برنامج (مداد)
للتحليل الصرفي للكلمات العربية

محمود محمد صابر **عبد المنعم فريد عبد المنعم**

شركة مداد لتقنية المعلومات

Abdo_fw@hotmail.com
Abdo_fa_M@yahoo.com

تقديم:

مر أكثر من أربعين عامًا تقريبًا على ابتكار الحاسوب الشخصي، وعلى الرغم من أنه كان حصيلة قرون من التطور المعرفي والتقني الذي شاركت فيه الإنسانية جمعاء، إلا أن مبتكريه في الولايات المتحدة الأمريكية لم يأخذوا بالاعتبار أثناء تصميمه سوى حاجات أسواقهم وثقافتهم ولغتهم، ولم يأخذوا في اعتبارهم فكرة انتشاره على مستوى العالم. ولكن هذا الانتشار قد حدث بالفعل، وذلك بعد دخول الحاسوب في كل المجالات وكل البلدان، وتقدم برمجيات الحوسبة اللغوية للغة الإنجليزية خاصة ولغات أخرى مثل: (الألمانية والفرنسية واليابانية... وغيرها). وتبعا لهذا فقد برزت الحاجة لتطويعه للّغات عالمية الأخرى منها اللغة العربية. ومن ثم ظهرت صعوبات في إمكانية تنفيذ ذلك على أنظمة لم تصمم أصلًا للغات مختلفة ولكنها تيسرت نسبيًا بإضافة برمجيات وتقنيات ملحقة خاصة بثقافات ولغات أخرى، كانت اللغة العربية واحدة منها. ولا ننكر محاولات شركات القطاع الخاص بالمنطقة العربية في عمل هذه المحللات بمهارة حيث يتوفر لها الدعم بكافة صوره إلا أن الجهود البحثية مازالت تحبو ببطء لحاجة هذه الأعمال للدعم المادي والمعنوي والبشري. وقد قمنا بعمل دراسة تقيمية للمنتجات المحللات اللغوية العربية السابقة الموجودة بالسوق العربي وحصر المشاكل والعيوب قدر الإمكان. ونقدم في هذا البحث المحلل الصرفي الذي قمنا بتصميمه لنتلافى العيوب وقد أطلاقنا عليه اسم «مداد». ولقد راعينا في هذا المحلل عند بنائه الإلمام التام بكافة قواعد الإعلال والإبدال، وكذلك راعينا مواضع الهمزات في الجذر الواحد وكيفية الاشتقاق منهما (الجذور التي بها إعلال أو إبدال في مشتقاتها أو همزات في أصولها)، مع مراعاة الجانب السماعي والذي بذل فيه فريق البحث والإعداد جهدًا مضنيًا ليصل للمستوى المطلوب.

برنامجنا المقترح

يهدف هذا المشروع إلى تصميم وإنجاز محلل صرفي للغة العربية قادر على تجزيء وتفكيك النصوص العربية إلى كلمات وتحليلها للتعرف أنواع الكلمة وذلك وفقًا لقسم الكلمة فهناك ثلاثة أقسام للكلمة كالتالي:

١. قسم الاسمية.

٢. قسم الفعلية.

٣. قسم الحروف ، والصفات، والأدوات.

ونتوخى في المرحلة الأولى من هذا المشروع معالجة الأقسام الثلاثة للكلمة انطلاقًا من نظام تحليلي يقوم على أساس التعرف على الكلمة المدخلة إليه وتحديد نوعها، ثم تحديد الزوائد من سوابق ولواحق وحروف مزيدة، ثم الوصل إلى الجذر، ومع عرض كافة التشكيلات الممكنة للكلمة. وقد أخذنا في الاعتبار مجموعة من الأسس التي كانت حتى يكون هذا المحلل قادر على التنافس أمام العديد من المحللات الأخرى من خلال ثلاثة أهداف وهي:

١.الدقة.

٢. صغر الحجم.

٣.السرعة.

ولقد قمنا بحصر قواعد صرفية منها:

١.قواعد اشتقاق الأفعال وتصريفها.	٢. قواعد معالجة الأفعال.
٣. قواعد اشتقاق الأسماء المشتقة وتصريفها.	٤. قواعد معالجة الأسماء المشتقة.
٥. قواعد اشتقاق المصادر وتصريفها.	٦. قواعد معالجة المصادر.
٧. قواعد الإعلال والإبدال.	٨. قواعد الهمزة.

