

إشكاليّات اللغة العربية في عصر الذكاء الاصطناعيّ

حبيب عبدالرب سروري

(محاضرتي الافتتاحية لمهرجان الأدب والثقافة العربية،

في الجامعة الكاثوليكية بميلانو. ١١ إبريل ٢٠٢٤)

«سلايد» المحاضرة¹

ملخص:

بعد تعريفٍ مقتضبٍ لمفهوم الذكاء الاصطناعي، ولتاريخه المرتبطٍ جينياً بتاريخ الكمبيوتر منذ ١٩٥٠، استعرضُ سريعاً المجالات التي يتفوقُ فيها الذكاء الاصطناعي عن الذكاء الإنساني، أو تلك التي ما زال الأول فيها بعيداً جداً عن مستوى الثاني وقدراته.

أوضّحُ في نهاية هذا المدخل أن الذكاء ليس مجالاً واحداً نمطياً خطياً، والتفاعل بين الذكاء الاصطناعي والذكاء الإنساني زجاجيٌّ يتطوّر على الدوام.

أناقش في ضوء ذلك سؤال المؤتمر: الذكاء الاصطناعيّ «قيّد للماضي أم أفق للمستقبل؟» موضّحاً أنّ عبارة «قيّد للماضي» لا محلّ لها من الإعراب في تقديري، مع ضرورة مقاومة تطبيقات الذكاء الاصطناعي المؤذية السامة كالتجسس الآلي؛ وضرورة جعل مدونات اللغة Corpus، ونماذجها LLM، شفافةً ومجانيةً، في تناول الجميع (في إطار «العلوم المفتوحة»)، لتلافي الكثير من التداعيات الضارة والمخاطر الافتراضية الممكنة جرّاء تعليم الكمبيوتر لغات الإنسان واستخدامها آلياً بدلاً منه.

انتقلُ بعد ذلك إلى الوضع الحالي لتقنيّتي الترجمة الآلية والتوليد الآلي للنصوص، منذ طفرة تقنية «التعلّم العميق بشبكات العصبونات الاصطناعية».

¹ <https://habibabdulrab.com/wp-content/uploads/2024/04/ConfMilan8.pdf>

أستعرضُ وضعَ اللغة العربية في هذين المجالين مُسلِّطاً الضوءَ على بعض النواقص والعقبات الموروثة منذ عدم اكتمال بناء الأسس التحتيّة لدخول العربية العصر الرقمي، وتداعيات ذلك اليوم في عصر الذكاء الاصطناعي.

أتناول في الأخير بعض المشاريع التي ينبغي رفعها سريعاً، والاتجاهات الضرورية لحضورٍ مرموقٍ وجبّارٍ للعربية يليق بتاريخها وإمكانياتها الهائلة.

فهرس

- ١) تعريف الذكاء الاصطناعي
- ٢) تاريخ الذكاء الاصطناعي بكلمتين
- ٣) مجالات تفوق الذكاء الاصطناعي وتأخره على الذكاء الإنسانيّ حالياً
- ٤) الذكاء الاصطناعي «قيّد للماضي أم أفق للمستقبل»؟
- ٥) النواقص والعقبات الموروثة منذ دخول العربية العصر الرقمي
- ٦) التوليد الآلي للنصوص، ووضع العربية
- ٧) الترجمة الآلية، ووضع العربية
- ٨) آفاق ومقترحات لحضور لائق ومرموق للعربية

١) تعريف الذكاء الاصطناعي

الذكاء الاصطناعي مجالٌ في علوم الكمبيوتر هدّفه محاكاةُ الكمبيوتر للذكاء الإنساني (في التعلّم، الاستيعاب، الاستنباط، البرهنة، حلّ الإشكاليات، واتخاذ القرارات...). تطبيقاته العملية تشملُ مختلف مجالات الحياة: الصحة، الإدارة، المواصلات، السياسة، التأليف، التعليم، أنظمة المعلومات...

٢) تاريخ الذكاء الاصطناعي بكلمتين

بعض أهمّ منعطفات هذا التاريخ:
← اختراع آلان تورنج، في عام ١٩٣٦، لماكينة نظريّة (تجربة ذهنية) يمكن استخدامها للحلّ الآلي لأي خوارزمية.

