A Knowledge-Migration-Based Multi-Population Cultural Algorithm to Solve Job Shop Scheduling

Mohammad R. Raeesi N. and Ziad Kobti
School of Computer Science, University of Windsor, Windsor, ON, Canada N9B 3P4
raeesim@uwindsor.ca and kobti@uwindsor.ca

Abstract
In this article, a multipopulation Cultural Algorithm (MP-CA) is proposed to solve Job Shop Scheduling Problems (JSSP). The idea of using multiple populations in a Cultural Algorithm is implemented for the first time in JSSP. The proposed method divides the whole population into a number of sub-populations. On each sub-population, a local CA is applied which includes its own population space as well as belief space. The local CAs use Evolutionary Programming (EP) to evolve their populations, and moreover they incorporate a local search approach to speed up their convergence rates. The local CAs communicate with each other using knowledge migration which is a novel concept in CA. The proposed method extracts two types of knowledge including normative and topographic knowledge and uses the extracted knowledge to guide the evolutionary process to generate better solutions. The MP-CA is evaluated using a well-known benchmark. The results show that the MP-CA outperforms some of the existing methods by offering better solutions as well as better convergence rates, and produces competitive solutions when compared to the state-of-the-art methods used to deal with JSSPs.

Introduction
Job Shop Scheduling Problem (JSSP) is a combinatorial optimization problem which is well-known in different areas, specially manufacturing systems. JSSP is the task of scheduling different operations to be processed on different machines. The main goal of this type of problems is minimizing the maximum completion time of all the operations. The maximum completion time of a schedule is also called makespan. JSSP is still an open problem. It is proved that the job shop scheduling systems with more than two machines are NP-complete (Garey, Johnson, and Sethi 1976), which means that there is no method capable to find the best solution for all the scheduling problems in an acceptable time.

There are various types of algorithms proposed to deal with JSSPs including heuristic approaches, meta-heuristic methods and Evolutionary Algorithms (EAs). In the area of Evolutionary Computation, there are various algorithms with different versions proposed to solve JSSPs including Genetic Algorithm (GA), hybrid GA, Ant Colony Optimization (ACO), Memetic Algorithm (MA), and Cultural Algorithm (CA). Each method has its own strengths and weaknesses. However, combinations of different types of algorithms work better.

In this article, a new CA is proposed to solve JSSPs. The proposed method incorporate a multipopulation design which is called multipopulation CA (MP-CA). In this design, there are a number of sub-populations incorporating local CAs to cooperate with each other to generate better solutions. The sub-populations communicate with each other by exchanging their extracted knowledge every predefined number of generations.

The structure of this article is as follows. Section present the existing Evolutionary Algorithm introduced in the area of Job Shop Scheduling, which is followed by the definition of the classical JSSPs in Section . Section describes the proposed MP-CA in details, and Section shows the results of evaluating the MP-CA. Finally, the conclusions are represented in section.

Related Work
The application of EAs in JSSPs is first introduced by Lawrence (1985) as a GA. Hasan et al. (2008) combined a GA with different priority rules including Partial Reordering, Gap Reduction, and Restricted Swapping. An ACO method is proposed by Wang, Cao, and Dai (2005), and recently we proposed a MA to solve JSSPs (Raeesi N. and Kobti 2011).

Becerra and Coello (2005) introduced the application of CA in JSSP for the first time. CA, developed by Reynolds (1994), is an EA which extracts knowledge to improve its search mechanism as well as its convergence rate. CA consists of population space and belief space. Population space contains individuals which are evolving to generate the optimal solution. The knowledge is extracted from the best individuals every generation. The extracted knowledge which is also called belief is recorded in the belief space to be used in the next generations to direct the evolutionary process. The link from population space to belief space which sends the best individuals in order to update the extracted knowledge is called acceptance function, and the link from belief space...
to population space which sends the updated knowledge to guide the evolution is called influence function.

The CA proposed by Becerra and Coello (2005) incorporates Evolutionary Programming (EP) for population space evolution which only uses the mutation operator to generate the offspring population. The belief space of their proposed method only records the best individual as the situational knowledge. An improved version of the first CA (Becerra and Coello 2005) in JSSP is proposed by Corts, Becerra, and Coello (2007). The newer method is similar to the first version with the main differences being in the implementation of the mutation operator and influence function.

