# A Fuzzy Genetic Algorithm for Real-World Job Shop Scheduling

Carole Fayad and Sanja Petrovic

School of Computer Science and Information Technology
University of Nottingham, Jubilee Campus, Wollaton Road, Nottingham NG8 1BB UK
{cxf,sxp}@cs.nott.ac.uk
http://www.cs.nott.ac.uk/~cxf,~sxp

**Abstract.** In this paper, a multi-objective genetic algorithm is proposed to deal with a real-world fuzzy job shop scheduling problem. Fuzzy sets are used to model uncertain due dates and processing times of jobs. The objectives considered are average tardiness and the number of tardy jobs. Fuzzy sets are used to represent satisfaction grades for the objectives taking into consideration the preferences of the decision maker. A genetic algorithm is developed to search for the solution with maximum satisfaction grades for the objectives. The developed algorithm is tested on real-world data from a printing company. The experiments include different aggregation operators for combining the objectives.

**Keywords:** job shop scheduling, fuzzy logic and fuzzy sets, genetic algorithms

## 1 Introduction

Scheduling is defined as the problem of allocation of machines over time to competing jobs [1]. The *m* x *n* job shop scheduling problem denotes a problem where a set of *n* jobs has to be processed on a set of *m* machines. Each job consists of a chain of operations, each of which requires a specified processing time on a specific machine.

Although production scheduling has attracted research interest of operational research and artificial intelligence community for decades, there still remains a gap between the academic research and real world problems. In the light of the drive to bridge this gap, we consider in this work a real-world application and focus on two aspects in particular, namely uncertainty inherent in scheduling and multi-objective scheduling.

Scheduling parameters are not always precise due to both human and machine resource factors [2]. As a result, classical approaches, within a deterministic scheduling theory, relying on precise data might not be suitable for representation of uncertain scenarios [3]. Consequently, the deterministic scheduling models and algorithms have been extended to the stochastic case, mainly to models that assume that processing times are random variables with specified probability distributions [1]. However, probabilistic characteristics of processing times and other scheduling parameters are often not available in manufacturing environments. That is the reason why standard stochastic methods based on probability are not appropriate to use. Fuzzy sets and fuzzy logic have been increasingly used to capture and process imprecise and uncertain information [4,5]. For example, Chanas et al. consider minimization of maximum lateness of jobs in a single machine scheduling problem [6] and minimization of maximal expected value of the fuzzy tardiness and minimization of the expected value of maximal fuzzy tardiness in a two-single machine scheduling problem [7]. Itoh et al. [8] represent the execution times and due dates as fuzzy sets to minimize the number of tardy jobs.

Real-world problems require the decision maker to consider multiple objectives prior to arriving at a decision [9, 10]. Recent years have seen an increasing number of publications handling multi-objective job shop scheduling problems [11]. A survey on available multi-objective literature is given in [9] and a review on most recent evolutionary algorithms for solving multi-objective problems is given in [12].

This paper deals with a real-world job shop scheduling problem faced by Sherwood Press, a printing company based in Nottingham, UK. It is a due date driven client-oriented company. This is reflected in the objectives of minimizing average tardiness and number of tardy jobs. The durations of operations on the machines, especially the ones involving humans are not known precisely. Also, due dates are rigid and can be relaxed up to a certain extent. Fuzzy sets are used to model imprecise scheduling parameters and also to represent satisfaction grades of each objective. A number of genetic algorithms with different components are developed and tested on real-world data.

The paper is organized as follows. In Section 2, the fuzzy job shop problem is introduced together with the objectives and constraints; then, the real-world problem at Sherwood Press is described. The fuzzy genetic algorithm with the fitness function, which aggregates multiple objectives, is given in Section 3. Experimental results obtained on real-world data are discussed in Section 4 followed by conclusions in Section 5.

## 2 Problem Statement

In the job shop problem considered in this research, $n$ jobs $J_1,...,J_n$ with given release dates $r_1,...,r_n$ and due dates $d_1,...,d_n$ have to be scheduled on a set of $m$ machines $M_1,...,M_m$. Each job $J_j$ $j$=1,…,$n$ consists of a chain of operations determined by a process plan that specifies precedence constraints imposed on the operations. Each operation is represented as an ordered pair $(i, j)$, $i$=1,..,$m$ and its processing time is denoted by $p_{ij}$.

