

2.2 Проблемът за безквадратните и безкубните думи

Една от първите задачи в теорията на формалните езици е поставена от Аксел Туе в началото на нашия век и е известна като *проблем за безквадратните и безкубните думи*.

Всяка безкрайна редица от букви на азбуката V ще наричаме *свърхдума* над V .

Ще казваме, че думата (или свърхдумата) α е *безквадратна*, ако тя няма поддума от вида x^2 , където x е дума над V , различна от ε . Аналогично α е *безкубна*, ако няма поддума от вида x^3 , където x е непразна дума над V .

Думата (или свърхдумата) α ще наричаме *силно безкубна*, ако не съдържа поддума от вида x^2a за някое $x \in V^*$ и a е първата буква на x . Очевидно безквадратните думи (или свърхдуми) са силно безкубни и силно безкубните са безкубни думи.

Проблемът на Туе се състои в построяване на безквадратна дума с максимална дължина, а ако е възможно безквадратна свърхдума над дадена азбука V . Ако това е невъзможно, то същата задача стои за силно безкубни или безкубни думи или свърхдуми. Очевидно ако V е азбука само от една буква, то всяка дума с дължина по-голяма от 2 не е безкубна.

За азбуките с две букви е в сила следната лема.

Лема 2.2.1 *Нито една дума с дължина по-голяма от 3 над азбуката $V = \{a, b\}$ не е безквадратна.*

Доказателство: В V^* има само две думи с дължина 3, които са безквадратни

$$aba \text{ и } bab,$$

тъй като всички други думи с дължина 3 са: $a^3, a^2b, ba^2, b^3, b^2a, ba^2$ и те не са безквадратни. Също така, ако към горните думи добавим коя да е буква, ще получим дума, която съдържа една от поддумите $a^2, b^2, (ab)^2$ или $(ba)^2$. \square

За азбуки с две букви обаче се оказва, че проблемът за силно безкубните думи има решение, а именно, че над всяка

такава азбука съществува силно безкубна свръхдума. От това твърдение следва, че над такава азбука ще съществува и силно безкубна дума с произволна дължина.

Нека $V = \{a, b\}$ е произволна азбука от две букви. За да построим силно безкубна свръхдума над V , ще използваме пораждането на формален език $L = L(h, a)$, чрез хомоморфизма $h : V^* \rightarrow V^*$, за който $h(a) = ab$ и $h(b) = ba$. Думите в този език са:

$$\begin{aligned} w_0 &= h^0(a) = a \\ w_1 &= h(a) = ab \\ w_2 &= h^2(a) = abba \\ &\vdots \\ w_n &= h^n(a) \\ &\vdots \end{aligned}$$

Характерно свойство на думите в езика L е, че те могат да се получат рекурентно посредством равенството

$$(1) \quad w_{i+1} = w_i w'_i,$$

където, чрез x' е означена думата, получена от x , като всички a в x са заменени с b и обратно всички b са заменени с a . Да забележим, че $h(x') = (h(x))'$.

Равенството (1) ще докажем индуктивно.

При $i = 1$ то е очевидно.

Да допуснем, че (1) е вярно за $i = k$ и да го докажем при $i = k + 1$.

$$\begin{aligned} w_{k+1+1} &= w_{k+2} = h(w_{k+1}) = h(w_k w'_k) = h(w_k) h(w'_k) = \\ &= w_{k+1} h(w'_k) = w_{k+1} (h(w_k))' = w_{k+1} w'_{k+1}. \end{aligned}$$

Да разгледаме свръхдумата

$$\alpha = a \ b \ ba \ baab \ baababba \ \dots$$

като празното място между буквите е оставено само за да стане ясно къде се добавя w'_i . Очевидно началото на α до i -тото

празно място е думата w_{i-1} , а поддумата на α от i -тото празно място до $i+1$ -вото празно място, е думата w'_{i-1} .

За да докажем, че α е силно безкубна свърхдума, ще изучим някои от свойствата на α .

Да въведем за буквите на α по общото означение $\alpha = c_1 c_2 \dots$, където c_i – е или a или b , съгласно по-горното равенство, с което се дефинира α .

