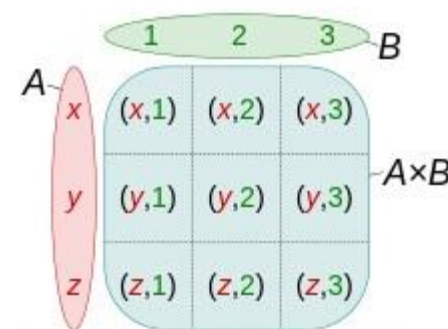
Нехай $w = a_1 a_2 \dots a_n$ – рядок над абеткою Σ . Автомат M приймає рядок w , якщо послідовність станів r_0, r_1, \dots, r_n в Q відповідає таким умовам:

1. $r_0 = q_0$
2. $r_{i+1} = \delta(r_i, a_{i+1})$ for $i=0, \dots, n-1$
3. $r_n \in Z$.

Словами, перша умова каже, що автомат починає з початкового стану q_0 . Друга умова каже, що з кожним наступним символом з рядка w автомат переходить зі стану в стан за функцією переходу δ . Остання умова каже, що автомат приймає w , якщо останній символ з w спричиняє перехід автомата в один з допустимих станів. Інакше, кажуть, що автомат відхилив рядок. Набір допустимих рядків M – це мова, що розпізнається автоматом M , і така мова позначається $L(M)$.



Однозначна функція f відображає область визначення X в цільову множину Y ; менший овал всередині Y — це область значень функції f



Декартів добуток $A \times B$ множин $A = \{x, y, z\}$ та $B = \{1, 2, 3\}$

Приклад 1. Побудувати автомат M типу розпізнавача, який приймає цілі числа без знаку, які діляться на 5.

Число складається з окремих цифр $N=d_1d_2\dots d_n$. Цифри читаємо по черзі окремо кожен. Для кожної цифри будуємо відповідну функцію переходу. Число прийняте, якщо остання цифра 5 або 0. (Рис.8)

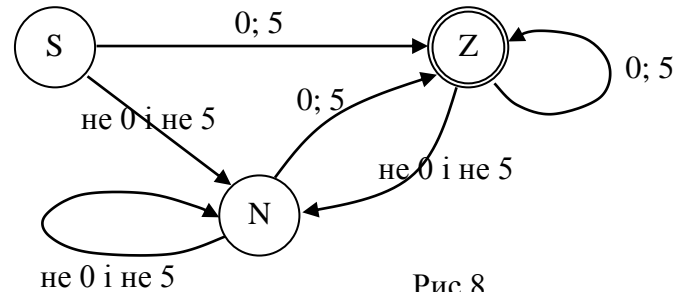


Рис.8

Приклад 2. Побудувати автомат M типу перетворювача для такої задачі. Задане число складається з окремих цифр $N=d_1d_2\dots d_n$. 1) Утворити з нього нове число R шляхом викреслення цифр 0 і 5. 2) Порахувати кількість цифр n_x заданого числа і кількість цифр n_y нового числа.

Приймемо за основу автомат, побудований для попереднього прикладу 1. Для будови структури автомата цього прикладу достатньо додати два нові стани до автомата попереднього прикладу. Кожне ребро будемо позначати парою умова/операції, розділяючи їх горизонтальною лінією.

Знаком Λ (пусто) позначаємо відсутність умови, а також відсутність операцій. Якщо з деякого стану виходять декілька ребер, то перехід за ребром Λ виконують лише тоді, коли не виконуються умови переходу жодного іншого ребра, тобто – "у всіх інших випадках". Це є доповнення до означення функції переходу δ .

Результат показано на рис.9.

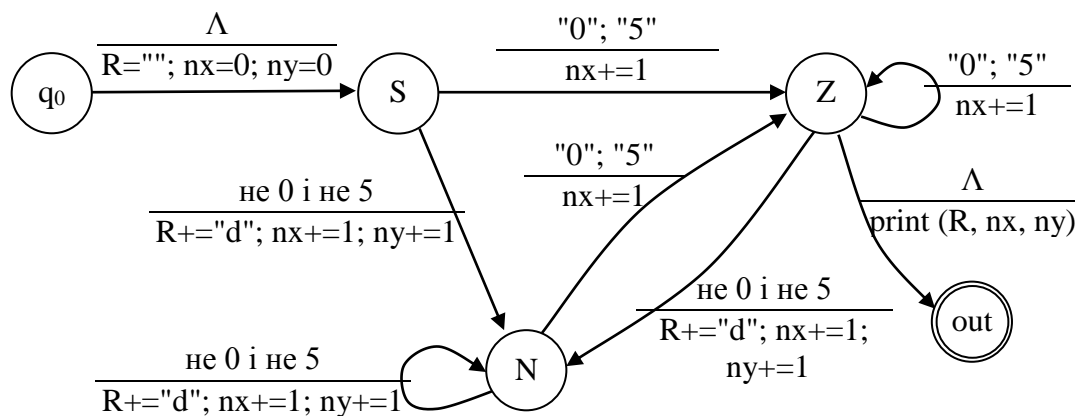


Рис.9

Табличний спосіб визначення автомата Мілі

Автомат Мілі типу розпізнавача

Нагадаємо, що детермінований скінченний автомат M – це п'ятірка $(Q, \Sigma, \delta, q_0, Z)$.

