

Compilers 2021/2022

## **Projekt 2**

Anna Kelm

110455

2 maja 2022

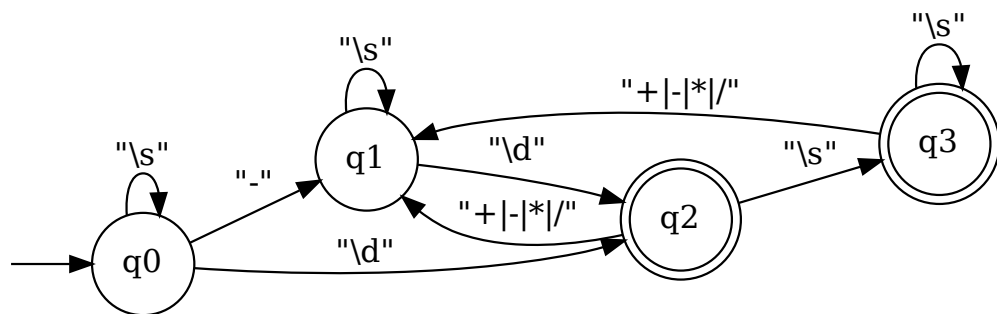
## 1 Cel

Celem projektu jest stworzenie programu obliczającego wyrażenie arytmetyczne (kalkulator).  
Założenia:

- liczby są typu `int`,
- pierwszą liczbę może poprzedzać znak `-`,
- parser ignoruje spacje między liczbami i operatorami,
- dostępne operatory to `+`, `-`, `*`, `/`,
- nie istnieje hierarchia operatorów.

Wyrażenie, które będzie możliwe do obliczenia przez program będzie więc postaci:  
`'-'?\s*\d+\s*(( '+' | '-' | '*' | '/' )\s*\d+\s*)*`.

## 2 Niedeterministyczny automat skończony



Rysunek 1: Automat skończony dla wyrażenia `'-'?\s*\d+\s*(( '+' | '-' | '*' | '/' )\s*\d+\s*)*`.

## 3 Tablica przejść

symbole wejściowe	stany			
	q0	q1	q2	q3
<code>\d</code>	q2	q2	q2	–
<code>'-'</code>	q1	–	–	–
<code>'+'   '-'   '*'   '/'</code>	–	–	q1	q1
<code>\s</code>	q0	q1	q3	q3
akceptujący?	N	N	A	A

Tabela 1: Tablica przejść dla automatu z rys. 1.

Dodanie cyfry do zapamiętywanej liczby następuje przy przejściach:

- $q0 \rightarrow q1$  (obsługa '-' jako pierwszego znaku),
- $q0 \rightarrow q2$ ,
- $q1 \rightarrow q2$ ,
- $q2 \rightarrow q2$ .

Przypisanie operatora zachodzi dla przejść:

- $q2 \rightarrow q1$ ,
- $q3 \rightarrow q1$ .

Ewaluacja wczytanego dotychczas wyrażenia zachodzi:

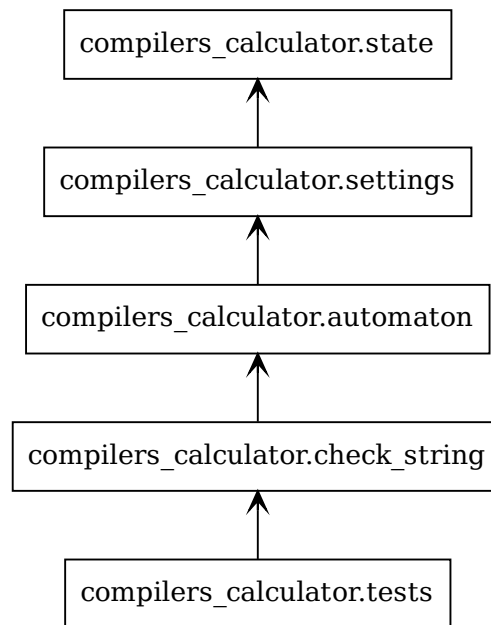
- dla przejść z przypisaniem operatora, przed jego przypisaniem,
- gdy cały tekst zostanie przetworzony, a stan automatu jest akceptujący.

