

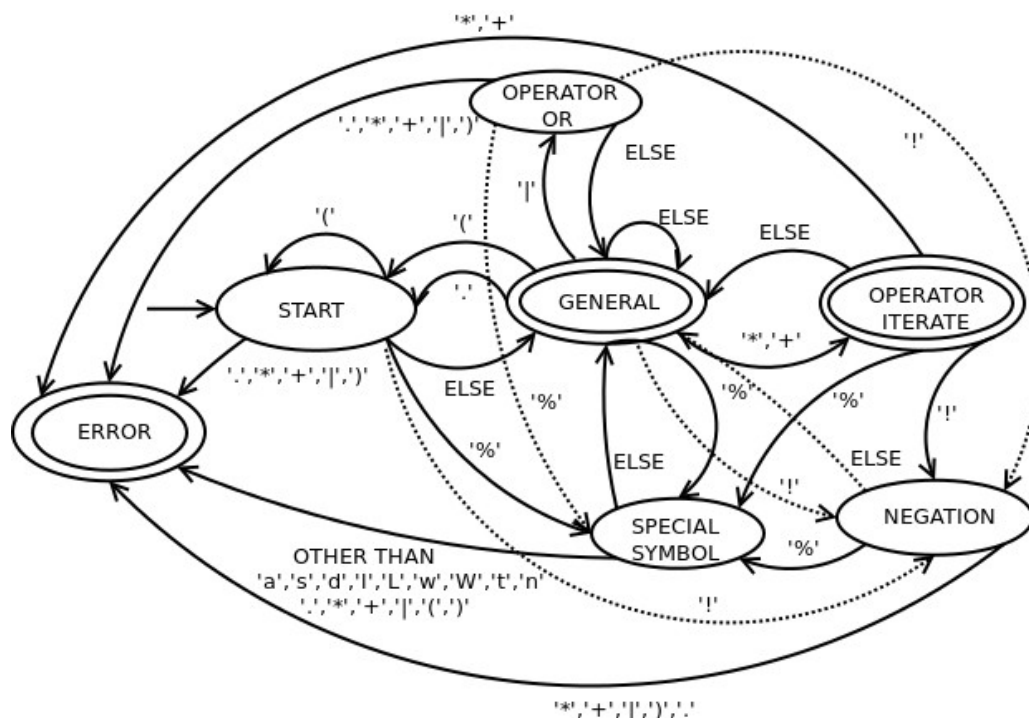
Cieľom projektu je vytvoriť skript, ktorý zvyrazňuje časti textu zadané regulárnymi výrazmi. Na jednotlivé časti textu sa aplikujú HTML elementy definované formátovacími príkazmi. Výstup môže byť presmerovaný na základe zadaných prepínačov do súboru, alebo na štandardný výstup.

Parametre sú spracovávané pomocou funkcie `getopts`, ktorá jednotlivým parametrom priradí hodnoty. Následne sa skontrolujú, ktoré parametre neboli zadané, prípadne zadané viac-krát cyklom prechádzajúcim cez zadané argumenty príkazového riadku.

Formátovací súbor je spracovávaný znak po znaku, pričom sa zapisuje do reťazcov `regexp` a `formatexp` odpovedajúce znaky. Podľa toho, či sa načítavaný znak nachádza pred alebo za oddelovačom `'\t'` sa rozhodne do ktorého reťazca sa bude zapisovať. Následne sa vytvorí tabuľka regulárnych výrazov a ku každému regulárnemu výrazu postupnosť formátovacích príkazov.

Vzhľadom na to, že definícia regulárnych výrazov zadaných projektom a výrazov v jazyku PHP je odlišná, bolo vstupný výraz potrebné transformovať na výraz vyhovujúci PHP. Na prevod bol použitý konečný automat so siedmimi stavmi. Ak automat skončí v nekonečnom stave, alebo v stave `ERROR` program končí s návratovou hodnotou 4. Počas prevodu sa kontroluje aj správne uzátvorkovanie výrazu pomocou zásobníku. Pre správnu funkcionálnosť bolo potrebné znaky, ktoré sú špeciálne pre regulárne výrazy v PHP a bežné pre výrazy zadané projektom správne ošetriť a to pridaním znaku `'\'` pred tieto znaky.

Pre vyhľadávanie častí textov vyhovujúcich regulárnym výrazom je použitá vstavaná funkcia `preg_match_all`, ktorá vráti nájdený reťazec a jeho pozíciu vo vstupnom súbore. Vloženie odpovedajúceho HTML elementu sa zmenia indexy výstupného reťazca, preto je pre riešenie použité pomocné pole `indexArray`, do ktorého sa ukladajú indexy do výstupného poľa obsahujúceho HTML elementy. S poľom `indexArray` je možné zistiť index do výstupného reťazca z pozície, ktorú nám vráti funkcia `preg_match_all`. Zadanie vyžaduje aby nové uzatvárané HTML elementy boli pred staršími. To je docielené overením rozdielu dvoch susedných indexov. Ak je rozdiel medzi indexami v `indexArray` väčší ako jeden, nachádza sa tam starší HTML element a aktuálny sa vloží na predchádzajúcu pozíciu.



Obrázky 1