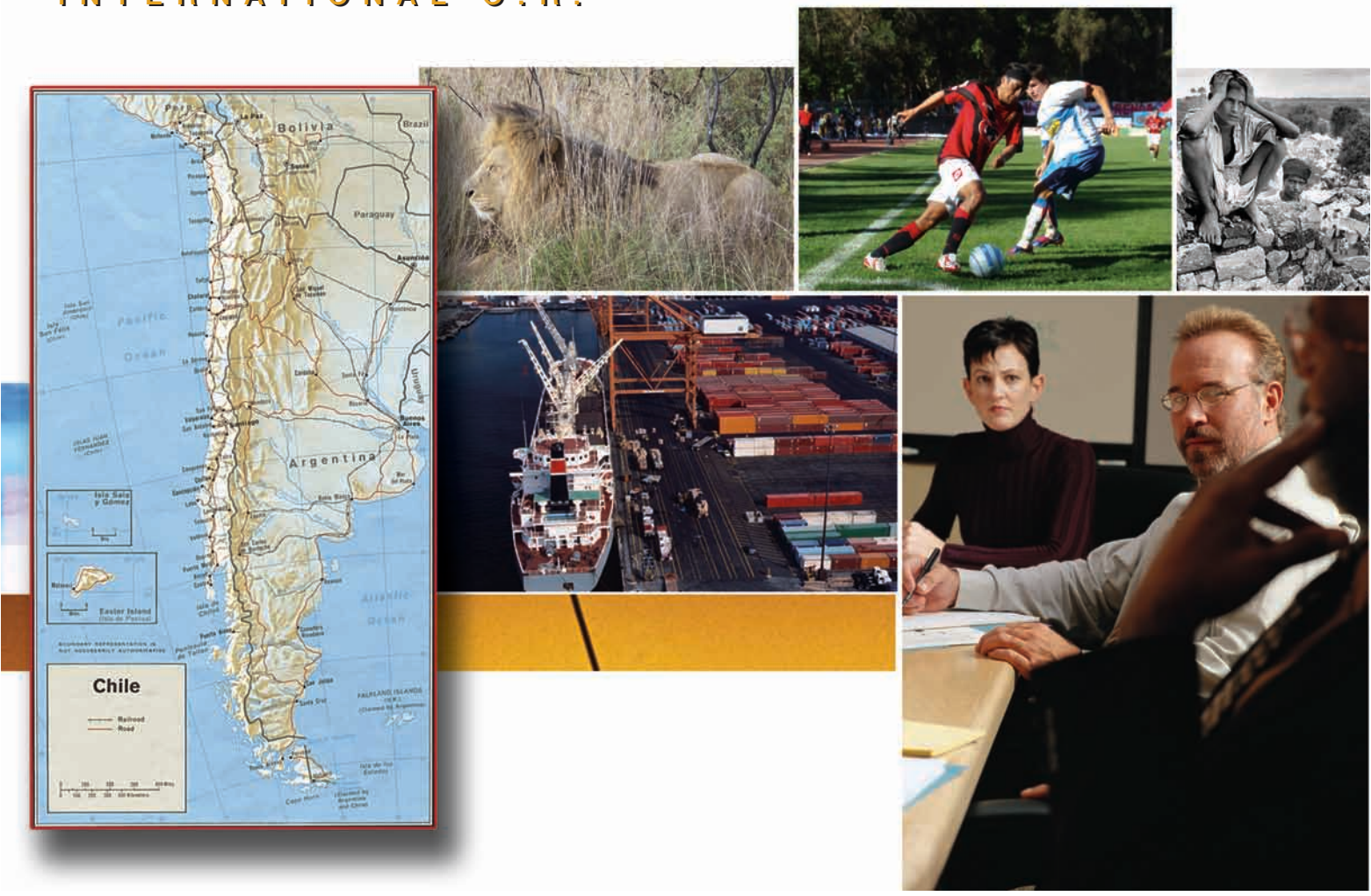


O.R. & SOCCER

Scheduling the Chilean League using mathematical programming.