- ← التصميم العملي للكمبيوتر، في عام ١٩٤٥. اشترك في صنع مداميكه التأسيسية: آلان تورنج، فان نيومان...
- ← آلان تورنج هو نفسه مؤسس علم الذكاء الاصطناعي. راهن على امتلاك الكمبيوتر لذكاء الإنسان في نهاية القرن العشرين (ووضع معيارا مثيرا لذلك، يسمّى: «امتحان تورنج»، وإن لا يخلو من بعض المآخذ).
- ← مؤتمر تأسيسي لِعِلْم الذكاء الاصطناعي في ١٩٥٦، حضره المؤسس مكارثي، بجانب مينسكي، شانون، وعددٌ ممّن صنعوا اللبنة الأولى لهذا العلم، الآتين جميعهم في الغالب من علوم الرياضيات أو الفيزياء...
- ← بدأت بعد ذلك عقود اللغات الذكية مثل LISP، ثمّ PROLOG، وبرامج «المنظومات الخيرة» Expert System.
- ← في نهاية الثمانينات، خيم ما يسمّى شتاء الذكاء الاصطناعي، بسبب البعد عن تحقيق أهداف العلم بمحاكاة ذكاء الإنسان، مع اقتراب نهاية القرن!
- ← دخول الإنترنت إلى حياة الناس، وازدياد جبروت وسرعة الكمبيوتر في التسعينات أعاد كثيرا من الأمل، وسمح لتقنيات نظرية قديمة نسبيّة في الذكاء الاصطناعي، وفي علوم البرمجة، أن تمتلك أخيرا تطبيقات كمبيوترية ممكنة فعالة وسريعة، مثل تقنية «التعلّم العميق» بـ «شبكة العصبونات الاصطناعية».
- ← في عام ٢٠١٢، برهنت هذه التقنية تفوّقها الكاسح، أثناء المباريات التنافسية بين برمجيات المختبرات العلمية للتعرف على الصور، في أحد المؤتمرات العلمية. بدأت بعد ذلك، وبفضل ذلك غالبا، الانتصارات المتتالية للذكاء الاصطناعي في مجالات عديدة مختلفة من حياة الإنسان.

(٣) مجالات تفوّق الذكاء الاصطناعي وتأخره على الذكاء الإنساني حاليا

يتفوّق الذكاء الاصطناعي على الإنسان اليوم في مجالات عدّة من «الذكاء الموجه» الذي تمّ تدريبه عليها، أهمّها: مقدرته على الترجمة الآلية من ٢٠٠ لغة إلى ٢٠٠ لغة؛ تشخيص الأمراض وطرائق علاجها انطلاقا من صور الكشافات الطبيّة؛ هزيمة بطل العالم جي سيدول، في ٢٠١٦، في أصعب لعبة عقلية: ألغو، بالذكاء المطلّق هذه المرّة (وليس على غرار هزيمة الكمبيوتر لكازباروف بالشطرنج في ١٩٩٧، بفضل تطوّر سرعة وذاكرة الكمبيوتر، في الجوهر). لذلك يشكّل هذا الحدث بدايةً لمرحلة جديدة من حياة البشرية، بدأ بعدها العدّ التنازلي لحضورٍ أساسيٍّ متصاعدٍ وشامل للذكاء الاصطناعي في حياة البشر.

لكن الذكاء الاصطناعي، بطبيعة الحال، ما زال متأخراً عن ذكاء الإنسان في مجالات جوهرية كثيرة من «الذكاء غير الموجه» كتعلم تخطيط وأداء المهام اليومية: الطباخة، السفر... وغيرها مما يمكن أحيانا لطفلٍ صغير أن يتعلمه خلال ساعات.

عموماً، الذكاء ليس مجالاً نمطياً خطياً. والتفاعل بين الذكاء الاصطناعي والإنساني ديكالكتيكيٌّ، زجاجيٌّ (علاقتهما ليست مانوية).

٤) سؤال المؤتمر: الذكاء الاصطناعي «قيّد للماضي أم أفق للمستقبل»؟

الذكاء الاصطناعي أفق علم الكمبيوتر منذ ولادته، لا يمكن فصله عن الكمبيوتر. لذلك الإجابة على سؤال: «هل الكمبيوتر قيّد للماضي؟» هي: بالطبع، لا! أي أن عبارة «قيّد للماضي» لا محلّ لها من الإعراب. بيد أنه علم ذو حدّين، يمكنه أن يقود إلى مخاطر (على غرار معادلة آينشتاين لتحويل المادة إلى طاقة التي قادت إلى القنبلة الذرية) إذا استخدم بطريقة ضارة.

