What is the Value of an Action in Ice Hockey? Learning a Q-function for the NHL

Oliver Schulte, Zeyu Zhao, and Kurt Routley

Simon Fraser University,
Burnaby, BC, Canada
{oschulte}@cs.sfu.ca

Abstract. Recent work has applied the Markov Game formalism from AI to model game dynamics for ice hockey, using a large state space. Dynamic programming is used to learn action-value functions that quantify the impact of actions on goal scoring. Learning is based on a massive dataset that contains over 2.8M events in the National Hockey League. As an application of the Markov model, we use the learned action values to measure the impact of player actions on goal scoring. Players are ranked according to the aggregate goal impact of their actions. We show that this ranking is consistent across across seasons, and compare it with previous player metrics, such as plus-minus and total points.

1 Introduction

This paper describes and extends a recent approach to sports analytics that applies advanced concepts from Artificial Intelligence to model game dynamics. This approach models game dynamics using Markov games [4], a multi-agent extension of Markov Decision Processes. A Markov game model can answer a fundamental question about a sport: Which actions contribute to winning in what situation?

We approach this question for ice hockey by learning an *action-value function*, or Q-function, for a Markov game model of the National Hockey League (NHL). In reinforcement learning (RL) notation, the expression $Q(s, a)$ denotes the expected reward of taking action a in state s. We learn a Q-function from a massive dataset about matches in the National Hockey League (NHL). This dataset comprises all play-by-play events from 2007 to 2014, for a total of over 2.8M events/actions and almost 600K play sequences. The Markov Game model comprises over 1.3M states. Whereas most previous works on Markov Game models aim to compute optimal strategies or policies [4], we learn a model of how hockey is actually played, and do not aim to compute optimal strategies. In RL terminology, we learn a Q-function for the *on-policy* setting [13]. Motivation for learning a Q-function for NHL hockey dynamics includes knowledge discovery and player evaluation, which is the application we focus on in this paper.

Knowledge Discovery. The Markov Game model provides information about the likely consequences of actions. The basic model and algorithms can easily be adapted to study different outcomes of interest, such as goals and penalties. For example, with goals as rewards, a Q-function specifies the impact of an action on future goals. With penalties as costs in the same model, the resulting Q-function specifies the impact of an action on future penalties.

Player Evaluation. One of the main tasks for sports statistics is evaluating the performance of players [11]. A common approach is to assign action values, and sum the corresponding values each time a player takes the respective action. A simple and widely used example in ice hockey is the +/- score: for each goal scored by (against) a player's team when he is on the ice, add +1 (-1) point. Researchers have developed several extensions of +/- for hockey [5, 12, 10]. The NHL has started publishing advanced player statistics such as the Corsi (Shot Attempts) and Fenwick (Unblocked Shot Attempts) ratings.[1] Using action values defined by the Q-function has two major advantages compared to the previous action count approaches used in ice hockey. (1) The Markov game model is aware of the *context* of actions within a game. In the Markov Game model, *context* = *state*. (2) An action may have medium-term and/or ripple effects rather than immediate consequences. For example, penalties have a short-term but not immediate effect on goal scoring. Therefore evaluating the impact of an action requires *look-ahead*, which the Q-function provides.

To evaluate player performance, we use the Q-function to quantify the value of a player's action in a context. The action values are then aggregated over games and seasons to get player impact scores. Player impact scores correlate with plausible alternative scores, such as a player's total points, but improve on these measures, as our impact score is based on many more events. A new finding is that player impact on goals correlates well across seasons ($r = 0.7$).

Contributions. We make our extensive dataset available on-line, in addition to our code and the learned Markov game model [9]. The main contributions of this paper may be summarized as follows:

1. We describe a set of recent developments that apply Markov game modelling to sports analytics. Learning a Q-function provides context-aware action values to score hockey player actions based on look-ahead from the current game context.
2. We show that the resulting player rankings are consistent across seasons.

2 Related Work

This paper is an extension of our previous work [8]. The closest predecessor to our work in ice hockey is the Total Hockey Rating (THoR) [10]. This assigns a value to all actions, not only goals. Actions were evaluated based on whether or not a goal occurred in the following 20 seconds after an action. This work used data from the 2006/2007 NHL season only. THoR assumes a fixed value for every action and does not account for the context in which an action takes

[1] `nhl.com`

place. Furthermore, the window of 20 seconds restricts the look-ahead value of each action.