BY GUILLERMO DURAN, MARIO GUAJARDO, ANDRES WEINTRAUB AND RODRIGO WOLF

Soccer photos by Juan Peñaloza

In 2004, George Nemhauser attended a conference on combinatorial optimization held in Santiago, Chile (LACGA'04), which was organized by the Center for Operations Management (CGO) at the University of Chile's Department of Industrial Engineering. After giving a talk on game scheduling for Major League Baseball (MLB), the Georgia Tech engineering professor granted an interview to a prominent local daily that caught the attention of the Chilean National Professional Soccer Association (ANFP). The Association contacted the CGO, setting in motion a working relationship between the two organizations that has so far produced the scheduling of all Chilean soccer league games for the First Division since 2005, the Second Division since 2007 and the youth divisions since 2008.

Using a manual system, attractive games used to be scheduled on inappropriate dates.

Chile's annual First Division soccer calendar is divided into two tournaments known as the Apertura ("opening"), held during the first half of the year, and the Clausura ("closing"), played during the second half. Until last year the organization of each tournament was as follows. There were 20 teams organized into four groups of five, and each team played the other 19 once over the course of the "regular season." At the end, the two top clubs in each group qualified for the playoffs to determine the tournament champion, similar to the Mexican Tournament. The current format is very similar but the First Division now has only 18 teams, each of which plays every other one once during the regular season of each tournament. The eight clubs with the most points at the end advance to the playoffs.

Before 2005, the ANFP drew up its game schedules using the manual system employed by most of the world's professional soccer leagues in which a random draw is held for places on a pre-established schedule template. If there are six teams and five rounds, for example, the template would be defined as in Fig. 1, and the teams would then draw for each place (see page 44). For the case shown, the result is that in the first round, Team 1 plays Team 2, Team 6 plays Team 3, Team 5 plays Team 4, and so on.

This method has a number of weaknesses, which in the case of the Chilean league led to problems such as the following:

- Attractive games were scheduled on inappropriate dates. For example, in 2004 the most important match of the season, which pits Universidad de Chile against Colo Colo, was set for the first round of the tournament when interest tends to be relatively low.

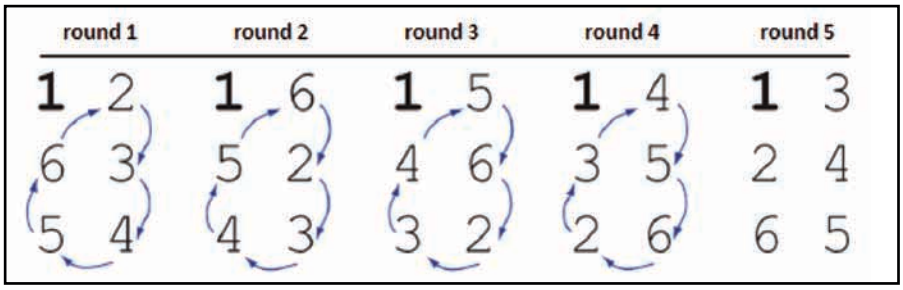


Figure 1: Schedule template for six teams under traditional manual scheduling method.

- Many games were scheduled for dates when the teams' stadiums were not in fact available, forcing the ANFP to suspend the matches or transfer them to other venues.
- Game dates dovetailed poorly with international matches. Teams participating in the Copa Libertadores de América (South American club championship cup) were sometimes scheduled to journey abroad for mid-week cup games after traveling long distances within Chile for the previous Sunday match. In such cases, either the players had to face international teams while in less-than-top condition or Chilean league games had to be postponed.
- Weaker teams on occasion were scheduled to play two or three games in a row against stronger teams, a situation considered to be undesirable.
- In 2004, a few rounds were held on Wednesdays when attendance, and therefore gate receipts (which go entirely to the home team), are lower than on Sundays. Differences among the teams in the number of Wednesday matches played at home meant that some were favored financially while others were hurt. Also, Wednesday and Sunday rounds were not combined intelligently to avoid unnecessarily tiring road trips. Since Chile is geographically long (some 2,700 miles) and narrow, the ideal arrangement would be for a team from the north with a Wednesday game in the south to play the previous or following Sunday in the same region, and vice versa for teams from the south.

