A Markov Game Model for Valuing Player Actions in Ice Hockey

by

Kurt Routley

B.Sc., Simon Fraser University, 2013

Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

Master of Science

in the
School of Computing Science
Faculty of Applied Sciences

APPROVAL

Name: Kurt Routley

Degree: Master of Science

Title: A Markov Game Model for Valuing Player Actions in Ice Hockey

Examining Committee: **Chair**: Dr. James Delgrande
Full Professor

Dr. Oliver Schulte
Senior Supervisor
Computing Science,
Simon Fraser University
Associate Professor

Dr. Tim Swartz
Supervisor
Statistics,
Simon Fraser University
Full Professor

Dr. Anoop Sarkar
Internal Examiner
Computing Science,
Simon Fraser University
Associate Professor

Date Approved: April 17th, 2015

Partial Copyright Licence

SFU

Abstract

Evaluating player actions is very important for general managers and coaches in the National Hockey League. Researchers have developed a variety of advanced statistics to assist general managers and coaches in evaluating player actions. These advanced statistics fail to account for the context in which an action occurs or to look ahead to the long-term effects of an action. I apply the Markov Game formalism to play-by-play events recorded in the National Hockey League to develop a novel approach to valuing player actions. The Markov Game formalism incorporates context and lookahead across play-by-play sequences. A dynamic programming algorithm for value iteration learns the values of Q-functions in different states of the Markov Game model. These Q-values quantify the impact of actions on goal scoring, receiving penalties, and winning games. Learning is based on a massive dataset that contains over 2.8 million events in the National Hockey League. The impact of player actions varies widely depending on the context, with possible positive and negative effects for the same action. My results show using context features and lookahead makes a substantial difference to the action impact scores. Accounting for context and lookahead also increases the information in the model. Players are ranked according to the aggregate impact of their actions, and compared with previous player metrics, such as plus-minus, total points, and salary, as well as recent advanced statistics metrics.

Keywords: Markov Game model; ice hockey; value iteration; player ranking;

Acknowledgements

I give thanks to God for giving me the strength to get through my degree, and to my Dad who gave me a love of hockey from a young age and is now with our Lord and Saviour.

I would like to thank my Mom, Kara, and Luke, for supporting me throughout my degree.

I would like to thank Dr. Oliver Schulte for allowing me to work on ice hockey statistics and for his close guidance in my research. I would also like to thank Dr. Tim Swartz for discussing various hockey and sports statistics with us on a regular basis. Thanks to Dr. Anoop Sarkar for agreeing to be my examiner, and to Dr. Jim Delgrande for chairing my examination.

Thanks to all my bros for understanding when I had to work and couldn't attend our treasured McDonald's runs.

Thanks also to my friends in the Computational Logic Lab, I have enjoyed our interesting talks during our lab coffee social meetings and working with you regularly.

Finally, a big thanks to Lydia Fang, Sunjeet Singh, and Angus Telfer for recommending me for graduate studies and pushing me to strive for excellence.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

A fundamental goal of sports statistics is to understand which actions contribute to winning in what situation. As sports have entered the world of big data, there is increasing opportunity for large-scale machine learning to model complex sports dynamics. Our research applies AI techniques to model the dynamics of ice hockey; specifically the Markov Game model formalism [12], and related computational techniques such as the dynamic programming value iteration algorithm. The Markov Game model makes use of a massive dataset of matches from the National Hockey League (NHL). This dataset comprises all play-by-play events from 2007 to 2014, as well as part of the 2014-2015 regular season, for a total of over 2.8 million events and actions and almost 600,000 play sequences. The Markov Game model comprises over 1.3 million states. Whereas most previous work on Markov Game models aim to compute optimal strategies or policies [12] (i.e., minimax or equilibrium strategies), this application learns a model of how hockey is actually played, and does not aim to compute optimal strategies. In reinforcement learning (RL) terminology, dynamic programming is used to compute a Q-function in the *"on policy"* setting [32]. In RL notation, the expression $Q(s, a)$ denotes the expected reward of taking action a in state s. There are many benefits to using the Markov Game model and value iteration for player evaluation. Our player evaluations for goals correlated with points, suggesting our Markov Game model captures assist information not present in the play-by-play event data. From a coaching perspective, the Q-values learned can be used for rapid post-game analysis of player performance. From a business perspective, player evaluations for wins can be used to identify bargain players and improve monetary valuations of players.

1.1 Motivation

Motivation for learning a Q-function for the dynamics of the NHL includes the following.

