

## المتتابعات الماركوفية مع التطبيق في المجال اللغوي

أ.د. باسل يونس ذنون الخياط  
كلية علوم الحاسبات والرياضيات  
جامعة الموصل

### Abstract

This research deals with a new concept called the Markov sequence. A Markov sequence is a series of symbols or letters or words having the Markov property. An other concept called the likelihood sequence is suggested . The likelihood sequence is the greatest probabilistic bond between the elements of a Markov sequence. After giving the mathematical definition of Markov sequence, two linguistic applications are considered in the field of Arabic language. Decision trees are driven from the element of Markov sequences are explained.

### المستخلص:

يتناول هذا البحث مفهوماً جديداً أطلق عليه تسمية المتتابعات الماركوفية Markovian Sequences. والمتتابعة الماركوفية هي عبارة عن سلسلة من الرموز أو الحروف أو الكلمات تتمتع بالخاصية الماركوفية. كما يُقترح مفهوماً آخر أطلق عليه المتتابعات المرجحة، والتي تمثل أقوى آصرة احتمالية موجودة بين عناصر المتتابعات الماركوفية. وبعد إعطاء التعريف الرياضي للمتتابعات الماركوفية وبعض خصائصها الرياضية يُقدم تطبيقان في مجال اللغة العربية، أحدهما يتعامل مع الحروف ويتناول كلمة "السلام"، في حين إن الآخر يتعامل مع الكلمات ويتناول البسمة "بسم الله الرحمن الرحيم". وتوضح شجرة القرار المشتقة من عناصر المتتابعات الماركوفية قيد الدراسة.

## 1- المقدمة:

لم يعد خافيا الأهمية الكبيرة للأفكار والمضامين البنيوية التي تتمتع بها السلاسل الماركوفية باعتبارها نموذجا تصادفيا كفؤا لتمثيل الظواهر والعمليات المختلفة ، انظر على سبيل المثال الخياط [2009a] . ونتيجة للاهتمامات المتزايدة بمضامين السلاسل الماركوفية فقد انبثق في ستينيات القرن العشرين أسلوبا رياضيا خاصا، مبني أساسا على فكرة الخاصية الماركوفية، هو أسلوب فضاء الحالة State Space Approach ، انظر [1960] Kalman. وفي سبعينيات القرن العشرين تم إجراء تطورا آخر على السلاسل الماركوفية هو نموذج ماركوف المخفي Hidden Markov Model، انظر [1989] Rabiner. وتم تطبيق النماذج الماركوفية في المجالات التقنية والحياتية والإدارية ومجالات متعددة أخرى. وفي المجال الديني استخدم الخياط والعجوز [2009] النموذج الماركوفي لنمذجة كلمة التوحيد. كذلك استخدم الخياط [2009b] النموذج الماركوفي لنمذجة عدد الركعات في الصلوات الخمس.

من السمات الخاصة التي يتمتع بها النموذج الماركوفي كون فضاء الحالة له يضم حالات States ، وهذه الحالات ممكن ان تكون أي صفة اسمية أو رتبية أو رمز أو حرف أو كلمة أو أي شيء آخر. أما ألسمه الأساسية الأخرى التي يتمتع بها النموذج الماركوفي فهي الخاصية الماركوفية The Markovian Property. وبايجاز فان المقصود بالخاصية الماركوفية هو ان الحالة التي تحدث في زمن ما ( أو موقع ما) تعتمد في حدوثها فقط على الحالة في الزمن (أو الموقع) الذي يسبقها مباشرة وبغض النظر عن ما كان كائن في الأزمنة (أو المواقع) السابقة الأخرى، الخياط [2009a] .

## 2-المتتابعات الماركوفية Markovian Sequences:

بافتراض ان فضاء الحالة يضم عددا محدودا من الحالات  $S = \{s_1, s_2, \dots, s_m\}$ . يرتكز الاهتمام عندما تمثل الحالات رموزا أو حروفا أو كلمات معينة إضافة إلى معناها المتمثل بكونها حالات معينة. ففي مجال اللغة العربية وباعتبار ان فضاء الحالة يكون مجموعة حروف اللغة العربية، فان المجموعة S تمثل مجموعة جميع حروف الأبجدية العربية (28 حرفا)، أي ان  $S = \{أ، ب، ت، ...، ي\}$ . وباعتبار ان فضاء الحالة يكون مجموعة حروف العلة، فان  $S = \{أ، و، ي\}$ . وباعتبار ان فضاء الحالة يكون مجموعة حروف الجر فان

$S = \{من، إلى، عن، على، في، ل، ب، حتى، ك، خلا، حاشا، عدا، مذ، منذ، رب، كي، واو القسم، تا القسم، متى\}$ . كذلك باعتبار ان فضاء الحالة يكون مجموعة كلمات البسمة فان  $S = \{بسم، الله، الرحمن، الرحيم\}$ .

وعلى فرض ان  $p_{s_i, s_j}$  تمثل الاحتمالية الانتقالية من الحالة  $s_i$  إلى الحالة  $s_j$  بخطوة واحدة. ولو كانت الحالات التي لدينا طاقية (طوريه) Ergodic باحتمالات مترنة  $p_{s_i}$ ،  $i = 1, 2, \dots, m$ . فان ما يهم الآن حساب احتمالية حدوث متتابعة معينة من هذه الحالات وفق تسلسل معين، ولتكن هذه المتتابعة  $O = (o_1 o_2 \dots o_k)$ ، اذ ان عناصر المتتابعة  $O$  هي جميع أو بعض عناصر فضاء الحالة  $S$ ، بحيث يكون العنصر  $o_i$  في الموقع  $i$  في المتتابعة  $O$ .

#### تعريف:

يقال للمتتابعة  $O = (o_1 o_2 \dots o_k)$  بأنها متتابعة ماركوفية إذا كانت تتحقق في عناصرها الخاصية الماركوفية. أي ان احتمالية حدوث عنصر معين فيها تعتمد فقط على العنصر الذي يسبقه.

#### تعريف:

يقال للمتتابعة الماركوفية بأنها ذات حجم  $k$  إذا كانت تضم  $k$  من الرموز (العناصر).

لذا فان المتتابعة  $(abc)$  هي ذات حجم  $k=3$ ، والمتتابعة  $(aabcc)$  هي ذات حجم  $k=5$ .

