

جامعة الملك فهد للبترول والمعادن
قسم علوم الحاسب الآلي والمعلومات



تقرير مشروع وَاسِمَ عَرْمُرْفَ

عمر الذهبي

11/7/2012

المحتويات

3.....	مقدمة.....
3.....	المرحلة ما قبل الأولى: المرحلة النظرية، فهم ومعالجة الوسوم.....
3.....	وصف عام.....
5.....	تفصيل الخطوات.....
5.....	الدراسة والشرح:.....
7.....	تقسيم الشروحات إلى عناصر تصنيف:.....
8.....	إعادة هيكلة الوسوم حسب الرئيس والفرع:.....
10.....	كتابة القواعد:.....
12.....	المرحلة الأولى: تجانس البيانات (الجزء).....
13.....	المرحلة الثانية: إثراء قاعدة البيانات من ذاتها.....
14.....	المرحلة الثالثة: إثراء قاعدة البيانات من خارجها.....
15.....	تطبيق المرحلة الثالثة (مشروع "واسم" الرئيسي).....
15.....	أجزاء التصميم.....
15.....	القواعد:.....
18.....	محرك القواعد:.....
18.....	المحققات:.....
20.....	المشكل الآلي:.....
20.....	المجدّع:.....
20.....	المجذر:.....
20.....	المترجم:.....
20.....	مقيّم الكلام:.....
20.....	ما بعد المعالجة:.....
20.....	المُدخل والمُخرَج:.....
22.....	مقترحات التطوير.....
23.....	النتائج.....

مقدمة

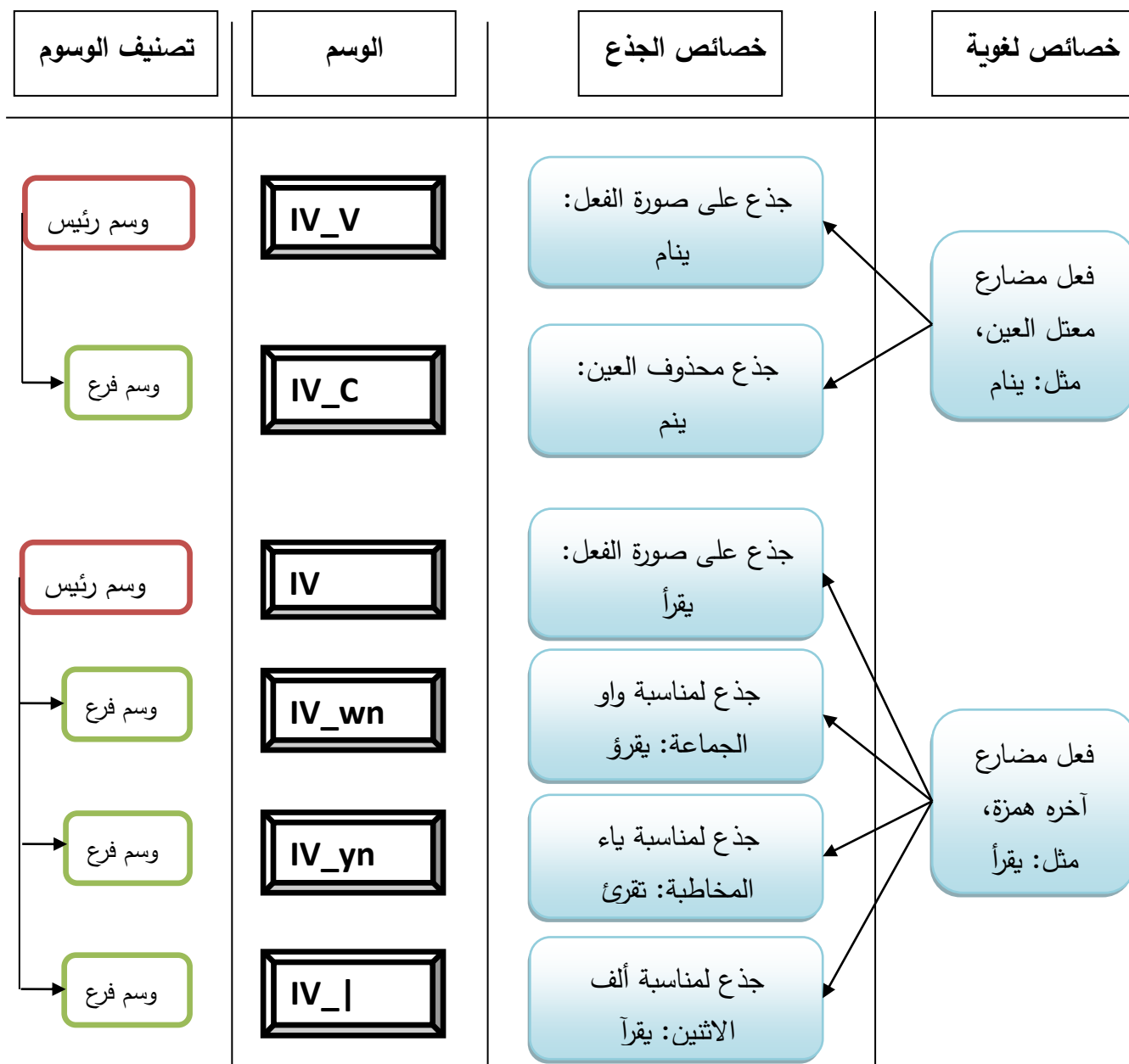
يقوم هذا المشروع على تطوير محلل عَرْمُرف الصرفي¹ بهدف زيادة دقته بإثراء قاعدة بياناته المعجمية. وفيما يلي تقسيم مراحل التطوير.

المرحلة ما قبل الأولى: المرحلة النظرية، فهم ومعالجة الوسوم

وصف عام

إن فهم الوسوم يتطلب دراسة عميقة وملاحظة دقيقة، وقد حاولنا قدر الإمكان تسهيل شرحها فكتبنا وصفا كاملا لكل وسم ثم صنفناها تصنيفا شجرياً. في التصنيف أمر مهم هو أنه يتم على نوعين من الخصائص: خصائص لغوية، وخصائص الجذع، فنجد أوسمة تحوي نفس الخصائص اللغوية ولكن تختلف فيما بينها بخصائص جذوعها، فمثلاً الأفعال المنتهية بهمزة -أي معتلة العين تأتي جذوعها على شكلين: الأصلي، والمجزوم محذوف العين، وكذلك الأسماء والأفعال المنتهية بهمزة -أي مهموزة اللام- تأتي جذوعها على أشكال الهمزة المختلفة (ئ، ء، ؤ، أ، إ، آ)، فهاتان درجتان من التصنيف تتمثلان في (شكل 1) حيث اخترنا مثالا لمعتلّ العين "ينام" ولمهموز اللام "يقرأ".

¹ <http://www.nongnu.org/aramorph/>



شكل 1: بيان تفرقة خصائص اللغة عن خصائص الجذع، وعلاقة النوعين بالوسوم.

وقد سمينّا كل وسم جذعه على صورته الأصلية وسمّا رئيسا (Primary Tag) وسمينا جذعه جذعا رئيسا (Primary Stem)، وسمينا كل وسم جذعه متغير عن الأصل وسمّا فرعا (Secondary Tag) وسمينا جذعه جذعا فرعا (Secondary Stem)، ثم أرجعنا كل وسم فرع إلى أبيه من الأوسمة الرئيسة.

تفصيل الخطوات

مضت هذه المرحلة على خطوات متدرجة هذا تفصيلها وبيان فائدة كلّ منها وما تم منه وما بقي:

الدراسة والشرح:

مرفق بالمشروع ملف شرح الوسوم "TagsComments.xlsx" تجذّه في شكل 2، درسنا كلّ وسم بالنظر في

عنوانه وجذوعه ووسوم عائلته، وهي خمس عوائل:

○ الاسم "N"، 49213 جذعا

○ الفعل الماضي "PV"، 17367 جذعا

○ الفعل المضارع "IV"، 14387 جذعا

○ الأداة "FW"، 1147 جذعا

○ فعل الأمر "CV"، 36 جذعا

شرحناها بتفصيل وتباين، ونقصد بالتباين أن إذا وجدنا وسمّا معتلّ العين مثلا حرصنا أن نبين أن غيره من وسوم عائلته - غير معتلّ العين، وقد أفادنا الشرح المتباين كثيرا في الخطوة التالية، كذلك ميزنا الخصائص اللغوية من خصائص الجذع؛ فبدأنا الشرح بالأولى وثبتينا بالثانية مثل "الجذع على صورة الفعل" أو "في الجذع كذا وكذا" كما ترى في شكل 2، وقد أفادنا هذا في الخطوة الثالثة كثيرا.

أهملنا عائلة الأدوات (FW) لأنها لا تحمل سمات صرفية ولغرافية اجتماع جذوعها تحت عائلة واحدة، فما فهمنا مراد واضعها منها، ولأنها قليلة الجذوع بالنسبة. توقفنا عند بعض الوسوم وجذوعها خيرة، ثم بعد دراسة وتحريّ اكتشفنا أخطاء جليلة فبينّاها في تعليقات بالملف، ولوّنا الخاطئ يقيّنا بالأحمر مع بيان الخطأ، وما يغلب على الظن خطؤه بالبرتقالي، وما نرجو إضافته بالأخضر، وكل ذلك مذكور في الملف.