لذلك يلزم مقاومة تطبيقات الذكاء الاصطناعي المؤذية السامة كالتجسس الآلي والروبوتات القاتلة؛ وما يهّمنا هنا على نحو خاص: ضرورة جعل مدونات اللغة Corpus، ونماذجها LLM (التي يستند عليهما تعليم الآلة للغات الإنسان، كما سنوضح ذلك لاحقاً) شفافةً ومجانيةً، في تناول الجميع (في إطار «العلوم المفتوحة»)، لتلافي الكثير من التدايعات الضارة والمخاطر الافتراضية الممكنة جزاءً تعليم الكمبيوتر لغات الإنسان واستخدامها آلياً بدلاً منه.

٥) أهم النواقص الموروثة اليوم منذ دخول العربية العصر الرقمي

في مقال لي في ٢٠٠٧، بعنوان: «اللغة العربية في الزمن الرقمي: ٦ فجائع و٣ مقترحات»²، تناولتُ بعض النواقص التي يلزم استكمالها سريعاً لإكمال البناء التحتي لدخول اللغة العربية العصر الرقمي. ما زال بعضها غائباً أو ناقصاً حتى الآن!

² https://my-last-articles-and-texts.blogspot.com/2009/03/blog-post_23.html

Corpus

مدوّنة أية لغة، (Corpus)، هي مجموعة هائلة (كانت تعدُّ كلماتها بالمليارات عند كتابة المقال، بآلاف المليارات اليوم) من عينات النصوص المكتوبة أو المنطوقة، الآتية من قطاع متنوّع عريض محايد من المصادر:

الكتب القديمة والحديثة، ويكيبيديا، نصوص المؤسسات الدولية، وما يرفده الإنترنت يوميا من صحف ومجلات، نقاشات، تقارير، مواقع إنترنت... تتلقّفها وتؤرشفها مشاريع دولية مثل paraCrawl.eu للغات الأوربية.

تعطي جميعها صورةً دقيقةً كاملةً عن اللغة في مختلف أشكالها واستعمالاتها اليومية والعلمية والعملية والأدبية، خلال مرحلةٍ زمنيةٍ معيّنة!...

تمتلك اللغات اليوم مدوّناتها، المسماة أحيانا «بنوك اللغة». ثمة بوابات على الإنترنت تسمح بالوصول لـ «قواعدها البيانية» الضخمة والبحث المحدّد في طياتها، أو معالجتها آليا بشكلٍ إجمالي!

من كنوزها (التي يتمُّ رفدها كل يوم) يستطيع الباحثون دراسة تطوّر المجتمع في مختلف المجالات. منها تُستخلصُ القواميس والمعاجم المتخصصة في المجالات اللغوية والعلمية والتقنية والعملية.

هي المختبر الذي تخرج منه الدراسات اللغوية المتنوّعة لبنيّة اللغة وظواهرها وشتى دلالات كلماتها، لنواقصها واحتياجاتها المتجدّدة، لتاريخ وأصول الكلمات وعلاقتها باللغات الأخرى... وتتكيّ عليه وسائل البحث المتقدّمة الحديثة مثل Ngram لغوغل (أنظر مقالي: نغرام غوغل ضرورة ثقافية غائبة عربيا)³، الموجود بلغات كثيرة بما فيها العبرية، والغائب كليّةً بالعربية، والذي يستخدمه الباحثون كثيرا في عددٍ من المجالات الاجتماعية والتاريخية واللغوية...

وفي عصر الذكاء الاصطناعي صارت مدوّنة اللغة (Corpus) المادة الخام الرئيسة التي يتدرّب عليها نموذج اللغة (LLM). يتعلّم منها توليد النصوص آليا للإجابة

³ <https://habibabdulrab.com/%d9%85%d9%82%d8%a7%d9%84%d8%a7%d8%aa-articles/>

على أي سؤال، كما هو حال تشات جي بي تي مثلاً. هي أيضا مرجعيته للترجمة الآلية.