Ho and Tay (2004) proposed a CA to solve Flexible Job Shop Scheduling which is called GENACE. Like (Becerra and Coello 2005), they used EP and situational knowledge. The next version of GENACE is also proposed by Ho and Tay (2005) which is called LEGA.

All of the proposed CAs in JSSP use a single population in their population space. Our proposed method would be the first attempt to use a CA with multiple populations which is called MP-CA. However, there are a number of CAs in different fields which are using multiple populations. Digalakis and Margaritis (2002) introduced the multipopulation concept in CA for the first time by proposing a master-slave design. The master processor generates the initial population and manages it, while the slave processors execute different CAs on different sub-populations. The communication among sub-populations is implemented using Message-Passing Interface (MPI). The latter also used the situational knowledge.

One of the main issues in MP-CAs is the communication among the sub-populations. Most of the existing methods consider exchanging the best individuals as the communication process. But there is a more powerful strategy which exchanges the extracted knowledge instead of the best individuals. The new strategy is called knowledge migration which was first proposed by Guo et al. (2009). The extracted knowledge can incorporate more useful information about the previous generations to be used to direct the evolutionary process. Consequently, it would be more effective to use knowledge migration as the sub-populations communication mechanism.

**Problem Definition**

Classical JSSP is defined as a process of assigning different jobs to be processed on different machines (Baker 1974). There are $M$ machines denoted by $m_k$ and $N$ jobs denoted by $J_i$, where $k$ is the machine index and $i$ is the job index. Each job is defined by a fixed sequence of operations. Each operation is denoted by $O_{ij}$ where $i$ is the job index and $j$ is the operation index in that job. Each operation can be processed on only one machine in a known processing time. In other words, the route of operations of each job through the given machines is predefined.

In classical JSSP, there are some assumptions which are as follows: All the jobs are available at the starting point which are independent from each other; the machine set up time and part movement time between machines are negligible; each job is processed only one time on each machine; the machines can process only one operation at a time which cannot be interrupted; and there is no due date for the jobs.

Table 1 presents a sample classical JSSP which contains 3 jobs to be processed on 3 machines. This table shows the applicable machine with the corresponding processing time for each operation. The third operation of the second job $O_{23}$, for example, is applicable on the third machine, and it needs 2 time units to be processed. A sample schedule for this example is shown in Figure 1.

It should be mentioned that we use the active schedule concept defined by Croce, Tadei, and Volta (1995), and used by Hasan et al. (2008) as Gap Reduction rule and by Becerra and Coello (2005) as permissible left shift.

### Proposed Cultural Algorithm

In this article, a MP-CA is proposed to deal with JSSPs. The proposed method is the first attempt to incorporate multiple populations in JSSP. In this method, the whole population is divided into a number of sub-populations. Each sub-population incorporates a local CA which contains its own belief space. The local CAs are cooperating to find the optimal solutions. They are communicating with each other by exchanging their own extracted knowledge which is called knowledge migration.

The architecture of the MP-CA is represented in Figure 2. As it is illustrated in the figure, like other CAs, each local CA has its own population space and belief space (Reynolds 1994).

Moreover, the overall framework of the proposed MP-CA is represented in Figure 3. The number of sub-populations, the number of iterations which is the termination criterion, and the migration frequency are denoted by the `SubPopulationsNo`, `IterationNo`, and `ExchangeRate` parameters, respectively. The parameters values which are used in our experiments are presented in Table 2.

### Population Space

Population space is a set of individuals which are evolving using EP. Our EP uses the selection and mutation as regular EP, and incorporates a local search heuristic to speed up the convergence rate.
PROCEDURE: MP-CA Framework
INPUT: Test Problems and Algorithm Parameters
OUTPUT: Optimal or Near-Optimal Schedules

Generate SubPopulationsNo sub-populations.
FOR (IterationNo)
    FOR (each subpopulation)
        Evaluate all individuals and sort them.
        Apply mutation and local search method to generate offspring population.
        Update belief space.
    END
    IF (IterationNo mod ExchangeRate = 0)
        Exchange knowledge.
    END
END
Output the best found individual so far.

