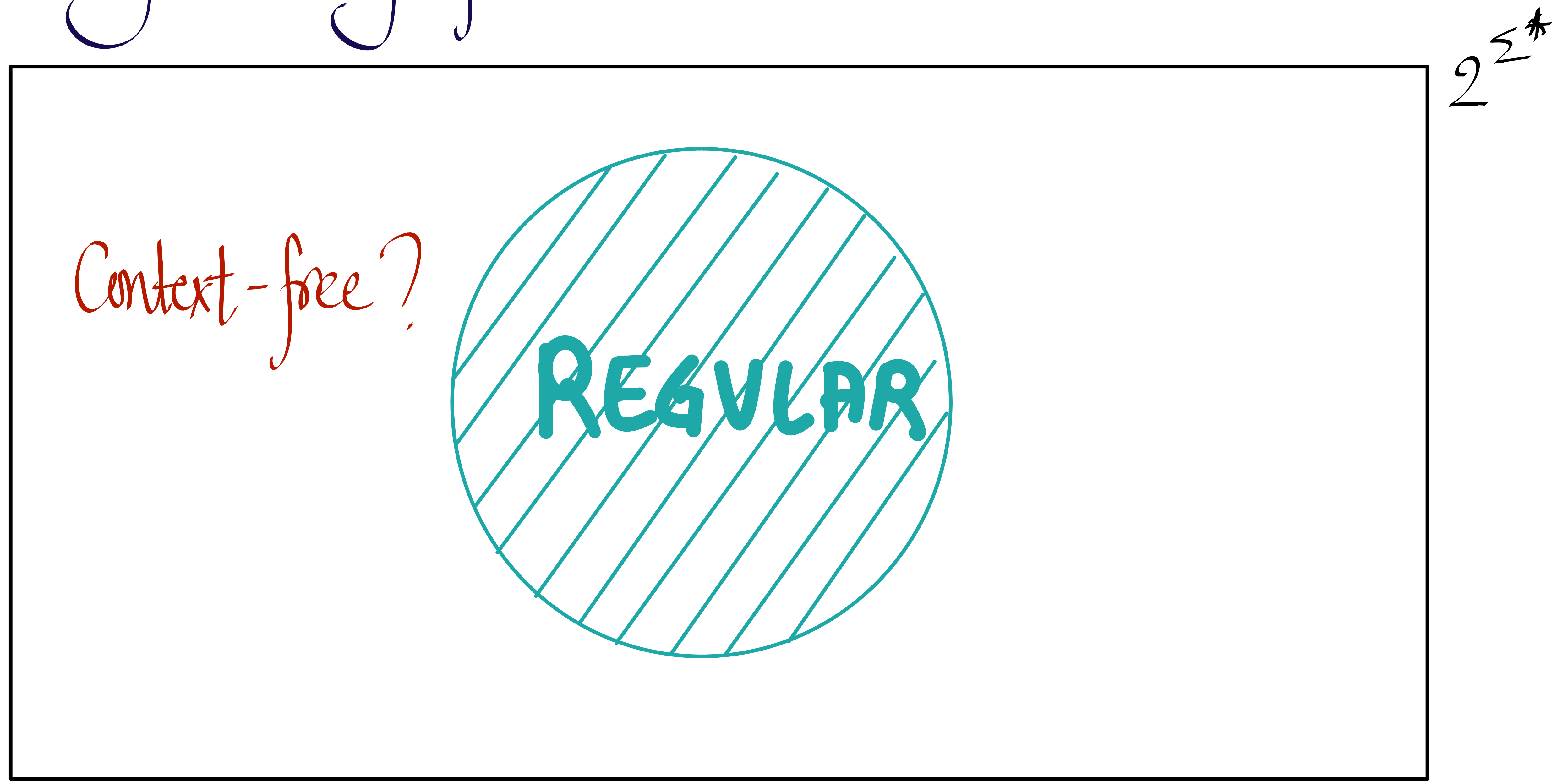
Non-Context Free Languages

Recall: We showed that some non-regular languages are context-free languages. These are generated by context-free grammars, and recognized by pushdown automata.