المفارقة المثيرة والمؤلمة أن اللغة العربية التي كانت أول من أسس القواميس والمعاجم اللغوية (منذ الخليل بن أحمد الفراهيدي صاحب قاموس العين، وربما الأصمعي قبل ذلك)، والتي قامت في عصرها الذهبي بدورٍ طليعيٍّ في تأسيس دراسات النحو والصرف العبرية، وتصنيف المفردات وترتيب جذورها واشتقاقاتها، وتأليف كل المعاجم (بما فيها معاجم الجنّ والشياطين!)، والتي انفتحت بشكلٍ مبكّرٍ على لغات العالم منذ العصر العباسي وحملة ترجماته الزاهرة، تفتقد إلى مدونةٍ تليق بها اليوم!

القارئ الضوئي OCR

مثالٌ آخر: القارئ الضوئي للأحرف OCR الذي يسمح بتحويل الكتب الورقية (لا سيّما التي سبقت عصر الإنترنت) إلى نصوص رقمية. ما زال أثره غائباً عربياً، في حين كانت اللغات الكبرى، عند نشر مقالي، قد استكملت رقمته ملايين الكتب القديمة.

سمح ذلك حينها مثلاً بتحويل كتاب «الكوميديا الإلهية» الورقيّ لدانتي إلى نصّ رقميٍّ، في حين لا يمكن العثور على «رسالة الغفران» للمعريّ إلى اليوم إلا كصورة PDF.

التأخر في إنجاز رقمنة الكتب القديمة بالعربية معضلةٌ قوميةٌ، قادت إلى ضعف موارد ذخائرها الرقمية، ومدوّنتها Corpus التي يستند تعلّم الكمبيوتر للغات عليها، وعلى ثراء تموينها المتنوع الدائم، كما أشرنا سابقاً.

انسحاب اللغة العربية من كتابة التكنولوجيا والعلوم:

أضحت المواد العلمية تُدرّسُ باللغات الأجنبية في كل المدارس الخاصة في العالم العربي، وفي كثير من المدارس الحكومية أيضاً. ناهيك عن غياب العربية شبه الكليّ عن تدريس المواد العلمية والتقنية والطبية في جميع الجامعات العربية تقريباً، بسبب عدم استخدامها لكتابة المعارف الحديثة!... ربما لذلك يُقال اليوم أكثر فأكثر إنّها «لغةٌ لا تصلح للحداثة، فقيرة في المصطلحات»!

المفارقة الكبرى: كانت العربية، قبل قرون، لغة العلوم بامتياز. قبل الخوارزمي مثلا، كانت الرياضيات في الجوهر هندسةً أقلديسية في الأساس، ورسوماتٍ لأشكال هندسية تُستخدم لحلّ هذه الإشكالية العملية أو تلك، لا غير. أدخل جلالته عليها اللغة الإنسانية كما لم يفعل أحد قبله، وحولها إلى لغةٍ ومنهجٍ يسمحان بتنظير تجريدي كليّ، بعد اختراعه لِعِلْم الجبر.

شرح مثلا نظرية معادلات الدرجة الأولى والثانية في الرياضيات بلغةٍ دقيقة. اخترع فكرة المتغيّر الرياضي x ، وأسماه «شيء»، قبل أن تصل هذه الكلمة العربية بدورها إلى إسبانيا وتلفظ في لغتها القديمة: "إكسي"، ثم تغزو أوروبا بعد ذلك بصيغتها النهائية: «إكس».

استخدم أيضا مصطلحات أخرى جديدة، مثل «الجذر» (الرقم الذي يحلّ المعادلة)، «الدرهم» (الرقم الثابت في المعادلة)...

ما المعادلات الرياضية، وما علم المنطق الرياضي، بل وما كلُّ العلوم، لو كانت بدون استخدام المتغيرات والمجاهيل ودمج اللغة الإنسانية في ثناياها؟

بيد أن لغة الخوارزمي تفتقر اليوم لِمِرادفٍ لكل جديدٍ في القاموس العلمي والتكنولوجي، وذلك منذ دهر. ويعلم الله أن هذا الجديد يهطل بغزارة من كل حدبٍ وصوب هذه الأيام، لا سيما في علوم التكنولوجيا الحديثة.

تفتقر لغة الضاد أيضا إلى صيغٍ تعبيرية موحّدة، تترجمُ بعض الصيغ التقليدية المتعارف عليها في سياقات الكتابة العلمية باللغات الأجنبية...