#### المبرهنة(1):

إذا كان فضاء حالة يضم  $m$  من العناصر المختلفة، فان عدد المتتابعات الماركوفية ذات العناصر المختلفة والتي كل منها بحجم  $k$  والتي يمكن تكوينها من عناصر فضاء حالة هو:  
 أ-  $m^k$  عندما يكون هناك انتقال من عنصر إلى نفسه (أي التكرار بالعناصر مسموح).  
 ب-  $m(m-1)^{k-1}$  عندما لا يكون هناك انتقال من عنصر إلى نفسه ويسمح بتكرار العنصر في باقي المواقع.

ت-  $\prod_{i=1}^k (m-i+1) = m(m-1)(m-2)\dots(m-k+1)$  عندما لا يكون هناك انتقال من عنصر إلى نفسه ولا يسمح بتكرار العنصر .

المثال(1):

لديك فضاء الحالة  $S=\{a,b,c,d\}$ . كونّ المتتابعات الماركوفية الممكنة ذوات الحجم  $k=2,3$  وبحيث يكون:

أ- التكرار بالعناصر مسموح.

ب- لا يكون هناك انتقال من عنصر إلى نفسه ويسمح بتكرار العنصر في باقي المواقع.

ت- لا يكون هناك انتقال من عنصر إلى نفسه ولا يسمح بتكرار العنصر في باقي المواقع.

الحل:

عندما  $k=2$ :

أ- يمكن تكوين  $m^k = 4^2 = 16$  متتابعة ماركوفية وذلك بعمل الجدول الآتي:

|   | a  | b  | c  | d  |
|---|----|----|----|----|
| a | aa | ab | ac | ad |
| b | ba | bb | bc | bd |
| c | ca | cb | cc | cd |
| d | da | db | dc | dd |

فتكون المتتابعات الممكنة في هذه الحالة هي العناصر الناتجة من الجدول السابق، وهي:

$$O_{\underline{1}} = (aa); O_{\underline{2}} = (ab); O_{\underline{3}} = (ac); O_{\underline{4}} = (ad);$$

$$O_{\underline{5}} = (ba); O_{\underline{6}} = (bb); O_{\underline{7}} = (bc); O_{\underline{8}} = (bd);$$

$$O_{\underline{9}} = (ca); O_{\underline{10}} = (cb); O_{\underline{11}} = (cc); O_{\underline{12}} = (cd);$$

$$O_{\underline{13}} = (da); O_{\underline{14}} = (db); O_{\underline{15}} = (dc); O_{\underline{16}} = (dd).$$

ب- يمكن تكوين  $m(m-1)^{k-1} = 4(3)^1 = 12$  متتابعة ماركوفية في هذه الحالة وهي:

$$O_{\underline{1}} = (ab); O_{\underline{2}} = (ac); O_{\underline{3}} = (ad);$$

$$O_{\underline{4}} = (ba); O_{\underline{5}} = (bc); O_{\underline{6}} = (bd);$$

$$O_{\underline{7}} = (ca); O_{\underline{8}} = (cb); O_{\underline{9}} = (cd);$$

$$O_{\underline{10}} = (da); O_{\underline{11}} = (db); O_{\underline{12}} = (dc).$$

ت- يمكن تكوين  $\prod_{i=1}^2 (4-i+1) = 4(3) = 12$  متتابعة ماركوفية في هذه الحالة وهي نفس

المتتابعات التي تم الحصول عليها في ب.

عندما  $k=3$ :

أ- يمكن تكوين  $m^k = 4^3 = 64$  متتابعة ماركوفية في هذه الحالة وذلك بعمل الجدول آخر، واعتمادا على الجدول السابق وكما يأتي:

|    | a   | b   | c   | d   |
|----|-----|-----|-----|-----|
| aa | aaa | aab | aac | aad |
| ab | aba | abb | abc | abd |
| ac | aca | acb | acc | acd |
| ad | ada | adb | adc | add |
| ba | baa | bab | bac | bad |
| bb | bba | bbb | bbc | bbd |
| bc | bca | bcb | bcc | bcd |
| bd | bda | bdb | bdc | bdd |
| ca | caa | cab | cac | cad |
| cb | cba | cbb | cbc | cbd |
| cc | cca | ccb | ccc | ccd |
| cd | cda | cdb | cdc | cdd |
| da | daa | dab | dac | dad |
| db | dba | dbb | dbc | dbd |
| dc | dca | dcb | dcc | dcd |
| dd | dda | ddb | ddc | ddd |

ويمكن إيجاد هذه المتتابعات من الجدول مباشرة:

$$O_1 = (aaa); O_2 = (aab); O_3 = (aac); O_4 = (aad); \dots; O_{64} = (ddd).$$

ب- يمكن تكوين  $m(m-1)^{k-1} = 4(3)^{3-1} = 36$  متتابعة ماركوفية في هذه الحالة وهي موضحة بالجدول الآتي:

|    | a   | b   | c   | d   |
|----|-----|-----|-----|-----|
| ab | aba |     | abc | abd |
| ac | aca | acb |     | acd |
| ad | ada | adb | adc |     |
| ba |     | bab | bac | bad |
| bc | bca | bcb |     | bcd |
| bd | bda | bdb | bdc |     |
| ca |     | cab | cac | cad |
| cb | cba |     | cbc | cbd |
| cd | cda | cdb | cdc |     |
| da |     | dab | dac | dad |
| db | dba |     | dbc | dbd |
| dc | dca | dcb |     | dcd |

ت- يمكن تكوين  $\prod_{i=1}^3 (4-i+1) = 4(3)(2) = 24$  متتابعة ماركوفية في هذه الحالة وهي

موضحة بالجدول الآتي:

|    | a   | b   | c   | d   |
|----|-----|-----|-----|-----|
| ab |     |     | abc | abd |
| ac |     | acb |     | acd |
| ad |     | adb | adc |     |
| ba |     |     | bac | bad |
| bc | bca |     |     | bcd |
| bd | bda |     | bdc |     |
| ca |     | cab |     | cad |
| cb | cba |     |     | cbd |
| cd | cda | cdb |     |     |
| da |     | dab | dac |     |
| db | dba |     | dbc |     |
| dc | dca | dcb |     |     |

ملاحظة:

في حالة وجود عناصر مكررة ومنتالية في المتتابعة الماركوفية فإنها تعامل معاملة عنصر واحد. فلو كان  $S=\{a,b,c\}$  وأردنا تكوين متتابعات ماركوفية ذات حجم 5 وبحيث يكون كل من الحرفين a و c مكرر ومنتالي، فان عدد المتتابعات الماركوفية الممكنة يحسب باعتبار ان  $m=3$  و  $k=3$ ، وهذا العدد هو:

$$\prod_{i=1}^3 (3-i+1) = 3(2)(1) = 6.$$

وهذه المتتابعات هي:

$$O_{-1} = (aabcc); O_{-2} = (aaccb); O_{-3} = (baacc); O_{-4} = (bccaa); O_{-5} = (ccaab); O_{-6} = (ccbaa).$$