راجعنا الشروح مرّات، وفي كلّ مرة نصّح ونغيّر، فهذا موطن تطوير في المستقبل؛ وهو مهم لأن عليه اعتماد صياغة القواعد. بالملف أيضا ملاحظات وتعليقات عديدة منها ما يخص جذوعا معينة، ومنها ما يبيّن عيوب تقسيم الوسوم وتغطيتها وما إلى ذلك.

A	B	C	D	E	F	G	H	I	J	K	L
19	IV_0_no-Pref-A	8	<p>أوزان المضارع المبني للمعلوم المضموم</p> <p>أوله (حرف المضارعة): يُفتح، يُغلق، يُنقل.</p> <p>تفاعيل: يُفتقل، يُغقل، وهي أوزان أفعال كل</p> <p>وسم آخره ياء</p>								
20	IV_0_Pass_yu	526	<p>أوله (حرف المضارعة): مضموم - الجذع هو الفعل ذاته على صورته مرفوعاً أو منصوباً</p> <p>فعل مضارع، معبّد أو لازم، معتل الآخر بألف مقصورة، فاؤه همزة، مبني للمعلوم، أوله (حرف المضارعة) مفتوح - الجذع هو الفعل ذاته على صورته مرفوعاً أو منصوباً</p> <p>فعل مضارع، معبّد أو لازم، معتل الآخر بألف أصلية ياء (تكتب مقصورة)، ليست فاؤه همزة، مبني للمجهول، أوله (حرف المضارعة) مضموم - الجذع هو الفعل ذاته على صورته مرفوعاً، وآخره ألف مقصورة فلا تلحقه لاحقة</p>								
21	IV_0_Pass_yu	1	<p>هذا الـوسم خطأ طباعاً، وجذعه الوحيد يجب أن ينتقل إلى IV_0_Pass_yu</p>								
22	IV_0hAnn	393	<p>فعل مضارع، معبّد، أحرفه الأصلية صحيحة، فاؤه همزة، مبني للمعلوم، أوله (حرف المضارعة) مفتوح - لا يجوز من أحرف المضارعة لهذا الجذع إلا الألف، وقد نُزعت منه، وكذلك نُزعت فاؤه (ألف)، وهاتان الألفان أدخيتا إلى "آ" لتدخل على الفعل سابقة.</p>								
46	IV_need-Pref-I	9	<p>فعل مضارع، معبّد، أحرفه الأصلية صحيحة، فاؤه واو صحيحة، مبني للمعلوم، أوله (حرف المضارعة) مضموم - الجذع: لا يجوز من أحرف المضارعة لهذا الجذع إلا الألف، وقد نُزعت منه</p>								
47	IV_need-Pref-A_yu	11	<p>فعل مضارع، معبّد، أحرفه الأصلية صحيحة، فاؤه واو صحيحة، مبني للمعلوم، أوله (حرف المضارعة) مضموم - الجذع: لا يجوز من أحرف المضارعة لهذا الجذع إلا الألف، وقد نُزعت منه</p>								
48	IV_no-Pref-A	130	<p>فعل مضارع، معبّد، أحرفه الأصلية صحيحة، غير مُضاعف (باليين واللام)، فاؤه همزة، ليس آخره (لامه) نونا، مبني للمعلوم، أوله (حرف المضارعة) مفتوح - الجذع هو الفعل ذاته على صورته مرفوعاً أو منصوباً</p>								
49	IV_no-Pref-A_yu	9	<p>فعل مضارع، معبّد، أحرفه الأصلية صحيحة، غير مُضاعف (باليين واللام)، فاؤه همزة، ليس آخره (لامه) نونا، مبني للمعلوم، أوله (حرف المضارعة) مضموم - الجذع هو الفعل ذاته على صورته مرفوعاً أو منصوباً</p>								
50	IV_Pass	1	<p>ليس لهذا الـوسم فائدة، بل هو غير منطقي، إذ ليس من مبني للمجهول غير مضموم حرف مضارعة. وجذعه الوحيد يجب أن ينتقل إلى IV_Pass_yu</p>								
59	IV_wn	109	<p>فعل مضارع، معبّد، آخره (لامه) همزة، مبني للمعلوم، أوله (حرف المضارعة) مفتوح، قبل آخره (صينه) إما صحيح غير مكسور أو علة قبلها غير مكسور - في الجذع همزة على واو فلا يجوز إلا أن تلحقها واو الجماعة، فالجذع إذا يوافق وزنين من الأفعال الخمسة هما "تفعلون" و"تفعلون"، وعليه يسبق الجذع من حروف المضارعة الياء والتاء دون الألف والنون</p>								
60	IV_wn_int_r_yu	1	<p>فعل مضارع، لازم، آخره (لامه) همزة، مبني للمعلوم، أوله (حرف المضارعة) مضموم، قبل آخره صحيح مكسور أو علة قبلها مكسور (أي ياء مد) - في الجذع همزة على نبرة فيجوز بألف الاثنين أو واو الجماعة أو ياء المخاطبة، فالجذع إذا يوافق خمسة الأفعال الخمسة</p>								
61	IV_wn_yu	17	<p>فعل مضارع، معبّد، آخره (لامه) همزة، مبني للمعلوم، أوله (حرف المضارعة) مفتوح، قبل آخره صحيح مكسور أو علة قبلها مكسور (أي ياء مد) - في الجذع همزة على نبرة فيجوز أن يلحقها ألف الاثنين أو واو الجماعة أو ياء المخاطبة، فالجذع إذا يوافق خمسة الأفعال الخمسة</p>								
62	IV_yn	116	<p>فعل مضارع، معبّد، آخره (لامه) همزة، مبني للمعلوم، أوله (حرف المضارعة) مفتوح، قبل آخره صحيح مكسور أو علة قبلها مكسور (أي ياء مد) - في الجذع همزة على نبرة فلا يجوز إلا أن تلحقها ياء المخاطبة، فالجذع إذا يوافق وزناً واحداً من الأفعال الخمسة هو "تفعلين"، وعليه فلا يسبق الجذع من حروف المضارعة إلا التاء</p>								
65	IV_yn_yu	6	<p>فعل مضارع، معبّد، آخره (لامه) همزة، مبني للمعلوم، أوله (حرف المضارعة) مفتوح، قبل آخره (صينه) علة - في الجذع همزة على نبرة فلا يجوز إلا أن يلحقها ياء المخاطبة أو ألف الاثنين، فالجذع إذا يوافق ثلاثة من الأفعال الخمسة هم "تفعلون" و"تفعلان" و"تفعلين"، وعليه فلا يسبق الجذع من حروف المضارعة إلا الياء والتاء</p>								
66	IV_ynAn	0	<p>فعل مضارع، لازم، آخره (لامه) همزة، مبني للمعلوم، أوله (حرف المضارعة) مفتوح، قبل آخره (صينه) علة - في الجذع همزة على نبرة فلا يجوز إلا أن يلحقها ياء المخاطبة أو ألف الاثنين، فالجذع إذا يوافق ثلاثة من الأفعال الخمسة هم "تفعلون" و"تفعلان" و"تفعلين"، وعليه فلا يسبق الجذع من حروف المضارعة إلا الياء والتاء</p>								
67	IV_ynAn_i_ntr	0	<p>فعل مضارع، معبّد، آخره (لامه) همزة، مبني للمعلوم، أوله (حرف المضارعة) مضموم، قبل آخره (صينه) علة - في الجذع همزة على نبرة فلا يجوز إلا أن يلحقها ياء المخاطبة أو ألف الاثنين، فالجذع إذا يوافق ثلاثة من الأفعال الخمسة هم "تفعلون" و"تفعلان" و"تفعلين"، وعليه فلا يسبق الجذع من حروف المضارعة إلا الياء والتاء</p>								
68	IV_ynAn_y_u	0	<p>فعل مضارع، معبّد، آخره (لامه) همزة، مبني للمعلوم، أوله (حرف المضارعة) مضموم، قبل آخره (صينه) علة - في الجذع همزة على نبرة فلا يجوز إلا أن يلحقها ياء المخاطبة أو ألف الاثنين، فالجذع إذا يوافق ثلاثة من الأفعال الخمسة هم "تفعلون" و"تفعلان" و"تفعلين"، وعليه فلا يسبق الجذع من حروف المضارعة إلا الياء والتاء</p>								
69	IV_ynAn_y_u	0	<p>فعل مضارع، لازم، آخره (لامه) همزة، مبني للمعلوم، أوله (حرف المضارعة) مضموم، قبل آخره (صينه) علة - في الجذع همزة على نبرة فلا يجوز إلا أن يلحقها ياء المخاطبة أو ألف الاثنين، فالجذع إذا يوافق ثلاثة من الأفعال الخمسة هم "تفعلون" و"تفعلان" و"تفعلين"، وعليه فلا يسبق الجذع من حروف المضارعة إلا الياء والتاء</p>								

شكل 2 مقتطف من ملف شرح الـوسم

تقسيم الشروحات إلى عناصر تصنيف:

نعني بعناصر التصنيف كل وجه تصنف به الوسوم، فمثلا يُصنّف الوسم بعائلته (اسم، فعل، أداة)، فإذا كان فعلا فبنوعه (ماضي، مضارع، أمر)، ثم يُصنّف من وجه تعديه أو لزومه، ومن وجه لاهه أمهموز أم معتل أم نوني... إلخ، ومن وجه عينه أمضعف أم معتل، وهكذا، فكل وجه من هذه عناصر تصنيف.

