Facilitating Timely Completion of a College Degree With Optimization Technology

AVI DECHTER

California State University, USA

adechter@csun.edu

Students who pursue a bachelor’s degree in four-year colleges and universities often take longer than four years to complete their degrees. The reasons for prolonging the time-to-degree seem to fall into three broad categories: part-time enrollment, deficiencies in academic readiness, and inadequate course planning. This article focuses on the latter cause which is due, in part, to the complexity of graduation requirements and institutional regulations that makes effective course planning a difficult task. The article presents an integer linear programming model for finding academic plans that would satisfy a given set of graduation requirements and other constraints in the shortest possible time. The article shows how such a model can enhance and improve existing academic advising tools that are aimed at helping students in decision making and planning for more timely completion of their degree programs.

Not surprisingly, students who enter so called “four-year” colleges and universities in the United States expect to complete their bachelor’s degree in four years. A recent large scale national survey conducted by UCLA’s Higher Education Research Institute shows that only 6.6% of first-time, full-time freshmen enrolled in four-year institutions think that there is very good chance that they will need extra time to complete their degree requirements (Pryor et al., 2006). Yet, data collected by the U.S. Department of Education’s National Center for Education Statistics shows that the 4-year gradu-
The reasons students take longer to graduate than the expected four years may be classified into three broad categories:

1. **Less than full-time enrollment.** Many students do not maintain the course load required to complete their degrees in four years either for financial and personal reasons (for example, they need to work to support themselves or their families) or for academic reasons (reduced course load is needed to maintain good grades). A typical degree program consists of 120 credits and, therefore, completing it in four years requires an average course load of at least 15 credits per semester. If the per semester course load is reduced to 12 credits, for example, completing the program would require at least 10 semesters or 5 years. An in-depth discussion of light course-load as a reason for extending time to graduation is provided by Volkwein and Lorang (1996).

2. **Less than complete academic readiness.** The American higher education system offers access (to some part of it) to almost anyone who wishes to attend college or university (Trow, 1988). As a result, there is wide variation in the levels of preparation for college of students both within and across higher education institutions. Studies have consistently shown that lower levels of college preparation are strongly associated with lower graduation rates and with extended time-to-degree (Astin & Oseguera, 2005, Horn, 2006). Students who are less prepared for college take longer to graduate because they need to take remedial courses that do not count toward graduation requirements; because they often need to take a course more than once before passing it; and because (as previously discussed) they are unable to maintain the full course load required for timely completion of their programs.

3. **Less than adequate course planning.** The ability of students to navigate the requirements for a degree, as well as other institutional rules and regulations, has long been viewed as an impediment for the timely completion of a college degree that can be mitigated by providing students with enhanced academic advisement (Kegley & Kennedy, 2002, Tinto, 2004). Satisfying degree requirements, course prerequisites, and other institutional rules within a given time frame is a complex planning and scheduling prob-
lem that most students are ill equipped to tackle. Pursuing a degree program without proper decision making often leads to taking more courses than required to satisfy the degree requirements as well as to taking courses too late or in the wrong sequence. Helping students to plan better has been shown to improve time-to-degree statistics (Capaldi, Lombardi, & Yellen, 2006).

This article is concerned with the third category of reasons for delayed graduation and with the means higher education institutions have employed to help students in the area of academic planning and decision making. Such help has mostly been provided through academic advising and counseling that, among other functions, is intended to help students navigate the degree requirements and guide them in course selection and scheduling.

The primary tool provided by many colleges and universities for advising on course planning and scheduling is the automated degree progress report (DPR). The DPR is generated by a computerized degree audit system (McCauley & Southard, 1996), which compares the student’s academic record with the requirements for the degree for the purpose of tracking the student’s progress towards degree completion. A DPR typically consists of a list of all the requirements for the degree, identifying those requirements that have been met and those that are not yet satisfied. For each requirement that is not completely satisfied, the report provides information about the deficiency in satisfying it (e.g., remaining number of credits) and a list of courses available to complete it.

