APPLYING GENETIC ALGORITHM TECHNIQUES FOR TIME-COST OPTIMIZATION

Daisy X.M. Zheng, S. Thomas Ng and Mohan M. Kumaraswamy

Department of Civil Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong

Completing a construction project with the least time and cost is critical in a competitive environment. However, shortening the duration of any construction activity usually requires the engagement of additional labour and plant or the adoption of alternative construction techniques, and these usually add extra cost liabilities to the contractor. To minimise the total costs associated with schedule compression, contractors are often urged to establish the optimal time-cost relationships for construction activities when planning decisions are made. Several analytical models have been developed for time-cost optimisation. Despite that, difficulties are still being encountered in construction Time-Cost Optimisation (TCO), as there is a lack of unique solutions for integrating constraints associated with the time and cost requirements. Being a powerful tool to locate the global optimum (rather than local optimum), the Genetic Algorithms (GA) could be used to establish the fitness of solutions by evaluating the objective function and its constraints. In this paper, the analytical power of the GA is compared with other techniques proposed for TCO. The results indicate that the GA techniques could generate the most optimal outcome for construction TCO, especially when the project is large and/or complex.

Keywords: construction cost, genetic algorithm, optimization, time-cost trade-off, time duration.

INTRODUCTION

Effective planning and estimation of project time and cost is paramount important to a contracting organisation, as the ability to establish an optimal time-cost equilibrium for a project could improve the chance of underbidding its competitors while attaining the greatest profitability in a competitive environment (Park and Chapin, 1992). In fact, time and cost are intricately related. Classical time-cost estimation concepts construe an inverse relationship between direct cost and activity duration, implying an increase in labour and plant costs, when project duration is shortened (Adrian, 1979). Indirect cost, as characterised by the project overhead, however, increases with the project duration.

In the construction industry, liquidated damages would be imposed on contractors for any non-excusable delays. Despite that, some clients might encourage contractors to maintain high productivity by offering incentives should the project be completed ahead of the targeted duration. Contractors are therefore faced with a dilemma of completing a project in the shortest duration (as encouraged by the incentive/penalty costs), although they would have to increase resources to achieve this. However, introducing extra resources not only leads to a sharp cost increase but could also yields a diminishing output per day (Ahuja, 1994). Both these scenarios are detrimental to the contractor, who strives to minimise its cost within the designated (or

Zheng, D X M, Ng, S T and Kumaraswamy, M M (2002) Applying genetic algorithm techniques for time-cost optimization. *In:* Greenwood, D (Ed.), *18th Annual ARCOM Conference*, 2-4 September 2002, University of Northumbria. Association of Researchers in Construction Management, Vol. 2, 801-10.

shortest) project duration. Therefore, it is important that, when planning a project, the optimum time and cost for each activity should be sought.

Over the years, researchers have developed various decision models for Time-Cost Optimisation (TCO). These models include Critical Path Methods (CPM) (Antill, 1994); mathematical methods, such as linear programming (Kelly, 1961; Reda and Carr, 1989), integer programming (Meyer and Shaffer, 1963; Liu *et al.*, 1995), and dynamic programming (Robinson, 1975); and heuristic approaches, such as Fondahl's method (Fondahl, 1961), structural model (Prager, 1963), effective-cost slope model (Siemens, 1971), and structural stiffness method (Moselhi, 1993).

Being regarded as a promising technique for locating the global optimum, Genetic Algorithms (GA) were adopted for construction TCO (Feng *et al.*, 1997; Li and Love, 1997, 1999). Despite its analytical abilities, the value of GA for TCO has not been fully explored. In this paper, the optimal time and cost generated by the GA techniques are compared with those produced by other analytical techniques through a hypothetical case.

TIME-COST OPTIMISATION TECHNIQUES

Heuristic Methods

Heuristic approaches are non-computer approaches that require less computational effort. Since they are based on the rules of thumb, they could produce reasonably good solutions although they do not ensure optimality. Heuristic approaches have been applied in solving a variety of problems due to their simplicity and ease of application. However, when applied to construction TCO, heuristic approaches are only suitable for small projects, as each iteration involves four steps: to recognise CPM; to combine different activities on CPM; to calculate respective cost and select the least one; to set reduction time and re-test CPM. Therefore, the number of trial may rise sharply with an increase in the number of activities, and the calculation would be prone to error. Moreover, the solutions produced do not provide a range of possible solutions and cannot guarantee to be the global optimum.

