Sports Scheduling with Genetic Algorithms

Keith Brady | Kerry Cook | Nick De Heras

What is Sports Scheduling?

- Problem - find an optimal schedule for a set of teams over a period of time that satisfy a particular set of constraints
- Applications - MLB
  - 30 teams where each team plays 162 games
  - How to find a season schedule that optimizes profit?

Our Problem - A General Case

General Round Robin

- Each team in the league must play every other team exactly once
- No team can play two games at the same time

Example

- Four teams : A, B, C, and D
- Below is an example for a feasible schedule

Round 1	Round 2	Round 3
A vs. B, C vs. D	A vs. C, B vs. D	A vs. D, B vs. C

Our Problem - NESCAC Men’s Soccer

10 Teams in the Conference

- Amherst, Williams College, Middlebury, Connecticut College, Tufts, Bates College, Bowdoin, Colby, Trinity, and Wesleyan
- In a season, each team plays each other exactly once

What are we trying to find?

- Build a schedule for the NESCAC men’s soccer season
    - Determines dates and Home/away teams
- Need an algorithm that runs quickly and finds a desirable schedule
    - What makes a schedule desirable?

Constraints - NESCAC Men's Soccer

Hard Constraints- Must be enforced

- Schedule Spans 8 weeks
  - 8 weekend games and 1 mid week
- Teams play each other exactly once
- 5 away games, 4 are home (or Vice Versa)
- Can't play more than two consecutive away games
- Teams must be home on given dates
- Rival triplets play one home and one away
  - Amherst, Bates, and Williams -Amherst is home for Bates and Away for Williams

Soft Constraints-makes schedules more desirable

- No team plays all opponents that are furthest away in one season
- Midweek games less than 2 hrs away
- Rival matches played during preferred part of season

Genetic Algorithm

- Starts with a random population of feasible solutions
- selects parent chromosomes based on their fitness values
- Parents mate and create a new population of n individuals

- Chromosomes - NESCAC Algorithm
    - 1st Chromosome- order of teams
    - 2nd Chromosome - determines the week team plays
- Steady State Reproduction
    - Keep an elite portion of the population steady
    - Mate elite with members of the non-weak portion of population

Direct Encoding

- The schedule itself is a chromosome
- In our case, a 10x10 array
- Would require extremely complex operators

Indirect Encoding

- Permutation of teams acts as chromosome
- Order of the teams influences the schedule produced
- Easy to perform genetic operations on

Applying the Algorithm

1. Schedule nine rounds of games so each team plays each other once
2. Assign each game a location, based on a series of constraints
3. Evaluate the “fitness” of the schedule
4. Generate a population and reproduce

1 Decoding the Chromosomes

Round 1 Schedule

```
('Bow', 'Tri', 'Bat', 'Amh', 'Wil', 'Col')
('Bat', 'Bow', 'Wil', 'Amh', 'Col', 'Tri')

Matchups:          ('Bow', 'Col'), ('Tri', 'Wil'), ('Bat', 'Amh')
Week Assignment: permutation2.index('Tri') = 5
```

Round 2 Schedule

```
('Bow', 'Col', 'Tri', 'Bat', 'Amh', 'Wil')
('Bat', 'Tri', 'Bow', 'Wil', 'Amh', 'Col')

Matchups:          ('Bow', 'Wil'), ('Col', 'Amh'), ('Tri', 'Bat')
Week Assignment: permutation2.index('Tri') = 2
```

2

Choosing the Home Teams

- Iterate through unscheduled season generated in step 1
- Apply a series of prioritized rules

1. If the game has been assigned a location already because it is part of a triplet, continue to next game
2. If the game is between two teams in a triplet (i.e., Amherst vs. Williams):
   a. If one of the other triplet games has been assigned a location (i.e., Amherst at Wesleyan), then set the home team of this game (i.e., Williams at Amherst) and the other game in the triplet (i.e., Wesleyan at Williams)
   b. If none of the games in the triplet have been scheduled yet, continue on to next rule
3. If team 1 requires a home game on this date, make it the home team
4. If team 2 requires a home game on this date, make it the home team
5. Let c_1 and c_2 be the number of consecutive home games for team 1 and team 2 respectively:
   a. If $c_1 > c_2$ and $c_1 \geq 2$, make team 1 the home team
   b. If $c_2 > c_1$ and $c_2 \geq 2$, make team 1 the home team
6. Set the team that has few home games scheduled to be the home team. If both have an equal number of home games, arbitrarily assign team one to play at home

