Probabilistic Reasoning for Enhancing Decision Making in Elite Sports

Bahadorreza Ofoghi[1]; John Zeleznikow[1,2]; Clare MacMahon[1,3]; Dan Dwyer[4]; Ann Nicholson[5]; Cathryn Pruscino[6]

[1] Institute of Sport, Exercise, and Active Living, Victoria University, Melbourne
[2] School of Management and Information Systems, Victoria University, Melbourne
[3] School of Sport and Exercise Science, Victoria University, Melbourne
[4] School of Exercise and Nutrition Sciences, Deakin University, Geelong
[5] Clayton School of Information Technology, Monash University, Melbourne
[6] Victorian Institute of Sport, Melbourne

November 27, 2012

Decision Making in Elite Sports

- Types of decisions
  - tactics, e.g., choosing certain formations in soccer
  - action choices, e.g., how to return certain serves in tennis
  - strategy
    - athlete selection and training, e.g., using certain combinations of gymnasts in a competition
    - race planning, e.g., how to finish each 500m sector of 2000m rowing races

Decision Making in Elite Sports

- There is uncertainty
  - complexity of interactions
  - environmental conditions
- There are heterogenous sources of evidence, e.g.,
  - time
  - ranking
  - psychological preparedness
  - pre-season training conditions
  - weather conditions
  - etc.

BNs in Track Cycling Omnium

- A six-event competition
  - flying time trial
  - time trial
  - individual pursuit
  - scratch race
  - points race
  - elimination race

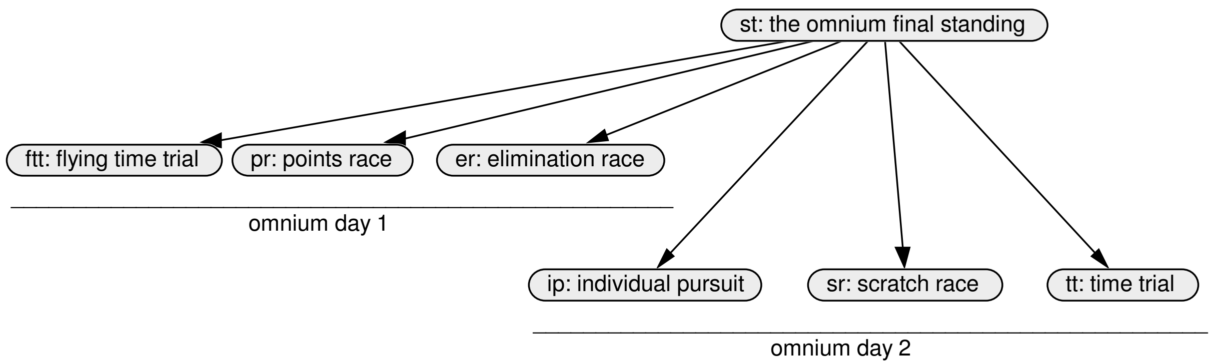
- Winning criterion
  - riders get scores according to their ranking in each individual component
  - winner is the rider with the least overall score

BNs in Track Cycling Omnium

- Research matters
  - what is the likelihood of finishing in certain overall places given the ranking of a rider in completed components?
    - BN modeling
  - what is the best possible combination of rankings in the upcoming components that maximizes the likelihood of finishing in certain overall places?
    - BN modeling + combinatorial optimization

BNs in Track Cycling Omnium

- Structure



- The joint probability function

$$p(e, \bar{h}) = \prod_i p(h_i|e)$$
$$e = fs : final\ standing$$
$$\bar{h} = \{h_1, h_2, \ldots, h_6\}$$

BNs in Track Cycling Omnium

- Evaluation
  - data: competition results since 2009
    - category 1: medal winners ranked 1-3
    - category 2: non-medal winners ranked 4-10
    - category 3: non-medal winners ranked >10
  - CPT learning: counting-learning
  - procedure: LOOCV
  - interface: Netica + Netica API (C#.Net)

