

الخوارزمية الجينية في جدولة العمليات مع عدم إمكانية القطع

رقية زيدان شعبان
مدرس مساعد-جامعة الموصل
كلية الطب-وحدة الحاسبة الالكترونية
swallow2_2006@yahoo.com

إسراء نذير الكلاك
مدرس مساعد-جامعة الموصل
كلية التمريض-فرع العلوم الأساسية
isra19kalak@yahoo.com

المستخلص

تطرقنا الدراسة إلى استخدام مفهوم الخوارزمية الجينية في جدولة المعالجات المتعددة ، من خلال دمج خوارزميتين من خوارزميات الجدولة مع عدم إمكانية قطع العمليات ، والمتمثلة بالبيان الحلقي المباشر ، فالخوارزمية الأولى جدولة المستوى الأعلى أولاً مع تقدير الزمن "HLFET"، والخوارزمية الثانية "جدولة المستوى الأصغر أولاً مع تقدير الزمن" SCFET. أظهرت الدراسة نتائج إيجابية من خلال إبراز الحل الأفضل، من بين عدد كبير من الحلول في تقليل طول الجدولة (تقليل زمن إنهاء العمل)، من خلال الجدولة المثلى والمتمثلة بالخوارزمية الجينية المقترحة، إذ أظهرت كفاءة عالية وسهولة مرنة مقارنة مع استخدام خوارزميات الجدولة لوحدها.

A Genetic Algorithm in Scheduling Processes with Non – Permeability

Isra N. Alkallak
Assistant lecturer
Department of Basic Sciences
College of Nursing
University of Mosul

Ruqaya Z. Sha'ban
Assistant lecturer
Computers Unit
College of Medicine
University of Mosul

Abstract

The current paper tackles the combination of the genetic Algorithm with scheduling the multiprocessor by browsing the two algorithms for dependent tasks and non preemptive. The first algorithm is the highest level first estimated time (HLFET), and the second algorithm is smallest co -level first with estimated time (SCFET). The genetic algorithm proved to be powerful, efficient and appeared the feasible solution among many solutions; that minimize the schedule length (execution time) to find the optimal scheduling. In this research, it is proposed a genetic algorithm that finds a good combination of two list algorithms to produce a schedule with shortest schedule length.

The results of the experiments showed that scheduling found with the proposed two list scheduling genetic algorithms outperform those found with each one of the two lists scheduling alone.

١-١ المقدمة Introduction

تعد الخوارزمية الجينية Genetic Algorithm أحد أساليب الذكاء الاصطناعي، إذ برزت أهميتها في حل مسائل معقدة، إذ تمتلك كما هائل من الحلول البديلة، والحل الناتج من تطبيق الخوارزمية الجينية يكون في أغلب الأحيان حلاً قريباً إلى الحل الأمثل، كما تعتمد الخوارزمية الجينية على آلية الانتقاء الطبيعي ونظام الجينات الطبيعية (Goldberg, 1989). إن فكرة العمل للخوارزمية الجينية تعتمد على أفكار الهندسة الوراثية، التي تتميز بالإنتاج المقصود للمجموعات الموروثة بهدف تكوين أفراد ذوي صفات جيدة، وعلى هذا الأساس تقوم الخوارزمية الجينية بانتخاب الحلول الفضلى من بين عدد كبير من الحلول وإجراء بعض التداخلات والتبديلات بين هذه الحلول بهدف تكوين حل أفضل. (Gold, 1989, Hou et al., 1994).

نظراً لكون الخوارزمية الجينية ثقلاً وتختصر الكثير من الجهد والزمن المطلوبين لدى مصممي الأنظمة والبرامج، وذلك من خلال إيجاد خوارزمية عامة يعتمد عليها في حل مختلف أنواع المسائل، بدلاً من بناء خوارزمية خاصة لكل مسألة، مع مراعاة التغييرات اللازمة التي تتناسب مع خصوصية كل مسألة من حيث حجم ونوع البيانات المستخدمة وطبيعة دالة الهدف والقيود لكل مسألة. درست مسألة الجدولة Scheduling Problem بشكل واسع في حقل علوم الحاسبات وبحوث العمليات، وتعدّ من مسائل الأمثلية Optimization Problem، التي عادة تصنف إلى مسائل من نوع NP-Complete (Auyeung et al., 2003, Peng et al., 1995, Thiebaut, 1997)، ومن هنا جاء هدف الدراسة.