المبرهنة (2):

احتمالية حدوث المتتابعة  $O = (o_1 o_2 \dots o_k)$  هي:

$$P(O) = P(o_1 o_2 o_3 \dots o_{k-1} o_k) \\ = P(o_1) P(o_2 / o_1) P(o_3 / o_2) \dots P(o_k / o_{k-1}).$$

البرهان:

$$\begin{aligned} P(\underline{O}) &= P(o_1 o_2 o_3 \dots o_{k-1} o_k) \\ &= P(o_k / o_1 o_2 \dots o_{k-1}) P(o_1 o_2 \dots o_{k-1}) \end{aligned}$$

حسب تعريف الاحتمالية الشرطية. وبلاستفادة من الخاصية الماركوفية تصبح

$$P(o_k / o_1 o_2 \dots o_{k-1}) = P(o_k / o_{k-1}).$$

وبإعادة نفس المعالجة على  $P(o_1 o_2 \dots o_{k-1})$  نجد ان:

$$P(o_1 o_2 \dots o_{k-1}) = P(o_{k-1} / o_{k-2}) P(o_1 o_2 \dots o_{k-2}).$$

وهكذا بتكرار المعالجة السابقة نحصل على ما يأتي:

$$\begin{aligned} P(\underline{O}) &= P(o_1 o_2 o_3 \dots o_{k-1} o_k) \\ &= P(o_k / o_{k-1}) P(o_{k-1} / o_{k-2}) \dots P(o_2 / o_1) P(o_1). \end{aligned}$$

ملاحظة:

عند عدم توافر قيمة  $P(o_1)$  فيمكن الاستعاضة عنها بالاحتمالية المتزنة المقابلة لها.

النتيجة (1):

إذا كانت الاحتمالية الانتقالية بين أي عنصرين في المتتالية الماركوفية صفراً، فإن احتمالية حدوث تلك المتتالية يكون صفراً أيضاً.

النتيجة (2):

احتمالية حدوث المتتالية  $\underline{O} = (o_1 o_2 \dots o_k)$  معطى بان الحالة  $o_1$  قد حدثت فعلا هي:

$$\begin{aligned} P(\underline{O} / o_1) &= P(o_1 o_2 o_3 \dots o_{k-1} o_k / o_1) \\ &= P(o_2 / o_1) P(o_3 / o_2) \dots P(o_k / o_{k-1}). \end{aligned}$$

المثال (2):

إذا كان فضاء الحالة يضم ثلاثة عناصر وكما يأتي،  $S=\{a,b,c\}$ . وكانت المصفوفة الانتقالية على النحو الآتي:

$$P = \begin{pmatrix} 0 & 0.4 & 0.6 \\ 0.2 & 0 & 0.8 \\ 0.3 & 0.7 & 0 \end{pmatrix}$$

كونّ المتتابعات الماركوفية الممكنة ذوات الحجم  $k=2,3$  واحسب احتمالية حدوث كل منها.

الحل:

عندما  $m=3$  و  $k=2$  فإن عدد المتتابعات الماركوفية الممكنة هو  $m^k = 3^2 = 9$  وهذه المتتابعات هي:

$$O_1 = (aa); O_2 = (ab); O_3 = (ac);$$

$$O_4 = (ba); O_5 = (bb); O_6 = (bc);$$

$$O_7 = (ca); O_8 = (cb); O_9 = (cc).$$

واحتمالية حدوث هذه المتتابعات يتم إيجادها من المصفوفة الانتقالية مباشرة وهي كالآتي:

$$P(O_{-1}) = 0; P(O_{-2}) = 0.4; P(O_{-3}) = 0.6;$$

$$P(O_{-4}) = 0.2; P(O_{-5}) = 0; P(O_{-6}) = 0.8;$$

$$P(O_{-7}) = 0.3; P(O_{-8}) = 0.7; P(O_{-9}) = 0.$$

أما عندما  $m=3$  و  $k=3$  فإن عدد المتتابعات الماركوفية الممكنة هو  $m^k = 3^3 = 27$  وما دام لا يوجد انتقال من العنصر إلى نفسه بخطوة واحدة، فيكون عدد المتتابعات الماركوفية الممكنة هو  $m(m-1)^{k-1} = 3(3-1)^2 = 3(4) = 12$  وهذه المتتابعات هي:

$$O_1 = (aba); O_2 = (abc); O_3 = (aca); O_4 = (acb);$$

$$O_5 = (bab); O_6 = (bac); O_7 = (bca); O_8 = (bcb);$$

$$O_9 = (cab); O_{10} = (cbc); O_{11} = (cba); O_{12} = (cac);$$

ولحساب احتمالات حدوث هذه المتتابعات، يتم إيجاد الاحتمالات المتزنة أولاً وهي كالآتي

$$p_c = 0.4220 \text{ و } p_b = 0.3761 \text{ و } p_a = 0.2018$$



$$P(O_1) = P(a)P(b/a)P(a/b) = 0.2018(0.4)(0.2) = 0.0338,$$

$$P(O_2) = P(a)P(b/a)P(c/b) = 0.2018(0.4)(0.8) = 0.0646,$$

$$P(O_3) = P(a)P(c/a)P(a/c) = 0.2018(0.6)(0.3) = 0.0363,$$

$$P(O_4) = P(a)P(c/a)P(b/c) = 0.2018(0.6)(0.7) = 0.0848,$$

وبنفس الطريقة يمكن حساب احتمالات حدوث باقي المتتابعات الماركوفية وهي كالآتي:

$$P(O_5) = 0.0301; P(O_6) = 0.0451; P(O_7) = 0.0903; P(O_8) = 0.2106;$$

$$P(O_9) = 0.0506; P(O_{10}) = 0.2363; P(O_{11}) = 0.0591; P(O_{12}) = 0.0760.$$

تعريف:

يقال للمتتابعة  $O^*$  بأنها المتتابعة المرّجحة **Likelihood Sequence** إذا كان لها أكبر احتمالية من بين أقرانها  $O_{-1}, O_{-2}, \dots$ . أي أن  $O^*$  تحقق  $P(O^*) = \max_i \{P(O_{-1}), P(O_{-2}), \dots\}$ .