BNs in Track Cycling Omnium

- Results

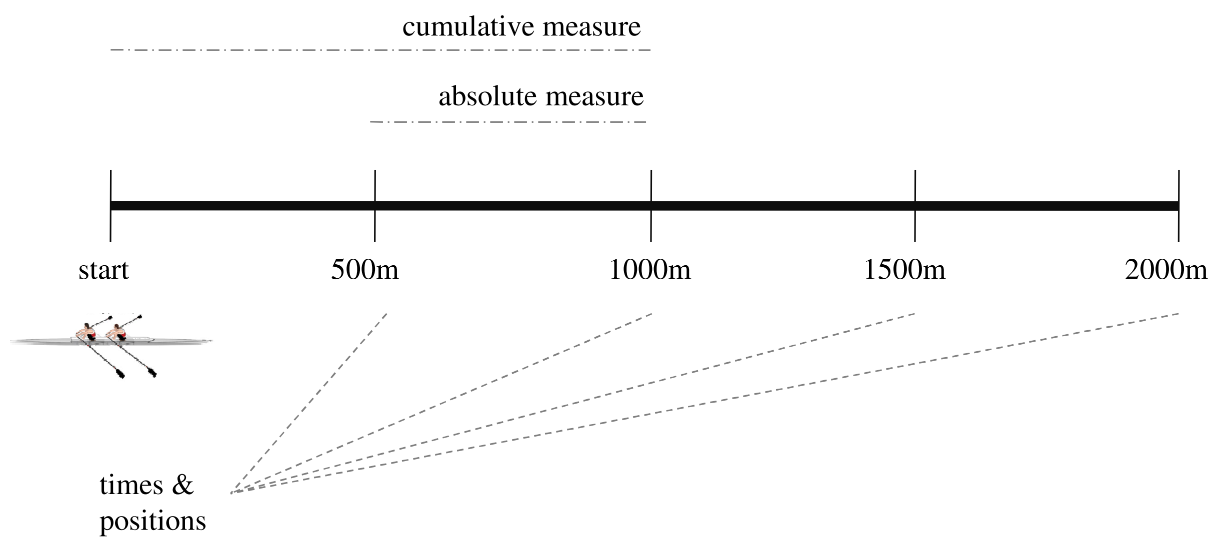
Table 1. Average accuracy measures (%) of the LOOCV procedure on the BNs constructed for both genders after each round of the six-event cycling omnium

Gender	Eval. criterion	r1	r2	r3	r4	r5	r6
male	$a == p$	1.7610	4.0750	4.0750	8.3370	12.2160	16.1740
	$a == p \pm 1$	7.6110	9.3710	11.0370	26.1340	32.1450	38.1670
	$a == p \pm 2$	14.1680	16.9580	19.0410	35.6160	42.9810	50.2090
	$a == p$ (cat.)	38.5070	40.9000	42.5660	56.1640	59.0290	66.8590
female	$a == p$	1.7450	5.6525	7.7487	9.0887	12.6137	11.0512
	$a == p \pm 1$	10.2612	18.7862	20.3362	33.2575	36.1287	35.5162
	$a == p \pm 2$	16.8525	26.5812	25.7525	46.7650	49.2175	49.3587
	$a == p$ (cat.)	43.5350	49.0075	49.6262	66.3125	62.5987	62.2012

Note. $a == p$ (cat.) represents $a == p$ for categories (1st-3rd,4th-10th,>10th)

BNs in Rowing

- A high profile Wolrd/Olympic competition
  - sprint: 500m
  - endurance: 160/185km
  - most common: 2000m

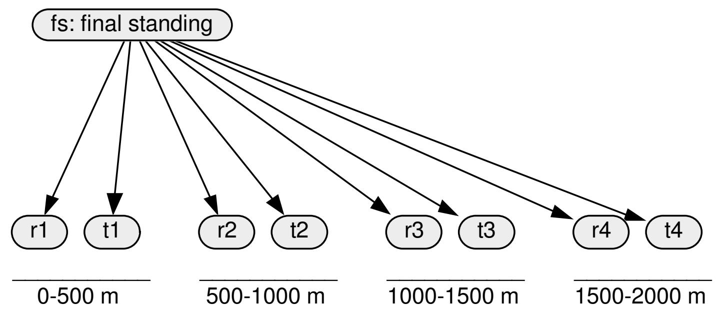


BNs in Rowing

- Research matters
  - strategizing:
    - considering measures of energy expenditure
    - understanding the level of performance required at the different sectors
  - we considered:
    - what is the chance of medal winning for rowers if they finish a sector in n^{th} position?
    - what performance during each 500m sector of rowing races may maximize the chance of finishing in certain positions?