Крім діаграми, можна мати еквівалентне визначення, яке називають таблицею переходів. *Таблицю переходів* будують за такими правилами:

- 1) стовпці таблиці позначають літерами $\{a_i\}$, які можуть зустрітись в рядку символів;
- 2) рядки таблиці позначають станами з набору станів $\{q_i\}$;
- 3) на перетині рядка q_k і стовпця a_p записують стан, до якого перейде автомат з стану q_k при наявності літери a_p ;
- 4) перший рядок позначають початковим станом q_0 ;
- 5) рядки, які відповідають допустимим станам $Z \subseteq Q$, позначаємо праворуч одиницями, а всі інші – нулями;
- 6) незаповнені клітинки таблиці означають відсутність переходу (такі клітинки часом позначають знаком Λ "пусто").

Покажемо таблицю переходів для задачі наведеного вище приклада: "Побудувати автомат M типу розпізнавача, який приймає цілі числа без знаку, які діляться на 5." Праворуч від таблиці подано для порівняння автомат як діаграму (рис.8 копія).

		0; 5	не 0 і не 5	
$q_0 \rightarrow$	S	Z	N	0
	N	Z	N	0
	Z	Z	N	1

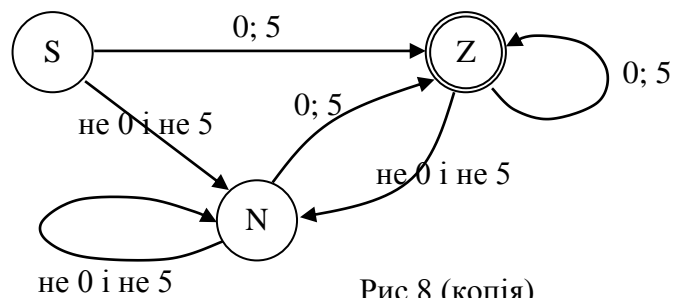


Рис.8 (копія)

Автомат Мілі типу перетворювача

Загальні правила будови таблиці переходів такі самі, як для автомата-розпізнавача. Проте, тепер переходи між станами супроводжуються операціями. Отже, в таблиці записуємо пари "куди перейти / операції", розділяючи їх горизонтальною лінією. Незаповнені клітинки таблиці означають відсутність переходу.

Таблицю переходів будуємо для наведеної вище задачі: "Задане число складається з окремих цифр $N=d_1d_2\dots d_n$. 1) Утворити з нього нове число R шляхом викреслення цифр 0 і 5. 2) Порахувати кількість цифр nx заданого числа і кількість цифр ny нового числа."

Поряд від таблиці подано для порівняння автомат як діаграму (рис.9 копія).

	0; 5	не 0 і не 5	Λ	
$q_0 \rightarrow$			S $R=""$; $nx=0$; $ny=0$	0
S	Z $nx+=1$	N $R+="d"$; $nx+=1$; $ny+=1$		0
N	Z $nx+=1$	N $R+="d"$; $nx+=1$; $ny+=1$		0
Z	Z $nx+=1$	N $R+="d"$; $nx+=1$; $ny+=1$	out $print(R, nx, ny)$	0
out				1

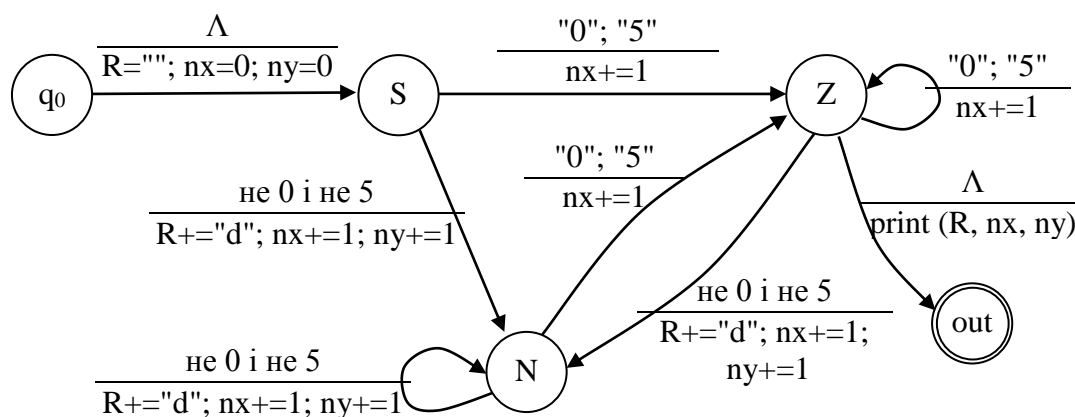


Рис.9 (копія)

Посилання за матеріалами теми

https://uk.wikipedia.org/wiki/Скінченний_автомат

- скінченний автомат — особливий різновид автомата — абстракції, що використовується для опису шляху зміни стану об'єкта в залежності від поточного стану та інформації отриманої ззовні

https://uk.wikipedia.org/wiki/Абстрактний_автомат

- теоретична модель апаратної або програмної обчислювальної системи, побудована на основі теорії автоматів

https://uk.wikipedia.org/wiki/Теорія_автоматів

- логіко-математична теорія, об'єктом дослідження якої є абстрактні автомати — покрокові перетворювачі інформації

https://uk.wikipedia.org/wiki/Детермінований_скінченний_автомат

- детермінований скінченний автомат (ДСА) — скінченний автомат, який приймає скінченний рядок символів; детермінованість означає наявність лише одного результату (тобто переходу в наступний стан для кожного символу ($S_0 \rightarrow S_i$) або повернення в той самий стан ($S_0 \rightarrow S_0$))

https://uk.wikipedia.org/wiki/Недетермінований_скінченний_автомат

- недетермінований автомат — абстрактний автомат, який при даному вхідному символі і внутрішньому стані може переходити в декілька різних внутрішніх станів

https://uk.wikipedia.org/wiki/Абстрактна_теорія_автоматів

- абстрактна теорія автоматів — напрям в теорії автоматів, який характеризується тим, що при вивченні автоматів не беруть до уваги їх структурні особливості