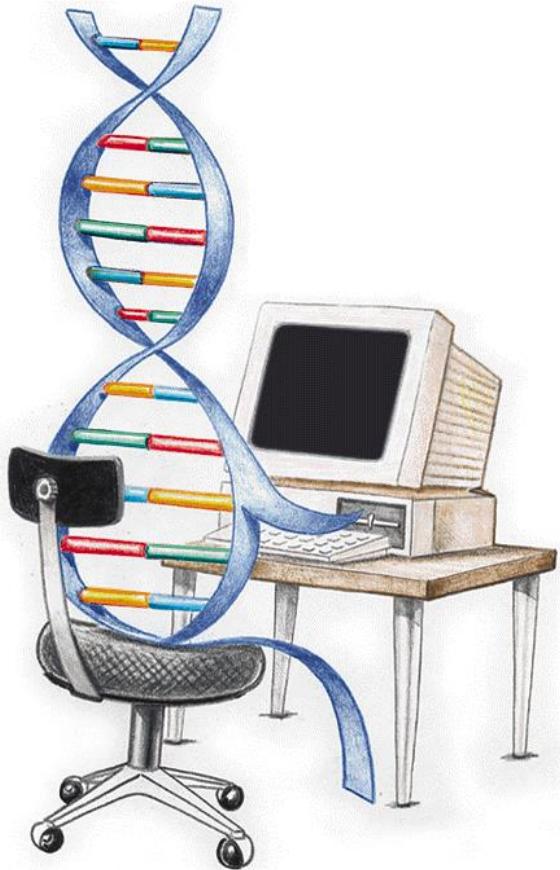


ذكاء صنعي 2

محاضرة 1

Genetic Algorithms

د. فادي متوج





What are Genetic Algorithms?

- **الخوارزميات الوراثية** **genetic algorithms** هي نماذج حسابية تعتمد على علم الوراثة والتطور في الكائنات الأحياء.
- تعتمد فلسفة الخوارزميات الوراثية على **توليد** عدد كبير من الحلول الممكنة لمشكلة معينة.
- بعد ذلك، يتم **تقييم** كل حل من هذه الحلول. و تكون للحلول الأفضل فرص أكبر لتوليد حلول أخرى في حين تقل فرص توالد الحلول السيئة.
- وبتكرار هذه العملية **تطور** نوعية الحلول المطروحة و تصل أو تقترب من الحل الأمثل للمشكلة تماماً مثلاً ما تتطور الكائنات الحية حتى تتكيف بشكل أفضل مع بيئتها.

Genetics in real life

- صفات كل كائن حي مشفرة في **جينات (مورثات)** هذا الكائن الحي، والتي بدورها ترتبط معاً في سلاسل طويلة تسمى **الكروموسومات (الصبغيات)**. **chromosomes**.
- يمثل كل جين سمة محددة للكائن الحي، مثل لون العين أو لون الشعر أو زمرة الدم، ويمكن أن يأخذ كل جين قيمة من بين عدة قيم مختلفة. على سبيل المثال، جين لون الشعر قد يكون أشقر، أسود أو بني محمر. وعادة ما يشار إلى هذه الجينات وقيمها **بالتراكيب الوراثي** للكائن.
- معظم قيم الجينات يتم توارثها من الأم والأب للطفل.
- في بعض الأحيان، تتغير بعض قيم الجينات (على سبيل المثال، من 0 إلى 1)، قد يحصل ذلك بسبب بعض المؤثرات الكيميائية أو الفيزيائية أو البيولوجية الغير العادية (على سبيل المثال، تعرض جين للإشعاعات الذرية). ويسمى هذا التغيير في قيمة الجين بـ **الطفرة mutation**.

- في أي نوع من أنواع الكائنات الحية، فإن الأفراد الذين يتكيفون بطريقة أفضل مع البيئة لديهم احتمالات أعلى للبقاء على قيد الحياة، وبالتالي لديهم احتمالات أعلى لإنجاب أطفال.
- على مدى أجيال، يتم تكرار هذه العملية، والنتيجة هي أن هؤلاء الأفراد والجينات التي تتكيف بشكل أفضل مع البيئة يميلون إلى **البقاء** بينما تلك التي لا تتكيف تميل إلى **الانقراض**.
- وتسمى نظرية الفرز الطبيعية هذه بالداروينية **(Darwinian evolution)**.

- الخوارزميات الوراثية **الحسابية** التي سوف ندرسها هي **نماذج مختصرة** من علم الوراثة الطبيعية وعملية التطور التي ناقشناها سابقاً.
- الخوارزميات الجينية تشمل مفاهيم مثل الكروموسومات chromosomes والجينات genes، التزاوج mating or crossover، الطفرة mutation، والتطور evolution.
- العملية الأساسية للخوارزمية الجينية هي على النحو التالي:
 - في البداية، نولد عشوائيا حلول أو "كروموسومات" للمشكلة.
 - بعد التوليد الأولي للحلول، نقوم بأدوار تكرارية iterations. ويكون كل تكرار من عدة خطوات - اختيار حلول جيدة نجري عملية تزاوج .
في بعض الأحيان قد تكون لدينا طفرات على بعض الحلول. من خلال اختيار حلول جيدة خلال التكرار، سيقوم الكمبيوتر بتطوير أفضل الحلول باستمرار كما في حالة من التطور الطبيعي.

- يمكن تطبيق الخوارزميات الجينية لحل كثير من المسائل مثل مسائل **إيجاد الحل الأمثل** لمشكلة مطروحة (**optimization**)

- فمثلاً لو أراد مندوب إحدى الشركات زيارة عشرين مدينة مختلفة باستعمال السيارة مع مراعاة تقليل وقت التنقل و عدد الكيلومترات المقطوعة فإنه سيجد نفسه أمام إشكالية رياضية معقدة.