وبالعودة إلى المثال الأخير عندما  $k=2$  فإن المتتابعة المرّجحة هي  $O_6 = (bc)$  وذلك لكون احتماليّتها هي الأعلى بين أقرانها، 80%. أما عندما  $k=3$  فإن المتتابعة المرّجحة هي  $O_{10} = (cbc)$  وذلك لكون احتماليّتها هي الأعلى بين أقرانها،  $P(O_{10}) = 0.2363$ . ان المتتابعة المرّجحة تمثل أقوى أصرة احتمالية موجودة بين الرموز.

### 3- التطبيق الأول ( كلمة "السلام"):

تتألف كلمة "السلام" من أربعة حروف أساسية هي الألف واللام والسين والميم. أي ان فضاء الحالة للحروف التي تتكون منها كلمة "السلام" هي المجموعة  $S = \{أ، ل، س، م\}$ . ويمكن ببسر حساب مشاهدات الانتقالات بين الحروف وهي مبيّنة في مصفوفة التكرارات الآتية، انظر الخياط [2009a]:

$$F = \begin{matrix} & \begin{matrix} أ & ل & س & م \end{matrix} \\ \begin{matrix} أ \\ ل \\ س \\ م \end{matrix} & \begin{pmatrix} - & 1 & - & 1 \\ 1 & - & 1 & - \\ - & 1 & - & - \\ 1 & - & - & - \end{pmatrix} \end{matrix}$$

ومن مصفوفة التكرارات يمكن إيجاد المصفوفة الانتقالية بين حروف كلمة "السلام" وهي كما مبيّنة في أدناه:

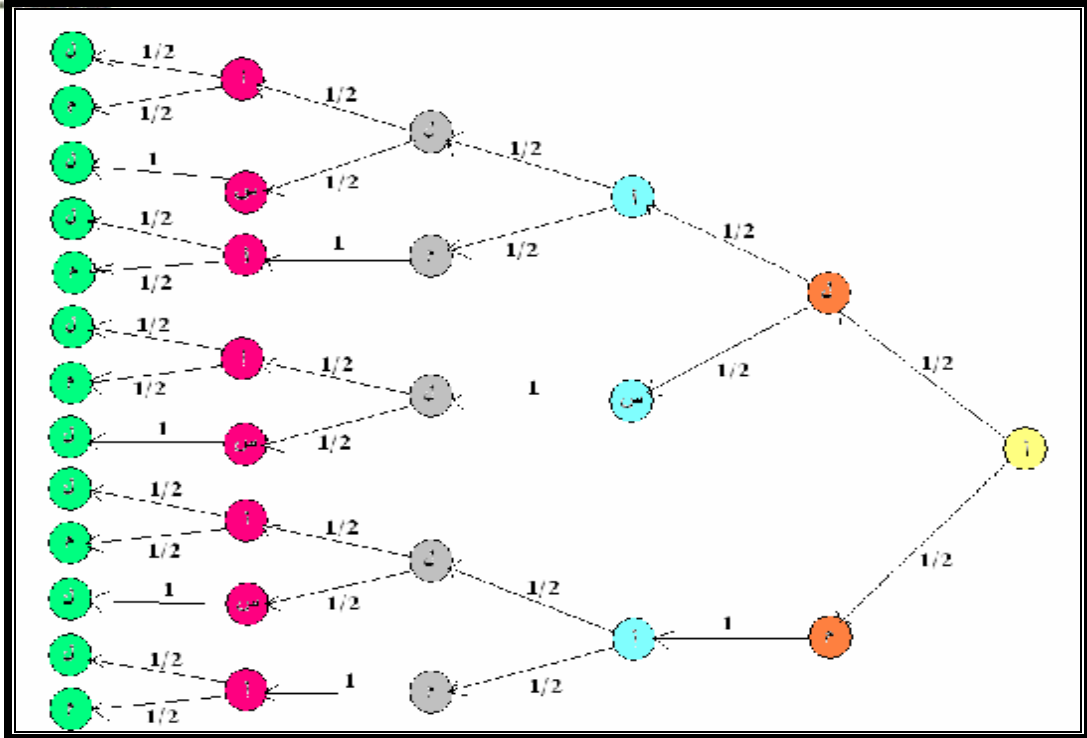
$$P = \begin{matrix} & \begin{matrix} أ & ل & س & م \end{matrix} \\ \begin{matrix} أ \\ ل \\ س \\ م \end{matrix} & \begin{pmatrix} - & \frac{1}{2} & - & \frac{1}{2} \\ \frac{1}{2} & - & \frac{1}{2} & - \\ - & 1 & - & - \\ 1 & - & - & - \end{pmatrix} \end{matrix}$$

ومن خلال دراسة تحليلية للحروف الأساسية الأربعة لكلمة "السلام" باعتبارها حالات States، تبين بأنها حالات طاقة (طوريه) Ergodic وتوزيعها الاحتمالي المتزن موجود هو كما مبيّن في أدناه:

$$p_A = 0.4444; p_L = 0.2222; p_S = 0.1111; p_M = 0.2222.$$

ولغرض دراسة المتتابعات الماركوفية الممكن تكوينها من حروف كلمة "السلام"، يُؤخذ كل حرف من حروفها الأربعة ويعتبر جذرا Root يُبدأ به تكوين المخطط الشجري Tree Diagram ولخمسة أجيال متعاقبة من الحروف الممكنة في فضاء كلمة "السلام" ووفق المصفوفة الانتقالية السابقة.

أولاً: - بالبدء بالحرف "ألف" باعتباره جذرا سيتم الحصول على المخطط الشجري المبيّن في الشكل الآتي.



الشكل (1): المخطط الشجري لخمس أجيال ناجمة من الجذر "ألف" في فضاء كلمة "السلام".  
من المخطط السابق يلاحظ الآتي:

1- هناك كلمتان يمكن تكوينهما كل منهما مؤلفة من حرفين فقط، وهما كلمتي "ال" و "ام"،  
وكل منهما باحتمالية 50%.

