

## Úvod

Úlohou je převést vstupní soubor ve formátu XML na sekvenci DDL příkazů s dalšími požadavky danými vstupními parametry. Zvolený jazyk Python dovoluje využívat jak objektově orientovaného paradigmatu, tak strukturovaného programování a k tomu bylo přihlíženo při návrhu. Také byl brán ohled na možnosti zvolených knihoven.

## Zpracování vstupních parametrů

Zpracování parametrů probíhá pomocí knihovny `argparse`. Tato knihovna nabízí velmi jednoduché přidávání parametrů, ale pro potřeby projektu bylo nutné práci s touto knihovnou rozšířit o jiné nastavení návratových kódů a ošetření unikátnosti parametru `--help`.

## Zpracování vstupního souboru

Pro zpracování vstupního XML byla zvolena knihovna `xml.etree.ElementTree`. Ta vytvoří strom z jednotlivých XML elementů a tím je pak postupně iterováno, přičemž pro každého rodiče se vytvoří seznam potomků a dvojice rodič-seznam potomků jsou uchovávány v datovém typu slovník, přičemž pro každý další stejný element jsou do seznamu potomků přidávány pouze ty podelementy, které jsou v kontextu předchozích výskytů navíc.

Dále se pro každý element vytvoří instance třídy `Table` a jsou naplněny atributy této třídy podle atributů příslušného elementu.

## Třída Table

Třída `Table` je umístěna v samostatném modulu `table` v balíku `modules`. Reprezentuje tabulku databáze. Její gettery a settery upravují jednotlivé atributy do požadovaného formátu. Uchovává v sobě jméno dané tabulky, seznam cizích klíčů, slovník atribut:hodnota, dále pro omezování počtu podelementů při generování cizích klíčů v sobě nese také toto omezení jako hodnotu `etc`, pro potřebu přidání cizího klíče předka objekt s předkem a posledním atributem je příznak pro generování klíče z předka. Metoda `getTableStr()` vrací řetězec DDL v požadovaném formátu reprezentující danou tabulku.

Dále modul `table` obsahuje funkci pro získání datového typu. To je prováděno pomocí regulárních výrazů. Tato funkce je volána z metody tabulky pro získání atributů a textových hodnot pro každý atribut zvlášť a její návratová hodnota, kterou je řetězec s názvem datového typu je porovnávána s předchozími atributy podle priorit.

## Výstup

Formát výstupu se řídí především podle vstupních parametrů. Pokud je zadán parametr s hlavičkou, je jako první úkon proveden výpis této hlavičky. Pro tisk jednotlivých tabulek se iteruje přes slovník mající jako hodnoty instance třídy `Table` a do výstupního proudu je poslán řetězec získaný metodou `getTableStr()`. Na základě parametru `-a` je při iteraci stromem elementů přeskočeno nastavování atributů, a proto se negenerují. Podobným způsobem parametr `-b` omezí počet generovaných podelementů na jediný.