The Distance-Optimal Inter-League Schedule for Japanese Pro Baseball

Richard Hoshino and Ken-ichi Kawarabayashi
National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan

Abstract

Nippon Professional Baseball (NPB) is Japan's largest and most well-known professional sports league, with over 22 million fans each season. Each NPB team plays 24 *inter-league* games during a five-week period each spring. In this paper, we solve the problem of determining the best possible NPB inter-league schedule that minimizes the sum total of distances traveled by all teams. Despite the NP-completeness of the general problem with $2n$ teams, we show that for the 12-team NPB, the distance-optimal inter-league schedule can be determined from two heuristics that drastically cut down the computation time. Using these heuristics, we generate the best possible schedule when the home game slots are *uniform* (i.e., every team in one league plays their home games at the same time), and when they are *non-uniform*. Compared to the 2010 NPB inter-league schedule, our optimal schedules reduce the total travel distance by 15.3% (7849 km) in the uniform case, and by 16.0% (8184 km) in the non-uniform case.

Introduction

Nippon Professional Baseball (NPB) is Japan's most popular professional sports league, with annual revenues topping one billion U.S. dollars. In terms of actual attendance, the NPB ranked second in the world among all professional sports leagues, ahead of the National Football League (NFL), the National Basketball Association (NBA), and the National Hockey League (Wikipedia 2011).

The NPB is split into the six-team Pacific League and the six-team Central League. Each team plays 144 games during the regular season, with 120 *intra*-league games (against teams from their own league) and 24 *inter*-league games (against teams from the other league). To complete these $\frac{1}{2} \times 12 \times 144 = 864$ games, the teams travel long distances from city to city, primarily by airplane or bullet-train. During the 2010 regular season, Pacific League teams traveled a total of 153940 kilometres to play intra-league games, while the more closely-situated Central League teams traveled 79067 kilometres (Hoshino and Kawarabayashi 2011b).

By reformulating intra-league optimization as a shortest path problem, the authors determined a distance-optimal schedule that retained all of the NPB constraints that ensured competitive balance, while reducing total Pacific League team travel by 25.8% (nearly 40000 kilometres) and Central League team travel by 26.8% (over 21000 kilometres).

The motivation for this paper is to extend our NPB analysis to inter-league play, to determine whether the 2010 schedule requiring 51134 kilometres of total team travel can be improved, and perhaps even minimized to optimality. The rationale for our paper is timely, given current global economic and environmental concerns. Implementing a distance-optimal schedule would help this billion-dollar sports league be more efficient and effective, saving money, time, and greenhouse gas emissions.

The paper proceeds as follows. We formalize the NPB inter-league problem in Section 2, providing the locations of the twelve teams as well as the inter-league schedule from the 2010 regular season. We use this to explain the concepts of *uniform* and *non-uniform* tournament scheduling, to motivate the *Bipartite Traveling Tournament Problem (BTTP)* in Section 3. Despite the NP-completeness of the general problem with $2n$ teams, we describe two heuristics that allow us to solve *BTTP* for the 12-team NPB, which we do in Section 4. In Sections 5 and 6, we present uniform and non-uniform inter-league schedules requiring 43285 and 42950 kilometres of total travel, respectively. We prove these schedules are optimal, achieving a 15.3% and 16.0% reduction in travel distance as compared to the 2010 NPB schedule.

The 2010 NPB Inter-League Schedule

Each NPB team plays in a home city somewhere within Japan, whose location is marked in Figure 1.

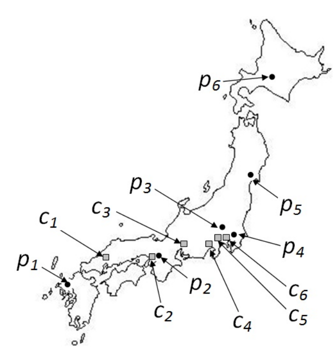

Figure 1: Location of the 12 teams in the NPB.

Let us explain the labelling. The Pacific League teams are p_1 (Fukuoka Hawks), p_2 (Orix Buffaloes), p_3 (Saitama Lions), p_4 (Chiba Marines), p_5 (Tohoku Eagles), and p_6 (Hokkaido Fighters). The Central League teams are c_1 (Hiroshima Carp), c_2 (Hanshin Tigers), c_3 (Chunichi Dragons), c_4 (Yokohama Baystars), c_5 (Yomiuri Giants), and c_6 (Yakult Swallows). For readability, we will refer to the teams by their labels rather than their full names.

Table 1 provides the 12×12 NPB *distance matrix* D, representing the distances between the home stadiums of each pair of teams in $\{p_1, \ldots, p_6, c_1, \ldots, c_6\}$. All distances are in kilometres. By symmetry, $D_{x,y} = D_{y,x}$ for any two teams x and y, and all diagonal entries are zero.

	c_1	c_2	c_3	c_4	c_5	c_6	p_1	p_2	p_3	p_4	p_5	p_6
c_1	0	323	488	808	827	829	258	341	870	857	895	1288
c_2		0	195	515	534	536	577	27	577	564	654	1099
c_3			0	334	353	355	742	213	396	383	511	984
c_4				0	37	35	916	533	63	58	364	886
c_5					0	7	926	552	51	37	331	896
c_6						0	923	554	48	39	333	893
p_1							0	595	958	934	1100	1466
p_2								0	595	582	670	1115
p_3									0	86	374	928
p_4										0	361	904
p_5											0	580
p_6												0

Table 1: The 12×12 NPB distance matrix.

In NPB inter-league play, each team in the Pacific League $P = \{p_i : 1 \leq i \leq 6\}$ plays four games against all six teams in the Central League $C = \{c_i : 1 \leq i \leq 6\}$, with one two-game set played at the home of the Pacific League team, and the other two-game set played at the home of the Central League team. All inter-league games take place during a five-week stretch between mid-May and mid-June, with no intra-league games occurring in that period.

Table 2 provides the 2010 NPB inter-league schedule, listing the twelve *sets*, with one set in each *time slot*, and two games in each set. In this table, as with all other schedules presented in this paper, home games are marked in bold.

	1	2	3	4	5	6	7	8	9	10	11	12
c_1	$\boldsymbol{p_5}$	$\boldsymbol{p_6}$	p_2	p_1	$\boldsymbol{p_3}$	$\boldsymbol{p_4}$	p_5	p_6	$\boldsymbol{p_1}$	$\boldsymbol{p_2}$	p_4	p_3
c_2	$\boldsymbol{p_6}$	$\boldsymbol{p_5}$	p_1	p_2	$\boldsymbol{p_4}$	$\boldsymbol{p_3}$	p_6	p_5	$\boldsymbol{p_2}$	$\boldsymbol{p_1}$	p_3	p_4
c_3	$\boldsymbol{p_1}$	$\boldsymbol{p_2}$	p_4	p_3	$\boldsymbol{p_5}$	$\boldsymbol{p_6}$	p_1	p_2	$\boldsymbol{p_4}$	$\boldsymbol{p_3}$	p_5	p_6
c_4	$\boldsymbol{p_3}$	$\boldsymbol{p_4}$	p_5	p_6	$\boldsymbol{p_1}$	$\boldsymbol{p_2}$	p_3	p_4	$\boldsymbol{p_5}$	$\boldsymbol{p_6}$	p_1	p_2
c_5	$\boldsymbol{p_4}$	$\boldsymbol{p_3}$	p_6	p_5	$\boldsymbol{p_2}$	$\boldsymbol{p_1}$	p_4	p_3	$\boldsymbol{p_6}$	$\boldsymbol{p_5}$	p_2	p_1
c_6	$\boldsymbol{p_2}$	$\boldsymbol{p_1}$	p_3	p_4	$\boldsymbol{p_6}$	$\boldsymbol{p_5}$	p_2	p_1	$\boldsymbol{p_3}$	$\boldsymbol{p_4}$	p_6	p_5
p_1	c_3	c_6	$\boldsymbol{c_2}$	$\boldsymbol{c_1}$	c_4	c_5	$\boldsymbol{c_3}$	$\boldsymbol{c_6}$	c_1	c_2	$\boldsymbol{c_4}$	$\boldsymbol{c_5}$
p_2	c_6	c_3	$\boldsymbol{c_1}$	$\boldsymbol{c_2}$	c_5	c_4	$\boldsymbol{c_6}$	$\boldsymbol{c_3}$	c_2	c_1	$\boldsymbol{c_5}$	$\boldsymbol{c_4}$
p_3	c_4	c_5	$\boldsymbol{c_6}$	$\boldsymbol{c_3}$	c_1	c_2	$\boldsymbol{c_4}$	$\boldsymbol{c_5}$	c_6	c_3	$\boldsymbol{c_2}$	$\boldsymbol{c_1}$
p_4	c_5	c_4	$\boldsymbol{c_3}$	$\boldsymbol{c_6}$	c_2	c_1	$\boldsymbol{c_5}$	$\boldsymbol{c_4}$	c_3	c_6	$\boldsymbol{c_1}$	$\boldsymbol{c_2}$
p_5	c_1	c_2	$\boldsymbol{c_4}$	$\boldsymbol{c_5}$	c_3	c_6	$\boldsymbol{c_1}$	$\boldsymbol{c_2}$	c_4	c_5	$\boldsymbol{c_3}$	$\boldsymbol{c_6}$
p_6	c_2	c_1	$\boldsymbol{c_5}$	$\boldsymbol{c_4}$	c_6	c_3	$\boldsymbol{c_2}$	$\boldsymbol{c_1}$	c_5	c_4	$\boldsymbol{c_6}$	$\boldsymbol{c_3}$

