

# Wyrażenia Regularne

Marek Piechula  
#inf.aei.polsl.pl@Freenode

# Historia

Były sobie wyrażenia regularne...  
oraz żyły długo i szczęśliwie...

# Do czego to jest

- Do szukania słów/powtórzeń
- Do masowej zamiany (ewentualnie zagłady)
- Sprawdzanie poprawności danych od użytkownika
- Do szpanowania

# W czym nie pomoże

- W zrobieniu obiadu
- W parsowaniu (google html regexp)
- Do nieregularnego tekstu

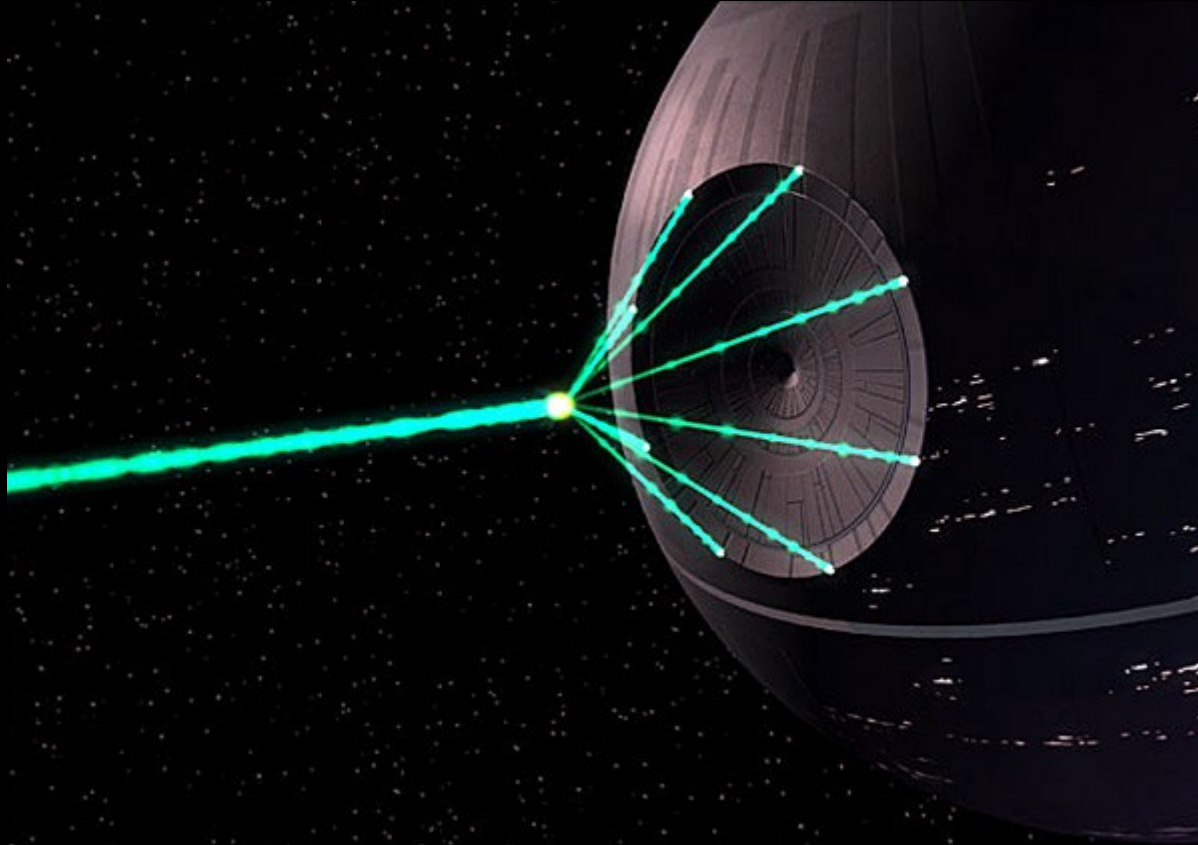
Co tu pisze

`[A-Z0-9._%+~]+@[A-Z0-9.-]+\.[a-z]{2,4}`

Co widziecie wy

是什麼樣的 @ 個無聊的演講

# Co widzę ja



## Kilka słów

- znak
- grupa
- meta
- metaznaki
- zachłanność
- leniwość
- odwołania  
wsteczne
- flagi



# Pomocnicze

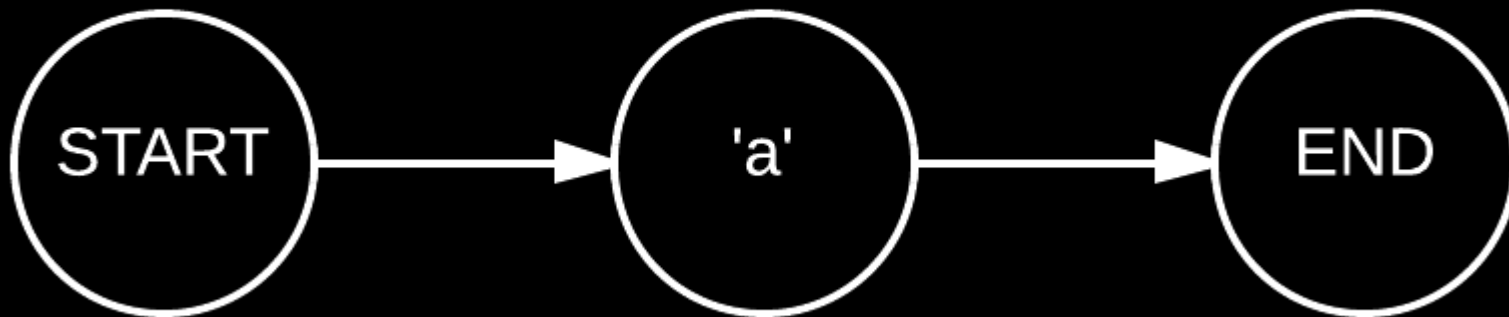
[

]

]

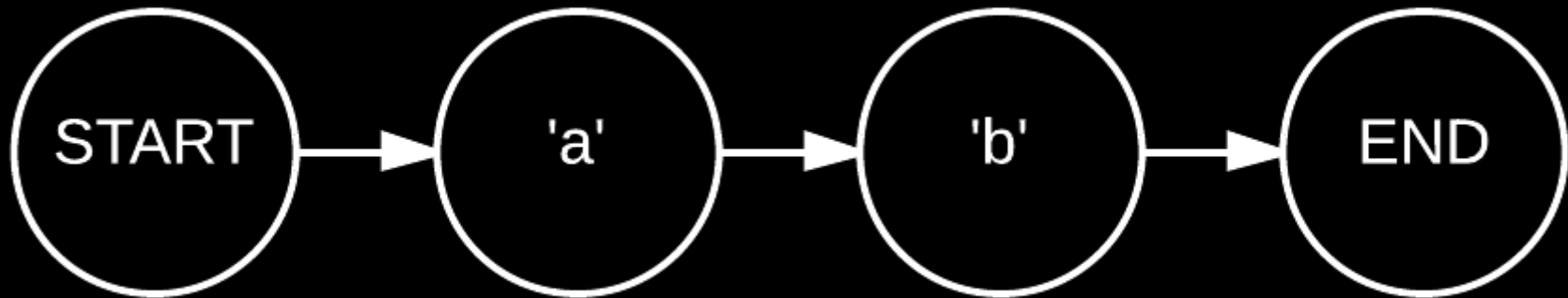
Pierwsze wyrażenie

[a]