- وهذه إشكالية كلاسيكية تعرف باسم "**مشكلة البائع المتجول**" (Traveling Salesman Problem) (TSP). لكن حل هذه المشكلة و مشكلات أخرى مماثلة أو أكثر تعقيداً يكون سهل باستعمال الخوارزميات الوراثية.



Representations of solutions

- A **solution** here means any value that is a candidate for a **correct solution** or a final answer.
- A solution may or may not be the correct solution.
- For example, suppose we want to maximize function $y = 5 - (x - 3)^2$. Then $x = 1$ is a solution, $x = 2.5$ is another solution, and $x = 3$ is the correct solution that maximizes y .

- قد يختلف تمثيل الحلول من مسألة إلى أخرى ولكن الشائع هو استعمال **سلسلة من الأرقام الثنائية** (binary numbers)
- فمثلاً إذا كانت الحلول المنشورة لمشكلة معينة أرقاماً صحيحة وتقع بين الصفر والخمسة عشر فيمكن ترميزها باستعمال أرقام ثنائية مثل 0000، 1010، 1110، إلى غير ذلك وتسمى هذه الأرقام **كروموسومات**.
- إذاً كل كروموسوم هو حل ممكن للمشكلة المطروحة

The fitness of a solution

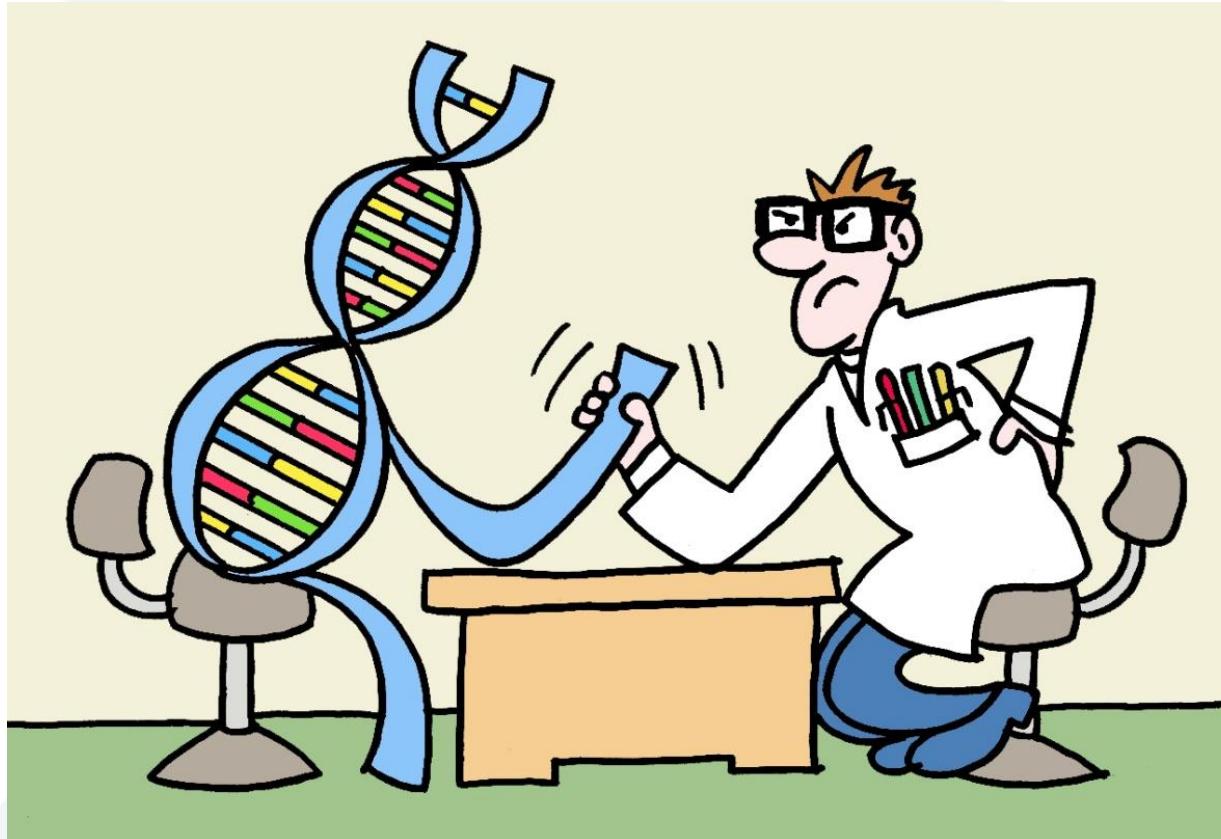
- لياقة الحل هو مقياس يمكن استخدامه لمقارنة الحلول لتحديد أيها أفضل.
- عادةً ما يسمى هذا المقياس "تابع اللياقة" (**Fitness Function**)
- على سبيل المثال، شركة ما تسعى لزيادة الربح. الربح نفسه يمكن أن يستخدم كمقياس اللياقة.
- يأخذ تابع اللياقة كل كروموسوم (حل) على حدة و يقيم مدى أدائه في حل المشكلة بإعطاء قيمة معينة.
- وكلما كانت هذه القيمة أكبر كلما كان الكروموسوم أكثر كفاءة.

- وبالتالي فإن الكروموسومات (الحلول) التي لها لياقة عالية تكون لها حظوظ أوفر في البقاء لأنها أكثر فائدة لحل المشكلة.
- وبتكرار هذه الخطوات **تزداد لياقة** الكروموسومات.
- بعد انتهاء عمليات الخوارزميات تصل الكروموسومات إلى أعلى مستوى من اللياقة مما يعكس حلولاً جيدة للمشكلة المراد حلها.

