

Приликом рада са реалним текстовима издвајање "поглавља-покољења" је понекад скопчано са тешкоћама, због чега смо се у таквим случајевима ограничавали само на приближно разбијање текста на делове.

Нека текст  $X$  описује догађаје на довољно дугом временском интервалу  $(A, B)$ , у чијем току се променило најмање неколико покољења учесника. Нека је текст  $X$  разбијен на "поглавља-покољења"  $X(T)$ , где је  $T$  редни број покољења (описаног у тексту  $X(T)$ ) у оној нумерацији "поглавља" која је забележена у датом тексту. Питање: да ли су исправно нумерисана (уређена) та "поглавља" у тексту, или пак, ако је та нумерација изгубљена или сумњива, како је реконструисати, како исправно уредити "поглавља" у времену? Формулишимо модел: пуно име = личност. Шта то значи? Ако је временски интервал који хроничар описује дуг, онда, како смо проверили анализом великог броја историјских текстова, у највећем броју случаја учесници имају у тексту различита пуна имена. Пуно име се може састојати од неколико речи (на пример, Александар Македонски). Дакле, број разних лица са истим именима је ништавно мали у односу на број свих личности текста. То је тачно за неколико десетина текстова које је аутор истражио, који описују историју Грчке, Немачке, Италије, Русије итд. Заиста, хроничар је заинтересован за разликовање разних личности своје хронике, ради избегавања забуне. Најједноставнији начин да се то достигне је дати разним лицима различита пуна имена. Ово једноставно психолошко објашњење се потврђује нашим прорачунима. Формулишимо сада принцип пригушења фреквенција, који описује хронолошки исправан поредак "поглавља".

Код правилне нумерације "поглавља-покољења" хроничар, прелазећи са једног покољења на следеће мења и личности, наиме приликом описа претходника покољења са редним бројем  $T_0$  он ништа не говори о личностима тог покољења (пошто се оне још нису родиле), затим приликом описа покољења  $T_0$  хроничар управо ту највише говори о личностима тог покољења, пошто су за њих везани догађаји које описује; најзад, прелазећи на опис следећих покољења хроничар све ређе помиње раније учеснике, пошто описује нове догађаје чији учесници потискују умрле.

Укратко: свако покољење рађа нове историјске личности, приликом смене генерација та се лица мењају, старе личности се заборављају. Пошто су учесници скоро једнозначно одређени пуним именима (име = личност), у даљем ћемо изучавати списак свих имена у тексту. Више од тога, испоставља се да је највећи део имена "једноставан", састављен од једне речи, зато приликом обраде великих текстова са великим бројем имена можемо уочити само "елементарна имена - цигле", разлажући ретка пуна имена на њихове појединачне компоненте.

Уочимо групу имена која су се први пут појавила у тексту у "поглављу-покољењу" са редним бројем  $T_0$ . Број свих помињања (са вишеструкостима) тих имена у "поглављу" обележимо са  $K(T_0, T_0)$ . Избројмо затим колико се пута иста та имена помињу у "поглављу" са редним бројем  $T$ . Добијени број обележимо са  $K(T_0, T)$ . При томе, ако се исто име понавља неколико пута (тј. вишеструко), бројимо сва таква појављивања. Нацртајмо график, носећи на