Figure 3: MP-CA Framework

Chromosome Representation Chromosome representation is one of the main characteristics of an EA. In the literature, there are various chromosome representations, with nine of them described by Cheng, Gen, and Tsujimura (1996). Recently, we introduced Machine Operation Lists (MOL) representation (Raeesi N. and Kobti 2011). MOL is an extended version of preference list-based representation such that it adds the concept of fixed list, the operation sequence of which cannot be changed unless due to the permissible left shift. In (Raeesi N. and Kobti 2011), we showed that MOL representation outperforms preference list-based representation by yielding better solutions.

In our proposed algorithm, we incorporate the MOL representation. MOL considers a list of operations for each machine determining the sequence of operations to be processed on that machine. Because each machine only processes one operation of each job, the operations in the list can be denoted by their job indices. For example, the sample schedule illustrated in Figure 1 is represented as follows.

\[
\{(1, 2, 3), (1, 2, 3), (2, 3, 1)\}
\]

Evolutionary Programming The proposed method uses EP to evolve the population space. The EP incorporates only the mutation operator as a genetic operator. The mutation operator uses the knowledge recorded in belief space to influence the direction of evolution. The EP applies the mutation operator on all the individuals to generate new ones. In our proposed MP-CA, there are two mutation operators which will be described in details in subsection.

Local Search After generating the offspring population, a number of best individuals will be selected to be investigated by a local search heuristic. We use the same local search method as we used before (Raeesi N. and Kobti 2011) which is compatible with MOL representation. The search method reassigns all the operations of a randomly selected job to decrease the makespan. It has been shown by the authors that the time complexity of the local search method is negligible compared to the number of fitness evaluations in each generation.

Belief Space Each local CA has its own belief space which space gets updated every generation using the acceptance function. The acceptance function passes a number of best individuals from the sub-population to the belief space. We consider the top 20 percent of the individuals in our implementation. The belief space extracts both normative and topographic knowledge from the individuals, and updates its own belief by the extracted knowledge. The knowledge stored in the belief space is incorporated to direct the mutation operator using the influence function.

The local CAs migrate their knowledge to each other to improve the search exploration in different sub-populations. The knowledge migration occurs every predefined number of generations. In knowledge migration, the sub-population which finds the best individual so far sends its own knowledge to others. The sub-populations which receive the migrated knowledge replace their own knowledge with migrated one.
Normative Knowledge We use normative knowledge to improve the search exploration of our proposed method. Since normative knowledge records the feasible search space, it is used to explore all the feasible search space uniformly. In our method, we consider the position of each operation in its corresponding machine list as the search variable. So we have $M \times N$ variables for a system with $M$ machines and $N$ jobs. For each operation $O_{ij}$, there are a lower position $L_{O_{ij}}$ and an upper one $U_{O_{ij}}$. The normative knowledge is initialized using the best individual as follows:

$$
L_{O_{ij}} = P_{O_{ij}} - 0.5 \\
U_{O_{ij}} = P_{O_{ij}} + 0.5
$$

where $P_{O_{ij}}$ denotes the position of operation $O_{ij}$ in its corresponding machine list in the best individual at the first iteration. The 0.5 is used to provide a range of 1 for each position. The normative knowledge gets updated every generation using the following function:

$$
L_{O_{ij}} = \min (L_{O_{ij}}, P_{O_{ij}} - 0.5) \\
U_{O_{ij}} = \max (U_{O_{ij}}, P_{O_{ij}} + 0.5)
$$

where $P_{O_{ij}}$ denotes the position of operation $O_{ij}$ in its corresponding machine list in that individual.

Topographic Knowledge We use the topographic knowledge to exploit certain regions which have more individuals. The topographic knowledge for each operation keeps the record of all positions used by all the individuals passed through the acceptance function. For each operation, there is a list of $N$ available positions where $N$ is the number of jobs. The topographic knowledge counts the number of occurrences for each position. Since the greater number of occurrences determines the existence of more good individuals in that region, the topographic knowledge is used to exploit those regions. Like the normative knowledge, topographic knowledge is also updated every generation, and it has its own mutation operator which influences the evolution to exploit certain regions.