Mathematical Methods

Mathematical methods convert heuristic rules into constraints and objective functions, and use linear programming (LP), integer programming (IP) or dynamic programming to solve the problem. Linear programming can only be used when assuming a linear relationship between time and cost for each activity in the network. Although these methods are more efficient and accurate than heuristic methods, they require a lot of computational effort when the number of activities becomes too large or the network becomes too complex. Moreover, formulating constraints and objective functions is time-consuming and prone to errors. In fact, popular software for mathematical methods usually has a limited number of variables and constraints. As a result, only a few construction managers use them to solve large and complex problems encountered in day-to-day practices.

Genetic Algorithms

The GA approaches codify the time and cost of each activity relevant to TCO as a gene along a finite-length string. The GA uses objective functions rather than derivatives or other auxiliary knowledge. In addition, the GA utilises probabilistic transition rules as compared to other deterministic models. All these should contribute to the robustness, and hence result in a more accurate TCO model over the heuristic or mathematical techniques. An extensive literature review has revealed that GA approaches have been used in solving construction TCO problems. For instance, Li and Love (1997) developed an improved GA model for representing the simple linear continuous time-cost relationship, while Feng *et al.* (1997) formulated a GA model for considering multi-criteria type TCO problems. Feng's study also introduced the distance method into the GA model to simultaneously minimise duration and direct cost of a project. Building upon their earlier model, Li and Love (1999) developed a refined GA model, which combined machine learning and GAs for solving problems involving nonlinear and continuous time-cost relationship. Finally, a practical GA model, one which implemented the GA procedures within the commercial project management software, i.e. Microsoft Project™, was formulated by Hegazy (1999).

CASE STUDY

In order to compare the analytical power of various analytical techniques, a case study was devised based on the project data from Liu *et al.* (1995). For practicality, the total project costs, including the direct, indirect and incentive/penalty costs, were considered. The activities, costs, durations and their relationships are presented in Table 1. Considering the difficulties of manual selection in the heuristic approaches and the problems associated with the formulation of constraints for the mathematical methods, the activities were only restricted to seven.

Table 1: Details of the case

Activity description	*Activity No.*	*Precedent Activity*	*Option*	*Duration (days)*	*Direct Cost ($)*	*Linear Relationship for L.P.*
Site Preparation	1		1 2 3	14 20 24	23000 18000 12000	a = –1100 b =38400
Forms and rebar	2	1	1 2 3 4 5	15 18 20 23 25	3000 2400 1800 1500 1000	a = –200 b = 6000
Excavation	3	1	1 2 3	15 22 33	4500 4000 3200	a = –72.2 b = 5583.3
Precast concrete girder	4	1	1 2 3	12 16 20	45000 35000 30000	a = –1875 b = 67500
Pour foundation and piers	5	2, 3	1 2 3 4	22 24 28 30	20000 17500 15000 10000	a = –1250 b =47500
Deliver PC girders	6	4	1 2 3	14 18 24	40000 32000 18000	a = –2200 b = 70800
Erect girders	7	5, 7	1 2 3	9 15 18	30000 24000 22000	a = –888.9 b = 38000

Note: a = gradient (determining the increase of direct cost)
b = constant (value that not varies with the duration)

The data shown in Table 1 was fitted into the CPM, LP, IP and GA methods respectively, and the forms of objective functions are as follows:

for the GA model:

$$\textit{Minimise} \quad C = \sum_{i=1}^{n} C(i) + C_i \cdot T_a - k(T_t - T_a) \qquad (1)$$

for the CPM, L.P. and I.P. methods:

$$\textit{Minimise} \quad C = \sum_{i=1}^{n} C(i) + C_i \cdot T_a - k_i (T_t - T_a) \left[\frac{\frac{T_t - T_a}{|T_t - T_a|} + 1}{2} \right] - k_p (T_t - T_a) \left[\frac{\frac{T_t - T_a}{|T_t - T_a|} - 1}{2} \right] + C_c \qquad (2)$$

where *C = project total cost; C(i) = actual direct cost of activity i;* C_i *= project daily indirect cost;* T_t *= project target duration;* T_a *= project actual duration; k = sharing ratio(k =* k_i *= incentive ratio when (*T_t *-* T_a*) > 0; k =* k_p *= penalty ratio when (*T_t *-* T_a*) < 0);* C_c *= constant cost which remains unchanged in terms of different solutions.*