Today:



Is every non-regular language context-free?

Closure properties:

→ We saw that CFLs are closed under intersection with regular languages
(Cross-product construction, touch the stack only for the CFL)

→ Are CFLs closed under union?
If $\mathcal{L}_1$ and $\mathcal{L}_2$ are CFLs, is $\mathcal{L}_1 \cup \mathcal{L}_2$ also a CFL?

$G_1 = (NT_1, T_1, R_1, S_1)$ $\quad$ $G_2 = (NT_2, T_2, R_2, S_2)$ $\quad$ $T = T_1 \cup T_2$

$G = (NT_1 \cup NT_2 \cup \{S\}, T, R_1 \cup R_2 \cup \{S \to S_1 | S_2\}, S)$ $\quad$ S "fresh"

* Which of the other regular operations is the class of CFLs closed under?

Construct a CFG or a PDA for $L = \{a^n b^n c^n \mid n \geq 0\}$

$L' = \{a^n b^{2n} a^n \mid n \geq 0\}$

$L_{sq} = \{a^{n^2} \mid n \geq 0\}$

$L_{2n} = \{a^{2n} \mid n \geq 0\}$

We said that DFAs could not "count"

What can PDAs not do?

What CAN they do? Broadly,

- They can count[1] Is $\{a^{n^2} \mid n \geq 0\}$ a CFL?
- They can keep track of pairs[2] Is $\{a^n b^n c^n \mid n \geq 0\}$ a CFL?

1) Not beyond a "linear" count

2) But not if distinct matches disturb the stack in between

How does one show that a language is not context-free?

The pumping lemma for CFLs.

Pumping lemma for CFLs :

Consider a language $\mathcal{L} \subseteq \Sigma^*$. If

$\forall$ for any $k \geq 0$,

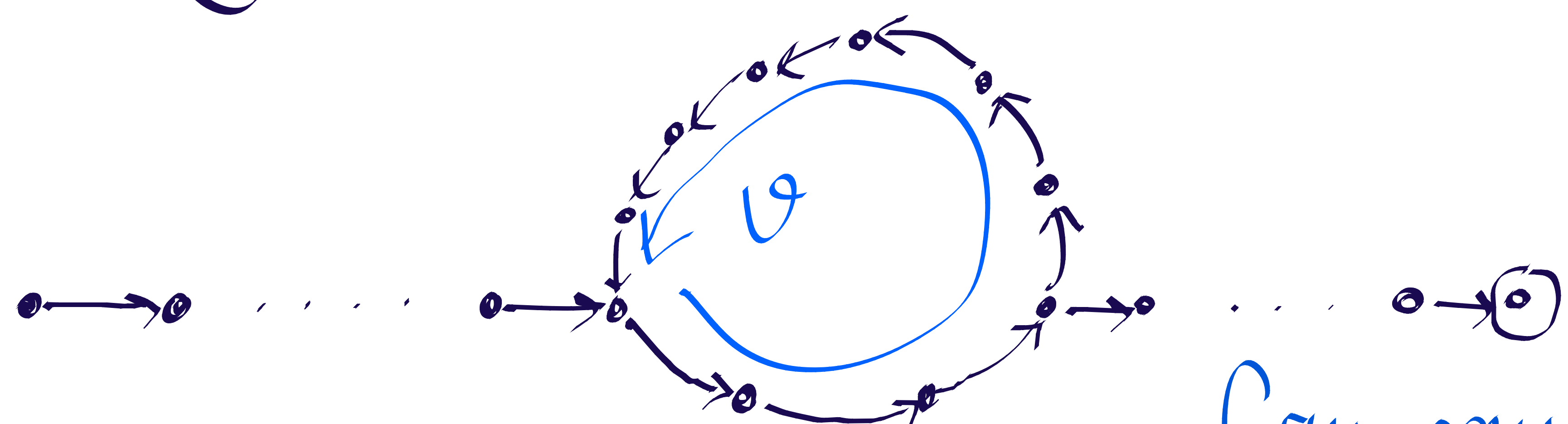
$\exists$ there is an $s \in \mathcal{L}$ s.t. $|s| \geq k$,