The Win-Probability-Added player metric [6, 3] scores the importance of a goal relative to the game context. This work uses a Markov model as well, which is a submodel of ours in the following sense: (1) The only actions considered are goals. (2) The context includes the current goal and manpower differentials, but not the period and not the recent play history. Their work does, however, include the current game time, unlike our model. The game time is especially important for propagating the impact of an action to a final win, since for instance the probability of a win is 1 if a goal puts a team ahead in the last minute.

In a finalist paper at the MIT Sloan conference, [2] Cervone *et al.* used spatial-temporal tracking data for basketball to build the POINTWISE model for valuing player decisions and player actions. Conceptually, their approach to defining action values is the closest predecessor to ours: The counterpart to the value of a state in a Markov game is called expected possession value (EPV). The counterpart to the impact of an action on this value is called EPV-added (EPVA). Cervone *et al.* emphasize the potential of the context-based impact definitions for knowledge discovery: "we assert that most questions that coaches, players, and fans have about basketball, particularly those that involve the offense, can be phrased and answered in terms of EPV [i.e., the Q-function]." While the definition of action impact is conceptually very similar, [2] use neither AI terminology nor AI techniques. Moreover, the NHL does not yet have and therefore we do not use spatial tracking data, which is the main focus of [2].

Cervone *et al.* [1] note that the Q-function approach for valuing actions can in principle be applied to different types of sports. Substantially smaller Markov Decision Process Models than ours have been used to model dynamics in various sports; for a review please see [2] and [8]. To our knowledge these models have not been applied to valuing actions.

3 Hockey Data Used

We assume familiarity with the basic rules of NHL play; [8] provides a brief summary. The NHL provides information about sequences of play-by-play events, which are scraped from `http://www.nhl.com` and stored in a relational database. The real-world dataset is formed from $2,827,467$ play-by-play events recorded by the NHL for the complete 2007-2014 seasons, regular season and playoff games, and the first 512 games of the 2014-2015 regular season. A breakdown of this dataset is shown in Table 1 (left). The type of events recorded by the NHL from the 2007-2008 regular season and onwards are listed in Table 1. There are two types of events: actions performed by players and start and end markers for each play sequence. Every event is marked with a continuous time stamp, and every action is also marked with a zone Z and which team, Home or Away, carries out the action.

Table 1: Left: Size of Dataset. Right: NHL Play-By-Play Events Recorded

Number of Teams	32
Number of Players	1,951
Number of Games	9,220
Number of Sequences	590,924
Number of Events	2,827,467

Action Event	**Start/End Event**
Face-off	Period Start
Shot	Period End
Missed Shot	Early Intermission Start
Blocked Shot	Penalty
Takeaway	Stoppage
Giveaway	Shootout Completed
Hit	Game End
Goal	Game Off
	Early Intermission End

4 Markov Games

A Markov Game [4], is defined by a set of states, S, and a collection of action sets, one for each agent in the environment. State transitions are controlled by the current state and one action from each agent. A state transition is associated with a reward for each agent. In our hockey Markov game model, there are two players, the Home Team H and the Away Team A. In each state, only one team performs an action, although not in a turn-based sequence. This reflects the way the NHL records actions.

A **state** comprises context features and play sequences that represent the recent trajectory of the game. A sequence in the NHL play-by-play data corresponds to an episode in Markov decision process terminology.

Context Features remain constant throughout a play sequence (episode). A **context state** lists the values of relevant features at a point in the game. These features are shown in Table 2(left), together with the range of integer values observed. Goal Differential GD is calculated as Number of Home Goals - Number of Away Goals. A positive (negative) goal differential means the home team is leading (trailing). Manpower Differential MD is calculated similarly. Period P represents the current period number the play sequence occurs in, typically ranging in value from 1 to 5. Periods 1 to 3 are the regular play of an ice hockey game, and periods 4/5 indicate overtime/shootout periods.

Play Sequences are sequences of actions. The basic action events are shown in Table 2. Each of these actions has two parameters: which team T performs the action and the zone Z where the action takes place. Zone Z represents the area of the ice rink in which an action takes place, relative to the team performing an action (Offensive, Neutral, Defensive). Table 2 shows an example of a NHL play-by-play action sequence in tabular form. A **state** is a pair $s = \langle \mathbf{x}, h \rangle$ where $\mathbf{x}$ is a list of context features and h a play/action sequence.

Table 2: The Markov Game State Space. Left: Context Features. Right: Play Sequences in Tabular Format. A **play sequence** h is a sequence of events starting with exactly one start marker, followed by a list of action events, and ended by at most one end marker. Start and end markers are shown in Table 1(Right), adding shots and face-offs as start markers, and goals as end markers.

