



4

18

2160 Sp

نظري

كلية الهندسة المعلوماتية

السنة الخامسة - الذكاء الصناعي

المعالجة الصرفية

د. ندى غنيم



RB Informatics; 09/10/2024

معالجة لغات طبيعية

مقدمة:

تعد دراسة هيكل وقواعد اللغات المختلفة أحد الفروع الأساسية في علم اللغويات، حيث تشكل المعالجة الصرفية جزءاً جوهرياً من فهم اللغة وتحليلها. وتتجلى أهمية المعالجة الصرفية في كونها تهدف إلى تحليل الكلمات وتحديد مكوناتها البنيوية مثل الجذور، واللاحق، والسوابق، والتغييرات التي تطرأ على الكلمة نتيجة القواعد الصرفية. بها أن كل لغة تمتلك نظاماً صرفياً خاصاً بها يعكس طبيعتها وتاريخها، فإن دراسة المعالجة الصرفية لكل لغة على حدة أمر ضروري لتوضيح الفروقات اللغوية وفهم أوجه التشابه والاختلاف بينها. في هذه المحاضرة، سنتناول المعالجة الصرفية في اللغة الإنجليزية، مع التركيز على بنيتها، وأهم قواعدها، وكيفية تحليلها. سيوفر هذا التحليل فهماً أعمق للنظام الصرفي الإنجليزي، مما يساعد على تطبيقه في مجالات متعددة مثل الترجمة الآلية ومعالجة اللغة الطبيعية.

أساسيات اللغة

تلعب القواعد دوراً محورياً في تنظيم اللغة وضمان اتساقها، حيث تحدد كيفية تجميع "الوحدات" اللغوية مثل الأصوات والكلمات والعبارات لإنشاء بنية لغوية أعقد تحمل معانٍ محددة. وتنقسم القواعد إلى ثلاثة أنواع رئيسية، وهي:

القواعد الدلالية:

تهتم القواعد الدلالية بتحديد كيفية ربط الكلمات والجمل بالمعاني التي تحملها. تساعد هذه القواعد على تفسير المعاني الدقيقة بناءً على العلاقات بين الكلمات وسياقها.

مثال:

الجملة "John gave Mary a book"
تظهر العلاقة الدلالية بين John (الفاعل)، و Mary (المستقبل)، و a book (المفعول).

القواعد النحوية:

تتعلق القواعد النحوية بكيفية ترتيب الكلمات لتكوين جمل صحيحة ومنطقية في اللغة. تحدد هذه القواعد بناء الجملة، مثل ترتيب الفاعل والفعل والمفعول به.

مثال:

الجملة "She is reading a book"
صحيحة نحوياً،
بينما "Reading she a book is"
غير صحيحة.

القواعد الصرفية:

تختص القواعد الصرفية بكيفية تشكيل الكلمات من مكوناتها الأساسية مثل الجذور، واللاحق، والسابقة، وذلك بهدف إنتاج كلمات جديدة أو تعديل صيغة الكلمة لتناسب مع السياق.

مثال:

إعادة تشكيل كلمة invest بإضافة السابقة re، واللاحقة ing للحصول على reinvesting التي تعني "إعادة الاستثمار".

الصرف

يُعنى علم الصرف بدراسة كيفية بناء الكلمات وتشكيلها من مكوناتها الأساسية. تعتبر الكلمة الوحدة الأساسية في اللغة، وتتألف عادةً من مجموعة من العناصر الصرفية التي تعمل معاً لتكوين المعنى. هذه المكونات تُعرف باسم الصرفيات (Morphemes)، وهي أصغر وحدة تحمل معنى أو وظيفة داخل الكلمة.

مكونات الكلمة:

1. الجذر: (Stem or Root)

الجذر هو العنصر الأساسي في الكلمة الذي يحمل معناها الأساسي.

مثال:

كلمة act هي الجذر في الكلمات المشتقة منها مثل react أو acting.

2. السوابق: (Prefixes)

هي عناصر تُضاف إلى بداية الجذر لتعديل معناه.

مثال:

re- في react تعني "إعادة".

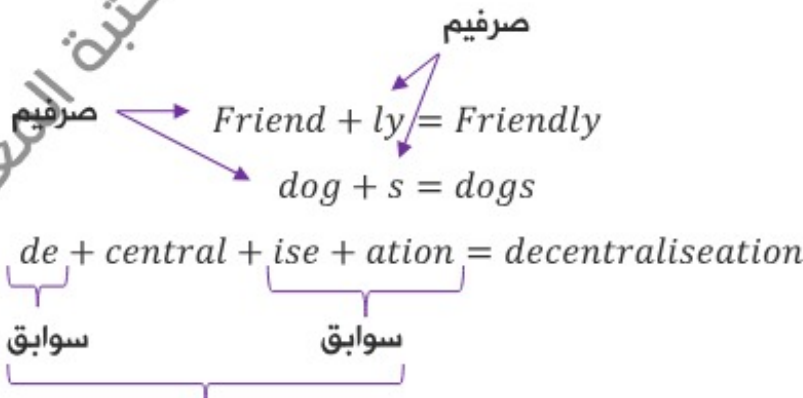
3. اللواحق: (Suffixes)

هي عناصر تُضاف إلى نهاية الجذر لتحديد السياق النحوي أو الدلالي للكلمة.

مثال:

-ing في acting تشير إلى الاستمرارية أو الزمن الحاضر.

