A Markov Game Model for Valuing Player Actions in Ice Hockey

Kurt Routley, Oliver Schulte

School of Computing Science, Simon Fraser University, Vancouver-Burnaby Canada

Introduction

Our vision: sports analytics = branch of reinforcement learning.

- Fundamental question: which actions contribute to winning in what situation?
- Answer: learn an **action-value** function or Q-function.

Motivation

Advantages over previous action-based analytics (plus-minus, Corsi, Fenwick).

- **Context-Awareness.** Action values depends on context = state.
  - Example: Goals are worth more with tied scores than with a 2 goal lead.
- **Lookahead.** Actions can have medium-term impact.
  - Example: Penalties can lead to goals but not immediately.

Related Work

- Expected Possession Value EPV: a Q-function for basketball [2]. Spatial-temporal model based on tracking data.
- Total Hockey Rating (THoR) [3] assigns a value to all ice hockey player actions. No context, fixed look-ahead window (20 sec).

Data Set

- 2.8M events, > 600K play sequences.

GameId	Period	Sequence Number	Event Number	Event
1	1	1	1	PERIOD START
1	1	1	2	faceoff(Home,Neutral)
1	1	1	3	hit(Away,Neutral)
1	1	1	4	takeaway(Home,Defensive)
1	1	1	5	missed_shot(Away,Offensive)
1	1	1	6	shot(Away,Offensive)
1	1	1	7	giveaway(Away,Defensive)
1	1	1	8	takeaway(Home,Offensive)
1	1	1	9	missed_shot(Away,Offensive)
1	1	1	10	goal(Home,Offensive)

References

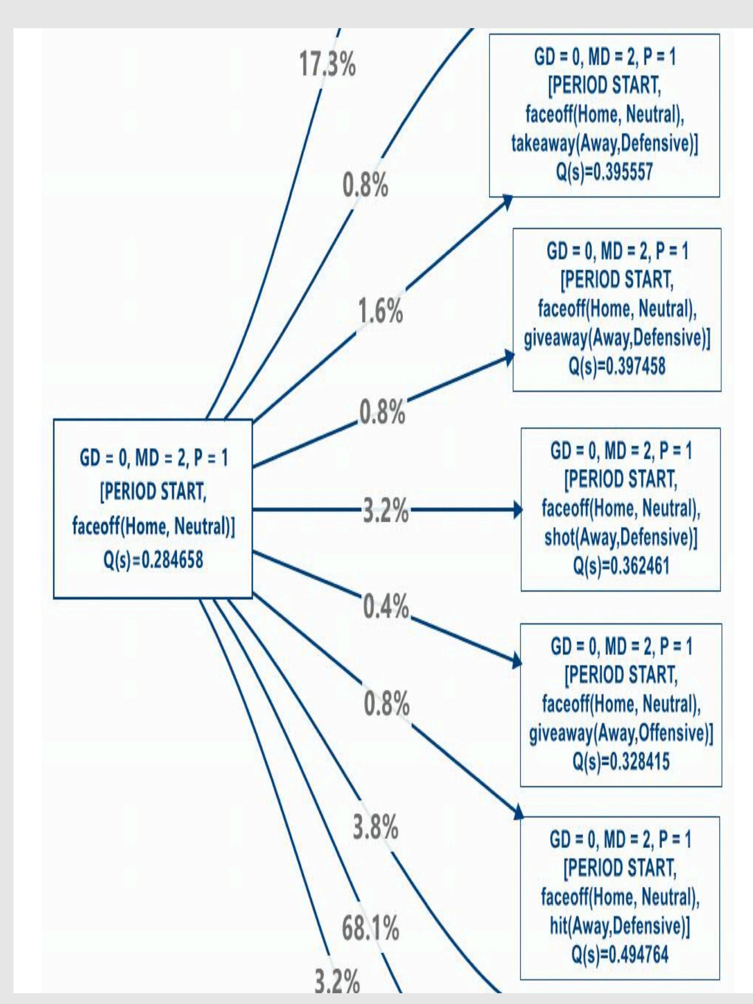
1. M. L. Littman. Markov games as a framework for multi-agent reinforcement learning. In ICML, pp. 157-163, 1994.
2. Cervone, D.; D'Amour, A.; Bornn, L. & Goldsberry, K. POINTWISE: Predicting points and valuing decisions in real time with nba optical tracking data. In MIT Sloan, 2014.
3. M. Schuckers and J.Curro. Total hockey rating (THoR): A comprehensive statistical rating of national hockey league forwards and defensement based upon all on-ice events. In MIT Sloan, 2013.