Table 2: The 2010 NPB inter-league schedule.

Whenever a team is scheduled for a road trip consisting of multiple away sets, the team doesn't return to their home city but rather proceeds directly to their next away venue. Furthermore, we assume that every team begins the tournament at home, and returns home after its last away game. For example, in Table 2, team c_1 would travel a total distance of $D_{c_1,p_2} + D_{p_2,p_1} + D_{p_1,c_1} + D_{c_1,p_5} + D_{p_5,p_6} + D_{p_6,c_1} + D_{c_1,p_4} + D_{p_4,p_3} + D_{p_3,c_1} = (D_{p_1,p_2} + D_{p_3,p_4} + D_{p_5,p_6}) + \sum_{j=1}^{6} D_{c_1,p_j} = 1261 + 4509 = 5770$.

Let $M_p = D_{p_1,p_2} + D_{p_3,p_4} + D_{p_5,p_6}$ and $M_c = D_{c_1,c_2} + D_{c_3,c_6} + D_{c_4,c_5}$. From Table 2, the total travel distance is $M_p + \sum_{j=1}^{6} D_{c_i,p_j}$ for each c_i and $M_c + \sum_{j=1}^{6} D_{p_i,c_j}$ for each p_i. In this inter-league schedule, the six teams in each league are grouped into three pairs based on geographic proximity, so that the teams in each league have three identical road trips lasting two sets (four games).

For example, each team c_i has road trips consisting of the pairs $\{p_1, p_2\}$, $\{p_3, p_4\}$, and $\{p_5, p_6\}$ in some order. From Table 1, we can show that the Central League teams travel 27205 kilometres, and the Pacific League teams travel 23929 kilometres, for a total of 51134 kilometres.

The sums $\sum_{j=1}^{6} D_{p_i,c_j}$ and $\sum_{j=1}^{6} D_{c_i,p_j}$ are fixed for each i, regardless of how the six teams are paired up. Thus, if the teams in each league play the same set of three road trips, the total distance is minimized whenever the sum $M_p + M_c$ is minimized. Note that both M_p and M_c represent the total edge weight of a perfect matching in a complete graph on 6 vertices.

Therefore, the optimal value of M_c is the minimum value of $D_{c_{\pi(1)},c_{\pi(2)}} + D_{c_{\pi(3)},c_{\pi(4)}} + D_{c_{\pi(5)},c_{\pi(6)}}$ over all permutations π of $\{1, 2, 3, 4, 5, 6\}$. The distance-optimal schedule occurs by making each team play their three road trips based on the other league's minimum-weight perfect matching.

From Figure 1 (or Table 1), the minimum-weight perfect matching for the Central League occurs when $\pi = (1, 2, 3, 4, 5, 6)$. By replacing the Pacific League road trips with the optimal matching $\{c_1, c_2\}$, $\{c_3, c_4\}$, and $\{c_5, c_6\}$, we can reduce each team's travel by $(D_{c_3,c_6} + D_{c_4,c_5}) - (D_{c_3,c_4} + D_{c_5,c_6}) = 51$ kilometres.

This "optimal" grouping was used by the NPB in 2009, with the twelve teams traveling a total of $51134 - 51 \times 6 = 50828$ kilometres. From Figure 1, it is clear that the perfect matching used in the other league, namely $\{p_1, p_2\}$, $\{p_3, p_4\}$, and $\{p_5, p_6\}$, has minimum weight.

Hence, if we adopt the framework of Table 2, where the teams play inter-league games by alternating home and away sets two at a time, then the optimal schedule requires 50828 kilometres of total team travel. We note that for any 12×12 distance matrix, we can rapidly generate the distance-optimal schedule by finding the best possible pairing of three road trips for each league and arranging the tournament schedule in the format of Table 2. This argument easily generalizes to the scenario where there are $2n$ teams in each league, since there is an $O(n^3)$ algorithm to determine a minimum-weight perfect matching for any weighted complete graph on $2n$ vertices (Lawler 1976).

In Table 2, every time slot has the property that the teams in each league either all play at home, or all play on the road. We say that such a schedule is *uniform*. By relaxing this constraint and including *non-uniform* schedules for consideration, we expand the search space.

However, the next result shows that if we impose the restriction that no team can play more than two consecutive sets at home or on the road, the distance-optimal schedule must be uniform, and have the same structure as Table 2.

Proposition 1 *Let P and C each consist of $2n$ teams, for some $n \geq 1$. Consider an inter-league tournament between P and C, where each pair of teams p_i and c_j plays two sets, with one set at each team's home venue. If every team can play at most two consecutive sets at home or on the road, then any distance-optimal inter-league schedule must be uniform, alternating home and away sets two at a time.*

Proof Let $P = \{p_i : 1 \leq i \leq 2n\}$ be the Pacific League teams and $C = \{c_i : 1 \leq i \leq 2n\}$ be the Central League teams. Define ILB_t to be the *individual lower bound* for team $t \in P \cup C$. This value represents the minimum possible distance that can be traveled by team t in order to complete its $4n$ sets, independent of the other teams' schedules.

Then a trivial lower bound for the total travel distance is $TLB = \sum_{t \in P \cup C} ILB_t$. If team p_i does not play their $2n$ road sets in pairs, there would be at least two single-set road trips which can be combined to reduce that team's travel distance, by the Triangle Inequality. Letting M_c be the minimum weight of a perfect matching for the Central League, we have $ILB_{p_i} = M_c + \sum_{j=1}^{2n} D_{p_i,c_j}$. Analogously, $ILB_{c_i} = M_p + \sum_{j=1}^{2n} D_{c_i,p_j}$. Therefore,

$$TLB = 2n(M_p + M_c) + 2\sum_{i=1}^{2n}\sum_{j=1}^{2n} D_{p_i,c_j}.$$

By the same construction as Table 2, we can construct a uniform inter-league schedule with total distance TLB, which equals 50828 in the case of the NPB distance matrix.

To complete the proof, we must show that any schedule having total distance TLB must be uniform, and have the structure of alternating home and away sets two at a time.

If a schedule has total distance TLB, by definition, each team t must travel a total distance of ILB_t. Suppose on the contrary that there exists a non-uniform distance-optimal schedule. By the Triangle Inequality, each team t must play their $2n$ road sets in n pairs, as otherwise that team's travel distance would exceed ILB_t.

For this schedule, let $H(P, s)$ and $R(P, s)$ be respectively the number of teams in P playing at home and on the road in time slot s. Similarly, define $H(C, s)$ and $R(C, s)$. For all $1 \leq s \leq 4n$, we have $H(P, s) + R(P, s) = H(C, s) + R(C, s) = 2n$, $H(P, s) = R(C, s)$ and $R(P, s) = H(C, s)$.

For each $a, b \in [H, R]$, let p_{ab} be the number of teams in P playing their first set at a and their second set at b. For example, p_{HR} is the number of teams in P playing their first set at home and their second set on the road. By definition, $p_{HH} + p_{HR} + p_{RH} + p_{RR} = 2n$. Similarly define c_{ab} so that $c_{HH} + c_{HR} + c_{RH} + c_{RR} = 2n$. Note that $p_{RH} = c_{RH} = 0$ since every team must play their road sets in pairs.