أمثلة:



وكل جزء من الكلمة يسمى صرفيم أيضاً



إلى جانب دراسة كيفية تكوين الكلمات من الجذور واللاحق والسابقة، يهتم الصرف بتحليل البنية الداخلية للكلمات، بما يشمل التغيرات التي تطرأ عليها نتيجة القواعد اللغوية أو السياق. يركز هذا الجانب من الصرف على العمليات الصرفية التي تساهم في تشكيل كلمات جديدة أو تغيير صيغ الكلمات الموجودة لتناسب مع متطلبات المعنى والنحو.

الرد إلى أسماء (nominalization)،

enjoy

مثال:

صفوف الكلمات (categories-classes)

■ الصفوف المفتوحة والصفوف المغلقة:

الصفوف المفتوحة

تشمل الأسماء، الأفعال، الصفات، والظروف، وهي تتسم بمرونة كبيرة حيث يمكن أن تضاف إليها كلمات جديدة باستمرار نتيجة للتغيرات الثقافية، الاجتماعية، أو التقنية.

أهمية معرفتها:

يعكس وجود الصفوف المفتوحة الطبيعة الحية والمتطورة للغة. يبرز تحدي بناء النماذج اللغوية القادرة على فهم الكلمات الجديدة أو معانيها في سياقات مختلفة. يساعد في تطوير أنظمة الذكاء الاصطناعي لتحليل النصوص وفهم الكلمات الجديدة بناءً على السياق.

الصفوف المغلقة

تشمل الأدوات مثل أدوات التعريف، الضمائر وحروف الجر، وحروف العطف وغيرها. هذه الصفوف ثابتة نسبياً ولا تشهد تغييرات كبيرة.

أهمية معرفتها:

تسهم في الحفاظ على بنية اللغة وقواعدها. تعد جزءاً أساسياً في تصنيف الكلمات وتنظيم الجملة.

بماذا يساعد تصنيف الكلمات إلى الصفوف؟

تصنيف الكلمات يساعد بمعرفة المعنى مثلاً (عين) إذا عرفت أنها اسم سيختصر هذا معاني كثيرة للكلمة مثل: عَيْن أو عَيْن لأنهم أفعال.

Excellence is to do a common thing in an uncommon way.

■ أنواع الصرف:

الصرف الاشتقاقي :(Derivational morphology)	الصرف التصريفي :(Inflectional morphology)
دمج جذر الكلمة مع morpheme لاشتقاق كلمات بأصناف مختلفة. مثال: Process, processing, processable, processor, ...	دمج جذر الكلمة مع morpheme للحصول على كلمة من الصف نفسه مثل تحويل الفعل من الجذر إلى المضارع. مثال: Give, given, gave, gives, ...

■ أهمية الصرف:

الصرف هو أحد الأركان الأساسية في علم اللغة لأنه يساعد على تحليل الكلمات إلى مكوناتها الصرفية (morphemes)، مما يتيح استخلاص معلومات إضافية على المستويين النحوي والدلالي. تعتمد هذه العملية على المعرفة السابقة ببنية الكلمة لتحديد أجزائها الأساسية ووظائفها ومعانيها في السياق.

■ مثال:

كلمة processable تحلل إلى 2 morphemes هما process و able ويكون:

Process	able	
Process	Possible	معنى كل مورفيم
Root	Suffix	Role
main	Less	Semantic information

■ ملاحظة:

يختلف تعقيد الصرف من لغة إلى أخرى.



We were born to be.
Young and free.



■ الجذور (stems or roots) والملحقات (affixes):

تقسم الكلمات إلى صرفيات: جذوع وملحقات.

هناك عدة أنواع من الملحقات:

- سوابق prefixes: in- -credible
- لواحق suffixes: incred- -ible
- دواخل infixes:

■ مثال في لغة Tagalog في الفلبين:

hingi (to borrow) → *humingi* (agent of the action)

- سوابق ولواحق معاً circumfixes: مثال في الإنكليزية:

unbelievably → *un + believe + able + ly*

■ ملاحظة:

نسمي الصرف بإضافة السوابق واللواحق بالصرف المتسلسل concatenative morphology.

■ اللغات السامية:

في بعض اللغات، مثل العربية والعبرية، لا يكون الصرف متسلسلاً، أي أنه لا يعتمد فقط على إضافة السوابق أو اللواحق إلى الجذر، بل يتم عبر عملية معقدة تشمل إدخال تغييرات داخلية على بنية الكلمة. يُعرف هذا النوع من الصرف بـ **الصرف غير المتسلسل (Non-concatenative Morphology)**، حيث يتم تعديل الجذر نفسه عبر إدخال صوائت (حروف متحركة) في داخله أو إجراء تغييرات صوتية أخرى لتشكيل الكلمات المختلفة مثل: الجذر فعل، فَعَلَ ← فُعِلَ.

■ الفرق بين جذر الكلمة وجذعها في اللغة العربية:

في اللغة العربية، هناك تمييز دقيق بين جذر الكلمة وجذعها. هذا التمييز نابع من الطبيعة الصرفية المعقدة للغة العربية، حيث تعتمد الكلمات على قواعد ونماذج صوتية وصرفية محددة، مما يجعل التحليل الصرفي يتطلب أكثر من مجرد إزالة السوابق واللواحق كما هو الحال في بعض اللغات الأخرى.