فمثلًا عند إدخال كلمة (أفستتذكّروننا) فنحصل على معظم سمات الكلمة وخصائصها من حيث وزنها وجذرها ومعرفة السوابق واللواحق على هذا الجذر ونوع الكلمة من ناحية كونها فعلًا أو اسمًا، وتصريف فعلها في الأزمنة الثلاثة والمصادر المحتملة للكلمة أي إظهار توصـــيف كامل لها بما يمكننا من إحاطة كاملة بالكلمة وكل ما يتعلق بها. أي أن دور هذا المحلل يتكون من شقين في حقيقة الأمر ؛ محلل صرفي ومولد اشتقاقي في نفس الوقت.

فكما في الشكل السابق حصلنا على النتيجة التالية:

الجذر: ذ ك ر

النوع: فعل مضارع ثلاثي مزيد بحرفين لازم ومتعدًّ

الوزن: أَفَسَتَتَفَعَّلُونَنَا

السابقة: أَفَسَ

اللاحقة: ونَنَا – 'ونَ' فاعل جمع مذكر مخاطب – 'نَا' مفعول به جمع متكلم

الحالة الإعرابية: مرفوع بثبوت النون

الباب التصريفي: تَفَعَّلَ – يَتَفَعَّلُ أزمنة الفعل: تَذَكَّرَ – يَتَذَكَّرُ – تَذَكَّرْ

وكما تلاحظ في الكلمة الســـابقة كان لها احتمال واحد، أما إذا كان لها أكثر من احتمال فيتم ظهور النتيجة كما في الشكل التالي:

نتائج تحليل كلمة " اطَّلَعُوا":

احتمال (١) **اطَّلَعُوا** الجذر: ط ل ع النوع: فعل ماضٍ ثلاثي مزيد بحرفين لازم ومتعدًّ الوزن: افْتَعَلُوا السابقة: اللاحقة: وا – 'وا' فاعل جمع مذكر غائب الحالة الإعرابية: مبني على الضم الباب التصريفي: افْتَعَلَ – يَفْتَعِلُ **أزمنة الفعل: اطَّلَعَ – يَطَّلِعُ – اطَّلِعْ**	**احتمال (٢)** **اطْلُعُوا** الجذر: ط ل ع النوع: فعل أمر ثلاثي مجرد لازم ومتعدًّ الوزن: افْعُلُوا السابقة: اللاحقة: وا – 'وا' فاعل جمع مذكر مخاطب الحالة الإعرابية: مبني على حذف النون الباب التصريفي: فَعَلَ – يَفْعُلُ **أزمنة الفعل: طَلَعَ – يَطْلُعُ – اطْلُعْ**
احتمال (٣) **اطْلَعُوا** الجذر: ط ل ع النوع: فعل أمر ثلاثي مجرد متعدٍّ لمفعول الوزن: افْعَلُوا السابقة: اللاحقة: وا – 'وا' فاعل جمع مذكر مخاطب الحالة الإعرابية: مبني على حذف النون الباب التصريفي: فَعَلَ – يَفْعَلُ **أزمنة الفعل: طَلَعَ – يَطْلَعُ – اطْلَعْ**	**احتمال (٤)** **اطْلَعُوا** الجذر: ط ل ع النوع: فعل أمر ثلاثي مجرد متعدٍّ لمفعول الوزن: افْعَلُوا السابقة: اللاحقة: وا – 'وا' فاعل جمع مذكر مخاطب الحالة الإعرابية: مبني على حذف النون الباب التصريفي: فَعِلَ – يَفْعَلُ **أزمنة الفعل: طَلِعَ – يَطْلَعُ – اطْلَعْ**

احتمال (٥) **اطَّلِعُوا** الجذر: ط ل ع النوع: فعل أمر ثلاثي مزيد بحرفين لازم ومتعدٍّ الوزن: افْتَعِلُوا السابقة: اللاحقة: وا – 'وا' فاعل جمع مذكر مخاطب الحالة الإعرابية: مبني على حذف النون الباب التصريفي: افْتَعَلَ – يَفْتَعِلُ **أزمنة الفعل: اطَّلَعَ – يَطَّلِعُ – اطَّلِعْ**	

كما كان هناك معيار مهم للتعامل مع هذا المحلل، وهو فصل البيانات عن الكود، فكان من أهم الأهداف أن تستطيع التعامل مع عنصري المحلل كلًا على حدا، فعندما يكون هناك خطأ ما سينحصر هذا الخطأ في البيانات المدخلة للبرنامج، أو عدم وضع قاعدة ما دخل الخوارزميات، فإذا كانت القاعدة البيانات تحتوي على كافة البيانات، وكانت القواعد حصر لكافة القواعد الصرفية، قد تصل النتيجة إلى ١٠٠% وهذه النتيجة عكسه فكلما كان هناك خطأ في أحد عنصري البرنامج كلما قلت هذه النسبة على حسب نسبة الخطأ.