Since $H(P, 1) = R(C, 1)$ and $H(P, 2) = R(C, 2)$, we have $p_{HH} + p_{HR} = c_{RR}$ and $p_{HH} = c_{HR} + c_{RR}$. This implies that $p_{HR} + c_{HR} = 0$, forcing $p_{HR} = c_{HR} = 0$. Thus, $p_{HH} = c_{RR}$ and $c_{HH} = p_{RR}$.

Each team counted in p_{HH} must play on the road in set 3, as otherwise some team would play three consecutive home sets. But since road sets take place in pairs, each of these teams must also play on the road in set 4. Thus, $R(P, 4) \geq p_{HH}$. Similarly, each team counted in c_{HH} must play on the road in sets 3 and 4, implying that $H(C, 4) \leq c_{RR}$. Since $R(P, 4) = H(C, 4)$ and $p_{HH} = c_{RR}$, we have $H(C, 4) = c_{RR}$. In other words, each team counted in c_{RR} plays at home in sets 3 and 4, forcing them to play on the road in sets 5 and 6. The same pattern holds for each team counted in p_{RR}. Thus, $R(P, 6) + R(C, 6) \geq p_{RR} + c_{RR}$.

Each team counted in p_{HH} and c_{HH} plays at home in set 5, as otherwise some team would play three consecutive road sets. Furthermore, each of these teams must also play at home in set 6, as otherwise $2n = R(P, 6) + R(C, 6) \geq p_{RR} + c_{RR} + 1 = p_{RR} + p_{HH} + 1 = 2n + 1$, a contradiction.

Repeating this argument, each team counted in p_{HH} and c_{HH} must play two home sets followed by two road sets, and alternating that pattern until the end of the tournament. Conversely, each team counted in p_{RR} and c_{RR} must play the inverse alternating schedule of two road sets followed by two home sets.

Without loss, assume $p_{HH} > 0$. Suppose $c_{HH} > 0$. Then there exist two teams p_i and c_j that have the exact same home-road pattern HH-RR-HH-RR-$\cdots$-HH-RR. This is a contradiction as these two teams would then be unable to play each other. Thus, $c_{HH} = 0$, implying that $p_{RR} = 0$. We therefore have $p_{HH} = 2n$ and $c_{RR} = 2n$, proving that the tournament must be uniform, with the teams in each league alternating home and away sets two at a time. ∎

Therefore, we have solved the inter-league optimization problem for any $4n \times 4n$ distance matrix, given the constraint that every team can play at most *two* consecutive sets at home or on the road. By our analysis, there exists an $O(n^3)$ algorithm to construct an optimal tournament schedule, by pairing each team's $2n$ road sets according to the minimum-weight perfect matching. As shown in Proposition 1, any distance-optimal schedule *must* be uniform.

However, if we consider the constraint that every team can play at most *three* consecutive sets at home or on the road, then the problem of generating a distance-optimal inter-league schedule becomes extremely difficult. In fact, as we will explain in the next section, this problem becomes NP-complete, even when restricting our search space to the set of uniform schedules!

We analyze this extension to three-set home stands and road trips, since both scenarios frequently occur during NPB *intra*-league play, where teams play nine consecutive games (3 sets of 3 games) either at home or on the road. Since inter-league sets consist of two games, even if a team played three consecutive sets at home or on the road, that would only correspond to six games. This simple change from "at most 2" to "at most 3" motivates the *Bipartite Traveling Tournament Problem (BTTP)*, which we now present.

The Bipartite Traveling Tournament Problem

The challenge of creating a distance-optimal intra-league schedule motivated the *Traveling Tournament Problem* (TTP), in which every pair of teams plays two sets (i.e., a double round-robin tournament). The output is an optimal schedule that minimizes the sum total of distances traveled by the n teams as they move from city to city, subject to three constraints that ensure competitive balance:

(a) *at-most-three*: No team may have a home stand or road trip lasting more than three sets.

(b) *no-repeat*: A team cannot play against the same opponent in two consecutive sets.

(c) *each-venue*: Each pair of teams plays twice, with one set at each team's home venue.

The TTP involves both *integer programming* to prevent excessive travel, as well as *constraint programming* to create a schedule of home and road sets that meet all the feasibility requirements. While each problem is simple to solve on its own, its combination has proven to be extremely challenging (Easton, Nemhauser, and Trick 2001), even for small cases such as $n = 6$ and $n = 8$. The TTP has emerged as a popular area of study (Kendall et al. 2010) within the operations research community due to its surprising complexity.

We introduce the *Bipartite Traveling Tournament Problem (BTTP)*, the inter-league analogue of the TTP. Let $X = \{x_1, x_2, \ldots, x_n\}$ and $Y = \{y_1, y_2, \ldots, y_n\}$ be two leagues, where each pair of teams x_i and y_j (with $1 \leq i, j \leq n$) plays two sets, with one set at each venue. Given a $2n \times 2n$ distance matrix, the solution to *BTTP* is a distance-optimal double round-robin inter-league schedule satisfying the *at-most-three*, *no-repeat*, and *each-venue* constraints.

Let *BTTP** be the restriction of *BTTP* to the set of *uniform* tournament schedules. By definition, the solution to *BTTP** has total travel distance at least that of *BTTP*.

We proved that both *BTTP* and *BTTP** are NP-complete (Hoshino and Kawarabayashi 2011a) by obtaining a reduction from 3-SAT, the well-known NP-complete problem on boolean satisfiability.

To explain the difficulty of *BTTP*, we provide a simple illustration for the case $n = 3$ by providing two feasible tournaments in Table 3, with one uniform schedule and one non-uniform schedule.

	1	2	3	4	5	6
x_1	y_1	y_2	y_3	$\boldsymbol{y_1}$	$\boldsymbol{y_2}$	$\boldsymbol{y_3}$
x_2	y_2	y_3	y_1	$\boldsymbol{y_2}$	$\boldsymbol{y_3}$	$\boldsymbol{y_1}$
x_3	y_3	y_1	y_2	$\boldsymbol{y_3}$	$\boldsymbol{y_1}$	$\boldsymbol{y_2}$
y_1	$\boldsymbol{x_1}$	$\boldsymbol{x_3}$	$\boldsymbol{x_2}$	x_1	x_3	x_2
y_2	$\boldsymbol{x_2}$	$\boldsymbol{x_1}$	$\boldsymbol{x_3}$	x_2	x_1	x_3
y_3	$\boldsymbol{x_3}$	$\boldsymbol{x_2}$	$\boldsymbol{x_1}$	x_3	x_2	x_1

	1	2	3	4	5	6
x_1	y_3	y_2	$\boldsymbol{y_1}$	$\boldsymbol{y_3}$	y_1	$\boldsymbol{y_2}$
x_2	$\boldsymbol{y_1}$	y_3	$\boldsymbol{y_2}$	y_1	y_2	$\boldsymbol{y_3}$
x_3	$\boldsymbol{y_2}$	y_1	$\boldsymbol{y_3}$	y_2	y_3	$\boldsymbol{y_1}$
y_1	x_2	$\boldsymbol{x_3}$	x_1	$\boldsymbol{x_2}$	$\boldsymbol{x_1}$	x_3
y_2	x_3	$\boldsymbol{x_1}$	x_2	$\boldsymbol{x_3}$	$\boldsymbol{x_2}$	x_1
y_3	$\boldsymbol{x_1}$	$\boldsymbol{x_2}$	x_3	x_1	$\boldsymbol{x_3}$	x_2

Table 3: Two feasible inter-league tournaments for $n = 3$.

For each team, define a *trip* to be a pair of consecutive sets where that team doesn't play at the same location in time slots s and $s + 1$. In Table 3, the top schedule has 24 total trips, while the bottom schedule has 32 total trips.

Now let the teams x_1, x_3, y_1, and y_2 be located at $(0, 0)$ and let x_2 and y_3 be located at $(1, 0)$. Then the top schedule has total distance 16 and the bottom schedule has total distance 12. So minimizing trips does not correlate to minimizing total travel distance; while the former is a trivial problem, the latter is extremely difficult, even for the case $n = 3$. Of course, a brute-force enumeration of all possible tournaments is one approach, but this only works for small cases, and not for the NPB, which has $n = 6$.