3 Fitness Evaluation

- Travel Time Rating
  - Rates the average time spent traveling by each team
- Consecutive Rating
  - Rates number of times consecutive away games exceeds 2
- Location Balance Rating
  - Rates number of teams not scheduled with 4/5 home games

$$r_1 = \sqrt{\frac{\sum_{i=0}^{N-1} (avg_i - act_i)^2}{N}}$$

$$E = \sum_{i=1}^{6} f_i r_i$$

4 Crossover

Order-based

Parent 1	A	B	C	D	E	F	G
Parent 2	F	C	G	A	D	B	E

⇒

Child 1	A	B	F	D	C	E	G
Child 2	F	C	B	A	D	G	E

Mutation-based

Parent 1	A	B	C	D	E	F	G
Parent 2	F	C	G	A	D	B	E

⇒

Child 1	A	C	G	E	D	B	F
Child 2	G	A	C	D	E	F	B

4 Mutation

Position-based

- Remove an element from one position and insert into another

Order-based

- Swap two random elements in a chromosome

Scramble Sublist

- Pick a sublist of consecutive elements in the chromosome and randomly shuffle them

4 NESCAC Reproduction

Parameters

- *n* = population size
- *e* = elite percentage
- *c* = culled percentage
- W = mutation weight
- M = mutation rate
- X = crossover rate

Our Values

- *n* = 50
- *e* = 0.2
- *c* = 0.4
- W = 0.2
- M = 0.2
- X = 0.1

1. Promote the top $n * e$ individuals to the next generation
2. Loop until $n(1-e)$ unique offspring are created:
   a. Select one parent from top $n(1-c)$ chromosomes using binary tournament selection
   b. For each permutation in the chromosome:
      i. Apply crossover or mutation
      ii. If crossover is chosen, select additional parent chromosome
      iii. Randomly choose operator and apply to chromosome
   c. If new offspring is unique, add it to the new generation

