الگوریتم یک جواب متنهای به تعداد نامتناهی سوال است.

مدارهای بولینی (Boolean circuit) می توانند فقط توابع متناهی (محدود) را محاسبه کنند. بنابراین آنها برای تعریف الگوریتم به عنوان یک تعریف واحد (single recipe) که توابع نامتناهی (نامحدود) را محاسبه کنند، مناسب نیستند.

ما در اینجا به سراغ اینکه چه چیزی را (What) باید محاسبه کنیم میرویم و اینکه چجوری (How) آن را محاسبه کنیم را در فصلهای بعد بررسی میکنیم.

: FUNCTIONS WITH INPUT OF UNBOUNDED LENGTH - 6.1

تعریف تابع XOR با ورودی نامحدود (نا متناهی):

$$XOR(x) = \sum_{i=0}^{|x|-1} x_i , x \in \{0,1\}^*$$

ما در این قسمت به سراغ توابعی به صورت $\{0,1\}^* \to \{0,1\}^* \to \{0,1\}^* \to \{0,1\}^*$ این به این معنی نیست که طول ورودی بی نهایت است، بلکه به این معنی است که طول ورودی می تواند متغیر باشد و هر مقداری باشد. (بنابراین نمی توان آن را به صورت یک جدول که برای هر ورودی یک خروجی را نشان می دهد، نوشت)

یک تابع به شکل $\{0,1\}^* \to \{0,1\}^* \to F$ یک تسک محاسباتی (computational task) یک تابع به شکل $F: \{0,1\}^* \to \{0,1\}^*$ یک تسک محاسباتی (pomputational task) یک تابع به شکل $X \in \{0,1\}^*$ به شکل $X \in \{0,1\}^*$ به شکل $X \in \{0,1\}^*$ به شکل نام دروجی تابع به خروجی تابع به خروجی تابع به نام دروجی تابع به خروجی به خروجی تابع به خروجی

هر چیزی می تواند به صورت رشته باینری (binary string) ذخیره شود.

تابع TWIN PRIME به صورت زیر تعریف می شود: (می گوید آیا در اعداد بزرگتر از x آیا دو عدد اول که با هم دو تا فاصله داشته باشند وجود دارد یا نه)

$$TWINP(x) = \left\{ egin{array}{lll} 1 & \exists_{p \in \mathbb{N}} & p > |x| & \varrho & p + 2 \\ 0 & & & \end{array}
ight.$$
 در غیر این صورت

این تابع به صورت ریاضی تعریف شدهاست و برای هر ورودی، یا مقدار صفر یا یک را اختیار می کند. این تعریف یک specification است. همچنین برنامه پایتونی که این را محاسبه کند وجود ندارد.

: Varying inputs and outputs - 6.1.1

خیلی از توابع، بیشتر از یک (دو یا بیشتر) ورودی دارند. مثل تابع ضرب که به صورت مقابل است: MULT(x,y)=x.y با توجه به اینکه ما می توانیم دوتایی (pair) از رشتههای باینری را به صورت یک رشته باینری در نظر بگیریم، پس میتوانیم بگوییم تابع ضرب یک نگاشت (map) از $\{0,1\}^* + \{0,1\}$ است. (یعنی: $\{0,1\}^* \to \{0,1\}^*$ است)

تابع پالیندورم (PALINDORM) به صورت زیر است:

$$PALINDORM(x) = \begin{cases} 1 & \forall_{i \in [|x|]} x_i = x_{|x|-i} \\ 0 & \text{equation} \end{cases}$$
در غیر این صورت

اینگونه توابع که فقط یک بیت خروجی دارند، توابع بولینی (Boolean Function) نامیده می شوند. (از انواع توابع نامحدود (نامتناهی) هم به شمار میروند طبیعتا چون ورودی شان، رشته ای با طول نامحدود است)