٢-١ نموذج البيان الحلقي المباشر (Directed Acyclic Graph Model DAG)

تعدّ معالجة العمليات المرتبطة والمتمثلة في نموذج البيان الحلقي المباشر ضمن لبنة البحوث في علوم الحاسبات، ويمكن تمثيل البرنامج المتوازي Parallel Program من خلال تمثيل العمليات بوساطة نموذج البيان الحلقي المباشر DAG، وسمي بذلك لعدم وجود حلقة (رجوع Cycle) في البيان بين العقدتين (Albala, 1977, Gonzalez, 1977, et al., 1997), إذ إن $G = (E, V)$ ، هي مجموعة من العقد Node، و E هي مجموعة من الحافات Edge، إذ تمثل العقد مجموعة العمليات التي يجب تنفيذها بالتتابع، في حين تمثل الحافات الاتصال بين تلك العمليات (Collins, 2001, Aronsson et al., 2002)، يرافق كل عقدة وزن Weight يمثل كلفة الحساب Computation Cost، ويقصد بها زمن التنفيذ للعملية. ويرافق كل حافة اتصال بين عقدتين وزن Weight يمثل كلفة الاتصال Communication Cost، وإذا تم تنفيذ جميع العمليات في المعالج نفسه، فإن كلفة الاتصال تفترض أن تكون صفراً أو تهمل (Kowk et al., 1999, Radulescu et al., 2002). إن أساس مفهوم الجدولة

لاينبغي اختراق علاقة الأسبقية Precedence Relation بين العقد (العمليات)، وإن العملية المستعدة للتنفيذ Ready هي العملية التي لا تمتلك أسلفاً أو تم تنفيذها (Gonzalez,1977, Baskiyar, 2002, Radulescu et .al., 2002).

١-٣ الخطوات الأساسية في الجدولة Basic Steps in Scheduling

تستند معظم خوارزميات الجدولة إلى أسلوب جدولة القائمة List Scheduling، التي تعتمد على عمل قائمة، أي أن يتم تخصيص أسبقية لكل عملية، وإن أغلب خوارزميات الجدولة تعمل بالأساس على ما يأتي:

١. تحديد الأسبقية لكل عملية.
 ٢. اختيار العملية ذات الأسبقية العالية.
 ٣. تخصيص العملية للمعالج مع تحديد المدة الزمنية للتنفيذ.
- أما في حالة وجود أكثر من عملية تمتلك الأسبقية نفسها، فسيتم الاختيار عشوائياً (كيفياً) Arbitrary أي ليس على وفق معايير حسابية (Arroyo,1996, Kowk et .al.,1999).