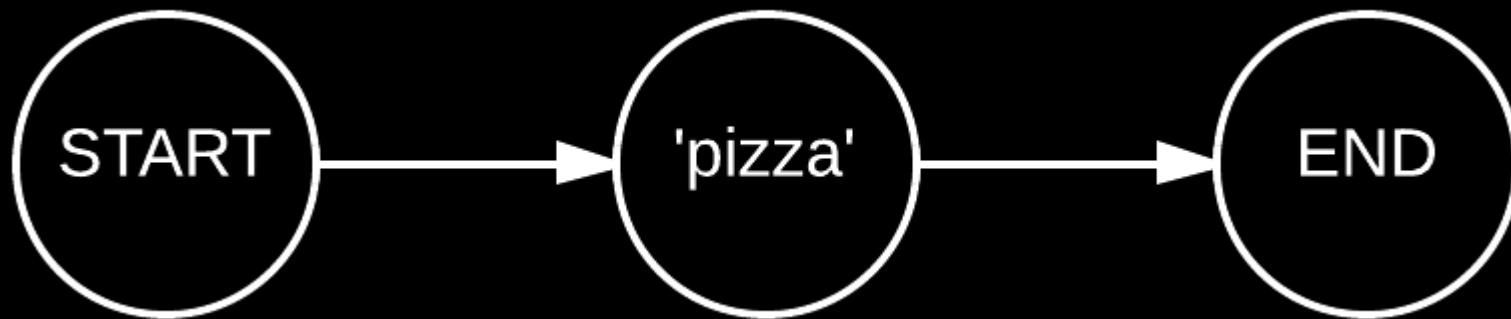


O Boże Drugie wyrażenie!