Лема 2.2.2 *Думите a^3 , b^3 , $ababa$, $babab$ не са поддуми на α . Всяка поддума x на α , за която $|x| = 5$ има за поддума или a^2 или b^2 .*

Доказателство: Ако допуснем, че a^3 или b^3 са поддуми на α , то съществува минимално i , $i \geq 1$ така, че a^3 или b^3 са поддуми на w_i . Но това е невъзможно, тъй като $w_i = h(w_{i-1})$ и буквите на w_i се получават от дописването на думите ab и ba към w_{i-1} . Да допуснем, че $ababa$ е поддума на α , като

$$c_j c_{j+1} c_{j+2} c_{j+3} c_{j+4} = ababa.$$

Нека $w_i \in L$ е такава, че $ababa$ е поддума на w_i , т.е. $|w_i| \geq j+4$.

Ако j е нечетно, то съгласно $w_i = h(w_{i-1})$ от $ababab = h(a)h(a)h(a)$ получаваме, че a^3 е поддума на w_{i-1} . Аналогично при j -четно се получава, че b^3 е поддума на w_{i-1} , което е противоречие. Следователно $ababa$ не е поддума на α .

По същия начин се доказва, че $babab$ не може да е поддума на α .

Очевидно единствените думи над V с дължина 5, които не съдържат нито a^2 нито b^2 са $ababa$ и $babab$, и понеже тези думи не са поддуми на α следва, че всяка поддума x на α , за която $|x| = 5$ съдържа като поддума a^2 или b^2 . \square

Лема 2.2.3 *Ако a^2 или b^2 са поддуми на α и началото на такива поддуми е от j -та позиция в α , то j е четно.*

Доказателство: Да допуснем, че

$$a^2 = c_j c_{j+1} \text{ или } b^2 = c_j c_{j+1}.$$

Да изберем i толкова голямо, че $|w_i| > j + 1$. От $w_i = h(w_{i-1})$ следва, че ако j е нечетно, то $c_j c_{j+1} = h(a)$ или $c_j c_{j+1} = h(b)$, което е невъзможно. Следователно j е четно. \square

Теорема 2.2.1 *Свърхдумата α е силно безжубна.*

Доказателство: Да допуснем, противното. Нека x е поддума на α с минимална дължина и s е първата буква на x , така че xxs е поддума на α .

Съгласно Лема 2.2.2 $|x| > 1$. Наистина, ако $|x| = 1$, то $x = a$ или $x = b$. От това следва, че a^3 или b^3 биха били поддуми на α , което противоречи на Лема 2.2.2.

Ако $|x| = 2$, то от Лема 2.2.2 следва, че $xxs = ababa$ или $xxs = babab$. Пак от Лема 2.2.2 следва, че и този случай е невъзможен.

Следователно $|x| = t \geq 3$.

Да допуснем, че xxs влиза в α започвайки от j -позиция, т.е.

$$(2) \quad xxs = c_j c_{j+1} \dots c_{j+2t}.$$

Понеже $|xxs| > 5$, то по Лема 2.2.2 следва, че a^2 или b^2 е поддума на xxs . Нека например a^2 е поддума на xxs . Ако a^2 е поддума на x , то a^2 два пъти се съдържа в xxs . Ако a^2 не е поддума на x , то x започва и извършва с a , и следователно a^2 пак се съдържа два пъти в xxs .

От това и от Лема 2.2.3 ще получим, че t е четно. Наистина да допуснем противното и нека първото влизане на a^2 в xxs започва от k -та буква на α , т.е.

$$c_k c_{k+1} = a^2, \quad j \leq k \leq j + t - 1.$$

Но тогава от $c_{k+t} c_{k+t+1} = c_k c_{k+1}$ следва, че $c_{k+t} c_{k+t+1} = a^2$ започва от $k + t$ -та буква на α , което противоречи на Лема 2.2.3. Следователно $t = 2u \geq 4$.

