

For office use only

T1 _____
T2 _____
T3 _____
T4 _____

Team Control Number

2015001

For office use only

F1 _____
F2 _____
F3 _____
F4 _____

2015

**The International Mathematical Modeling Challenge (IM²C)
Summary Sheet**

(Attach a copy of this page to the front of your solution paper.)

While most people have watched movies before, many do not know about the immense amounts of planning that go into the production of a film. Specifically, the creation of a filming schedule involves the consideration of multiple constraints. These include the availability of actors, resources, and locations, as well as the preparation of essential elements and the time it takes to shoot on each site. Our method takes these constraints as well as industry regulations into consideration, which no model in the literature to date has done.

We went through multiple strategies before settling on our final model. First we created a model that optimizes time and runs through all possible permutations. This uses logic easily portable into a programming language like Python. Next, we created a cost-optimizing model using Microsoft Excel which had a well-developed and versatile matrix component. We combined the powerful matrix setup of the Excel model with the programmatic aspects of the first model to create an enhanced, primary model.

Our main model uses Integral Linear Programming (ILP) with two heuristics we developed. The first allows for local optimization using pairwise interchange while the second maximizes matches between the actor/resource/prerequisite availability matrix and requirement matrix. This intelligently assigns filming days to calendar dates in a way that capitalizes on human and physical resources as efficiently as possible. We then use pairwise interchange to optimize cost without significantly increasing processing time by only considering cases that are likely to turn out to be both feasible and optimal. We confirm the validity of our model through test cases that showcase the capabilities of both our match maximization and pairwise interchange heuristics.

A very important consideration for the studio is the sensitivity of our model/schedule to unexpected problems, changes, or delays. Our model helps movie studios minimize the effect of such disruptions. We employ sensitivity analysis to determine that actor and resource availability rather than set construction and content generation requirements is the most sensitive part of the model. We can quantify this sensitivity for the movie studio, which has many applications. For example, when the studio insures different parts of their production, it knows how much value to assign to each one.

Another critical consideration is real-world conditions. One of our model's greatest strengths is its cognizance of real-world conditions, which allow it to function not just in a purely mathematical context, but in the context of the film industry, where the regulations of unions, for example, could render another model useless.

Our model thus quickly, efficiently, and accurately produces a schedule movie studios can feel confident using as they plan their future productions.

From Script to Screen: Scheduling the Production of a Movie

Team 2015001

Introduction

The glamour of the entertainment industry often hides the extensive amount of planning required for the filming of a movie. Multiple constraints must be met to construct the most optimal and feasible schedule. These constraints include:

- Availability of actors, locations, and resources
- Preparation of key elements, such as the constructions of sets, that must be completed before shooting certain scenes
- Time necessary to film at specific sites and locations
- Time for reshooting after the primary production period ends

But a movie studio must do much more than just satisfy these constraints: it must also do so in the most cost-effective way possible.

This problem can be classified as a resource constrained project scheduling problem (RCPSP) because it considers resources of limited availability and activities of known duration. RCPSP problems are considered NP-hard, meaning that a solution cannot be found in polynomial time [Artigues 2008]. This justifies the use of heuristics in order to come to good solutions in a reasonable amount of time.

In order to create a model that best satisfies realistic circumstances, it is essential to account for the real-world conditions of the industry. For example, the Screen Actors Guild (SAG), representing 160,000 movie actors all over the world, imposes very stringent regulations on movie producers. Most prominent actors are affiliated with SAG, and their affiliation requires them to only work with movie producers that sign agreements with SAG. As a result, the vast majority of movie companies sign these agreements to gain access to these actors and therefore must comply with SAG regulations. A model ignoring SAG requirements would therefore be of little use to anyone, and one of our model's strengths is its consideration of these requirements. These requirements are detailed in the Assumptions section.

Cheng, et al. [1993] pioneered the treatment of the actor availability problem by seeking to reduce talent hold time (that is, the time actors are paid but are simply being "held" and not used) and Shyu, et al. [2000] followed up on Cheng's work. But both models have substantial problems. Neither takes into account enough real-world conditions to make it practically useful, they can turn up empty-handed when a solution is there, and the computational power required to carry out the PVM model of Shyu is beyond the scope of the resources of a high school environment, as well as those of the department of a movie studio dedicated to scheduling.

Assumptions

1. The film is neither a documentary nor an animated film.

Justification: Documentaries and animated films require different resources and planning from what is presented in the problem. For example, actor availability is not especially relevant to documentaries, and set construction is irrelevant to animated films.

2. Screenwriting is finished, and we are only focusing on production.

Justification: It would not be appropriate to design a schedule for the filming of a movie that has not yet been written. Considerations like actor availability and resource requirements change between the initial and final drafts of the script.