「ab」

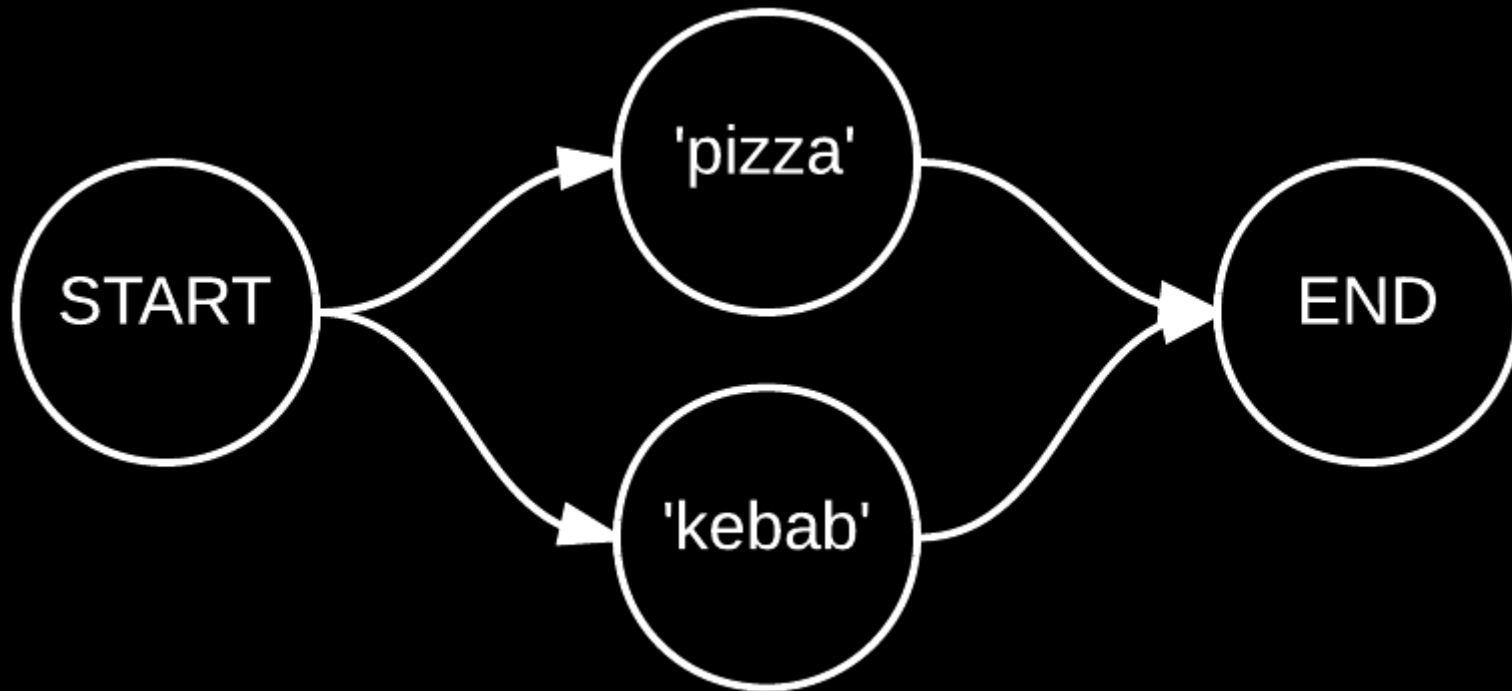


# 「pizza」



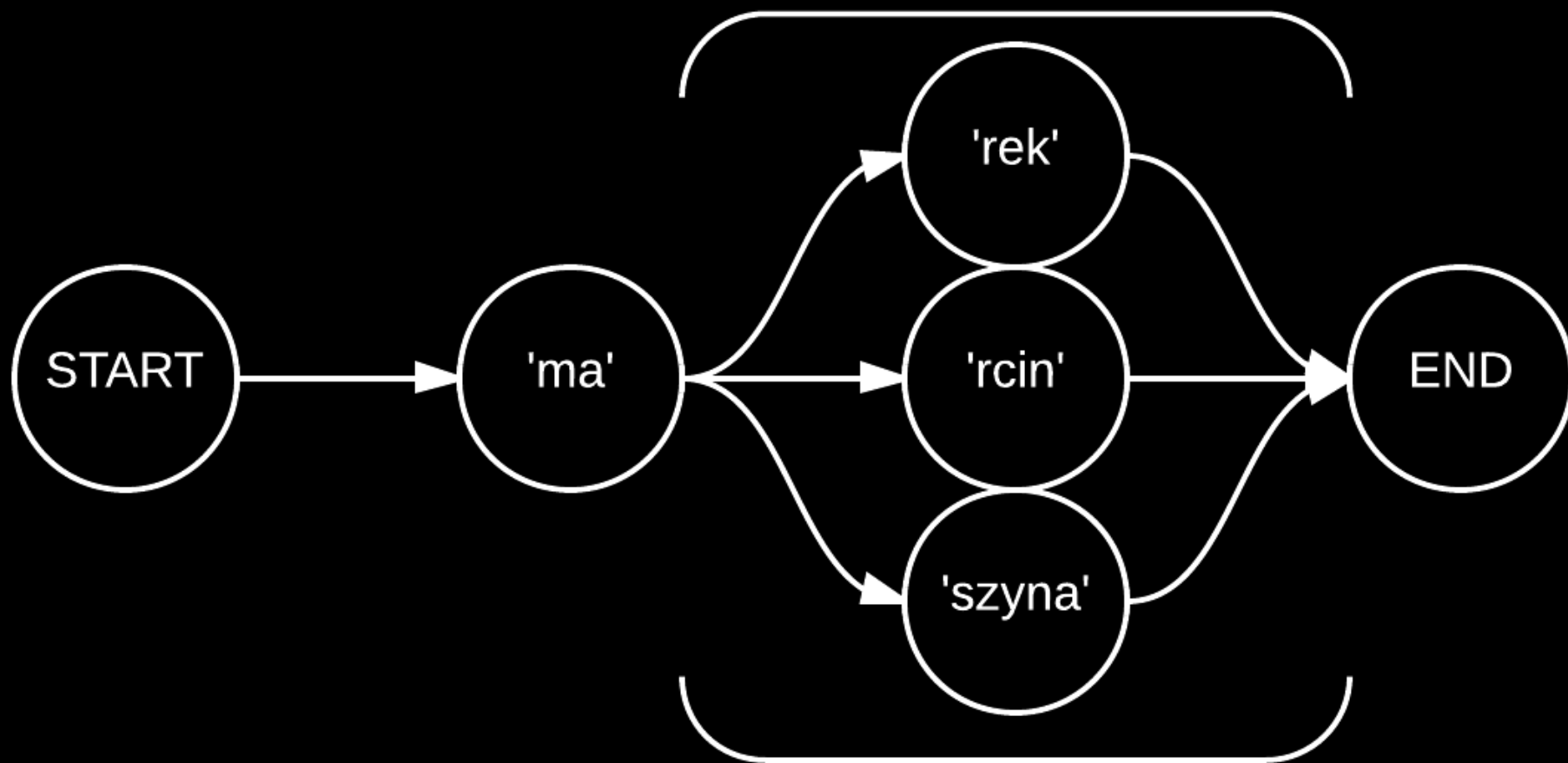
alternatywa

[pizza|kebab]



grupy

「ma(rek|rcin|szyna)」



Znaki początku i końca

`「^pizza」`

`「pizza$」`

`「^pizza$」`

# Klasy znaków

$[ID:(a|b|c|d|e|f)(a|b|c|d|e|f)]$



# Klasy znaków

[ID:[abcdef][abcdef]]

# Klasy znaków

[ID:[a-f][a-f]]

[ID:[a-fA-F][a-fA-F]]

## Klasy znaków

`\d`  $\leftrightarrow$  `[0-9]`

`\w`  $\leftrightarrow$  `[a-zA-Z0-9_]`

`\s`  $\leftrightarrow$  `[\t\n\r\f\v]`

## Negacije klas

$\backslash D \leftrightarrow [^0-9]$

$\backslash W \leftrightarrow [^a-zA-Z0-9_]$

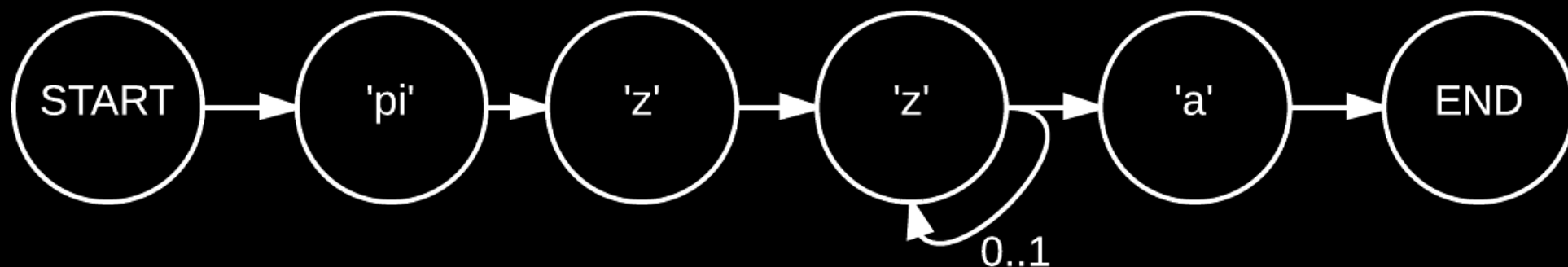
$\backslash S \leftrightarrow [^_\backslash t\backslash n\backslash r\backslash f\backslash v]$