The task is to find a sequence of operations of $n$ jobs on each of $m$ machines with the following objectives:

(1) to minimize the average tardiness $C_{AT}$:

$$C_{AT} = \frac{1}{n}\sum_{j=1}^{n} T_j\,;\quad T_j = \max\{0, C_j - d_j\};\quad j = 1,\ldots,n \text{ and } C_j \text{ is the comple-} \tag{1}$$

tion time of job $J_j$ on the last machine on which it requires processing.

(2) to minimize the number of tardy jobs $C_{NT}$:

$$C_{NT} = \sum_{j=1}^{n} u_j\,;\, u_j = 1 \text{ if } T_j > 0, \text{ otherwise } u_j = 0 \tag{2}$$

The resulting schedule is subject to the following constraints: (1) the precedence constraints which serve in ensuring that the processing sequence of operations of each job conforms to the predetermined order, (2) the capacity constraints which ensure that a machine processes only one job at a time and its processing cannot be interrupted.

Any solution satisfying all above listed constraints is called a feasible schedule.

### 2.1 A Real-World Job Shop Problem

In this section, a job shop problem faced by a printing company, Sherwood Press Ltd, is described. There exist 18 machines in the shopfloor, which are grouped within 7 work centers: Printing, Cutting, Folding, Card-inserting, Embossing and Debossing, Gathering, Stitching and Trimming and Packaging. Jobs are processed in the work centres, following a pre-determined order. A 'Job Bag' is assigned to each order to record the quantity in units to be produced and the 'Promised delivery date' of the order (referred to as due date).

Processing times of jobs are uncertain due to both machine and human factors. Consequently, the completion time of each job is uncertain. In addition, as it is not always possible to construct a schedule in which all jobs are completed before their due dates, some of the jobs may be tardy. The model should allow the decision maker to express his/her preference to the tardiness of each job. Fuzzy sets are used to model uncertain processing times of jobs and the decision maker's preference to the tardiness of each job.

Unlike a conventional crisp set, which enforces either membership or non-membership of an object in a set, a fuzzy set allows grades of membership in the set. A fuzzy set $\tilde{A}$ is defined by a membership function $\mu_{\tilde{A}}(x)$ which assigns to each object *x* in the universe of discourse *X*, a value representing its grade of membership in this fuzzy set [13]:

$$\mu_{\tilde{A}}(x) : X \to [0,1] \tag{3}$$

A variety of shapes can be used for memberships such as triangular, trapezoidal, bell curves and s-curves [13]. Conventionally, the choice of the shape is subjective and allows the decision maker to express his/her preferences.

The ‘estimation’ of processing time of each operation is obtained taking into consideration the nature of the machines in use. While some machines are automated and can be operated at different speeds, others are staff-operated and therefore the processing times are staff-dependent. Uncertain processing times $\tilde{p}_{ij}$ are modeled by triangular membership functions represented by a triplet ( $p_{ij}^1, p_{ij}^2, p_{ij}^3$), where $p_{ij}^1$ and $p_{ij}^3$ are lower and upper bounds of the processing time while $p_{ij}^2$ is so-called modal point [13]. An example of fuzzy processing time is shown in Fig.1. A trapezoidal fuzzy set (TrFS) is used to model the due date $\tilde{d}_j$ of each job, represented by a doublet ( $d_j^1, d_j^2$ ), where $d_j^1$ is the crisp due date and the upper bound $d_j^2$ of the trapezoid exceeds $d_j^1$ by 10%, following the policy of the company. An example of a fuzzy due date is given in Fig.2.