■ جذر الكلمة (Root):

الجذر هو الوحدة الأساسية للكلمة، وهو يتألف غالباً من ثلاثة صوائت (Consonants) تمثل المعنى الأساسي.

■ الخصائص:

يمثل المعنى العام للكلمة.

يُعد قاعدة لجميع الكلمات المشتقة.

لا يُظهر صوائت (حروف علة) أو أي تغييرات صرفية.

أمثلة:

الجذر "كتب" يعبر عن معنى الكتابة.

الجذر "أكل" يعبر عن معنى الأكل.

جذع الكلمة (Stem):

الجذع هو الكلمة الناتجة عن إزالة السوابق واللواحق من الكلمة الأصلية، مع الاحتفاظ بالتعديلات البنيوية التي تحدثها النماذج الصرفية في الكلمة.

الخصائص:

يمثل الشكل الفعلي للكلمة بعد تطبيق قواعد الصرف.

يأخذ في الحسبان النماذج الصرفية التي تقحم الصوائت أو تغييرات داخلية في الجذر.

قد يكون معقداً ويحتاج إلى تحليل صرفي عميق.

أمثلة:

في كلمة "كاتب"، الجذع هو "كاتب" (وهو ناتج عن الجذر "كتب" وتطبيق صيغة فاعل).

في كلمة "مكتوب"، الجذع هو "مكتوب" (ناتج عن الجذر "كتب" وتطبيق صيغة اسم المفعول).

لماذا لا يكفي إزالة السوابق واللواحق فقط؟

اعتماد اللغة العربية على النماذج الصرفية:

الكلمات العربية غالباً ما تُبنى بإقحام صوائت داخل الجذر، وليس فقط بإضافة ملحقات.

مثال:

كلمة "مكتوب" لا يمكن تحليلها بمجرد إزالة "م" و"وب"، لأن التغير الصوتي الذي طرأ على الجذر يتبع نموذجاً صرفياً (اسم المفعول).

■ التغيرات الداخلية للجذر:

الأوزان العربية تؤدي إلى تغييرات داخلية في الجذر لإنتاج معاني ودلالات جديدة.

مثال:

الجذر "علم" يعطي كلمات مثل:

"عالم" (اسم فاعل).

"معلوم" (اسم مفعول).

"تعليم" (مصدر مزيد).

■ التعديلات الصرفية الأخرى:

في كثير من الحالات، تكون الكلمات العربية مشتقة من الجذر عبر إضافة سوابق، لواحق، أو تغييرات صوتية، وكلها تحتاج إلى تحليل معمق لتحديد الجذع.

مثال:

كلمة "استكتب" (طلب الكتابة) تبنى على الجذر "كتب" باستخدام صيغة مزيدة بـ "است".

الصرف الحاسوبي (computational morphology):

الصرف الحاسوبي هو مجال ضمن معالجة اللغة الطبيعية (NLP) يهدف إلى تحليل الكلمات وتحويلها بين أشكالها المختلفة بناءً على قواعد صرفية. يمكن تقسيم العملية إلى مرحلتين رئيسيتين:

التحليل (Analysis):

تحويل الشكل الظاهري للكلمة (Surface Form) إلى الشكل القاعدي أو المعجمي (Lexical Form).

مثال:

الكلمة "is" تحلل إلى التمثيل القاعدي be+3+S+Ind+Pres، مما يعني أنها تصريف لفعل "to be" في الزمن الحاضر مع ضمير الغائب المفرد.

التركيب (Generation):

تحويل التمثيل القاعدي إلى الشكل الظاهري للكلمة بناءً على القواعد الصرفية.

مثال:

التمثيل be+3+S+Ind+Pres يُركب ليصبح "is".

■ أمثلة على الصرف التصريفي في المعالجة الحاسوبية:

1. جمع الأسماء:

القاعدة العامة:

تضاف لاحقة "s" للأسماء لتكوين الجمع.

مثال:

book → books

car → cars

الشواذ: (Irregular Plurals)

بعض الأسماء لا تتبع القاعدة العامة للجمع، وتحتاج إلى معالجة استثنائية.

مثال:

mouse → mice

child → children

fox → foxes (إضافة "es" بسبب القاعدة الإملائية للأسماء التي تنتهي بـ "x")

The truth is always exciting. Speak it, then.

Life is dull without it.



2. تصريف الأفعال:

الأفعال النظامية: (Regular Verbs)

تتبع قاعدة عامة لإضافة لاحقات التصريف:



التحدي في الصرف الحاسوبي:

التحدي الأساسي في الصرف الحاسوبي هو تنفيذ عمليتي التحليل والتركيب بدقة وفعالية. يتطلب ذلك التعامل مع مجموعة واسعة من الحالات والقواعد الصرفية التي تشمل القواعد العامة، والشواذ، والسياقات المتنوعة. لتحقيق ذلك، نحتاج إلى أدوات قادرة على معالجة هذه العمليات بشكل منهجي وسريع، وأحد الأدوات المهمة المستخدمة هو المحولات (Transducers).

المحولات (Transducers) في الصرف الحاسوبي:

المحولات هي أنظمة تستخدم لتنفيذ عملية التحويل بين:

التمثيل القاعدي أو النظامي
(Lexical Form):
الشكل المجرد أو المعجمي
لللمة، الذي يتضمن معلومات
نحوية ودلالية.

الشكل الظاهري لللمة
(Surface Form):
الشكل الفعلي لللمة كما
تظهر في النصوص.