The case $n = 6$ requires the optimal scheduling of $6 \times 12 = 72$ matches. The 12-team *BTTP* is comparable in difficulty to the 8-team TTP (with 56 total matches) and the 10-team TTP (with 90 total matches), both of which were solved recently on a benchmark data set using the aid of powerful computers running calculations over multiple processors (Trick 2011). Nonetheless, with the aid of two clever heuristics that we present in the next section, we can quickly solve the NP-complete problems *BTTP** and *BTTP* for the 12×12 NPB distance matrix.

Two Heuristics for *BTTP** and *BTTP*

We now present two theorems that tackle the Bipartite Traveling Tournament Problem. The first theorem is applicable for any n, and the second theorem is specific for the case $n = 6$. These results form the theoretical basis for our heuristics that solve *BTTP** and *BTTP* for the NPB. Before presenting our results, we provide several key definitions.

For each $t \in X \cup Y$, let S_t be the set of possible schedules that can be played by team t satisfying the *at-most-three* and *each-venue* constraints. Let $\pi_t \in S_t$ be a possible schedule for team t. For each π_t, we just list the opponents of the six *road* sets, and ignore the home sets, since we can determine the total distance traveled by team t just from the road sets. To give an example, below is a feasible schedule $\pi_{x_1} \in S_{x_1}$ for the case $n = 6$:

	1	2	3	4	5	6	7	8	9	10	11	12
x_1	y_1	y_6	$\boldsymbol{y}$	$\boldsymbol{y}$	y_3	y_5	y_4	$\boldsymbol{y}$	$\boldsymbol{y}$	$\boldsymbol{y}$	y_2	$\boldsymbol{y}$

In the above team schedule π_{x_1}, each $\boldsymbol{y}$ represents a home set played by x_1 against a unique opponent in Y. Note that π_{x_1} satisfies the *at-most-three* and *each-venue* constraints.

Let $\Phi = (\pi_{x_1}, \pi_{x_2}, \ldots, \pi_{x_n}, \pi_{y_1}, \pi_{y_2}, \ldots, \pi_{y_n})$, where $\pi_t \in S_t$ for each $t \in X \cup Y$. Since road sets of X correspond to home sets of Y and vice-versa, it suffices to list just the time slots and opponents of the n road sets in each π_t, since we can then uniquely determine the full schedule of $2n$ sets for every team $t \in X \cup Y$, thus producing an inter-league tournament schedule Φ. We note that Φ is a feasible solution to *BTTP* iff each team plays a unique opponent in every time slot, and no team schedule π_t violates the *no-repeat* constraint.

In the following sections, we will frequently refer to *team* schedules π_t and *tournament* schedules Φ. From the context it will be clear whether the schedule is for an individual team $t \in X \cup Y$, or for all $2n$ teams in $X \cup Y$.

As before, define ILB_t to be the individual lower bound of team t, the minimum possible distance that can be traveled by team t in order to complete its $2n$ sets.

For each $\pi_t \in S_t$, let $d(\pi_t)$ be the integer for which $d(\pi_t) + ILB_t$ equals the total distance traveled by team t when playing the schedule π_t. By definition, $d(\pi_t) \geq 0$.

For each $\Phi = (\pi_{x_1}, \ldots, \pi_{x_n}, \pi_{y_1}, \ldots, \pi_{y_n})$, define

$$d(\Phi) = \sum_{t \in X \cup Y} d(\pi_t).$$

Since $\sum ILB_t$ is fixed, the optimal solution to *BTTP* is the schedule Φ for which $d(\Phi)$ is minimized. This is the motivation for the function $d(\Phi)$.

For each subset $S_t^* \subseteq S_t$, define the *lower bound* function

$$B(S_t^*) = \min_{\pi_t \in S_t^*} d(\pi_t).$$

If $S_t^* = S_t$, then $B(S_t^*) = 0$ by the definition of ILB_t. For each subset S_t^*, we define $|S_t^*|$ to be its cardinality.

If n is a multiple of 3, we define for each team the set R_3^t as the subset of schedules in S_t for which the n road sets occur in $\frac{n}{3}$ blocks of three (i.e., team t takes $\frac{n}{3}$ three-set road trips). For example, in the top schedule of Table 3 (which has $n = 3$), every team t plays a schedule $\pi_t \in R_3^t$.

Finally, we define Γ to be a *global constraint* that fixes some subset of matches, and S_t^Γ to be the subset of schedules in S_t which are consistent with that global constraint. For example, if Γ is the simple constraint that forces y_2 to play against x_1 at home in time slot 3, then $S_{x_1}^\Gamma$ would only consist of the team schedules where slot 3 is a road set against y_2. If Γ is a much more complex global constraint (e.g. where the number of fixed matches is large), then each $|S_t^\Gamma|$ will be significantly less than $|S_t|$.

We present the first theorem, a powerful reduction heuristic that drastically cuts down the computation time.

Theorem 1 *Let M be a fixed positive integer. For any global constraint Γ, define for each $t \in X \cup Y$,*

$$Z_t = \left\{ \pi_t \in S_t^\Gamma : d(\pi_t) \le M + B(S_t^\Gamma) - \sum_{u \in X \cup Y} B(S_u^\Gamma) \right\}.$$

If $\Phi = (\pi_{x_1}, \ldots, \pi_{x_n}, \pi_{y_1}, \ldots, \pi_{y_n})$ is a feasible tournament schedule consistent with Γ so that $d(\Phi) \le M$, then for each $t \in X \cup Y$, team t's schedule π_t appears in Z_t.

Proof Consider all tournament schedules consistent with Γ. If there is no Φ with $d(\Phi) \le M$, then there is nothing to prove. So assume some schedule Φ satisfies $d(\Phi) \le M$. Letting $Q = \sum_{u \in X \cup Y} B(S_u^\Gamma)$, we have $M \ge d(\Phi) = \sum_{u \in X \cup Y} d(\pi_u) \ge \sum_{u \in X \cup Y} B(S_u^\Gamma)$, so that $M \ge Q$.

If $\pi_t \in Z_t$, then $Z_t \subseteq S_t^\Gamma$ implying that $d(\pi_t) \ge B(S_t^\Gamma)$.

Now suppose there exists some $v \in X \cup Y$ with $\pi_v \notin Z_v$. Since π_v is consistent with Γ, $\pi_v \in S_v^\Gamma$ and $d(\pi_v) > M + B(S_v^\Gamma) - Q \ge B(S_v^\Gamma)$. This is a contradiction, as

$$\begin{aligned} d(\Phi) &= d(\pi_v) + \sum_{u \in X \cup Y, u \neq v} d(\pi_u) \\ &> (M + B(S_v^\Gamma) - Q) + \sum_{u \in X \cup Y, u \neq v} B(S_u^\Gamma) \\ &= (M + B(S_v^\Gamma) - Q) + (Q - B(S_v^\Gamma)) \\ &= M. \end{aligned}$$

Hence, if $\Phi = (\pi_{x_1}, \ldots, \pi_{x_n}, \pi_{y_1}, \ldots, \pi_{y_n})$ is a feasible tournament schedule consistent with Γ so that $d(\Phi) \le M$, then $\pi_t \in Z_t$ for all $t \in X \cup Y$. ∎

Theorem 1 shows how to perform some reduction *prior* to propagation, and may be applicable to other problems. To apply this theorem, we will reduce *BTTP* to k scenarios where in each scenario the six home sets for four of the Pacific League teams are pre-determined. Expressing these scenarios as the global constraints $\Gamma_1, \Gamma_2, \ldots, \Gamma_k$, each Γ_i fixes 24 of the 72 total matches.

For every Γ_i, we determine Z_{c_j} for the Central League teams and by setting a low threshold M, we show that each $|Z_{c_j}|$ is considerably smaller than $|S_{c_j}^\Gamma|$, thus reducing the search space to an amount that can be quickly analyzed.

From there, we run a simple six-loop that generates all 6-tuples $(\pi_{c_1}, \pi_{c_2}, \pi_{c_3}, \pi_{c_4}, \pi_{c_5}, \pi_{c_6})$ that can appear in a feasible schedule Φ with $d(\Phi) \le M$. By Theorem 1, each $\pi_{c_j} \in Z_{c_j}$ for $1 \le j \le 6$. From this list of possible 6-tuples, we can quickly find the optimal schedule Φ which corresponds to the solution to *BTTP*.