Name	Range
Goal Diff.	[-8,8]
Manpower Diff.	[-3,3]
Period	[1,7]

GameId	**Period**	**Sequence #**	**Event #**	**Event**
1	1	1	1	PERIOD START
1	1	1	2	(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)
1	1	2	11	face-off(Away,Neutral)
...				

State Transitions. If h is an incomplete play sequence, we write $h \star a$ for the play sequence that results from appending a to h, where a is an action event or an end marker. Similarly if $s = \langle \mathbf{x}, h \rangle$, then $s \star a \equiv \langle \mathbf{x}, h \star a \rangle$ denotes the unique successor state that results from executing action a in s. State transition examples are shown in Figure 1. Since the complete action history is encoded in the state, action-state pairs are equivalent to state pairs.

Rewards and the Q-Function. Important reward/cost events for hockey include goals, penalties [8], and final wins [7]. This paper examines goal scoring, represented by the following reward function. Any state ending with a home (away) goal is an absorbing state, where the home (away) team receives a reward of 1. For other states the reward is 0. With this Next Goal reward function, the expected reward $Q_H(s)$ represents the probability that if play starts in state s, a random walk through the state space of unbounded length ends with a goal for the Home team resp. the Away team. Given the transition probabilities in the Markov game model, the Q-function values can be computed by dynamic programming [8].

5 Action Values and Player Impact Rankings.

The **impact** of an action is a function of context (= Markov state), defined as follows:

$$impact(s, a) \equiv Q_T(s \star a) - Q_T(s) \tag{1}$$

where T is the team executing the action a. The impact quantity measures how performing an action in a state affects the expected reward difference. Figure 1 shows a "Q-value ticker" representation of how the Q-values for the Next

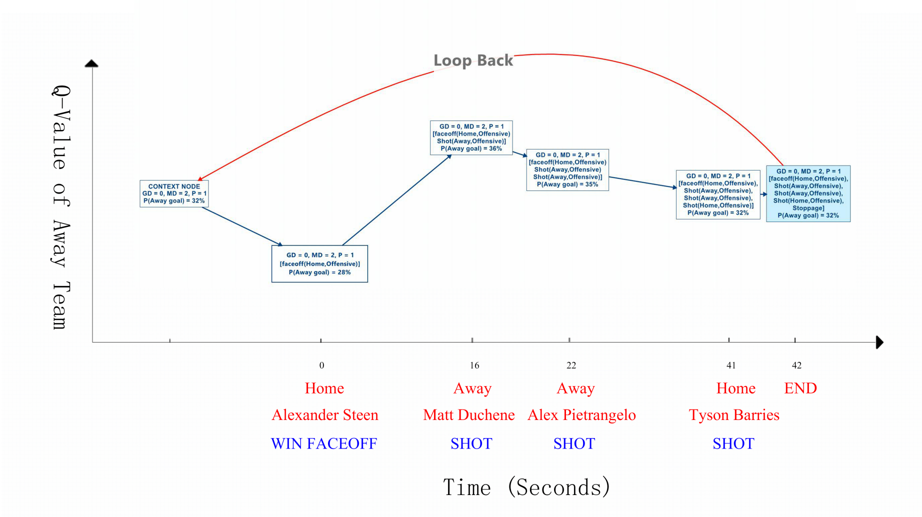

Fig. 1: The Q-Value Ticker tracks the change in state value during a game. In this example, Colorado (Home) plays St. Louis (Away), 1st period, 8th play sequence. The Q-value represents the probability that St. Louis scores the next goal.

Goal reward change with state transitions in a specific game sequence [2]. The impact measures the magnitude of the Q-value change when an action takes the match from one state to another. This change measures the impact of an action on the *chance* of scoring the next goal. The Next Goal Q-values are very different from simply counting goals, in at least two respects. (1) The Next Goal Q-values reflect both offensive and defensive contributions. For example, if a player wins a face-off in his defensive zone, this decreases the chance that the opposing team will score the next goal, and therefore increases the chance that his team will score the next goal. (2) The look-ahead of the Q-value computation means that actions that lead to goals, but are not goals themselves, receive high impact counts. Routley and Schulte report that averaged over states, shots have the highest goal impact compared to other actions [8], as one would expect. They show that *depending on the context and event history, the value of an action can vary greatly.* All actions, including penalties, but excluding goals and face-offs won in the offensive zone, have at least one conext (state) where the action has a positive impact, and another context with a negative impact.