لتكن Σ_1 و Σ_2 مجموعتان معدودتان (حروف أبجدية)، و $\Sigma((\Sigma_1 \cup \{\varepsilon\}) \times (\Sigma_2 \cup \{\varepsilon\}) \setminus \{(\varepsilon, \varepsilon)\})$ المحوّل هو أوتومات منته الحالات حتمي Deterministic Finite-State Automata.

Σ_1 : "left" language

: upper language

: input language

Σ_2 : "right" language

: lower language

: output language

مثال:

الحالات: (States)

الحالة 0 (الحالة الابتدائية): تبدأ عملية التحويل من هذه الحالة.

الحالة 1: حالة انتقالية.

الحالة 2 (الحالة النهائية): تنهي عملية التحويل إذا وصلت إلى هذه الحالة.

الانتقالات: (Transitions)

كل انتقال مُحدد بين الحالات يتم تمثيله بالصورة:

"a:b" حيث:

a هو المدخل (Input).

b هو المخرج (Output).

b:ε تعني أن المدخل هو b، بينما لا يوجد مخرج ε (تشير إلى الخرج الفارغ).

الرموز الفردية (رمز دخل دون رمز خرج) تعني أن الرمز يمثل الدخل والخرج.

أمثلة عن تحويلات:

1. التحويل الأول [0,0,2] (bb, b) :

نبدأ في الحالة 0.

المدخل b: يبقى النظام في الحالة 0.

المدخل الثاني b: ينتقل النظام إلى الحالة 2.

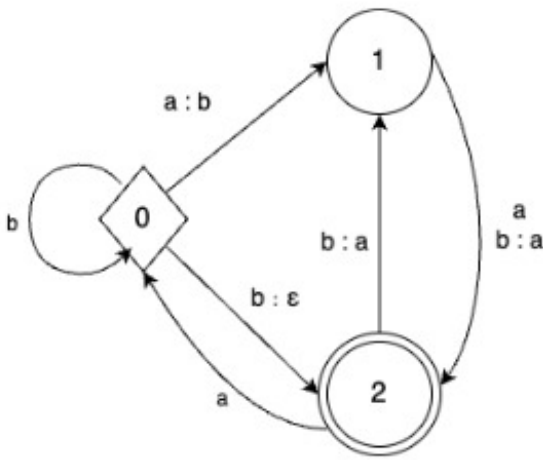
2. التحويل الثاني [0,1,2,0,0,2] (ababb, baab) :

نبدأ في الحالة 0.

المدخل الأول: a المخرج b، ينتقل إلى الحالة 1.

المدخل الثاني: b المخرج a، ينتقل إلى الحالة 2.

المدخل الثالث: a المخرج b، ينتقل إلى الحالة 0.



Initial state

Final state(s)



المدخل الرابع: b يبقى في الحالة 0.

المدخل الخامس: b ينتقل إلى الحالة 2.

استخدامات المحولات:

1. التحقق من الارتباط (association checking):

التحقق من وجود علاقة بين سلسلة مدخل ($string_1$) وسلسلة مخرج ($string_2$) باستخدام المحول.

$$(abba, baaa) \in \Sigma^*$$

2. توليد (generation): $string_1 \rightarrow string_2$

توليد سلسلة مخرج ($string_2$) من سلسلة مدخل ($string_1$).

$$bbab \rightarrow ?$$

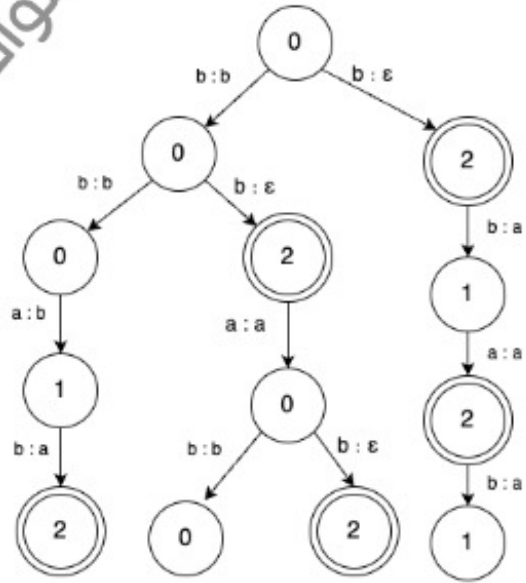
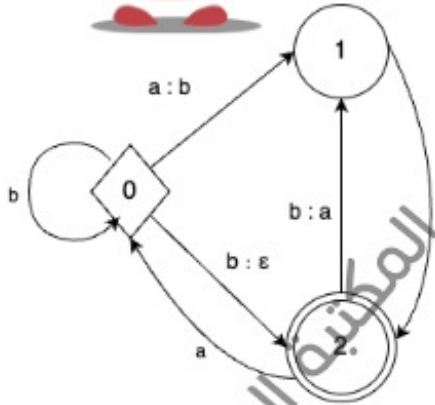
3. تحليل (analysis): $string_2 \rightarrow string_1$

تحليل سلسلة مخرج ($string_2$) وإيجاد سلسلة المدخل الأصلية ($string_1$).

