

استمارة مستخلصات رسائل و أطاريح الماجستير و الدكتوراه في جامعة البصرة

اسم الطالب: مصطفى عيسى لطيف عبدالله
اسم المشرف: أ.م.د سلمى عبدالباقي محمود
الشهادة: بكالوريوس علوم حاسبات

الكلية: كلية العلوم
القسم: علوم الحاسبات
التخصص: معالجة لغات طبيعية
عنوان الرسالة أو الإطروحة:

استخدام نظرية التمرکز لتحليل الإحالة المشتركة في النصوص العربية

ملخص الرسالة أو الإطروحة:

هذا البحث هو برمجة الإحالة المشتركة لنصوص مكتوبة باللغة العربية واقتراحت طريقتان: طريقة الاحالة لأقرب محال يتطابق بالخصائص, وهو مبدأ معتمد في النصوص العربية, والطريقة الأخرى هي باستخدام نظرية التمرکز لمعالجة الإحالة في النصوص العربية. كشف تطبيق النظرية على النصوص العربية عدد من نقاط الضعف والمشاكل وجرت محاولة اقتراح حلول مناسبة لها, وهذا ما دُرِس في الفصل الثالث بالتفصيل. طُبِق النظام على مدونتين نصيتين: مدونة قصصية تشمل 100 قصة أطفال. ومدونة رياضية تشمل 100 نص وجرت برمجة النظام باستعمال لغة Visual basic.net لما تقدمه من إمكانات برمجية ودوال جاهزة لمعالجة الخيوط الرمزية. بشكل عام كانت نتائج النظام مقبولة ولكنها متفاوتة ومختلفة بين الطريقتين من نص إلى آخر لعدة أسباب جرت مناقشتها في الفصل الرابع. وأخيراً, لا يخلو أي عمل من نقاط ضعف ولا يمكن أن يكون أي عمل متكامل من كل الوجوه فلا بد من وجود نقص يقوم الآخرون بدراسته والبحث فيه وهذا ما يساهم في استمرارية البحث العلمي.

College: Science college

Name of student: Mustafa Issa Lateef

Dept: Computer science

Name of supervisor: Ass.Prof. Salma Abdulkali

Mahmood

Specialization: Natural Language Processing
science

Certificate: B.Sc. of computer

Title of Thesis:

Using Centering Theory for Coreference Resolution of Arabic texts

Abstracts of Thesis:

Current research topic is programming the coreference resolution for Arabic written texts. Two method were proposed the first method is reference to closest element that matches the features, and other method is by using centering theory. When applied the centering theory of Arabic texts revealed a number of weaknesses and problems and had suggested suitable solutions to them. This is discussed in detail in chapter 3. The system has been applied to two text corpora – the story corpora include 100 children's stories, and a sports corpora include 100 text, the system was programmed using Visual basic.net, because it offers the potential of software and ready-made functions to deal with strings. In general the results of the system were acceptable but varied and different between the two methods and between the texts for several reasons discussed in chapter 4. Finally, there is no work without weaknesses and there is no integrated work in all respects. There must be a shortage of others to study and research. This contributes to the continuity of scientific research.

