Лемма о разрастании

Также называется «леммой о накачке» - **pumping lemma** и является важной леммой в теории автоматов позоляющая проверить (очень часто) если определённый язык является или нет автоматным или нерегулярным.

Термин накачка (разрастание), в названии леммы отражает возможность многократного повторения некоторой подстроки в любой строке, подходящей длины любого бесконечного автоматного языка.

Лемма о разрастании

Лемма: Для любого регулярного языка L существует константа n, так, чтобы для любого слова z из L длина которого больше или равно чем n, верны следующие свойства:

1. Слово z можно разбить на 3 подслова u, v, w;
2. Длина подслова uv должна быть меньше или равна чем n, а длина подслова v должна быть больше единицы, $|uv|<=n$, а $|v|>=1$;
3. Слово $uv^i w$ тоже принадлежит языку L, где i=0,1,2...

Лемма о разрастании

В качестве константы n возьмем количество состояний автомата КА. Пусть будет слово z из L и длина слова $z \geq n$. Если язык регулярный для него можно построить эквивалентный КА. Пусть будет детерминированный КА. Это означает что слово z проходит через КА, то есть $(q_0, z) \vdash^* (q_f, \varepsilon)$. Напишем состояния через которые проходит автомат при проверки слова z:

$q_0, q_1, q_2, \ldots, q_j, q_{j+1}, \ldots, q_k, q_{k+1}, \ldots, q_m = q_f$.

Лемма о разрастании

Находим первое повторение состояний и выделяем их. Пусть эти повторения будут q_j и q_k. Проверяем свойства:

1) $z=uvx$. Из вышесказанного,
   - u это та часть слова z, которая проходит через автомат до первого повторения, q_j;
   - v это та часть слова z, которая проходит через автомат от состояния q_j до q_k, включая q_{k-1};
   - x это та часть слова z, проходящая через состояния q_k до q_m.

То есть : $(q_0, z)=(q_0, uvx) \vdash^* (q_j, vx) \vdash^* (q_k, x) \vdash^* (q_m, \varepsilon)$.

Лемма о разрастании

Находим первое повторение состояний и выделяем их. Пусть эти повторения будут q_j и q_k. Проверяем свойства:

1) $z=uvx$. Из вышесказанного,
   - u это та часть слова z, которая проходит через автомат до первого повторения, q_j;
   - v это та часть слова z, которая проходит через автомат от состояния q_j до q_k, включая q_{k-1};
   - x это та часть слова z, проходящая через состояния q_k до q_m.

То есть : $(q_0, z)=(q_0, uvx) \vdash^* (q_j, vx) \vdash^* (q_k, x) \vdash^* (q_m, \varepsilon)$.

Лемма о разрастании

Если n — число состояний автомата, а длина слова больше или равно чем n, то путь проверки слова проходит по меньшей мере два раза через одно и то же состояние, т. е. в автомате должен присутствовать хоть один цикл (петля).

Лемма о разрастании

2) При распознавании подслов u и v участвуют разные состояния КА, поэтому их длина не может превзойти n.

3) $uv^ix \in L$, где $i=0, 1, 2, \dots$

Пусть $i=0$, тогда $ux \in L$. Значит $(q_0, ux) \vdash^* (q_i, x) = (q_i, x) \vdash^* (q_f, \varepsilon)$ – слово проходит и значит оно принадлежит языку L.

Если $i>0$, тогда $uv\dots vx \in L$, значит можно пройти автомат так: $(q_0, uvvv\dots vx) \vdash^* (q_i, vvv\dots vx) \vdash^* (q_i, vv\dots vx) = (q_i, vv\dots vx) \vdash^* (q_i, v\dots vx) \dots (q_i, x) \vdash^* (q_f, \varepsilon)$.

Лемма о разрастании

Замечание.

1) Слова u и x, из леммы, могут быть и пустыми, а v не может быть пустым.

Условия, сформулированные в лемме являются необходимыми для доказательства автоматности, но не достаточными.

Применение леммы о разрастании

1) Обратная формулировка леммы о разрастании

Пусть L некоторый язык над алфавитом V. Если для $\forall n \in N$, существует $\exists$ $z \in L$, так что бы $|z| \geq n$ и для любых подстрок $u, v, x \in V^*$ для которых $z=uvx$, $|v| \geq 1$, существует $i \in N \cup \{0\}$ для которого $uv^ix \notin L$ ⇒ то L – не автоматный.

Используется для доказательства регулярности или нерегулярности языка. $L_{01} = \{0^n1^n / n \geq 1\}$ не является регулярным.

2)

Теорема: Регулярный язык L не ограничен тогда и только тогда, когда существует слово z из L, имеющее длину $n \leq |z| < 2n$, где n — константа из леммы о разрастании. (значит существуют повторы)

Задания:

Используя лемму о разрастании разбить на 3 части, uvw, слова проходящие через следующие автоматы. Проверить выполнение свойств леммы.

1. $AF=(\Sigma, Q, q_0, \delta, F)$, $\Sigma=\{a, b, c\}$, $Q=\{q_0, q_1, q_2, q_3, q_4\}$, $F = \{q_4\}$,
δ: $\delta(q_0,a)=\{q_1\}$, $\delta(q_1,b)=\{q_2\}$, $\delta(q_2,c)=\{q_0\}$, $\delta(q_1,c)=\{q_3\}$, $\delta(q_3, c)=\{q_3\}$, $\delta(q_3, a)=\{q_4\}$, $\delta(q_4,c)=\{q_4\}$.

2. $AF = (\Sigma, Q, q_0, \delta, F)$, $\Sigma = \{0, 1\}$,
$Q = \{q_0, q_1, q_2, q_3, q_4\}$, $F = \{q_4\}$,
δ: $\delta(q_3, 1)=\{q_4\}$, $\delta(q_3, 0)=\{q_2\}$, $\delta(q_2, 0)=\{q_2\}$,
$\delta(q_2, 1)=\{q_0\}$, $\delta(q_0,0)=\{q_3\}$, $\delta(q_0,1)=\{q_1\}$, $\delta(q_1,1)=\{q_2\}$, $\delta(q_1,0)=\{q_4\}$, $\delta(q_4, 1)=\{q_4\}$.