١-٤ نموذج لوحة جانت Gantt Chart Model

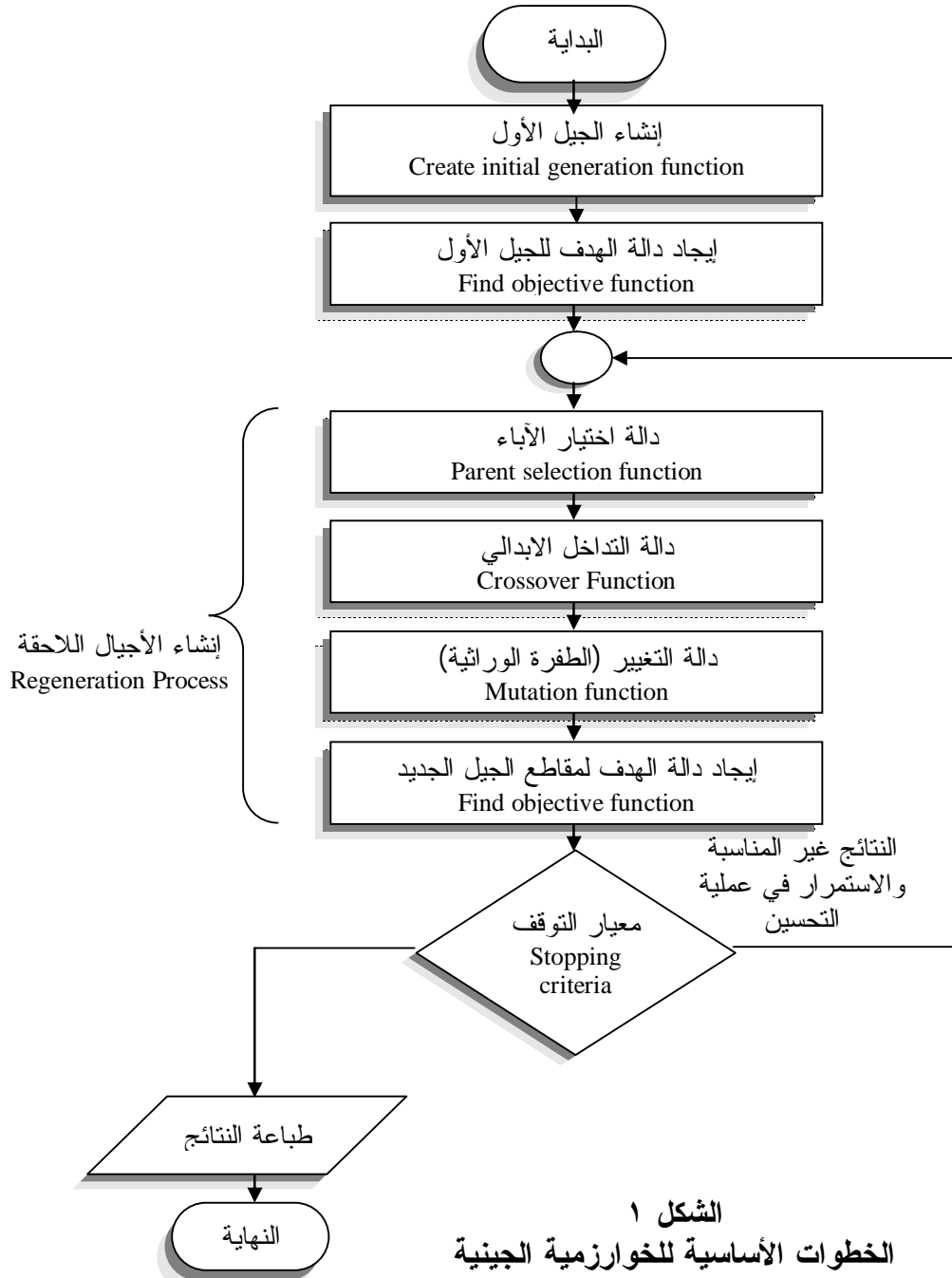
يتم تمثيل جدولة العمليات من خلال مخطط توضيحي يدعى لوحة كانت Gantt Chart، إذ يقيس أداء الخوارزمية ويحدد كفاءتها وأفضليتها من خلال طول الجدولة، إذ يعد مقياساً لمنفعة المعالج ومعياراً مهماً لتصميم الخوارزمية، (Thiebaut,1995)، كما يوضح الزمن المستغرق لإنهاء تنفيذ مجموعة من المهام على معالجات متعددة، ويستخدم الرمز (ϕ) للمعالج الساكن (العاطل) Idle Period (Arroyo,1996).

١-٥ خوارزمية جدولة العمليات المرتبطة Dependent Tasks Scheduling Algorithm

أشير في المصادر العلمية إلى استخدام خوارزميتين للجدولة الأولى جدولة المستوى الأعلى أولاً مع تقدير الزمن Highest Level First With Estimated Times (HLFET)، والثانية جدولة المستوى الأصغر أولاً مع تقدير الزمن Smallest Co-Level First With Estimated Time (SCFET)، وتحدد الأسبقية للعملية في الخوارزمية الأولى من خلال المستوى level، ويقصد بالمستوى مجموع زمن التنفيذ لكل العمليات على طول المسار الحرج Critical Path، أي من العملية المعينة إلى الأخيرة، وبعد ذلك ترتب المستويات للعمليات في نموذج البيان الحلقي المباشر ترتيباً تنازلياً. أما الخوارزمية الثانية تحدد الأسبقية للعملية من خلال Co-level يقصد به مجموع زمن التنفيذ لكل العمليات على طول المسار الحرج من عملية البدء Entry Node إلى العملية المعينة (Kowk et .al.,1999).

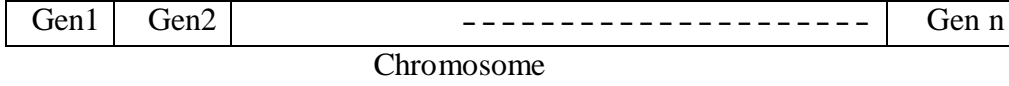
٢-١ طرائق العمل

إن طرائق العمل المتبعة في الدراسة الحالية لخصت بالخطوات التحضيرية للخوارزمية الجينية وكما موضحة في الشكل الآتي:



٢-٢ المجتمع Population

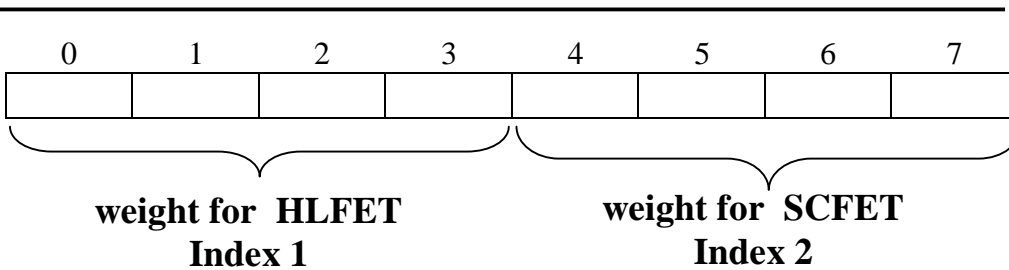
يتكون المجتمع من الأفراد Individuals ويحدد حجم المجتمع حسب المسألة المراد حلها. وتم تمثيل أبجدية الكروموسومات بالصيغة الثنائية وكل كروموسوم يتكون من عدد من الجينات genes.



n تمثل طول الكروموسوم (Auyeung *et. al.* 2003, Grajcar, 1999, Mahmood, Mitchell, 1998) 2000, Nossal, 2000, تم تحديد طول الكروموسوم في الدراسة الحالية بـ ٨ خانة 8bit، وعليه يكون حجم المجتمع للدراسة الحالية مساو ١٦ فرد ٢٤، ومن ثم يتم توليد مجموعة من الأرقام العشوائية Randomly Function لتمثيل القيم الابتدائية لجينات الأفراد في المجتمع.

٢-٣ التشفير Encoding Solution

يقصد به إيجاد تمثيل مناسب للحل ويعتمد على المسألة المراد حلها. هناك عدة طرائق للتشفير منها Binary Encoding التي اعتمدت في مسألة الدراسة، إذ يمتلك كل جين موقعا في الكروموسوم من خلال تكوين وزن weight، لكل من الخوارزميتين المذكورتين، والهدف منه إيجاد الوزن المثالي، ولهذا تم تمثيل الكروموسوم على شكل مجموعتين، وكل مجموعة تمثل بسلسلة من ٤ خانات 4bits، بحيث كل جزء من الكروموسوم يشير إلى خوارزمية من إحدى الخوارزميتين حسب ترتيبها في الكروموسوم. كما استخدمت في الدراسة الحالية مصفوفة ذات بعد واحد مكون من ٨ خانات 8bits، تمثل الكروموسوم، وضحت في الشكل ٢. إذ تمثل الخانات من (0-7) الوزن، أما تسلسل الجزء Index يشير إلى الوزن الخاص بالخوارزمية المعنية.



الشكل ٢

عملية تشفير المسألة

٢-٤ دالة اللياقة Fitness Function

تعد دالة اللياقة دالة هدف المسألة المراد حلها، وتم حسابها لكل مقطع من مقاطع الجيل، من خلال إيجاد الأسبقية لكل عقدة (عملية) في البيان الحلقي المباشر كما يأتي:

$$Pr(n)=(W_{hlfet}*P_{hlfet})+(W_{scfet}*P_{scfet})$$

إذ تمثل :

$pr(n)$ الأسبقية المخصصة لكل عملية في البيان الحلقي المباشر.

W_{hlfet} الوزن المخصص لخوارزمية HLFET الموضح في Index1 في الشكل ٢.

W_{scfet} الوزن المخصص لخوارزمية SCFET، الموضح في Index2 في الشكل ٢. P_{hlfet} الأسبقية المخصصة لكل عملية أو المستوى level حسب خوارزمية HLFET في DAG.

P_{scfet} الأسبقية المخصصة لكل عملية أو Co-level حسب خوارزمية SCFET في DAG.

٢-٥ الانتقاء أو الانتخاب Selection

وهي عملية اختيار الأباء Parent في المجتمع لأجل التزاوج لإنتاج الجيل الجديد، وان عملية الانتقاء لها دور في تطور الخوارزمية الجينية، وذلك لاختيار أفضل الأفراد في المجتمع، وقد تم اعتماد طريقة الانتخاب المستخدمة في البحث على وفق أسلوب عجلة الروليت Selection Wheel Roulette (Mitchell,1998). تعتمد هذه الطريقة على اختيار أفضل ما في المجتمع من أفراد، وذلك بعد حساب مجموع الجودة للمجتمع وتستخدم القيم الناتجة احتمالية لاختيار الأفراد في الأجيال اللاحقة (Goldberg,1989). كما في الشكل ٣، وأدناه الخوارزمية المقترحة لإجراء عملية الانتقاء لمسألة الدراسة الحالية، إذ تعدّ الأولوية محلياً في هذا المجال كما يأتي:

١. البداية.