5.1 Single Season Player Valuations.

To calculate player valuations, we apply the impact of an action to the player as they perform the action. Next, we sum the impact scores of a player's actions over a single game, and then over a single season, to compute a net season impact score for the player. This procedure compares the actions taken by a specific player to those of the league-average player, similar to previous work

Table 3: 2014-2015 Top-25 Player Impact Scores For Goals

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

[6, 2]. Since a shot has a high impact on goal chances, it contributes strongly to the goal impact value, regardless of whether the shot leads to a goal or not. In this regard our goal impact measure agrees with the intuition behind other hockey statistics such as the Corsi (Shot Attempts) and Fenwick (Unblocked Shot Attempts) ratings that reward shot attempts, not goals. The Q-function provides a principled method for assigning weights to different types of shots attempts. It also takes into account actions other than shots, such as winning face-offs.

Table 3 compares impact on Next Goal Scored with three other player ranking metrics: points earned, salary, and +/-. Player impact scores are shown in Table 3, for the first 512 games of the 2014-2015 season. ables for all seasons are available as well [7]. Figure 2(left) shows that next goal impact correlates well with points earned. A point is earned for each goal or assist by a player. The players with a high impact on goals, also tend to have a positive +/- rating.

5.2 Case Studies

We discuss our findings for some individual players of interest. Table 3 appears to pass the “eye test” in that it lists top offensive players of the NHL, including goal-getters such as Sidney Crosby, Steven Samkos, and Alex Ovechkin. The fact that these high scores are not ranked at the top illustrates the difference between goal impact and goals (cf. [6]). All three members of St. Louis’ famed “STL” line

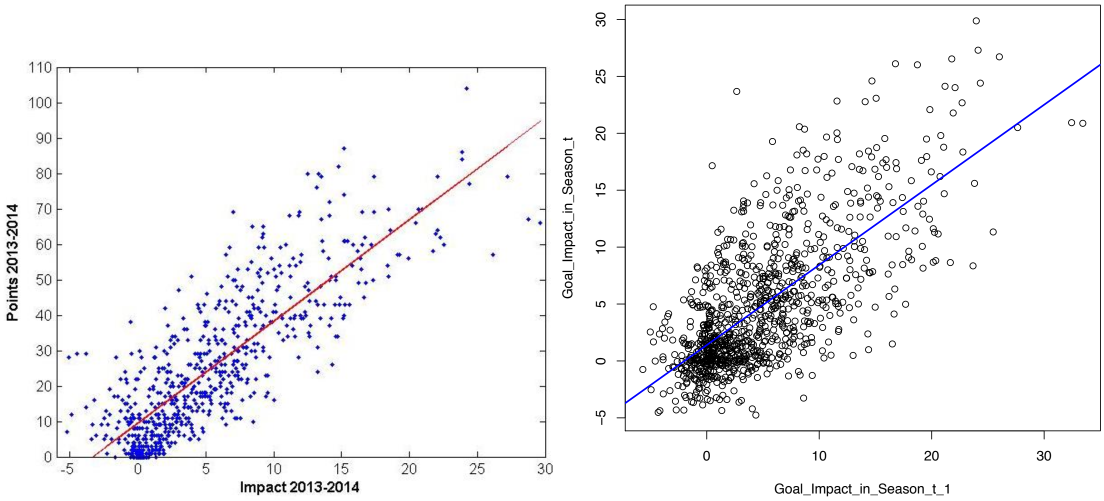

Fig. 2: Left: 2013-2014 Correlation between Player Goal Impact and Points Earned. Right: Correlation between Goal Impact in one season and the next, for Seasons 2007-2014.

(Schwartz, Tarasenko, Lehtera)[2] are among the top 20 in our list. In fact, Jori Lehtera tops our goal impact list, although his linemate Tarasenko outscored him by far. Our analysis suggests that Lehtera's actions create the opportunities that Tarasenko exploits. This pattern fits the traditional division of labor between a center and a wing player. Tarasenko is also the most undervalued player in our list. Starting with the 2015-2016 season, St. Louis has signed him for an annual average of 7.5M contract, which our analysis strongly supports.

Jason Spezza is an anomaly, as he has the highest impact score but a negative +/- score. This may be due to a lack of defensive contributions (defensive actions are underrepresented in the NHL data). Another explanation is that while Spezza generally performs useful actions, he happens to play on a relatively poor line. A strong example of this scenario was the 2013-2014 season, where Spezza topped our goal impact list, but had a very low +/- score of -26 [8]. This reason for this score is that his Ottawa team performed poorly overall in the 2013-2014 season, with a goal differential of -29. Spezza requested a trade, which our analysis would recommend. At the Dallas Stars, his season total +/- score was more in line with his goal impact as we would predict. (-7 compared to Dallas' +1 overall).