However, according to the inherent procedure of the heuristic and mathematical methods, actual duration should not exceed target duration, and hence k_p will be neglected in this case. This paper only concentrates on the TCO whereby the project must be completed within the target duration. The parameters set for this study are as follows:

Project desired (deadline) duration: T_t = 70 days
Incentive bonus for early completion: k_i = \$1,000/day
Project indirect cost: C_i = \$1500/day

CPM TIME-COST OPTIMISATION ALGORITHMS

CPM is a traditional networking method widely accepted and applied by construction professionals in simple projects. Involving a form of heuristic approach, the CPM TCO algorithms require one to identify the critical path and then select which activities to 'crash' (reduce time), according to some heuristics (decision or thumb-rules). Each alternative represents a feasible solution in which selected activities could be compressed to meet the required reduction of the project target duration. The choice of activities is presented for expedition; the heuristic rule for this method is to select a set of activities that can be expedited at a lower cost.

In the case study, starting from the all-normal set with 105 days duration at a project cost of \$253,700, the heuristic solution was derived manually in an iterative process. Initially, activities 1, 3, 5 and 7 were on the critical path with activity 3 having the least cost slope (4,500 – 3,200) ÷ (33 – 15) = \$72.70/day (assume linear relationship between extreme points). Activity 10 was then crashed by 8 days. CPM was recalculated, and the project duration became 97 days. The process was continued based on crashing activities 2, 3, 7, and then 1 consecutively. The project duration dipped below the desired level of 70 days (to 68 days) at a total cost of $218,499.70. If the iteration continues, another two feasible solutions can also be derived, and the respective durations and total costs are 67 days at $217,249.70 and 60 days at $221,624.70. In summary, simple CPM TCO algorithms established three feasible solutions well within the target duration of 70 days, and the solution with a total cost of $217,249.70 outperforms others to become the final optimal solution derived from this approach. The output of each iteration is listed in Table 2.

Table 2: TCO results with simple CPM algorithms

Iteration No.	Overall Duration (days)	Total Cost ($)	CRASHED ACTIVITIES	DAYS Crashed
1	105	253,700.00	–	–
2	97	242,277.60	3	8
3	87	229,999.60	2 3	10 10
4	78	224,499.70	7	1
5	68	218,499.70	1	10
6	67	217,249.70	5	1
7	60	221,624.70	5 4	7 7

LINEAR PROGRAMMING

In the LP model, the time-cost curve of each activity is identified as linear continuous. If the relationship is linear, it is possible that the duration of each activity is nondiscrete and noninteger, such as 3.5 days, implying that the duration is continuous. However, only integer durations will be considered in construction applications. Therefore, the duration on each curve is defined as successive integers and the relationship between duration points is linear.

The objective functions and constraints of all activities contribute to the LP model for the TCO and are presented as follows:

$$Minimize: \quad C = \sum_{i=1}^{n} a_i \cdot D_i + C_i \cdot (S_n + D_n) + 0 \cdot \sum_{i=1}^{n-1} S_i - k_i \cdot [T_t - (S_n + D_n)] + C_c \quad (3)$$

$$Subject\ \ to: \quad S_1 = 0\ ; \quad (4)$$

$$S_2 \cdots S_7 \geq 0\ ; \quad (5)$$

$$S_i + D_i \leq T_t\ ; \quad i = 1,2,\cdots 7 \quad (6)$$

$$S_a + D_a \leq S_b\ ; \text{ for each precedence a→b, a,b = 1,…,7} \quad (7)$$

$$D_{ic} \leq D_i \leq D_{in}; \quad i = 1,2,\cdots 7 \quad (8)$$

where *C = project total cost; a_i = cost slope of activity i; D_i = actual duration of activity i; C_i = project daily indirect cost; S_i = start time of activity i; n = number of activities in the project; k_i = incentive ratio; T_t = project target duration; C_c = sum of intercept of cost (y-axis) for each activity i; D_{ic} = the crash duration of activity i; D_{in} = the normal duration of activity i*

The objective function shows that optimisation is sought to minimise the total project cost. The constraints (4) and (5) ensure that the project starts from time zero and all activities will be carried out. Constraint (6) is set to ensure project completion within the target duration. Constraint (7) is the constraint from the precedence relationships of the network. Constraint (8) sets the effective scope of activity duration.