Various other problematic or unfair situations were common under the former scheduling system, thus accounting for the ANFP's interest in Nemhauser's comments and its decision to approach the CGO. However, though the Association realized that using operations research techniques could improve the design of tournament schedules, its lack of familiarity with such tools meant that it did not initially appreciate the magnitude of the possible benefits and how these would be achieved. Also a factor was that mathematical scheduling methods were not yet widely used in the world of soccer (unlike the main professional and amateur sports in the United States).

Fortunately, the two organizations quickly reached an understanding and an agreement was arrived at on the design of a schedule for the 2005 *Apertura*. But the results would have to meet the ANFP's requirements, and with the deadline less than two months away, time was very short.

The working method adopted was as follows. Once they understood the wide margin allowed by the scheduling techniques for setting criteria, ANFP officials defined the requirements it wished the schedules to satisfy. To determine these conditions the Association first analyzed in detail the problems with manual scheduling such as those described above. It then decided on other elements that would help produce schedules that were fairer to the teams and more attractive to soccer fans. The process included a series of consultations with First Division team executives. One of the new constraints was that certain attractive games be scheduled in tourist spots during the summer to take advantage of the influx of vacationers. An attractive match is considered to be one in which a local team hosts one of the three most popular clubs at the national level, such as when Everton plays at home in Viña del Mar against Colo Colo.

The ANFP was extremely pleased that important game dates would no longer be left to chance and the resulting tournaments would generate greater interest. The assurance that the season would run smoothly as planned without the need to move up or postpone matches because no stadium was available was also highly appreciated. In general, the Association was very content with its new ability to define a schedule that would guarantee all the conditions for an exciting and profitable tournament.

The CGO also put forward some of its own ideas for making the schedule more appealing, subject to the approval of the ANFP. All that then remained was to translate everything into mathematical language. This task in the end was not overly difficult, and the integer linear programming model was completed in a single day. The next step was to make the transition from pencil and paper to the computer and the generation of solutions. Writing the code for the model was also relatively straightforward, but obtaining the solutions turned out not to be.

The original design of the model contained 7,900 binary variables and some 3,000 constraints; not a simple problem by any definition. After running it in CPLEX for almost two days with no result, it was clear that the model as it stood was practically impossible to solve. It was therefore decided to introduce patterns consisting of sequences that determined whether a team was playing at home or away for the various rounds. These patterns imposed the condition, for example, that teams could not play at home three rounds in a row, or that teams with rela-

The original model contained 7,900 binary variables and some 3,000 constraints; not a simple problem by any definition.

tively limited stadium availability would have at least one pattern that assigned home games on dates their grounds were free. Thus, there were a number of considerations to be taken into account.

After the patterns had been generated, certain restrictions were placed on them and the model was run with only the most important constraints included. Some of the teams were also obliged to follow a given pattern. The size of the scheduling problem was thus reduced and the feasible solution space made considerably smaller. Once this modified problem was solved, the excluded constraints were gradually reintroduced and work continued on the search for a definitive solution that met all the conditions.

An attractive match is considered one in which a local team hosts one of the three most popular clubs at the national level.

However, using patterns in this way runs the risk of shrinking the feasible solution space to the point where certain possible solutions are no longer attainable. A balance must therefore be found between narrowing the space and excluding some solutions. The idea is to start by using as many patterns as there are teams and constrain each team to associate with a pattern. If this still does not yield a feasible solution, the solution space can be enlarged by generating additional patterns, which will then outnumber the set of teams. Another option that assures even greater margin is to free some teams from the obligation to adopt a pattern. Under this approach, some clubs would have to abide by predefined home-away patterns while others would have them generated as the model is solved.

The objective function we chose maximized the number of games between teams of the same group played toward the end of the tournament. This ensured all of the final rounds, when teams are fighting for a place in the playoffs, would be important ones. Since in our case the principal concern is feasibility, other objective functions could have been used, but we opted for a design that scheduled important games during the closing weeks in order to heighten the excitement of the tournament and increase its attractiveness.

Chicago

The International Symposium on Mathematical Programming is a scientific meeting held every three years on behalf of the Mathematical Programming Society. The symposium marks the 60th anniversary of the Zero th Symposium organized by the Cowles Commission at the University of Chicago in June, 1949. Further information can be found on the conference web site:

ismp2009.org

ISMP

International Symposium on Mathematical Programming
August 23–28, 2009

Register Today!
Early registration by May 29

informs® mps® GAMS Microsoft Solver Foundation ILOG An IBM Company mosek
CHICAGO BOOTH Gurobi Optimization LINDO SYSTEMS INC. AMPL

Thanks to the scheduling changes, gate receipts have increased dramatically, and in some instances, have tripled.