Да разгледаме най-напред случая, когато j в (2) е четно. Да изберем i толкова голямо, че $|w_i| > j + 2t$, т.е. така, че

$$xxc = c_j c_{j+1} \dots c_{j+t} \dots c_{j+2t-1} c_{j+2t}$$

е поддума на w_i . Понеже $w_i = h(w_{i-1})$, то следва, че

$$c_{j-1} c_j = h(a) \text{ или } c_{j-1} c_j = h(b),$$

както и

$$c_{j+t-1} c_{j+t} = h(a) \text{ или } c_{j+t-1} c_{j+t} = h(b)$$

понеже t е четно. От $c_j = c_{j+t}$, следва че $c_{j-1} = c_{j+t-1}$.

Следователно, за всяко n , $j-1 \leq n \leq j+t$ е в сила $c_n = c_{n+t}$.

Да положим $l = \frac{j}{2}$. Тогава, за всяко r , $0 \leq r \leq t$,

$$c_{j-1+2r} c_{j+2r} = h(d_{r+l}),$$

където d_{r+l} е или a или b , и d_{r+l} е буква от w_{i-1} . Следователно $c_{j-1} xxc$ е образ на поддумата $d_l d_{l+1} \dots d_{l+t}$ на w_{i-1} при h . Поради $c_{j-1} = c_{j-1+t}$ следва, че за всяко s , $0 \leq s \leq u$

$$d_{s+l} = d_{s+u+l}.$$

Ако положим $y = d_l d_{l+1} \dots d_{l+u-1}$, то тогава думата yud , където $d = d_l$ е първата буква на y , е поддума на w_{i-1} и следователно на α . Поради $|y| = u < t$ достигаем до противоречие с избора на x . Следователно в този случай xxc не е поддума на α . В случая, когато j е нечетно по аналогичен начин се доказва, че поддумата на w_i

$$c_j c_{j+1} \dots c_{j+t-1} c_{j+t} \dots c_{j+2t-1} c_{j+2t} c_{j+2t+1} = xxc c_{j+2t+1}$$

е образ при h на поддумата на w_{i-1}

$$d_l d_{l+1} \dots d_{l+u-1} d_{l+u} d_{l+u+1} \dots d_{l+2u} = yud$$

при $l = \frac{j+1}{2}$ и d е първата буква на y . Понеже $|y| = u$ отново достигаем до противоречие с избора на x , и в този случай xxc не е поддума на α . \square

По-нататък, ще докаже, че над азбука с четири букви може да се построи безквадратна свърхдума. За целта отново ще използваме свърхдумата α от предишната теорема.

Теорема 2.2.2 *Съществува безквадратна свърхдума над азбука от четири букви.*

Доказателство: Нека $V_1 = \{[aa], [ab], [ba], [bb]\}$.

Конструираме следната свърхдума над V_1

$$\beta = d_1 d_2 d_3 \dots$$

като

$$d_i = [c_i c_{i+1}], \quad i = 1, 2, \dots,$$

където $c_i c_{i+1}$ е поддума на α започваща от i -та позиция в α . Началото на β изглежда, така

$$\beta = [ab][bb][ba][ab][ba][aa][ab] \dots$$

Да допуснем, че β не е безквадратна и нека y^2 е поддума на β . Да положим

$$y = d_{j+1} d_{j+2} \dots d_{j+t} = d_{j+t+1} d_{j+t+2} \dots d_{j+2t},$$

за някое t , $t \geq 1$. Следователно

$$\begin{aligned} & [c_{j+1} c_{j+2}] [c_{j+2} c_{j+3}] \dots [c_{j+t} c_{j+t+1}] = \\ & = [c_{j+t+1} c_{j+t+2}] [c_{j+t+2} c_{j+t+3}] \dots [c_{j+2t} c_{j+2t+1}]. \end{aligned}$$

Ето защо $c_{j+1} = c_{j+t+1} = c_{j+2t+1}$ и за всяко r , $1 \leq r \leq t$

$$c_{j+r} = c_{j+r+t}.$$

Това значи, че думата

$$c_{j+1} c_{j+2} \dots c_{j+t} c_{j+1} c_{j+2} \dots c_{j+t} c_{j+1} = xxc,$$

със $c = c_{j+1}$, е поддума на α , което е противоречие с предходната теорема. \square

За окончателното решаване на проблема на Туе е необходимо да установим, че съществува безквадратна свърхдума над азбука от три букви. Тук отново ще използваме свърхдумите α и β от предходните две теореми.

