Базовые задания

Дано описание языка (два — словесных, одно — атрибутное).

1. Если язык регулярный — привести или регулярку, или автомат (4 балла) и проверить префикс-свойство (1 балл)
2. Если язык детерминированный КС — привести DPDA или LL(k)-грамматику (4 балла) и доказать нерегулярность (1 балл).
3. Если язык недетерминированный КС — привести произвольный PDA или грамматику (2 балла) и доказать недетерминированность (3 балла).
4. Если язык не КС, описан атрибутно — доказать, что не КС (3 балла), а также привести его словесное описание (2 балла).
5. Если язык не КС, описан не атрибутно — доказать, что не КС (3 балла), привести атрибутную грамматику (2 балла).

Дополнительные задания

Оцениваются каждое до +5 баллов. Если базовая задача лёгкая (например, регулярка очень простая), то максимум может и не достигаться.

1. Если язык регулярный — проанализировать размер минимального НКА; построить 1-однозначную регулярку (не всегда это возможно); проверить минимальное число классов эквивалентности, если бы язык был VPL.
2. Если язык детерминированный КС — проанализировать на LL-свойство, проанализировать на префикс-свойство, проанализировать, является ли он VPL.
3. Если язык недетерминированный КС — проверить, является ли он линейным; проверить, выполняется ли префикс-свойство; проверить, является ли КС-языком его дополнение.
4. Если язык не КС — проверить префикс-свойство, привести альтернативные доказательства не КС, построить расширенный regex или конъюнктивную грамматику, проверить на КС-свойство его дополнение.

Лайфхаки

- Если язык КС, но не понятно, детерминированный или нет — строить произвольную грамматику или PDA и в любом случае получить за это 2 балла.
- Если префикс-свойство не совсем тривиальное — проверить его и точно добавить себе 1 балл.
- Доказательство нерегулярности языка, если он КС — тоже гарантированный 1 балл (но 0 баллов, если язык — не КС).
- Если язык не является DCFL — анализ дополнения принесёт допбаллы почти всегда, за исключением случая, когда он совсем тривиальный.
- Если язык судя по всему не КС, но не удаётся это доказать, зато удалось доказать, что он не DCFL — всё равно будет 1 балл.
- Если удалось доказать, что язык не LL, но не удаётся доказать, что он не DCFL — это потеря всего 1 балла из 3 возможных.
- Если удалось доказать, что язык не VPL, но не удаётся понять, DCFL ли он — 1 балл получен всё равно.

Размер минимального НКА

Уточнённая **теорема Глайстера-Шаллита**: если существуют N префиксов $\gamma_1,\dots,\gamma_N$ и N суффиксов $\omega_1,\dots,\omega_N$ таких, что $\forall i,j(\gamma_i\omega_i \in \mathscr{L} \;\&\; (j > 0 \Rightarrow \gamma_{i-j}\omega_i \notin \mathscr{L}))$, то размер минимального НКА не меньше, чем N.

Пример — $a(a|b)^*a|b(a|b)^*b$.

КЭ по Майхиллу–Нероде:

	ε	a	b	aa	bb
ε	0	0	0	1	1
a	0	1	0	1	0
b	0	0	1	0	1
aa	1	1	0	1	0
bb	1	0	1	0	1

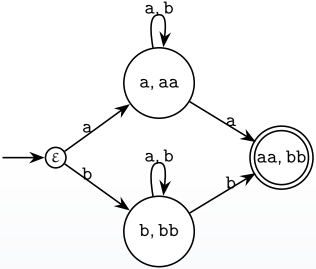
Приближённые КЭ НКА:

	aa	a	b	ε
ε	1	0	0	0
a	1	1	0	0
b	0	0	1	0
aa	1	1	0	1

NB: бывают языки с неточной оценкой КЭ НКА по этой теореме.

Размер минимального НКА

Пусть нашлась треугольная матрица таких $\gamma_1,\dots,\gamma_N$; $\omega_1,\dots,\omega_n$. Тогда, если префиксы γ_i и γ_{i+j} будут всегда вместе присутствовать в одном и том же состоянии НКА, то из него станет возможно распознать слово $\gamma_i\omega_{i+j}$, которое согласно таблице языку не принадлежит. То k-ая строка в таблице определяет префиксный КЭ, достигаемый хотя бы в одном состоянии, не достижимом по всем классам из $1 \dots k-1$ строк.



Если построить НКА с 4 состояниями, распознающими $a(a|b)^*a|b(a|b)^*b$, можно увидеть, что финальное состояние достижимо только по префиксу $\mathtt{aa}$. То, что по нему также достижимо состояние, соответствующее префиксу $\mathtt{a}$, показывает, что в таблице приближённых КЭ $\mathtt{a}$ может стоять только выше, чем $\mathtt{aa}$.

Заметим, что префикс $\mathtt{bb}$ не определяет никакого состояния, не достижимого по остальным префиксам, поэтому на размер НКА его наличие в таблице КЭ по Майхиллу-Нероде не влияет.