بولینی کردن توابع (Booleanizing" functions") یعنی تبدیل یک تابع به شکل تابع بولینی آن. مثلا یعنی برای تابع ضرب، تابع بولینی آن، تابعی است که خروجی، بیت i أم را می دهد. یعنی:

$$MULT(x,y,i) = \begin{cases} i^{th} \ bit \ of \ x.y & i < |x.y| \\ 0 &$$
در غیر این صورت

نحوه کلی بولینی کردن توابع:

اگر تابع $\{0,1\}^* o \{0,1\}^*$ داشته باشیم، تابع بولینی آن به صورت زیر است:

$$BF(x,i,b) = \begin{cases} F(x)_i & i < |F(x)| & , b = 0 \\ 1 & i < |F(x)| & , b = 0 \\ 0 & i < |F(x)| \end{cases}$$

دلیل وجود b در این تابع چیست؟ دلیل آن فقز برای این است که در برنامه پایتون متناظر با این کار، بتواند طول ورودی را تشخیص دهد. (اگر B را برابر با یک بگذاریم، خروجی این تابع از اندیس صفر تا طول ورودی برابر با یک است و در بقیه موارد برابر با صفر است. بدین شکل میتوان از این تابع برای محاسبه طول ورودی هم استفاده کرد)

: Formal Languages – 6.1.2

برای هر تابع بولینیای (Boolean Function) به صورت $\{0,1\}^* \to \{0,1\}^* \to \{0,1\}^*$ ما می توانیم مجموعهای از رشتهها (Boolean Function) که خروجی تابع به ازای آنها یک است را تعریف کنیم (تعریف ریاضی آن به صورت $\{x \mid F(x) = 1\}$ است) و به آن زبان (Language) می گوییم. یک زبان رسمی (formal language) یک زیرمجموعه از $\{0,1\}^* \to L$ است. (اگر بخواهیم به صورت جامعتر بر روی الفبای (محدود) مورد استفاده (کاراکترهای مورد استفاده که تعدادشان محدود است) تعریف کنیم، می شود $\{0,1\}^*$

L متعلق به زبان، در اصل تسک مشخص کردن اینکه ورودی $x \in \{0,1\}^*$ متعلق به زبان در اصل تسک مشخص کردن اینکه ورودی $x \in \{0,1\}^*$ متعلق به زبان $x \in L$) هست یا خیر می شود.

حالا اگر بتوانیم تابع F را محاسبه کنیم، میتوانیم مسئله تصمیم عضویت (آیا ورودی متعلق به زبان پذیرش تایع هست یا نه) را حل کنیم و برعکس. (اگر بتونیم مسئله تصمیم عضویت یک ورودی به زبان پذیرش تابع را حل کنیم، میتوانیم تابع را محاسبه کنیم)

: Restriction of Function - 6.3.1

 F_n اگر تابع F_n بر روی طول ورودی برابر با $n \in \mathbb{N}$ باشد و $n \in \mathbb{N}$ باشد، آنگاه محدودشده تابع $f:\{0,1\}^* \to \{0,1\}^* \to \{0,1\}^*$ یک تابع بولینی باشد و به صورت $f:\{0,1\}^n \to \{0,1\}^n$ است به طوریکه برای هر رشته ای مانند x به طول x مقدار توابع نشان داده می شود، یک تابع محدود است که به صورت $x \in \{0,1\}^n$ باشد به ازای تمامی $x \in \{0,1\}^n$ باشد به ازای تمامی با باشد به ازای تمامی باشد با باشد با بازای تمامی بازای بازای تمامی بازای باز

: Circuit collection for every infinite function – Theorem 6.1

اگر تابع $\{C_n\}_{n\in\{1,2,...\}}^* o \{C_n\}_{n\in\{1,2,...\}}$ او تابع $F:\{0,1\}^* o \{0,1\}$ وجود دارد که به ازای هر $F:\{0,1\}^* o \{0,1\}$ وجود دارد که به ازای هر C_n محدود شده تابع F را بر روی ورودیهایی به طول C_n محاسبه می کند.