Next, we present a special result that works only for the case $n = 6$, when two teams from one league are located quite far from the other 10 teams, forcing the distance-optimal schedule Φ to have a particular structure.

Theorem 2 *Let M be a fixed positive integer, and define $S_t^* = \{\pi_t \in S_t : d(\pi_t) \le M\}$. Suppose there exist two teams $x_i, x_j \in X = \{x_1, x_2, \ldots, x_6\}$ for which $S_{x_i}^* \subseteq R_3^{x_i}$, $S_{x_j}^* \subseteq R_3^{x_j}$, and for each team $y_k \in Y$, every schedule in $S_{y_k}^*$ has the property that y_k plays their road sets against x_i and x_j in two consecutive time slots.*

If $\Phi = (\pi_{x_1}, \ldots, \pi_{x_6}, \pi_{y_1}, \ldots, \pi_{y_6})$ is a feasible tournament schedule with $d(\Phi) \le M$ where each $\pi_t \in S_t^*$, then the team schedules π_{x_i} and π_{x_j} both have the home-road pattern HH-RRR-HH-RRR-HH; moreover, each team's six home slots must have the following structure for some permutation (a, b, c, d, e, f) of $\{1, 2, 3, 4, 5, 6\}$:

	1	2	3	4	5	6	7	8	9	10	11	12
x_i	$\boldsymbol{y_a}$	$\boldsymbol{y_b}$	y	y	y	$\boldsymbol{y_c}$	$\boldsymbol{y_d}$	y	y	y	$\boldsymbol{y_e}$	$\boldsymbol{y_f}$
x_j	$\boldsymbol{y_b}$	$\boldsymbol{y_a}$	y	y	y	$\boldsymbol{y_d}$	$\boldsymbol{y_c}$	y	y	y	$\boldsymbol{y_f}$	$\boldsymbol{y_e}$

Proof We first note that if π_{x_i} and π_{x_j} have the above structure, they satisfy all the given conditions since $\pi_{x_i} \in R_3^{x_i}$, $\pi_{x_j} \in R_3^{x_j}$, and every team $y_k \in Y$ plays road sets against x_i and x_j in two consecutive time slots. For example, y_d plays road sets against x_j in slot 6 and against x_i in slot 7. We now prove that π_{x_i} and π_{x_j} must have this structure.

For each team $x_t \in X$ and time slot $s \in [1, 12]$, define $O(x_t, s)$ to be the *opponent* of team x_t in set s. We define $O(x_t, s)$ only when x_t is playing at *home*; for the sets when x_t plays on the road, $O(x_t, s)$ is undefined.

Since $\pi_{x_i} \in S_{x_i}^*$ and $S_{x_i}^* \subseteq R_3^{x_i}$, there are four possible cases to consider:

(1) x_i plays set 1 at home, and sets 2 to 4 on the road.
(2) x_i plays sets 1 and 2 at home, and sets 3 to 5 on the road.
(3) x_i plays sets 1 to 3 at home, and sets 4 to 6 on the road.
(4) x_i plays sets 1 to 3 on the road, and set 4 at home.

We examine the cases one by one. In each, suppose there exists a feasible schedule Φ satisfying all the given conditions. We finish with case (2).

In (1), let $O(x_i, 1) = y_a$. Then $O(x_j, 2) = y_a$, since y_a must play road sets against x_i and x_j in consecutive time slots. Since $\pi_{x_j} \in R_3^{x_j}$ and x_j plays at home in set 2, x_j must also play at home in set 1. Thus, $O(x_j, 1) = y_b$ for some y_b, which forces $O(x_i, 2) = y_b$. This is a contradiction as x_i plays set 2 on the road.

In (3), let $O(x_i, 1) = y_a$, $O(x_i, 2) = y_b$, and $O(x_i, 3) = y_c$. Then $O(x_j, 2) = y_a$ and $O(x_j, 4) = y_c$. Either $O(x_j, 1) = y_b$ or $O(x_j, 3) = y_b$. In either case, we violate the *at-most-three* constraint or the condition that $\pi_{x_j} \in R_3^{x_j}$.

In (4), team x_i starts with a three-set road trip. In order to satisfy the *at-most-three* constraint, π_{x_i} must have the pattern RRR-HHH-RRR-HHH. Then this reduces to case (3), as we can read the schedule Φ backwards, letting $O(x_i, 12) = y_a$, $O(x_i, 11) = y_b$, $O(x_i, 10) = y_c$, and applying the argument in the previous paragraph.

In (2), let $O(x_i, 1) = y_a$ and $O(x_i, 2) = y_b$. Then $O(x_j, 2) = y_a$ and $O(x_j, 1) = y_b$. If $O(x_j, 3) = y_c$ for some y_c, then $O(x_i, 4) = y_c$, forcing x_i to play a single road set in slot 3. Thus, x_j must play on the road in set 3, and therefore also in sets 4 and 5. Hence, both x_i and x_j start with two home sets followed by three road sets. Since this is the only case remaining, by symmetry x_i and x_j must end with two home sets preceded by three road sets. Thus, these two teams must have the pattern HH-RRR-HH-RRR-HH.

In order for each y_k to play their road sets against x_i and x_j in two consecutive time slots, we must have $O(x_i, 6) = O(x_j, 7)$, $O(x_i, 7) = O(x_j, 6)$, $O(x_i, 11) = O(x_j, 12)$, and $O(x_i, 12) = O(x_j, 11)$. This completes the proof. ∎

We will use Theorem 2 to solve *BTTP*, since teams p_5 and p_6 are located quite far from the other ten teams (see Figure 1). This heuristic of isolating two teams and finding its common structure significantly reduces the search space and enables us to solve *BTTP* for the 12-team NPB in hours rather than weeks.

By applying these results, we do not require weeks of computation time on multiple processors. With these two heuristics, *BTTP** can be solved in *two minutes* and *BTTP* can be solved in less than *ten hours*, each on a single laptop. All of the code was written in Maple and compiled using Maplesoft 13 using a single Toshiba laptop under Windows with a single 2.10 GHz processor and 2.75 GB RAM. In the next two sections, we present the optimal uniform and non-uniform schedules for the NPB and justify their optimality.

Solution to *BTTP** for the NPB

We first compute ILB_t for each team $t \in P \cup C$, based on the distance matrix given in Table 1. We find that for each team t, a schedule $\pi_t \in S_t$ satisfies $d(\pi_t) = ILB_t$ only if $\pi_t \in R_3^t$, i.e., π_t plays its six road sets in two blocks of three.

We determine that $\sum ILB_t = \sum ILB_{p_i} + \sum ILB_{c_i} = 16686 + 26077 = 42763$. For any feasible inter-league tournament schedule Φ, the total distance traveled by the teams is $d(\Phi) + \sum ILB_t = d(\Phi) + 42763$.

Table 4 presents a uniform inter-league schedule Φ^* that is a feasible solution of *BTTP** with $d(\Phi^*) = (0 + 0 + 1 + 11 + 0 + 0) + (0 + 1 + 153 + 339 + 11 + 6) = 522$.