2- الكلمات الممكن تكوينها والمؤلفة كل منها من ثلاثة حروف مبيّنة في الجدول الآتي هي  
واحتمالاتها.

| الكلمة   | الا | السا | اما | المجموع |
|----------|-----|------|-----|---------|
| احتمالها | 1/4 | 1/4  | 1/2 | 1       |

3- الكلمات الممكن تكوينها والمؤلفة كل منها من أربعة حروف مبيّنة في الجدول الآتي هي  
واحتمالاتها.

| الكلمة   | الال | الام | السل | امال | امام | المجموع |
|----------|------|------|------|------|------|---------|
| احتمالها | 1/8  | 1/8  | 1/4  | 1/4  | 1/4  | 1       |

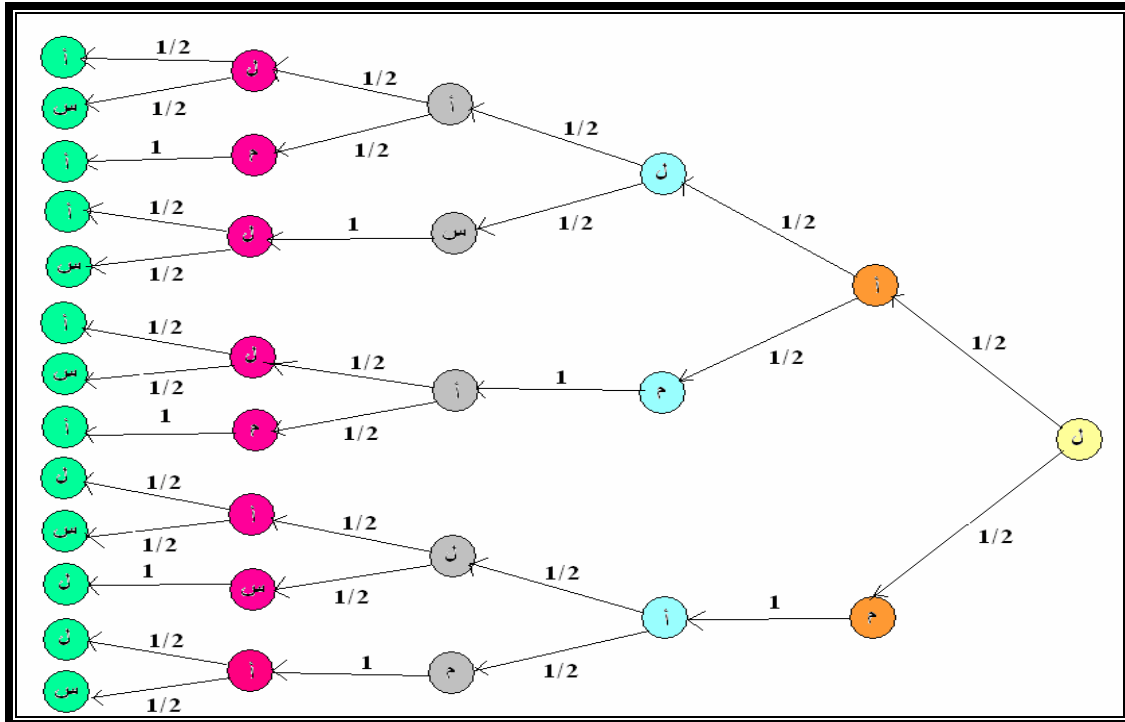
4- الكلمات الممكنة تكوينها والمؤلفة كل منها من خمسة حروف مبينة في الجدول الآتي هي واحتمالاتها.

| الكلمة   | الالا | الالس | الاما | السلس | امالا | امالس | اماما | المجموع |
|----------|-------|-------|-------|-------|-------|-------|-------|---------|
| احتمالها | 1/16  | 1/16  | 1/8   | 1/8   | 1/8   | 1/8   | 1/4   | 1       |

5- الكلمات الممكنة تكوينها والمؤلفة كل منها من ستة حروف مبينة في الجدول الآتي هي واحتمالاتها.

| الكلمة   | الالال | الالام | الالسل | الامال | الامام | السالل | السلام  |
|----------|--------|--------|--------|--------|--------|--------|---------|
| احتمالها | 1/32   | 1/32   | 1/16   | 1/16   | 1/16   | 1/16   | 1/16    |
| الكلمة   | السلسل | امالال | امالام | امالس  | امامال | امامام | المجموع |
| احتمالها | 1/8    | 1/16   | 1/16   | 1/8    | 1/8    | 1/8    | 1       |

ثانياً:- بالبدء بالحرف "لام" باعتباره جذراً سيتم الحصول على المخطط الشجري المبين في الشكل الآتي.



الشكل (2): المخطط الشجري لخمسة أجيال ناجمة من الجذر "لام" في فضاء كلمة "السلام".

من المخطط السابق يلاحظ الآتي:

1- هناك كلمتان يمكن تكوينهما كل منهما مؤلفة من حرفين فقط، وهما كلمتي "لا" و "لم"، وكل منهما باحتمالية 50%.

2- الكلمات الممكنة تكوينها والمؤلفة كل منها من ثلاثة حروف مبيّنة في الجدول الآتي هي واحتمالاتها.

| الكلمة   | لال | لام | لما | المجموع |
|----------|-----|-----|-----|---------|
| احتمالها | 1/4 | 1/4 | 1/2 | 1       |

3- الكلمات الممكنة تكوينها والمؤلفة كل منها من أربعة حروف مبيّنة في الجدول الآتي هي واحتمالاتها.

| الكلمة   | لالا | لالس | لاما | لامال | لامام | المجموع |
|----------|------|------|------|-------|-------|---------|
| احتمالها | 1/8  | 1/8  | 1/4  | 1/4   | 1/4   | 1       |