Basic steps of a genetic algorithm



تفاصيل الخوارزميات الوراثية



Fitness Evaluation تقييم اللياقة

- يتم تقييم اللياقة عن طريق تابع رياضي و يحتاج التقييم إلى ثلاث مراحل :
 - في **المراحل الأولى** يتم فك ترميز الكروموسوم (decoding)
 - و في **المراحل الثانية** يتم تحويل الأرقام الثنائية إلى أرقام عشرية بين حدود معينتين.
 - **أما المراحل الأخيرة** فمهمتها تقييم هذه الأرقام العشرية لاعطاء قيمة تعكس لياقة الكروموسوم.

$f(x,y)$

- لنفترض أننا نريد البحث عن القيمة العظمى لتابع ما

على أن تكون قيمة x و y بين 0 و 15

- لأخذ عشوائياً إثنين من الكروموسومات :

□ الكروموسوم الأول : 10110110

□ الكروموسوم الثاني : 01010000

- عندما نفك ترميز هذين الكروموسومين و نحولهما إلى أرقام عشرية سنحصل على:

- الكروموسوم الأول $x_1 = 1011 = 11, y_1 = 0110 = 6$

- الكروموسوم الثاني $x_2 = 0101 = 5, y_2 = 0000 = 0$

- لم يبقى إلا تقييم اللياقة:

$f(x_1, y_1) = 3.214$ الكروموسوم الأول:

$f(x_2, y_2) = 0.146$ الكروموسوم الثاني:

ومن هذه النتائج يمكننا أن نستخلص أن الكروموسوم الأول أكثر لياقة و أقرب للحل الأمثل

- تجدر الإشارة هنا إلى أن الخوارزميات الوراثية لا ترتبط بالمشكلة المراد حلها إلا عن طريق **تابع التقييم**.
- ولهذا يمكن استعمال برنامج الخوارزميات الوراثية في حل عدد كبير من المسائل دون إعادة كتابته أو تغيير أي من أجزائه باستثناء **تابع التقييم** (تابع اللياقة).
- كذلك نلاحظ أننا استعملنا أربعة برات لتتمثيل الأرقام الكبيرة يجب استعمال عدداً أكبر من البتات.

التجمع السكاني للكروموسومات (Population)

- نقطة البداية تكون دائماً بإعداد التجمع السكاني المبدئي و ذلك بـ**توليد عدد كبير من الكروموسومات بطريقة عشوائية**.
كأن نولد مثلاً 100 كروموسوم.
- تعطى الكروموسومات التي لها لياقة أكبر فرصة للتزاوج و التوالد و البقاء و بهذا يتم الحصول على 100 كروموسوم جديد و الاستغناء عن الكروموسومات القديمة. و نكون بذلك قد كوننا جيلاً جديداً (Generation) أكثر لياقةً من الجيل الذي سبقه.
- و مع مرور الأجيال تصل لياقة الكروموسومات إلى مستوى مرتفع

تولد الكروموسومات (Reproduction)

• تمر عملية تولد الكروموسومات بثلاث مراحل مهمة وهي :

- انتقاء الوالدين (Parent Selection)
- العبور الكروموسومي (Crossover)
- وأخيراً التغير المفاجئ أو الطفرة (Mutation)

قبل شرح هذه المراحل، تجدر الإشارة إلى أن الطريقة التي تتبعها الكرومومسومات هي نقطة قوة الخوارزميات الوراثية :

- لأنها تصل بعمليّة الاستقصاء إلى **حلول شمولية** (global solution)
- ولا تعلق في **الحلول الموضعية** (local solutions) كما هو الشأن في معظم طرق الاستقصاء المعروفة.

انتقاء الوالدين

Parents Selection

- تبدأ عملية التوالد **باختيار الوالدين** و المقياس الوحيد في هذا الاختيار هو **اللياقة**.
- و رغم أن عملية الانتقاء، كغيرها من العمليات، تتم عشوائياً إلا أن فرص انتقاء كروموسوم معين مرتبطة ارتباطاً وثيقاً بلياقته.
- فأكثر الكروموسومات لياقةً مرشح للانتقاء مرات عديدة في حين أن الكروموسوم الضعيف قد لا ينتقى إطلاقاً.
- إن الهدف من عملية انتقاء الوالدين هو إعطاء **الكروموسومات الجيدة** فرصة أكبر للتوالد وبالمقابل، تقليل فرص التوالد أمام الكروموسومات ضعيفة اللياقة

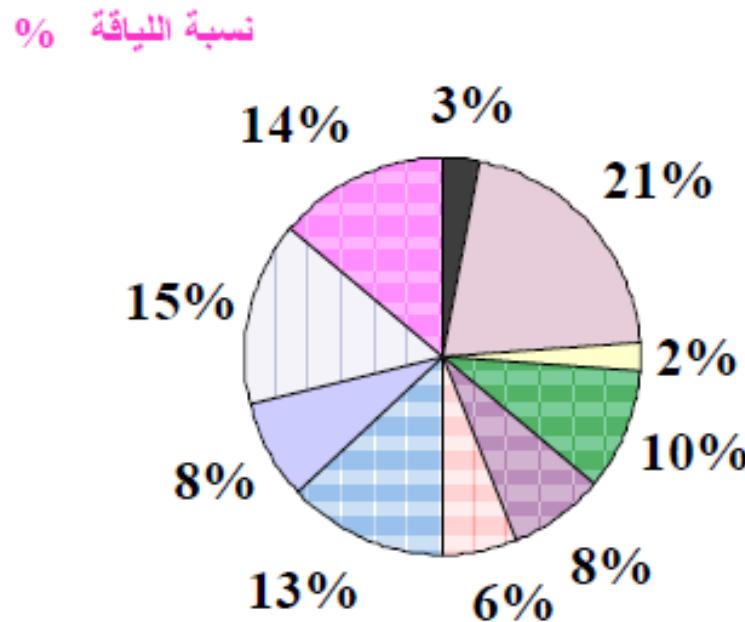
• للقيام بهذه العملية توجد طرق عديدة لكن أهمها وأكثرها استخداماً هي الطريقة المسمى **بالعجلة الدوارة** (roulette wheel) وهي مفصلة كالتالي:

- تجمع قيم لياقة كل الكروموسومات الموجودة ونحصل على اللياقة الإجمالية
- يولد رقم عشوائي شريطة أن يقع بين الصفر و قيمة اللياقة الإجمالية
- يتم انتقاء الكروموسوم الذي إذا ما جمعت لياقته مع لياقة الكروموسومات التي تسبقه ساوت أو تعدت قيمة الرقم العشوائي المولد

- لتوسيع هذه الخطوات، نفترض أننا نود القيام بعملية الانتقاء من تجمع سكاني يضم 10 كروموسومات بعد تقييم لياقتها

كروموسوم	10	9	8	7	6	5	4	3	2	1
اللياقة	6	4	9	6	11	10	2	15	1	7
المجموع الجاري	71	65	61	52	46	35	25	23	8	7

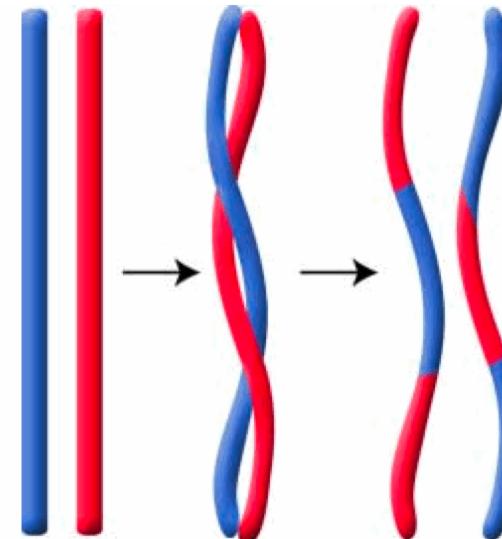
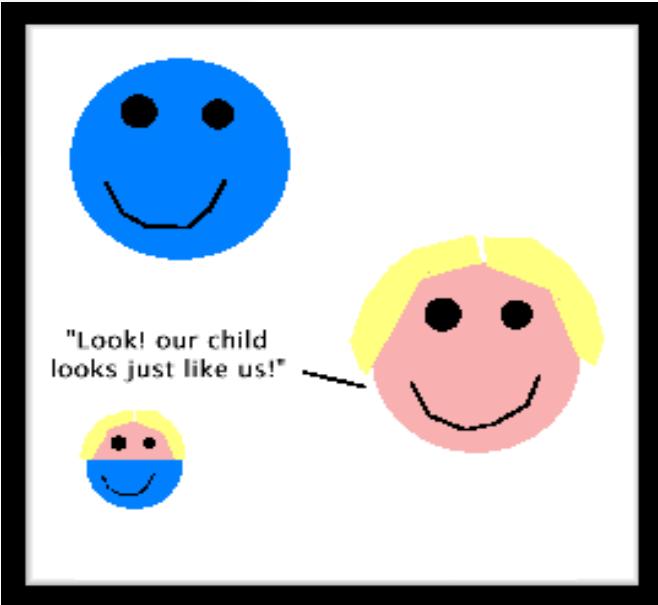
الرقم العشوائي	37	17	61	5	26	11	49
الكروموسوم المنتقى	6	3	8	1	5	3	7



- فـكما هو موضح في الشـكل، إـذا ما أـدرنا هـذه المسـاحة عـلى غـرار عـجلـة الحـظ فـحظـوظ المسـاحـات الأـكـبر من الطـبـيعـي أـن تكون أـوـفر مـن غـيرـها

العبور الكروموسومي (Crossover)

- هي التي تجسد عملية التوالي.
- وبعد اختيار الوالدين (يمثلان سلسلتين من الأرقام الثنائية) يقع **تبادل** جزء من السلسلة.
- تكون نقطة التبادل عشوائية تماماً. كأن يأخذ الكروموسوم الأول آخر ثلاثة برات من الكروموسوم الثاني ويعطيه آخر ثلاثة برات من سلسلته.
- ولا تتم هذه العملية مع كل والدين، ولكنها تتم بنسبة محددة يختارها واضع البرنامج كأن تطبق على 50% من الكروموسومات ولا تطبق على البقية.