المفارقة الأهم: تمتلك العربية إمكانيات مرنة ومتميّزة لاحتضان جديد العلم وابتكار أجمل الكلمات الجديدة مثل: «رقمنة» التي تترجمُ (numérisation (digitization) انطلاقا من صيغة «فعلنة» التي تعكس التوليد والتحوّل: عقلنة، رهبنة، تيسنة...)،

أو «استرقام» التي يمكنها أن تترجمَ (demetarialization) dématérialisation (انطلاقاً من صيغة «استفعل» التي تعني طلبَ الشيء واستدعاءه: استسقاء، استحضار...).

ولأن النصوص التوليدية والترجمة الآلية لا يمكنهما اختراع كلماتٍ وصيغٍ جملٍ من العدم، ففقرُ مدوِّنة اللغة العربية عموماً، وعلى الصعيد العلمي والتكنولوجي خصوصاً، معضلةٌ جوهرية تمنع دخول اللغة العربية عصر الذكاء الاصطناعي من أوسع أبوابه.

٦) التوليد الآلي للنصوص، ووضع العربية

توليد النصوص الآلي يتكئ على راسيتين: مدوِّنة اللغة التي ينبغي أن تكون ثريةً متنوّعة، ونموذج اللغة LLM الذي يتعلّم من المدوِّنة، عبر تقنية «التعلّم العميق بشبكات العصبونات الاصطناعية».

يسمح ذلك لمولّد النصوص أن يجد أكثر الكلمات احتمالاً لاستطراد نصٍّ ما أو مجموعة كلمات، ثمّ أكثر الكلمات احتمالاً لمواصلة النص بعدها، وهكذا دواليك. مثال لما يمكن كتابته له، وردّه على ذلك:
ميلانو إحدى أروع...
← ... ميلانو إحدى أروع المدن...

يبدو جلياً عند توجيه الأسئلة المرتبطة بالثقافة العربية وجود فجوة هائلة بسبب فقر المدوِّنة العربية في بعض أنظمة التوليد الآلي كحال تشات جي بي تي، أو عدم اكتمال تعليم نموذج اللغة عليها كحال أربيك جي بي تي (جيس).

كمثال، عند طلب تشات جي بي تي كتابة نصٍّ بأسلوبٍ دانتي، أو عند توجيه أي سؤال عن شخصيات وأحداث وهوامش «الكوميديا الإلهية» نرى الردّ دقيقاً وكافياً.

وغالباً ما تكون ردود تشات جي بي تي، في كل شيء ولا شيء، مدهشة. ليس الحال كذلك عند التعرض لقضايا ثقافية عربية، حيث مدوِّنة اللغة «ضعيفة التموين» mal doté، بسبب النواقص التي استعرضناها أعلاه.

كمثال: عند استفساره عن رأي المعري بالمتنبي أو عن بعض تفاصيل كتابه رسالة الغفران، نجد ردودا لا يُغفرُ لها مثل: «من الصعب تحديد الآراء الدقيقة للمعري بشأن المتنبي بدقة، نظرا لنقص المعلومات».

عند سؤاله عن الجنة في رسالة الغفران أو عن حوار ابن القارح وآدم فيها، يأتي الرد صادما تماما: «لم يذكر المعري تفاصيل محدّدة عن الجنة»، «لا توجد إشارة إلى حوارات في رسالة الغفران مع شخصيات تاريخية كآدم!»

يختلف الأمر كما يبدو مع أربيك جي بي تي (جيس) الذي تمّ رفضه بنصوص ثقافية عربية عن المعري ورسالة الغفران. ردّه جيد حول رأي المعري بالمتنبي. لكن نموذج اللغة الخاص به ما زال غير مكتمل التعليم، كما يبدو.

لسؤالٍ تضليليٍّ مثل: «كيف وجد ابن القارح جمال عبد الناصر في جنة رسالة الغفران؟» يردّ على نحوٍ يستقيم له شعر الرأس: «في رسالة الغفران ينطلق ابن القارح في رحلة رائعة إلى الآخرة يواجه فيها شخصيات أدبية مختلفة، وخلال رحلته يلاقي جمال عبد الناصر أحد الشخصيات البارزة التي كان لها تأثير على العالم العربي... يعكس إدراج جمال عبد الناصر في النص تقدير المعري للشخصيات السياسية الهامة وتأييده لها...!»

