---

٤٣٦ علم اللغة الحاسوبي

أستاذ المقرر: د. صلاح الناجم

البريد الإلكتروني:

alnajem@alnajem.com

الموقع على الإنترنت:

www.alnajem.com

تعريف علم اللغة الحاسوبي:

عِلْم اللغة الحاسوبي Computational Linguistics ومعالجة اللغةِ الطبيعيةِ Natural Language Processing يعرفان بأنهما دراسة الجوانب الحاسوبية للغة والمشاكل الشائعة التي تواجه المعالجة الحاسوبية للغة المكتوبة والمنطوقة. يعرف علم اللغة الحاسوبي أيضا بأنه علم دراسة أنظمة الحاسوب لغرض فهم وتَوليد اللغةِ الطّبيعيةِ

تطبيقات عِلْمِ اللغة الحاسوبي:

هنالك ثلاثة مِنْ أصنافِ من التّطبيقاتِ تمثل ركيزة في تطورِ عِلْمِ اللغة الحاسوبي:

التّرجمة الإلكترونية:

بدأ العمل في التّرجمة الإلكترونيةِ في أواخر الخمسيناتِ بآمال كبيرةِ وإدراك قليل للصعوبات التي تتضمنها هذه العملية. المشاكل في التّرجمة الإلكترونيةِ حفّزتْ العمل في كل من عِلْم اللغة وعِلْم اللغة الحاسوبي. العمل الأكثر كان في أوائل السّتّينات، لكن مع قليل من النّجاحِ، ومع إدراك بأنَّ تحقيق ترجمة آلية بشكل كامل بجودة عالية لا يمكن أن يتم دون عملِ مكثف من أجل 'فَهْم' النَّصِّ. أدّى هذا إلى تقليص في التّمويل. عدد قليل فقط من المشاريعِ الحاليةِ في عِلْمِ اللغة الحاسوبي في الولايات المتّحدةِ موجهة نحو التّرجمة الإلكترونيةِ، بالرغم من أن هناك مشاريع كبيرة في أوروبا واليابان.

استرجاع المعلومات:

كثير مِنْ المعلوماتِ التي نستعملها تظهر في شكلِ لغة طبيعية (كُتُب، مجلات، تقارير). أحد التطبيقات الهامة ذات العلاقة بعلم اللغة الحاسوبي هو الاسترجاع الآلي للمعلوماتَ مِنْ نصوصِ اللغةِ الطّبيعيةِ. استجابة إلى استفسار يدخله المستخدم. يَنتزعَ النظام النَّصَّ ذا العلاقةَ مِنْ مجموعةِ من

النصوص ويقوم بعرض النَّصَّ أَو يَستعملُ النَّصَّ للإجابة عن الاستفسار مباشرة. لأن النّصوصَ في أكثر مجالاتِ الاهتمام يكون مركبا، لم يكن هنالك نجاحُ فوريُ إلى حدٍّ ما في هذه المنطقة. استرجاع المعلومات الآلي الآن يلقى اهتماما من بضعة مجموعات للأبحاث.

الأنظمة التفاعلية:

تَبْدو اللغةُ الطّبيعيةُ النّمطُ الأكثر سهولةً للاتصال بالأنظمةِ التّفاعليةِ (مثل قواعد البيانات وتطبيقاتِ الحاسوب المرتبطة باللغة الإنسانية)، خصوصاً لغير المتخصصين في الحاسوبِ. لعلم اللغة الحاسوبي عِدّة فوائد على مستوى التفاعل بين الإنسان والحاسوب. أولاً، يمكن هذا العلم المستخدمين من إدخال مدخلاتهم inputs إلى أنظمة الحاسوب بشكل أكثر سهولة نحويا ودلاليا دون الحاجة إلى استخدام لغة مصطنعة خاصة. ثانيا، الطّبيعة التّفاعلية لأنظمة الحاسوب التي تسمح بالتحاور مع الإنسان تَسْمحُ للنّظامِ لِكي تَكُونَ قابلا للاستخدام حتى لو كانت المدخلات مرفوضة لسبب أو لآخر. فالنظام المبني على تطبيقات علم اللغة الحاسوبي يمكنه القيام بالتصحيح الآلي للمدخلات أو طلب تصحيح المدخلات الخاطئة من المستخدم في حال عدم قدرته على القيام بالتصحيح التلقائي. حيث يمكن لهذا النظام أن يقول للمستخدم "أرجو إعادة صياغة هذه الجملة". تم تحقيق نجاح ملحوظ في هذا المجال على مستوى اللغة الإنجليزية، حيث بدأ استخدام هذه الأنظمة كأنظمة تجارية فعلية وليست أنظمة تجريبية. أكثر إنجازات عِلْمِ اللغة الحاسوبي منذ أوائِل السّبعيناتِ كانت مرتبطة بالأنظمة التفاعلية التي تسمح بتفاعل الإنسان والحاسوب بلغة طبيعية.