Commercially available L.P. software-TORA[TM], version 2.0 was used to work out the optimal solutions. The following is the output of one feasible solution, with an overall duration of 67 (S7+d7) and a total cost of $217,250.20 (i.e. objective value + 273,783.30 – 1000T_t) (see Table 3)

Table 3: Optimum solution summary with linear programming

Title: TCT_L.P. (14 variables and 28 constraints)
Final iteration No: 6
Objective value (min) = 13466.8984

Variable	Value	Obj Coeff	Obj Val Contrib
X1– D1	14.0000	-1100.0000	-15400.0000
X2– D2	15.0000	-200.0000	-3000.0000
X3– D3	15.0000	-72.0000	-1083.0000
X4– D4	20.0000	-1875.0000	-37500.0000
X5– D5	29.0000	-1250.0000	-36250.0000
X6– D6	24.0000	-2200.0000	-52800.0000
X7– D7	9.0000	1611.1000	14499.8984
X8– S1	0.0000	0.0000	0.0000
X9– S2	14.0000	0.0000	0.0000
X10– S3	14.0000	0.0000	0.0000
X11– S4	14.0000	0.0000	0.0000
X12– S5	29.0000	0.0000	0.0000
X13– S6	34.0000	0.0000	0.0000
X14– S7	58.0000	2500.0000	145000.0000

INTEGER PROGRAMMING

Similar to LP, IP also involves an objective function and constraints. The mathematical model for TCT problem are described as the following:

$$Minimize: C = \sum_{i=1}^{n}\sum_{j=1}^{m_i} C_{ij}\cdot X_{ij} + C_i\cdot[S_n + \sum_{j=1}^{m_n}(D_{nj}\cdot X_{nj})] - k_i\cdot\{T_t - [S_n + \sum_{j=1}^{m_n}(D_{nj}\cdot X_{nj})]\} + 0\cdot\sum_{i=1}^{n-1} S_i \quad (9)$$

$$Subject\ \ to: S_1 = 0\ ; \quad (10)$$

$$S_7 + \sum_{j=1}^{m_7} D_{7j}\cdot X_{7j} \le T_t\ ; \quad (11)$$

$$S_a + \sum_{j=1}^{m_a} D_{aj}\cdot X_{aj} \le S_b\ ;\ \text{for each precedence a}\rightarrow\text{b, a,b} = 1,\dots,7 \quad (12)$$

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i} X_{ij} = 1\ ; \quad (13)$$

where: the variable X_{ij} is assigned to the option j within activity i; m_i = the number of options within activity i; C_{ij} = the direct cost of option j, activity i; C_i = project daily indirect cost; D_{ij} = duration of option j, activity i; k_i = incentive ratio; T_t = project target duration; S_i = start time of activity i; n = number of activities in the project; m_i = the number of options of activity I

Constraints (10) and (11) guarantee that the project will be finished within the targeted duration. Constraint (12) is derived from the precedence relationships of the network. Constraint (13) is to ensure that only one option is selected as the optimal option for each activity. After the objective functions and constraints are formulated, TORA^{TM} can also be used to find the optimal solutions. The output of TORA^{TM} computation is presented in Table 4.

Table 4: TCO solutions based on integer programming

Optimal solution	Iteration no.	Overall duration (days)	***TOTAL COST*** ($)
1	4	60	230,500
2	7	67	225,500
3	9	63	221,000
4	14	63	220,400
5	29	63	218,500

GENETIC ALGORITHMS

GAs provide a set of tools based on natural selection process and mechanisms of population genetics. GAs have been applied to a diverse range of engineering and construction management searching problems. Their robustness is due to their capacity to locate the global optimum in a multimodal landscape. Therefore, a GA is less likely to restrict the search to a local optimum compared with heuristic and mathematical methods.

GAs employ a random but yet directed search for locating the globally optimal solution. Typically, a GA encodes feasible solutions to a linear string called a chromosome, which can be viewed as boxed arranged in a linear manner, as indicated in Figure 1. Each box (gene) in the string represents an activity in the project. The number on the top of the box indicates its activity number in the network and the content in the box is the selected option for the corresponding activity.