3. We are only concerned about the availability of lead actors, and the availability of secondary actors is immaterial. Similarly, there will not be a shortage of workers to do work like the preparation of sets.

Justifications: 1. While a movie might need specific stars for popularity, the identity of the other actors is not important to the movie. Therefore, the movie studio can and will simply choose whichever actors for the secondary role are available, and as a result secondary actors are available all of the time from a mathematical standpoint. There are always agencies for large studios to find people to prepare sets and act in minor roles. 2. The problem asks us to consider the availability dates of the “stars” of the film only.

4. Actors are employed in “blocks” of time and paid during the entire duration of shooting, regardless of whether they work the whole time.

Justification: This is a result of the SAG constraints mentioned in the introduction. Since the movie company in the problem is described as a “large studio,” we make the assumption that it is bound by SAG rules (because it is virtually impossible to become a large studio without the help of guild-affiliated actors). The most relevant consideration in the SAG rules is that *actors must be paid for hold time*; that is, if an actor is needed on Monday, Tuesday, and Friday, the movie producer must pay the actor as if s/he had been working on Wednesday and Thursday, too, even if the actor did nothing for the producer during that time. SAG covers much of the world, although Bollywood actors in India, for example, organize under their own union, FWICE, and enjoy similar benefits, so we assume for the purposes of this problem that the proportion of actors not covered by SAG, FWICE, or other similar unions and guilds is negligible [Tellychakkar 2010].

5. Time needed to shoot the movie is related to the cost of production. Therefore, minimizing time saves money.

Justification: As production drags on, more money is needed to pay for basic necessities such as food and housing as well as compensating for the crew members’ time.

6. The filming of the scenes must be disjunctive, meaning two scenes cannot be filmed at the same time.

Justification: There is only allowed to be one director for each movie [Directors’ Guild of America 2011] and the director must be present at the filming of each scene so that he or she can direct the artistic and dramatic aspects of the film. Furthermore, the director is responsible for directing the actors and cinematographers and thus the director must always be on set to do so [Media College, nd.]. If there can only be one director and the director must be on set to direct, only one scene can be filmed at a time.

7. The amount of time required to film each scene is generally fixed before filming begins. Therefore, we should aim to shorten the the total production time by reducing the delay in between the filming of scenes and not the scenes themselves.

Justification: The time required to film each scene is dependent upon the circumstances of the shot, which we cannot control. It is reasonable that the expected time for each scene is an accurate prediction for the actual filming time. Therefore, we take the time required to film each scene as constant.

Model

Overview:

The scheduling problem is not new to the literature; it has been attempted before. In this paper, we take two preliminary approaches to the problem presented and then combine them to produce a third, primary model. Of the two preliminary models, one optimizes time and takes an approach relying on computer programming while the other optimizes cost and takes a more algebraic approach using Microsoft Excel Solver. Both preliminary models are available for the studio to use (for instance, if optimizing time is more important than minimizing cost, the time-based model may be utilized), but we focus on the final, primary model, which is a synthesis of the two preliminary models. We combine the programming aspect of the time-based preliminary model with the matrix-based portions of the cost-based model to produce a new model based on integral linear programming using local optimization. We create two heuristics that vastly improve processing time while still ensuring we have a high chance of producing the most optimal solution.

Preliminary Model #1 (Time):

Question 1:

There are many significant factors that affect the total duration of production, including time spent traveling, time spent creating sets, and time windows during which certain resources are not available, creating windows of time where nothing can be done (lag time).

The problem asks us to consider the availability of 6 types of necessities: actors, sites, sets, resources, computer generated content, and physically constructed items. Given the nature of the availabilities, we can categorize them into 2 groups.

First, there are necessities that will be available in an on-and-off basis over time; these include actors, sites, and items such as helicopters, which can only be rented at certain times. We will label all such necessities, with the exception of actors, “resources.” The second group of items will be termed “prerequisites,” because they must be prepared before the scenes requiring them can be shot. Furthermore, they will always be available once enough time has elapsed. This category of necessities will include sets, computer generated content, and physically constructed props.

The expression that gives the total time for the production of a film is:

$$\Sigma \text{ time} = \Sigma \text{ total time filming} + \Sigma \text{ lag time}$$

We wrote pseudocode for this model, which can be represented by the flowchart shown in Figure 1.