Теорема 2.2.3 *Съществува безквадратна свърхдума над азбука от три букви.*

Доказателство: Преди да конструираме необходимата ни свърхдума, да отбележим някои от свойствата на свърхдумата β . За целта нека да опростим означението на буквите в β , чрез следните полагания

$$[aa] = 1, [ab] = 2, [ba] = 3, [bb] = 4.$$

Тогава началото на думата β , в тази нова азбука $V' = \{1, 2, 3, 4\}$ изглежда по следния начин

$$\beta' = 243231243123243231232431\dots$$

От горното полагане следва, че след 1 в β' могат да бъдат само 1 или 2. Ако 1 следва след 1, то тогава $[aa][aa]$ е поддума на β и следователно $aaaa = a^4$ е поддума на α , което е невъзможно. Следователно, след 1 в β' може да следва само 2. От друга страна 1 в β' може да се предшества единствено от 1 или 3. Както по-горе се установява, че 1 не може да предшества 1 в β' и следователно 1 може да следва в β' само след 3.

С аналогични разсъждения се доказва, че 4 може да се предшества в β' само от 2 и след 4 може да се появява само 3.

Да разгледаме азбуката $V_2 = \{1, 2, 3\}$.

Нека γ е свърхдума над V_2 , получена от β' , като в нея навсякъде символът 4 се замени с 1. Така началото на γ изглежда по следния начин:

$$\gamma = 2132312131232132312321\dots$$

Ще покажем, че γ е безквадратна свърхдума.

Да допуснем, че γ не е безквадратна. Следователно съществува дума $x \in V_2^*$ така, че x^2 е поддума на γ и $|x| \geq 1$. Тогава съществуват поддуми y_1 и y_2 на β' така, че $|y_1| = |y_2| = |x| = t$ и ако

$$y_1 = d_{j+1}d_{j+2}\dots d_{j+t}, \quad y_2 = d_{j+t+1}d_{j+t+2}\dots d_{j+2t},$$

то y_1 и y_2 се различават евентуално само в буквите d_i и d_{i+t} когато те са 1 или 4. Ако в тях заменим 4 с 1, то y_1 и y_2 ще съвпадат. Ако $t = 1$, то понеже 11, 14, 41 и 44 не са поддуми на β , следва, че $y_1 = y_2$ и от Теорема 2.2.2 достигаем до $t \geq 2$.

Да допуснем, без ограничение на общността, че за някое r , $1 \leq r < t$ е в сила $d_{j+r} = 1$ и $d_{j+t+r} = 4$.

Съгласно направените по-горе разсъждения следва, че $d_{j+r+1} = 2$ и $d_{j+t+r+1} = 3$, което е противоречие. Следователно за всяко такова r ще е в сила $d_{j+r} = d_{j+t+r}$.

Ако приемем, че $d_{j+t} = 1$, а $d_{j+2t} = 4$, то тогава d_{j+t-1} трябва да е 3, а d_{j+2t-1} да е 2, което отново е противоречие. Следователно y_1 и y_2 съвпадат, като поддуми в β' , което противоречи на Теорема 2.2.2. \square

З а д а ч и

1. Да изменим определението за силно безкубна дума по следния начин. Думата (или свърхдумата) α ще наричаме силно безкубна, ако не съдържа поддума от вида sxx , където $|x| \geq 1$ и s е последната буква на x . Да се покаже, че при това определение съществува силно безкубна свърхдума над азбука от две букви.
2. Нека $w = a_0a_1a_2\dots$ е свърхдума над азбуката $\{a, b\}$, за която

$$a_n = \begin{cases} a, & \text{ако броят на единиците в} \\ & \text{двоичното представяне на } n \text{ е четно;} \\ b, & \text{в останалите случаи.} \end{cases}$$

Да се докаже, че w съвпада със свърхдумата α , която се използва в Теорема 2.2.1.

3. Да наречем *двойно безкрайни* думите от вида

$$\alpha = \dots c_{-2}c_{-1}c_0c_1c_2\dots$$

Да се намерят примери за безквадратни двойно безкрайни думи над азбука от три букви и примери за силно безкубни двойно безкрайни думи над азбука от две букви.