ملاحظة: عدم معرفة محتوى ومعطيات المدونات اللغوية يمنع الباحثين من دراستها علميًا، وتحليل جذور الخطأ في عملها.

(٧) الترجمة الآلية اليوم، ووضع العربية

الترجمة، من لغة ل ١ إلى لغة ل ٢، كانت حتىّ أمد قريب تنطلق من القواعد النحويّة للغتين:

أي من القاعدة النحوية للجملة ج ١ التي نريد ترجمتها في ل ١، إلى القاعدة النحوية المقابلة لها في ل ٢،

ثمّ بثُّ ترجمة مفردات الجملة ج ١ في مواضعها في قاعدة ل ٢، للحصول على ج ٢.

منذ ٢٠١٧، اكتسحت تقنية المحوّل الآلي (Transformer : Encoder, Decoder) كل تقنيات الترجمة الآلية، وفرضت نفسها على الجميع. تستند التقنية على استخدام شبكة العصبونات الاصطناعية والتعلّم العميق، وعلى بناء نموذج اللغة الذي يتمّ تدريبه طويلا على مدوّنة اللغة، وعلى تقنية الإصغاء Attention للجملة، لأخذ السياق بعين الاعتبار عند الترجمة.

النتيجة:

في اللغات ذات المدوّنات الغنية والنماذج المدرّبة: النتائج مذهلة، تفوق توقّعات معايير التقييم التقليدية.

ترجمة ١٠٠ لغة إلى ١٠٠ لغة، كما هو حال غوغل ترانسليت، لا يحتاج إلى ١٠٠ * ٩٩ منظومة ترجمة، لكن إلى ٩٩ فقط: يكفي استخدام لغةٍ وسيطة واحدة! كذلك حال مشروع فيسبوك nllb-200 لترجمة ٢٠٠ لغة إلى ٢٠٠ لغة.

في العربية: النتائج ضعيفة نسبيا لمن يقارنها باللغات ذات التموين العالي، وإنّ يمكن ملاحظة تطوّر سنويّ. أمثلة:

(١) يمكن تمييز السياق بفضل تقنية الإصغاء، كما في هذه الترجمة: «فصل الربيع الجميل، فصل من الكتاب المبين».

The beautiful spring **season**, a **chapter** from the clear book.

(٢) ترجمَ غوغل «جوزي قال لي» قبل ٢٠٢٢، كما لو كان الحديث عن ثمرة الجوز:
Ma muscade m'a dit

في ٢٠٢٢، ترجمها معتبرا أن «جوزي» اسم إنسان.

Josie m'a dit

(٣) واليوم يترجمها كما يلزم باعتبار جوز مرادف كلمة زوج:

Mon mari m'a dit

سؤال: ماذا عن ترجمة: «فتنة الفتنة»، عندما ترتبط الكلمة الثانية بالجمال والإغواء، والأولى بما يقع به الناس من شركٍ إثر ذلك.

٨ آفاق ومقترحات لحضورٍ لائقٍ ومرموقٍ للعربية

إذا نريد الحفاظ على ثقافتنا العربية في عصر الذكاء الاصطناعي، واللاحق بالعصر، فيلزمنا سريعاً:

← ردم الهوة الشاسعة في مجال ترجمة العلوم وكتابتها، عبر تطوير وإثراء القاموس العلمي العربي، واستخدام اللغة العربية في كتابة وتدريس العلوم والمعارف عامة في المدارس والجامعات، وتطوير مشاريع قومية لبوابات علمية نموذجية على الإنترنت في مختلف الدروس العلمية والتكنولوجية.

← بناء LLM مدربة على مدونة ثرية تضم كل ما كتب في الثقافة والأدب بالعربية (وباللغات الأخرى لبلداننا العربية، كالأمازيغية والسريانية والكرديّة) وما يضمن أيضاً حفظ اللغات القديمة كالسقطرية والشحرية والمهرية وغيرهم. ذلك يعني ضرورة إدخال الكتب القديمة والحديثة، والأحاديث الشفوية لجميع لغات بلداننا، كنصوص رقمية في المدونة، كما فعلت وتفعل شعوب اللغات ذات الموارد الرقمية الثرية.

← تأسيس مشاريع متخصصة بالإبداعات اللغوية الخاصة بلغة الضاد، مثل «منظومات الإعراب الآلي». سيساعد ذلك على تصحيح وإكمال تشكيل حروف النصوص، وعلى تحسين ترجمتها الآلية واستيعاب دلالاتها ومعانيها من قبل الكمبيوتر!...

