A Sample Regex to NFA Conversion

UMD CMSC330 - Kauffman

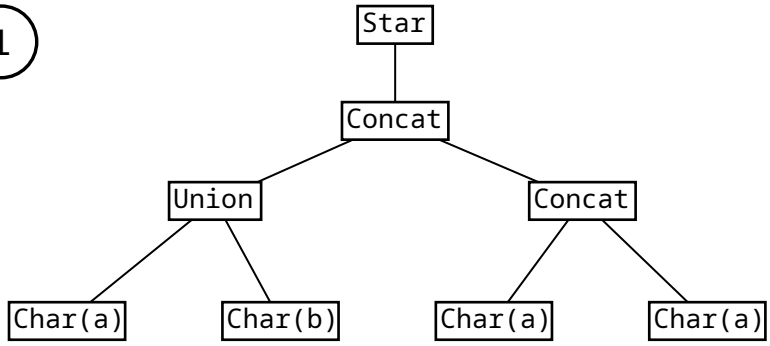
The parse tree for following formal regex is shown nearby.

```
((a|b)aa)*
```

In a program, it would likely be written with some shorthand conventions like this:

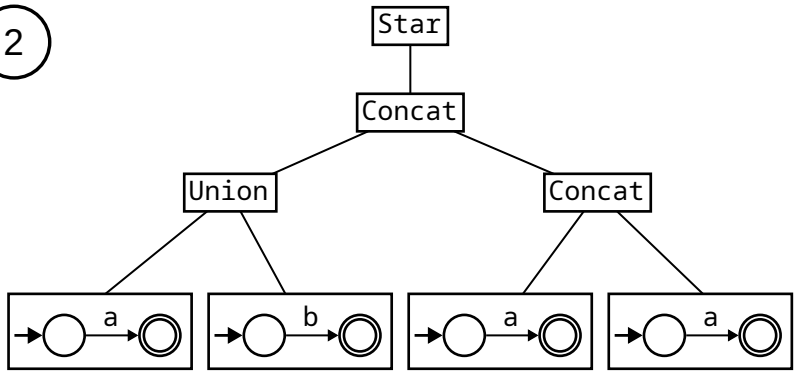
```
([ab]aa)*
```

1



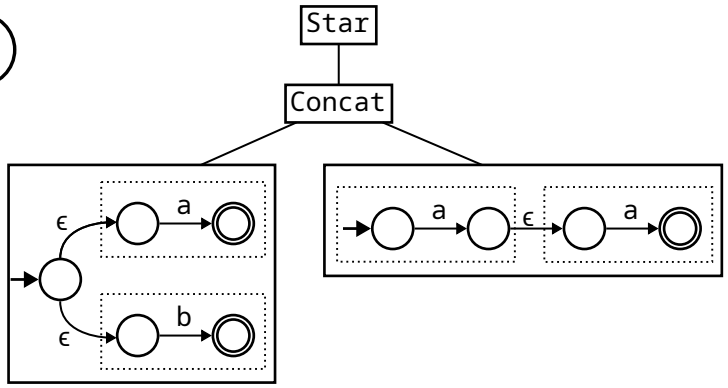
2

In a bottom up conversion, the leaf nodes which are Char() parts of the Regex can be converted to 2-state NFAs which Accept after reading the single input character indicated



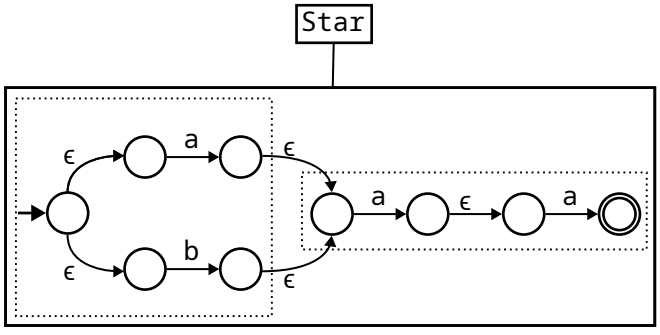
3

(Left branch) The Union of two NFAs is constructed by introducing a new start state with ϵ-edges to the two other NFA start states.
(Right branch) Concatenation switches all of the first NFA's accept states non-accepting, then connects them to the second NFA's start state with an ϵ-edge.



4

A second application of concatenation follows.



5

Star (Kleen Closure) introduces a new Start state which is also an Accept state. This is connected to the sub-NFA's start state with an ϵ-edge. Finally, all Accept states are connected to the original Start state with an ϵ-edge.