المحاضرة الأولى 13/ 10 /2016

- نظرية الاتومات و اللغات الصورية.

- Automata Theory and Formal Languages.

مقدمة: تعتبر الاتومات و اللغات الصورية من علوم الحاسوب النظرية و تدرس النمذجة الرياضية للحاسوب، و علاقة الاتومات باللغات الصورية و كيفية توليد اللغة و التعرف عليها و للمادة تطبيقات عديدة و هامة في مجال تصميم المترجمات (الموجودة في لغات البرمجة) و معالجة النصوص.

إذاً: الاتومات هو نمذجة رياضية للحاسوب و اللغات الصورية سميت كذلك لأننا نصورها و ليست طبيعية (لغة الـ Java ليست طبيعية).

مفاهيم أساسية:

الرمز : Symbol

تعريف: هو كائن غير قابل للتجزئة مثل:

الحروف اللاتينية a, b, ..., z ,

A, B, ..., Z

الحروف العربية أ, ب, ت, ...

الأرقام العربية 0, 1, ..., 9

الأبجدية : Alphabet

هي مجموعة منتهية و محددة و غير خالية من الرموز تتميز بأنه لا يمكن توليد أي رمز منها بواسطة بقية الرموز و نرمز لها عادةً بـ Σ .

مثال:

Alamal

/ /

الأبجدية في الدارات المنطقية $\Sigma = \{0, 1\}$

الأبجدية في اللغة العربية $\Sigma = \{$ أ, ب, ت, ..., ي $\}$

الأبجدية في اللغة الانكليزية $\Sigma = \{a, b, \dots, z, A, B, \dots, Z\}$

الأبجدية في النظام العشري $\Sigma = \{0, 1, \dots, 9\}$

السلسلة String :

هي تسلسل منتهي من رموز الأبجدية متوضعة بجانب بعضها دون فراغات أو فواصل ونرمز عادةً للسلاسل برموز صغيرة

$w, u, v, x, y, \dots$

مثال: لتكن لدينا الأبجدية $\Sigma = \{a, b\}$ فإنّ $w = abaaabbb$

هي سلسلة مولدة بالأبجدية Σ.

إنّ $x = a$ هي سلسلة ، $v = bbb$ سلسلة

$u = aaaaa$ أيضاً سلسلة مولدة بـ Σ

طول السلسلة Lenght of string :

هو عدد الرموز المشكلة للسلسلة ونرمز لطول السلسلة w بـ

$|w|$ فمثلاً: $w = aaabbab$ طولها $|w| = 7$ ،

السلسلة الفارغة Empty string :

هي السلسلة التي لا تحوي أية رموز وطولها يساوي الصفر ونرمز لها بـ ε .

$w = \varepsilon \Rightarrow |w| = 0$

/ /

ملاحظة: ε هي سلسلة معرفة على أية أبجدية، مثل المجموعة الفارغة.

العمليات على السلاسل:

التعاقب Concatenation :

تعاقب سلسلتين x و y هو سلسلة مشكلة من توضع رموز السلسلة الأولى x متبوعة مباشرة برموز السلسلة y و نرمز لعملية التعاقب بـ • و تحذف في معظم الأيام.

مثال: ليكن لدينا الأبجدية $\Sigma=\{0,1\}$ و لنأخذ السلسلتين $x=000$، $y=0100$. عندئذٍ:

$$x.y=xy=000\,0100$$

$$y.x=yx=0100\,000$$

ملاحظة: إنّ عملية التعاقب هي عملية غير تبديلية و لكنها تجميعية و العنصر الحيادي فيها هو ε.

$$y.\varepsilon=0100\varepsilon=0100=y$$

بادئة سلسلة Prefix :

