

# XML, czyli drzewa z danymi

## Pierwsza praca domowa

Filip Murlak      Paweł Parys

Utworzone: 11 kwietnia 2011. Termin: **22 kwietnia** 2011

Każde zadanie będzie ocenione w skali 0-5. Odpowiedź bez komentarzy (dowodu) da nie więcej niż 3 punkty.

1. Podaj najlepsze ograniczenie górne (algorytm) na złożoność następującego problemu: dla danych DTD  $D_1$  i  $D_2$  sprawdzić, czy każde drzewo spełniające  $D_1$  spełnia  $D_2$ .
2. *Plonem* drzewa  $t$ , ozn.  $\text{pl}(t)$ , nazywamy słowo powstałe przez wypisanie etykiet jego liści kolejno od lewej do prawej. Dla języka drzew  $L$  definiujemy  $\text{pl}(L) = \{\text{pl}(t) \mid t \in L\}$ .
  - (a) Dla danego języka bezkontekstowego  $P \subseteq \Sigma^*$  skonstruuj taki automat skończony na drzewach nieurangowanych  $\mathcal{A}$ , że  $\text{pl}(L(\mathcal{A})) = P$ .
  - (b) Dla danego języka regularnego  $R \subseteq \Sigma^*$  skonstruuj taki automat skończony na drzewach nieurangowanych  $\mathcal{A}$ , że  $L(\mathcal{A}) = \{t \mid \text{pl}(t) \in R\}$ .
  - (c) Podaj przykład takiego języka bezkontekstowego  $S \subseteq \Sigma^*$ , że język drzew binarnych nad  $\Sigma$ , których plon jest w  $S$ , jest nieregularny.
3. Zaprojektuj wielomianowy algorytm rozstrzygający następujący problem: dla danego DTD  $D$  i zapytań boole'owskich  $q, q' \in \text{XPath}(/, //, *)$  rozstrzygnij, czy  $q \subseteq_D q'$  (tzn. czy każde drzewo zgodne z  $D$  i spełniające  $q$  spełnia również  $q'$ ).

Wskazówka: zapytania  $\text{XPath}(/, //, *)$  nie używają spójników, ani filtrów, więc mają postać pojedynczej ścieżki.
4. Czy dla każdego zapytania  $q \in \text{NavXPath}$  istnieje równoważne zapytanie  $q' \in \text{NavXPath}$ , które nie korzysta z osi `previous-sibling` ani `next-sibling` (uwaga: osie `preceding-sibling` oraz `following-sibling`, chodzące w poziomie o wiele kroków, nadal są dostępne; niedostępne są tylko osie chodzące o 1 krok).
5. Zaprojektuj algorytm liniowy ze względu na  $|t| + |S|$  oraz wielomianowy ze względu na  $|q|$  rozstrzygający następujący problem: dla danego dokumentu  $t$  oraz wyrażenia ścieżkowego  $q \in \text{NavXPath}$  wyznacz zbiór  $S$  wszystkich par wierzchołków spełniających  $q$ .