The main benefit of the DPR is that it summarizes, in a single document, all the requirements for the degree often scattered in various places, and that it clarifies the options available for the student to complete the requirements. However, it does not provide guidance on which of the available courses should be selected or when the selected courses should be taken. These decisions are critical for the timely completion of the requirements, in particular in the early stage of the program when the number of unfulfilled requirements is large. To provide such guidance the system must be able to reason about the relationships between course selection and scheduling decisions and the total time it would take to complete the requirements. This capability calls for solving a scheduling problem that is far more complex than just checking which of the degree requirements have been satisfied and which have not, and is not part of current degree audit systems.
This article describes how an integer linear programming (ILP) optimization model of the course planning problem may be used to supplement the DPR when making course enrollment decisions. The objective of the model is to find a course schedule that would complete the requirements in the shortest possible number of terms subject to a maximum course load per term. The modeling of this problem, using both integer programming and constraint programming, is discussed in Dechter (2007).

Since degree requirements vary widely among higher education institutions, the next section defines a generic format of degree requirements that captures the basic common features of requirement specification that will be assumed for the remainder of the article. The article then proceeds to develop the proposed ILP model for course planning and then illustrates how such a model might be used to support course enrollment decisions. The article concludes with some suggestions for further study.

**The Generic Course Planning Problem**

This article assumes that earning a college degree entails satisfying a set of degree requirements (requirements, in short) by completing successfully a set of courses selected from a potentially large number of courses offered by the academic institution.

All the courses take a single term (semester, quarter) to complete. (This and some of the other assumptions are made to simplify the exposition; the ILP framework is flexible enough to accommodate many other possibilities, for example, courses that span more than one term.) Courses carry a varying number of credits (also called credit hours or units) which reflect the amount of workload required to complete them. A course may have one or more prerequisites, that is, courses that must be completed prior to taking the course. The prerequisite requirements restrict the way students may schedule courses because a course may only be scheduled for a term subsequent to the term when its prerequisites were taken or are planned to be taken.

A requirement is defined in terms of a minimum number of credits the student must earn to satisfy it and a list of courses from which the student may select to earn the required credit. When all the courses in the requirement’s course list must be taken to satisfy it (i.e., they are “required” courses), the minimum credit required to satisfy it is equal to the total credits
of the courses in the requirement’s course list. When the minimum credit required to satisfy it is smaller than the total credits of the requirement’s course list, the student has a choice of which subset of courses to take to satisfy it (i.e., they are “elective” courses). The time to complete the program may be affected by the student’s selection of elective courses.

Under these assumptions, the requirements for a degree may be summarized in table form as in Table 1, which shows a program consisting of ten courses and three degree requirements. An X in a cell corresponding to a course and a requirement indicates that the course is in the requirement’s course list.

As Table 1 demonstrates, a course may appear on the course lists of more than one requirement (e.g., C5). Unless specifically prohibited, it is assumed that in such a case completing the course would count toward satisfying all the requirements where it is listed. Similarly, a course may be a prerequisite for more than one course (e.g., C2), and finally, a course may be a prerequisite for another course and at the same time be on the course lists of one or more requirements (e.g., C4).

### Table 1
Sample Degree Requirements

<table>
<thead>
<tr>
<th>Course</th>
<th>Credits</th>
<th>Prerequisites</th>
<th>R1</th>
<th>R2</th>
<th>R3</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>3</td>
<td>None</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>4</td>
<td>None</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>3</td>
<td>C1, C2</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>2</td>
<td>None</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>3</td>
<td>C2, C4</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>C6</td>
<td>5</td>
<td>C4</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C7</td>
<td>3</td>
<td>C4</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>C8</td>
<td>2</td>
<td>C6, C7</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C9</td>
<td>3</td>
<td>C8</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>C10</td>
<td>4</td>
<td>C3, C9</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td><strong>Min Credits</strong></td>
<td><strong>15</strong></td>
<td><strong>9</strong></td>
<td><strong>7</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
When course enrollment decisions are made, students would normally have completed some coursework, and be in the process of completing one or more courses, all of which may count toward satisfying degree requirements and prerequisite requirements. When planning the completion of the degree program, the student needs to consider a reduced set of requirements derived from the original requirements by propagating the effects of having completed these courses. Specifically, any requirement that has been met may be removed from the list of requirements and all completed courses may be removed from the master list of available courses, from the course lists of all unmet requirements, and from the prerequisite lists of all remaining courses. The number of credits required to satisfy each of the unmet requirements must be reduced by the number of credits already earned and counted toward satisfying them. For example, if courses C1 through C5 have been completed the reduced requirements are given in Table 2.