٢. نفرض حجم المجتمع p .

٣. نفرض المجموع الكلي لدوال الهدف لجميع الكروموسومات $sumall$.

٤. $I=1$.

٥. توليد رقم عشوائي R randomly ضمن المدى $0 < R \leq sumall$.

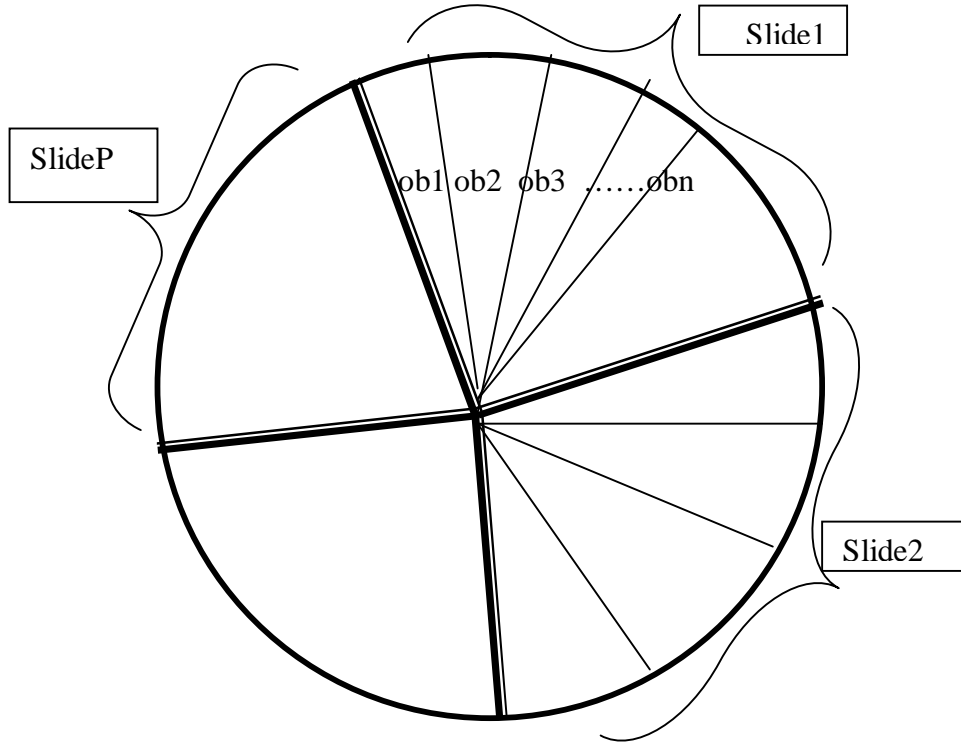
٦. $I=I+1$.

٧. إيجاد دالة الهدف لكل كروموسوم وليكن Ob .

٨. إيجاد مجموع دوال الهدف لكل كروموسوم $Slide=Slide+Ob$.

٩. إيجاد مجموع دوال هدف لكل كروموسوم مع مجموع دوال هدف للكروموسوم الذي يليه $Partsum=Partsum+Slide$.

١٠. مقارنة $Partsum > R$ إذا نعم اذهب إلى الخطوة ١١. لا اذهب إلى الخطوة ٦.
١١. توقف الخوارزمية، عندئذ رقم الكروموسوم يمثل الكروموسوم المنتخب في إنشاء الجيل اللاحق.
١٢. النهاية.

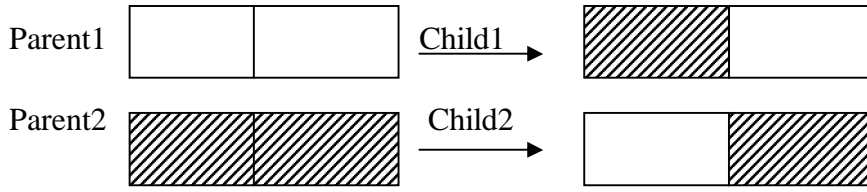


الشكل ٣
عجلة الروليت في مسألة الجدولة

٢-٦ التداخل الابدالي Crossover

تتمثل هذه العملية بإجراء تبديل بين قيم متقابلة من مقطعي الأبوين المنتخبين لغرض إنشاء المقطع الجديد، (Mahmood, 2000) (Auyeung et. al, 2003) (Mitchell, 1998) (Nossal, 2000) (Grajcar, 1999) (الحديثي وآخرون، ٢٠٠١).