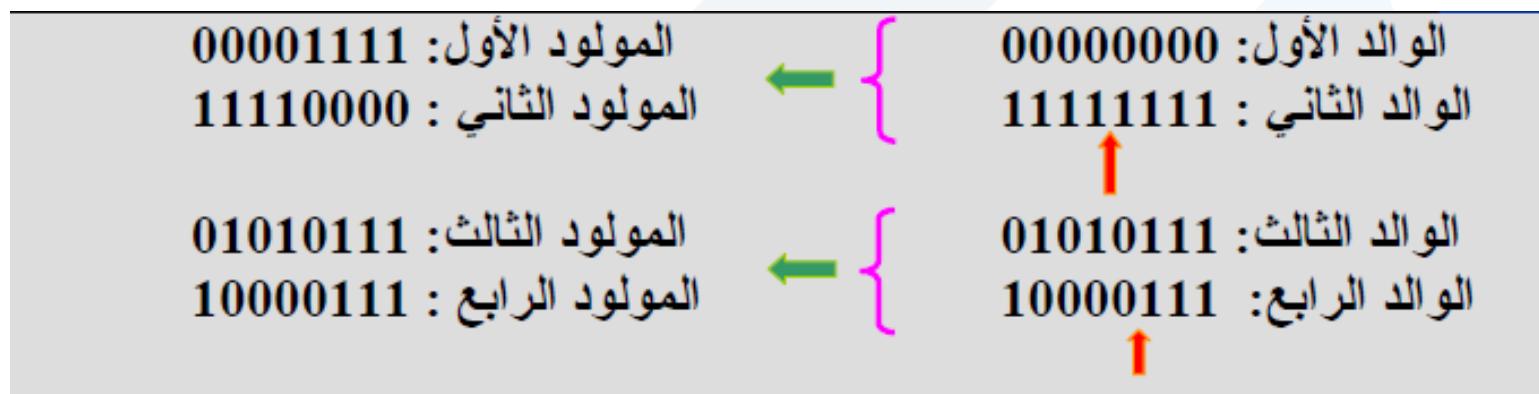


GGCTAATCTAAT**G**CTACCGTATAACAATCGAT
TATACGACCTA**G**GTACT**G**A**T**ATACAATCGAT

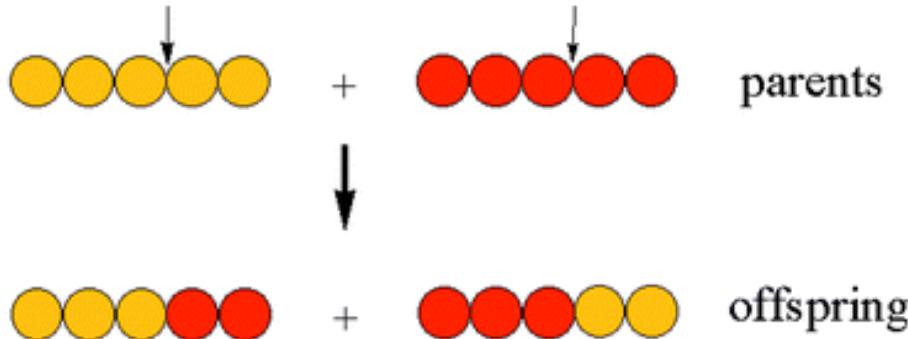
↓

GGCTAATCTAA**GGT**ACT**G**A**T**ATACAATCGAT
TATACGACCTA**T****G**CTAC**C****G**CTACCAGGGACG

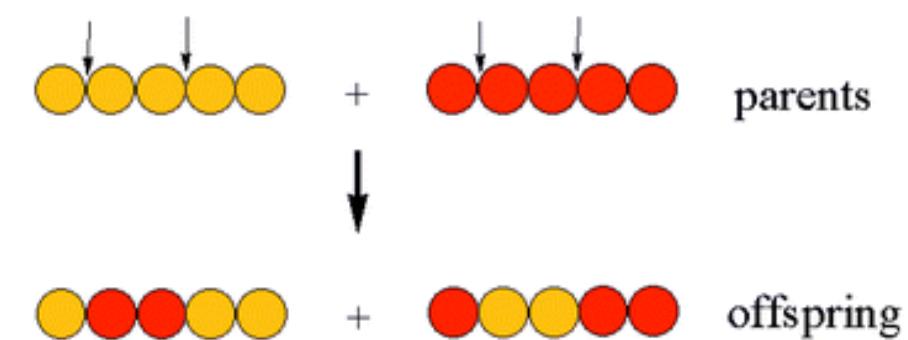
- بعد أن يتم تحديد نقطة التبادل عشوائياً يتبادل الوالدان جزءاً من سلسلة أرقامهما الثنائية الواقع بعد نقطة العبور الكروموسومي



(a)



(b)



الطفرة

Mutation

- تأتي عملية الطفرة أو التغير المفاجئ مباشرة بعد عملية العبور الكروموموني
- يمر الكرومومون الجديـد بـ**تـغـيـرـمـفـاجـعـيـ عـشـوـائـيـ** فـتـغـيـرـ إـحـدىـ بـتـاتـهـ منـ صـفـرـ إـلـىـ وـاحـدـ أوـ العـكـسـ
- وـ هـذـهـ الـعـمـلـيـةـ مـهـمـةـ لـأـنـهـاـ تـضـيـفـ بـعـضـ الـخـصـائـصـ الـجـديـدـةـ الـتـيـ قـدـ لاـ تـوـجـدـ فـيـ الـوـالـدـيـنـ
- كـمـاـ أـنـهـاـ تـمـنـعـ مـنـ الـوـقـوعـ فـيـ الـحـلـولـ الـمـحـلـيـةـ
- وـ بـعـبـارـةـ أـخـرىـ،ـ يـتـمـ اـسـتـخـدـامـ الـطـفـرـةـ لـلـحـفـاظـ عـلـىـ **الـتـنـوـعـ الـجـينـيـ**ـ بـيـنـ الـأـجـيـالـ.
- لـكـنـهـاـ لـاـ تـحـدـثـ إـلـاـ بـنـسـبـةـ صـغـيرـةـ جـداـ (ـمـنـ 0.5ـ%ـ إـلـىـ 1ـ%ـ)ـ لـأـنـهـ إـذـاـ زـادـتـ النـسـبـةـ عـنـ ذـلـكـ فـإـنـ الـخـوارـزمـيـةـ الـجـينـيـةـ تـصـبـحـ مـجـرـدـ خـوارـزمـيـةـ بـحـثـ عـشـوـائـيـ.

