## الملخص

يهدف علم الذكاء الاصطناعي (Artificial Intelligence) إلى بناء أنظمة حاسوبية ذكية تحاكي الى حدٍ ما السلوك الذكي للإنسان ، ومن بين جميع أنواع السلوك الذكي تبرز اللغة كأكثرها تجلياً ووضوحاً من حيث إدخالها وإخراجها ، فكل ما ينطق به الإنسان و ما يسمعه محسوس ومدرك بصورة مباشرة. أدى هذا الهدف الى ظهور علم اللسانيات الحاسوبية (Computational Linguistics) كأحد فروع علم الذكاء الاصطناعي ، هدفه الأساس وضع وبناء نماذج حاسوبية لتوليد وتحليل (فهم) اللغة الطبيعية ، وأيضاً ظهور علم معالجة اللغات الطبيعية (Natural Language Processing) كتطبيق عملي له ، حيث يهتم الأخير بكيفية إيجاد السبل والوسائل التمثيلية والعلاجية الحاسوبية للغات الطبيعية ، بمعنى أخر ، التفاعل مع الحاسوب باللغة الطبيعية وإسقاط الحاجز اللغوي بين الحاسوب ومستخدميه. ويمكن ملاحظة نوعين من الدوافع وراء التطور السريع لهذا العلم هما :

- الدافع التكنولوجي المتمثل في تطور نظم الحاسوب بشكل مذهل في نظم عتادها وبرامجياتها وتطور شبكات الاتصال والإنترنت وثورة انتقال معلومات ، كان له الأثر الكبير في زيادة الاهتمام ببناء نظم حاسوبية ذكية لمعالجة اللغة الطبيعية مثل أنظمة الترجمة الآلية ، أنظمة فهم وتحليل النصوص ، أنظمة فهم وتحليل الكلام وغيرها من أنظمة التفاعل مع الحاسوب باللغة الطبيعية.

- والدافع المعرفي للحصول على تبصر لكيفية اتصال البشر مع بعضهم باستخدام لغاتهم الطبيعية. ولأن اللغة هي الركيزة الأساسية لوحدة العلوم ووحدة الفكر والقاسم المشترك لعلوم الإنسانيات لذا فأن دخول تطبيقات الحاسوب في مجال الإنسانيات لابد ان يستند أساساً الى إخضاع اللغة للسيطرة الآلية.

تعني المعالجة الحاسوبية للغة الطبيعية بشكل عام بناء نظم حاسوبية لتوليد وتحليل اللغة. وبسبب تعقد نظام اللغة الطبيعية فان بناء نظام شامل لها تعد عملية غاية الصعوبة والتعقيد وتتطلب تضافر جهود الكثير من التخصصات العلمية في الذكاء الاصطناعي ، اللسانيات الحاسوبية ، علم اللغة ، علم النفس ، علم الاجتماع ، علم المعرفة وعلم الصناعة المعجمية ولفترات زمنية طويلة.

يشترط عند تطوير أي نظام للمعالجة الحاسوبية للغة الطبيعية وجود معجم للغة يعمل كعمود فقري للنظام لتزويده بمعلومات عن المفردات اللغوية ومعانيها. حيث يستخدم المعجم أنواع مختلفة من العلاقات الدلالية ، السياقية ،التصنيفية وأمثلة الاستعمال وغيرها من العلاقات لوصف وتوضيح معنى المفردات اللغوية. ويعد المعجم من أعقد أجزاء نظام معالجة اللغة وذلك لضخامة مادته المعجمية وتعقد وتشابك علاقاتها. ونظراً لاعتماد المعجم على المعرفة الإنسانية في شرح مفرداته لذا فأنه يعد قاصراً عند استخدامه في نظام حاسوبي لمعالجة اللغة ، بذلك يجب أن يستند المعجم الى قاعدة معرفة تضم مفاهيم العالم الحقيقي التي تقابل المفردات اللغوية.