$$? \rightarrow ba$$

مثال (1):

$$bbab \rightarrow ?$$

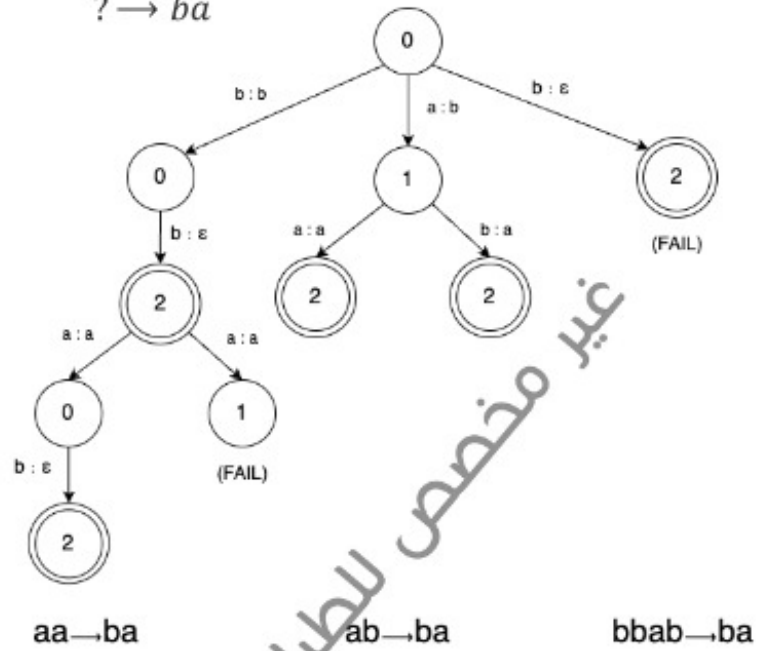
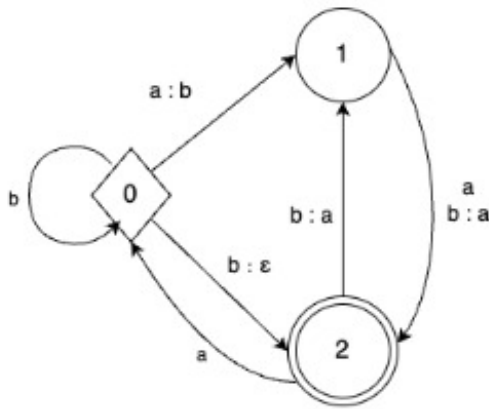


$$bbab \rightarrow bbba$$

$$bbab \rightarrow ba$$

هذا مثال عن توليد سلسلة من سلسلة.

نبدأ من الحالة الأولى ونتحقق من الانتقالات الممكنة فنجد انتقالين ممكنين (للحالة 0 وللحالة 2)، ثم نكرر العملية ذاتها حتى نصل للحالة النهائية عند نهاية الدخل، وتكون السلاسل الناتجة هي السلاسل التي يمكن توليدها من السلسلة الدخل.

$? \rightarrow ba$ 

هذا مثال عن التحليل حيث نتحقق من السلسلة الدخلى التي تنتج عنها السلسلة الخارج. نبدأ بالحالة الابتدائية ونتحقق من الانتقالات التي تتوافق مع الرمز الحالي في الخارج، فنجد انه يمكننا الانتقال من 0 إلى 0 عبر $b:b$ أو الانتقال من 0 إلى 1 عبر $a:b$ أو الانتقال من 0 إلى 2 عبر $b:\epsilon$ ونحصل على سلاسل صحيحة عندما تنتهي رموز سلسلة الخارج مع الوصول لحالة نهائية.

بعض تطبيقات المحولات:

- التحويل من نص مكتوب إلى رموز صوتية grapheme-to-phoneme transduction في نظم تركيب الكلام.
- المرشحات (إضافة، حذف، تعديل بعض العلامات؛ مثل HTML).
- تقطيع النصوص.

الصرف التصريفى

يهتم بدراسة التغيرات التي تطرأ على شكل الكلمات لتلائم الوظائف النحوية في الجملة، دون تغيير معناها الأساسي أو نوعها (مثل تحويل الكلمة من مفرد إلى جمع، أو من زمن المضارع إلى الماضي).

الجمع:

	Regular Nouns		Irregulars Nouns	
Singular	Cat	Thrush	Mouse	Ox
Plural	Cats	Thrushes	Mice	Oxen

Singular	Plural
Ibis	Ibises
Waltz	Waltzes
Finch	Finches
Box	Boxes
Butterfly	Butterflies

الملكية:

الكلمات التي لا تنتهي بحرف الـ "s" يكون الملحق (s).

الكلمات التي تنتهي بحرف الـ "s" يكون الملحق (').

مثال:

llama/ llama's
llamas/ llamas'

الأفعال الشاذة	الأفعال النظامية
Eat, eats, eating, ate, eaten Catch, catches, catching, caught, caught Cut, cuts, cutting, cut, cut	Walk, walks, walking, walked, walked

الصرف الاشتقاقي

الأفعال:

الصرف الاشتقاقي يُعنى بتكوين كلمات جديدة من خلال إضافة ملحقات تُغيّر وسم الكلمة أو معناها.

Suffix	Base Verb/ Adjective	Derived Noun
-ation	Computerize (V)	Computerization
-ee	Appoint (v)	Appointee
-er	Kill (V)	Killer
-ness	Fuzzy (V)	Fuzziness

Suffix	Base Noun/ Verb	Derived Adjective
-al	Computation (N)	Computational

-able	Embrace (v)	Embraceable
-less	Clue (V)	clueless

التحليل الصرفي

التحليل الصرفي هو عملية تعنى بتحليل الكلمات إلى مكوناتها الأساسية (المورفيمات - Morphemes) مع تحديد السمات الصرفية لكل مكون، مثل:

- نوع الكلمة (اسم، فعل، صفة... إلخ).
- العدد (مفرد أو جمع).
- الزمن (مضارع، ماضٍ... إلخ).
- الصيغة (ملكية، تصريف الشخص الثالث، أو تصريف تام...).