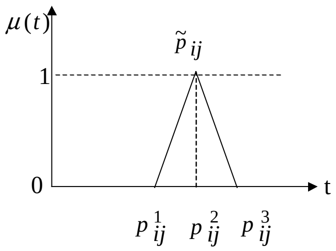


**Fig. 1.** Fuzzy processing time

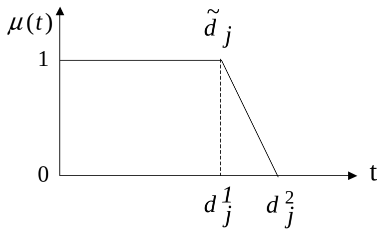


**Fig. 2.** Fuzzy due date

## 3 A Fuzzy Genetic Algorithm for the Job Shop Scheduling Problem

A genetic algorithm (GA) is an iterative search procedure widely used in solving optimization problems, motivated by biological models of evolution [14]. In each iteration, a population of candidate solutions is maintained. Genetic operators such as mutation and crossover are applied to evolve the solutions and to find the good solutions that have a high probability to survive for the next iteration.

The main characteristics of the fuzzy GA developed for job shop scheduling are described below:

- Chromosome: Each chromosome is made of two sub-chromosomes of length *m*, named machines sub-chromosome and dispatching rules sub-chromosome. The genes of the first sub-chromosome contain machines, while genes of the second sub-chromosome contain the dispatching rules to be used for sequencing operations on the corresponding machines.

- Initialization: The machine sub-chromosome is filled in by randomly choosing machines $i$, $i=1,..,m$. The initialization of the dispatching rules sub-chromosome is done by choosing randomly one among the following four rules: Early Due Date First, Shortest Processing Time First, Longest Processing Time First, Longest Remaining Processing Time First.
- Crossover operator: This operator is applied with a certain probability in order to combine genes from two parent sub-chromosomes and create new children sub-chromosomes, taking care that machines are not duplicated in the machine sub-chromosome [11].
- Mutation operator: A randomly chosen pair of genes exchange their positions in a sub-chromosome. Mutation is applied independently in both sub-chromosomes.
- Selection: A roulette-wheel-selection technique is used for selection of chromosomes to survive to the next iteration. The probability of a survival of the chromosome is proportional to its fitness.
- Elitist strategy: In each generation, the best chromosome is automatically transferred to the next generation.
- Fitness function: The genetic algorithm searches for the schedule with highest fitness, where the fitness function is used to assess the quality of a given schedule within the population. The fitness function aggregates the Satisfaction Grade (SG) of two objectives. The satisfaction grades are calculated taking into consideration the completion times of the jobs. Fuzzy processing times of job operations imply fuzzy completion times of jobs. The question arises how to compare a fuzzy completion time of a job with its fuzzy due date, i.e. how to calculate the likelihood that a job is tardy. Two approaches are investigated: (1) based on the possibility measure introduced by Dubois et al [5] and also used by Itoh et al [8] to handle tardy jobs in a job shop problem, and (2) based on the area of intersection introduced by Sakawa in [2].

1. The possibility measure $\pi_{\tilde{C}_j}(\tilde{d}_j)$ evaluates the possibility of a fuzzy event, $\tilde{C}j$, occurring within the fuzzy set $\tilde{d}_j$ [8]. It is used to measure the satisfaction grade of a fuzzy completion time $SG_T(\tilde{C}_j)$ of job $J_j$:

$$SG_T(\tilde{C}_j) = \pi_{\tilde{C}_j}(\tilde{d}_j) = \sup \min\{\mu_{\tilde{C}_j}(t),\ \mu_{\tilde{d}_j}(t)\} \qquad j=1,..,n \tag{4}$$

where $\mu_{\tilde{C}_j}(t)$ and $\mu_{\tilde{d}_j}(t)$ are the membership functions of fuzzy sets $\tilde{C}_j$ and $\tilde{d}_j$ respectively. An example of a possibility measure of fuzzy set $\tilde{C}_j$ with respect to fuzzy set $\tilde{d}_j$ is given in Fig.3.