كما جعلنا قاعدة البيانات تحتوي على الجذور والكلمات الشاذة فقط، أما باقي الكلمات التي تأتي على حسب القواعد فيستطيع البرنامج التعرف عليه دون الرجوع إلى قاعدة البيانات، وهذه الطريقة هي المؤثر الرئيس في صغر الحجم، والسرعة.

أهدفنا المستقبلية

من أهم الأحلام التي نسعى إليها جاهدين، هو التعامل مع اللغة العربية من حيثُ المعنى، فمثلًا عندما يكتب المستخدم «فن حديث»... لن يخرج له النتائج التي تتحدث عن «الأحداث الحالية»؛ لأن المحلل مزود بخاصية «البحث بالمعاني» التي تفهم معاني الكلمات قبل أن تخرج النتائج، وهذه الجزئية نعمل عليها الآن للوصول إلى المحلل المطلوب، وقد نرى من هذه المعالجة أنه يتم تضييق نطاق النتيجة لما هو أقرب للمعنى المطلوب من المستخدم، وبذلك يتم الحصول على المعلومات المطلوبة بدقة أكثر كثيرًا من استخدام القواعد الصرفية للغة فقط. ويقوم النظام المطوَّر بعرض المعاني المختلفة لكل كلمة بحث يتم إدخالها وعرضها على المستخدم، ليقوم باختيار المعنى الذي يقصده.

ولتحقيق هذه الخطوة، يتم تطوير برنامج يقوم بمعالجة معاني الكلمات الموجودة في المعجم الذي يستخدمه المحلل الصرفي ويقوم بالتفاعل مع المحلل الصرفي، وهذه المعالجة سيكون لها الأثر العظيم في برامج: (التصحيح الآلي – والمشكل الآلي—والترجمة الآلية – وفك اللبس...) الخ.

أهداف المحلل

١. تقديم أداة معلوماتية للتعامل مع النصوص العربية (غير المشكولة) وتفكيكها إلى عناصرها الأساسية.

٢. تمكين المتعلمين والباحثين من استعمال هذا المحلل في تعلم الصرف العربي والبحث فيه من خلال التعرف على الجذوع وما يلحقها من سوابق ولواحق.

٣. وضع هذا المحلل رهن إشارة المؤسسات التعليمية بمختلف أسلاكها للاستفادة منه.

٤. المساهمة في تطوير البحث المعلومياتي في اللغة العربية بجعل هذا المحلل أداة أساسية للمحللات التركيبية والدلالية والمترجمات وغيرها.

٥. الإسهام في تحصين اللغة العربية، وإتاحتها الفرصة للاستفادة من الأدوات والتكنولوجيا الخاصة بالصناعة اللغوية.

الانعكاسات العلمية

١. يعتبر هذا المحلل بمثابة أداة للقيام بتطبيقات أخرى في ميادين مختلفة: المترجم الآلي، المولد الآلي، المفهرس الآلي، المشكل الآلي، الواسم الآلي،...

٢. أداة لمساعدة المتعلمين للغة العربية من خلال وضع هذا المحلل رهن إشارة المؤسسات التعليمية بمختلف أسلاكها للاستفادة منه.

٣. توفير أداة جديدة وعصرية للتعامل مع الكلمة في اللغة العربية.

وبعد؛ فيُعدُّ هذا المحلل بحق نقلة كبيرة جدًّا في مجال معالجة اللغة العربية حاسوبيًّا؛ حيث إننا عند الشروع في هذا العمل كان الهدف الرئيسي لنا هو تحويل القواعد الصرفية إلى قواعد رياضية منطقية، مما يتيح للحاسب أن يتعامل مع اللغة بشكل يحاكي الطريقة التي يستخدمها الإنسان عند تحليله للمفردات، وهذا الأسلوب في التفكير هو الذي جعلنا نستطيع أن نختزل البيانات لدينا لتكون عبارة عن جذور اللغة وشواذها فقط، مما وفر علينا إنشاء قاعدة بيانات ضخمة تحوي معظم مفردات اللغة، وقد ترتب على ذلك صغر حجم البرنامج بدرجة كبيرة جدًّا لم يسبقنا إليها أحد مع الدقة المتناهية في النتائج بالإضافة إلى السرعة الفائقة في استخراج النتائج، مما يجعله الأفضل بين قرنائه حتى الآن على حد علمنا.