

Pro první část kurzu IPP – projekt v PHP mi bylo zadáno XQR, neboli XML query. Principem je aplikace SQL dotazu, definovaného buď jako parametr při volání nebo předaného v souboru, na vstupní XML soubor. Tyto data jsou programu předána na standardní vstup či parametrem a po zpracování jsou poslána buď na standardní výstup nebo do souboru. Dále program dovoluje generovat (nebo generování zakázat) XML hlavičky a obalovat výsledná data do tzv. ROOT tagu. Z časových důvodů nebyla implementována žádná rozšíření.

load\_file(), dump\_output(), extended\_die(), print\_help().

Pro zpracování parametrů z důvodů bezpečného zacházení s vícenásobným použitím téhož parametru není využita žádné prefabrikované řešení, ale vytvořeno vlastní. Parametry zpracovává funkce `arg_parse()`: vytvoří kopii globálního pole s argumenty a v cyklu je postupně od posledního zpracovává a odebírá z pole. Zpracování končí, zbývají-li v poli dva prvky: celé volání a název programu. Zpracované parametry jsou uloženy v asociativním poli. Nabízelo se také řešení vytvořit pro tato data samostatnou strukturu, ale z důvodů jednoduchosti bylo použito pole.

Důležitou součástí skriptu je analýza SQL dotazu. To zabezpečuje funkce `query_parse()`, která na základě gramatiky definované v zadání rozdělí data ve SQL dotazu a uloží je do asociativního pole či případně vypíše chybové hlášení. Analýza gramatiky, správné sémantiky a syntaxe je vyřešena regulárními výrazy. Nejdřív se zkontroluje správné pořadí všech částí dotazu a tyto se následně vyhodnocují každá zvlášť:

Příslušná část SQL dotazu	Regulární výraz
Kontrola syntaxe	/ (SELECT.* (LIMIT)? .* FROM.* (WHERE)? .* (ORDER BY)? .*) /
SELECT	/ (?<=SELECT\s) (?:\s*).+ ? (?=\s+LIMIT\s+FROM) /
SELECT ověření	^w*/
LIMIT	/ (?<=LIMIT\s) (?:\s*).+ (?=\s+FROM) /
LIMIT ověření	^d*/
FROM	/ (?<=FROM\s) (?:\s*).+ ? (?=\s \$) /
FROM ověření	^w*(?:\.)? \w+ /
WHERE	/ (?<=WHERE\s) (?:\s*).+ (?=\sORDER BY \$) /
WHERE jednotlivé složky	^(?:(\s*)(?:NOT\s+)* (\w+) ?(?:\.\(\w+\)?) ?\s+(CONTAINS = > <) \s+(\\\\\"(?:.*)\\\\\\" -?'d+ b \\'\"'\"')\$)
ORDER BY	/ (?<=ORDER BY\s) \w+ (?:(?:\.\(\w+\) (?:(ASC DESC))\$)

Za druhou nosnou část skriptu je zodpovědná funkce `proceed_query()`. Ta se stará o správnou filtraci dat ve vstupním XML. K tomu využívá velmi schopné funkce z knihovny SimpleXML – funkce `xpath()`, která umožňuje adresování do XML. Postupně z XML tedy filtruji data, nejdřív část FROM, poté SELECT, WHERE a nakonec LIMIT – tedy volám `xpath()` čtyřikrát. Obzvláště skládání WHERE části je složité a také naráží na limitace `xpath()` tak, jak je v SimpleXML (XPath 1.0) definováno. Například neumí porovnávat řetězce. U LIMIT jsem zvažoval dvě různé možnosti řešení, buď zastavit hledání a generování po určitém počtu výsledků a nebo vyfiltrovat vše, a pak výsledek jen omezit na požadovaný počet a ten vypsát. Druhé řešení není vůbec optimální, protože při větším objemu bude pomalejší, nicméně je to řešení jednoduché na implementaci a proto bylo použito.