أمثلة:

Input	Morphological Parsed Output
Cats	Cat +N +PL
Cat	Cat +N +SG
Cities	City +N +PL
Geese	Goose +N +PL
Goose	(Goose +N +SG) or (Goose +V)
Gooses	Goose +V +3SG
Merging	Merge +V +PRES-PART
caught	(catch +V +PAST-PART) or (catch +V +PAST)

■ المعجم:

هو مجموعة من الجذور والملحقات يتضمن:

Suffix-s

الملحقات

Irreg-sg-nouns

أسماء شاذة في حالة المفرد

Irreg-pl-nouns

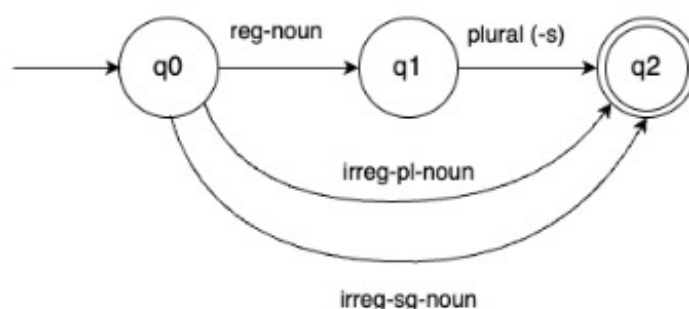
أسماء شاذة في حالة الجمع.

Regular nouns

أسماء نظامية.

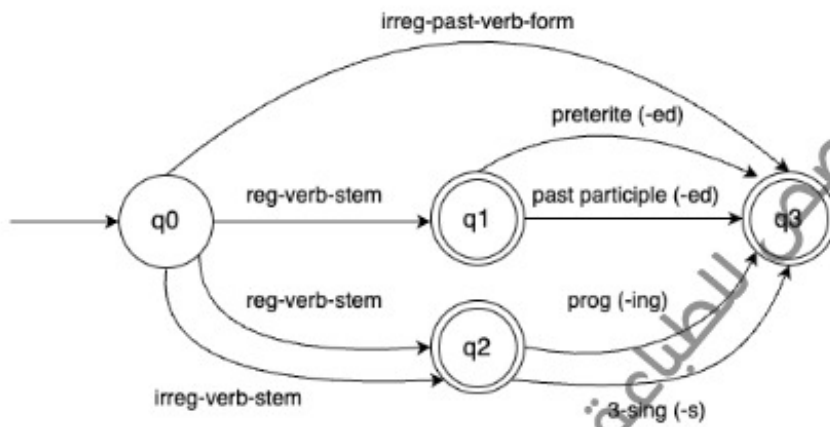
أمثلة لاوتومات التحليل الصرفي في اللغة الإنكليزية:

مثال (1):



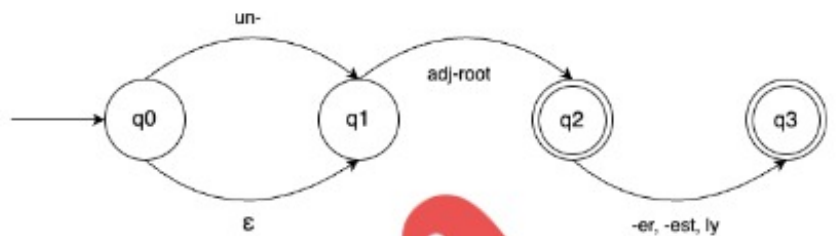
Reg-noun	Irreg-pl-noun	Irreg-sg-noun	Plural
Fox	Geese	Goose	-s
Cat	Sheep	Sheep	
Dog	Mice	Mouse	
aardvark			

مثال (2): FSA لتصريف الأفعال باللغة الإنكليزية:



reg-verb-stem	Irreg-verb-stem	Irreg-past-verb	past	Past-part	Pres-part	3sg
Walk	Cut	Caught	-ed	-ed	-ing	-s
Fry	Speak	Ate				
Talk	Sing	Eaten				
Impeach	Sang					
	spoken					

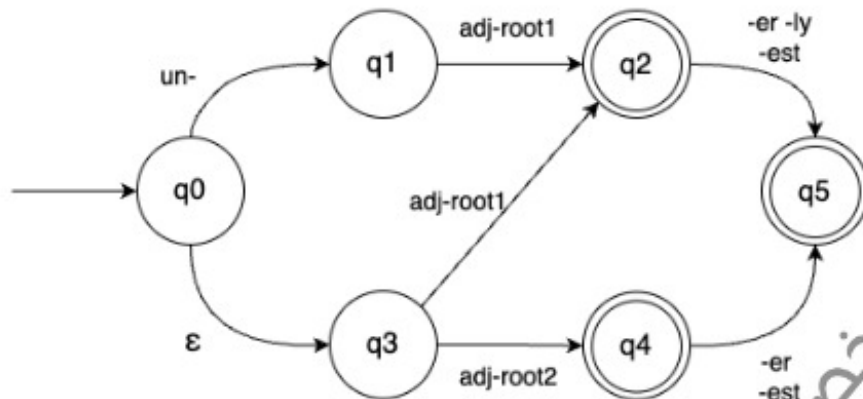
مثال (3): بعض الصفات في اللغة الإنكليزية:



Big, bigger, biggest.
Cool, cooler, coolest, coolly.
Red, redder, reddest.
Clear, clearer, clearest, clearly, unclear, unclearly.
Happy, happier, happiest, happily, unhappy, unhappier, unhappiest, unhappily.
Real, unreal, really.



