

---

Предполагаемую структуру свёртки можно описать КС-грамматикой. При этом можно использовать более выразительные конструкции, чем нормальная форма Хомского.

- Регулярные выражения

`any : (A | U | G | C)+` // любой из символов {A,U,G,C} не менее

- **Повторения** позволяют ограничивать количество однотипных элементов (в отличии от `+``*`).

`s : any*[2..5]` // повторение any от 2 до 5 раз включительно

- Метаконструкции

```
not_empty_list<item sep>: item (sep item)*
s1: not_empty_list<NUM COLON>
s2: not_empty_list<s1 DOT>
```

`[<Start>]`

`full: any folded any`

`folded: stem<(a3 stem<a10> a3 stem<a7> a5 stem<a7> a2)>`

`stem<s>:`

```
    [0.24999975] (A stem<s> U)
    | [0.24999975] (U stem<s> A)
    | [0.24999975] (C stem<s> G)
    | [0.24999975] (G stem<s> C)
    | [0.24999975] (G stem<s> U)
    | [0.24999975] (U stem<s> G)
    | [0.000001] (s)
```

```
a1: [0.24999] (A)
    | [0.24999] (U)
    | [0.24999] (G)
    | [0.24999] (C)
```

a2: a1 | a1 a1  
a3: a2 | a1 a2  
a4: a3 | a1 a3  
a5: a4 | a1 a4  
a6: a5 | a1 a5  
a7: a6 | a1 a6  
a8: a7 | a1 a7  
a9: a8 | a1 a8  
a10: a9 | a1 a9

- A B C D — терминалы, атомы входных цепочек
- any, simple1, simple2, s — нетерминалы
- s — стартовый нетерминал

Далее, можно искать всё, что совпадает с таким шаблоном, отбирать более вероятные.

Можно пытаться восстанавливать грамматики из известных структур свёрток, можно описывать их руками.

Отдаём эту грамматику и граф нашему тулу. Он строит некую структуру, из которой можно понять, например, что он нашёл следующие вхождения (надеюсь, я ничего не пропустил).

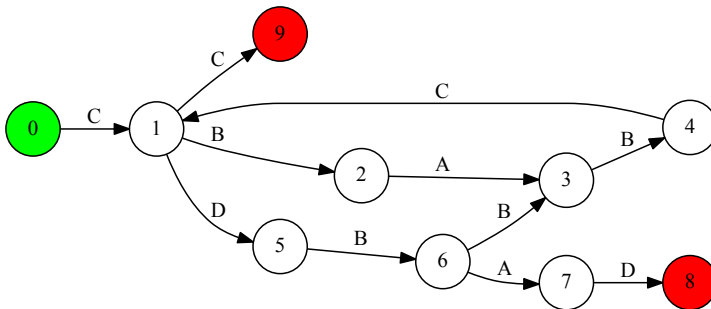


Рис. 1: Входной граф

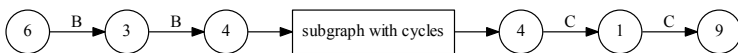


Рис. 2: Вхождение 1

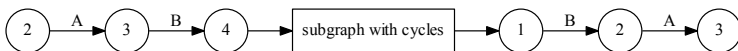


Рис. 3: Вхождение 2

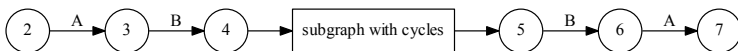


Рис. 4: Вхождение 3