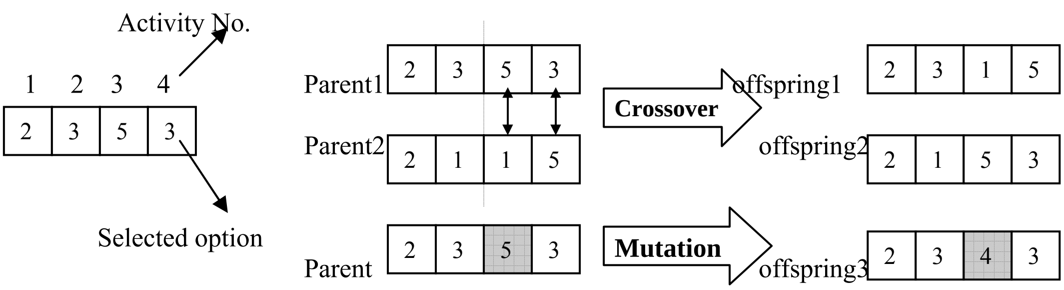

Figure 1: The concept of chromosome, crossover and mutation in the GA

Initially, a group of feasible solutions are randomly assigened to the GA system. The fitness of solutions are evaluated by the objective function. This gives each solution different fitness values according to Roulette wheel method, which represents the possibility of selection into next generation. Then, the selected parent chromosomes undergo stochastic transformations by means of genetic operations to form offsprings. There are two types of transformation: *crossover,* which creates new individuals by exchanging parts of information at randomly selected crossing point(s) from two parents, and *mutation,* which creates new individuals by making a sudden change in a single parent (Figure 1). New individuals are then evaluated. The better/“fitter” offsprings from the current generation then form the parent population for the next generation. After several generations, the algorithm converges to the best solution, which is expected to represent the optimal solution to the problem.

A macro as developed by Hegazy (1999) based upon commercial scheduling software – Microsoft Project™ – was used for establishing the TCO. This macro originally emphasized a random search with no feasible constraints. However, due to the model conditions of heuristic and mathematical methods as mentioned earlier, a penalty ratio in the macro had to be set to infinity to guarantee that the duration of each solution will be within the target duration for comparability among different models. Several runs using the macro resulted in the following solutions (Figure 2 and Table 5):

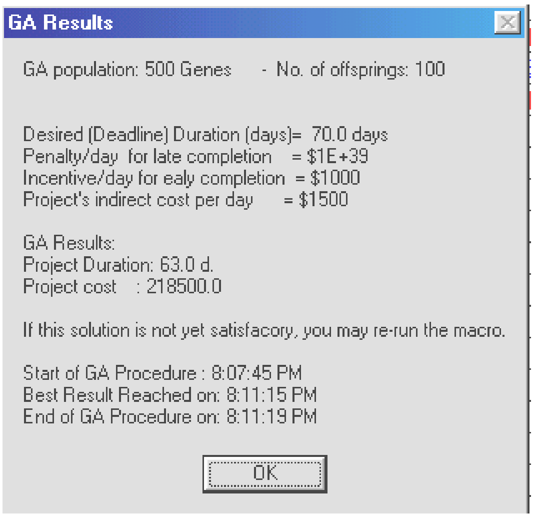

Figure 2: The best solution worked out by genetic algorithms

Table 5: TCO results based on genetic algorithms

	Solutions		
	1	*2*	*3*
Total Duration	63 days	63days	67 days
Total Cost	$220,400	$218,500	$222,300
No. of Generation	100	100	100
No. of Population	100	500	300

DISCUSSION

Having examined the optimal time and cost generated by various analytical techniques for TCO, it is seen that the optimal solutions achieved by a simple CPM algorithm and LP were very much the same, and similarly the results produced by IP and GA were very close. It is also seen that when the relationships between the time and cost were assumed to be linear and continuous, LP outperforms CPM for its efficiency and directness in optimising. Both IP and GA also have the ability to deal with linear and continuous problems if the scope is divided into smaller portions. However, since time-cost relationships are usually nonlinear and discrete in construction practice, IP and GA are more suitable for practical usage.