Influence Function Each type of knowledge has its own mutation operator, one mutation for normative knowledge and another one based on topographic knowledge. To generate a new individual, each machine list of an existing individual is mutated using one of both mutation operators which is selected randomly with the same chance.

The normative-knowledge-based mutation operator calculates a position for each operation using the following formula, and then it sequences the operations in their operation lists based on their calculated position values.

$$
P_{O_{ij}} = \begin{cases} 
    P_{O_{ij}} + R \times (U_{O_{ij}} - L_{O_{ij}}) & P_{O_{ij}} < L_{O_{ij}}, \\
    P_{O_{ij}} - R \times (U_{O_{ij}} - L_{O_{ij}}) & P_{O_{ij}} > U_{O_{ij}}, \\
    P_{O_{ij}} + G & \text{otherwise},
\end{cases}
$$

where $R$ is a random number uniformly distributed between 0 and 1, and $G$ is a random number in a normal distribution with mean 0 and variance 1.

In topographic knowledge, there is a list of position occurrences for each operation such that we need to use the position with greater occurrence to generate new individuals. Here we use roulette-wheel selection strategy to choose the new position for each operation.

### Results

The proposed algorithm is implemented and evaluated using the java programming language version 1.6.0.18 on a system with Intel(R) Core(TM)2Quad 2.50GHz CPU and 8.00GB RAM. Table 2 presents the parameters used in our experiments which are adjusted using extensive experiments. In our experiments, we generate 7 sub-populations with 142 individuals, approximately 1000 individuals in total. We run the method for 200 iterations such that the knowledge migration occurs every 20 iterations.

In our experiments, we consider a well-known benchmark (Lawrence 1984) for classical JSSP. The benchmark consists of 40 different problems with different size. All the experiments are done 10 times independently.

The experiment results show that the proposed MP-CA finds the optimal solution for 28 test problems out of 40. To show the performance of the proposed method, we compare our method with another CA recently published by Corts, Becerra, and Coello (2007). The authors claim that their method finds the optimal solution for 26 test problems. However, our MP-CA outperforms their method by finding the best solution for 2 more test problems as well as by offering better statistical results (lower median and worst solutions) for those 26 test problems which means that the convergence rate of our proposed method is better than theirs. Moreover, for the remaining test problems, our method offers better solutions comparing to theirs, for all the best, median and worst solutions. Table 3 shows three sample results of our proposed MP-CA compared to the result of Corts, Becerra, and Coello (2007). The results show that incorporating multi-population and using knowledge exchange offers better solution and improves the convergence rate.

To have a fair comparison, we compare our results with their results for 200,000 fitness evaluations, not for more than 2,000,000 evaluations (Corts, Becerra, and Coello 2007). The maximum number of evaluations in our method is 200,000 which is 1000 evaluations in 200 iterations.

1Please refer to http://cs.uwindsor.ca/~raeesim/Flairs-25/Statistical-Analysis.pdf to see the complete results.
Furthermore, different Evolutionary Algorithms are considered to be compared with our MP-CA including a hybrid GA (Goncalves, de Magalhaes Mendes, and Resende 2002), our published MA (Raeesi N. and Kobti 2011), and the CA (Corts, Becerra, and Coello 2007). All the four algorithms find the optimal solution for certain 26 test problems including la01-la19, la23, and la30-la35. Table 4 represents the results for the remaining 14 test problems. The Error Rate (ER) is considered to be able to compare the results over all the 14 problems which is defined as follows.

\[ ER = \frac{C - LB}{LB} \times 100\% \]

where \( LB \) is the best-known solution, and \( C \) is the best solution found by the algorithms. The ER values are represented in brackets in Table 4. To compare the 4 algorithms on the 14 problems, we used the non-parametric procedure incorporated by Garca et al. (2009) with the same levels of significance as theirs which are \( \alpha = 0.05 \) and \( \alpha = 0.10 \). To do so, first the algorithms are ranked using Friedman’s test. The \( p \)-value for this test is less than 0.0001 which is also less than both significance levels; this means there are significant differences between the compared algorithms. Then, the algorithm with the minimum average ranking is selected as the control algorithm which is our MP-CA. Afterwards, the critical differences (CD) for control algorithm MP-CA are calculated using Bonferroni-Dunns test which are:

\[ CD = 1.17 \quad \text{for} \quad \alpha = 0.05 \]
\[ CD = 1.09 \quad \text{for} \quad \alpha = 0.10 \]