Our NESCAC Schedule

```
Chromosome: (('Con', 'Tri', 'Bat', 'Mid', 'Wil', 'Wes', 'Tuf', 'Col', 'Bow', 'Amh'), ('Bow', 'Mid', 'Tri', 'Bat', 'Col', 'Tuf', 'Amh', 'Wes', 'Con', 'Wil'))

Team Schedule:

        Tri {'Amh': 'away', 'Bat': 'away', 'Wil': 'home', 'Mid': 'home', 'Bow': 'home', 'Wes': 'away', 'Col': 'home', 'Tuf': 'home', 'Con': 'away'}
        Amh {'Tri': 'home', 'Bat': 'home', 'Wil': 'away', 'Mid': 'away', 'Bow': 'home', 'Wes': 'home', 'Col': 'away', 'Tuf': 'away', 'Con': 'home'}
        Bat {'Tri': 'home', 'Amh': 'away', 'Wil': 'away', 'Mid': 'home', 'Bow': 'away', 'Wes': 'away', 'Col': 'away', 'Tuf': 'home', 'Con': 'home'}
        Wil {'Tri': 'away', 'Amh': 'home', 'Bat': 'home', 'Mid': 'away', 'Bow': 'home', 'Wes': 'home', 'Col': 'home', 'Tuf': 'away', 'Con': 'away'}
        Mid {'Tri': 'away', 'Amh': 'home', 'Bat': 'away', 'Wil': 'home', 'Bow': 'home', 'Wes': 'away', 'Col': 'away', 'Tuf': 'home', 'Con': 'home'}
        Bow {'Tri': 'away', 'Amh': 'away', 'Bat': 'home', 'Wil': 'away', 'Mid': 'away', 'Wes': 'home', 'Col': 'away', 'Tuf': 'home', 'Con': 'home'}
        Wes {'Tri': 'home', 'Amh': 'away', 'Bat': 'home', 'Wil': 'away', 'Mid': 'home', 'Bow': 'away', 'Col': 'home', 'Tuf': 'away', 'Con': 'home'}
        Col {'Tri': 'away', 'Amh': 'home', 'Bat': 'home', 'Wil': 'away', 'Mid': 'home', 'Bow': 'home', 'Wes': 'away', 'Tuf': 'home', 'Con': 'away'}
        Tuf {'Tri': 'away', 'Amh': 'home', 'Bat': 'away', 'Wil': 'home', 'Mid': 'away', 'Bow': 'away', 'Wes': 'home', 'Col': 'away', 'Con': 'home'}
        Con {'Tri': 'home', 'Amh': 'away', 'Bat': 'away', 'Wil': 'home', 'Mid': 'away', 'Bow': 'away', 'Wes': 'away', 'Col': 'home', 'Tuf': 'away'}

Season:

        1 [['Con', 'home', 'Col', 'away'], ['Bow', 'home', 'Tuf', 'away'], ['Amh', 'home', 'Wes', 'away'], ['Tri', 'home', 'Wil', 'away'], ['Bat', 'home', 'Mid', 'away']]
        2 [['Con', 'away', 'Tuf', 'home'], ['Col', 'away', 'Wes', 'home'], ['Bow', 'away', 'Wil', 'home'], ['Amh', 'away', 'Mid', 'home'], ['Tri', 'away', 'Bat', 'home']]
        3 [['Con', 'away', 'Wes', 'home'], ['Tuf', 'home', 'Wil', 'away'], ['Col', 'home', 'Mid', 'away'], ['Bow', 'home', 'Bat', 'away'], ['Amh', 'home', 'Tri', 'away']]
        4 [['Con', 'home', 'Wil', 'away'], ['Wes', 'home', 'Mid', 'away'], ['Tuf', 'away', 'Bat', 'home'], ['Col', 'away', 'Tri', 'home'], ['Bow', 'away', 'Amh', 'home']]
        5 [['Con', 'away', 'Mid', 'home'], ['Wil', 'home', 'Bat', 'away'], ['Wes', 'home', 'Tri', 'away'], ['Tuf', 'home', 'Amh', 'away'], ['Col', 'home', 'Bow', 'away']]
        6 [['Con', 'away', 'Bat', 'home'], ['Mid', 'away', 'Tri', 'home'], ['Wil', 'home', 'Amh', 'away'], ['Wes', 'away', 'Bow', 'home'], ['Tuf', 'away', 'Col', 'home']]
        7 [['Con', 'home', 'Tri', 'away'], ['Bat', 'away', 'Amh', 'home'], ['Mid', 'home', 'Bow', 'away'], ['Wil', 'home', 'Col', 'away'], ['Wes', 'away', 'Tuf', 'home']]
        8 [['Con', 'away', 'Amh', 'home'], ['Tri', 'home', 'Bow', 'away'], ['Bat', 'away', 'Col', 'home'], ['Mid', 'home', 'Tuf', 'away'], ['Wil', 'home', 'Wes', 'away']]
        9 [['Con', 'away', 'Bow', 'home'], ['Amh', 'away', 'Col', 'home'], ['Tri', 'home', 'Tuf', 'away'], ['Bat', 'away', 'Wes', 'home'], ['Mid', 'home', 'Wil', 'away']]

Rating:       33.1961891371

Mean Fitness: 2.06539637903

Time Rating:        3.19618913708
Consecutive Rating: 2
Location Balance:   1
```

Trinity's 2015 Schedule

SEPTEMBER	
Tue. 8	at Worcester St.
Wed. 9	Rivier
Sat. 12	Williams *
Tue. 15	Westfield St.
Sat. 19	Hamilton *
Sun. 20	at Tufts *
Sat. 26	at Colby *
Tue. 29	at Connecticut Col. *

OCTOBER	
Sat. 3	at Bates *
Tue. 6	at Wheaton (Mass.)
Sat. 10	Bowdoin *
Tue. 13	at Western Conn. St.
Sat. 17	at Middlebury *
Wed. 21	Wesleyan (Conn.) *
Wed. 28	Amherst *
Sat. 31	at Middlebury %

Trinity's Schedule Last Season

- Trinity traveled 16.5 hours
- Played 5 away games and 4 home games

Under our Optimal Schedule

- Trinity traveled 7.25 hours
- Played 4 away games and 5 home games

Conclusions

- Stochastic nature has heavy influence on fitness of each iteration
  - Increasing number of generations and increasing sample size helped optimums converge
  - High variance in running time
- Trade-offs abound
  - Schedule with low “Time Rating” but high “Consecutive Rating” vs. schedule with higher “Time Rating” but low “Consecutive Rating”
- Ultimately, far superior to system of rearranging index cards