	1	2	3	4	5	6	7	8	9	10	11	12
p_1	c_2	c_3	c_1	$\boldsymbol{c_2}$	$\boldsymbol{c_3}$	$\boldsymbol{c_6}$	c_4	c_5	c_6	$\boldsymbol{c_1}$	$\boldsymbol{c_4}$	$\boldsymbol{c_5}$
p_2	c_4	c_6	c_5	$\boldsymbol{c_3}$	$\boldsymbol{c_6}$	$\boldsymbol{c_1}$	c_2	c_3	c_1	$\boldsymbol{c_2}$	$\boldsymbol{c_5}$	$\boldsymbol{c_4}$
p_3	c_3	c_1	c_2	$\boldsymbol{c_6}$	$\boldsymbol{c_4}$	$\boldsymbol{c_2}$	c_5	c_6	c_4	$\boldsymbol{c_5}$	$\boldsymbol{c_3}$	$\boldsymbol{c_1}$
p_4	c_5	c_4	c_6	$\boldsymbol{c_5}$	$\boldsymbol{c_2}$	$\boldsymbol{c_4}$	c_3	c_1	c_2	$\boldsymbol{c_6}$	$\boldsymbol{c_1}$	$\boldsymbol{c_3}$
p_5	c_1	c_2	c_3	$\boldsymbol{c_4}$	$\boldsymbol{c_1}$	$\boldsymbol{c_5}$	c_6	c_4	c_5	$\boldsymbol{c_3}$	$\boldsymbol{c_2}$	$\boldsymbol{c_6}$
p_6	c_6	c_5	c_4	$\boldsymbol{c_1}$	$\boldsymbol{c_5}$	$\boldsymbol{c_3}$	c_1	c_2	c_3	$\boldsymbol{c_4}$	$\boldsymbol{c_6}$	$\boldsymbol{c_2}$
c_1	$\boldsymbol{p_5}$	$\boldsymbol{p_3}$	$\boldsymbol{p_1}$	p_6	p_5	p_2	$\boldsymbol{p_6}$	$\boldsymbol{p_4}$	$\boldsymbol{p_2}$	p_1	p_4	p_3
c_2	$\boldsymbol{p_1}$	$\boldsymbol{p_5}$	$\boldsymbol{p_3}$	p_1	p_4	p_3	$\boldsymbol{p_2}$	$\boldsymbol{p_6}$	$\boldsymbol{p_4}$	p_2	p_5	p_6
c_3	$\boldsymbol{p_3}$	$\boldsymbol{p_1}$	$\boldsymbol{p_5}$	p_2	p_1	p_6	$\boldsymbol{p_4}$	$\boldsymbol{p_2}$	$\boldsymbol{p_6}$	p_5	p_3	p_4
c_4	$\boldsymbol{p_2}$	$\boldsymbol{p_4}$	$\boldsymbol{p_6}$	p_5	p_3	p_4	$\boldsymbol{p_1}$	$\boldsymbol{p_5}$	$\boldsymbol{p_3}$	p_6	p_1	p_2
c_5	$\boldsymbol{p_4}$	$\boldsymbol{p_6}$	$\boldsymbol{p_2}$	p_4	p_6	p_5	$\boldsymbol{p_3}$	$\boldsymbol{p_1}$	$\boldsymbol{p_5}$	p_3	p_2	p_1
c_6	$\boldsymbol{p_6}$	$\boldsymbol{p_2}$	$\boldsymbol{p_4}$	p_3	p_2	p_1	$\boldsymbol{p_5}$	$\boldsymbol{p_3}$	$\boldsymbol{p_1}$	p_4	p_6	p_5

Table 4: Solution to *BTTP** with total distance 43285 km.

We claim that Φ^* is an optimal solution, with total distance $d(\Phi^*) + \sum ILB_t = 522 + 42763 = 43285$. To establish this claim, we show that $d(\Phi) \geq 522$ for any uniform schedule $\Phi = (\pi_{p_1}, \pi_{p_2}, \ldots, \pi_{p_6}, \pi_{c_1}, \pi_{c_2}, \ldots, \pi_{c_6})$.

We exploit the uniformity of Φ and split *BTTP** into two separate optimization problems: first, we show that in any tournament schedule Φ, the six-tuple $(\pi_{p_1}, \pi_{p_2}, \ldots, \pi_{p_6})$ must satisfy $\sum d(\pi_{p_i}) \geq 12$. Then we show that for any Φ, the six-tuple $(\pi_{c_1}, \pi_{c_2}, \ldots, \pi_{c_6})$ must satisfy $\sum d(\pi_{c_i}) \geq 510$. This will prove that $d(\Phi) \geq 12 + 510 = 522$.

First we explain why an optimal uniform schedule Φ must have the home-road pattern HHH-RRR-HHH-RRR for one league, and RRR-HHH-RRR-HHH for the other league. Note that no league can play its six road sets in a way that a single road set is sandwiched between two home sets (e.g. HH-RR-HH-R-HH-RRR), since the time slots can be reordered to reduce the total travel distance by the Triangle Inequality. Thus, each league's road sets must take place in two blocks of three (as in Table 4) or three blocks of two (as in Table 2).

In Section 2, we showed that if the six road sets occur in three blocks of two, the minimum total travel distance is $23929 > \sum ILB_{p_i} + 522$ for the Pacific League and $26899 > \sum ILB_{c_i} + 522$ for the Central League, for a total of 50828 kilometres. Thus, if Φ is distance-optimal, then neither league can have its six road sets in three blocks of two, as otherwise $d(\Phi) > 522$. Hence, $\pi_t \in R_3^t$ for each team $t \in P \cup C$. Without loss, assume the Central League teams play at home in set 1.

Let Γ be the global constraint that every Pacific League team plays its road sets in time slots 1, 2, 3, 7, 8, and 9. Then for each t, S_t^Γ is the set of team schedules π_t consistent with Γ. Recall that each π_t just lists the time slots and opponents for the six road sets. Since π_t must have a fixed home-road pattern (either HHH-RRR-HHH-RRR or RRR-HHH-RRR-HHH), there are only 6! possible options for π_t. Hence, $|S_t^\Gamma| = 720$ for each $t \in P \cup C$.

Since each team's ILB_t is attained by some schedule $\pi_t \in R_3^t$, we have $B(S_t^\Gamma) = 0$. By Theorem 1, if Φ is a feasible schedule with $\sum d(\pi_{p_i}) \leq 12$, then $\pi_{p_i} \in Z_{p_i}$ where $Z_{p_i} = \{\pi_{p_i} \in S_{p_i}^\Gamma : d(\pi_{p_i}) \leq 12\}$.

By this reduction heuristic, we determine that $(|Z_{p_1}|, |Z_{p_2}|, \ldots, |Z_{p_6}|) = (32, 32, 32, 48, 24, 16)$; for example, only 32 of the 720 schedules in $S_{p_1}^\Gamma$ satisfy $d(\pi_{p_1}) \leq 12$. This reduces the search space considerably.

From there, we apply a simple six-loop procedure, where on the k^{th} step, we examine all possible schedules $\pi_{p_k} \in Z_{p_k}$ for team p_k and compare them to the set of feasible schedules $\{\pi_{p_1}, \pi_{p_2}, \ldots, \pi_{p_{k-1}}\}$, adding π_{p_k} to each feasible set if no two teams play against the same opponent c_j in the same time slot. Since the cardinality of each set Z_{p_i} is so small, Maplesoft can rapidly perform the six-loop computation (in just 7.8 seconds), generating 128 possible 6-tuples $(\pi_{p_1}, \pi_{p_2}, \ldots, \pi_{p_6})$ that can appear in Φ with $\sum d(\pi_{p_i}) = 12$. Furthermore, this computation shows that there is no 6-tuple with $\sum d(\pi_{p_i}) < 12$.

Thus, $\Phi = (\pi_{p_1}, \pi_{p_2}, \ldots, \pi_{p_n}, \pi_{c_1}, \pi_{c_2}, \ldots, \pi_{c_n})$ must satisfy $\sum d(\pi_{p_i}) \geq 12$. We now show that $\sum d(\pi_{c_i}) \geq 510$.

To establish this much-harder bound for the Central League teams, we can repeat the same process as above by setting the bound $M = 510$ to determine that $(|Z_{c_1}|, |Z_{c_2}|, \ldots, |Z_{c_6}|) = (160, 224, 272, 144, 152, 152)$. But a six-loop computation here would take far too long; instead, we provide a fast algorithm inspired by Theorem 2.

For each Central League team c_j, let $T_{c_j} \subseteq S^{\Gamma}_{c_j}$ be the subset of schedules for which c_j does *not* play road sets against p_5 and p_6 in two consecutive time slots. We find that $B(T_{c_1}) = 365$, i.e., any schedule $\pi_{c_1} \in T_{c_1}$ satisfies $d(\pi_{c_1}) \geq 365$. Similarly we have $B(T_{c_2}) = 314$, $B(T_{c_3}) = 153$, $B(T_{c_4}) = 313$, $B(T_{c_5}) = 324$, and $B(T_{c_6}) = 319$.

If each c_j has the pattern HHH-RRR-HHH-RRR, it is easy to see that (at least) two Central League teams cannot play road sets against p_5 and p_6 in two consecutive time slots. For example, if c_1 and c_2 are two such teams, then $\sum d(\pi_{c_i}) \geq d(c_1) + d(c_2) \geq B(T_{c_1}) + B(T_{c_2}) = 365 + 314 > 510$. Thus, if there exists Φ for which $\sum d(\pi_{c_i}) \leq 510$, then there can only be two such teams, and one of them must be c_3.