Wszystkie znaki

$$\lceil [abc\dots] \rceil \leftrightarrow \lceil . \rceil$$

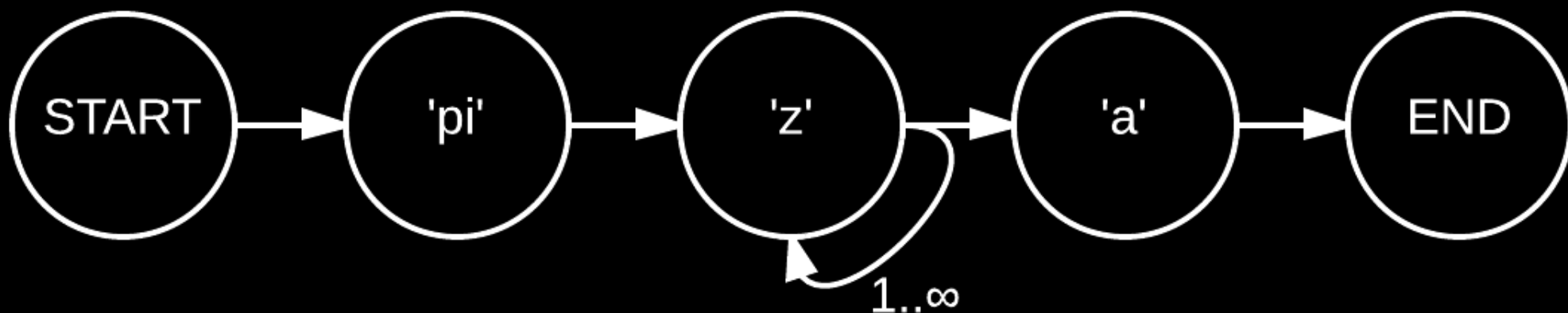
Zakresy

「pizz?a」



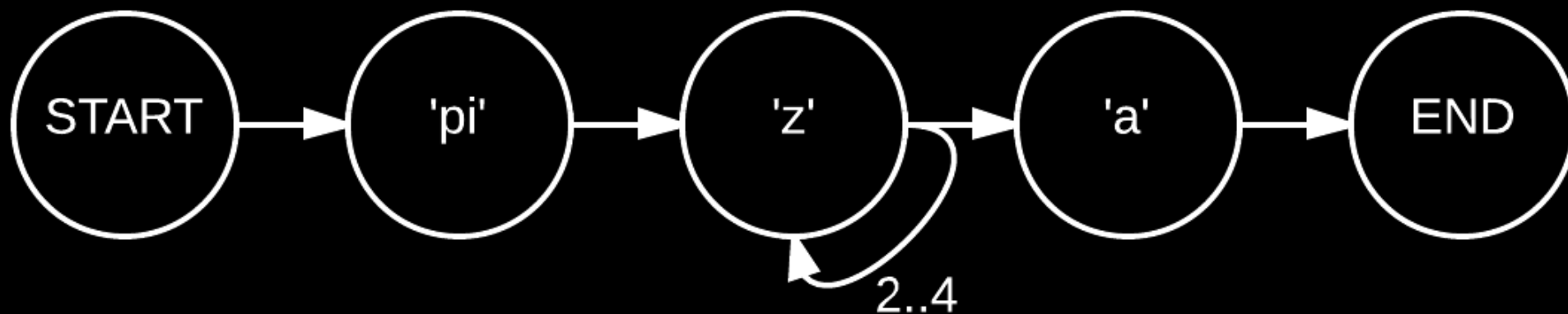
Zakresy

「piz+a」



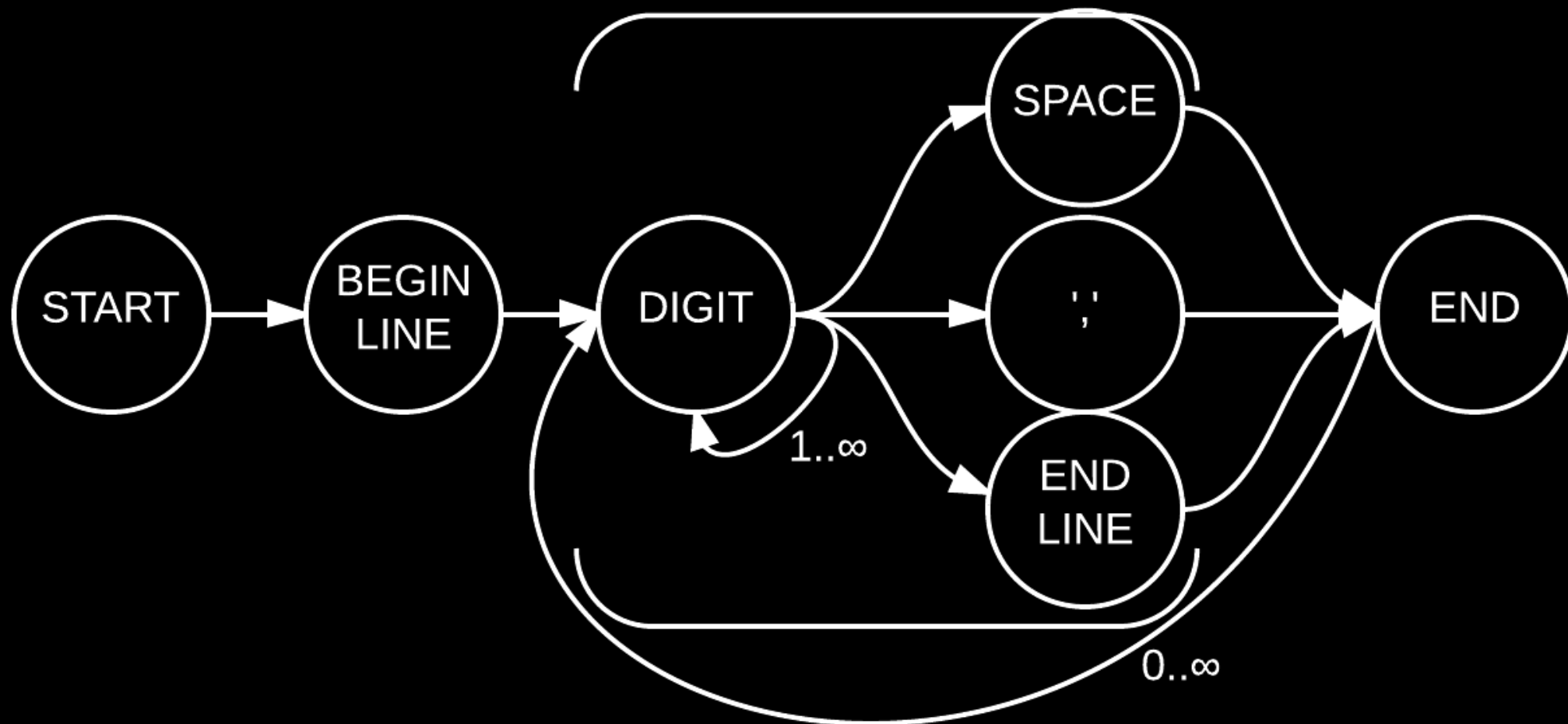
Zakresy

$[piz\{2,4\}a]$

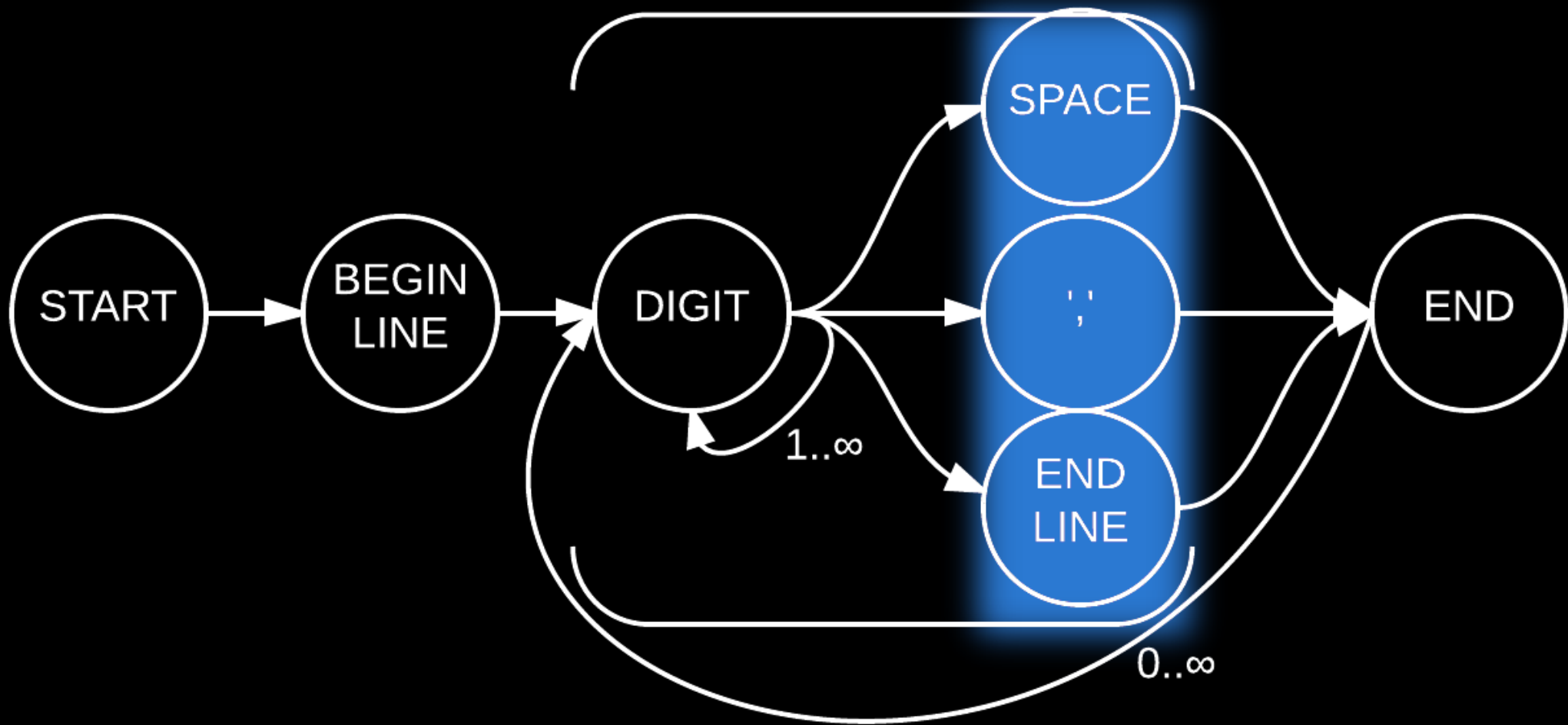