Markov Game Model

A Markov Game Model [1] consists of 4 components:
State Space, Transition Graph + Probabilities, Rewards

- **Players =** Home, Away.
- **State** = (Goal Differential, ManPower Differential, Period, Action History within play sequence)
- Transition probabablities estimated from the number of observed occurrences.
- >1.3 M states with >0 occurrences.

State Transition Examples.



Rewards/Costs

- Score Goal/Incur Penalty.

Value Iteration for Q-Learning

Since states encode action histories, the expected value of states is equivalent to learning a Q-function (V = Q).

$$Q_{i+1}(s) = R(s) + \frac{1}{Occ(s)} \sum_{(s,s')\in E} \left(Occ(s,s') \times Q_i(s') \right)$$

Applications of the Q-function

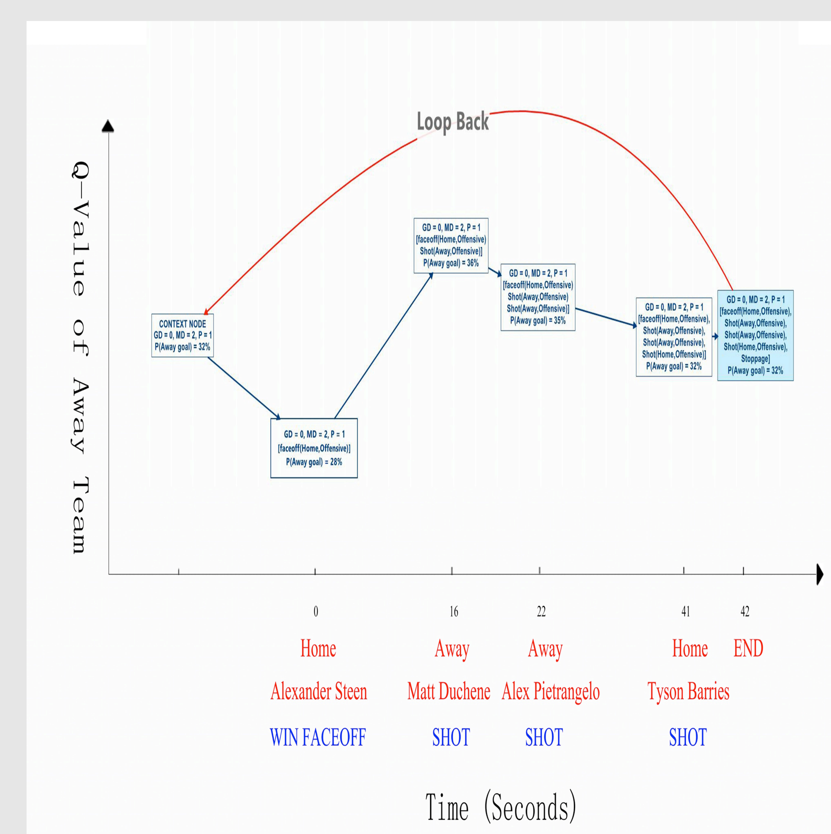
- *Knowledge Discovery*. Cervone et al. [2]: "We assert that most questions that coaches, players, and fans have ...can be phrased and answered in terms of EPV [i.e., the Q-function]."
- *Player ranking*. Add up the total **impact** of a player's actions.