مثال: $x=0110001$

تكون السلسلة u هي بادئة للسلسلة v إذا وجدت سلسلة w بحيث يتحقق لدينا مايلي $v=uw$

/ /

مثال: لتكن السلسلة v بحيث $v = ababbbaa$ عندئذٍ مجموعة بادئات هذه السلسلة هي

$\{\varepsilon, a, ab, aba, abab, ababb, ababbb, ababbba, ababbbaa\}$

لاحقة سلسلة : Suffix

تكون السلسلة u لاحقة للسلسلة v إذا وجدت سلسلة w بحيث يتحقق $v = wu$ ،

مثال: لتكن $v = ababbbaa$ عندئذٍ مجموعة لاحقات هذه السلسلة هي

$\{\varepsilon, a, aa, baa, bbaa, bbbaa, abbbaa, babbbaa, ababbbaa\}$

السلسلة الجزئية : Substring

تكون السلسلة u سلسلة جزئية من السلسلة v إذا كانت كل رموز السلسلة u موجودة في v مع المحافظة على ترتيب الرموز

مثال: في المثال السابق على سبيل المثال لدينا كل من السلاسل التالية

ε , b , a , baa , bb , ba

تعتبر سلسلة جزئية من السلسلة v

قوة أبجدية : Power of Alphabet

هي مجموعة كل السلاسل التي علينا تركيبها (أو تشكيلها) باستخدام

/ /

عملية التعاقب بين رموز الأبجدية Σ و نرمز لمجموعة السلاسل المولدة على الأبجدية Σ ذات الطول n بـ Σ^n

مثال: لتكن لدينا الأبجدية $\Sigma=\{0,1\}$

لدينا:

عبارة عن السلاسل المولدة من Σ و من الطول الصفر $\Sigma^0=\{\varepsilon\}$

$$\Sigma^1=\{0,1\}=\Sigma$$

$$\Sigma^2=\Sigma\Sigma=\{0,1\}\{0,1\}=\{00,01,10,11\}$$

$$\Sigma^3=\Sigma\Sigma\Sigma=\Sigma^2\Sigma=\Sigma\Sigma^2$$

$$=\{0,1\}\{00,01,10,11\}$$

$$=\{000,001,010,011,100,101,110,111\}$$

وهكذا ...

نعرف Σ^* على أنها مجموعة جميع السلاسل التي يمكن توليدها من رموز الأبجدية Σ أي

$$\Sigma^*=\Sigma^0\cup\Sigma^1\cup\Sigma^2\cup\ldots$$

وهي مجموعة غير منتهية و تحوي ε دوماً.

ملاحظة: / أحياناً يلزمنا تعريف كل السلاسل المولدة من الأبجدية Σ عدا السلسلة الفارغة عندئذ نرمز لذلك بـ Σ^+ و

$$\Sigma^+=\Sigma^*-\{\varepsilon\}=\Sigma^*-\Sigma^0$$

ومنه $\Sigma^*=\Sigma^+\cup\Sigma^0$

/ /

اللغة : Language

هي مجموعة السلاسل المختارة من المجموعة Σ^* والمولدة من الأبجدية Σ ونرمز لها عادةً بـ L أي أنه إذا كانت Σ أبجدية و كانت L لغة مولدة (مبنية) من Σ فإن L ستكون محتواة بـ Σ^* دائماً

$L \subseteq \Sigma^*$

/ملاحظة : / أي أن أي لغة معرفة على أبجدية ما لا تحوي بالضرورة جميع السلاسل المولدة بالأبجدية .

مثال: إذا كانت $\Sigma = \{a, b\}$ أبجدية و كانت $L \subseteq \Sigma^*$ لغة مولدة من الأبجدية Σ وتحوي التعاقب aa

$\Sigma^* = \{\epsilon, a, b, ab, ba, aab, aba, \ldots\}$

$L \subseteq \Sigma^* = \{aab, aa, baa, aaa, baaa, babaabb, \ldots\}$

وفي هذه الحالة أصبح لدينا

aab كلمة

aa كلمة

⋮

انتهت .