یعنی برای هر تابع نامحدود میتوان توابع محدود شده با طول مشخص وروی برایش آورد.

است. (چون F_n یک نگاشت: این قضیه، یک نتیجه بدیهی از قضیه جامعیت مدارهای بولینی (Universality of Boolean circuit) است. (چون F_n یک نگاشت $c.\frac{2^n}{n}$ است، پس یک تابع بولینیای وجود دارد که آن را محاسبه کند. در حقیقت سایز آن مدار حداکثر (mapping) است)

طبق این قضیه گفته شده، حتی برای تابع TWINP هم یک همچین مجموعه C_n ای وجود دارد؛ حتی با وجود اینکه نمیدانیم آیا برای این تابع، برنامهای (program) ای (به زبان پایتون یا ...) وجود دارد یا خیر. مشکل آن هم این است که برای هر (همه) مقدار n ای نمیدانیم که حلقه (circuit) موجود در C_n چیست. (برای یک مقدار میدانیم)

: DETERMINISTIC FINITE AUTOMATA -6.2

با DFA ها از قبل آشنا هستيم. اين DFA ها در توان محاسباتي برابر هستند با regular expressions. اين Regular expressions ها يك مكانيزم قوى براي تشخيص الگو (pattern) هستند.

در سطحهای بالا (high level)، الگوریتم یک دستور العمل (recipe) رسیدن به خروجی از روی یک ورودی با انجام ترکیبی از قدمهای زیر است:

۱- خواندن یک بیت از ورودی ۲- بروزرسانی کردن حالت (state) کنونی ۳- تموم کردن و تولید خروجی

تعریف single-pass constant-memory algorithm تعریف

یک بار ورودی را پیمایش می کنیم (single pass over the input) و همچنین حافظه ای (مموری ای) که با آن کار می کند، محدود DFA است. این الگوریتم به DFA یا همان Deterministic Finite Automata هم معروف است. (یک نام دیگر برای این DFA ها، DFA ها، DFA هم DFA می STA ها، STA ها، STA ها،

این الگوریتمها را می توان مانند ماشینی که C تا وضعیت (state) دارند هم در نظر گرفت که در یک حالت شروعی قرار دارد و شروع به خواندن ورودی این الگوریتمها می کند و وضعیت (state) خود را بروزرسانی می کند. این بروزرسانی با توجه به وضعیت قبلی و ورودی انجام می شود) خروجی این ماشین نیز وابسته به وضعیتی است که در انتها در آن قرار می گیرد.

اگر الگوریتم c تا بیت از حافظه را استفاده کند، در آن صورت محتوای داخل آن در یک رشته به طول c قابل نمایش است. بنابراین این الگوریتم میتواند حداکثر در یکی از c تا وصعیت (state) مختلف در هر مرحله از اجرا قرار بگیرد.

از این رو، ما می توانیم یک DFA که DFA که DFA تا وضعیت دارد را با لیستی از C . D تا قاعده (قانون یا rule) مشخص کنیم که به آن قاعده حرکت (c این رو، ما می توانیم یک V (state) هم گفته می شود. (هر قاعده به این صورت است که اگر در وضعیت V (state) هم گفته می شود. (هر قاعده به این صورت است: اگر در وضعیت نهایی در یکی از وضعیتهای وضعیت از وضعیت از

ما می توانیم یک DFA با C با DFA با یک گراف لیبل دار (state) هم توصیف کنیم. (برای هر وضعیت (state) یک خروجی به ازای هر عضو Σ مثل σ قرار می دهیم. همچنین یک سری از گرهها را هم با عنوان گره پایانی قابل قبول (accepting states) مشخص می کنیم.)