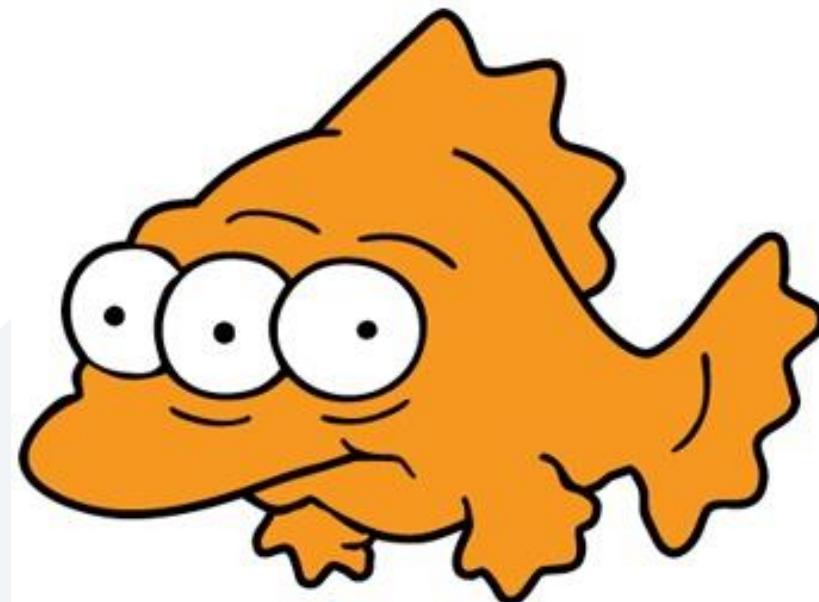
Mutation



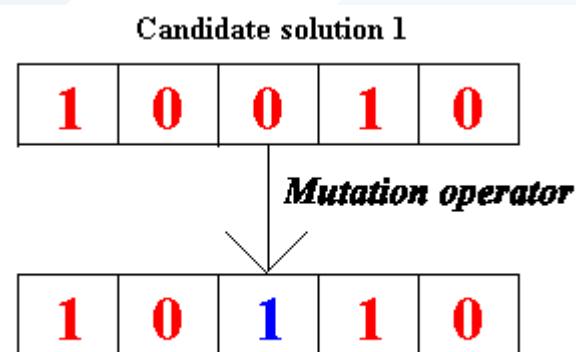
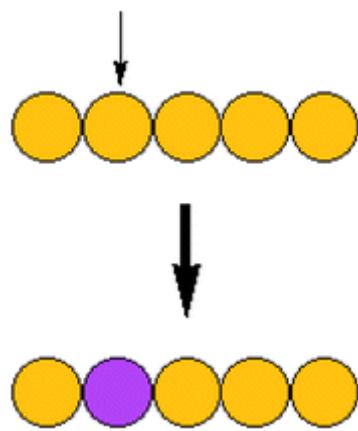
عند حدوث طفرة فإنها تتحول إلى