بإمكان هذا الأوتومات تعرف أغلب الصفات، إلا أنه سيتعرف أيضاً بعض الصيغ غير القواعدية، مثل:
(unbig, redly, realest)



■ ملاحظة:

نقوم باعتماد النماذج اللغوية (Automata) من الخبراء المتخصصين في اللغة، ثم نخضعها لاختبارات دقيقة للتأكد من قدرتها على رفض الصيغ غير القواعدية. فعلى سبيل المثال، فيما يتعلق بالصفات، لا يكون عكس كل صفة بإضافة "un-": وبالتالي، يجب تقديم ملاحظات (Feedback) للخبراء اللغويين بهدف تحسين النموذج وتعزيز دقته.

التحليل الصرفي باستخدام FST

يتم إجراء التحليل الصرفي باستخدام المحولات النهائية محددة الحالة (Finite State Transducers - FST)، حيث يعمل المحول على الربط بين مجموعتين من الرموز، إحداها تمثل المدخلات والأخرى تمثل المخرجات. تبني هذه المحولات على مفهوم الأوتومات المنتهية الحالة (Finite State Automata - FSA).

	c	a	t	+N	+PL			Lexical
	c	a	t	s				Surface

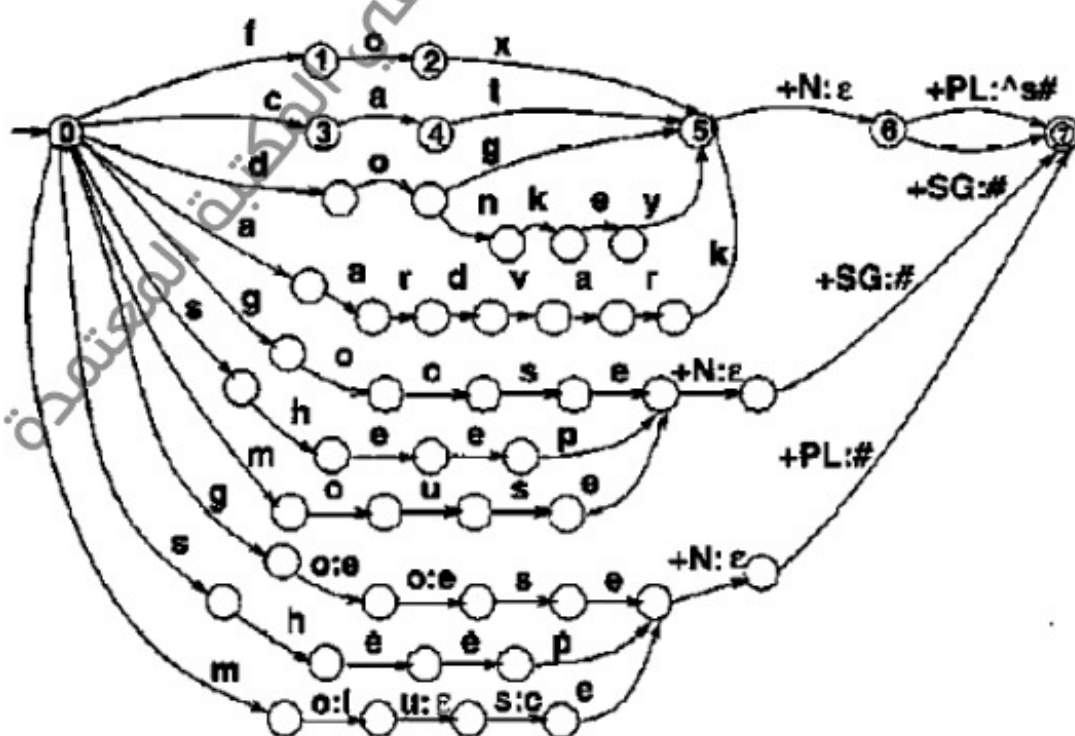
■ FST والغموض:

عند استخدام الأوتومات غير الحتمية المنتهية (ND-FSA) في عملية التعرف، قد تتعدد المسارات الممكنة للوصول إلى الحل النهائي. في هذه الحالة، لا يهم أي مسار يتم اختياره إذا كنا نقتصر على عملية التعرف فقط.

على سبيل المثال، يمكن تحليل كلمات مثل "Unionizable" باستخدام عدة مسارات داخل النموذج.

يمكن تحقيق معالجة القواعد الإملائية باستخدام محولات متعاقبة. في هذه العملية، يتم استخدام مخرجات FST كمدخلات لمحولات أخرى.

على سبيل المثال، لمعالجة الأخطاء الإملائية، يتم تصميم تسلسل من المحولات حيث يتم تصحيح الأخطاء تدريجياً وفقاً لقواعد محددة، مما يعزز دقة التحليل ويضمن توافق النص مع القواعد الإملائية.