In the end, the model produced a schedule that met all the requirements of the ANFP. The Association was very satisfied, and despite the election of a new board of directors at the end of 2006 who opposed the former members, the alliance with the CGO has continued unaffected. This clearly demonstrates the acceptance and appreciation by the clubs of the new scheduling method.

Another point of significance is that many different schedules are generated before the definitive one is arrived at. In 2007, for example, some 70 versions were produced. This occurs because new conditions often arise after the initial ones have been submitted and the first model defined. In some cases, stadiums that were thought to be available on certain dates turned out not to be; in others, the ANFP decided after reviewing provisional schedules that certain matches would be more attractive if moved to different rounds, and asked that these changes be incorporated. This ongoing presentation of new requests in effect turns the scheduling procedure into an iterative process that continues for some two weeks, finally ending with a definitive version that satisfies all, or almost all, of the Association's demands.

In the years since the CGO's first soccer season schedules, new constraints have been gradually incorporated into the model that make it increasingly attractive to the public while saving money for the teams and even for the ANFP-owned TV channel owning the broadcasting rights. One of these restrictions is designed to hold down the high cost of sending mobile television units to opposite ends of the country on the same date. Since three of the four games televised for each round involve Chile's most popular clubs, the constraint is designed to avoid schedules that have them playing on the same day in both the north and the south. This results in significant savings for the broadcaster.

To appreciate the impact of the new method, consider the following data on average stadium attendance. For the 2004 *Apertura,* the last one to be manually scheduled, the figure was 3,756, whereas two years later it had jumped more than 30 percent to 4,953. Gate receipts, meanwhile, rose by more than 100 percent.

Another telling statistic is the number of spectators at the Universidad de Chile-Colo Colo game, always the most important one of the season. In 2004, under manual scheduling methods, the figures were 26,000 for the *Apertura* and 22,000 for the *Clausura,* the match in both cases being the first round of the tournament. In 2005, using the mathematical scheduling model, attendance at the two top matches was 45,000 for the *Apertura* and 37,000 for the *Clausura.*

Further proof is the case of Everton, whose home ground as noted is in Viña del Mar, a major tourist town. In 2004 the team played Universidad Católica, one of the most popular clubs, in late April, whereas in 2006 the same matchup was set for the end of January, the heart of the Southern Hemisphere summer. Thanks to this scheduling change, gate receipts tripled. Similar improvements have been experienced by all clubs located in tourist areas.

The foregoing examples are illustrative of the quantifiable benefits reaped since the tournaments began scheduling using the CGO model. To be sure, the use of operations research techniques is not the only factor behind these outcomes, but there can be no doubt that it has made a significant contribution.

A number of anecdotes over the four years of working with the ANFP reinforce the evidence for the advantages of this type of model. During the *Apertura* of 2005, for instance, a player expressed considerable surprise that four of his team's last five games of the tournament were "just by chance" against its group rivals, noting with enthusiasm that this "fluke" made the final games much more attractive. Clearly, he had no idea that operations research was behind this result, which in fact was the very aim of the objective function.

A potentially major problem arose after the schedule for the 2005 *Clausura* had already been handed over to the ANFP. An official with the Association discovered that Colo Colo, the country's most popular team, was to play away against Coquimbo United on the same day the town of Coquimbo was hosting a major religious festival. This meant not only that relatively few people would attend the very game the home club was counting on to bring in the most money, but also that the festival's security demands would cut into the large police presence normally required for such an important match.

It was clear, then, that the game date had to be changed, but the schedule for the year was already official and the key features (season openers, classic matchups) had been published in the Internet versions of the country's newspapers. Rescheduling a single match might seem a simple task, but in fact is far from it.

He had no idea that operations research was behind this result, which in fact was the very aim of the objective function.

The new scheduling method has been well-received by both the ANFP and the various teams.