وفي الدراسة الحالية تم استخدام طريقة التداخل الابدالي ذي نقطة القطع الواحدة Single Point Crossover، إذ يتم اختيار عشوائي لنقطة واحدة ضمن مقطعي الأبوين بعدها يتم إجراء تبديل بين قيم متقابلة من مقطعي الأبوين المنتخبين لغرض إنشاء المقطع الجديد كما موضح في الشكل ٤.



الشكل ٤
عملية التداخل

٧-٢ الطفرة Mutation

تعدّ هذه العملية تغييراً أو تبديلاً بين جينات محددة ضمن الكروموسوم الواحد، لإنشاء كروموسومات تعطي حلولاً جديدة في الجيل اللاحق (Mahmood, 1989) (Nossal, 2000) (Goldberg, 2000). أما في الدراسة الحالية فقد تم استخدام طريقة عكس قيمة الجين Bit Inversion الأكثر استخداماً في حالة التشفير الثنائي للكروموسوم، إذ يتم اختيار موقع معين في الكروموسوم وعكس قيمته (أي إذا كان 0 يصبح 1 وبالعكس 10001001 11001001).

كما تم استخدام احتمالية الطفرة بنسبة ٠,٢، وهنا يتم إنشاء كروموسومات تعطي حلولاً جديدة في الجيل اللاحق، التي لم يسبق تكوينها في الأجيال السابقة وذلك بهدف توسيع أفضل الحلول الممكنة، والمتمثلة بتكوين أكبر عدد من الكروموسومات المختلفة ضمن الجيل، وهذا يؤدي للوصول إلى الاقتراب من الحل الأمثل.

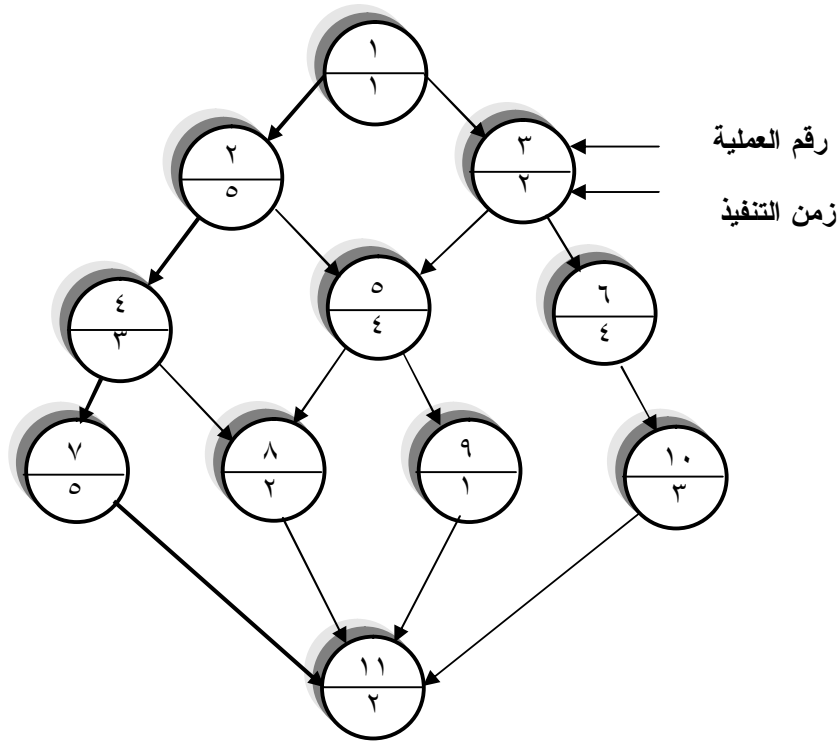
٨-٢ معيار التوقف Stop Criteria

يستمر إنشاء الأجيال المتعاقبة بهدف تحسين الحل أي جعله أكثر اقتراباً إلى الحل الأمثل، إلى أن يتحقق شرط التوقف الذي يعتمد على مقياس توقف الخوارزمية الجينية، ويختلف هذا المقياس باختلاف المسألة، ومن أهم مقاييس التوقف هو المقاييس التقليدية Classical Measures، الذي اعتمد في الدراسة الحالية ويعني التحديد المسبق لعدد الأجيال التي لم يحدث تكوينها أي تحسن للحل، إذ تتوقف الخوارزمية، إذا لم يطرأ أي تحسن للجيل الحالي بعد هذا العدد المحدد من الأجيال وفقاً لما أشير إليه في الدراسات السابقة ثابت، ٢٠٠٥ (Goldberg, 1989; Nossal, 2000 ;