المحول بين Intermediate و lexical

المحول بين Surface و Intermediate:

في التحليل اللغوي باستخدام FST، يتم الانتقال بين المستوي الوسيط (Intermediate) والمستوى السطحي (Surface) عبر تطبيق قواعد تحويل محددة. هذه القواعد تُعرّف العلاقة بين الشكل المجرد للكلمة والشكل المحقق في الاستخدام الفعلي، بناءً على السياق اللغوي.

• قاعدة تحويل تشومسكي:

وفقاً لقاعدة التحويل التي قدمها تشومسكي:

$$a \rightarrow b / c _ d$$

a تتحول إلى b فقط إذا كانت a موجودة بين c (السياق السابق) و d (السياق اللاحق).

هذه القاعدة تُستخدم لضمان أن التغيير يتم في السياقات التي تتطلب ذلك فقط.

مثال عملي على قاعدة تشومسكي:

إذا كان لدينا الحرف a ويظهر بين c و d، فإن القاعدة ستطبق كما يلي:

المدخل: cad

المخرج بعد التطبيق: cbd

• قاعدة إضافة الحرف e:

هذه القاعدة توضح كيفية إدخال الحرف e في مواقع معينة، بناءً على السياق:

$$\varepsilon \rightarrow e / \{x, s, z\} \wedge _ s\#$$

الرمز ε يشير إلى عدم وجود أي حرف في موضع معين.

يتم إدخال e فقط إذا كان الحرف السابق x أو s أو z، وكان الحرف التالي هو s، مع وجود الحرف s في نهاية الكلمة.

(#).

يضمن هذا النهج أن التغيير لا يتم إلا عند تحقق الشروط المحددة للسياق.

مثال عملي على قاعدة إضافة e:

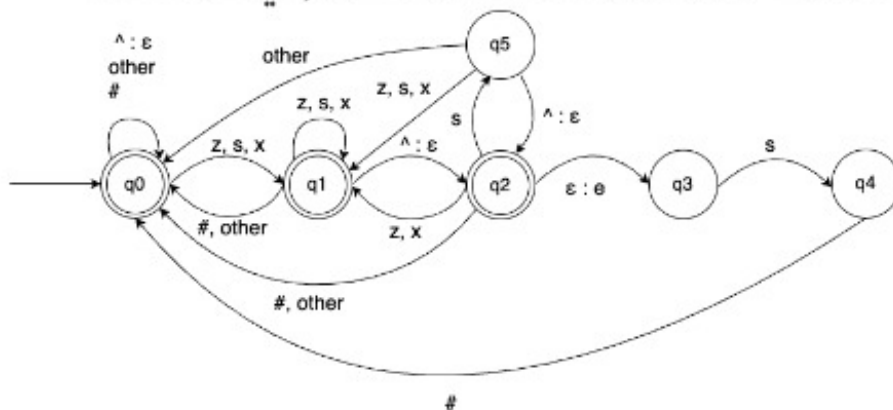
المدخل: xs.

السياق: الحرف السابق هو x، والحرف التالي هو s الموجود في نهاية الكلمة.

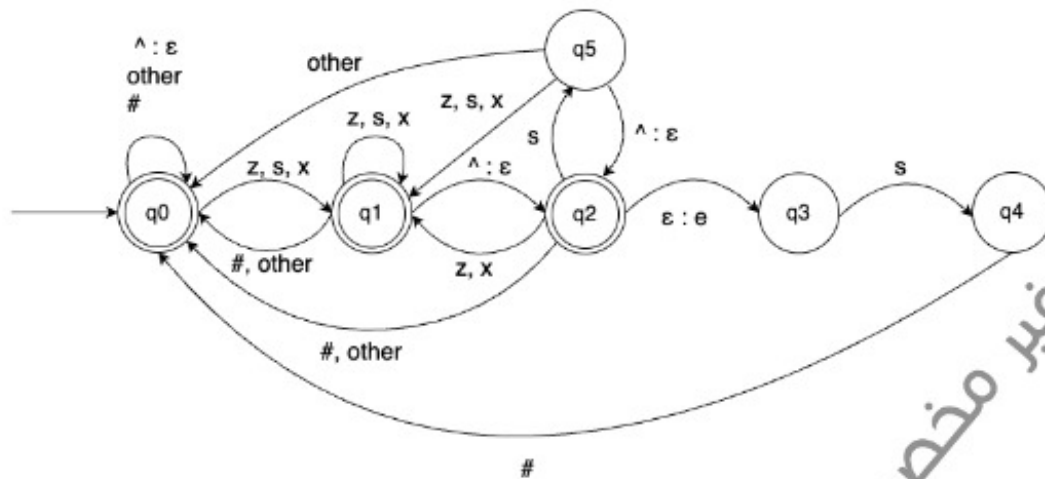
المخرج: xes.

المحول المتعلق بقاعدة إضافة e:

نحتاج إلى أوتومات لإضافة e بعد x, s, z وقبل s عند حدود الصرفيم في نهاية الكلمة.



جدول التحويلات لقاعدة إضافة e:



State/input	s: s	x: x	z: z	\wedge : ϵ	ϵ : e	#	other
q_0	1	1	1	0	—	0	0
q_1	1	1	1	2	—	0	0
q_2	5	1	1	0	3	0	0
q_3	4	—	—	—	—	—	—
q_4	—	—	—	—	—	0	—
q_5	1	1	1	2	—	—	0

The End.

Appreciation can make a day, even change a life. Your willingness to put it into words is all that is necessary.