An attempt by the Association to fix the situation manually met with no success. The very night the schedule conflict was noticed, a model was run that left the first rounds and classic rivalries unchanged, met all of the other ANFP conditions, set a new date for the Coquimbo meeting and minimized the number of other games that had to be rescheduled. By the next morning an optimal solution had been found, with dates altered for only 19 out of a total of 190 tournament matches. Both the Association and Coquimbo United were happy with the solution and the new tournament schedule was approved.

Another scheduling difficulty cropped up in 2007. The tournament calendar had already been finalized when the ANFP's disciplinary tribunal, which functions autonomously, ruled that a team that had earlier been relegated due to salary payment problems could return to the First Division. A new schedule for 21 clubs instead of 20 suddenly had to be generated within a very few days.

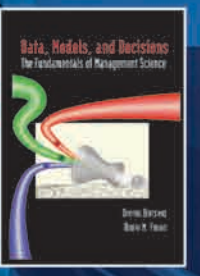
All in all, the experience with the new scheduling method has been very positive and both the ANFP and the various member teams are more than satisfied with the results. Work is currently underway on a mathematical programming model to define referee assignments that would incorporate the same conditions the Association currently applies to the problem from round to round. As for scheduling the Second and youth divisions, the procedure is the same as for the First: the ANFP submits its requirements and conditions to the CGO, which then translates them into mathematical language before running the resulting model on the computer. The Second Division tournaments have a relatively small number of restrictions and those of the youth divisions fewer still, making them easier to schedule than the First Division calendar. Even so, incorporating the ANFP's various requirements for these divisions would be impossible to achieve with manual techniques.

Finally, the soccer scheduling project has also been an enriching experience for the CGO, giving many academics and students the opportunity to combine their two great passions: operations research and soccer. IORMS

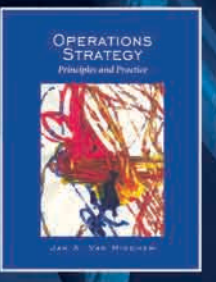
__Guillermo Duran__ is an assistant professor and __Andres Weintraub__ (aweintra@dii.uchile.cl) is a professor in the Department of Industrial Engineering, University of Chile in Santiago, Chile, where __Rodrigo Wolf__ and __Mario Guajardo__ are grad students and research assistants..

DYNAMIC IDEAS
publishes books in Operations Research and Applied Mathematics that aspire to make a difference in the education of the next generation.

Several of its books are used as the main text book in many leading universities in the world.

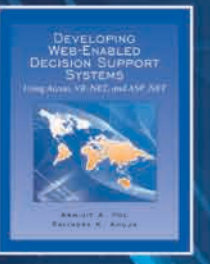


Data, Models and Decisions: The Fundamentals of Management Science
By Dimitris Bertsimas and Robert Freund
ISBN 0-9759146-0-X $89.99

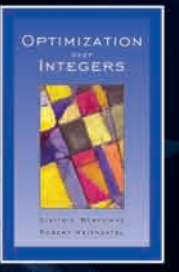


Operations Strategy: Principles and Practice
By Jan A. Van Mieghem
ISBN: 0-9759146-6-9 $89.99

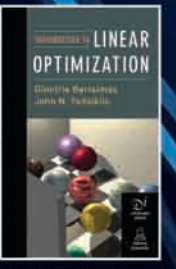
Developing Web-Enabled Decision Support Systems: Using Access, VB. NET, and ASP. NET
By Abhijit A. Pol and Ravindra K. Ahuja
ISBN: 0-9759146-4-2 $79.99



Developing Spreadsheet-Based Decision Support Systems: Using Excel and VBA for Excel
By Michelle M.H. Seref, Ravindra K. Ahuja, and Wayne L. Winston
ISBN: 0-9759146-5-0 $79.99



Optimization Over Integers
By Dimitris Bertsimas and Robert Weismantel
ISBN: 0-9759146-2-6 $89.99



Introduction to Linear Optimization
By Dimitris Bertsimas and John Tsitsiklis, co-published with Athena Scientific.
ISBN: 978-1-886529-19-9 $89.99

Urban Operations Research
By Richard Larson and Amedeo Odoni
ISBN: 0-9759146-3-4 $89.99

Ordering information: www.dynamic-ideas.com