:Deterministic Finite Automata – Definition – 6.2

Definition 6.2 — **Deterministic Finite Automaton.** A deterministic finite automaton (DFA) with C states over $\{0,1\}$ is a pair (T,\mathcal{S}) with $T:[C]\times\{0,1\}\to[C]$ and $\mathcal{S}\subseteq[C]$. The finite function T is known as the *transition function* of the DFA. The set \mathcal{S} is known as the set of accepting states.

Let $F:\{0,1\}^* \to \{0,1\}$ be a Boolean function with the infinite domain $\{0,1\}^*$. We say that (T,\mathcal{S}) computes a function $F:\{0,1\}^* \to \{0,1\}$ if for every $n\in\mathbb{N}$ and $x\in\{0,1\}^n$, if we define $s_0=0$ and $s_{i+1}=T(s_i,x_i)$ for every $i\in[n]$, then

 $s_n \in \mathcal{S} \Leftrightarrow F(x) = 1$

توجه داشته باشید که تابع حرکت (transition function) یک تابع محدود است که با یک جدول از قواعد (table of rules) قابل نمایش است.

چندین روش برای تعریف DFA ها وجود دارد. مثلا می توان DFA را با یک پنج تایی $(Q, \sum, \delta, q_0, F)$ تعریف کرد. که در آن Q همان مجموعه وضعیتها، Z همان الفبا (alphabet) است، δ همان تابع حرکت (transition function)، وضعیتهای پذیرش (accepting state) است.

(هر وقت که از ما خواستند یک DFA بکشیم، کمک کننده ست که با کشیدن یک single-pass constant-memory شروع کنیم و الگوریتم آن را بدهیم و این کار را با روشهای کلی تر مصل شبه کد یا کد پایتون یا ... انجام دهیم.)

Figure -6.4 : جدولی که برای کشیدن تابع حرکت (transition function) است، شامل سه ستون است که ستون اول وضعیت (state) کنونی، ستون دوم حرف الفبار ورودی و ستون سوم وضعیت جدید است.

: Anatomy of an automaton (finite vs. unbounded) -6.2.1

کمیتهایی با سایز محدود در DFA ها که به سایز ورودی وابسته نیستند:

• تعداد وضعیتها • تابع حرکت (Transition function) که C تا ورودی دارد و مجموعه C تابع حرکت (Transition function) هی تابع حرکت (State) تا عداد وضعیتهای می تواند با جدولی با تعداد سطر C مدل شود که خروجی هر کدام از این سطرها یک عدد در C است کدام از این سطرها یک عدد در C است کدام از این سطرها یک عدد در C است کدام از این سطرها یک عدد در C است کدام از این سطرها یک عدد در C است کدام از این سطرها یک عدد در C است کدام از این سطرها یک عدد در C است کدام از این سطرها یک عدد در C است کدام از این سطرها یک عدد در C است کدام از این سطرها یک عدد در C است کدام از این سطرها یک عدد در C است کدام از این سطرها یک عدد در C است کدام از این سطرها یک عدد در C است کدام از این سطرها یک کدام از این سطرها یک عدد در C است کدام از این سطرها یک کدام از

پس فارغ از توضيحات داده شده در رابطه با الگوريتمها، بايد بتوان توصيف (description) يک DFA را به صورت كامل نوشت.

کمیتهایی با سایز نامحدود در DFA که به هیچ عدد ثابتی محدود نمی شوند:

عداد مراحل (step) هایی که DFA طی می کند که می تواند با توجه به اینکه
 توجه به طول ورودی رشد کند. در اصل با توجه به اینکه single-pass
 است، پس تعداد مراحل، دقیقا برابر با سایز ورودی است.