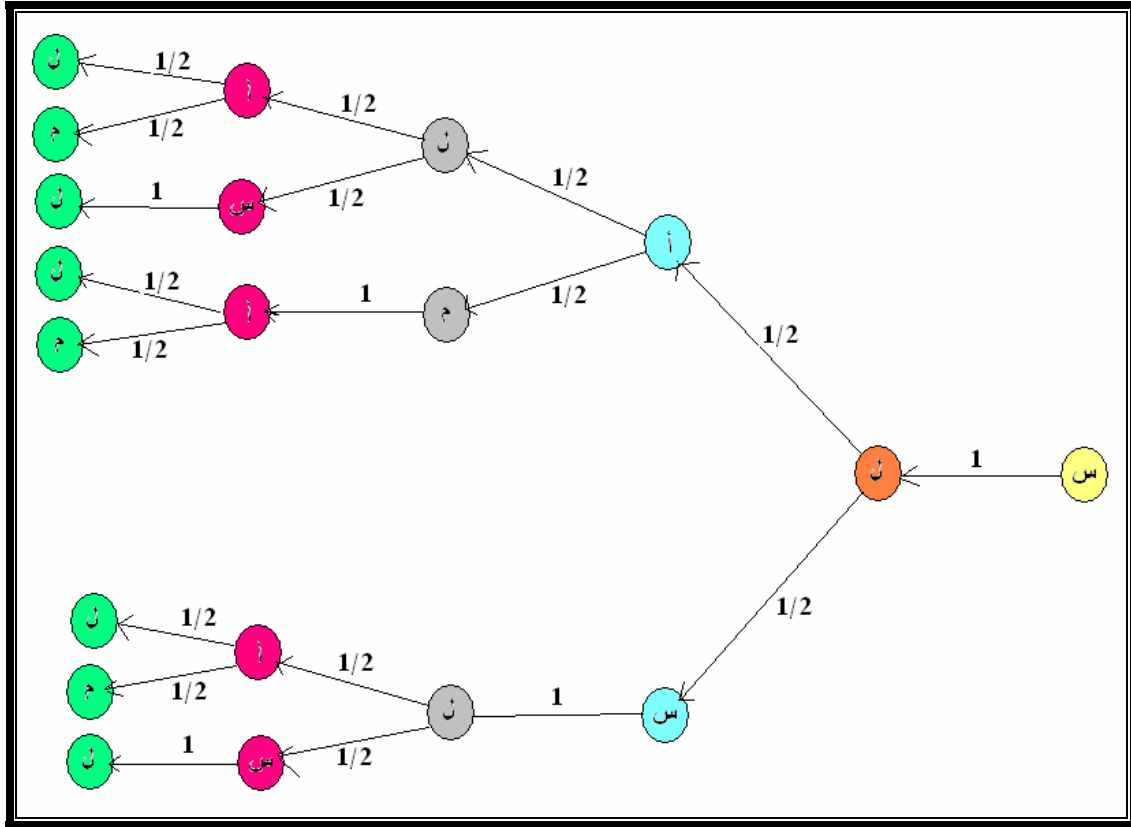
4- الكلمات الممكنة تكوينها والمؤلفة كل منها من خمسة حروف مبيّنة في الجدول الآتي هي واحتمالاتها.

| الكلمة   | لالال | لالام | لالسل | لامال | لامام | لامالا | لامالس | لاماما | المجموع |
|----------|-------|-------|-------|-------|-------|--------|--------|--------|---------|
| احتمالها | 1/16  | 1/16  | 1/8   | 1/8   | 1/8   | 1/8    | 1/8    | 1/4    | 1       |

5- الكلمات الممكنة تكوينها والمؤلفة كل منها من ستة حروف مبيّنة في الجدول الآتي هي واحتمالاتها.

| الكلمة   | لالالا | لالالاس | لالالاسلا | لالالاسلا | لالالاسلا | لالالاسلا | لالالاسلا | لالالاسلا | لالالاسلا | المجموع |
|----------|--------|---------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|---------|
| احتمالها | 1/32   | 1/32    | 1/16      | 1/16      | 1/16      | 1/16      | 1/16      | 1/16      | 1/16      | 1/16    |
| الكلمة   | لامالا | لامالاس | لامالاسلا | لامالاسلا | لامالاسلا | لامالاسلا | لامالاسلا | لامالاسلا | لامالاسلا | المجموع |
| احتمالها | 1/8    | 1/16    | 1/16      | 1/16      | 1/16      | 1/16      | 1/16      | 1/8       | 1/8       | 1       |

ثالثاً:- بالبدء بالحرف "سين" باعتباراه جذرا سيتم الحصول على المخطط الشجري المبين في الشكل الآتي.



الشكل(3):المخطط الشجري لخمسة أجيال ناجمة من الجذر "سين" في فضاء كلمة "السلام".  
من المخطط السابق يلاحظ الآتي:

1- هناك كلمة واحدة فقط مؤلفة من حرفين يمكن تكوينها ، وهي كلمة "سل" باحتمالية %100.

2- هناك كلمتان فقط كل منهما مؤلفة من ثلاثة حروف مبيّنة في الجدول الآتي هي واحتمالاتها.

| الكلمة   | سلا | سلس | المجموع |
|----------|-----|-----|---------|
| احتمالها | 1/2 | 1/2 | 1       |

3- هناك ثلاث الكلمات مؤلفة كل منها من أربعة حروف يمكن تكوينها مبينة في الجدول الآتي هي واحتمالاتها.

| الكلمة   | سلا | سلام | سلسل | المجموع |
|----------|-----|------|------|---------|
| احتمالها | 1/4 | 1/4  | 1/2  | 1       |

4- الكلمات الممكن تكوينها والمؤلفة كل منها من خمسة حروف مبينة في الجدول الآتي هي واحتمالاتها.

| الكلمة   | سلا | سلا | سلا | سلا | سلا | المجموع |
|----------|-----|-----|-----|-----|-----|---------|
| احتمالها | 1/8 | 1/8 | 1/4 | 1/4 | 1/4 | 1       |

5- الكلمات الممكن تكوينها والمؤلفة كل منها من ستة حروف مبينة في الجدول الآتي هي واحتمالاتها.

| الكلمة   | سلا  | سلا     | سلا | سلا | سلا | سلا | المجموع |
|----------|------|---------|-----|-----|-----|-----|---------|
| احتمالها | 1/16 | 1/16    | 1/8 | 1/8 | 1/8 | 1/8 | 1/8     |
| الكلمة   | سلسل | المجموع |     |     |     |     |         |
| احتمالها | 1/4  | 1       |     |     |     |     |         |

رابعا: - لو بدأنا بالحرف "ميم" باعتباره جذرا فإننا سنحصل على المخطط الشجري المبين في الشكل الآتي.





5- الكلمات الممكنة تكوينها والمؤلفة كل منها من ستة حروف مبينة في الجدول الآتي هي واحتمالاتها.

|          |        |         |        |        |        |        |        |
|----------|--------|---------|--------|--------|--------|--------|--------|
| الكلمة   | مالالا | مالالس  | مالاما | مالسلا | مالسلس | مامالا | مامالس |
| احتمالها | 1/16   | 1/16    | 1/8    | 1/8    | 1/8    | 1/8    | 1/8    |
| الكلمة   | ماماما | المجموع |        |        |        |        |        |
| احتمالها | 1/4    | 1       |        |        |        |        |        |

والجدول الآتي يبين الكلمات المرجحة المتولدة من الجذور الأربعة في فضاء كلمة "السلام" والمؤلفة كل منها من 2،3،4،5 و6 حروف.