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, penalties, and winning. 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 [28]. If advanced statistics were useful for accurately measuring and predicting player performance, general managers could effectively use advanced statistics to buy wins and increase the entertainment value of sports. Unfortunately, the predictive accuracy of current and advanced statistics are often low, and do not form very informative features for predicting match outcomes [27, 34]. A common approach in advanced statistics is to assign values to each action, and sum the corresponding values each time a player takes the respective action. An advantage of this additive approach is that it provides highly interpretable player rankings. 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 [14, 31, 25]. The NHL has started publishing advanced player statistics such as the Corsi (Shot Attempts) and Fenwick (Unblocked Shot Attempts) ratings[1].

There are three major problems with the current action valuation approaches. (1) They are unaware of the *context* of actions within a game. For example, a goal is more valuable in a tied-game situation close to the end of the match than earlier in the match, or when the scorer's team is already four goals ahead [20]. Another example is that if a team manages two successive shots on goal, the second attempt typically has a higher chance of success. In the Markov Game model, *context = state.* Formally, the Q-function depends *both* on the state s and the action a. Richer state spaces therefore capture more of the context

[1] http://www.nhl.com/stats/advancedstats

of an action. (2) Previous action scores are based on immediate positive consequences of an action (e.g. goals following a shot). However, an action may have medium-term and/or ripple effects rather than immediate consequences in terms of visible rewards like goals. Therefore evaluating the impact of an action requires *lookahead*. Long-term lookahead is especially important in ice hockey because evident rewards like goals occur infrequently [13]. For example, if a player receives a penalty, this leads to a manpower disadvantage for his team, known as a powerplay for the other team. It is easier to score a goal during a powerplay, but this does not mean that a goal will be scored immediately after the penalty that causes the powerplay. For another example, if a team loses the puck in their offensive zone, the resulting counterattack by the other team may lead to a goal eventually but not immediately. The dynamic programming value iteration algorithm of Markov Decision Processes provides a computationally efficient way to perform unbounded lookahead, without assuming a bound on how many other events occur between the action and the reward. (3) Many advanced statistics, such as Corsi, Fenwick, and Added Goal Value (AGV) [20] only account for goals, shots, blocked shots, and missed shots. Other actions performed by players, such as hits, faceoffs, takeaways, giveaways, and penalties, are largely ignored by these advanced statistics. As such, the contribution or detriment to a team's performance as a result of these player actions is neglected in current advanced statistics. Since ice hockey is by nature a low-scoring game [13], a significant portion of ice hockey events are not considered when only examining goals, shots, blocked shots, and missed shots. The Markov Game model I describe uses all recorded NHL play-by-play events and is applied with the dynamic programming value iteration algorithm to learn the values of all player actions.

1.2 Implementation

The main computational challenge is to build a data structure for managing the large state space. The state space is large because each subsequence and complete sequence of actions defines a new state, along with the context features of the play sequence. Since I model the actual hockey dynamics in the "on policy" setting for the Markov Game model, only action sequences and subsequences that are actually observed in some NHL match need to be considered, rather than the much larger space of all possible action sequences.

As such, the general approach I take is to map all observed NHL play-by-play events into a tree of events for each game context, where each play sequence forms a branch of the tree under each game context. The classic AD-tree structure [17] is used to compute and store sufficient statistics over observed action sequences. Additional edges model further state transitions; for example, a new action sequence is started after a goal. Thus the state transition graph essentially superimposes additional edges on an AD-tree that represents action histories. The AD-tree compactly manages sufficient statistics, in this case state transition probabilities. This data structure also supports value iteration updates very efficiently, and the reward values of each state can be easily modified to model different objectives.

1.3 Evaluation

Model evaluation is performed through two lesion studies, where I remove different features from the Markov Game model to examine the benefit of retaining or removing the features. The first lesion study examines the benefits of including more features as context. Examples of context dependence give a qualitative sense of how the Markov Game model accounts for context. I compare the uncertainty (entropy) of models with little to no context with the entropy of the full Markov Game model including all context features. We measure uncertainty with respect to the probability of scoring the next goal. The second lesion study examines the benefits of propagating effects across sequences by adding and removing edges in different scenarios, forming multiple state transition graphs. To evaluate the impact of propagating action effects, I provide evidence that including state transitions across play sequences reduces the uncertainty about which team scores the next goal. To evaluate player performance, the Q-function quantifies the value of a player's action in a context. The action values are then aggregated over games and seasons to compute player impact scores. Value iteration learns Q-function values for each model state for scoring the next goal, receiving the next penalty, and winning the match. As validation, I compare my computed action values for scoring the next goal with the action values learned in THoR (Total Hockey Rating) [25]. Action impact values have a wide variance of the impact of actions with respect to states, showing context makes a substantial difference. To validate my player impact scores, I compare my rankings with player values learned in previous works. Player impact scores with respect to goal scoring correlate with plausible alternative scores, such as a player's total points, but improve on these measures, as the impact score is based on many more events. Learning player impact on receiving a penalty is a novel problem, and results for this are presented.

