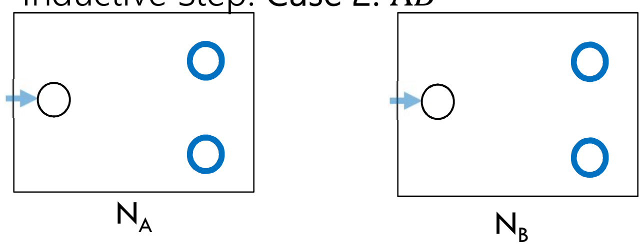
Let $P(A)$ be "There is an NFA whose language is the same as the language for A."

Let R be a regex not covered by the base cases. By the exclusion rule, $R = A \cup B$ or AB or A^* from some regexes A, B
Inductive Hypothesis: Suppose $P(A)$ and $P(B)$.

Inductive Step: Case 2: AB



Want a machine that accepts exactly strings matched by AB.

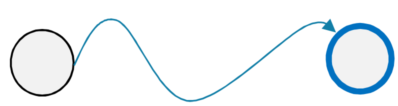
Forcing a Mistake

How do we know x, y must be in different states?

Well if one would be accepted and the other rejected, that would be a clear sign.

Or if there's some string z where xz is accepted but yz is rejected (or vice versa).

The machine is deterministic! If x and y take you to the same state, then xz and yz are also in the same state!



A Proof Outline

Claim: $\{0^k 1^k : k \geq 0\}$ is an irregular language.

…

Let $S =$ [TODO]. *S is an infinite set of strings.*

Because the DFA is finite, there are two (different) strings x, y in S such that x and y go to the same state. *We don't get to choose x, y*

Consider the string $z =$[TODO] *We do get to choose z depending on x, y*

Since x, y led to the same state and M is deterministic, xz and yz will also lead to the same state q in M. Observe that $xz \in \{0^k 1^k : k \geq 0\}$ but $yz \notin \{0^k 1^k : k \geq 0\}$. Since q is can be only one of an accept or reject state, M does not actually recognize $\{0^k 1^k : k \geq 0\}$. That's a contradiction!

Therefore, $\{0^k 1^k : k \geq 0\}$ is an irregular language.

Claim: $\{0^k 1^k : k \geq 0\}$ is an irregular language.

Proof:

Suppose, for the sake of contradiction, that $\{0^k 1^k : k \geq 0\}$ is regular.

Then there is a DFA M such that M accepts exactly $\{0^k 1^k : k \geq 0\}$.

Let $S = \{0^k : k \geq 0\}$.

Because the DFA is finite and S is infinite, there are two (different) strings x, y in S such that x and y go to the same state when read by M. Since both are in S, $x = 0^a$ for some integer a, and $y = 0^b$ for some integer b, with $a \neq b$.

Consider the string $z = 1^a$. $xz = 0^a 1^a \in \{0^k 1^k : k \geq 0\}$ but $yz = 0^b 1^a \notin \{0^k 1^k : k \geq 0\}$.

Since x, y both end up in the same state, and we appended the same z, both xz and yz end up in the same state of M.

Since $xz \in \{0^k 1^k : k \geq 0\}$ and $yz \notin \{0^k 1^k : k \geq 0\}$, M does not recognize$\{0^k 1^k : k \geq 0\}$. But that's a contradiction!

So$\{0^k 1^k : k \geq 0\}$ must be an irregular language.