$\forall$ for any $x, u, v, w, z \in \Sigma^*$ s.t. $xuvwz = s$, $|uvw| \leq k$, $|uw| > 0$,

$\exists$ there is some $i \geq 0$ s.t. $xu^i v w^i z \notin \mathcal{L}$

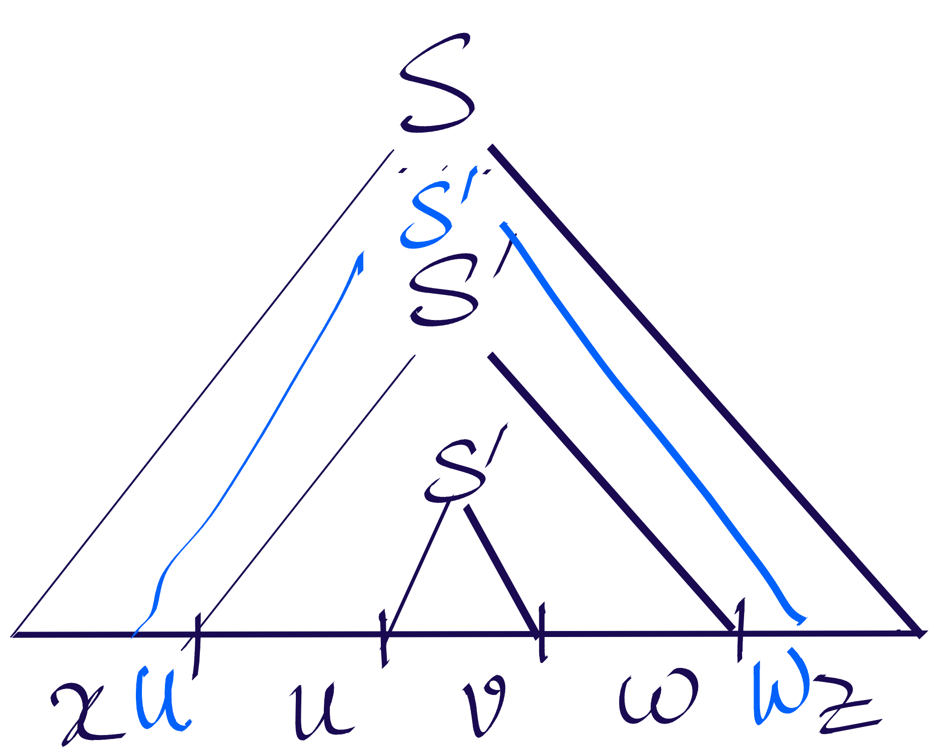
Then, $\mathcal{L}$ is not context-free.

For regular languages, we were basically exploiting the following structure of a DFA path



Can run the loop any number of times, and still stay in the language

For CFLs, we exploit the structure of parse trees



Cut and paste the section generated by the repeating non-terminal symbol any number of times, and still get a valid parse tree.

