

# Architektura systému



## Modulární struktura

Systém je rozdělen do několika modulů, zajišťujících flexibilitu a efektivitu.



## Analýza textu

Modul pro syntaktickou analýzu a generování stromových struktur z textů.

- **Výstupy:** Generované stromové struktury.
- **Použité technologie:** .NET



## Detekce vzorců

Tento modul identifikuje opakující se vzory ve stromových strukturách, které lze komprimovat.

- **Výstupy:** Seznam vzorců pro kompresi.
- **Použité technologie:** Algoritmy pro detekci vzorců.



## Kompresní algoritmy

Implementace kompresních metod pro optimalizaci stromových struktur.

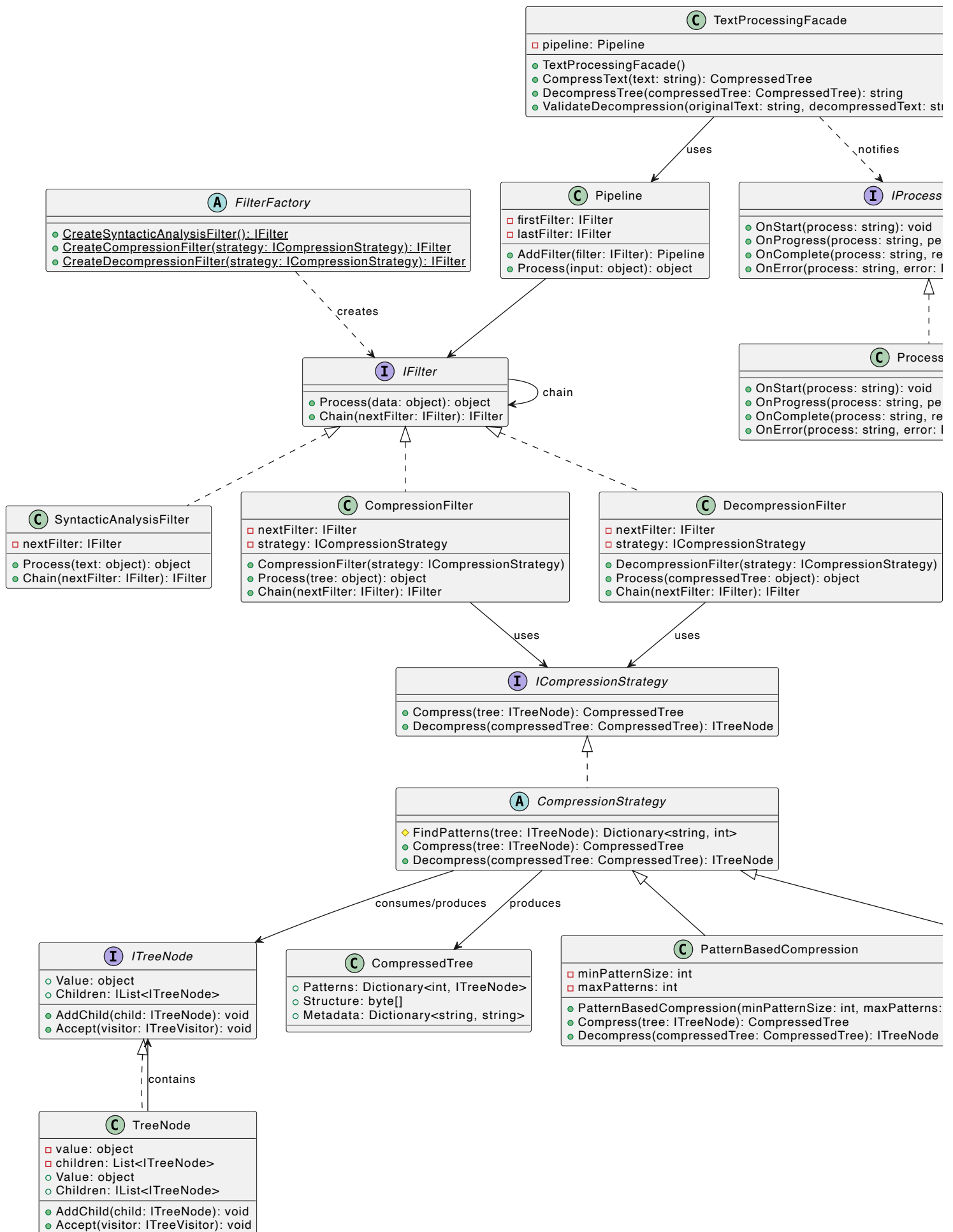
- **Typy algoritmů:** Huffmanovo kódování, LZW.
- **Výstupy:** Komprimované datové struktury.