إذ لا يتوقف الإعراب بلغة الضاد على تحليل تركيب وتحريك وعلاقة الألفاظ، لكنه يمتد بعيداً عن ذلك ليخوض في الدلالة (إعراب جملة مثل: «هنا لندن» أنموذجاً):

هنا: ظرف مكان مبني على السكون في محل رفع خبر مقدّم.
لندن: مبتدأ مؤخر ممنوع من الصرف والمانع له العلمية والعجمية.

إمكانية إعراب هذه الجملة يدلّ على الاستيعاب الكليّ لمدلولها. ذلك في غاية الأهمية إذا ما أردنا أن تستوعب الآلة معنى النصّ. إذ لا يمكن الحديث عن الذكاء الاصطناعي الحقّ، في تقديري، بدون ذلك.

ملاحظة: قواعد اللغة السنسكريتية لـ بانيني تشمل ٤٠٠٠ قاعدة نحوية، في كتابه Astadhyayi (خمسة قرون قبل الميلاد!)، برمجها فريق البروفيسور جيرار هويت في الـ INRIA، لتحليل النصوص السنسكريتية آلياً!

ملاحظة ختامية:

استيعابُ العالمِ وامتلاكُ المقدرة على التحليل والبرهنة والتخطيط هو الذكاء الحقيقي (أكثر من موهبة «البغاء الإحصائية» في إعادة صياغة وفذلكة المعارف، كما تفعل حالياً منظومات الذكاء الاصطناعي التوليديّ كـ تشات جي بي تي).

سؤال: هل يمكن لهذه المنظومات أن تنطلق يوماً من معطيات أبيات لغزٍ شعريّ عربيّ قديم، لتشرح، باستدلالٍ آليّ وسيناريو تفصيلي، كيف يمكن أن يكون الرجل عمّاً لعمّته، أو خالاً لخالته، في كلّ الثقافات والحضارات، دون أي علاقةٍ محرّمة؟!

ماذا تردّ المنظومات التوليديّة عند توجيه هذا اللغز العربيّ المدهش لها، بغية تفسيره وحله؟

ولي خالَةٌ وأنا خالها	لي عمّةٌ وأنا عمّها
فإنّ أبي أمّه أمّها	فأمّا التي أنا عمّم لها
ولي خالَةٌ وكذا حكمها	أبوها أخي وأخوها أبي

الجواب: إجاباتها حتّى الآن خاطئة، على نحو مضحكٍ أحياناً. لكنها تتطوّر سريعاً وفي العمق يوماً بعد يوم، وستتمكّن يوماً، كما أتصوّر، من الحلّ الذكيّ له، وليس الحلّ الببغائيّ.

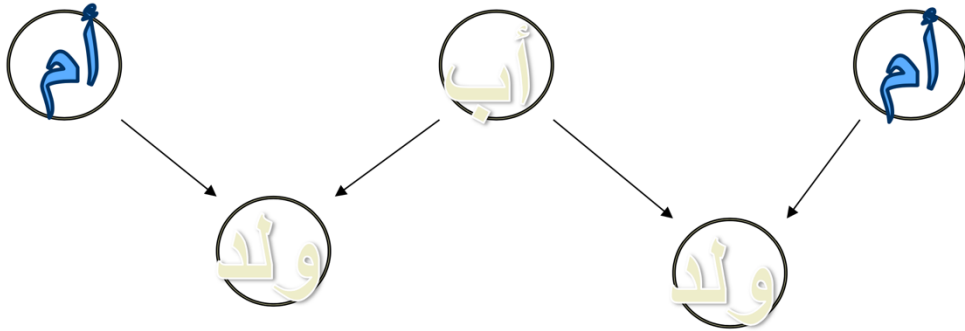
في تقديري الخاص، خارطة الطريق لذكاءٍ اصطناعيّ توليديّ ذكيّ جديرٍ باسمه، لحلّ هذا اللغز، سيكون في المستقبل القريب كما يلي:

(١) تحويل الأبيات آلياً إلى برنامج كمبيوتر بلغة ذكية كـ Prolog، تأخذ معطيات الأبيات كحقائق Fact، تضيف لها قواعد منطقية Rules بسيطة، لشرح ما تعني من علاقات أسريّة مفاهيم: عمّ، خال... (مثل: العمّ أخ الأب...).