Now, the algorithms with the average rankings greater than the summation of the control algorithm ranking and CDs are considered as the algorithms with worse performance than the control algorithm. In other words, the summation of the average ranking of the MP-CA and the CD for each significance level is a threshold to find the worse algorithms. So the thresholds are as follows.

\[ \text{Threshold} = 2.53 \quad \text{for} \quad \alpha = 0.05 \]
\[ \text{Threshold} = 2.45 \quad \text{for} \quad \alpha = 0.10 \]

Table 4: Comparison among Different Evolutionary Algorithms proposed recently to solve JSSP on LA Benchmark

<table>
<thead>
<tr>
<th>Problem</th>
<th>Size</th>
<th>Best Known</th>
<th>Hybrid GA</th>
<th>MA</th>
<th>CA</th>
<th>Proposed MP-CA</th>
</tr>
</thead>
<tbody>
<tr>
<td>LA20</td>
<td>10×10</td>
<td>902</td>
<td>907 (0.55%)</td>
<td>907 (0.55%)</td>
<td>907 (0.55%)</td>
<td>902 (0.00%)</td>
</tr>
<tr>
<td>LA21</td>
<td>15×10</td>
<td>1046</td>
<td>1046 (0.00%)</td>
<td>1057 (1.05%)</td>
<td>1059 (1.24%)</td>
<td>1048 (0.19%)</td>
</tr>
<tr>
<td>LA22</td>
<td>15×10</td>
<td>927</td>
<td>935 (0.86%)</td>
<td>935 (0.86%)</td>
<td>947 (2.16%)</td>
<td>932 (0.54%)</td>
</tr>
<tr>
<td>LA24</td>
<td>15×10</td>
<td>935</td>
<td>953 (1.93%)</td>
<td>944 (0.96%)</td>
<td>950 (1.60%)</td>
<td>943 (0.86%)</td>
</tr>
<tr>
<td>LA25</td>
<td>15×10</td>
<td>977</td>
<td>986 (0.92%)</td>
<td>983 (0.61%)</td>
<td>998 (2.15%)</td>
<td>983 (0.61%)</td>
</tr>
<tr>
<td>LA26</td>
<td>20×10</td>
<td>1218</td>
<td>1218 (0.00%)</td>
<td>1218 (0.00%)</td>
<td>1219 (0.08%)</td>
<td>1218 (0.00%)</td>
</tr>
<tr>
<td>LA27</td>
<td>20×10</td>
<td>1235</td>
<td>1236 (1.70%)</td>
<td>1269 (2.75%)</td>
<td>1279 (3.56%)</td>
<td>1264 (2.35%)</td>
</tr>
<tr>
<td>LA28</td>
<td>20×10</td>
<td>1216</td>
<td>1232 (1.32%)</td>
<td>1223 (0.58%)</td>
<td>1236 (1.64%)</td>
<td>1219 (0.25%)</td>
</tr>
<tr>
<td>LA29</td>
<td>20×10</td>
<td>1157</td>
<td>1196 (3.37%)</td>
<td>1191 (2.94%)</td>
<td>1219 (5.36%)</td>
<td>1182 (2.16%)</td>
</tr>
<tr>
<td>LA36</td>
<td>15×15</td>
<td>1268</td>
<td>1279 (0.87%)</td>
<td>1281 (1.03%)</td>
<td>1296 (2.21%)</td>
<td>1274 (0.47%)</td>
</tr>
<tr>
<td>LA37</td>
<td>15×15</td>
<td>1397</td>
<td>1408 (0.79%)</td>
<td>1429 (2.29%)</td>
<td>1416 (1.36%)</td>
<td>1415 (1.29%)</td>
</tr>
<tr>
<td>LA38</td>
<td>15×15</td>
<td>1196</td>
<td>1219 (1.92%)</td>
<td>1208 (1.00%)</td>
<td>1231 (2.93%)</td>
<td>1202 (0.50%)</td>
</tr>
<tr>
<td>LA39</td>
<td>15×15</td>
<td>1233</td>
<td>1246 (1.05%)</td>
<td>1248 (1.22%)</td>
<td>1269 (2.92%)</td>
<td>1240 (0.57%)</td>
</tr>
<tr>
<td>LA40</td>
<td>15×15</td>
<td>1222</td>
<td>1241 (1.55%)</td>
<td>1234 (0.98%)</td>
<td>1256 (2.78%)</td>
<td>1228 (0.49%)</td>
</tr>
</tbody>
</table>

| Average Error Rate | 1.20% | 1.20% | 2.18% | 0.73% |
| Average Ranking (Friedman) | 2.39 | 2.46 | 3.79 | 1.36 |