تُستخرج العناصر بالنظر في الشروح وما تتباين فيه وما تجتمع، لذلك -كما ذكرنا في الخطوة السابقة- سهل علينا الشرح المتباين هذه الخطوة، شكل 3 مخرَج هذه الخطوة، انظر تجد كل عنصر عمودا، لاحظ أيضا أن بعض العناصر لا تناسب جميع الوسوم؛ فمثلا عنصر نوع الفعل (ماضي، مضارع، أمر) لا يناسب الأسماء! إذاً نجعل لكل طائفة من الوسوم مشتركة في العناصر جدولاً.

في عملنا استخرجنا عناصر وسوم الفعل الماضي والفعل المضارع، وقد اخترناهما لأنهما أصعب الوسوم وأشغيبها، ولأن جذوعهما أكثر من 60% من مجموع الجذوع، وليكونا مثالاً جيداً يحتذي به كل من يكمل العمل.

الملف المرفق لهذه الخطوة "TagsClassified.xlsx"، والجدولان في شكل 3 من ورقتي "IV_Class" و"PV_Class"، علينا أن ننبه أن أخطاءه أكثر من أخطاء أبيه -ملف الخطوة السابقة-؛ لأننا في الخطوة الثالثة اكتشفنا أخطاء في ملف الخطوة الثالثة ورثها من ملف الثانية فأصلحنا ما في ملف الثالثة وأصلحنا أصل الخطأ في ملف الأولى لأنه الأهم، وقصرنا عن ملف الثانية، وأنه إنما أرفق مرجعاً ليُعلم كيف انتقلنا خطوة خطوة، وليُرى لتنظيمه، لا لصحته، وكل هذا مذكور في الملف.

العمود الأخير من جدول الأفعال المضارعة في شكل 3 المعنون "Stem Modification" متعلق بالخطوة الثالثة فتركنا شرحه لها.

	A	C	D	E	F	G	H	I	J	K
1	Tag	Class	Sub-class	Last	Before last	First	حرف المصارع	Active/Pass	tr/intr	Stem Modification
8	IV_0hAnn_no-Pref-A	فعل	مضارع	آخره ألف مضبورة	آخره ألف مضبورة	ما تفرق	مفوح	مبني للمعلوم	منحدي	ألف آخره أعيدت باء
9	IV_h_no-Pref-A	فعل	مضارع	آخره ألف مضبورة	آخره ألف مضبورة	ما تفرق	مفوح	مبني للمعلوم	منحدي	ألف آخره تُكتب ممدودة
10	IV_yn_yu	فعل	مضارع	آخره ألف مضبورة	آخره ألف مضبورة	ما تفرق	مضموم	مبني للمعلوم	منحدي	حذف آخره الممثل، والهززة (فعل منحدي)
11										
12	IV_0hwnyn	فعل	مضارع	آخره ألف مضبورة أو واو	آخره ألف مضبورة أو واو	ما تفرق	مفوح	مبني للمعلوم	منحدي	حذف آخره الممثل
13	IV_0hwnyn_yu	فعل	مضارع	آخره ألف مضبورة أو واو	آخره ألف مضبورة أو واو	ما تفرق	مضموم	مبني للمعلوم	منحدي	حذف آخره الممثل
14	IV_0Pass_yu	فعل	مضارع	آخره ألف مضبورة أو واو	آخره ألف مضبورة أو واو	ما تفرق	مضموم	مبني للمجهول	منحدي	حذف آخره الممثل
15	IV_Ann	فعل	مضارع	آخره ألف مضبورة أو واو	آخره ألف مضبورة أو واو	ما تفرق	مفوح	مبني للمعلوم	منحدي	أعيدت ألف آخره إلى أصلها -واو- ما تفرق
16	IV_Ann_Pass_yu	فعل	مضارع	آخره ألف مضبورة أو واو	آخره ألف مضبورة أو واو	ما تفرق	مضموم	مبني للمجهول	منحدي	أعيدت أو كُتبت ألف آخره باء
17	IV_awn_Pass_yu	فعل	مضارع	آخره ألف مضبورة أو واو	آخره ألف مضبورة أو واو	ما تفرق	مضموم	مبني للمجهول	منحدي	أعيدت أو كُتبت ألف آخره واو
18	IV_0hwnyn_no-Pref-A	فعل	مضارع	آخره ألف مضبورة أو واو	آخره ألف مضبورة أو واو	ما تفرق	مفوح	مبني للمعلوم	منحدي	حذف آخره الممثل
19	IV_Ann_no-Pref-A	فعل	مضارع	آخره ألف مضبورة أو واو	آخره ألف مضبورة أو واو	ما تفرق	مفوح	مبني للمعلوم	منحدي	أعيدت ألف آخره إلى أصلها -واو- ما تفرق
20										
21	IV_Cn	فعل	مضارع	لامه نون	لامه نون	ما تفرق	مفوح	مبني للمعلوم	منحدي	مفوكك التصغير أو محذوف الد منحدي
22	IV_Cn_intr	فعل	مضارع	لامه نون	لامه نون	ما تفرق	مفوح	مبني للمعلوم	منحدي	مفوكك التصغير أو محذوف الد لازم
23	IV-n	فعل	مضارع	لامه نون	لامه نون	ما تفرق	مفوح	مبني للمعلوم	منحدي	سليم
24	IV-n_intr	فعل	مضارع	لامه نون	لامه نون	ما تفرق	مفوح	مبني للمعلوم	منحدي	سليم
25	IV-n_no-Pref-A	فعل	مضارع	لامه نون	لامه نون	ما تفرق	مفوح	مبني للمعلوم	منحدي	سليم
26	IV-n_no-Pref-A_yu	فعل	مضارع	لامه نون	لامه نون	ما تفرق	مضموم	مبني للمجهول	منحدي	سليم
27	IV-n_Pass_yu	فعل	مضارع	لامه نون	لامه نون	ما تفرق	مضموم	مبني للمجهول	منحدي	سليم
28	IV-n_yu	فعل	مضارع	لامه نون	لامه نون	ما تفرق	مضموم	مبني للمجهول	منحدي	سليم
29										
30	IV_wn	فعل	مضارع	لامه هززة	لامه هززة	ما تفرق	مفوح	مبني للمعلوم	منحدي	آخره هززة على واو، قبلها ليس منحدي
31	IV_yn	فعل	مضارع	لامه هززة	لامه هززة	ما تفرق	مفوح	مبني للمعلوم	منحدي	آخره هززة على نبرة، قبلها ليس منحدي

	A	C	D	E	F	G	H	I
1	Tag	Class	Sub-class	Last	Before last	First	Active/Pass	tr/intr
2	PV	فعل	ملحني	آخره صحيح	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمعلوم
3	PV_0	فعل	ملحني	آخره ألف مضبورة	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمعلوم
4	PV_0h	فعل	ملحني	آخره ألف ممدودة	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمعلوم
5	PV_Atn	فعل	ملحني	آخره ألف مضبورة	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمعلوم
6	PV_C	فعل	ملحني	آخره صحيح	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمعلوم
7	PV_C_intr	فعل	ملحني	آخره صحيح	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمجهول
8	PV_C_Pass	فعل	ملحني	آخره صحيح	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمجهول
9	PV_Cn	فعل	ملحني	آخره نون	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمعلوم
10	PV_Cn_intr	فعل	ملحني	آخره نون	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمعلوم
11	PV_Ct	فعل	ملحني	آخره ناء	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمعلوم
12	PV_Ct_intr	فعل	ملحني	آخره ناء	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمعلوم
13	PV_h	فعل	ملحني	آخره ألف مضبورة	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمعلوم
14	PV_intr	فعل	ملحني	آخره صحيح	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمعلوم
15	PV_no-w	فعل	ملحني	آخره باء صحيحة	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمعلوم
16	PV_no-w_intr	فعل	ملحني	آخره باء صحيحة	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمجهول
17	PV_no-w_Pass	فعل	ملحني	آخره باء صحيحة	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمجهول
18	PV_Pass	فعل	ملحني	آخره صحيح	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمجهول
19	PV_Pass-a	فعل	ملحني	آخره معنل	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمجهول
20	PV_Pass-aAat	فعل	ملحني	آخره معنل	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمجهول
21	PV_ttAw	فعل	ملحني	آخره معنل	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمجهول
22	PV_ttAw_intr	فعل	ملحني	آخره معنل	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمجهول
23	PV_V	فعل	ملحني	آخره صحيح	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمجهول
24	PV_V_intr	فعل	ملحني	آخره صحيح	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمجهول
25	PV_V_Pass	فعل	ملحني	آخره صحيح	آخره ليست مضبغة ولا معنلة	ما تفرق	منحدي	مبني للمجهول

شكل 3 مقتطف استخراج عناصر التصنيف

إعادة هيكلة الوسوم حسب الرئيس والفرع:

هذه الخطوة وليدة سابقتها وقد عُمِلت في نفس ملف الخطوة السابقة "TagsClassified.xlsx" في ورقة "IV_Class^2"، وشكل 4 صورة منه.