بالأضافة إلى هذه الدوافع الهندسية المهتمة بالتطبيقات العملية والتي دفعت إلى البحث في عِلْمِ اللغة الحاسوبي، كان لأكثر الباحثين أهدافاً بحثية علمية مستقلة عن أي أهداف تطبيقية معيّنة. أحد الوظائف الأساسية لعلم اللغة الحاسوبي هو اختبار القواعدِ النحوية المقترحة مِن قِبل اللغويين النّظريين. بسبب تعقيد هذه القواعد (كالقواعد التحويلية)، صار من الهام استخدام الحاسوبِ للتحقق من مدى عمل مجموعةِ مُقتَرَحةِ مِنْ القواعدِ بشكل صحيح أم لا. في هذا السياق تم فحص قواعدِ فريدمان التّحويلية باستخدام نظام قام بتوليد الجُمَلَ بموجب قواعد تحويلية مُقتَرَحة، لكي يتمكن اللغويون من التأكد من أَنْ تلك القواعد قامت بتوليد كل الجمل الصحيحة نحويا. على أية حال، لم يكن اهتمام عِلْمِ اللغة النّظريِ منصبا على بناء القواعدِ التي يمكن اختبارها حاسوبيا، وهذا هو السبب في الاستعمال المحدود للحاسوب كأداة إختبارِ للنّظرياتِ اللغويةِ.

من الناحية الأخرى، الحاجة لتَطوير الأنظمةِ التفاعلية الكاملةِ أجبرتْ لغويين حاسوبيين على تَطوير مجالات بحث كَانَت قَدْ استُكشفَت بشكل ناقص من قِبَل العُلومِ التّقليديةِ. من هذه المجالات: النّماذج الإجرائية للعملياتِ النّفسيةِ لفَهْم اللغةِ. حاول بعض اللغويون الحاسوبيون النَّظْر إلى عملية الفهم الآلي للغة بشكل عام. مع استغلال هذه العملياتِ النفسية لجعل الحاسوب قادرا على محاكاة بعْض سماتِ الأداء البشري. هذه الجُهودِ، إضافة إلى جهود العلماء النّفسانيين والباحثين الآخرينِ، أدت إلى خَلْقِ حقلِ ثانويِ جديدِ وهو علم الإدراك.

عادة ما يتم أخذ الأهداف الهندسية والأهداف العلمية بشكل متواز. الحاجة إلى الأنظمةِ العمليةِ قَدْ تُؤدّي إلى البحث في فَهْم العملياتِ اللغويةِ. وهذا يؤدي بالتالي إلى إنتاج أنظمةَ أفضل لمعالجة اللغةِ الطّبيعيةِ. في بَعْض المجالات، مثل التّحليلِ النّحوي، يُمْكِنُ أَنْ يتم التفريق بين الأنظمةِ المصممة للتجارب الخاصة بعلم النفس وتلك الخاصة بتطبيقات عملية معينة. في المجالات الأخرى، التي لم تدرس بشكل مركز، لا يوجد إلى الآن تفريق واضح بين المناهج النفسية والمناهج الخاصة بتطبيقات معينة.

