

## Становище

от доц. Христо Александров Ганчев  
относно дисертацията на ас. Стефан Владимиров Герджиков  
*„Ефективен алгоритъм за приближено търсене в регулярни множества”*,  
представена за придобиване на образователната и научна степен „доктор”

Представената дисертация е в обем 197 страници и съдържа увод, основен текст (8 глави с обем 164 страници), апендикс, заключение и библиография. Основният проблем, разгледан в дисертацията, е ефективното намиране на всички думи от даден език, близки до дадена дума. Този проблем се появява в най-различни приложни задачи – от разпознаването на човешка реч, през автоматичното коригиране на текст, до сигурното предаване на информация. Обичайно две думи се наричат близки, ако те могат да се получат една от друга чрез достатъчно малък брой поправки, включващи триене, заместване или прибавяне на букви. Точно колко малък е допустимият брой на поправките зависи от конкретната задача, като в зависимост от спецификата на задачата всяка корекция носи определен брой наказателни точки. Обичайно при общото математическо разглеждане се използва така нареченото Левенщайн разстояние, което брой минималния брой наказателни точки, необходими за получаването на една дума от друга, като две думи се считат за близки, ако Левенщайн разстоянието е по-малко от дадена фиксирана константа или ако е по-малко от дадено число, пропорционално на дължините на думите.

Първите пет глави от дисертацията са посветени на представянето на алгоритъм, който намира всички думи от даден регулярен език, намиращи се на Левенщайн разстояние от дадена дума, по-малко или равно на фиксиран процент от дължината на думата, като в първа и втора глава се въвеждат необходимите понятия, в трета глава се дава общата идея за алгоритъма, в четвърта глава се прави анализ на свойствата на подравняванията между две думи, а в пета глава е изложен самият алгоритъм. В шеста глава се разглежда едно обобщение на Левенщайн разстоянието, в което освен отделни замествания на буква с буква се позволява и заместването на дадена последователност с друга, като е предложена модификация на алгоритъма от предната глава, която да работи върху обобщеното разстояние. В седма глава се анализира времевата сложност на алгоритъма от шеста глава (което разбира се дава и съответния анализ на алгоритъма от пета глава). Установява се, че при фиксирани регулярен език и коригиращи правила времето за генериране на всички думи от регулярния език, близки до дадена дума, е пропорционално на дължината на думата, като коефициентът на пропорционалност зависи от сложността на крайния автомат, представляващ езика, коригиращите правила и понятието за близост, което се използва. В така направената оценка, в някои ситуации коефициентът на пропорционалност би могъл да бъде и неопределен. За това дисертантът дава някои достатъчни условия, за това изразът, определящ този коефициент, да има смисъл.

В последната глава е представен алгоритъм с чиято помощ от едно мулти-множество с двойки от думи, за които знаем, че са близки, се извличат коригиращи правила и съответните им цени. Допълнително е представен и алгоритъм, който въз основа на извлечените правила генерира списък от думи, близки до дадена дума, като е пресметнато че времето за изпълнението на алгоритъма зависи линейно от дължината на думата.

Като цяло

Резултатите от дисертацията са публикувани в три статии, от които една самостоятелна в списание с импакт фактор, а другите две са в съавторство, като едната е публикувана в рецензирания сборник от статии на конференцията ICDAR 2013, а другата е съкратен вариант на чернова на статия, описваща подробно алгоритъма за приближено търсене, който съкратен вариант авторите представят на международното състезание String Similarity Search/Join Competition, Genoa 2013. Забележително е, че този алгоритъм печели първо място на състезанието, с което доказва напълно своята ефективност.

Авторефератът към дисертацията пълно и правилно отразява приносите в нея.

**Заключение:** Считаю, че представената дисертация съдържа високи научни постижения и напълно отговаря на критериите на ЗРАСРБ, Правилника за неговото прилагане, Правилника на СУ „Св. Кл. Охридски” за условията и реда за придобиване на научни степени и съответният правилник на ФМИ за придобиване на научна степен „доктор”. Ето защо уверено препоръчвам на уважаемото жури да присъди на ас. Стефан Владимиров Герджиков образователната и научна степен „доктор”.

София,  
02.01.2014 г.

Христо Ганчев