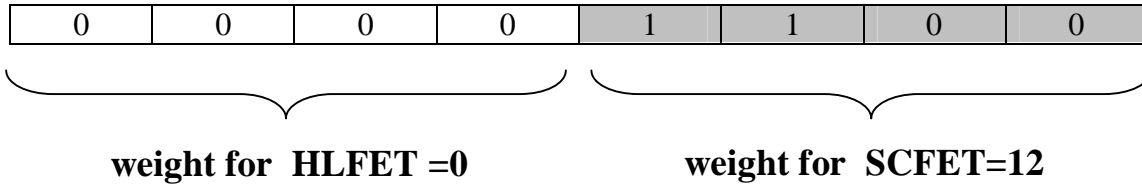
١-٣ النتائج Results

تطرق الجانب العملي في الدراسة الحالية إلى تنفيذ الخوارزمية الجينية المقترحة في جدول المعالجات المتعددة، مع عدم قطع العمليات، ويوضح الشكل ٥ أنموذج البيان الحلقي المباشر بمدى (١-١) و(١-٥) لعدد العمليات وزمن التنفيذ

على التوالي وتوفر معالجين . كما اعتمد عدد الأجيال وطول الكروموسوم وطول الجيل بـ ٧ و ٨ و ١٦ على التوالي واعدت برامج حاسوبية بلغة C++ لتنفيذ الخوارزمية المقترحة في الدراسة، وتم تنفيذها على أمثلة عديدة ومنها الإنموذج الموضح في الشكل ٥، أما الشكل ٦ فيبين مثالا لإيجاد أنموذج وزن تابع لأي من الخوارزميتين، فإذا كان الوزن المخصص للخوارزمية الأولى HLFET أكبر من الصفر والوزن المخصص للخوارزمية الثانية SCFET مساو للصفر فإن التنفيذ يتجه إلى الخوارزمية الأولى، في حين إذا كان الوزن المخصص للخوارزمتين أكبر من الصفر عندئذ يتم إيجاد الأسبقية من خلال حاصل جمع الوزن المخصص لكل من الخوارزميتين لأيجاد احتمالية أسبقية العمليات في التنفيذ. ويبين الجدول ١ عدد المهام وحساب المستوى Level و Co_level إزاء كل مهمة للمثال في الشكل ٥.



البيان الحلقي المباشر



الشكل ٦
أنموذج وزن تابع

الجدول ١

عدد المهام وحساب المستوى للخوارزميتين

TASK	1	2	3	4	5	6	7	8	9	10	11
LEVEL	16	15	11	10	8	9	5	4	3	5	2
CO_LEVEL	1	6	3	9	10	7	14	12	11	10	16

وطياً الجداول الخاصة بخطوات التنفيذ للمثال المذكور في الشكل ٥ للدراسة الحالية في الملحق، التي توضح عدد الحلول في كل جيل للاقتراب إلى الحل الأفضل.

ويوضح الجدول ٢ نتيجة تنفيذ الخوارزمية الجينية المقترحة في الدراسة الحالية على أنموذج البيان الحلقي المباشر في الشكل ٥ متمثلة في أنموذج لوحة كانت:

الجدول ٢

جدولة العمليات باستخدام الخوارزمية الجينية المقترحة

معالج ١	1	2	2	2	2	2	4	4	4	5	5	5	5	10	10	10	9	11	11
معالج ٢	φ	3	3	0	0	0	6	6	6	6	7	7	7	7	7	8	8	0	0

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

٣-٢ الاستنتاجات Conclusions

تمخضت الدراسة الحالية عن الاستنتاجات الآتية :

إن استخدام الخوارزمية الجينية لحل مسألة جدولة العمليات للمعالجات المتعددة ، نتج عنه مجموعة من الحلول ليس بحجم المجتمع ، لأن بعض الحلول الناتجة من تنفيذ الخوارزمية الجينية، لا يمكن أن تؤدي إلى حل المسألة بسبب اختراق

- علاقة الأسبقية بين العمليات (حلاقة الأب والابن)، ففي هذه الحالة تم إهمال تلك الحلول والتأكيد على الحلول المثلثة الناتجة إلى الحل الأفضل.
- ٢ تم اقتراح دالة لياقة جديدة وإضافة بعض التحويرات والإضافات إلى روتين عجلة الروليت، من أجل تطبيق مفهوم الخوارزمية الجينية والوصول إلى الحل الأفضل.
٣. أعطت الخوارزمية الجينية المقترحة، نتائج إيجابية خلال دمج خوارزميتي ن معينتين، مقارنة مع استخدام إحدى الخوارزميتين .