We split the analysis into five cases. For each $j \in \{1, 2, 4, 5, 6\}$, Φ is generated from selecting π_{c_3} and π_{c_j} from T_{c_3} and T_{c_j} respectively. For the other four teams in the Central League, we let $Z_{c_i} = \{\pi_{c_i} \in S^{\Gamma}_{c_i} : d(\pi_{c_i}) \leq 510 - B(T_{c_3}) - B(T_{c_j}) \leq 44\}$. We then run our six-loop, computing all possible 6-tuples $(\pi_{c_1}, \pi_{c_2}, \ldots, \pi_{c_6})$ satisfying the given conditions.

If $j = 4$, then $(|Z_{c_1}|, |Z_{c_2}|, |T_{c_3}|, |T_{c_4}|, |Z_{c_5}|, |Z_{c_6}|) = (16, 48, 16, 32, 64, 64)$, We find that in this case, there are 512 different 6-tuples with $\sum d(\pi_{c_i}) = 510$ and none with $\sum d(\pi_{c_i}) < 510$. As for the other values of j, there is no feasible 6-tuple with $\sum d(\pi_{c_i}) \leq 510$. Maplesoft is exceptionally fast; all five cases run in a total of 119 seconds.

Therefore, we have shown that in any uniform schedule $\Phi = (\pi_{p_1}, \pi_{p_2}, \ldots, \pi_{p_6}, \pi_{c_1}, \pi_{c_2}, \ldots, \pi_{c_6})$, we must have $\sum d(\pi_{p_i}) \geq 12$ and $\sum d(\pi_{c_i}) \geq 510$. Since Table 4 is a uniform inter-league schedule with $d(\Phi) = 12 + 510 = 522$, our proof is complete; our optimal solution to *BTTP** reduces the total travel distance by 7849 kilometres, or 15.3%, compared to the 2010 NPB schedule.

Solution to *BTTP* for the NPB

Table 5 presents an inter-league tournament schedule Φ that is a solution to *BTTP* with $d(\Phi) = (0+4+0+0+1+1) + (51+9+31+58+19+13) = 187$.

	1	2	3	4	5	6	7	8	9	10	11	12
p_1	$\boldsymbol{c_3}$	$\boldsymbol{c_5}$	$\boldsymbol{c_1}$	c_3	c_2	c_1	$\boldsymbol{c_6}$	$\boldsymbol{c_2}$	c_4	c_5	c_6	$\boldsymbol{c_4}$
p_2	$\boldsymbol{c_5}$	$\boldsymbol{c_3}$	c_2	c_1	c_3	$\boldsymbol{c_6}$	$\boldsymbol{c_1}$	c_4	c_5	c_6	$\boldsymbol{c_4}$	$\boldsymbol{c_2}$
p_3	$\boldsymbol{c_4}$	$\boldsymbol{c_2}$	c_6	c_5	c_4	$\boldsymbol{c_3}$	$\boldsymbol{c_5}$	$\boldsymbol{c_1}$	c_3	c_2	c_1	$\boldsymbol{c_6}$
p_4	$\boldsymbol{c_2}$	$\boldsymbol{c_4}$	$\boldsymbol{c_5}$	c_4	c_6	c_5	$\boldsymbol{c_3}$	$\boldsymbol{c_6}$	$\boldsymbol{c_1}$	c_3	c_2	c_1
p_5	$\boldsymbol{c_1}$	$\boldsymbol{c_6}$	c_4	c_6	c_5	$\boldsymbol{c_2}$	$\boldsymbol{c_4}$	c_3	c_2	c_1	$\boldsymbol{c_5}$	$\boldsymbol{c_3}$
p_6	$\boldsymbol{c_6}$	$\boldsymbol{c_1}$	c_3	c_2	c_1	$\boldsymbol{c_4}$	$\boldsymbol{c_2}$	c_5	c_6	c_4	$\boldsymbol{c_3}$	$\boldsymbol{c_5}$
c_1	p_5	p_6	p_1	$\boldsymbol{p_2}$	$\boldsymbol{p_6}$	$\boldsymbol{p_1}$	p_2	p_3	p_4	$\boldsymbol{p_5}$	$\boldsymbol{p_3}$	$\boldsymbol{p_4}$
c_2	p_4	p_3	$\boldsymbol{p_2}$	$\boldsymbol{p_6}$	$\boldsymbol{p_1}$	p_5	p_6	p_1	$\boldsymbol{p_5}$	$\boldsymbol{p_3}$	$\boldsymbol{p_4}$	p_2
c_3	p_1	p_2	$\boldsymbol{p_6}$	$\boldsymbol{p_1}$	$\boldsymbol{p_2}$	p_3	p_4	$\boldsymbol{p_5}$	$\boldsymbol{p_3}$	$\boldsymbol{p_4}$	p_6	p_5
c_4	p_3	p_4	$\boldsymbol{p_5}$	$\boldsymbol{p_4}$	$\boldsymbol{p_3}$	p_6	p_5	$\boldsymbol{p_2}$	$\boldsymbol{p_1}$	$\boldsymbol{p_6}$	p_2	p_1
c_5	p_2	p_1	p_4	$\boldsymbol{p_3}$	$\boldsymbol{p_5}$	$\boldsymbol{p_4}$	p_3	$\boldsymbol{p_6}$	$\boldsymbol{p_2}$	$\boldsymbol{p_1}$	p_5	p_6
c_6	p_6	p_5	$\boldsymbol{p_3}$	$\boldsymbol{p_5}$	$\boldsymbol{p_4}$	p_2	p_1	p_4	$\boldsymbol{p_6}$	$\boldsymbol{p_2}$	$\boldsymbol{p_1}$	p_3

Table 5: Solution to *BTTP* with total distance 42950 km.

In Table 5, we see that only seven of the twelve teams satisfy $\pi_t \in R^t_3$, namely c_1 and all six of the Pacific League teams. However, unlike Table 4, every Central League team in this schedule plays road sets against p_5 and p_6 in consecutive time slots. This explains why each $d(c_j)$ in Φ is small.

We claim that Φ is an optimal solution, with total distance $d(\Phi) + \sum ILB_t = 187 + 42763 = 42950$. To prove this, we set $M = 187$. Define $S^*_t = \{\pi_t \in S_t : d(\pi_t) \leq M\}$, from which we determine that $S^*_{p_5} \subseteq R^{p_5}_3$ and $S^*_{p_6} \subseteq R^{p_6}_3$.

In the previous section, we defined $T_{c_i} \subseteq S^{\Gamma}_{c_i}$ to be the subset of schedules for which c_i does *not* play their road sets against p_5 and p_6 in two consecutive time slots. From this, we showed that $B(T_{c_3}) = 153$, and that $B(T_{c_j}) > M = 187$ for $j \in \{1, 2, 4, 5, 6\}$. We claim that if Φ satisfies $d(\Phi) \leq 187$, then $\pi_{c_j} \notin T_{c_j}$ for all $1 \leq j \leq 6$.

It suffices to prove the claim for $j = 3$. There are 144 schedules in T_{c_3}, all of which belong to the set $R^{c_3}_3$. For example, one such schedule π_{c_3} is

	1	2	3	4	5	6	7	8	9	10	11	12
c_3	$\boldsymbol{p}$	p_2	p_1	p_6	$\boldsymbol{p}$	$\boldsymbol{p}$	$\boldsymbol{p}$	p_3	p_4	p_5	$\boldsymbol{p}$	$\boldsymbol{p}$

Suppose there exists a tournament schedule Φ with $d(\Phi) \leq 187$ and $\pi_{c_3} \in T_{c_3}$. There are nine possible home-road patterns for $\pi_{p_5} \in R^{p_5}_3$ (e.g. HHH-RRR-H-RRR-HH and H-RRR-HHH-RRR-HH), each of which gives rise to $6! = 720$ possible orderings for the six home sets. Thus, there are $9 \times 720 = 6480$ ways we can select the time slots and opponents for the six home sets in π_{p_5}. Similarly, there are 6480 ways to do this for π_{p_6}. A simple Maplesoft procedure shows that only 140 of the 6480^2 possible pairs (π_{p_5}, π_{p_6}) are consistent with at least one $\pi_{c_3} \in T_{c_3}$.