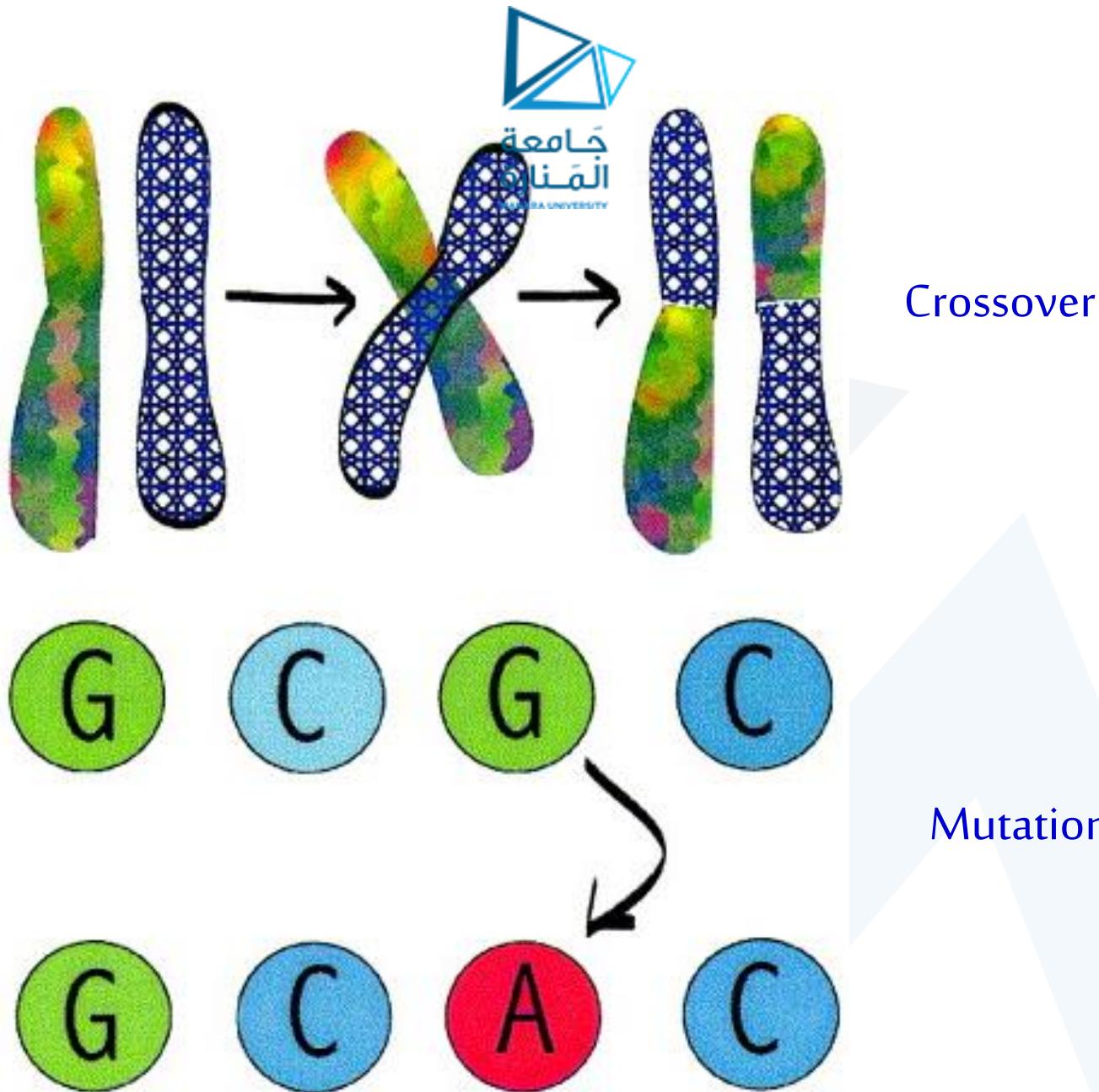
Mutation

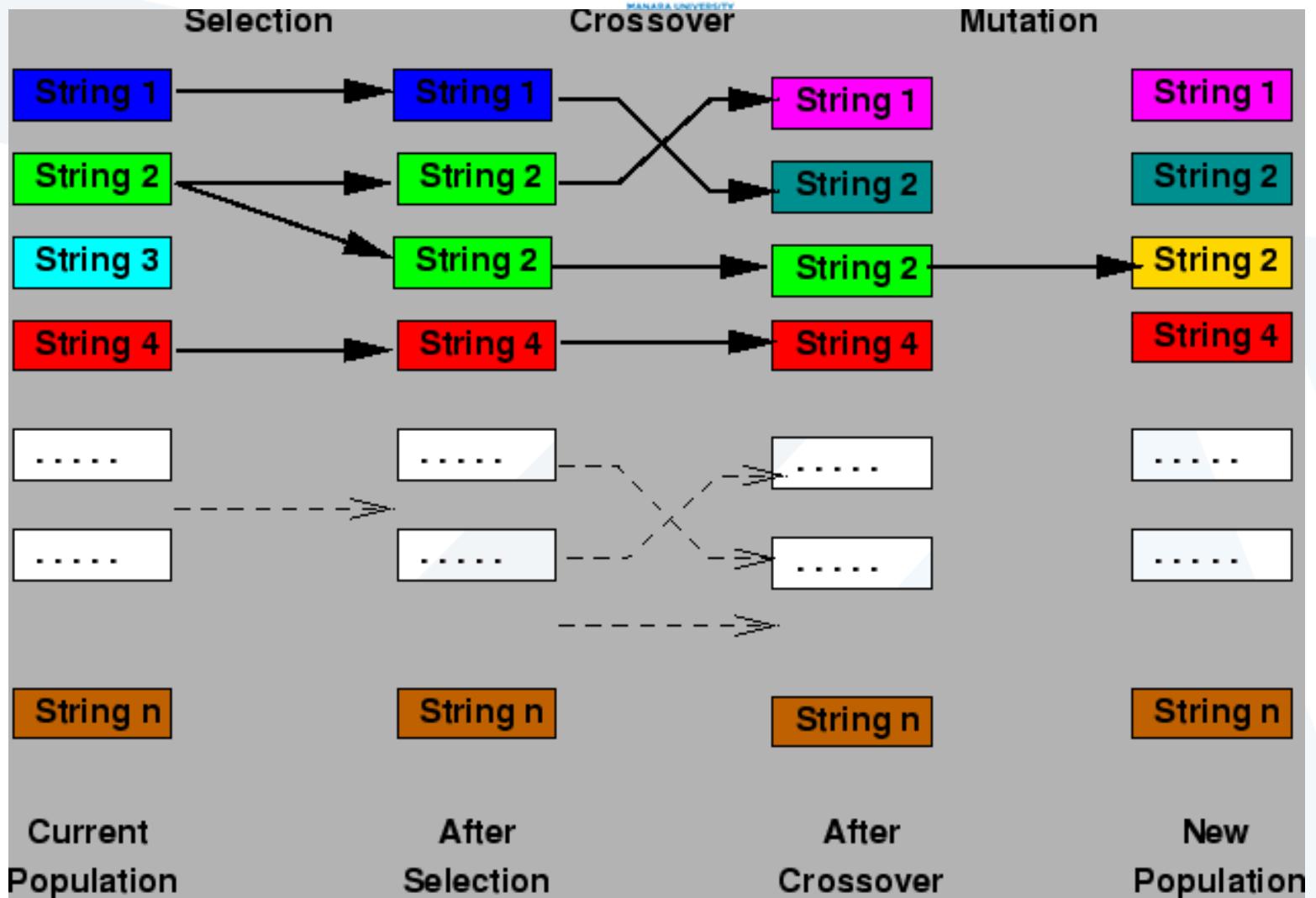
..... هذه



- تطبق هذه العملية على نسبة ضئيلة جداً من الكروموسومات المولدة(في حدود 1%)







مثال إيضاحي للخوارزميات الجينية

- يود أحد التجار استعمال شاحنة تبلغ سعة حمولتها القصوى **11000 كغ** و ذلك لنقل عدد من :