عناصر تصنيف الخطوة السابقة -رغم نظامها- ينقصها أمران: أنها ليست بمستوى من البساطة يسمح بتطبيقها في البرنامج، والثاني أنها لا تميز خصائص اللغة من خصائص الجذوع إلا من كلمة (يخص الجذع) في بعض الخانات، والواجب أن تُعاد هيكلة الوسوم بحيث يوضع كل وسم فرع تحت رئيسه، وهذا بالضبط ما فعلناه في هذه الخطوة، وقد شرحنا هذين المصطلحين في المرحلة ما قبل الأولى: المرحلة النظرية، فهم ومعالجة الوسوم.

والآن لنأمل في شكل 4 لنعرف كيف تتم هذه الخطوة، لاحظ الأمور التالية:

- أولاً أن الأعمدة هي نفسها التي في الخطوة السابقة في شكل 3 باستثناء آخر أربعها سنشرحه قريباً.
- في العمود الأول: كلّ وسمٍ أحمر وسمٍ رئيسٍ، وكلّ يرتقاليّ فرعٍ.
- يلحق العمود الأول أعمدة منقولة من الخطوة السابقة كما ذكرنا، سمينها مواصفات أو شروطاً أو قيوداً، فلنختار "شروط".
- كل وسم رئيسٍ يحمل كلمة "سليم" تحت عمود "Stem Modification"، والعكس صحيح، وإنما يُعرف الوسم رئيساً بسلامته من التحريف، وهذا تعريفه.
- نجد وسماً رئيساً -بالأحمر- تحته وسم فرع واحد، ونجد ما تحته اثنان، وما تحته ثلاثة، وأربعة، وفي الملف أكثر من ذلك، ونجد ما ليس تحته شيء مثل "IV-n_yu".
- الشروط الإضافية بالأصفر، تُشرح في النقطة التالية
- تُطابق مواصفات (شروط) كلّ وسمٍ فرعٍ رئيسه كلاً أو جزءاً، المهم أن لا تخالفه، وتبعية الفرع رئيسه في شرط من الشروط على ثلاث حالات:
- أن يوافق تماماً مثل أول وسم في شكل 4 "فأوه همزة" وابنه كذلك "فأوه همزة".
- أن يمثل الرئيس مجموعة جزئية من الفرع، أي أن يكون شرط الرئيس أكثر تقييداً من شرط الفرع مع شمول شرط الفرع شرط الرئيس، نرى هذا في آخر وسم رئيس في شكل 4 رقم 42: "IV_no-Pref-A" وابنه آخر وسم فرع رقم 46: "IV_need-Pref-I" تحت عمود "لامه نون"؛ نجد شرط الرئيس بقيده من أن تكون لامه نونا أو علة بينما شرط الفرع مطلق أي أنه غير مقيد، ففي هذه الحالة يجوز للوسم الرئيس أن يتبنى الوسم الفرع دون أي شرط إضافي.
- أن يمثل الفرع مجموعة جزئية من الرئيس، أي أن يكون شرط الفرع أكثر تقييداً من شرط الرئيس مع شمول شرط الرئيس شرط الفرع، فمثلاً نجد تحت "IV" ثلاثة أوسمة "لامها همزة" مع أن رئيسها لا يشترط فيه ذلك بل يشترط ما يشمل ذلك؛ إذ ليست الهمزة نونا ولا علة، ففي هذه الحالة يجوز للوسم الرئيس أن يتبنى الوسم الفرع بشرط إضافي الذي هو بطبيعة الحال الأكثر تقييداً وهو شرط الفرع.
- الشروط باللون الأبيض لا يُؤخذ بها في الخطوة التالية (كتابة القواعد) وإنما تُركت مرجعاً وبياناً.
- عمود "Under unmod" يحمل اسم رئيس كل وسمٍ فرعٍ.
- عمود "Stem Modification" يحمل وصفاً للتغيير الذي تعرض له الجذع لينتقل من الوسم الرئيس إلى الوسم الفرع، انظر الوسم الرئيس في شكل 4 رقم 31: "IV" والأوسمة الثلاثة تحته فهو نفس ما يجري في مثال "يقراً" في شكل 1.

	A	C	D	E	F	G	H	I	J	K	L	M	N
1	Tag	Class	Sub-class	Last	Before last	First	حرف المصلح	Active/Pass	tr/Intr	Stem Modification	Under unmod	cond1	cond2
24	IV-n_no-Pref-A	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
25	IV_need-Pref-	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
26													
27	IV-n_no-Pref-A_yu	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
28													
29	IV-n_yu	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
30													
31	IV	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
32	IV_wn	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
33	IV_yn	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
34	IV-	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
35													
36	IV_intr	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
37													
38	IV_intr_yu	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
39	IV_yn_intr_yu	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
40	IV_wn_intr_yu	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
41													
42	IV_no-Pref-A	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
43	IV_wn	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
44	IV_yn	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
45	IV-	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
46	IV_need-Pref-	فعل	مضارع	لازم	لازم	فعل	فعل	مضارع	مضارع	مضارع	مضارع	مضارع	
47													

شكل 4 مقتطف إعادة هيكلة الوسوم

كتابة القواعد:

قررنا شرح هذه الخطوة في فصل "أجزاء التصميم" تحت باب "

تطبيق المرحلة الثالثة (مشروع "واسم" الرئيسي) لشدة ارتباطها بالتطبيق وبعدها عن النظرية.

المرحلة الأولى: تجانس البيانات (الجدوع)

في هذه المرحلة تُفحص البيانات لاكتشاف أخطاء عدم التجانس ثم تُصحح هذه الأخطاء، والمقصود بالتجانس أن جميع الجذوع التي تتشارك نفس الخاصية اللغوية متوفرة في قاعدة البيانات، ولنستخدم نفس المثال السابق الموضح في (شكل 1)؛ فلو توفرت الجذوع "يقرأ" و"يقرؤ" و"تقرئ" دون "يقرأ" قلنا بعدم تجانس بيانات الفعل المضارع "يقرأ". إن تنفيذ هذه المرحلة يتطلب فهم الوسوم -المرحلة السابقة- فهما عميقا شاملا، لحل المشكلة قسمنا أولا الوسوم إلى مجموعات؛ لكل مجموعة خاصية لغوية مثل خاصية "فعل مضارع معتل العين" أو "فعل ماضي معتل الآخر" أو خاصية أعم "فعل معتل الآخر"، بسبب التباين في درجة العموم والخصوص لكل مجموعة نجد بعضها يتقاطع؛ أي يشترك في وسم أو أكثر. بعد ذلك كتبنا مجموعات من القواعد واحدة لكل مجموعة وسوم، ثم وضعنا خوارزمية تستخدم القواعد لاكتشاف أخطاء عدم التجانس.

عدم التجانس ينقسم حسب اعتماده على الخصائص؛ فهو لغوي إن اعتمد على خصائص اللغة، وجذعي إن اعتمد على خصائص الجذوع، فالأول تصححه المرحلة الأولى والثاني الثانية، والنوعان منفصلان مستقلان؛ لذلك فالمرحلتان مستقلتان؛ أي أن إحداها لا تعتمد على الأخرى مطلقا، وإنما قُدمت الأولى لبساطة قواعدها.

المرحلة الثانية: إثراء قاعدة البيانات من ذاتها

تسمى كذلك "تجانس البيانات اللغوي" لأنها مشابهة للمرحلة الأولى إلا أنها تعالج خصائص اللغة بدل خصائص الجذوع. وإنما سميت إثراء لأنها بمثابة إضافة مفردات جديدة لا صوراً مختلفة لنفس المفردة فقط كما المرحلة الأولى.

إن من عدم التجانس اللغوي يسير وضخم، أما اليسير فيُصحّح حال اكتشافه، ومثاله ما اكتشفته خوارزمتنا فعلى ماضيين لا مضارع لهما هما "طهق" و"شتر"، وأما الضخم فمثل الأفعال التي لا أمر لها؛ وهي تقريباً كل الأفعال عدا حوالي عشرين، ومثل هذه تتطلب قواعد أعقد وأشمل.

المرحلة الثالثة: إثراء قاعدة البيانات من خارجها

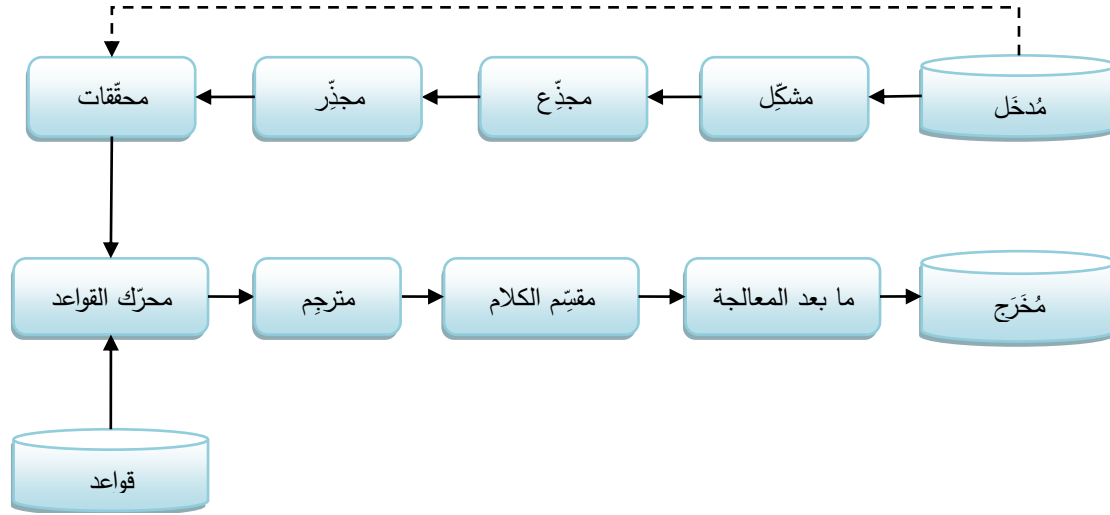
تعتمد هذه المرحلة على إتمام سابقتها بل على إتقانها لسببين؛ اشتراكهم في الأساس وهو كتابة القواعد، والثاني أن لا يُضاف إلى بيانات ناقصة بها أخطاء، ومن أجل السبب الثاني نحبّذ أن يُبقى على مُخرج هذه المرحلة منفصلاً بغير دمج مع البيانات الأصلية. في هذه المرحلة يُتعرّف على الخصائص اللغوية لكلمة غير متوفرة بأي صورة في قاعدة البيانات، لنستعن بشكل 1 مرة ثالثة؛ تحتوي هذه العملية خطوتين رئيسيتين:

- الخطوة الأولى: تصنيف الكلمة لغوياً، وهي عملية غير يسيرة يتعدى إطار بحثها مجرد استخدام القواعد إلى التعلم الآلي، الأكيد أن القواعد مهما كثرت وفصلت لن تكفي لتحديد جميع الخصائص اللغوية. ويمكن لتسهيل هذه الخطوة أن يُؤتى بالكلمات الجديدة من قواعد بيانات منمقة ومعدة لغرض الحوسبة -مثل قاعدة بيانات محلل الخليل الصرفي- بدل النصوص غير المعالجة.
- الخطوة الثانية: هب أن خوارزمية الخطوة الأولى حسّبت الكلمة الجديدة فعلاً مضارعاً معتلاً العين، إذاً عليها إضافة الكلمة بصورتين كلّ جذع: الأول صورة الفعل مرفوعاً فتسمه بالوسم الأول، والثاني صورته مجزوماً فتسمه بالوسم الثاني. لاحظ أن هذه الخطوة لا تحتاج فوق القواعد المستخدمة في المرحلتين السابقتين شيئاً، ولكنها تحتاج اكتمالها لتغطي جميع الحالات فتصير هذه الخطوة آلية.

ثم نُبّهنا أن يمكننا بدء الخطوة الثانية على افتراض تمام الخطوة الأولى، فبدأ العمل فيها حتى انتهت إلى مشروع مستقل، وفيما يلي تفصيل هذا المشروع الذي هو تمديد للمشروع الأساسي.

تطبيق المرحلة الثالثة (مشروع "واسم" الرئيسي)

هدف هذا المشروع بناء البنية الأساسية لأداة تنثري قاعدة البيانات من خارجها، في شكل 5 هيكل المشروع يليه بيان كل جزء وهدفه النهائي وتطبيقه الحالي وكيفية تطويره في المستقبل، صُمم الهيكل على ضوء دراسة المرحلة ما قبل الأولى وعلى ضوء ما أطلعنا عليه عمل المشروع الأساسي من خفايا النظام وهندسة باطنه محركا وبيانات.



شكل 5 أجزاء مشروع واسم

أجزاء التصميم

هذا بيان كل جزء ظهر في شكل 5 وهدفه النهائي وتطبيقه الحالي وكيفية تطويره في المستقبل.

القواعد:

ليست القواعد مفهوما معقدا، لكن وضعها وترتيبها ونقلها إلى مستوى فهم الحاسوب عملية طويلة؛ إن لم يدخلها رفق وتأنٍ كثرت فيها الأخطاء وخربت المشروع كله. يجب على القواعد أن تكون في أبسط صورة حتى يفهمها الحاسوب؛ افترض أنا نريد فك التضعيف من الفعل "يمدّ" ليصير "يمدد"، فلا تكون القاعدة مثلا "إذا كان عين الفعل مضعفا؛ فكّ التضعيف"، بل تكون: "إذا كان آخر حرف مشددا؛ أزل الشدة وكرر الحرف"، فكلما كانت القاعدة أبسط كانت أقرب من مستوى لغة البرمجة وكان تطبيقها بلغة البرمجة أسهل.

كما لاحظت تتكون القاعدة من شرط وفعل، فإذا تحقق الشرط فُعل الفعل، وإلا فلا، بهذه البساطة، تماما مثل "إذا فإن" أو "If Then" المشهورة في لغات البرمجة. فلماذا إذاً لا نطبقها بالبرمجة على شكل عبارات "If Then" ؟ لأن القواعد كثيرة ذات رتب ودرجات منها مباشر بسيط ومنها مركب، فهي تحتاج ترتيبا خاصا لتسهيل كتابتها ومعاينتها ومراجعتها وتعديلها ثم يسهل على الغير فهمها وإثراؤها بعد.

اعلم أن وراء كل شرط من شروط القواعد محقق لها، ولكل محقق دالة في المصدر (source code)، ومن هذي الدوال المحققة سهل ومعقد، فالمعقد ما لا يمكن تبسيطه بحال مثل شرط "فعل لازم"، فهذا صعب على الحاسوب فهمه ولا يمكن

تبسيطه بطرق حسابية أو غيرها لينزل إلى مستوى فهم الحاسوب، لذلك تجد في البرنامج هذه المحققات مرصوصة كلها تبدأ بـ "estim"، وهي محل تطوير رئيسي، وسنتحدث عنها تحت جزء المحققات من أجزاء التصميم.

وقد قمنا بهذا الترتيب الخاص لتحقيق السهولة المرجوة، فنظّمنا القواعد في ملف إكسل كما في شكل 6 هو "BuckTagRules.xls"، وفضلنا صيغة XLS على CSV لأن الثانية لا تدعم خطوطاً مختلفة ولا ألوان ولا أي وسائل إيضاح، وهذه وسائل يحتاجها كاتب القواعد لتسهيل عليه التركيز خصوصاً في المستقبل مع زيادة عدد الجداول والقواعد. فهُم نظام الملف مهمٌّ لإضافة قواعد في المستقبل، ففي النقاط التالية شرّح وافٍ لنظامه، فإن قَصُر شرحنا هنا عن أمر غفلناه فلنرجع إلى [توثيق Drools](#) وليبحث عن Decision Table.

	A	B	C	D	E	F	G	H	I	J	K	L	M	
1														
2		RuleSet	com.amt.rules						omar:					
3		Import	com.amt.rules.*						Blank = Don't Care					
4		Notes	This decision table is for classifying Buckwalter Tags											
5														
6		RuleTable BuckwalterPrimaryTags												
7		NAME	CONDITION	CONDITION	CONDITION	CONDITION	CONDITION	CONDITION	CONDITION	CONDITION	CONDITION	CONDITION	ACTION	
8			word1: VocStem	getStemClass()	getStemSubClass()	getLast() in (\$param)	getLast() not in (\$param)	isBeforeLastDblOrVowel()	isBeforeLastHamza()	isFirstHamza()	isImpYa()	isPassive() != \$param	isTransitive()	TagResult.setTag(\$param); word1.setPrimaryTag(\$param);
9														Resulting Primary Tag
10	Rules	Rule Names	Class	Sub-class	Last Letter in	Last Letter not in	Is Before Last == Last OR a vowel ?	Is Before Last a hamza ?	Is First Hamza	ya/ya	Actv/Pass	tr/intr		
11		IV_P_Rule 1	verb	imperfect	"ي"			false	false	true	true			IV_0
12		IV_P_Rule 2	verb	imperfect	"ي"			true	true	true	true			IV_0_no-Pref-A
13		IV_P_Rule 3	verb	imperfect	"ي"			false	false	false	true	true		IV_0hAnn_yu
14		IV_P_Rule 4	verb	imperfect	"ي", "و"			false	false	true	true	true		IV_0hAnn
15		IV_P_Rule 5	verb	imperfect	"ن"			false	false	true	true	true		IV-n
16		IV_P_Rule 6	verb	imperfect	"ن"			false	false	true	true	false		IV-n_intr
17		IV_P_Rule 7	verb	imperfect	"ن"			false	true	true	true	true		IV-n_no-Pref-A