For each of these 140 cases, define the global constraints $\Gamma_1, \Gamma_2, \ldots, \Gamma_{140}$, obtained from fixing the twelve home sets in $\{\pi_{p_5}, \pi_{p_6}\}$. For each Γ_k, define for each $j \in \{1, 2, 4, 5, 6\}$ the set $Z_{c_j} = \{\pi_{c_j} \in S^{\Gamma_k}_{c_j} : d(\pi_{c_j}) \leq M - B(T_{c_3}) = 34\}$. Then we run our six-loop to compute all possible 6-tuples $(\pi_{c_1}, \pi_{c_2}, \ldots, \pi_{c_6})$ satisfying the given conditions with $\pi_{c_3} \in T_{c_3}$ and $\pi_{c_j} \in Z_{c_j}$ for $j \neq 3$. Within twenty minutes, Maplesoft solves all 140 cases and returns no feasible 6-tuples that can appear in a schedule Φ with $d(\Phi) \leq 187$.

Therefore, in Φ, each c_j must play road sets against p_5 and p_6 in consecutive time slots. Thus, teams p_5 and p_6 satisfy the conditions of Theorem 2. Hence, the home-road pattern of π_{p_5} and π_{p_6} in Φ must be HH-RRR-HH-RRR-HH.

Without loss, assume that p_5 plays a home set against c_1 within the first *six* time slots; otherwise we can read the schedule backwards by symmetry. Thus, there are $\frac{6!}{2} = 360$ ways to assign opponents to the six home sets in π_{p_5}. By Theorem 2, each of these 360 arrangements uniquely determines the six home sets in π_{p_6}.

A short calculation shows that in order for $d(\Phi) \leq M = 187$, teams p_1 and p_3 must also play their six road sets in two blocks of three. In other words, $\pi_{p_1} \in R_3^{p_1}$ and $\pi_{p_3} \in R_3^{p_3}$. As mentioned earlier, there are $9 \times 6!$ possible ways to select the six home sets for each of π_{p_1} and π_{p_3}.

Thus, there are $360 \times (9 \cdot 6!) \times (9 \cdot 6!)$ ways we can select the 24 home sets played by the teams in $\{p_1, p_3, p_5, p_6\}$. We eliminate all scenarios in which some p_i and p_j play against some c_k in the same time slot. For the possibilities that remain, we create a global constraint to apply Theorem 1.

Let $\{\Gamma_1, \Gamma_2, \ldots, \Gamma_k\}$ be the complete set of global constraints derived from the above process, where each Γ_i fixes 24 of the 72 matches, corresponding to the home sets of $\{p_1, p_3, p_5, p_6\}$. The reduction heuristic of Theorem 1 allows us to quickly verify the existence of feasible tournament schedules Φ consistent with Γ_i for which $d(\Phi) \leq M$.

To explain this procedure, let us illustrate with the inter-league schedule in Table 5. Let Γ be the constraint that fixes the 24 home sets of teams p_1, p_3, p_5, and p_6 in Table 5. Then $S_{c_5}^{\Gamma}$, defined as the subset of schedules in S_{c_5} consistent with Γ, consists only of team schedules π_{c_5} for which c_5 plays road sets against p_1 in slot 2, p_3 in slot 7, p_5 in slot 11, and p_6 in slot 12.

We find that there are only 11 schedules $\pi_{c_5} \in S_{c_5}^{\Gamma}$ with $d(\pi_{c_5}) \leq M$ that are consistent with Γ. Furthermore, each $d(\pi_{c_5}) \in \{19, 41, 46, 48\}$, implying that $B(S_{c_5}^{\Gamma}) = 19$. Similarly, we can calculate the other values of $B(S_{c_j}^{\Gamma})$.

We find that $\sum_{j=1}^{6} B(S_{p_j}^{\Gamma}) = 0$ and $\sum_{j=1}^{6} B(S_{c_j}^{\Gamma}) = 51 + 9 + 31 + 58 + 19 + 13 = 181$, implying that $Z_{c_5} = \{\pi_{c_5} \in S_{c_5}^{\Gamma} : d(\pi_{c_5}) \leq 187 + 19 - 181 = 25\}$. Hence, Z_{c_5} reduces to just the *two* schedules with $d(\pi_{c_5}) = 19$, including the team schedule π_{c_5} in Table 5.

By Theorem 1, any schedule Φ consistent with Γ satisfying $d(\Phi) \leq M$ must have the property that $\pi_t \in Z_t$ for each team t. Since each $|Z_{c_j}|$ is small, the calculation is extremely fast. Of course, if any $|Z_{c_j}| = 0$, then no schedule Φ can exist.

This algorithm, based on Theorems 1 and 2, runs in 34716 seconds in Maplesoft (just under 10 hours). Maplesoft generates zero inter-league schedules with $d(\Phi) < 187$ and 14 inter-league schedules with $d(\Phi) = 187$, including the schedule given in Table 5. Since we made the assumption that p_5 plays a home set against c_1 within the first six time slots, there are actually twice as many distance-optimal schedules by reading each schedule Φ backwards.

In each of the 28 distance-optimal schedules Φ, we find that $(d(\pi_{p_1}), d(\pi_{p_2}), \ldots, d(\pi_{p_6})) = (0, 4, 0, 0, 1, 1)$ and $(d(\pi_{c_1}), d(\pi_{c_2}), \ldots, d(\pi_{c_6})) = (51, 9, 31, 58, 19, 13)$.

Therefore, we have proven that Table 5 is an optimal inter-league schedule for the NPB, reducing the total travel distance by 8184 kilometres, or 16.0%, compared to the 2010 NPB schedule.

Conclusion

In this paper, we introduced the Bipartite Traveling Tournament Problem and applied it to the Nippon Professional Baseball (NPB) league, illustrating the richness and complexity of bipartite tournament scheduling. There may be other sports leagues for which *BTTP* is applicable. We can also expand our analysis to model *tripartite* and *multipartite* tournament scheduling, where a league is divided into three or more conferences. A specific example of this is the newly-created Super 15 Rugby League, consisting of five teams from South Africa, Australia, and New Zealand. In addition to intra-country games, each team plays four games (two home and two away) against teams from each of the other two countries. It would be interesting to see whether we can determine the distance-optimal tripartite tournament schedule for this rugby league using our two heuristics.

While the solution to the uniform *BTTP** was relatively simple, our solution to the non-uniform *BTTP* required 10 hours of computations. Furthermore, we were only able to solve *BTTP* by applying Theorem 2, whose requirements would not hold for a randomly-selected 12×12 distance matrix. As a result, we require a more sophisticated technique that improves upon our two heuristics, perhaps using methods in constraint programming and integer programming, such as a hybrid CP/IP. We wonder if there exists a general algorithm that would solve *BTTP* given any distance matrix, for "small" values of n such as $n = 6$, $n = 7$, and $n = 8$. We leave this as a challenge for the interested reader.

Acknowledgements

This research has been partially supported by the Japan Society for the Promotion of Science (Grant-in-Aid for Scientific Research), the C & C Foundation, the Kayamori Foundation, and the Inoue Research Award for Young Scientists.

References

Easton, K.; Nemhauser, G.; and Trick, M. 2001. The traveling tournament problem: description and benchmarks. *Proceedings of the 7th International Conference on Principles and Practice of Constraint Programming* 580–584.

Hoshino, R., and Kawarabayashi, K. 2011a. The inter-league extension of the traveling tournament problem and its application to sports scheduling. *Proceedings of the 25th AAAI Conference on Artificial Intelligence* (to appear).

Hoshino, R., and Kawarabayashi, K. 2011b. The multi-round balanced traveling tournament problem. *Proceedings of the 21st International Conference on Automated Planning and Scheduling* (to appear).

Kendall, G.; Knust, S.; Ribeiro, C.; and Urrutia, S. 2010. Scheduling in sports: An annotated bibliography. *Computers and Operations Research* 37:1–19.

Lawler, E. 1976. *Combinatorial Optimization: Networks and Matroids*. New York: Holt, Rinehart, and Winston.

Trick, M. 2011. Challenge traveling tournament problems. [Online; accessed 22-April-2011].

Wikipedia. 2011. List of sports attendance figures. [Online; accessed 22-April-2011].