1. $\mathscr{L}_1 = \left\{v_1 z v_2 z \mid |z| \geqslant 2 \ \&\ z, v_2 \in \{a, b, c\}\star \ \&\ v_1 \in \{a, b\}\star\right\}$

2. Множество троичных чисел, кратных 5.

3. Множество трасс грамматик с правилами вида $N_i \to \gamma N_j$, $N_i \to N_j\gamma$, $N_i \to \varepsilon$.

4. Множество трасс грамматики G_k с правилами вида $S \to S_1 S_1$, $S_1 \to S_2 S_2 \ldots$, $S_k \to a$, $S_k \to SS$.

Решение задачи I

«Интуитивно» язык не регулярный — есть требование вхождения одинаковых подслов, которые «по идее» можно как-то отделить от аморфных значений v_1 и v_2. Попробуем провести это отделение более последовательно.

Два момента в данном языке обращают на себя внимание:

- словарь (язык) v_1 отличается от словаря смежного с ним слова z. Это значит, что если взять первую букву z из разности языка z и языка v_1, то никакое значение v_1 не сможет её поглотить.

- словарь (язык) v_2 точно такой же, как у смежных вхождений z. Поэтому в принципе v_2 может поглотить любой суффикс первого вхождения z, а также любой префикс второго вхождения.

Чтобы слово v_2 не могло поглотить префикс второго вхождения z, надо сделать так, чтобы его положение определялось однозначно. Мы уже знаем, что выгодно взять z начинающимся с буквы c (это исключит поглощение префикса z значением v_1). Значит, если в слове будет всего две буквы c, то ровно одна из них должна начинать первое вхождение z, и ровно одна должна начинать второе вхождение. Осталось дополнить z достаточно длинными суффиксами, исключающими возможность «накачки» их одновременно.

Кандидат на контрпример — серия слов $ca^{n+2}ca^{n+2}$, где n — длина накачки.

Теперь можно доказать нерегулярность языка посредством теоремы Майхилла–Нероуда. Действительно, слова вида $ca^m ca^{m+k+1}$ языку не принадлежат, а $ca^{m+k}ca^m$ — точно принадлежат (при этом $z = ca^m$, $v_2 = a^k$), что порождает нижнетреугольную матрицу принадлежности: наименования строк здесь — это префиксы слов, а столбцов — соответствующие суффиксы.

	$\mathtt{ca}$	$\mathtt{ca}^2$	$\mathtt{ca}^3$	$\dots$
$\mathtt{ca}$	$+$	$-$	$-$	$-$
$\mathtt{ca}^2$	$+$	$+$	$-$	$-$
$\mathtt{ca}^3$	$+$	$+$	$+$	$-$
$\dots$		$\dots$		

Также можно использовать этот контрпример для построения короткого доказательства нерегулярности $\mathscr{L}_1$, если пересечь его с языком (регулярным) $R = \mathtt{ca}^*\mathtt{ca}^*$. В слове $\mathtt{ca}^{n+2}\mathtt{ca}^{n+2}$, принадлежащем пересечению $R \cap \mathscr{L}_1$, можно накачивать лишь фрагмент, состоящий только из букв $\mathtt{a}$, чтобы не выйти из языка R. Пусть такой фрагмент имеет длину k (где $k > 0$). Тогда при отрицательной накачке получится слово $\mathtt{ca}^{n-k+2}\mathtt{ca}^{n+2}$, которое не входит в $\mathscr{L}_1$.

Наиболее неприятный путь — прямое применение леммы о накачке к слову $\mathtt{ca}^{n+2}\mathtt{ca}^{n+2}$ без использования свойств замыканий. На этом пути придётся разобрать два случая.

- Фрагмент накачки имеет вид $\mathtt{a}^k$. Этот случай аналогичен уже рассмотренному в решении с пересечением.
- Фрагмент накачки имеет вид $\mathtt{ca}^k$. Отметим, что $k < n$. Тогда при положительной накачке в одну итерацию получим слово $\mathtt{ca}^k\mathtt{ca}^{n+2}\mathtt{ca}^{n+2}$. Поскольку это слово начинается с $\mathtt{c}$, то значение z должно начинаться с $\mathtt{c}$, а значит, должно заканчиваться на $\mathtt{a}^{n+2}$ (потому что первое с конца вхождение $\mathtt{c}$ уж точно будет относиться к z). Но это значит, что z должна содержать также и фрагмент $\mathtt{ca}^k\mathtt{c}$ (иначе первое вхождение z не сможет заканчиваться на $\mathtt{a}^{n+2}$), а он в этом слове не повторяется.

В этой задаче у большинства возникла одна из двух проблем:

- Или взято значение z с префиксом в языке $\{\mathtt{a}, \mathtt{b}\}$, из-за чего оно смешалось со значением v_1.
- Или взято значение z, равное $\mathtt{c}^n$ (очевидно, вы заметили, что в противном случае анализу мешает v_1). Тогда мы можем весь суффикс z, кроме двух первых букв, положить в v_2. Вообще, в этой задаче не стоит брать значения z, значение префикс-функция у которых больше, чем 0 (и осторожнее с такими словами в других задачах).