<table>
<thead>
<tr>
<th>Course</th>
<th>Credits</th>
<th>Prerequisites</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>C6</td>
<td>5</td>
<td>None</td>
<td>R2 R3</td>
</tr>
<tr>
<td>C7</td>
<td>3</td>
<td>None</td>
<td>X X</td>
</tr>
<tr>
<td>C8</td>
<td>2</td>
<td>C6, C7</td>
<td>X</td>
</tr>
<tr>
<td>C9</td>
<td>3</td>
<td>C8</td>
<td>X</td>
</tr>
<tr>
<td>C10</td>
<td>4</td>
<td>C9</td>
<td>X</td>
</tr>
</tbody>
</table>

| Min Credits | 4 | 7 |

When planning for the completion of a degree in a timely manner, students must consider not only the degree requirements and the prerequisite restrictions, but also the size of the course load they are prepared to take per term. In contrast with the degree requirements that are “hard” constraints over which students have no control, the course load is a “soft” constraint that may be adjusted by the student and therefore is part of the course planning decision. There is an obvious trade-off involved: reducing the course load per term may lead to extending the time to complete the requirements. Conversely, shortening the time to complete the degree may require taking on a larger course load.
With this tradeoff in mind, it is assumed that students wish to find a course plan that would allow completing the degree requirements within the shortest possible time subject to a chosen limit on the per-term course load.

**Modeling the Generic Course Planning Problem by Integer Linear Programming**

Let \( t = 1, 2, \ldots T \) be an index identifying the terms included in the problem’s planning horizon. Let \( C \) be a set of all the courses a student may take to earn a degree, \( u_c \) be the number of credits earned by completing each course \( c \in C \), and \( P_c \) be the set of prerequisites for course \( c \in C \). Let \( R \) be the set of all the course requirements for the student’s degree program, \( C_r \) be the set of courses available to satisfy requirement \( r \in R \) and \( U_r \) be the minimum number of credits required to satisfy requirement \( r \in R \). Finally, let \( MaxLoad \) be the maximum number of units the student may take during any term.

The decision variables are defined as follows:

\[
\begin{align*}
x_c^t &= \begin{cases} 
1 & \text{if course } c \text{ is scheduled for term } t, \quad c \in C, \quad t = 0, 1, \ldots, T \\
0 & \text{otherwise}
\end{cases} \\
y_c^r &= \begin{cases} 
1 & \text{if course } c \text{ is used to satisfy requirement } r, \quad r \in R, \quad c \in C, \\
0 & \text{otherwise}
\end{cases}
\end{align*}
\]

LastTerm = The index of the term when the degree requirements are completed

The objective function makes sure that the plan is completed at the earliest possible term. The constraints in (1) guarantee that no course is scheduled beyond the last term of the plan and the constraints in (2) guarantee that the course load limit per term is not violated during any of the terms. The constraints in (3) state that no course may be scheduled more than once and the constraints in (4) make sure that courses used to satisfy course requirements are included in the plan (i.e., are scheduled for some term). The constraints in (5) guarantee that all the requirements are satisfied. Finally, the constraints in (6) make sure that the prerequisites for each course that is included in the
plan are scheduled for an earlier term than the course for which they are prerequisites.

The course planning problem is represented by following integer linear programming model:

Minimize $LastTerm$
Subject to:

1. $\sum_{t=1}^{T} x_{ct} \leq LastTerm \quad c \in C$
2. $\sum_{t=1}^{T} u_{c} x_{ct} \leq MaxLoad \quad t = 1, 2, ..., T$
3. $\sum_{t=1}^{T} x_{ct} \leq 1 \quad c \in C$
4. $\sum_{t=0}^{T} x_{ct} \geq y_{cr} \quad r \in R, c \in C_r$
5. $\sum_{c \in C} u_{c} y_{cr} \geq U_r \quad r \in R$
6. $\sum_{t=0}^{T} (t+1)x_{ct} - \sum_{t=1}^{T} t x_{ct} \leq (1 - \sum_{t=1}^{T} x_{ct})T \quad c \in C, P \in P_c$

$x_{ct} \in \{0,1\} \quad c \in C, \ t = 1, 2, ..., T$
$y_{cr} \in \{0,1\} \quad r \in R, c \in C_r$
$LastTerm \in \{0,1, ..., T\}$

Executing the course planning model with the degree requirements of Table 1, with the additional specification $MaxLoad = 6$, reveals that this program can be completed in as few as 5 terms. A five-term course plan is shown in Table 3:

### Table 3
A 5-Term Course Plan

<table>
<thead>
<tr>
<th>Term</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Courses</td>
<td>C2, C4</td>
<td>C6</td>
<td>C1, C7</td>
<td>C5, C8</td>
<td>C3, C9</td>
</tr>
<tr>
<td>Credits</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>
Using the same model, one can verify that the shortest time to complete the reduced requirements of Table 2 with no more than 6 credits per term is 4 terms. Since courses C1 through C5 (that the student has already completed) represent a total of 15 credits, they would have taken at least three terms to complete. Therefore, the total time to complete the program if C1 through C5 are taken first would be at least seven terms, that is, two terms longer than the five terms this program should take to complete if planned correctly from the start.