- ❖ السيارات
- ❖ و البرادات
- ❖ وأحواض مطبخ

الوزن (كغ)	القيمة (يورو)	
4000	3000	السيارة
400	280	الثلاجة
100	50	حوض المطبخ

- تتمثل **المشكلة** في إيجاد **العدد** الذي يجب نقله من كل صنف حتى:
 - نحصل على أكبر قيمة مالية
 - دون أن يتعدى الوزن الإجمالي سعة الحمولة القصوى
 - مع مراعاة أن تكون هذه الأعداد صحيحة دون كسور لأنه ليس من المعقول أن ننقل سيارتين ونصف مثلاً

- لنفترض :
 - ❖ التعداد السكاني للكروموسومات (الحلول) = 50
 - ❖ عدد الأجيال (مرات التكرار)= 30
 - ❖ نسبة العبور الكروموسومي = %60
 - ❖ نسبة الطفرة= %5
 - ❖ عدد البتات لكل كروموسوم= 4
- ليس هناك قانون واضح لتحديد هذه القيم كما أنها لا تحتاج إلى دراسة عميقة لاختيارها.
- فبإمكاننا اختيار قيم أخرى حسب الخطوط العريضة والتي سبق ذكرها دون أن تؤثر كثيراً على نتائج الخوارزميات

- ما يحتاج فعلاً إلى دراسة متأنية هو **تابع التقييم** (*Fitness Function*) (تابع اللياقة)
- إن اختيار هذا التابع مرتبط ارتباطاً وثيقاً بفعالية الخوارزميات التي لا تعرف شيئاً عن طبيعة المسألة إلا عن طريق تابع التقييم
- وبالتالي يجب أن يعكس هذا التابع و بكل دقة **طبيعة المسألة** و ما تتطلبه من **ضوابط (قيود)**

- قبل عرض التابع الذي تم اختياره لنعرف بعض المتغيرات و الثوابت و هي :

P_r = سعر الثلاجة = 280

W_a = وزن السيارة = 4000

W_k = وزن الحوض = 100

n_a = عدد السيارات

n_k = عدد الأحواض

P_a = سعر السيارة = 3000

P_k = سعر الحوض = 50

W_r = وزن الثلاجة = 400

\max_w = الحمولة = 11000

n_r = عدد الثلاجات

- إن القيم الثلاثة الأخيرة هي ما نسعى لإيجاده .
- حين تتوفر هذه القيم يمكننا حساب القيمة المالية للمواد المشحونة حسب المعادلة التالية:
- كما يمكننا حساب الوزن الإجمالي حسب المعادلة التالية:

$$\text{Value} = (n_a) (P_a) + (n_r) (P_r) + (n_k) (P_k)$$

$$\text{Weight} = (n_a) (W_a) + (n_r) (W_r) + (n_k) (W_k)$$

- من هنا يصبح الهدف هو البحث عن قيم $\text{Na}, \text{Nr}, \text{Nk}$ التي تعطينا أكبر قيمة مالية ممكنة و التي سميناها **Value**
- شريطة ألا يتعدى الوزن الإجمالي (Weight) الحمولة القصوى \max_w والمحددة ب 11000
- حسب هذه المعطيات هناك عدد كبير من توابع التقييم التي تفي بهذا الغرض وربما أبسطها و ليس بالضرورة أحسنها هو التابع التالي:

$$\frac{\text{Value}}{1 + (\max_w - \text{Weight})^2} = \text{Fitness}$$

$$\frac{\text{Value}}{1 + (\max_{\text{W}} - \text{Weight})^2} = \text{Fitness}$$

- وبهذا تصل قيمة اللياقة (Fitness) أعلى مستوياتها عندما نحصل (كما هو المطلوب) على:

- أعلى قيمة للمتغير Value

- وأقرب وزن إجمالي Weight من الحمولة القصوى

ترميز الكروموسومات

- أولاًً اخترنا أن تكون الحلول من أربعة باتاً فقط بحيث يكون الحد الأقصى لعدد السيارات أو البرادات أو الأحواض هو 1111 أي 15
- باختيار 4 باتاً لكل متغير يكون الكروموسوم **101100101001** عبارة عن ترميز للحل التالي

$$n_a = 1011 = 11$$

$$n_r = 0010 = 2$$

$$n_k = 1001 = 9$$

- بـتوليد تجمع سكاني يضم **50** كروموسوماً كما اخترنا لهـذه المسـألـة:
 - نـقـيم لـيـاقـة كـل مـنـهـا
 - نـقـوم بـعـمـلـيـة اـخـتـيـار الـوـالـدـين
 - فـعـمـلـيـة الـعـبـور الـكـرـوـمـوـسـومـي
 - ثـم الـطـفـرة
 - و تـنـتـهـي بـتـولـيد جـيل جـديـد
- بـعـد ذـلـك يـتم تـقـيـيم الـلـيـاقـة مـجـدـداً و تـتـكـرـر هـذـه الـعـمـلـيـات لـمـدة **30** جـيلاً وـهـو عـدـد الـأـجيـال الـذـي تم تـحـديـده آـنـفـاً
- عـنـد اـنـتـهـاء الـخـواـرـزـمـيـات الـوـرـاثـيـة مـن كـل هـذـه الـعـمـلـيـات نـحـصـل عـلـى لـيـاقـة كـل جـيل وـالـحـل الـأـمـثـل لـهـذـه المسـألـة

لياقة الأجيال

الجيل	الأول	الثاني	الثالث	من الرابع إلى الثلاثين
اللياقة	0	0.2	0.2	8060

الحل الأمثل

الوزن الإجمالي	القيمة الإجمالية	عدد الأحواض	عدد الثلاجات	عدد السيارات
11000	8060	2	7	2