الاتومات واللغات الصورية

المحاضرة الثالثة — الثلاثاء 17/10/2017

مثال عن اللغات: لتكن $\Sigma = \{a, b, c\}$

* L_1 هي اللغة المكونة من كل الكلمات المولدة من Σ والتي تحوي c مرتين على الأقل «ليس شرط التعاقب»

$L_1 = \{cc, cabc, abcc, cacb, accc, cccc, cbcccba, \dots\}$

$bbabbcb \notin L_1$, $abc \notin L_1$, $aa \notin L_1$, $\epsilon \notin L_1$

* L_2 هي اللغة المكونة من كل الكلمات المولدة من Σ والتي تحوي c مرتين فقط

$L_2 = \{cc, cabc, abcc, cacb, acc, bcbaca, \dots\}$

$acbcbac \notin L_2, abbc \notin L_2, ccc \notin L_2, cccc \notin L_2, \epsilon \notin L_2$

* L_3 هي اللغة المكونة من كل الكلمات المولدة من Σ والتي تبدأ بـ a

$L_3 = \{a, aba, abbcaa, aabccacbb, \dots\}$

$cabbaa \notin L_3, babcbb \notin L_3, \epsilon \notin L_3$

العمليات على اللغات: set operations

1. اتحاد لغتين: Union
L_1 و L_2 يمثل على الشكل التالي $L_1 \cup L_2$ وهو عبارة عن اللغة التي تحوي جميع السلاسل الموجودة في L_1 أو L_2 أو كلاهما

$L_1 \cup L_2 = \{x / x \in L_1 \text{ OR } x \in L_2\}$

2. تقاطع لغتين: Intersection
L_1 و L_2 يمثل على الشكل التالي $L_1 \cap L_2$ وهو عبارة عن اللغة التي تحوي جميع السلاسل الموجودة في L_1 و L_2 معاً

$L_1 \cap L_2 = \{x / x \in L_1 \text{ and } x \in L_2\}$

3. فرق لغتين: Difference
L_1 و L_2 يمثل على الشكل التالي $L_1 - L_2$ وهو عبارة عن اللغة التي تحوي السلاسل الموجودة في L_1 وغير موجودة في L_2

$L_1 - L_2 = \{x / x \in L_1 \;\&\; x \notin L_2\}$

4. متمم لغة: Complement
L معرفة على الأبجدية Σ وهو عبارة عن اللغة التي تحوي جميع السلاسل الموجودة في Σ^* وغير موجودة في L ويرمز لمتمم اللغة بالرمز $\overline{L}$

$\overline{L} = \{x / x \in \Sigma^* \;\&\; x \notin L\}$

مثال: لتكن لدينا اللغتين L_1 و L_2 حيث:

$L_1 = \{a, ab, aaaa\}$

$L_2 = \{bb, ab\}$

* $L_1 \cap L_2 = \{ab\}$
* $L_1 \cup L_2 = \{a, ab, aaaa, bb\}$
* $L_1 - L_2 = \{a, aaaa\}$
* $L_2 - L_1 = \{bb\}$
* $\overline{L_1} = \Sigma^* - L_1$

  $= \{\epsilon, aa, aaa, abb, bb, bbba, \dots\}$
* $\overline{L_2} = \Sigma^* - \{bb, ab\}$

  $= \{\epsilon, a, aba, ba, aaaa, bbb, \dots\}$

5- تعاقب لغتين: Concatenation L_1 و L_2 نرمز له ب L_1L_2 وهو اللغة المكونة من السلاسل المشكلة من تعاقب سلاسل L_1 متبوعة ب L_2

$L_1L_2 = \{xy \,/\, x \in L_1 \;\&\; y \in L_2\}$

مثال:

$L_1 = \{a, ab, aaaa\}$

$L_2 = \{bb, ab\}$

$L_1L_2 = \{abb, aab, abbb, abab, aaaabb, aaaaab\}$

$= \{ab^2, a^2b, ab^3, (ab)^2, a^4b^2, a^5b\}$

$L_2L_1 = \{bb, ab\}\{a, ab, aaaa\}$

$= \{bba, bbab, bbaaaa, aba, abab, abaaaa\}$

$= \{b^2a, b^2ab, b^2a^4, aba, (ab)^2, aba^4\}$

6- انغلاق لغة: kleene star, kleene closure إذا كانت L لغة فإن انغلاق اللغة L نرمز له ب L^* وهي اللغة المشكلة من مجموعة كل السلاسل الممكنة الناتجة من تعاقب السلاسل في L

$L^* = L^0 \cup L^1 \cup L^2 \cup L^3 \cup L^4 \cup \dots$

$L^+ = L^* - \{\epsilon\}$

$L^0 = \{\epsilon\}, L^2 = LL, L^3 = LLL = L^2L = LL^2$

مثال: $L=\{bb, ab\}$

$L^0=\{\epsilon\}$, $L^1=\{bb, ab\}$

$L^2 = LL = \{bb,ab\}\{bb,ab\} = \{b^4, b^2ab, ab^3, (ab)^2\}$

$L^3 = LLL = L^2L = LL^2 = \{bb,ab\}\{b^4, b^2ab, ab^3, (ab)^2\}$

$= \{b^6, b^4ab, b^2ab^3, b^2(ab)^2, ab^5, ab^3ab, abab^3, (ab)^3\}$

$(ab)^3 \neq a^3b^3$, $(ab)^3 = ababab$

$L^* = \{\epsilon, b^2, ab, b^4, b^2ab, \dots\} = L^0 \cup L^1 \cup L^2 \cup L^3 \cup L^4 \cup \dots$