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	
35		RuleTable BuckwalterSecondaryTags															
36		NAME	CONDITION	CONDITION	CONDITION	CONDITION	CONDITION	CONDITION	CONDITION	CONDITION	CONDITION	ACTION	ACTION	ACTION	ACTION		
37			word1: VocStem	getPrimaryTag() in (\$param)	isTransitive()	isLastHamza()	getLast() in (\$param)	getLast() not in (\$param)	isBeforeLastHamza()	isBeforeLastDiacritic(\$param)	isBeforeLastDiacritic(\$param)	isBeforeLastDoubled()	getBeforeLast() in (\$param)	word1.replaceFirst(\$param);	word1.replaceBeforeLast(\$param);	word1.replaceLast(\$param);	word1.getSecondaryTagStems().add(new TagStem(\$param); word1.getModStem());
38																	
39	Rules	Rule Names	Primary Tag	Is Transitive	Is Last Hamza ?	Is Last Letter any of	Last Letter not any of	Is Before Last a Hamza ?	Is Before Last Diacritic	Is Before Last Diacritic not	Is Before Last Doubled ?	Is Before Last	Replace First	Replace Before Last	Replace Last	Resulting Secondary Tag	
40		IV_S_Rule 1	IV_0	true											"ي"	IV_h	
41		IV_S_Rule 2	IV_0	true											"ي"	IV_0hwnyn	
42		IV_S_Rule 3	IV_0												"ي"	IV_Ann	
43		IV_S_Rule 8	IV_0hAnn_yu					true							"ي"	IV_yn_yu	
44		IV_S_Rule 9	IV_0hAnn_yu												"ي"	IV_0hwnyn_yu	
45		IV_S_Rule 10	IV_0hAnn												"ي"	IV_0hwnyn	
46		IV_S_Rule 11	IV-n_no-Pref-A												"ي"	IV_need-Pref	
47		IV_S_Rule 12	IV		true				KASRA						"و"	IV_wn	
48		IV_S_Rule 13	IV		true				KASRA						"و"	IV_yn	
49		IV_S_Rule 14	IV		true				FATHA						"ي"	IV	
50		IV_S_Rule 15	IV_intr_yu		true					KASRA					"و"	IV_yn_intr_yu	
51		IV_S_Rule 16	IV_intr_yu		true				KASRA						"و"	IV_wn_intr_yu	
52		IV_S_Rule 17	IV-no-Pref-A		true					KASRA					"و"	IV_wn	
53		IV_S_Rule 18	IV-no-Pref-A		true					KASRA					"و"	IV_yn	

شكل 6 جدول القواعد المستخدم مباشرة في البرنامج

نقاط شرح ملف القواعد (يرجع في هذا الشرح إلى شكل 6) :

- موقع الملف في مرفق المشروع src/main/rules/BuckTagRules.xls
- لا اعتبار لأي ورقة عدا الأولى، وفي الملف المرفق ورقتان كنا نستعين بالثانية لنخزّن تجاربنا بينما نترك الأولى لما نعلم أنه صحيح سليم.

- يمكن أن تحمل الورقة عدة جداول، وهي الآن تحمل جدولين، الشرط أن يفصلهما صف خالي.
- الصفوف الثلاثة الأولى بالأسود ليست اختيارية، بل يجب أن تكون بهذه الطريقة ويجب أن يفصلها عن أول جدول صفً خالي، وإن كان محتوى آخر صف "Notes" اختياري إلا أنا نحبذ وجوده للإيضاح.
- أول صف في الجدول (صف 6) يحتوي تعريفه، ويجب أن يحتوي كلمتين لا أكثر بينهما مسافة لا أكثر، ويجب أن تكون الأولى "RuleTable" والثانية اسم الجدول في النظام حسب اختيارك.
- ثاني صف في الجدول (صف 7 في الشكل) لا يسمح إلا بأحد القيم الثلاثة نفسها التي في الشكل، ولا يسمح بها إلا بهذا الترتيب؛ فعمود "NAME" يحمل اسم كل قاعدة في النظام، وعمود "CONDITION" و "ACTION" يحددان ما إذا حمل العمود شرطاً أو قضاءً.
- ثالث صف في الجدول (صف 8 في الشكل) يحوي متغيرات "Variables"، ويجب أن يحمل قيمة لكل عمود شرط، ويمكن أن يحمل قيمة في عمود قضاء، فإذا أردت استخدام نفس المتغير في أكثر من شرط، فيجب أن تكون خلايا المتغير تحت هذه الشروط مجتمعة (Merged)، وهذا ما فعلناه كما ترى في الشكل. صيغة قيمة خلية في هذا الصف: اسم المتغير أولاً (يسمى بقواعد لغة البرمجة)، ثم رأسيتان ":", ثم مسافة، ثم نوع المتغير (Variable Type).
- رابع صف في الجدول (صف 9 في الشكل) هو لبّ الجدول وسحره وأفضل ميزة فيه، هو أنماط الشروط وأنماط القضاء.
- نبدأ بالشروط، في هذا الصف نكتب صيغة الشرط بلغة قواعد Drools، وهي تسمح ببعض العمليات مثل "in" بمعنى "محتوى في" و "not in" نفيها، وكذلك عمليات المقارنة الأساسية "==" و ">" و "<" ومشتقاتها، وغير ذلك، والأهم أنها تسمح ببناء دوال من برنامجك وتستحضر منها أي قيمة سواء عددا صحيحا (Integer) أو نصا (String) أو Boolean. والأفضل من كل ذلك أنها تسمح باحتواء متغيرات (Parameters) في تعبيرها - تُكتب "\$param" - تُعوّض من الخانات أسفلها، فتكون عبارةً فريدة لكل خلية، ثم تقارن القيمة النهائية الناتجة من دمج نمط العبارة بال parameters فهي صائبة أم خاطئة، إذ يجب أن تعود في النهاية -كأي عبارة منطقية- إلى أحد هذين. فإذا أردت استخدام أكثر من parameter فما عليك إلا أن تسميها بدل "\$param" بـ "1" فـ "2" فـ "3" وهكذا، ثم تفصل قيمها التي تعوّضها بفاصلة ",", في الخلايا أسفلها.
- والآن نشرح أنماط القضاء، تكون عبارات برمجية "Statements" تُكتب تماما كما تُكتب في لغة البرمجة؛ لذلك تنتهي بـ ";"، ويمكن أن تستخدم المتغيرات المعرفة تحت الشروط كما استخدمنا "word1" في جدولنا.
- خامس صف في الجدول (صف 10 في الشكل) به تعليقات ليس لها اعتبار عند البرنامج، وإنما وُضعت لتسهيل تمييز الأعمدة بعبارات مفهومة.
- سائر صفوف الجدول إما أن تحوي القيمة التي تعوض ال "\$param" في نمط ذلك العمود، وإما أن تحمل صائب "true" أو خاطئ "false" النتيجة النهائية لتقييم العبارة، وهذه تؤخذ من مخرج تفصيل الخطوات (شكل 4) بعد تحديد الدوال المحققة لكل من أنماط الشروط -كما ذكرنا بداية هذا القسم وكما سيأتي في شرح المحققات-.
- الجدول الثاني يتبع نفس طريقة الجدول الأول، ولكن علينا أن نبين أن الجدول الأول يحدد الوسم الرئيس، وشروطه نفسها التي بالصفوف الحمراء في شكل 4، ومن بعد ذلك يحدد الجدول الثاني الأوسمة الأفرع، بالتحقق من الشروط الإضافية (بالأصفر في شكل 4) ثم تطبيق ما تحت "Stem Modification" في شكل 4 الذي هو أعمدة القضاء في جدولنا الثاني. فهذا هو الربط بين تلك الجداول النظرية وهذه الجداول العملية.

وبالنقطة الأخيرة تكون منهجية وضع القواعد قد تمت أخيراً، وهي عملية طويلة تبدأ من مرحلة فهم الوسوم على خطوات ذكرنا ثلاثاً منها هناك وتركنا الأخيرة هنا، وإنما قسمنا خطوات هذه المنهجية وكتبناها في مكانين منفصلين في التقرير لأن ثلاث الخطوات الأولى أكثر ارتباطاً بالنظرية من التطبيق، وهذه الخطوة الأخيرة أكثر ارتباطاً بالتطبيق من النظرية.

الهدف النهائي أن تغطي القواعد جميع الوسوم جدولين لكل عائلة منها (جدولاً للرئيسة وجدولاً للفروع)، فقد طبقنا جدولين لعائلة الفعل المضارع "IV"، وبقي مثلهما لعائلة وسوم الفعل الماضي "PV" ومثلهما للأمر "CV"، ومثلهما للأسماء "N"، مع أننا نحسب الأسماء أكثر من عائلة لوجود جموع التكسير والممنوع من الصرف ونحو ذلك مما يزيد تفصيل الأسماء.

محرك القواعد:

بدل أن نكتب قواعدنا بصيغتنا الخاصة رأينا أن نتبع المقاييس المشهورة في عالم البرمجة، ومن هذا الرأي قررنا أن لا نبرمج محرك القواعد بأنفسنا بل نستعمل محركاً متوفراً للجميع؛ لفوائد:

- شهرته وكثرة عدد مستخدميه تعني قلة أخطائه وضمان استمرارية تحديثاته، واتباعه مواصفات ومقاييس عالمية.
- توفير وقتنا ولنركز على القواعد بدل محرك القواعد.