:DFA-Computable functions -6.2.2

. تابع $F:\{0,1\}^* o DFA$ است اگر یک DFA است اگر یک محاسبه با محاسبه کند. قابل محاسبه کند

: DFA computable functions are countable - Theorem 6.4

اگر DFACOMP مجموعه تمام تابعهای بولینی مانند $F:\{0,1\}^* o \{0,1\}^*$ باشد به طوریکه یک DFA ای وجود داشته باشد که $F:\{0,1\}^*$ محاسبه کند. آن وقت DFACOMP قابل شمارش (countable) است.

onto) ایده اثبات: هر DFA را می توان با یک رشته با طول محدود (finite length string) نمایش داد. که منجیر به یک تناظر پوشا (map بین *DFA را با تابعی که آن DFA آن را محاسبه می کند. (map بین *DFA را با تابعی که آن DFA

(accepting states) آن و استیتهای پذیرش (transition function) آن و استیتهای پذیرش (DFA ای را میتوان با یک رشته که شامل تابع حرکت (TFA یک تابع $T:\{0,1\}^* \to \{0,1\}^* \to DFA$ را محاسبه میکند. بنابراین ما تابع آن است، نمایش داد از طرفی هر TFA یک تابع TFACOMP را به صورت زیر تعریف کنیم:

$$StDC(a) = egin{cases} F & ext{ solution solution} \ ONE & ext{ constant} \end{cases}$$
 در غیر این صورت

تابع ONE تابع ثابت یک است. حالا با توجه به اینکه $ONE: \{0,1\}^* \to \{0,1\}^* \to ONE$ تابع ثابت یک است. حالا با توجه به اینکه هر تابع ONE تابع ثابت یک است که دامنه آن ONE با اتوماتایی قابل محاسبه است، پس ONE یک تابع پوشا (onto) است که دامنه آن ONE و برد آن ONE است. که یعنی ONE قابل شمارش (Countable) است.

حالا با توجه به اینکه تمام توابع بولینی غیرقابل شمارش (Uncountable) هستند، پس میتوانیم نتیجه بدیهی زیر را بگیریم:

: Existence of DFA-uncomputable functions - Theorem 6.5

تابع بولینی وجود دارد که با هیچ DFA ای قابل محاسبه نیست.

ALL ها بودند، آن وقت DFA مساوی مجموعه DFA مساوی مجموعه DFA ما این مجموعه ناشماراست (uncountable) است و با قضیههای قبلی در تضاد است.

: REGULAR EXPRESSIONS -6.3

ما مى توانيم به كاربر اجازه دهيم كه الگو را با تعين كردن يک تابع قابل محاسبه (computable function $F:\{0,1\}^* \to \{0,1\}$) مشخص كند؛ كه اين تابع F(x)=1 متناظر با پيدا كردن (matching) باشد.

ما نمیخواهیم اجازه بدهیم که در حلقه بینهایت (infinite loop) قرار بگیریم، برای همین به کاربر اجازه استفاده از ابزارهای کامل زبان برنامهنویسی را نمیدهند.

در regular expression ها دو عمليات | که برابر با چسباندن (concatenation) و * که نمايانگر تکرار است را داريم. بقيه موارد عموما با syntactic sugar ها از همينا بدست ميان.

: Regular Expression - Definition 6.6

یک regular expression در الفبای (alphabet) کی از فرمهای زیر است: Σ U {(,), |,*, Ø, ""}

- اگر σ عضوی از $e = \sigma$ -
- باشند regular expression و e'' و و e'' و الحي که e = (e'|e'') -
- در حالی که ' e و " e خودشان regular expression باشند (عموما پرانتز را نمی گذاریم) e در حالی که ' e و " e خودشان e اللی که ' e در حالی که ' e خودشان e خودشان e در حالی که ' e در حالی که ' e خودشان e خودشان e در حالی که ' e خودشان e در حالی که نمونی خودشان e در حالی خودشان e در خودشان
 - باشدپ regular expressions باشدپ e' به طوریکه e' به طوریکه e'

no) مربوط به پذیرش هیچ رشتهای ($e=\emptyset$) مربوط به پذیرش هیچ رشتهای ($e=\emptyset$) مربوط به پذیرش هیچ رشتهای (empty string) است.