1.4 Contributions

The main contributions may be summarized as follows:

1. The first Markov Game model for a large ice hockey state space (over 1.3 million states), based on play-by-play sequence data.

2. Learning a Q-function that models play dynamics in the National Hockey League from a massive data set (over 2.8 million events). This application introduces a variant of AD-Trees as a data structure to (1) compute and store the large number of sufficient statistics required [17], and (2) efficiently support value iteration updates.

3. Applying the Q-function to define a context-aware look-ahead measure of the value of an action, over configurable objective functions (rewards).

4. Applying the context-aware action values to score player contributions, including how players affect penalties as well as goals. This impact score is a novel AI-based alternative to existing player scoring methods such as the +/- score.

1.5 Paper Organization

I begin in Chapter 2 with a review of related work in measuring player contributions and machine learning in sports. Next, some background information on the ice hockey domain and NHL play-by-play sequences data is given in Chapter 3. In Chapter 4, I give an overview of the Markov Game model and explain how the Markov Game model translates the hockey domain features into the Markov formalism. Next, I demonstrate how to construct a Markov Game model from ice hockey play-by-play events in Chapter 5. The implementation of a scalable value iteration algorithm for the ice hockey domain is then discussed in Chapter 6. Chapter 7 describes how action values and player values are computed from the results of the value iteration. The evaluation in Chapter 8 addresses the impact of context and lookahead, the two main advantages of the Markov Game model. The Markov Game model is applied to rank the aggregate performance of players and describe the resulting player rankings in Chapter 9. This work is viewed as taking the first step, not the last step, in applying AI modelling techniques to ice hockey, and is concluded with a number of potential extensions and open problems for future work in Chapter 10.

Chapter 2

Related Work

I use a Markov Game to model ice hockey dynamics. A **Markov Game** is a multi-agent variation of a Markov process [12]. A **Markov Process** is a stochastic transition model satisfying the Markov assumption, that is, where "the current state depends on only a finite fixed number of previous states" [22]. In my work, the Markov Game for ice hockey is a multi-agent variant of a first-order Markov Process. Related works discussed will cover the initial work on Markov Games with reinforcement learning. Recent advanced statistical methods for evaluating players in the NHL are also discussed. These advanced statistics form the basis for comparing action values and player values. Finally, I will discuss Markov Process models in ice hockey, as well as other sports.

2.1 Markov Games

[12] was the first to create Markov Games for reinforcement learning. Littman creates a Markov Game as a two-agent Markov Decision Process, where the two agents have opposing goals. He also uses Q-learning as the reinforcement learning technique to find optimal policies for each agent. I follow a similar approach for ice hockey, where the two agents are the Home and Away teams, and value iteration is used as a reinforcement learning technique to evaluate states. I use a Markov Game model because it can capture the opposing objectives of both teams (e.g. only the Home or Away team will win the game). The dynamic programming algorithm for value iteration learns the values of states in the Markov Game model to evaluate the actions of both teams simultaneously, something that can not be done with a single-agent Markov Decision Process. Rather than using

value iteration to determine an optimal policy for each team, value iteration learns following an "on policy" method [32], as the focus is on evaluating player actions rather than team strategies.

2.2 Evaluating Actions and Players

+/- is a statistic used in ice hockey and is calculated for each player in a game. +/- is calculated as the number of goals scored for a player's team while the player is on the ice, minus the number of goals scored by the opposing team while the player is on the ice. For example, if a player's team scores a goal while he is on the ice, the player's +/- will increase by 1. Conversely, if the player's team is scored on while he is on the ice, the player's +/- will decrease by 1. This represents the goal differential while the player was present on the ice, and is calculated for all players on the ice during a goal.