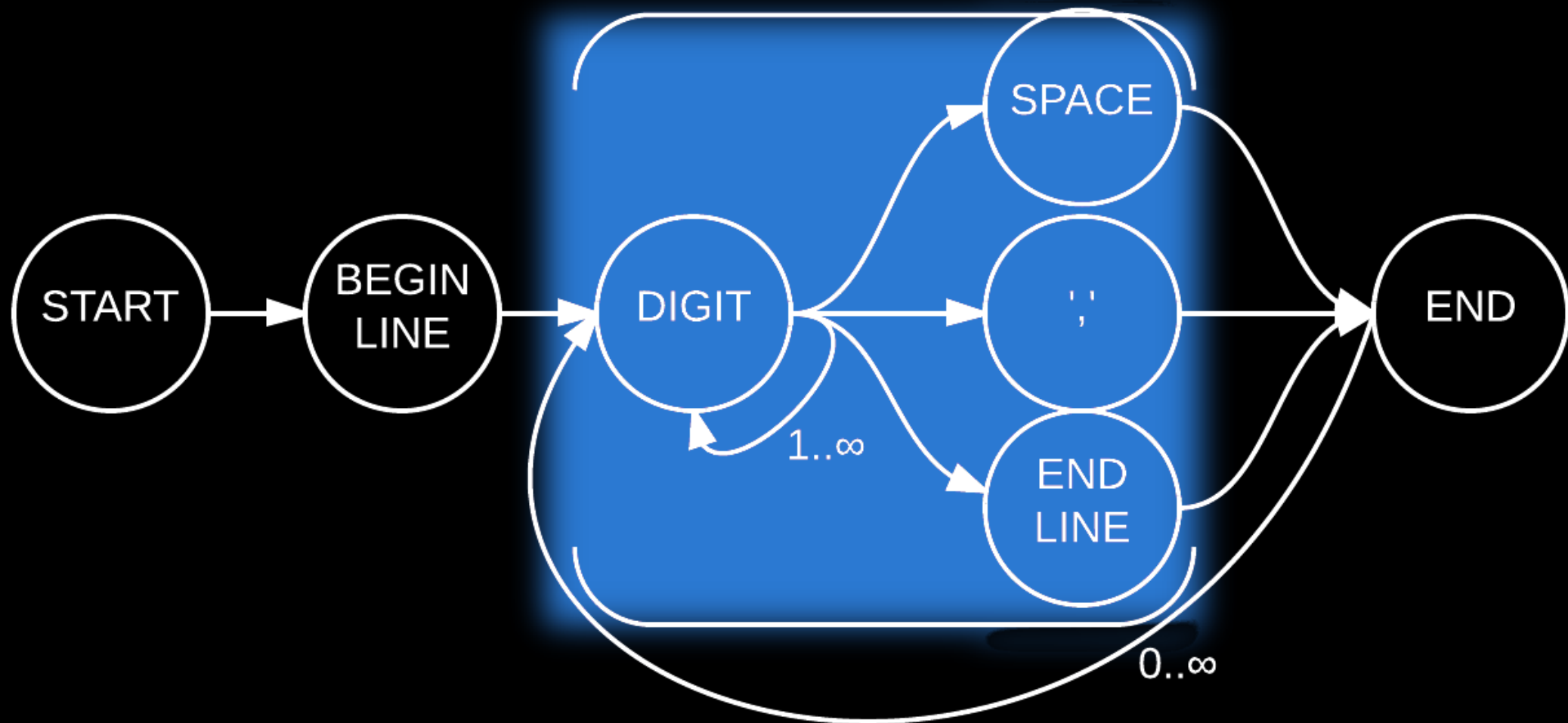
# Zakresy



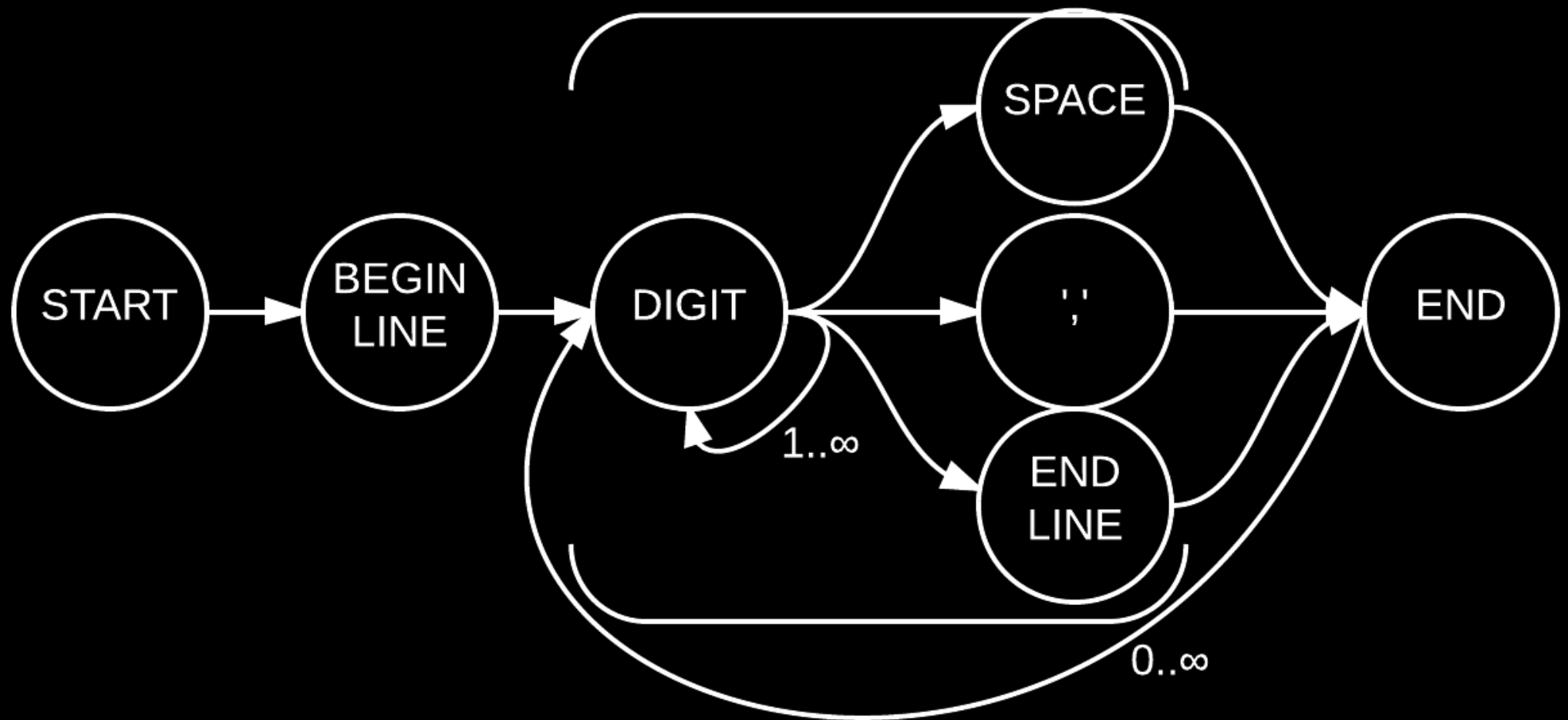
「([\_ , ] | \$)」



$\text{[}(\backslash\text{d}+(\text{[}\_\text{,}\text{]}|\$))\text{]}$



$\text{[}^\wedge(\backslash\text{d}+(\text{[}\_\text{,}\text{]}|\$))\text{]*}$



# Level 2!

Co nam zostało:

- odwołania wsteczne
- Leniwość
- Nowe grupy