الجدول(2): الكلمات المرجحة المتولدة من الجذور الأربعة في فضاء كلمة "السلام" .

| الاحتمالات | الكلمات المرجحة                    | عدد الحروف | الجذر |
|------------|------------------------------------|------------|-------|
| 1/2        | ال ، ام                            | 2          |       |
| 1/2        | اما                                | 3          |       |
| 1/4        | امال ، امام                        | 4          | الف   |
| 1/4        | اماما                              | 5          |       |
| 1/8        | امالسلس ، امامال ، امامام          | 6          |       |
| 1/2        | لا ، لم                            | 2          |       |
| 1/2        | لما                                | 3          |       |
| 1/4        | لاما ، لمال ، لمام                 | 4          | لام   |
| 1/4        | لماما                              | 5          |       |
| 1/8        | لاماما ، لمالسلس ، لمامال ، لماماس | 6          |       |
| 1          | سل                                 | 2          |       |
| 1/2        | سلا ، سلس                          | 3          |       |
| 1/2        | سلسلس                              | 4          | سين   |
| 1/4        | سلاما ، سلسلا ، سلسلس              | 5          |       |
| 1/4        | سلسلسلس                            | 6          |       |
| 1          | ما                                 | 2          |       |
| 1/2        | مال ، مام                          | 3          |       |
| 1/2        | ماما                               | 4          | ميم   |
| 1/4        | مالسلس ، مامال ، مامام             | 5          |       |
| 1/4        | ماماما                             | 6          |       |

كذلك يمكننا حساب احتمالية حدوث أي متتابعة مولدة في هذا الفضاء. فعلى سبيل المثال لو أردنا حساب احتمالية حدوث الجملة " السلام مال الامال" في هذا الفضاء فإننا نحسب أولاً احتمالية حدوث كلمة من هذه الكلمات الثلاثة وكما يأتي:

$$P(\text{السلام}) = p_i \frac{1}{16} = 0.4444 \frac{1}{16} = 0.028,$$

$$P(\text{مال}) = p_r \frac{1}{2} = 0.2222 \frac{1}{2} = 0.111,$$

$$P(\text{الامال}) = p_i \frac{1}{16} = 0.4444 \frac{1}{16} = 0.028.$$

فتكون احتمالية حدوث الجملة " السلام مال الامال" في هذا الفضاء هي حاصل ضرب الاحتمالات الثلاثة، أي ان:

$$\begin{aligned} P(\text{السلام مال الامال}) &= P(\text{السلام}) P(\text{مال}) P(\text{الامال}) \\ &= 0.028 (0.111) (0.028) \\ &= 0.000087. \end{aligned}$$

#### 4- التطبيق الثاني (البسمة) :

نتناول في هذا التطبيق فضاء يتألف من مجموعة كلمات، هذه الكلمات تتضمن البسمة " بسم الله الرحمن الرحيم". تتألف البسمة من أربعة كلمات هي " بسم" و "الله" و "الرحمن" و "الرحيم". أي ان فضاء الحالة للكلمات التي تتكون منها البسمة هي المجموعة  $S = \{\text{بسم ، الله ، الرحمن ، الرحيم}\}$ . لقد تم الاستعانة بالحاسوب، بكل من البرنامجين الجاهزين المصحف الرقمي والمصحف المساعد، لحساب مشاهدات الانتقالات بين عناصر البسمة كما وردت في القرآن الكريم وهي مبينة في مصفوفة التكرارات الآتية، انظر الخياط [2009a] :

$$F = \begin{matrix} & \text{بسم} & \text{الله} & \text{الرحمن} & \text{الرحيم} \\ \text{بسم} & (-) & 3 & - & - \\ \text{الله} & (-) & 1 & 2 & - \\ \text{الرحمن} & (-) & - & - & 6 \\ \text{الرحيم} & (-) & - & - & - \end{matrix}$$

وكما هو واضح من مصفوفة التكرارات فقد وردت في المصحف ثلاثة انتقالات فقط من "بسم" إلى "الله"، وانتقال واحد من "الله" إلى "الله"، وانتقالان من "الله" إلى "الرحمن"، وستة انتقالات من "الرحمن" إلى "الرحيم"، ولا يوجد أي انتقال مباشر من "الرحيم" إلى أي من الكلمات الثلاثة الأخرى. وبعد تحويل التكرارات إلى احتمالات ينتج المصفوفة الانتقالية الآتية:





الجدول(3): الجمل المرجحة المتولدة من الجذور الثلاثة في فضاء البسمة.

| الاحتمالات | الجمل المرجحة                    | عدد الأجيال | الجذر  |
|------------|----------------------------------|-------------|--------|
| 1          | بسم الله                         | 2           |        |
| 2/3        | بسم الله الرحمن                  | 3           |        |
| 2/9        | بسم الله الرحمن الرحيم           | 4           | بسم    |
| 2/27       | بسم الله الله الرحمن الرحيم      | 5           |        |
| 2/81       | بسم الله الله الله الرحمن الرحيم | 6           |        |
| 2/3        | الله الرحمن                      | 2           |        |
| 2/3        | الله الرحمن الرحيم               | 3           |        |
| 2/9        | الله الله الرحمن الرحيم          | 4           | الله   |
| 2/27       | الله الله الله الرحمن الرحيم     | 5           |        |
| 2/81       | الله الله الله الرحمن الرحيم     | 6           |        |
| 1          | الرحمن الرحيم                    | 2           |        |
| -          | -                                | 3           |        |
| -          | -                                | 4           | الرحمن |
| -          | -                                | 5           |        |
| -          | -                                | 6           |        |

والشيء الملفت للنظر في الجدول الأخير ان الجمل المرجحة المتولدة من فضاء البسمة جميعها هي جمل محكمة لغويا ( ذات معنى تام لا يقبل التأويل )، كما ان تكرار لفظ أجلاله " الله " يفيد التأكيد.

## 20 - مناقشة:

تضمن هذا البحث مفهوما جديدا هو مفهوم المتتابعات الماركوفية، اذ يمكن باستخدام هذا المفهوم إنشاء شجرة قرار تستخدم لتوليد عدد محدد من العناصر اعتمادا على عناصر أساسية. كما تم تعريف مفهوم جديد آخر هو مفهوم المتتابعات المرجحة. وفي هذه المناقشة يتم حساب عدد المتتابعات الماركوفية ذوات الحروف المختلفة بحجم  $k=1,2,3,4,5$  والتي يمكن تكوينها في أبجدية كل من اللغة العربية واللغة الانكليزية، وذلك باستخدام المبرهنة (1)،

علما بان اللغة العربية تتألف من 28 حرفا أبجديا، في حين ان اللغة الانكليزية تتألف من 26 حرفا. والجدول الآتي يوضح النتائج النهائية إضافة إلى النسبة المؤية للزيادة في عدد المتتابعات الماركوفية في اللغة العربية بالنسبة إلى تلك في اللغة الانكليزية.