ولا يكتمل نظام المعالجة إلا بوجود معالجات إجرائية صرفية ، قواعدية ودلالية. الغرض منها تجميع معاني مكونات الجملة لإنتاج أو لتحليل معنى الجملة (الجمل) اللغوية. ولضمان كفاءة النظام وسهولة تطويره يجب أن يعتمد النظام معالجة اللغة الطبيعية على نموذج نظري (أو نظرية)

يصف خصائص اللغة ويستوعب معالجاتها.

واللغة العربية أعقد اللغات السامية وأغناها صوتاً وصرفاً ومعجماً وهي لغة القرآن الكريم والشعر والبلاغة والأبداع تفتقر الى جهود بحثية لأتمتتها تناسب وتكافئ أهميتها وموقعها بين اللغات الطبيعية . ولاشك ان وراء تأخر المعالجة العلمية الجادة للغة العربية هو تعقدها. ويرجع مصدر التعقيد الى عدد من المميزات منها مرونة اللغة وأتساع مفرداتها وتعدد معانيها وتباين استخدامها وصعوبة التعامل مع حالات اللبس والغموض والحذف وتفشي مجازها والمعاني الضمنية لجملها ونصوصها.

البحث الحالي هو محاولة في اتجاه لمعالجة اللغة العربية حاسوبياً، حيث يقدم أنموذجاً جديداً لهيكل تمثيلي لمعاني المفردات اللغوية العربية ، يتكون الأنموذج من معجم لمفردات اللغة العربية مستند الى قاعدة معرفة. تم بناء نموذج المعجم باعتماد محاكاة للطرق الحقيقية التي تمثل فيها المعاجم مادتها المعجمية وذلك باعتماد علاقات الترادف والتضاد (وهي العلاقات الأكثر شيوعاً في بناء المعاجم) واعتماد السياقات كمميز رئيسي لمعاني المفردات ، حيث تقابل هذه المفردات مفاهيم حقيقية لأحداث وكيانات العالم الحقيقي.

في حين اعتمد بناء هيكل قاعدة المعرفة على دراسة ثلاث نظريات دلالية هي :

1. نظرية الحقول الدلالية ، اهتمت هذه النظرية بشكل رئيسي بالعلاقات بين المفردات المعجمية وأكدت أن معاني المفردات محددة فقط بارتباطاتها و/أو اختلافاتها مع الكلمات الأخرى في المعجم اللغوي. اعتمد صاحب النظرية ترير (Trier) طريقة التحليل "أعلى-أسفل" لتحليل المعجم ، حيث وضع تصنيفاً عاماً للمفاهيم اللغوية وحاول تحديد انتماء كل مفردة لغوية الى مفهوم معين في هذا التصنيف. زودت هذه النظرية النظام المقترح بالهيكل التصنيفي العام لمفاهيم العالم الحقيقي.

2. نظرية قواعد الحالة ، زودت هذه النظرية بإطار الحالة (Case Frame) وهو قالب يرتبط من خلاله الفعل مع قائمة المعلمات التي يمكن أن ترتبط معه في الجملة ويستخدم هذا الإطار في تفسير الجملة. ومثلت قيود الدلالة القواعدية للأحداث الحقيقية .

3. ونظرية السمات الدلالية والقيود الانتقائية لتحديد السمات الرئيسية المميزة للمفاهيم

الحقيقية. استخدم كارز وفودور منهجاً معاكساً لمنهج ترير في تحليل المعجم إذ أعتمد التحليل "أسفل-أعلى" وطورا النظرية الدلالية التفسيرية (Interpretative Semantic Theory) لوصف هيكل المفردات المعجمية بدلالة مجموعة صغيرة نسبياً من عناصر المعنى العامة. وزودت هذه النظرية بطريقة لتحديد مواصفات المفاهيم والكيانات التي تنتمي اليها.