- PHP, Python, JS, C++ (!)

## Nowe grupy

「(?:niewazna)」

「(?:!zakazana)」

「(?:P<nazwa>\w)」

## Nowe grupy

「(?=niewidoczna)」

「(?<!zakazana)」

「(?<=niewidoczna)」

「(?#komentarz)」

leniwość

「.\*?」

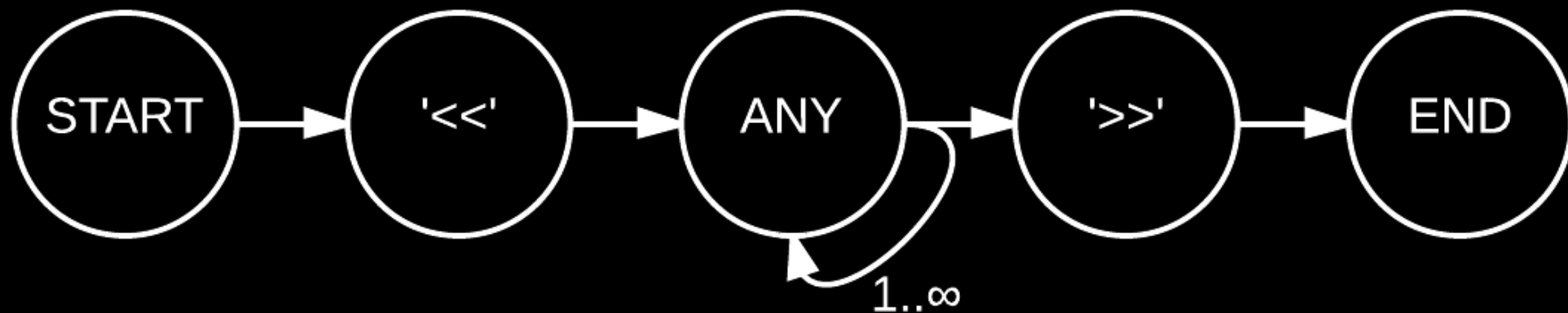
「.??」

「.+?」



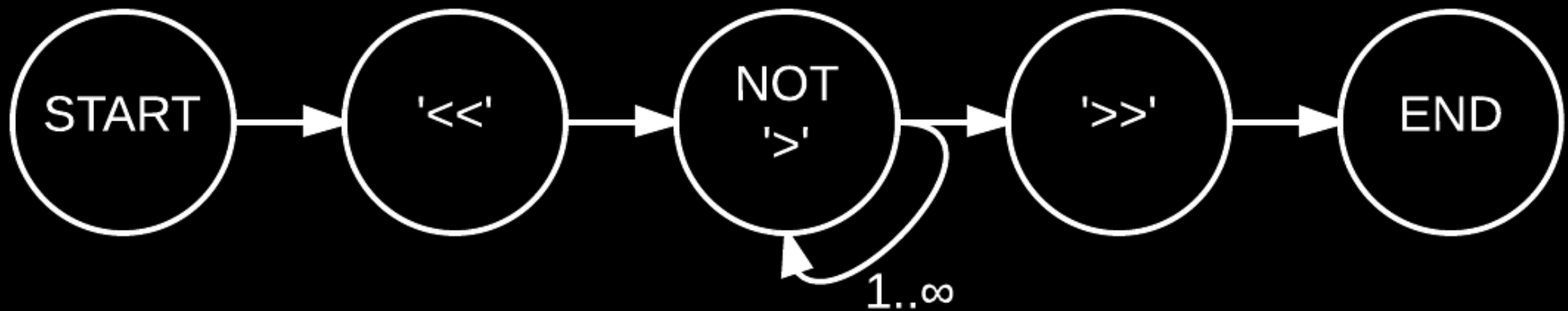
# leniwość

[<<.+>>]