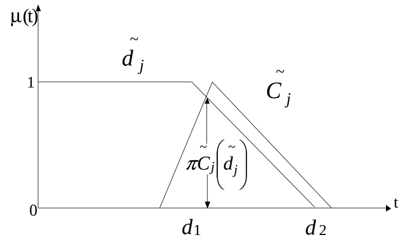


**Fig. 3.** Satisfaction Grade of completion time using possibility measure

2. Area of Intersection measures the portion of $\tilde{C}j$ that is completed by the due date $\tilde{d}_j$ (Fig. 4). The satisfaction grade of a fuzzy completion time of job $J_j$ is defined:

$$SG_T(\tilde{C}_j) = (\text{area } \tilde{C}_j \cap \tilde{d}_j)/(\text{area } \tilde{C}j) \tag{5}$$

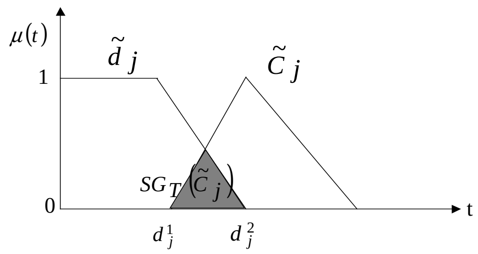


**Fig. 4.** Satisfaction Grade of completion time using area of intersection

The objectives given in (1) and (2) are transformed into the objectives to maximize their corresponding satisfaction grades:

(1) Satisfaction grade of Average Tardiness: $S_{AT} = \frac{1}{n} \sum_{j=1}^{n} SG_T(\tilde{C}_j)$ **(6)**

(2) Satisfaction grade of number of tardy jobs: A parameter $\lambda$ is introduced such that a job $J_j$ $j=1,\ldots,n$ is considered to be tardy if $SG_T(\tilde{C}_j) \leq \lambda$, $0 \leq \lambda \leq 1$. After calculating the number of tardy jobs *nTardy*, the satisfaction grade $S_{NT}$ is evaluated as:

$$S_{NT} = \begin{cases} 1 & \text{if} \quad nTardy = 0 \\ (n'' - nTardy)/n'' & \text{if} \quad 0 < nTardy < n'' \\ 0 & \text{if} \quad nTardy > n'' \end{cases} \tag{7}$$

*n*”= 15% of *n*, where *n* is the total number of jobs.

We investigate three different aggregation operators, which combine the satisfaction grades of the objectives:

1. Average of the satisfaction grades: $F_1$=1/2 ($S_{AT}$ + $S_{NT}$)
2. Minimum of the satisfaction grades: $F_2$=Min ($S_{AT}$ , $S_{NT}$)
3. Average Weighted Sum of the satisfaction grades: $F_3$= 1/2 ($w_1S_{AT}$ + $w_2S_{NT}$), where $w_k \in [0,1]$, $k$=1,2, are normalized weights randomly chosen used in the GA and changed in every iteration in order to explore different areas of the search space [10].

Apart from handling imprecise and uncertain data, fuzzy sets and fuzzy logic enable multi-objective optimization in which multiple objectives that are non-commensurable are simultaneously taken into consideration. In this problem, objectives, the number of tardy jobs and the average tardiness of jobs are measured in different units but have to be used simultaneously to assess the quality of schedules. Values of objectives are mapped into satisfaction grades, which take values from [0,1] interval and can be combined in an overall satisfaction grade.

## 4 Performance of the GA on Real-world Data

The developed GA algorithms were tested on real-world data collected at Sherwood Press over the period of three months denoted here by February, March and April. The load of each month is given in Table 1.

**Table 1.** Datasets

| Month | Number of Jobs | Number of Operations |
|---|---|---|
| February | 64 | 214 |
| March | 159 | 549 |
| April | 39 | 109 |

The experiments were run on a PC Pentium with 2 GHz and 512 MB of RAM, using Visual C++ .Net. The parameters used in the GAs are given in Table 2.

**Table 2.** Genetic algorithm parameters

| | |
|---|---|
| Population size | 50 |
| Length of the chromosome | 2*m*, where *m* = number of machines |
| Probability crossover | 0.8 |
| Mutation crossover | 0.3 |
| Termination condition | 250 iterations |

### 4.1 Experiments with Different Values of λ

The first sets of experiments are conducted with the aim of investigating what an effect changing the value $\lambda$ has on the solution. A higher value of λ leads to higher number of tardy jobs. This is illustrated in Fig. 5, in which two values are used for λ:

$\lambda$ =0.3 and $\lambda$ =0.7. Let $J_j$ be a job with a fuzzy due date $\tilde{d}_j$ that could complete at $\tilde{C}_j'$, $\tilde{C}_j''$ or $\tilde{C}_j'''$ : If it completes at $\tilde{C}_j'$, then $\pi_{\tilde{C}_j}(\tilde{d}_j) \geq 0.7$; therefore, $J_j$ is not tardy for both $\lambda$=0.3 & $\lambda$=0.7. If job $J_j$ completes at $\tilde{C}_j''$, it is tardy if $\lambda$=0.7 and not tardy if $\lambda$=0.3. If it completes at $\tilde{C}_j'''$, $J_j$ is tardy for both $\lambda$=0.3 & 0.7.