استخدمت هذه النظريات في تصميم وبناء هيكل تمثيلي تصنيفي تجريدي لمفاهيم العالم الحقيقي ، يتم الاستفادة منه لتعزيز وتحديد المعنى المعجمي . وتم ربط الهيكليين التمثليين للمعجم-قاعدة المعرفة ربطاً ثنائي الاتجاه ، حيث يحتوي مدخل المفردة اللغوية في المعجم اللغوي على مدخل تضم أرقام مفاهيم قاعدة المعرفة التي تترادف معها و مدخل أخر يضم أرقام المفاهيم المتضادة معها. في حين يحوي كل مفهوم في قاعدة المعرفة على مجموعة الترادف وهي مجموعة المفردات اللغوية المترادفة مع بعضها والتي تكون المفهوم.

للنموذج المقترح العديد من الميزات المهمة من بينها ، يفيد الهيكل المقترح في عمليتي توليد وفهم اللغة العربية ، ففي عملية التوليد ، وبأنتقاء مفاهيم معينة يمكن الوصول الى الصيغ السطحية مناسبة لأحتواء المفاهيم على مجاميع الترادف ، في حين تضم عملية الفهم تحليل الصيغة السطحية الى مجموعة المفاهيم المقابلة لها.

كما ان فصل المعجم اللغوي عن قاعدة المعرفة يؤدي الى فصل اللغة المستخدمة عن معلومات العالم الحقيقي وبالتالي يمكن إضافة أرتباطات لعدد من المعاجم للغات أخرى (وحسب استخدامها لمفرداتها اللغوية) بقاعدة المعرفة التي ستمثل معرفة عامة موضوعية عن العالم الحقيقي مستقلة عن أي لغة طبيعية.

والميزة المهمة الأخرى للهيكل المقترح هي التغلب على الخاصية التعددية التي هي من أصعب المشاكل التي تواجه أي نظام لمعالجة اللغة الطبيعية.

أن تجريد كلاً من الإطار الدلالي والسمات الدلالية من اللغة الى قاعدة المعرفة يقلل بشكل كبير من الفائضية ويسهل العمل في بناء المعاجم ويعكس سلوك أنساني فطري حيث يمكن ملاحظة أن الأنسان الذي لا يعرف القواعد اللغوية يمكنه ببساطة بناء جمل لغوية صحيحة قواعدياً.

ولإثبات كفاءة النموذج ، استخدم النموذج كجزء أساسي في نظام لفهم الجمل العربية البسيطة التركيب. هدف نظام الفهم هو هدف بسيط يتمثل باختبار صحة الجملة المدخلة دلالياً ، ولتحقيق هذا الهدف تم تطوير عدة معالجات صرفية ، قواعدية ودلالية ومعالجات لحل اللبس النحوي والمعجمي ومعالجات أخرى لإدارة وفهرسة المعجم. معظم المعالجة التي تمت على الجملة هي معالجة دلالية بالاستفادة من سياق الجملة وصولاً الى إثبات صحة الجملة المدخلة قواعديا ودلالياً.

تتكون عملية تحليل الجملة من عدة أطوار رئيسية هي على التوالي : طور البحث المعجمي ، طور التحليل الصرفي ، طور التحليل القواعدي ثم طور التحليل الدلالي ، تنفذ هذه الأطوار بشكل متتالٍ ، بحيث يضمن نجاح العمل في أحد الأطوار الانتقال الى الطور الذي يليه في حين يؤدي الفشل الى إنهاء عملية تحليل الجملة بأكملها.

ولان اللغة العربية تعاني من كثير من حالات اللبس المعجمي والنحوي فقد تم إضافة مكونين مهمين هما ، مؤشر الصنف القواعدي (Part of Speech Tagger) لحل اللبس القواعدي ، ومؤشر المعنى (Sense Tagger) لحل اللبس المعجمي.

استخدمت لغة البرمجة الموجهة الهدف C++ كأداة حديثة وكفوءة لبناء النظام المقترح حيث تم أستغلال أمكانياتها البرمجية من حيث سهولة بناء الفئات وأمكانية بناء الهياكل التدريجية المعقدة وأمكانيات التوارث والتوارث المتعددة.