To understand why this plan would require at least two extra terms to complete the degree, consider the graph shown in Figure 1, where the nodes represent the courses and the arrows represent prerequisite relationships in the obvious way. As the graph shows, both courses C6 and C7 are prerequisites for C8 which is, in turn, a prerequisite for C9.

Thus, to complete the program within five terms both C6 and C7 must be taken no later than the third term. Since they cannot be taken in the same term, at least one of them must be taken no later than the second term. If the first three terms are used to take courses C1 through C5 (thus satisfying requirement R1), the earliest either C6 or C7 may be taken is term 3, resulting in the lengthening of the program by at least two terms. As demonstrated in the next section, this type of scheduling error would be avoided by using the model in the course planning process.

![Figure 1. The prerequisite graph.](image-url)
Using the Models to Improve Course Selection and Scheduling Decisions

This section illustrates how the ILP models presented in the previous section may enhance the information provided by the DPR and assist the students in making sound course planning decisions.

Table 4 shows the essential information that would be provided by a DPR based on the requirements given in Table 1. It consists of a list of the degree requirements, an indication of whether each requirement has been met, the courses credited thus far towards satisfying each requirement, the remaining credits needed to satisfy each requirement, and a list of courses available to satisfy each of them. Bold type denotes courses that do not have prerequisites or whose prerequisites have already been satisfied and thus can be taken in the upcoming term.

Table 4
A Degree Progress Report

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Met?</th>
<th>Course Taken</th>
<th>Credits Remaining</th>
<th>Courses Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>No</td>
<td>None</td>
<td>15</td>
<td>(C1(3), C2(4), C3(3), C4(2), C5(3))</td>
</tr>
<tr>
<td>R2</td>
<td>No</td>
<td>None</td>
<td>9</td>
<td>(C4(2), C5(3), C6(5), C7(3))</td>
</tr>
<tr>
<td>R3</td>
<td>No</td>
<td>None</td>
<td>7</td>
<td>(C7(3), C8(2), C9(3), C10(4))</td>
</tr>
</tbody>
</table>

The information contained in Table 4 is helpful to the student as it points to the unmet requirements and to the courses that are available to satisfy them. In particular, it helps the student avoid taking courses that do not count toward graduation requirements. However, it does not provide any information that would associate the student’s decisions with the time to complete the requirements. A good starting point would be exploring the tradeoff between the maximum course load and the number of terms needed to complete the program with the aid of the proposed optimization models.

The previous section used the ILP model to show that if the course load per term is capped at 6 credits, the least number of terms required to complete the program is 5 terms. Repeating this analysis for varying levels of course load limits Table 5 is obtained:
As expected, the completion time goes down when course load limit increases. The shaded cells contain the “efficient” pairs of course load limit and shortest completion time (i.e., when any decrease in the course load limit would cause the corresponding shortest completion time to increase). The efficient pairs are plotted in Figure 2 to show the problem’s efficient frontier. As the graph illustrates, the effect of increasing the course load limit on the time-to-degree may not be linear. This type of information provides a basis for the student to decide on a realistic time target for completing the degree.

<table>
<thead>
<tr>
<th>Course Load Limit:</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shortest Completion Time:</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Figure 2. The course load/time-to-degree efficient frontier.

Suppose that based on this information the student decides to plan toward completing the program within 5 terms while enrolling in no more than 6 credits in any term (corresponding to the center point in Figure 2). Next, the student must decide on which courses to take in the upcoming term.

As Table 4 shows, there are three courses (C1, C2, and C4) available for taking in the upcoming term. Because of the 6 credit course load limit, there are five course combinations that may be selected: {C1}, {C2}, {C4}, {C1, C4}, and {C2, C4}. The effect of selecting each of these course com-
binations on the time to complete the degree may be found by executing the model with additional constraints representing each choice. For example, to represent taking just C1 in the first term the following constraints are added:

\[ x_{C1,1} = 1, \quad x_{C2,1} = 0, \quad x_{C4,1} = 0 \]

Executing the model with the added constraints shows that the program would take at least 6 terms to complete. Therefore, this choice is not compatible with the goal of completing the program in 5 terms. Repeating this analysis for the other course combinations shows that the only selection consistent with the target completion time of five terms is to schedule both C2 and C4 for the upcoming term.