بالاترين اولويت با * است، بعدش با چسباندن (concatenation) و بعدش با |. همچنين عملگر يا left associative است.

هر regular expression ای متناظر با یک تابع $\phi_e(x)=1$ است به طوریکه $\phi_e(x)=1$ است به طابق (match) باشد با regular expression

: Matching a regular expression - Definition 6.7

یک تابع بولینی (regular (Boolean function) نامیده میشود اگر خروجیاش دقیقا در مجموعهای از رشتهها که با یک regular یک تابع بولینی (match expression شوند، یک باشد.

: Regular functions/languages - Definition 6.8

: Algorithm for matching regular expression -6.3.1

: Regular Expression matching - Algorithm 6.10

```
Algorithm 6.10 — Regular expression matching.
Input: Regular expression e over \Sigma^*, x \in \Sigma^*
Output: \Phi_e(x)
 1: procedure MATCH(e,x)
        if e = \emptyset then return 0;
        if x = "" then return MATCHEMPTY(e);
        if e \in \Sigma then return 1 iff x = e;
        if e = (e'|e'') then return Match(e', x) or Match(e'', x)
        if e = (e')(e'') then
            for i \in [|x|] do
               if Match(e', x_0 \cdots x_{i-1}) and Match(e'', x_i \cdots x_{|x|-1})
    then return 1:
            end for
10:
      if e = (e')^* then
11:
            if e' = "" then return MATCH("", x);
12:
13:
                                          # ("")* is the same as ""
                                    # x_0 \cdots x_{i-1} is shorter than x
15:
               if Match(e, x_0 \cdots x_{i-1}) and Match(e', x_i \cdots x_{|x|-1})
           end for
18:
        return 0
20: end procedure
```

e' نکته کلیدی این است که در تعریف بازگشتی regular expression (recursive) ها، هر وقت e از دو تا regular expression مثل e'' و e'' تشکیل شده بود، آن وقت e'' و e'' و e'' مستند. و وقتی که سایز آنها به یک برسد، باید مطابق با یکی از حالتهای غیربازگشتی باشند. بنابراین با توجه به این نکته، فراخوانیهای بازگشتی الگوریتم بالا همیشه منجر به کوچکتر شدن e'' منجر به کوچکتر شدن طول ورودی می شود.

: Matching Empty String - Solved Exercise 6.3

مى توانيم براى اين كاريك الگوريتم بازگشتى با مشاهدات زير پيدا كنيم:

- را عبارتی به شکل "" یا $(e')^*$ همواره رشته خالی را پذیرش (match) می کند.
- را پذیرش نمی کند. (match) مبارتی به فرم $\sigma \in \Sigma$ در حالیکه $\sigma \in \Sigma$ باشد (در الفبای ما باشد)، هیچگاه رشته خالی $\sigma \in \Sigma$
 - رشته خالی را پذیرش (match) ای با فرم \emptyset رشته خالی را پذیرش (Regular expression .۳
- عبارتی به فرم "e'|e'' ، رشته خالی را پذیرش می کند (match) اگر و تنها اگر e'' یا e'' رشته خالی را پذیرش کنند.
- عبارتی به فرم (e')(e'') شته خالی را پذیرش می کند اگر و تنها اگر هر دو عبارت e'' و e'' شته خالی را پذیرش کنند.

```
Algorithm 6.11 — Check for empty string.

Input: Regular expression e over \Sigma^*, x \in \Sigma^*
Output: 1 iff e matches the emptry string.

1: procedure Matchempty(e)

2: if e = " then return 1;

3: if e = \emptyset or e \in \Sigma then return 0;

4: if e = (e'|e'') then return Matchempty(e') or Matchempty(e'');

5: if e = (e')(e'') then return Matchempty(e') and Matchempty(e'');

6: if e = (e')^* then return 1;

7: end procedure
```