تسمى (Rule Engine) و (Business Rule Engine)، وقد بحثنا حتى اصطفينا من محركات القواعد Drools. ويلي ما اطلعنا عليه من محركات القواعد مع روابطها مرجعاً:

- [Drools](#). يعمل على جاوة، مناسب جداً لتطبيقنا، تنصيبه سهل، إلا أن تعلمه يأخذ وقتاً وقراءة و كثير تجربة.
- [Drools.NET 3.0](#). يعمل على C#، لكنه متأخر عن Drools فما زال لا يدعم الجداول.
- [NxBRE](#). يعمل على C#، أعقد مما نحتاج، يتعامل مع ملفات Visio.
- [SRE \(Simple Rule Engine\)](#). يعمل على C#، أبسط مما نحتاج، لا يدعم الجداول ولا أنماط القواعد.
- [OpenRules](#). يمكن دمجه مع جاوة أو C#، لكنه لا يدعم الجداول.
- [Jess](#). غير مفتوح المصدر.
- [Jena](#).
- [Yajre](#) متوقف من 2002 ولا توثيق.
- [Ridor](#) يعمل على WEKA.
- [BIKE](#) يعمل على C++.
- [وهذه قائمة من محركات مفتوحة المصدر تعمل على جاوة](#)
- وغير ذلك كثير لكنه غير مجاني

المحققات:

عرّفنا المحققات في حديثنا عن القواعد، ونعرّفها هنا، هي دوالٌ كتبناها في برنامجنا تحقق شروط القواعد، لكل شرط دالة تحققه، ومعنى تحققه تخبر إذا كان الشرط صحيحاً أم خاطئاً، وإن كان هذا الدور فعلياً دور محرك القواعد وليس دور المحققات؛ إلا أن التسمية مناسبة بنهاية المطاف، ولتمام الفائدة نبين الدور الفعلي للمحققات: هي تخرج قيمة تمررها إلى محرك القواعد فيقارنها بخانات الجدول في العمود المراد تحقيقه (الشكل) ليخبرنا إن كان الشرط صحيحاً أم خاطئاً لكل

خانة، وتجد في أغلب الأعمدة القيم لا تحتاج مقارنة فهي "صحيح" "true" و"خاطئ" "false"، ففي مثل هذه الأعمدة يقع عبء التحقق كله على الدالة المحققة، وفي غيرها يقع أغلبه عليها، فالأمر في الحالتين على المحققات، لذا وجب التأني والحرص في كتابتها.

ذكرنا أن المحققات محل تطوير رئيسي، وذلك لكثرتها في التطبيق الحالي وصعوبة تطبيق بعضها، ولأنها ستكثر مع وضع قواعد جديدة، لعلها لن تزيد مع وضع قواعد الأفعال الماضية وأفعال الأمر لشيء حالات هاتين العائلتين بعائلة الفعل المضارع -التي طبقناها-؛ ولكنها حتما ستتضاعف مع وضع قواعد عائلة الأسماء. لو نظرت إلى المحققات في المصدر (source code) لرأيت في كل واحدة تحتاج تطويرا تعليقا أوله TODO، وهاك رصها:

- محققة "هل الفعل متعدي؟": والجواب طبعا عكس جواب "هل الفعل لازم؟"، تحتاج هذه الدالة مشروعا خارجيا، وأول ما خطر ببالنا استعانة بمحولات صرفية أخرى مثل "الخليل"، ولكك تسأل: إن كانت بالخليل حاجتنا فلم نستوردها إلى عَرْمُرف بدل أن نستخدم الخليل ؟ والجواب: ربما كانت بيانات الخليل أفضل ترتيبا وأشمل من بيانات عَرْمُرف، لكن محرك عَرْمُرف أفضل من محرك الخليل، وإنما نحاول أن نكسب أكبر قدر من المعلومات لنضعه في بيانات عَرْمُرف، ولذلك نحتاج -مستقبلا- إضافة معلومة الجذر لكل جذع في بيانات عَرْمُرف وغيرها من المعلومات ليصبح عَرْمُرف أشمل من الخليل حتى في البيانات، تبقى نقطة أن طريقة السوابق واللاحق غير عملية، وهذه جوابها أن عند توفر البديل سنعمل على نقل كل ما جمعنا ونمقنا من بيانات عَرْمُرف إلى البديل، وإنما التطور مراحل. لنعد إلى الموضوع، تطبيقنا الحالي لهذه الدالة تطبيق متواضع بسيط، هو ملف نص (Text file) به عمودان، الأول الفعل، والثاني معلومة أمتعدي هو أم لازم. لاحظ أننا نعامل المتعدي لمفعول واحد والمتعدي لمفعولين سواء؛ لأن عَرْمُرف كذلك، فهو لا يتعرف مثلا "أعطيتكه" أو "زوجتُكها"، وتصحيح هذا الأمر يتطلب تعديلا في ملفات عَرْمُرف؛ تحديدا ملف اللواحق وملف التوافق توافقي اللواحق مع الجذوع، وهما "dictSuffixes" و"tableBC"، وعلى العموم فقد عملنا على أساس أن هذه الفكرة طبقت فجعنا اعتبارا في المصدر وفي ملف النص للمتعدي لمفعولين بل وثلاثة. وإنما كان تطبيقنا هذه المحققة متواضعا لأنها خارج إطار مشروعنا، وليتم بناء أركان المشروع وبنيتها الأساسية، وهذا هدفنا كما ذكرنا في مقدمة المشروع. بقي أن تعلم أن هذه المحققة تستفيد من المُدخل قبل تشكيله أو تجذيعه أن إذا لحقه ضمير نصب قضت بالتعدي -ولذا تجد في شكل 5 أجزاء مشروع سهمها متقطعا من المُدخل مباشرة إلى المحققات-، إلا أن هذه الميزة غير مفعلة حاليا لتطلبها توافقا بينها وبين المجدع خارج مشروعنا، وإنما ذكرناها لتتم من بعدنا.
- محققة "هل الفعل مبني للمعلوم؟" وهو طبعا عكس "هل الفعل مبني للمجهول؟" وهذه تعتمد على التشكيل فإذا كان حرف المضارعة مفتوحا وما قبل الآخر مفتوحا -أو ألفا إن كان معتلا- قضت بأنه مبني للمجهول، وإلا فمبني للمعلوم، وهي تحتاج تطويرا لأنها لم تُدرس بما فيه الكفاية، فرب شذوذ عن القاعدة. لا ننسى أن اعتمادها على التشكيل يعني أن إتمامها يتطلب إتمام المشكل.

- محققة "هل فائوه همزة؟"، وهذه تحتاج مجذعا يضمن إزالة السوابق ولا يطغى، أو مؤصلا يستخرج الجذر فيعلم مباشرة فائوه وبالطبع عينه ولامه، وهذا يفيد أيضا إذ يزيد يقيننا في مخرج بقية المحققات التي تتحرى أحرف الأصل وهي:

- "هل فائوه همزة؟"
- "هل عينه همزة؟"
- "هل لامه همزة؟"
- "هل عينه مضعفة؟"
- "هل عينه معتلة؟"

ولنا حديث في المجذع لاحق.

المشكل الآلي:

كما يقول اسمه يأخذ المشكل كلمة غير مشكّلة فيرجعها مشكّلة، نحتاج المشكل كما ذكرنا في بعض المحققات مثل "هل الفعل مبني للمعلوم؟" و"هل حرف المضارعة مفتوح؟". وهو سابق المجذع ترتيبا في المشروع لأن تشكيل كلمة كاملة أسهل من تشكيل جذع، وأبسط مثال عليه أن أل التعريف تخبرنا أن الكلمة اسم فتزيل الشبهة بينه وبين فعل بنفس الأحرف. وهذا الجزء محلّ تطوير، نقترح له دراسة المشكّلات المتوفرة واختيار أحدها أو حتى الاستعانة بعدة منها مجتمعة عن طريق التصويت أو ما هو أدنى.

المجذع:

يُزيل المجذع سوابق الكلمة ولواحقها، وهو مفيد في كثير من المحققات كما ذكرنا ولا غنى عنه، رأينا مشاريع تستخدم ما يسمى "مجدّعات خفيفة"، وسميت خفيفة لأنها لا تحمل قاعدة بيانات، وهذه غير دقيقة وأحيانا تغطي فتزيل من أصل الكلمة؛ مثلا تزيل الفاء من "فارس" فتصبح "ارس" ! لذلك نحبّ استعمال قاعدة بيانات مجهزة لهذا الغرض أو محلا صرفيا آخر، وقد ناقشنا هذا الأمر تحت شرح محققة "هل الفعل متعدّ؟".

المجذر:

أو المؤصل أو مستخرج الجذر، يُعطي جذر الكلمة (أصلها الثلاثي المجرد أو الرباعي المجرد)، وهذا جزء لا نحتاجه إذا أردنا مُخرجا على صيغة معجم بكونتر، وإنما نريد به زيادة معلومات المعجم، ولعله يفيد في أجزاء المشروع المعقدة مثل المجذع ومقسّم الكلام. لم يُطبّق المجذر ولا حتى بصورة بسيطة، وإنما هو اقتراح.

المترجم:

يُعطي هذا الجزء معنى الجذع بالإنجليزية، وتطبيقنا له بسيط ملفّ نصّي (Text file) به عمودان، الأول الجذر، والثاني معنى الجذع بالإنجليزية، وهذا محلّ تطوير من المشروع، يمكن الاستعانة فيه بأي من قواميس آلية أو مترجمات أو معاجم ونحو ذلك.

مقسّم الكلام:

يُعطي هذا الجزء قسم الكلام المنتمي إليه جذعنا، وتطبيقنا له بسيط ملفّ نصّي (Text file) به عمودان، الأول الجذر، والثاني قسمه من الكلام. هذا الجزء محلّ تطوير حتما، ويمكن الاستعانة بمحلات أو قواعد بيانات تخبر بهذه المعلومة. وُضع في التصميم بعد تطبيق القواعد لأنه يمكن أن يستفيد من تحديد الوسم.

ما بعد المعالجة:

نحتاج أن يكون المُخرَج على صيغة مُعجم بكونتر المعدّل (معرّب الجذوع)، فما طبقناه في هذا الجزء يزيل حرف المضارعة من جذوع المضارع، ثم يرصّ معلومات المُخرَج، وهي مبيّنة في جزء المُخرَج.

المُدخل والمُخرَج:

تسمح الواجهة بإدخال كلمة واحدة أو ملف نصّي (Text file) بعدة كلمات يفرق كل اثنين سطر جديد، ننبه على ترميز ملف المُدخّل بـ"UTF-8".