Figure 4: The Graphical Representation of Statistical Result of Friedman’s test and Bonferroni-Dunn’s method

Furthermore, different Evolutionary Algorithms are considered to be compared with our MP-CA including a hybrid GA (Goncalves, de Magalhaes Mendes, and Resende 2002), our published MA (Raeesi N. and Kobti 2011), and the CA (Corts, Becerra, and Coello 2007). All the four algorithms find the optimal solution for certain 26 test problems including la01-la19, la23, and la30-la35. Table 4 represents the results for the remaining 14 test problems. The Error Rate (ER) is considered to be able to compare the results over all the 14 problems which is defined as follows.

\[ ER = \frac{C - LB}{LB} \times 100\% \]

where \( LB \) is the best-known solution, and \( C \) is the best solution found by the algorithms. The ER values are represented in brackets in Table 4. To compare the 4 algorithms on the 14 problems, we used the non-parametric procedure incorporated by Garca et al. (2009) with the same levels of significance as theirs which are \( \alpha = 0.05 \) and \( \alpha = 0.10 \). To do so, first the algorithms are ranked using Friedman’s test. The \( p \)-value for this test is less than 0.0001 which is also less than both significance levels; this means there are significant differences between the compared algorithms. Then, the algorithm with the minimum average ranking is selected as the control algorithm which is our MP-CA. Afterwards, the critical differences (CD) for control algorithm MP-CA are calculated using Bonferroni-Dunns test which are:

\[ CD = 1.17 \quad \text{for} \quad \alpha = 0.05 \]
\[ CD = 1.09 \quad \text{for} \quad \alpha = 0.10 \]

Now, the algorithms with the average rankings greater than the summation of the control algorithm ranking and CDs are considered as the algorithms with worse performance than the control algorithm. In other words, the summation of the average ranking of the MP-CA and the CD for each significance level is a threshold to find the worse algorithms. So the thresholds are as follows.

\[ \text{Threshold} = 2.53 \quad \text{for} \quad \alpha = 0.05 \]
\[ \text{Threshold} = 2.45 \quad \text{for} \quad \alpha = 0.10 \]
The proposed method is also applicable for other types of JSSPs, but since it uses the MOL representation and this representation does not cover other JSSP types, we need to use another representation or provide an extension for MOL, in future work.

**Conclusions**

In this article, we propose a MP-CA to solve JSSP which is a combinatorial optimization problem proved to be NP-Complete. This article introduces application of the MP-CA in JSSPs for the first time. Using multipopulation we improve the search exploration. The proposed method incorporates local CAs on the sub-populations cooperating to find the optimal solution. In addition, it uses knowledge migration among sub-populations as the communication link. In the proposed method, we consider the extraction of normative and topographic knowledge.

The experiments show that the proposed MP-CA offers better solutions. Moreover, it improves the convergence rate. The comparison of the MP-CA with the traditional CA shows that using multiple populations as well as incorporating knowledge migration enhance the search space exploration and help the algorithm to avoid trapping into local optimal solutions.

The statistical comparisons reveal that the proposed method outperforms some existing methods very well and offers competitive solutions compared to other state-of-the-art methods. Furthermore, while the statistical comparisons do not show the significant differences between the proposed MP-CA and those methods, the results show that it outperforms others by offering lower error rates.

The proposed method is also applicable for other types of JSSPs, but since it uses the MOL representation and this representation does not cover other JSSP types, we need to use another representation or provide an extension for MOL, in future work.

**Acknowledgments**

This work is made possible by a grant from the National Science Foundation and NSERC Discovery No. 327482.

**References**