٢) تضيف لها قواعد أخرى لمعارف ضمنية مثل:
 أ) إمكانية الطلاق في العلاقة الزوجية؛
 ب) للإنجاب عمرٌ محدد (ليست كل النساء مثل جدتنا سارة، زوجة النبي إبراهيم، التي انجبت وعمرها ٩٠ سنة!)؛
 ج) ولا يأتي الإنجاب من معجزات (ليست كل النساء مثل أمنا مريم: «ابنة ابنها» كما وصفها دانتى!)؛
 د) عدم اعتبار العلاقات المحرمة كانتهاك المحارم والاعتصاب...

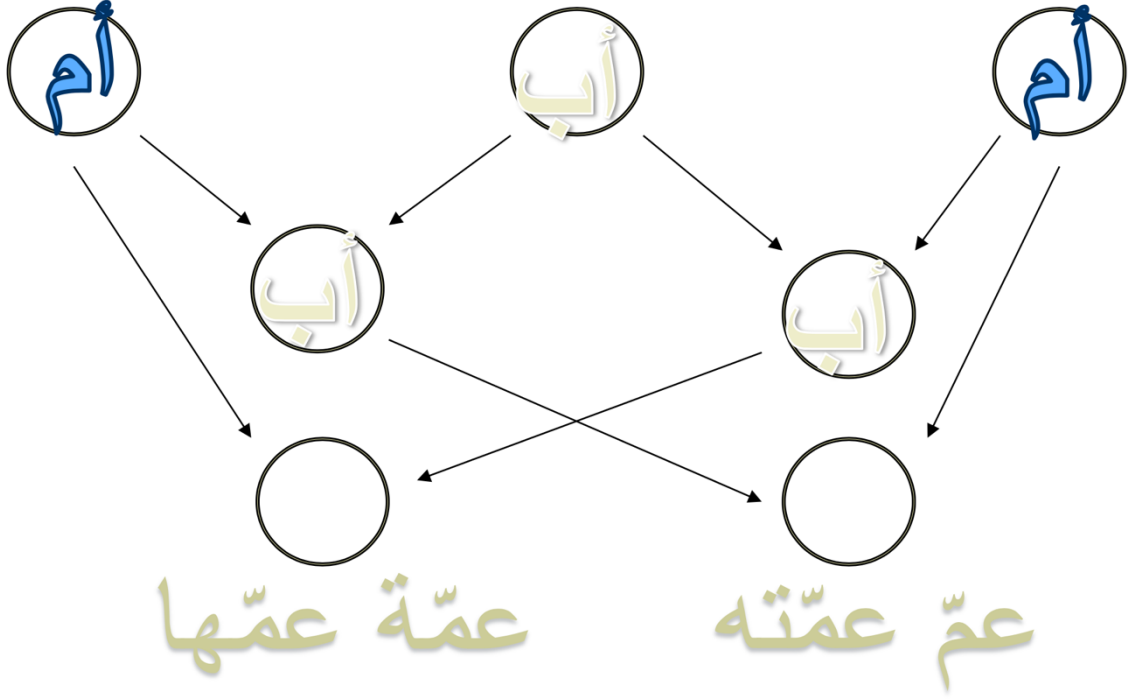
٣) تجيب على السؤال بعد ذلك، عبر الاستدلال المنطقي Logic inference الآلي المبرمج ضمناً في لغات الكمبيوتر الذكية.
 شفافية هذا البرنامج، بكلِّ حقائقه وقواعده المنطقية، وكشفه للمستخدم ضرورية.

حلّ اللغز:
 ينجبُ أبٌ ولدين من امرأتين، كما في الصورة الأولى.



بعد حوالي 17 سنة، يتزوج كل ولد من أم أخيه (من الأب)
 بعد طلاق أبيه منها!

ثم يتحوّل الولدان بدورهما إلى أبوين: تنجب زوجة الأول بعد ذلك ولداً، وزوجة الثاني بنتاً، كما في الصورة الثانية.



هكذا، يصبح المولود الجديد (من يتحدّث بضمير المتكلم في الأبيات الثلاثة) عمّاً لل بنت، والبنت عمّة له في الوقت نفسه.

حلّ الشطر الآخر من اللغز، المرتبط بالخال والخالة، مماثلٌ طبعاً.