كيف يتم بناء تطبيقات علم اللغة الحاسوبي وأنظمة معالجة اللغة الطبيعية

بناء نظام لمعالجة اللغة الطبيعية يعتبر مهمة صعبة ومعقدة. في الوقت الحالي، على أية حال، يتمثل التحدي الرئيس الذي يواجه تصميم هذه الأنظمة في دمج العديد مِنْ الأنواعِ المختلفةِ للمعرفةِ – المعرفةالنحوية، المعرفة الدلالية، معرفة الكلام (النص) – واستخدامها عملياً في معالجة اللغةِ. في هذا المجال. كَمَا هو الحَال عند بناء أي نظام حاسوبي، هنالك بَعْض التّقنياتِ العامّةِ التي يُمْكِنُ أَنْ نَستعملَها لتَسهيل عملنا. إحدى هذه التقنيات هو النمذجة modularity وهي تَقسيم نظامنا إلى مكونات مستقلة نسبياً. تَقسيم المشكلةِ يَسْمحُ لنا بمهاجمة المشاكل الفرعية الأصغر -sub problems بشكل مستقل ، وبذلك لا تغمرنا المهمة التي أمامنا تماما. إذا ما تم تصميم المكونات بعناية، فسوف نجد أن ذلك التقسيم يقلل ليس فقط من حجم الأقسام المفردة بل من حجم النظام الحاسوبي بأكمله.

التّقنية الأخرى المستخدمة لتَبْسيط الأنظمةِ المعقّدةِ هي استخدام النماذج الصورية formal models. البرامج الكبيرة صعبة التصميم، والتعديل ، والفهم.قدرتنا على تصميم نظام ناجح تزداد

إذا استطعنا تصميم نموذج مُجرّد بسيط نسبياً ومن ثم تطوير نظامنا لكي يكون تطبيقا لذلك النّموذجِ. استعمال نموذجِ بسيطِ سَيَزِيدُ فرص فهم عملنا من قِبَل الباحثين الآخرين وبالتالي يسمح لنا بالمساهمة الإيجابية في تطوير الحقل العلمي الذي نعمل فيه.

أجزاء أنظمة معالجة اللغة الطبيعية

التّحليل والتوليد

تَتطلّبُ أكثر أنظمةِ معالجة اللغةِ الطّبيعيةِ توفر القدرة على تحليل وتوليد اللغة الطبيعية. يعتبر التحليل المكوّنَ الأكثرَ حسماً في الأنظمة العملية التطبيقيةِ. لكي يكون نظام اللغة الطبيعية ناجحا، يَجِبُ أَنْ يَكُونَ قادرا على معْرِفة وتحليل الكثيرِ من العبارات والصيغ التي تعبر عن نفس الأمر أو المعلومة. من ناحية أخرى، يكفي هذا النظام أن يكوم قادرا على توليد صيغة واحدة من هذه الصيغ.

تحليل الجمل وتحليل النص

بشكل عام، يُنظَر إلى تحليل الجمل على أنه معالجة لجمل فردية وليس تحديدا لبنية الكلام (النص) ككل. أي تحليل لبنية الكلام يتطلب تحليلا لمعنى الجمل المفردة في الكلام. في كثير من التطبيقات البسيطة لا يكون التحليل الكامل للكلام (النص) ضروريا حيث يمكن أن تفهم الجملة حتى ولو كانت الجملة تشير إلى جمل مذكورة سابقا في سياق الكلام. التحليل الكامل لسياق الكلام يعتبر من أصعب المهام التي تواجه علم اللغة الحاسوبي.

التحليل النحوي و التحليل الدلالي:

الهدف العام من تحليل الجمل هو تحديد معنى الجملة. يتطلب هذا الأمر ترجمة الجمل المدخلة إلى الحاسوب بلغة طبيعية إلى جمل تحمل دلالة بسيطة باستخدام ما يعرف بالمنطق الرمزي Formal Semantics. كذلك يمكن ترجمة الجمل المدخلة إلى لغة مصطنعة خاصة يمكن أن يفهمها الحاسوب (مثل لغة SQL الخاصة بأنظمة إدارة قواعد البيانات). في أكثر الأنظمة، أول خطوة لهذه الترجمة الخاصة بالدلالة هي التحليل النحوي. التحليل النحوي هو عملية تحديد بنية الجملة. وهذه الخطوة هي أول خطوة قام بتنفيذها علماء اللغة الحاسوبيون.

j