$\phi^0 = \phi^* = \{\epsilon\}$

التعابير المنتظمة

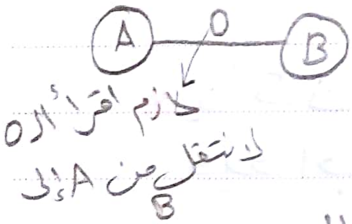
Regular Expressions

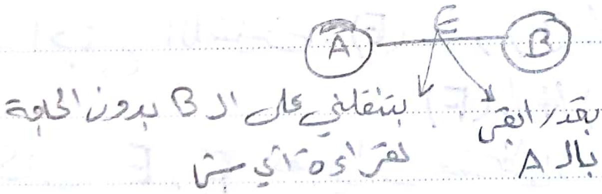
يوجد مجموعة من اللغات سنتعرف عليها لاحقاً ولكن سنبدأ بوصف اللغات المنتظمة.

في صف اللغات المنتظمة نستطيع أن نعبر عن اللغة المنتظمة بتعبير منتظم أو اتومات منتهي أو بقواعد منتظمة.

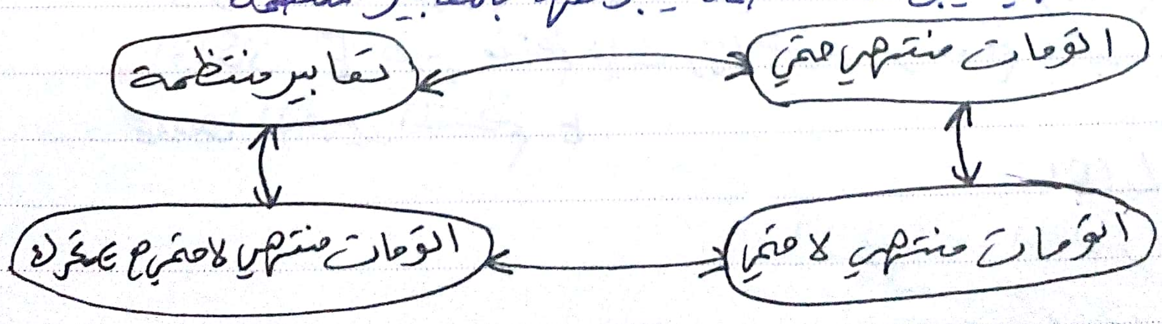
يوجد عدة أنواع من الاتومات المنتهي والتي جميعها متكافئة

1- اتومات منتهي حتمي
2- اتومات منتهي لا حتمي
3- اتومات منتهي لا حتمي مع ϵ - تحرك





حيث نستطيع تحويل كل اتومات منتهي لا حتمي مع ϵ - تحرك إلى اتومات منتهي لا حتمي ونستطيع تحويل كل اتومات منتهي لا حتمي إلى اتومات منتهي حتمي وبالتالي كل الاتومات المنتهية السابقة متكافئة
وكل اتومات منتهي يقبل لغة منتظمة يعبر عنها بالتعابير المنتظمة



* نستطيع تعريف اللغة المنتظمة على أبجدية Σ كما يلي:

$\{\epsilon\}$ لغة منتظمة على Σ

إذا كان a حرف من حروف الأبجدية Σ عندئذٍ تكون $\{a\}$ لغة منتظمة على Σ

إذا كانت L لغة منتظمة على Σ فإن L^* و L^n لغات منتظمة على Σ

إذا كانت كل من L_1 و L_2 لغات منتظمة فإن $L_1 \cup L_2$ ، $L_1 \cap L_2$ ، $L_1 L_2$ هي لغات منتظمة على Σ

لا يوجد لغات منتظمة أخرى

التعابير المنتظمة: تستخدم التعابير المنتظمة لتعريف اللغات المنتظمة

لتكن Σ أبجدية من الرموز عندئذٍ:

* ϕ هو التعبير المنتظم الذي يحدد اللغة الفارغة

اللغة التي يحددها (يعرفها) التعبير المنتظم ϕ

$$\underbrace{L(\underset{\text{تعبير منتظم}}{\phi})} = \phi$$

* ϵ هو التعبير المنتظم الذي يحدد اللغة التي تحوي السلسلة الفارغة

$$L(\epsilon) = \{\epsilon\}$$

* من أجل أي رمز a من الأبجدية Σ فإن a هو تعبير منتظم يحدد (يعين) اللغة

$$L(a) = \{a\}$$

* إذا كان E, F تعبيران منتظمان فإن التعبير $E+F$ هو تعبير منتظم يحدد اجتماع اللغتين: $L(E)$ ، $L(F)$

$$L(E+F) = L(E) \cup L(F)$$

* إذا كان E, F تعبيران منتظمان فإن التعبير EF هو تعبير منتظم يحدد تعاقب اللغتين: $L(E)$ ، $L(F)$

$$L(EF) = L(E)L(F)$$

* إذا كان E تعبير منتظم فإن E^* هو تعبير منتظم (تكرار لـ E من الصفر إلى تكرارات غير منتهية) واللغة التي يحددها

$$L(E^*) = (L(E))^*$$

* إذا كان E تعبير منتظم فإن (E) هو تعبير منتظم ويحدد نفس اللغة التي يحددها التعبير المنتظم E

$$L((E)) = L(E)$$

انتهت المحاضرة