BNs in Rowing

- Structure



- The joint probability function

$$p(e, \bar{h}) = \prod_i p(h_i|e)$$

$$\bar{h} = \{h_1, h_2, \ldots, h_8\} = \{t_1, t_2, t_3, t_4, r_1, r_2, r_3, r_4\}$$

BNs in Rowing

- Evaluation
  - data
    - competition results from 1996 to 2009
    - only top six teams in finals
    - each record: absolute and cumulative times at each 500m split
    - we extracted rankings at each 500m split
  - CPT learning: counting-learning
  - procedure: BNs + combinatorial optimization
  - interface: Netica + Netica API (C#.Net)

BNs in Rowing

- Results

Table 2. Maximal solutions found for certain rowing final rankings. The variables r_1 to r_4 show the rankings in the first to the fourth sectors and the variables t_1 to t_4 represent the finish times in the same sectors.

Race	r_1	r_2	r_3	r_4	t_1	t_2	t_3	t_4	Final rank	Prob. (%)
M4-	1	1	1	3	82-84	86-88	86-88	88-90	1	99.04
M4-	2	2	2	6	100-102	102-104	100-102	100-102	2	94.80
M4-	6	3	3	3	84-86	88-90	84-86	84-86	3	86.80
W2-	1	1	1	3	98-100	102-104	104-106	102-104	1	99.92
W2-	5	2	2	2	102-104	108-110	108-110	104-106	2	99.40
W2-	6	3	3	5	104-106	104-106	104-106	106-108	3	98.47

BNs in Rowing

- Results (cont.)

Table 3. Classification analysis of the rowing data with split measures

Data set	Classifier	Precision	Recall	F-measure
M4-	SVM	0.524	0.521	0.517
	C4.5	0.391	0.397	0.390
	Random Forest	0.429	0.432	0.428
	RBF	0.153	0.214	0.160
	NB	0.475	0.483	0.475
	KNN (K=10)	0.474	0.474	0.466
W2-	SVM	0.600	0.601	0.600
	C4.5	0.557	0.563	0.557
	Random Forest	0.554	0.550	0.551
	RBF	0.142	0.189	0.157
	NB	0.609	0.609	0.609
	KNN (K=10)	0.610	0.618	0.611

Limitations of BNs in Sports

- Unseen sports performances
  - record performances
  - outliers
    - poor performances, e.g., $t_1 > max(t_1)$ in rowing
    - unseen combinations, e.g., 24,24,24,24,24,24 or 1,1,1,1,1,1 in cycling omnium

Limitations of BNs in Sports

- Example scenario 1:
  - BN: constructed for track cycling omnium
  - input: 1,1,1,1,1,1 for the six nodes of the six individual events
  - observation: likelihood of final standing 1

Table 4. Likelihood analysis of unseen record performances in track cycling omnium using BNs

Sex	Learning algorithm	Likelihood of fs=1 (%)
female	CL	77.2
	EM	0.16
	GD	14.0
male	CL	69.1
	EM	84.9
	GD	89.7

Limitations of BNs in Sports

- Example scenario 2:
  - BN: constructed for rowing
  - input: $t_1 < min(t_1)$ for the first 500 m sector
  - observation: likelihood of final standing 1

Table 5. Likelihood analysis of unseen record performances in rowing using BNs

Race	$min(t_1)$	t_1 Entered	Likelihood of fs=1 (%)		Learning algorithm
			$min(t_1)$	t_1 Entered	
M4-	01:22.40	01:21.00	39.65	17.20	CL
			70.70	22.38	EM
			63.39	17.37	GD
W2-	01:38.70	01:37.00	25.66	17.26	CL
			39.36	17.15	EM
			38.42	16.12	GD

Concluding Remarks

- BNs (+ CO techniques) successfully applied to decision support in some sports, including:
  - track cycling omnium
  - rowing

- According to our observations, BNs fall short in dealing with (modeling) unseen performances

Thank you!

Bahadorreza OFOGHI
email: bahadorreza.ofoghi@vu.edu.au
phone: +61 3 9919 1820