Now assume that the student has selected and completed courses C2 and C4 in the first term and is ready to select courses for the second term. The updated DPr is given in Table 6.

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Met?</th>
<th>Course Taken</th>
<th>Units Remaining</th>
<th>Courses Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>R1</td>
<td>No</td>
<td>None</td>
<td>9</td>
<td>C1(3), C3(3), C5(3)</td>
</tr>
<tr>
<td>R2</td>
<td>No</td>
<td>None</td>
<td>7</td>
<td>C5(3), C6(5), C7(3)</td>
</tr>
<tr>
<td>R3</td>
<td>No</td>
<td>None</td>
<td>7</td>
<td>C7(3), C8(2), C9(3), C10(4)</td>
</tr>
</tbody>
</table>

The courses available now for enrollment are C1, C5, C6, and C7. Under the course load limit of 6 credits, there are 6 course combinations available for the next term: C1, C5, C6, C7, C1+C5, C1 + C7, C5 + C7. Repeating the method described shows that only three of these combinations are consistent with the goal of completing the remaining requirements in four terms: C6, C1 + C7, and C5 + C7. Selecting any of these options would enable the student to complete the program within five terms as planned.

There are many other decisions and choices (in addition to the maximum course load and the selection of courses in the upcoming term) that may affect the time to complete the program. Some examples are choosing to take a given course during a specific term, choosing not to take a given (elective) course at all, and reducing (or increasing) the maximum course
load for a given term. Knowing the effect of such changes on the time to complete the program, which can easily be determined by amending the model with appropriate constrains, would help the student to make more informed decisions and avoid making choices that would lengthen the time-to-degree.

Conclusions and Suggestions for Further Research

This article argues that course planning tools currently available for college students could be enhanced by employing optimization technology to address the underlying scheduling problem students are facing when making course selection decisions. Using a simplified set of degree requirements, the article illustrates how enrollment decisions guided by a standard degree progress report can lead to lengthening the time to complete the requirements. Also shown is how an optimization model that represents the degree requirements, the prerequisite requirements, and course load constraints might be used in conjunction with the DPR to help the student make enrollment decisions that would promote more timely completion of the degree program.

Since “real-life” degree requirements are usually far more complex than our example, it is conceivable that poor course planning contributes significantly to delays in degree completion. Colleges and universities invest heavily in efforts to help students better plan their studies in an effort to improve graduation rates and reduce time-to-degree. However, more research is needed on the exact role of inadequate planning in delaying degree completion. Using models like the one proposed in this article, the actual academic records of students can be analyzed to determine the degree to which “incorrect” enrollment decisions may have contributed to delaying their graduation. Ultimately, the value of the approach proposed in this article should be judged by its effect on actual timely completion statistics. This would require the implementation of the model in a real university setting and tracking its performance for a number of years.

The approach proposed in this article would be practical only if the scheduling model can be reliably solved quickly. While all the solutions for the discussion in the article were obtained in less than 2 seconds, the size and complexity of real degree requirements may present a computation challenge because the course planning problem is an inherently hard
combinatorial optimization problem. In fact, it is closely related to the notoriously hard problems of assembly line balancing (Boysen, Fliedner, & Scholl, 2007) and resource constrained project scheduling (Brucker, Drexel, Mohring, Neumann, & Pesch, 1999). The computational complexity of the problem should be studied and modeling techniques other than integer programming should be explored.

Lastly, the effect of graduation requirements themselves on graduation statistics should be studied. Graduation rates and average time-to-degree completion vary widely among higher education institutions. Studies aimed at explaining the variation have focused on the characteristics of the students populations: socio-economic background, high school GPA, need for financial support, and so forth (DesJardins, Ahlburg, & McCall, 2002; Astin & Oseguera, 2005). However, universities also vary widely in how graduation requirements and policies are structured and it would be very useful to study whether shorter degree completion times may be explained by the way requirements are set by the institution. It is quite plausible, for example, that by simplifying the requirements institutions may make it easier for their students to plan toward more timely degree completion.

References