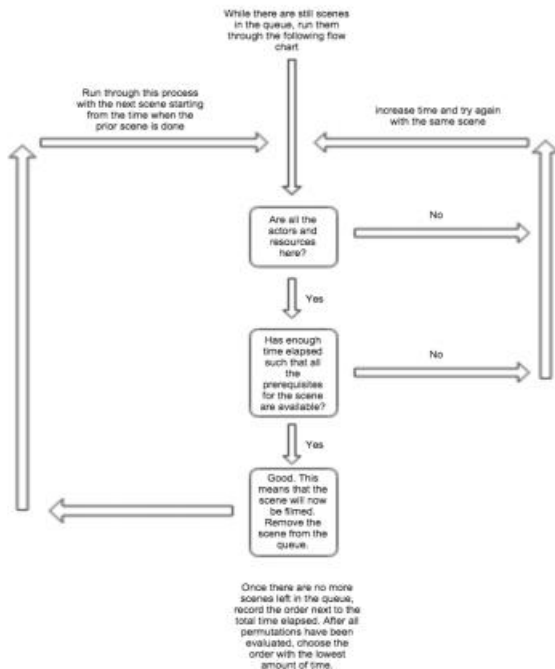


Figure 1

While the logic of this pseudocode holds no matter the number of scenes, actors, etc., using it at scale can become computationally impractical. The average movie has 40-60 scenes, and if there are 50 scenes, the algorithm would have to check 50! sequences. This would take very long for even the fastest computers.

This algorithm, once given all the required inputs, such as actor availability as a function of time, etc., will be able to effectively produce the most time-efficient order for filming the scenes by reducing lag time.

Let us apply this to a simple scenario to display the strengths of this model. Assume a movie has two scenes, two actors, and a single set that must be built. We assign random values to the availability functions and the times.

Let:

$$A_1(t) = \begin{cases} 0, & t = 2, 3, 5, 7, 9, 10 \\ 1, & t = 1, 4, 6, 8 \end{cases} \quad R_1(t) = \begin{cases} 0, & t \text{ is odd} \\ 1, & t \text{ is even} \end{cases}$$

$$N_{s1} = \{ \}$$

$$P_{s1} = \{A_1\}$$

$$P_{s_2} = \{A_1, R_1\}$$

$$N_{s_2} = \{f_1\}$$

$$T_{f_1} = 3$$

$$q = \{s_1, s_2\}$$

$$T_{s_1} = T_{s_2} = 1$$

$$t = 0$$

If one films s_1 first, then the requirements of P_{s_1} will be fulfilled on day 1, and the requirements for P_{s_2} , including N_{s_2} , will be fulfilled on day 4 because N_{s_2} requires 3 preparation days. Thus, the value printed by the algorithm will be $(s_1 s_2, 4)$, with “4” indicating a four day shoot and “ $s_1 s_2$ ” indicating that s_1 should come before s_2 .

However, if one starts with s_2 , then the requirements of R_{s_2} can only be fulfilled on day 4, leaving s_1 to be filmed on day 6 because the first day for which actor 1 is available after day 4 is day 6. This yields $(s_2 s_1, 6)$, with 6 indicating a six day shoot and “ $s_2 s_1$ ” indicating that s_2 comes before s_1 .

Clearly, given the constraints of this scenario, filming s_1 first is advantageous. By finding the lag times of all permutations of scene filming order, the algorithm will eventually achieve the most time efficient sequence. This logic will hold as long as the number of permutations the computer must calculate is reasonable. Because the time it takes to film scenes is constant,

$$\text{delay time} = t - \text{total time filming}$$

$$\text{total time filming} = T_{s_1} + T_{s_2} = 2$$

The total delay time for this scenario would thus be 4 days. By minimizing t , we are minimizing delay time, so the lowest t will provide the lowest delay time (see Assumption 7).

A Gantt Chart for the optimal schedule would then resemble that shown in Figure 2:

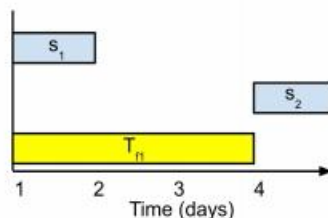


Figure 2

S_1 represents scene 1, and its location in the Gantt Chart (above $t = 1$) shows that it is filmed during day 1. T_n represents the preparation time required for f_1 , which is shown by its length in the chart ($\Delta t = 4 - 1 = 3$, meaning its preparation time is 3 days). Because S_2 cannot be filmed until f_1 is completed, S_2 must wait until the required preparation time (3 days) has elapsed. Therefore, S_2 appears above $t = 4$, after T_n ends, and is filmed during day 4.

Question 2:

This model is also capable of effectively dealing with delays. If an actor or resource has an unexpected change in availability, it would need to be reflected in a new availability function. However, scenes that have already been filmed would not be affected and therefore would no longer remain in the queue for the new algorithm that takes the delays into account. Furthermore, the amount of time elapsed before the delay/change in availability must be taken into account when calculating whether preparation time for the prerequisites has already been fulfilled, assuming the prerequisites have been in development the whole time.