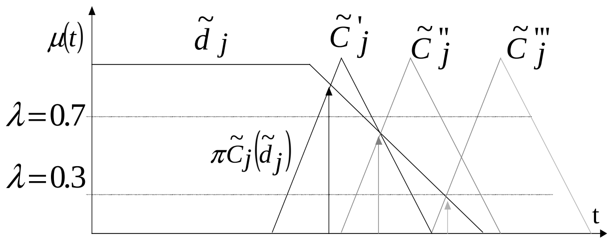


**Fig. 5.** Assessment of job tardiness with different completion times and values of λ

As an illustration, Fig. 6 shows the satisfaction grades of the objectives obtained on the March data, where the aggregation operator Average is used together with the possibility measure to determine tardy jobs. It can be seen that, $S_{NT}$ converges to a higher value ($S_{NT}$ =0.54) faster when $\lambda$=0.3 then when $\lambda$=0.7.

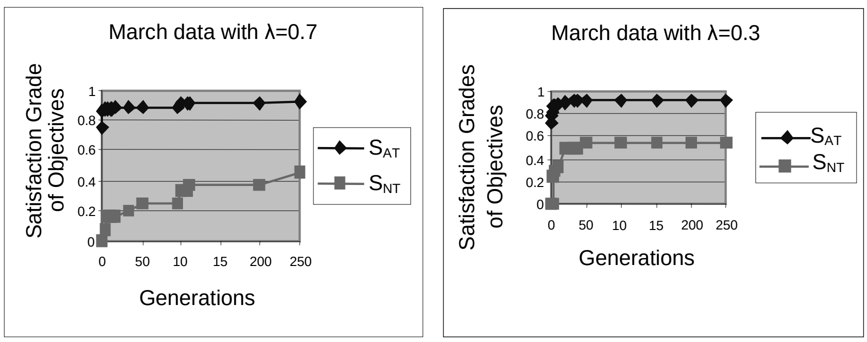


**Fig. 6.** The convergence of the values of objectives for different $\lambda$ values

### 4.2 Experiments with Different Variations of the Genetic Algorithm

Six different variations of the genetic algorithm were developed at $\lambda$ =0.7, with different approaches to determine tardy jobs, using possibility measure (Pos) and area of intersection (Area), and different aggregation operators, using Average, Min and WSum. For illustration purposes, results obtained on March data set running the different GA variations 20 times each are given in Table 3. The data in column FF shows the best/average value of the Fitness Function; of course, these values cannot be compared due to the difference in the nature of the aggregation operators. The columns $S_{NT}$ and $S_{AT}$ show the corresponding best/average values of satisfaction grades of the objective functions, while $C_{NT}$ shows the corresponding values of the objective function of number of tardy jobs. However, three different aggregation operators enable the decision maker to express his/her preferences. Average aggregation operator allows compensation for a *bad* value of one objective, namely a higher satisfaction grade of one objective can compensate, to a certain extent, for a lower satisfaction grade of another objective. On the contrary, Minimum operator is non-compensatory, which means that the solution with a *bad* performance with respect to one objective will not be highly evaluated no matter how *good* is its performance with respect to another objective.

The possibility measure reflects more *optimistic* attitude to the jobs' tardiness than the area of intersection because the former measure considers the highest point of intersection of the two fuzzy sets regardless of their overall dimensions, while the area of intersection considers the proportion of the fuzzy completion time that falls within the fuzzy due date.