6 Goal Impact Is Consistent Across Seasons.

A desirable feature of a player ranking score is temporal consistency [6], for at least two reasons. First, generally the skill of a player does not change greatly

[2] http://www.nhl.com/ice/news.htm?id=738943

from one season to the next. Therefore a good quality metric should show consistency between seasons. Second, a consistent ranking is useful because it supports predicting future performance in the next season from past performance in previous seasons. To assess consistency across seasons, we follow Pettigrew's methodology [6]. (1) For each pair of successive seasons, for each player who plays in both seasons, list the player score in each season. (2) Compute the correlation between players' impact score in season t and the score in season $t + 1$. Table 4 shows the season-to-season correlation for the goal impact scores and related measures. *Goal impact as defined by the Markov game model is well correlated across seasons*, $r = 0.71$. The trend line in the scatter plot of Figure 2 (right) shows the correlation graphically. In contrast, the traditional +/- score varies across seasons, $r = 0.345$. Table 4 also shows cross-season correlations for a number of adjusted goal impact (GI) metrics: goal season impact per total games played in season (similar to [6]), impact per minutes played, and impact per actions taken. All of the adjusted metrics are substantially less consistent than the summed goal impact metric.

An interesting computation for future work would be to focus the correlations on players who change teams. We would predict that Goal Impact remains consistent across seasons for such players, as it reflects their individual achievement, whereas goal-based scores like +/- and points depend heavily on a player's teammates and should therefore be less consistent when players change teams.[3]

Table 4: Season-to-Season Correlations for Different Player Performance Metrics.

Goal Impact	PlusMinus	GI/Games	GI/Actions	GI/TimePlayed
0.703	0.345	0.508	0.141	0.325

7 Conclusion

We have described a Markov Game Model for a massive set of NHL play-by-play events with a rich state space. Compared to previous work that assigns a single value to actions, the model's action-value Q-function incorporates two powerful sources of information for valuing hockey actions: (1) It takes into account the context of the action, represented by the Markov Game state. (2) It models the medium-term impact of an action by propagating its effect to future states. We applied our model to evaluate the performance of players in terms of their actions' total impact. We showed that players' goal impact values are consistent and hence predictable across seasons. In sum, the Q-function is a powerful AI concept that captures much information about hockey dynamics as the game is played in the NHL. While player ranking is one important application for hockey analytics, we expect that the Q-function concept will facilitate many others.

[3] We are indebted to an anonymous workshop reviewer for this suggestion.

References

1. D. Cervone, A. Amour, L. Bornn, and K. Goldsberry. A Multiresolution Stochastic Process Model for Predicting Basketball Possession Outcomes. *ArXiv e-prints*, August 2014.
2. D. Cervone, A. DAmour, L. Bornn, and K. Goldsberry. Pointwise: Predicting points and valuing decisions in real time with NBA optical tracking data. In *8th Annual MIT Sloan Sports Analytics Conference, February*, volume 28, 2014.
3. E. H. Kaplan, K. Mongeon, and J. T. Ryan. A markov model for hockey: Manpower differential and win probability added. *INFOR: Information Systems and Operational Research*, 52(2):39–50, 2014.
4. M. L. Littman. Markov games as a framework for multi-agent reinforcement learning. In *Proceedings of the eleventh international conference on machine learning*, volume 157, pages 157–163, 1994.
5. B. Macdonald. An improved adjusted plus-minus statistic for nhl players, March 2011.
6. S. Pettigrew. Assessing the offensive productivity of nhl players using in-game win probabilities. In *9th Annual MIT Sloan Sports Analytics Conference*, 2015.
7. K. Routley. A markov game model for valuing player actions in ice hockey. Master's thesis, Simon Fraser University, April 2015.
8. K. Routley and O. Schulte. A markov game model for valuing player actions in ice hockey. In *Uncertainty in Artificial Intelligence (UAI). To Appear.*, 2015.
9. K. Routley, O. Schulte, and Z. Zhao. Q-learning for the nhl. `http://www.cs.sfu.ca/~oschulte/sports/`, June 2015.
10. M. Schuckers and J. Curro. Total hockey rating (thor): A comprehensive statistical rating of national hockey league forwards and defensemen based upon all on-ice events. In *7th Annual MIT Sloan Sports Analytics Conference*, 2013.
11. R. P. Schumaker, O. K. Solieman, and H. Chen. Research in sports statistics. In *Sports Data Mining*, volume 26 of *Integrated Series in Information Systems*, pages 29–44. Springer US, 2010.
12. N. Spagnola. The complete plus-minus: A case study of the columbus blue jackets. Master's thesis, University of South Carolina, 2013.
13. R. S. Sutton and A. G. Barto. *Reinforcement learning : an introduction.* MIT Press, Cambridge, Mass., 1998.