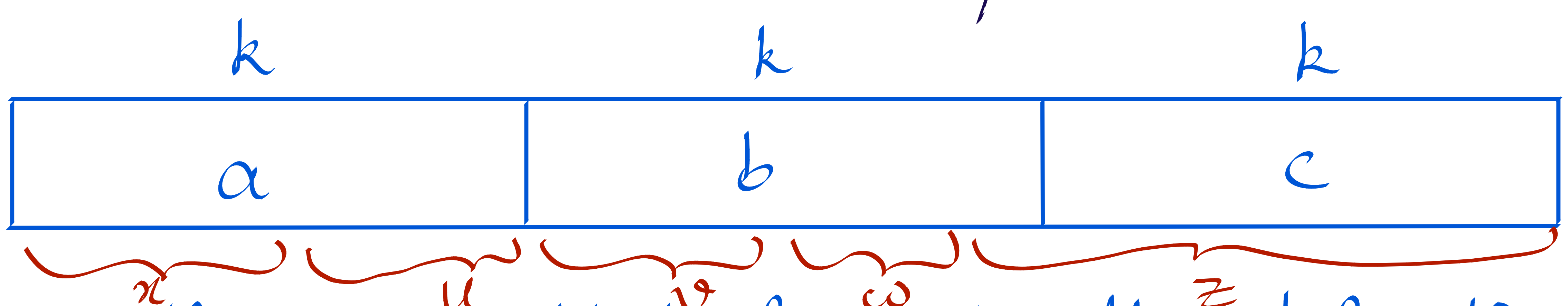
Use the pumping lemma to show that $\mathcal{L} = \{a^n b^n c^n \mid n \geq 0\}$ is not cf.

$\forall$ for any $k \geq 0$,

$\exists$ there is an $s \in \mathcal{L}$ s.t. $|s| \geq k$, $a^k b^k c^k$

$\forall$ for any $x, u, v, w, z \in \Sigma^*$ s.t. $xuvwz = s$, $|uvw| \leq k$, $|uw| > 0$,

$\exists$ there is some $i \geq 0$ s.t. $xu^i v w^i z \notin \mathcal{L}$ $i = 2$



① uvw falls exactly inside a block of size k: Mismatch with other two blocks

② uvw spans blocks: Can only span two blocks, not three. Choosing $i = 2$ will cause a mismatch with the third block.

So $\mathcal{L}$ is not context-free.

Use the pumping lemma to show that $\mathcal{L} = \{ww \mid w \in \{a,b\}^*\}$ is not CF.

$\forall$ for any $k \geq 0$,

$\exists$ there is an $s \in \mathcal{L}$ s.t. $|s| \geq k$, $a^k b^k a^k b^k$

$\forall$ for any $x, u, v, w, z \in \Sigma^*$ s.t. $xuvwz = s$, $|uvw| \leq k$, $|uw| > 0$,

$\exists$ there is some $i \geq 0$ s.t. $xu^i v w^i z \in \mathcal{L}$

Use the pumping lemma to show that $\mathcal{L} = \{ww \mid w \in \{a,b\}^*\}$ is not CF.

$\forall$ for any $k \geq 0$,

$\exists$ there is an $s \in \mathcal{L}$ s.t. $|s| \geq k$, $s = a^k b^k a^k b^k$

$\forall$ for any $x, u, v, w, z \in \Sigma^*$ s.t. $xuvwz = s$, $|uvw| \leq k$, $|uw| > 0$,

$\exists$ there is some $i \geq 0$ s.t. $xu^ivw^iz \notin \mathcal{L}$

k	k	k	k
a	b	a	b

Similar reasoning as earlier, set $i=2$.

Circle back to closure under the regular operations.

Are CFLs closed under intersection?

$\mathcal{L} = \{a^n b^n c^n \mid n \geq 0\}$ is not context-free

If we can express $\mathcal{L}$ as an intersection of two CFLs, that would prove non-closure under intersection.

$\{a^m b^n c^n \mid m \geq 0, n \geq 0\}$

$\{a^n b^n c^m \mid m \geq 0, n \geq 0\}$

Circle back to closure under the regular operations.

Are CFLs closed under intersection?

$\mathcal{L} = \{a^n b^n c^n \mid n \geq 0\}$ is not context-free

If we can express $\mathcal{L}$ as an intersection of two CFLs, that would prove non-closure under intersection.

$$\mathcal{L}_1 = \{a^m b^n c^n \mid m, n \geq 0\}$$

$$\mathcal{L}_2 = \{a^n b^n c^m \mid m, n \geq 0\}$$

* What about closure under negation?