يُخَزَّن المَخْرَج في ملف نصي (Text file)، لكل كلمة مُدْخَلَةٍ يُخْرَج سطرٌ على صيغة مُعْجَم بَكُولْتَر المعدل (معْرَب الجذوع) يضم المعلومات التالية على الترتيب:

- الجذع غير مشكل
- الجذع مشكلاً
- الوسم
- المعنى بالإنجليزية
- قسمه من الكلام

انظر في النتائج شكل 7 للمُدخل وشكل 8 للمُخرج.

مقترحات التطوير

لم تكتمل عملية نقل الوسوم كاملة من شروح إلى قواعد، ففي كل خطوة من خطوات العملية أهملت وسوم، وهذا تفصيلها:

- الشروح: ينقصها شروح الأدوات "FW".
- استخراج عناصر التصنيف: ينقصها عن سابقتها الأسماء "N" وأفعال الأمر "CV".
- إعادة هيكلة الوسوم: ينقصها عن سابقتها الأفعال الماضية "PV".
- كتابة القواعد: لا ينقصها عن سابقتها شيء.

لم تتم جميع أجزاء النظام (شكل 5) ، بقي المجدّع والمشكل الآلي والمترجم والمجدّر ومقسّم الكلام.

فكلّ منها محلّ تطوير، وقد ذكرنا بتفصيل طرق التطوير واقتراحات كلّ في مكانه من التقرير كالتالي:

- القواعد: في قسم "تفصيل الخطوات" من "المرحلة ما قبل الأولى: المرحلة النظرية، فهم ومعالجة الوسوم" كتبنا دليلاً معلّماً لنقل بقية الوسوم من الشرح إلى التطبيق.
- المحققات: في قسم "المحققات" رصصنا المحققات التي تحتاج تطويراً بتفصيل واقتراحات.
- أجزاء النظام: في "أجزاء التصميم" بيّنا الأجزاء المحتاجة مع اقتراحات.

ونرى الأولوية للبدء في تنفيذ ملف الشروح، أولاً لإضافة ما غفله بكونتر من وسوم، وكل ما اكتشفناه من نقص سجّلناه في

ملف شرح الوسوم "TagsComments.xlsx" في ورقة "نقص".

النتائج

اختبرنا برنامجنا بكلمات متنوعة وكانت النتيجة كالتالي:

نتيجة صحيحة



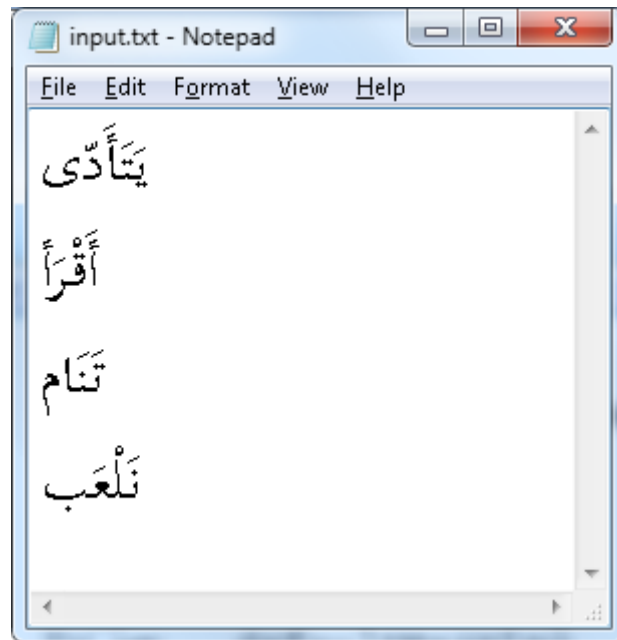
نتيجة خاطئة



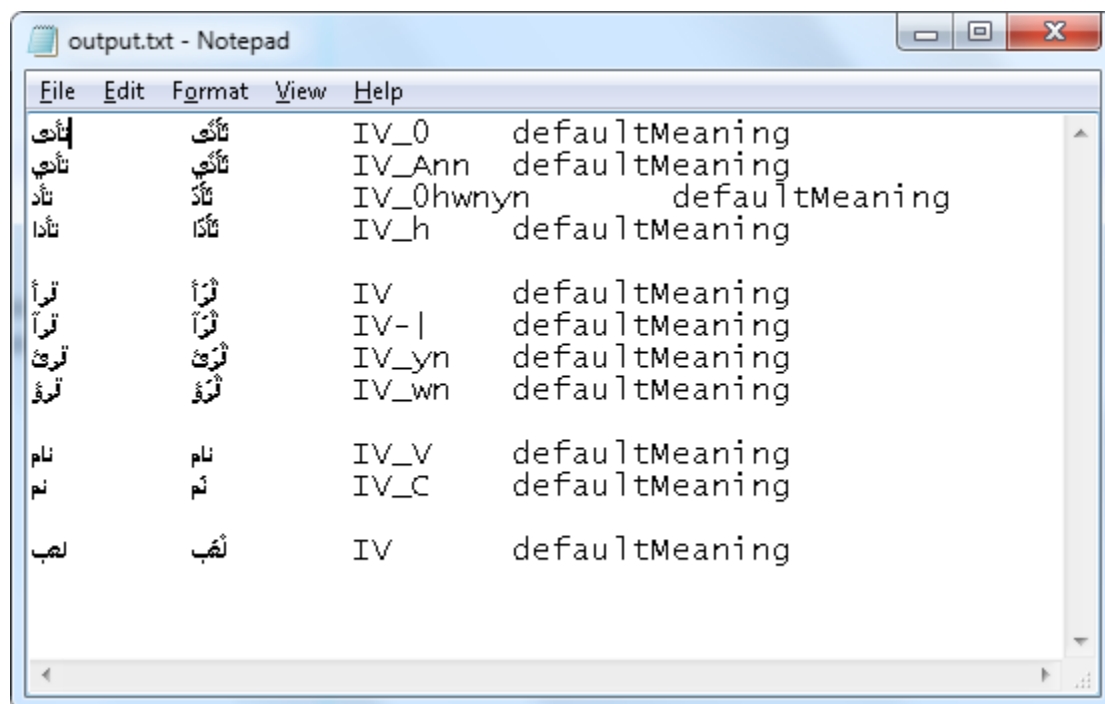
رقم	مُدخَل	الصحيح حسب مُعجم بكوتر	نتيجة برنامجنا
1	يَقُول	IV_V قول	IV_V قول
2	يَحْيَى	IV_0 حَيَّ	IV_0 حَيَّ
3	يُعْنِي	IV_0hAnn عُنِي	IV_0hAnn عُنِي
4	يَزِيد	IV_V زيد	IV_V زيد
5	أَنْظُرُ	IV نُظِرْ	IV نُظِرْ
6	تُكُون	IV_V كُون	IV_V كُون
7	أُلْحِقْ	IV_yu لِحِقْ	IV_yu لِحِقْ
8	تُقَدِّم	IV_Pass_yu قَدِّم	IV_Pass_yu قَدِّم
9	يَرَى	IV_0 رَى	IV_0 رَى
10	يُقَال	IV_V_Pass_yu قَال	IV_V_Pass_yu قَال
11	يَرْجُو	IV_0hAnn رَجُو	IV_0hAnn رَجُو
12	يُصَلِّي	IV_0hAnn_yu صَلِّي	IV_0hAnn_yu صَلِّي
13	يُعْرِفُ	IV_yu عَرَفَ	IV_yu عَرَفَ
14	يَأْتِي	IV_0hAnn_no-Pref-A أَتَى	IV_0hAnn_no-Pref-A أَتَى
15	يُخْرِجُ	IV_intr خَرَجَ	IV خَرَجَ
16	يُفْتَحُ	IV فُتِحَ	IV فُتِحَ
17	أُفْضِلُ	IV_yu فَضِّلَ	IV_yu فَضِّلَ
18	يَنْعَ	IV نَعَّ	IV نَعَّ
19	يُدْخَلُ	IV_Pass_yu دَخَلَ	IV_Pass_yu دَخَلَ
20	تَرَى	IV_0 رَى	IV_0 رَى

19 جذعا من 20 صُنِفَتْ بشكل صحيح كما أخرجت وسوما فرعية وحورتها بما يناسبها تماما كما في شكل 1، وأخطأ في الجذع رقم 15 لسبب بسيط خارج عن إطار مشروعنا هو عدم اكتمال محققة "هل الفعل متعدٍ؟"، فظن برنامجنا "يُخرج" متعديا وهو لازم، ولما أُخبر أنه لازم عن طريق الواجهة وسمه بالوسم الصحيح مما يثبت صحة القواعد.

وفيما يلي نتائج وسوم الفروع وتحويل جذورها مع بيان المُدخلات في ملف نص (شكل 7 عينة مُدخلشكل 7) والمُخرجات كذلك (شكل 8):



شكل 7 عينة مدخل



شكل 8 عينة مخرج

النتائج صحيحة ومطابقة لمعجم بكونتر المعرب، بل هي أصح مما في المعجم؛ إذ تجد فيه جذعا زائدا عما في مخرجنا هو (قرأ) موسوما بـ"IV_wn"، أي يسمح بكلمة "يقرأون" وهذا خطأ إملائي لا يقع فيه برنامجنا، والحمد لله.