Although the best solutions derived from IP and GA approaches were very similar, IP is inferior to GA in that its success depends on the limits on the number of variables and constraints inherent in the IP software. That is to say, IP is more suitable for smaller projects, and it cannot adequately cater for both the incentive and penalty costs simultaneously. Moreover, IP cannot guarantee the global searching as in the case of the GA, e.g. the solution produced with the GA (i.e. 67 days at $222,300) is outside the range of optimal solutions generated by IP and it is a better solution than that generated by IP as (i.e. 67 days at $225,500).

CONCLUSIONS

This paper illustrates how GAs can be applied in establishing an optimal time-cost equilibrium for construction projects. By taking incentive and penalty costs into the consideration, the time-cost relationship would become more complex and dependent on concrete data. However, GAs show better overall analytical power in solving this kind of problem when compared with other techniques, such as CPM, LP and IP methods.

In summary, the GA techniques surpass their more traditional cousins for TCO in several ways:

- GAs work with a coding of the parameter set, not the parameters themselves.
- GAs search from a population of solutions rather than a single solution, which guarantees the global exploration.
- GAs use objective functions only instead of derivatives, which avoid the complex formulation of mathematical models.
- GAs use random choice as a mechanism to guide the search toward regions of the search space with likely improvement, not deterministic transition rules.
- GAs facilitate easy experimentation with different scenarios for what-if analysis.

The GA techniques show robustness and result in advantages over other commonly used techniques in achieving TCO. More research efforts appear desirable to validate this initial findings and ensure continuous improvement in this area. The authors will soon apply the latest concepts and techniques of GA ‘technology’ to the construction TCO problem on a more extensive basis, and the results will be disseminated in due course.

REFERENCES

Adeli, H. and Karim, A. (1997) Scheduling/cost optimization and neural dynamics model for construction, *Journal of Construction Engineering and Management*, ASCE, 123(4), 450-458

Adrian, J.J. (1979) *Construction Accounting: Financial, Managerial, Auditing and Tax*, Prentice-Hall, Reston, Virginia.

Ahuja, H.N., Dozzi, S.P. and Abourizk, S.M. (1994) *Project Management: Techniques in Planning and Controlling Construction Projects*, John Wiley and Sons, New York.

Antill, J.M. (1982) Critical Path Methods in Construction Practice, John Wiley and Sons, New York, 1982.

Feng, C., Liu, L. and Burns, S. (1997) Using genetic algorithms to solve construction time-cost trade-off problems, *Journal of Computing in Civil Engineering*, ASCE, 11(3), 184-189.

Fondahl, J.W. (1961) *A Non-Computer Approach to the Critical Path Method for the Construction Industry*, Technical Report No. 9, The Construction Institute, Department of Civil Engineering, Stanford University, Stanford, California.

Goldberg, D.E. (1989) *Genetic Algorithms in Search, Optimisation and Machine Learning*, Addision-Wesley Publishing Co., Reading, Mass.

Hegazy, T. (1999) Optimization of construction time-cost trade-off analysis using genetic algorithms, *Canadian Journal of Civil Engineering*, 26, 685-697

Kelly, J.E. (1961) Critical path planning and scheduling: mathematical basis, *Operational Research*, 9(3), 296-320.

Liu, L., Burns, S. and Feng. C. (1995) Construction time-cost trade-off analysis using LP/IP, *Journal of Construction Engineering and Management*, ASCE, 121(4), 446-454

Meyer, W.L. and Shaffer, L.R. (1963) *Extensions of Critical Path Method Through the Application of Integer Programming*, Civil Engineering and Construction Research Series No. 2, University of Illinois, Urbana-Champaign, Illinois.

Moselhi, O. (1993) Scheduling compression using the direct stiffness method, *Canadian Journal of Civil Engineering*, 20(1), 65-72.

Park, W.R. and Chapin, W.B. (1992) *Construction Bidding: Strategic Pricing for Profit*, John Wiley & Sons, New York.

Prager, W. (1963) A structured method of computing project cost polygons, *Management Science*, 9(3), 394-404.

Reda, R.M. and Carr, R.I. (1989) Time-cost trade-off among related activities, *Journal of Construction Engineering and Management*, ASCE, 115(3), 475-486.

Robinson, D.R. (1975) A dynamic programming solution to cost-time trade-off for CPM, *Management Science*, 22(2), 158-166.

Siemens, N. (1971) A simple CPM time-cost trade-off algorithm, *Management Science*, 17(6), 354-363.