Player Impact Scores

The impact of an action in a state is defined by

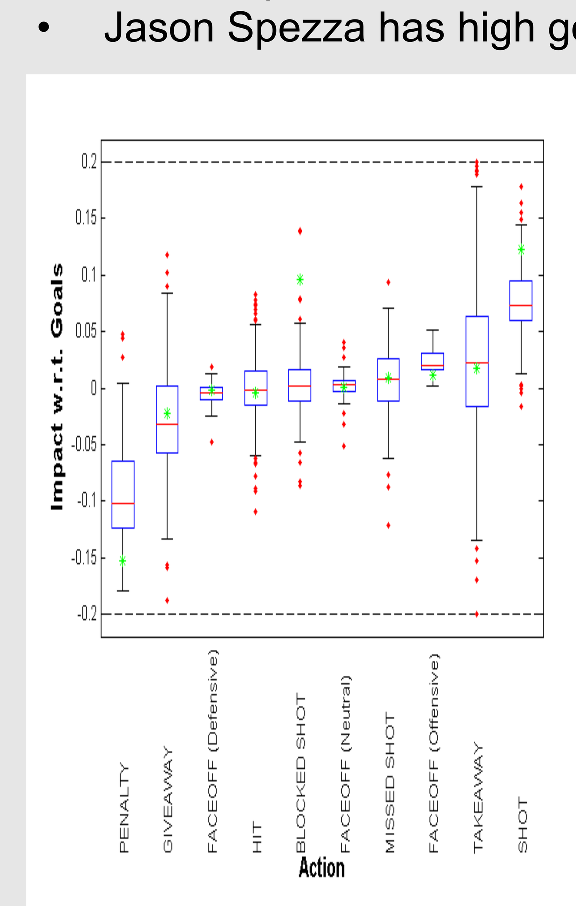
$$impact(s,a) = Q(s * a) - Q(s)$$

The Q-Value Ticker for Colorado vs. St. Louis



Action Goal Impact Depends on Context

- Boxplot of action value for each state.
- * = THoR Action Values [3].
- Player Total Goal Impact (2014-2015 Season 1st half)
  - Jason Spezza has high goal impact, low plus-minus.



Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Jori Lehtera	C	17.29	8	25	13	21	$3,250,000
Henrik Zetterberg	LW	14.54	7	30	-1	21	$7,500,000
Jason Spezza	C	14.33	6	25	-11	25	$4,000,000
Vladimir Tarasenko	RW	12.78	20	37	18	20	$900,000
Jonathan Toews	C	12.60	13	29	9	19	$6,500,000
Joe Pavelski	C	12.22	16	29	5	22	$6,000,000
Kyle Okposo	RW	11.79	8	29	-4	18	$3,500,000
Brent Burns	D	11.56	10	27	-3	16	$5,760,000
Gustav Nyquist	RW	11.47	14	22	-7	15	$1,050,000
Joe Thornton	C	11.44	8	30	2	28	$6,750,000
Ryan Kesler	C	10.99	12	27	-1	20	$5,000,000
Tomas Plekanec	C	10.50	10	23	6	15	$5,000,000
Sidney Crosby	C	10.43	10	37	12	18	$12,000,000
Patrick Marleau	LW	9.96	7	27	-2	19	$7,000,000
Martin Hanzal	C	9.76	6	17	1	16	$3,250,000
Jaden Schwartz	LW	9.57	11	27	10	21	$2,000,000
Pavel Datsyuk	C	9.51	13	25	4	16	$10,000,000
Steven Stamkos	C	9.44	16	33	-2	14	$8,000,000
Alex Ovechkin	RW	9.43	16	28	5	18	$10,000,000
Rick Nash	LW	9.35	23	36	16	32	$7,900,000
Sean Monahan	C	8.92	11	22	6	23	$925,000
Phil Kessel	RW	8.70	17	38	-4	14	$10,000,000
Jaromir Jagr	RW	8.68	5	20	-12	25	$3,500,000
Frans Nielsen	C	8.64	6	17	-1	23	$3,000,000
Nikita Kucherov	RW	8.60	14	31	20	13	$743,000

Conclusion

- The Q-function is a powerful AI concept that captures much information about hockey dynamics (or other sports).
- Novel player ranking method based on reinforcement learning.
- The Q-impact of an action varies greatly with context, and medium-term ripple effects make a difference.
- Goal Impact scores correlate with points.