**Tip:** Každý modul je navržen pro snadnou rozšiřitelnost a testování nových metod.



## Výhody modulární struktury

- **Flexibilita:** Snadno přidáváte nové moduly.
- **Škálovatelnost:** Podporuje práci s různými velikostmi dat.
- **Paralelizace:** Každý modul lze paralelizovat pro zrychlení výpočtů.



# Požadavky na systém

## Funkční požadavky

- **Analýza textu:** Systém musí generovat stromové struktury.
- **Detekce vzorců:** Identifikace opakujících se vzorců ve stromových datech.
- **Kompresní algoritmy:** Aplikace kompresních metod.
- **Dekomprese:** Obnovení komprimovaných stromů.

## Systémové požadavky

Komponenta	Požadavek
Hardware	8 GB RAM, více-jádrový procesor
OS	Linux, Windows, macOS

## Ne-funkční požadavky

- **Výkon:** Schopnost zpracovávat soubory o velikosti několika GB.
- **Modularita:** Snadná rozšiřitelnost o nové moduly.
- **Bezpečnost:** Zajištění ochrany dat.

**Poznámka:** Tyto požadavky zajišťují výkon a flexibilitu pro různé velikosti dat.

# Journal - Komprese stromových struktur

V tomto deníku dokumentuji postup a experimenty, které provádím při vývoji metody pro kompresi stromových struktur v rámci mého projektu. Cílem je optimalizovat způsob, jakým jsou syntaktické stromy reprezentovány, aby bylo možné efektivněji ukládat a zpracovávat jazyková data.





## Cíl projektu

Projekt se zaměřuje na vývoj a testování různých metod pro kompresi syntaktických stromů, které jsou výsledkem analýzy textu. Věnuji se zkoumání různých nástrojů pro **dependency parsing**, zpracování textu a tvorbě kompresních algoritmů, jako je gramatická komprese nebo metoda **RePair**.

## Jak deník pomáhá

Každý zápis v deníku se zaměřuje na konkrétní denní pokrok, nové výzvy, řešení a testování nových metod. Je to pro mě způsob, jak sledovat vývoj projektu krok za krokem a zároveň poskytnout ostatním náhled na to, jak se projekt vyvíjí.




Zápisky obsahují:

-  **Experimenty** s různými nástroji pro syntaktickou analýzu.
-  **Testování metod komprese**, jako je gramatická komprese a RePair.
-  **Problémy**, na které jsem narazil při implementaci, a způsoby, jak je řešit.
-  **Plány do budoucna**, jak projekt posunout dál.

Každý zápis je podrobný a dokumentuje jak technické kroky, tak teoretické úvahy a rozhodnutí.

## Prozkoumejte jednotlivé záznamy

Pokud máte zájem o podrobnosti, můžete si prohlédnout jednotlivé zápisky z deníku podle data. Každý zápis je zaměřen na konkrétní fázi vývoje a řešení, která jsem v daný den řešil.

-  [2025-02-19](#)
-  [2025-02-20](#)
-  [2025-02-23](#)

Tento deník je nejen záznamem pro mě, ale také způsobem, jak sdílet moje pokroky s ostatními a získat zpětnou vazbu na implementované metody.

# Namespace ConsoleApp

## Classes

[Program](#)