المراجع

أولاً - المراجع باللغة العربية

١. الحديثي، جبير جوامير وعلي هند عبد الرزاق محمد، تحليل التراميز الكتائية الخطية باستخدام الخوارزمية الجينية، مجلة أبحاث الحاسوب، المجلد الخامس، العدد الثاني، ٢٠٠١.
٢. ثابت، همسة معن محمد، بعض تطبيقات الخوارزمية الجينية في حل مسائل الأمثلية، رسالة ماجستير، كلية علوم الحاسبات والرياضيات، جامعة الموصل، ٢٠٠٥.

ثانياً - المراجع باللغة الأجنبية

1. Albala, A., Chan, W. Y., Chow, A., Fan, X. and Stephan, V., Data Structure and Algorithms Topic #30 :Directed A cyclic Graph, Class Notes for 308-251.School of Computer Science, Mc Gill University,1997.
<http://www.cs.mcgill.ca/~cs251/old courses/1997/topic30/>
2. Aronsson, P. and Fritzson, P. Multiprocessor Scheduling of Simulation Code from Modelica Models, 2nd International Modelica Conference, 2002.
<http://www.modelica.org/Conference2002/paper.s.html>.
3. Arroyo, D. O., STS A Simple Tool for Scheduling, School of Computer Science, McGill University, 1996.
<http://www.cs.mcgill.ca/~cs251/oldcourses/1997/>
4. Ayeung, A., Gondra, I. And Dai, H.K., Multi-Heuristic List Scheduling Genetic Algorithm for Task Scheduling. ACM.,2003.
<http://www.cs.okstate.edu/~wingha/sac03-ga.pdf>
5. Baskiyar, S., Scheduling DAGs on Message Passing m-Processor System, IEICE Trans. Inf. & Syst., E 83-D (7), 2000.
6. Collins, J. B., An Approach to Scheduling Task Graphs with Contention in communication, 2001.
<http://www.ait.nrl.navy.mil/pgmt-for-web/pdf/an-approach-to-scheduling.pdf>
7. Goldberg, D.E, Genetic Algorithms in Search Optimization and Machine Learning, Addison-Wesly. 1989.
8. Gonzalez, M. J., Deterministic Processor Scheduling, ACM. Computing Surveys, 9(3) , 1977.
9. Grajcar, M., Genetic List Scheduling for Scheduling and Allocation on a Loosely Coupled Heterogeneous Multiprocessor System, 1999.
<http://portal.acm.org/citation.cfm?id=309931&coll>.
10. Hou, E. S. H. ,Ansari, N., Ren, H., A Genetic Algorithm for Multiprocessor Scheduling, IEEE Transaction on parallel and Distributed systems,15(2) , 1994.

11. Kowk, Y. K. and Ahmad, I., Static Scheduling Algorithm for Allocating Directed Task Graphs to Multiprocessors, ACM Computing Surveys, Vol. 31, No. 4, 1999.
12. Mahmood ,A., A Hybrid Genetic Algorithm for Task Scheduling in Multiprocessor Real-Time Systems, 2000.
<http://www.ici.ro/ici/revista/sic2000-3/art05.html>
13. Mitchell, M., An Introduction to Genetic Algorithms, MIT press, London, 1998.
14. Nossal, R., An Evolutionary Approach to Multiprocessor Scheduling of Dependent Task, 2000.
[http:// www.citesser.ist.psu.edu/context/377390](http://www.citesser.ist.psu.edu/context/377390)
15. Peng,D.T., Shin, K.G and Abdel Zaher, T. F., Assignment and Scheduling Communicating Periodic Tasks in Distributed Real-Time Systems, IEEE. Transactions on Software Engineering ,23(12) ,1997.
16. Radulescu, A. and Gemund, A. J. C. V. Low-Cost Task Scheduling for Distributed-Memory Machines, IEEE Transactions on Parallel and Distributed Systems, Vol. 13, No. 6, 2002.
17. Thiebaut, D. Problem Decomposition on a multiprocessor Network, Parallel Programming in C for the Transputer, chap. IX, 1995.
<http://cs.smith.edu/~thiebaut/transputer/chapter9/chap9-1.html>