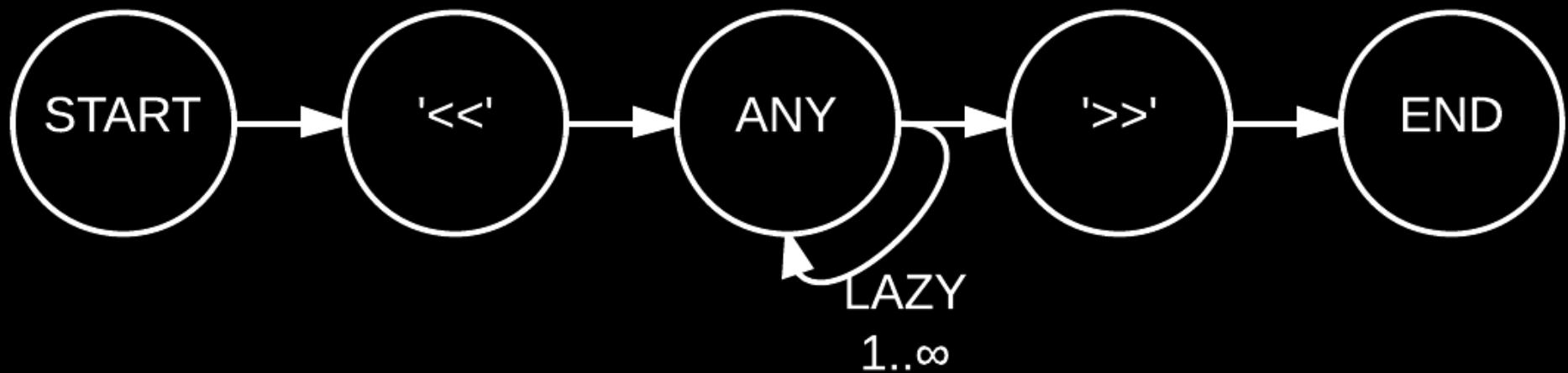
leniwość

[<<[^>]>>]



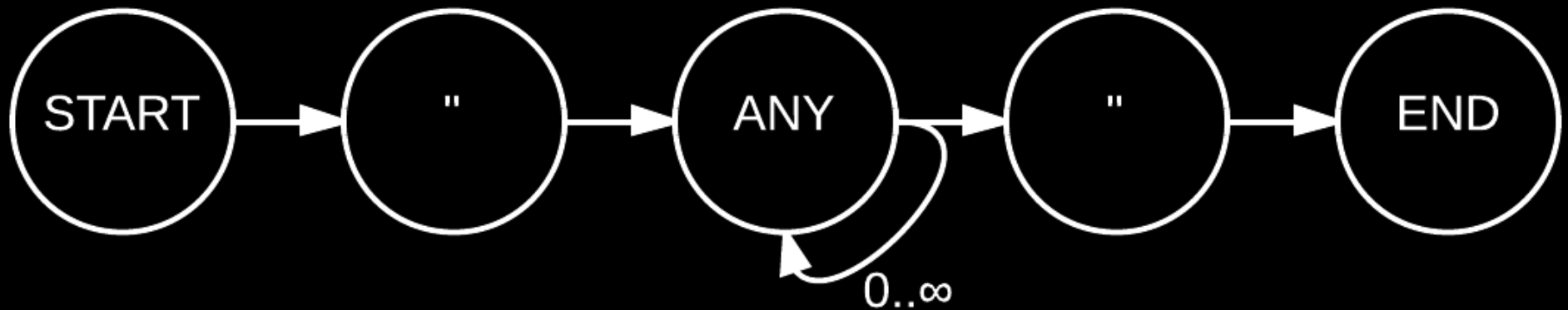
# leniwość

[<<.+?>>]



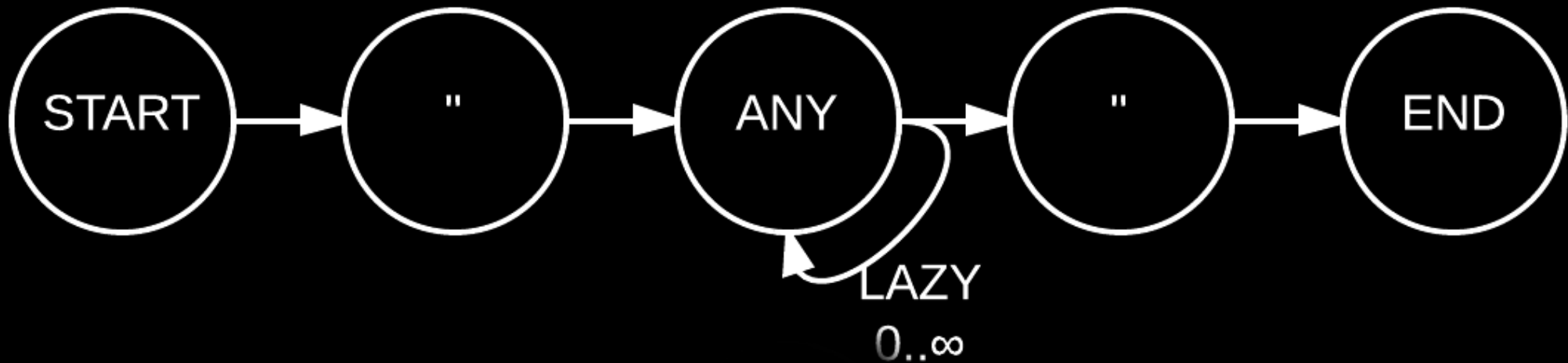
# Inny przykład

[ ".\*"]