To illustrate, assume that after day 2, actor 2 injures himself and requires 4 days of rest in order to recover. His new availability function would be:

$$A_2(t) = \begin{cases} 0, & t = 1, 3, 4, 5, 6, 7, 9 \\ 1, & t = 2, 8, 10 \end{cases}$$

Now, by revising the algorithm to meet these new needs, we get:

$$\begin{aligned} t &= 2 \\ q &= \{s_2\} \end{aligned}$$

This means that two days have passed, and the only scene in the queue left to film is scene 2. Now, s_2 can only be shot only on day 8. The Gantt chart for this new scenario is shown in Figure 3.

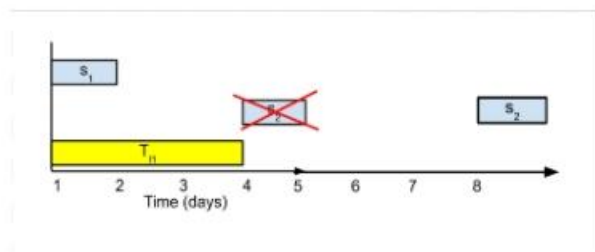


Figure 3

Notice that S_2 has been moved to $t = 8$ because 2 days have already elapsed (meaning the prerequisite is already ready), but the next day for which all requirements of S_2 are fulfilled is day 8 (since actor 2, required for S_2 , is not ready until $t = 8$).

Question 3: The investigation of the most important constraints in the schedule is not directly related to the model itself. Thus, we leave a sensitivity analysis to our discussion of the primary model.

Preliminary Model #2 (Cost with Microsoft Excel):

Question 1:

In this model, we take the more standard approach of optimizing cost. Our first method for modeling the total cost involves creating a table of givens with actor availability, resource availability,

actors required, filming locations, resources required, preparation time (for prerequisites) required, and average cost of changing filming locations.

We generate a matrix of the days each actor gets paid using Microsoft Excel functions. Each column of the matrix represents an actor and each row represents a day. The matrix is binary with 1 indicating a paid day and 0 indicating an unpaid day (all according to SAG constraints; see Assumption 4). For the first and last rows of the matrix only, the actor only gets paid if s/he is required to be filming that day (as opposed to all other days, in which the actor might be paid even if s/he is not working), so the function for those days is

$$=IF([cell\ reference\ of\ actor\ requirement\ for\ day\ 1]=1,1,0),$$

returning 1 if true and 0 if false.

From the second to the second-to-last row of the matrix, an actor gets paid for all days during which they film and days in between filming days (hold days). Therefore, the function becomes

$$=IF(OR([cell\ reference\ of\ actor\ requirement\ for\ day]=1,AND(SUM([cell\ reference\ of\ actor\ requirement\ on\ first\ day]:[cell\ reference\ of\ actor\ requirement\ day\ before])>=1,SUM([cell\ reference\ of\ actor\ requirement\ day\ after]:[cell\ reference\ of\ actor\ requirement\ on\ last\ day])>=1)),1,0),$$

again returning 1 if true (either the actor is required to work on that date or the actor is required to work during days both before and after that date, so must be paid anyway) and 0 if false.

The product of the matrix of actor pay days (dimensions: $a \times b$, with a being the number of days and b being the number of actors) and the column matrix of actor pay rate (dimensions: $b \times 1$, with b being the number of actors) yields the matrix showing the total cost of hiring actors each day (dimensions: $a \times 1$). The sum of this column matrix is the total cost of hiring actors given a certain schedule.

To determine the additional cost resulting from changing locations, the movie studio enters the average cost for moving from one site to another. For the location column in our table, each number represents a different location. We create a column for moving costs based on the fact that each change in location (represented by a change in number in our model) entails added cost, so the Excel function is

$$=IF([cell\ reference\ for\ cell\ before\ current\ one\ in\ site\ column]<>[current\ cell\ reference\ in\ ,[average\ cost\ of\ a\ change\ in\ shooting\ location]],0),$$

outputting the average cost of changing location if true (there is a change in location) and 0 if false (no change in location).

To determine if doing a day of filming on a certain date is possible, we must ensure that all required actors and resources are available and that any time requirement for prior prop construction/computer content generation has been met. In our table for actor availability, for each day, "1" denotes available while "0" denotes unavailable. In our table for actors required per day, "1" denotes required while "0" denotes not required. Thus, if the values of the cells in our table of actor availability are greater than or equal to those of the corresponding cells in our table of actors required, the actors are able to perform on the days in which they are needed to do so. Our tables of resource availability and