## 4 Struktura projektu

Pliki:

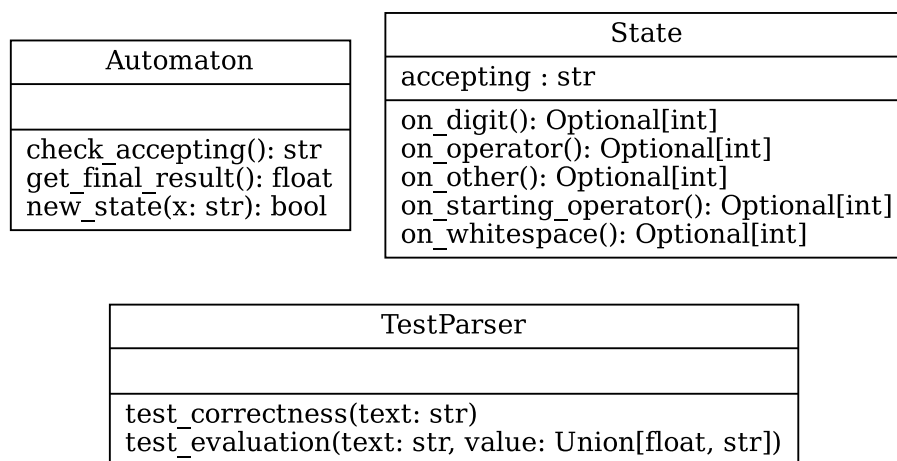
- automaton.py – klasa Automaton, która przyjmuje listę stanów i wykonuje tranzycje.
- check\_string.py – główny plik z funkcjami sprawdzającymi i obliczającymi zawartość łańcucha wejściowego,
- dokumentacja.pdf
- README.md
- requirements.txt
- settings.py – definicje listy stanów i alfabetu,
- state.py – klasa reprezentująca stan z dozwolonymi przejściami,
- tests.py – testy jednostkowe.

Diagram zależności między plikami w projekcie przedstawia rys. 2.



Rysunek 2: Diagram zależności między składnikami pakietu.

## 5 Klasy



Rysunek 3: Diagram klas stworzonych w projekcie.

### 5.1 State

Pola publiczne:

- `accepting` : 'N' lub 'A' w zależności od tego, czy jest to stan akceptujący.

Metody publiczne:

- `__init__(on_digit, on_starting_operator, on_operator, on_whitespace, on_other, accepting)` – konstruktor klasy przyjmujący numer kolejnego stanu w odpowiedzi na

znak odpowiadający, odpowiednio, cyfrze, operatorowi, który może występować na początku tekstu, pozostałym operatorom, białym znakiem oraz pozostałym znakom. Jeśli przejście dla danego znaku nie istnieje, numer zawiera wartość `None`. Parametr `accepting` o możliwych wartościach 'A' lub 'N' wskazuje na to, czy stan jest akceptujący.

- `on_digit` – zwraca numer kolejnego stanu dla cyfry,
- `on_starting_operator` – zwraca numer kolejnego stanu dla operatora, który może występować na początku tekstu,
- `on_operator` – zwraca numer kolejnego stanu dla pozostałych operatorów,
- `on_whitespace` – zwraca numer kolejnego stanu dla białego znaku,
- `on_other` – zwraca numer kolejnego stanu dla symbolu odpowiadającego innym znakom.

## 5.2 Automaton

Metody publiczne:

- `new_state(x)` – próbuje wykonać przejście - zmienić stan automatu dla symbolu `x`. Zwraca 'true' dla sukcesu, 'false' dla porażki (gdy przejście nie istnieje) oraz oblicza na bieżąco wartość wyrażenia,
- `check_accepting` – zwraca wartość pola `State.accepting` dla bieżącego stanu
- `get_final_result` – aktualizuje (jeśli są nieprzetworzone cyfry) oraz zwraca wartość wyrażenia.

## 6 Algorytm

1. Wczytanie tekstu.
2. Zmiana stanu automatu w odpowiedzi na kolejne znaki ze słowa. Cyfry bieżącej liczby, operator oraz bieżący wynik są trzymane w zmiennych. Przy parsowaniu kolejnego operatora, aktualizowana jest wartość: wywoływany jest dotychczasowy operator dla nowej liczby i poprzedniego wyniku. Jeśli dla danego znaku nie istnieje przejście, program zwraca "error" zamiast wyniku.
3. Po przetworzeniu całego tekstu, jeśli automat nie jest w stanie akceptującym, zwracany jest "error". Jeśli jest w stanie akceptującym, wyrażenie jest aktualizowane dla ostatniego operatora i ostatniej liczby, a następnie zwracane.