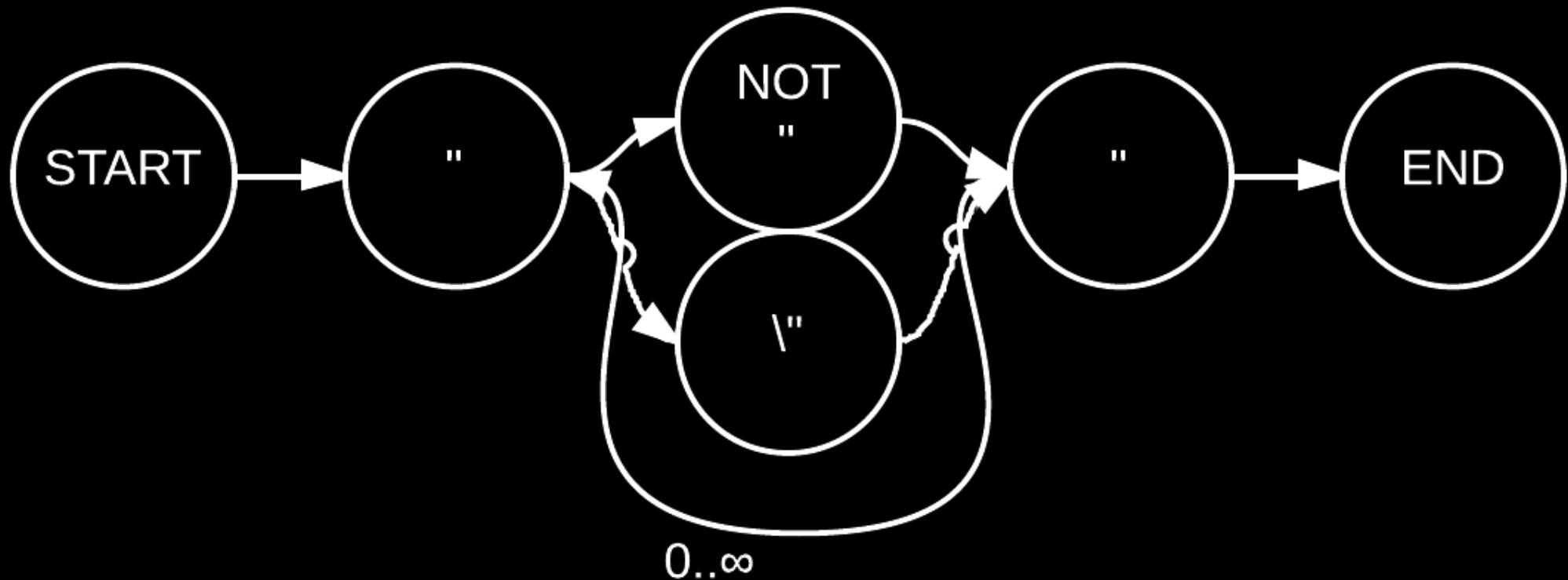
# Inny przykład

[".\*?"]



# Inny przykład

`[\"(\\\"|[^\"])*\"]`

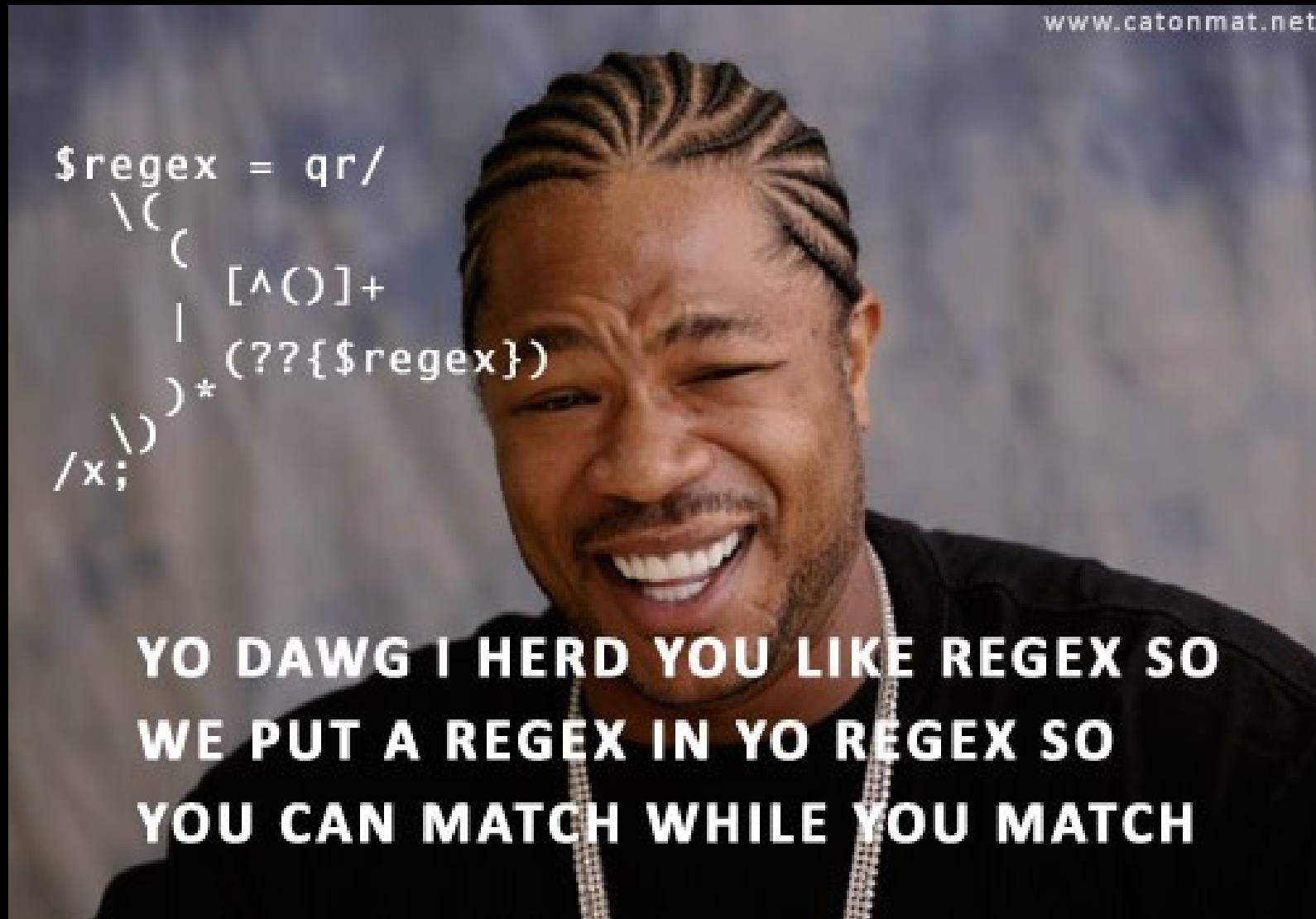


# Odwołania wsteczne

「<\w+>\w+<\^1>」

「ID:([a-f])\1」

# rekursja





# Nie przesadzajmy

[illegible]







Koniec

Pytania?