**Table 3.** Best and average values of satisfaction grades

| Variations of GA | FF | $S_{AT}$ | $S_{NT}$ | $C_{NT}$ |
|---|---|---|---|---|
| AverageArea | 0.641/0.62 | 0.911/0.908 | 0.371/0.331 | 14/15 |
| MinArea | 0.371/0.264 | 0.904/0.893 | 0.371/0.264 | 14/17 |
| WSumArea | 0.43/0.415 | 0.907/0.893 | 0.371/0.26 | 14/17 |
| AveragePos | 0.69/0.66 | 0.923/0.913 | 0.455/0.41 | 12/13 |
| MinPos | 0.455/0.342 | 0.914/0.903 | 0.455/0.342 | 12/15 |
| WSumPos | 0.435/0.425 | 0.919/0.902 | 0.455/0.316 | 12/15 |

## 5 Conclusion

This paper deals with a multi-objective fuzzy job shop scheduling problem, where uncertain processing times and due dates are represented by fuzzy sets. The objectives considered are average tardiness and the number of tardy jobs. Six variations of the genetic algorithm are developed combining three aggregation operators for objectives and two different methods to determine tardiness of jobs. The results obtained highlight the differences of these aggregation operators in terms of compensation of objectives and the influence of the parameter $\lambda$ in expressing an attitude toward the tardiness of jobs.

Our future research work will be focused on investigation of splitting jobs into lots and combining two or more jobs to be processed at the same time on the machine and processing different jobs of the same category, one after the other to reduce cost of set-up times.

## Acknowledgments

The authors would like to thank the Engineering and Physics Science Research Council (EPSRC), UK, Grant No. GR/R95319/01 for supporting this research. We also acknowledge the support of the industrial collaborator Sherwood Press Ltd, Nottingham.

## References

1. Pinedo, M., 'Scheduling Theory, Algorithms, and Systems,' Prentice Hall, Second Edition, (2002).
2. Sakawa, M. and Kubota, R., 'Fuzzy Programming for Multiobjective Job Shop Scheduling with Fuzzy Processing Time and Fuzzy Duedate through Genetic Algorithms', European Journal of Operational Research 120 2 (2000) 393-407.
3. Foretemps, P., 'Jobshop Scheduling with Imprecise Durations: A Fuzzy Approach', IEEE Transactions on Fuzzy Systems 5 4 (1997) 557-569.
4. Slowinski, R. and Hapke (Eds), M., Scheduling Under Fuzziness, Physica-Verlag, New York (2000).
5. Dubois, D. and Prade H., Possibility Theory: an Approach to Computerized Processing of Uncertainty, New York (1988).
6. Chanas, S. and Kasperski, A., 'Minimizing Maximum Lateness in a Single Machine Scheduling Problem with Fuzzy Processing times and Fuzzy Due Dates,' Engineering Applications of Artificial Intelligence 14 3 (2001) 377-386.
7. Chanas, S. and Kasperski, 'On Two Single Machine Scheduling Problems with Fuzzy Processing Times and Fuzzy Due Dates, European Journal of Operational Research, 147 2 (2003) 281-296.
8. Itoh, T. and Ishii, H., 'Fuzzy Due-date Scheduling Problem with Fuzzy Processing Time,' International Transactions in Operations Research 6 (1999) 639-647.
9. Nagar, A., Haddock, J. and Heragu, S., 'Multiple and Bi-criteria Scheduling: A Literature Survey', European Journal of Operational Research, 81 (1995) 88-104.
10. Murata, T., Ishibuchi, H. and Tanaka, H., 'Multi-Objective Genetic Algorithm and its Applications to Flowshop Scheduling,' Computers Industrial Engineering, 30 4 (1996) 957-968.
11. Bagchi, T., Multi-Objective Scheduling by Genetic Algorithms, Kluwer Academic Publish ers (1999).
12. Coelho, Carlos., Van Veldhuizen, D. and Lamont, G., Evolutionary Algorithms for Solving Multi-Objective Problems, Kluwer Acedmic Publishers (2002).
13. Klir, G. and Folger, T., Fuzzy Sets, Uncertainty and Information, Prentice Hall, New Jersey, (1988).
14. Reeves, C., Genetic Algorithms, in V.J. Rayward-Smith (Eds), Modern Heuristic Techniques for Combinatorial Problems, McGraw-Hill International, UK, ISBN 0-07-709239-2 (1995) 151-196.