الجدول(1): المتتابعات الماركوفية الممكنة بحجم k في اللغتين العربية والانكليزية.

| 5        | 4      | 3     | 2       | 1  | k                | عندما يكون هناك انتقال من حرف إلى نفسه |
|----------|--------|-------|---------|----|------------------|--|
| 17210368 | 614656 | 2195  | 78<br>4 | 28 | اللغة العربية    |  |
| 11881376 | 456976 | 17576 | 67<br>6 | 26 | اللغة الانكليزية |  |
| 45       | 35     | 25    | 16      | 8  | % للزيادة        |  |

| 5        | 4      | 3     | 2       | 1  | k                | عندما لا يكون هناك انتقال من حرف إلى نفسه ويسمح بتكرار الحرف في باقي المواقع |
|----------|--------|-------|---------|----|------------------|--|
| 14880348 | 551124 | 20412 | 75<br>6 | 28 | اللغة العربية    |  |
| 10156250 | 406250 | 16250 | 65<br>0 | 26 | اللغة الانكليزية |  |
| 47       | 35     | 25    | 16      | 8  | % للزيادة        |  |

| 5        | 4      | 3     | 2   | 1  | k                | عندما لا يكون هناك انتقال من حرف إلى نفسه ولا يسمح بتكرار الحرف في باقي المواقع |
|----------|--------|-------|-----|----|------------------|---|
| 11793600 | 491400 | 19656 | 756 | 28 | اللغة العربية    |   |
| 7893600  | 358800 | 15600 | 650 | 26 | اللغة الانكليزية |   |
| 49       | 37     | 26    | 16  | 8  | % للزيادة        |   |

وكما هو واضح من الجدول السابق فان اللغة العربية تتفوق على اللغة الانكليزية في عدد المتتابعات الماركوفية التي يمكن تكوينها. وحيث ان المتتابعات الماركوفية الممكنة تكوينها من لغة معينة تمثل مفردات تلك اللغة، فان هذا يعني بان عدد مفردات اللغة العربية أكثر وبشكل ملحوظ من تلك في اللغة الانكليزية.

لقد تم إيضاح مفهومي المتتابعات الماركوفية والمتتابعات المرجحة من خلال مثالين، كما تم إجراء تطبيقين في مجال اللغة العربية لتوليد كلمات من حروف كلمة "السلام". والجدول الآتي يبين الكلمات المرجحة المتولدة من الجذور الأربعة في فضاء كلمة "السلام" والمؤلفة كل منها من 6،2،3،4،5 حروف.

الجدول(2): الكلمات المرجحة المتولدة من الجذور الأربعة في فضاء كلمة "السلام".

| الاحتمالات | الكلمات المرجحة                    | عدد الحروف | الجذر |
|------------|------------------------------------|------------|-------|
| 1/2        | ال، ام                             | 2          |       |
| 1/2        | اما                                | 3          |       |
| 1/4        | امال، امام                         | 4          | ألف   |
| 1/4        | اماما                              | 5          |       |
| 1/8        | امالسل، امامال، امامام             | 6          |       |
| 1/2        | لا، لم                             | 2          |       |
| 1/2        | لما                                | 3          |       |
| 1/4        | لاما، لمال، لمام                   | 4          | لام   |
| 1/4        | لاماما                             | 5          |       |
| 1/8        | لامامال، لمالسل، لمامال،<br>لماماس | 6          |       |
| 1          | سل                                 | 2          |       |
| 1/2        | سلا، سلس                           | 3          |       |
| 1/2        | سلسل                               | 4          | سين   |
| 1/4        | سلاما، سلسلا، سلسلس                | 5          |       |
| 1/4        | سلسلسل                             | 6          |       |
| 1          | ما                                 | 2          |       |
| 1/2        | مال، مام                           | 3          |       |
| 1/2        | ماما                               | 4          | ميم   |
| 1/4        | مالسل، مامال، مامال                | 5          |       |
| 1/4        | ماماما                             | 6          |       |

كذلك يمكن حساب احتمالية حدوث أي متتابعة مولدة في هذا الفضاء. فعلى سبيل المثال لو أُريد حساب احتمالية حدوث الجملة "السلام مال الامال" في هذا الفضاء فإنه تُحسب أولاً احتمالية حدوث كل كلمة من هذه الكلمات الثلاثة وكما يأتي:

$$P(\text{السلام}) = p_1 \frac{1}{16} = 0.4444 \frac{1}{16} = 0.028.$$

$$P(\text{مال}) = p_1 \frac{1}{2} = 0.2222 \frac{1}{2} = 0.111.$$

$$P(\text{الامال}) = p_1 \frac{1}{16} = 0.4444 \frac{1}{16} = 0.028.$$

فتكون احتمالية حدوث الجملة "السلام مال الامال" في هذا الفضاء هي حاصل ضرب الاحتمالات الثلاثة، أي ان:

$$\begin{aligned} P(\text{الامال}) P(\text{مال}) P(\text{السلام}) &= p(\text{السلام}) P(\text{مال}) P(\text{الامال}) \\ &= 0.028(0.111)(0.028) \\ &= 0.000087. \end{aligned}$$

## المصادر

- 1- الخياط، باسل يونس ذنون و العجوز، زينة فالح صالح(2009)، "النمذجة الماركوفية لكلمة التوحيد"، المجلة العراقية للعلوم الإحصائية، العدد .
- 2- الخياط، باسل يونس ذنون (2009)، "النمذجة الرياضية لعدد الركعات في الصلوات الخمس"، المجلة العراقية للعلوم الإحصائية، العدد .
- 3- الخياط، باسل يونس ذنون (2010)، "مدخل إلى الاحتمالية وتطبيقاتها باستخدام MATLAB"، كتاب قيد الإعداد.

4-Kalman, R.E. (1960), "A new approach to linear filtering and prediction problems", Trans. ASME J. Basic Engtg., Series D, 28, 35-45.

5-Rabiner, L.R. (1989), "A tutorial on hidden Markov models and selected applications in speech recognition", Proceeding of IEEE, Vol. 77, No. 2, 257-286.