Several papers aim to improve the basic +/- score with statistical techniques [13, 16, 14, 6, 31]. These approaches are motivated by an adjusted +/- statistic used in the NBA [21]. A common approach used in these previous works is regression techniques where an indicator variable for each player is used as a regressor for a goal-related quantity (e.g., log-odds of a goal for the player's team vs. the opposing team). The regression weight measures the extent to which the presence of a player contributes to goals for his team or prevents goals for the other team. These approaches usually only look at goals, and sometimes shots, but no other player actions. They also do not adjust for home team advantage, as advocated by [25]. The only context these previous works take into account is which players are on the ice when a goal is scored. No other features, such as goal differential, manpower differential, or game time are used. Typically these improvements of +/- have either only examined a single season, therefore using a small dataset, or fixed player values across multiple seasons, when in reality player performance is subject to change across seasons. My model construction and player evaluation covers the entire 2007-2008 season through to part of the 2014-2015 season, and assign player values for each season, showing that player values change across different seasons.

In [20], goals scored by players are evaluated by examining how the goal affected the change in the probability of winning, in a metric called Added Goal Value (AGV). AGV accounts for the goal differential, manpower differential, and time remaining in a game to determine the impact of scoring a goal. It also uses a beta prior distribution and a third-order polynomial to smooth winning probabilities, and is the first work to examine modelling winning probabilities in the NHL. The Markov Game model I construct differs in that it not only evaluates goals, but all other player actions and their effects on winning the game, and therefore captures more information of hockey matches. Given that goals are rare occurrences relative to other actions, only evaluating players scoring goals reduces the valuation capability of AGV compared to the Impact rating I present. The Markov Game model also examines the effects of actions on other objectives, such as goals and penalties. While the Markov Game model construction algorithm I present does not encode timestamps for each action, the period is preserved as a context feature to capture some temporal information. As an extension to basic action labels (e.g. 'shot(Home)'), I include zone information as an additional action feature (e.g. 'shot(Home,Offensive)'). [20] also did not include the scoring rates for manpower differentials of 6-on-5, 6-on-4, and 6-on-3, which can occur when the goalie is pulled, however, these situations are included in my Markov Game model construction. [20] makes an assertion that the home and away teams have even odds of winning in overtime. Our contingency table (ref. Table 4.2) shows this assertion to be false, with home teams 5.7% more likely to score a goal in overtime than the away team.

The closest predecessor to my work is the Total Hockey Rating (THoR) [25]. THoR assigns a value to all actions, not only goals. Actions are valued by observing the net difference in goal scoring 20 seconds after the action occurred between the player's team performing the action and his opponent. For penalties, the duration of the penalty is used as the lookahead window. THoR uses data from the $2006/2007$ NHL season only. Without the context of an action, THoR assumes a fixed value for every action, which gives a natural bias for actions to be valued in favor of a player's team or in favor of the opposing team. In contrast, I show that most actions can have both positive and negative impact on the team performing the action, depending on the context. Furthermore, the lookahead window of 20 seconds restricts the lookahead value of each action. Q-learning on the other hand is not restricted to any particular time window for lookahead, allowing greater flexibility and more accurate evaluation of player actions.

2.3 Markov Process Models for Ice Hockey

A number of Markov Process models have been developed for ice hockey [33, 4]. The main difference to my work is these previous models do not include actions, and hence cannot model the impact of actions. In [4], special teams situations are analyzed to account for different scoring rates. Expected goals are then generated for matches and used to predict the outcome of matches. The Markov Game model I present is similar in that it can project expected goals and the outcome of matches (i.e. win/loss). The extension I take beyond the context feature space is to account for a larger feature space of action sequences, which is more beneficial for analyzing players.

In [33], only even strength situations are analyzed, and the model is only dependent on a few indicators. These indicators encode whether the home team is leading, away team is leading, or the teams are tied. The Markov Game model I present extends this to include the specific goal differential rather than a leading/trailing indicator for each team. I also incorporate both even strength and special teams situations, using the exact manpower differential. I use more available play-by-play data and analyze all player actions and contributions over all gameplay situations.

2.4 Markov Decision Process Models for Other Sports

MDP-type models have been applied in a number of sports settings, such as soccer [9] and baseball [30]. My work is similar in that it uses value iteration on a Markovian state space, however, previous Markov models in sports use a much smaller state space. For example, the soccer model of [9] uses only 4 states, and the baseball model of [30] utilizes only 12 states. To effectively model ice hockey dynamics, $1,325,809$ states are constructed when forming a Markov Game model from our NHL play-by-play data, a significant increase in the level of modelling detail. The goal of Markov Game models is traditionally to find an optimal policy for a critical situation in a sport or game. In contrast, the Markov Game model I present learns in the "on policy" setting whose aim is to model hockey dynamics as it is actually played. As such, I use the Markov Game model to evaluate player actions, states, and players, rather than generate team strategies. A potential application for improving play and advising coaches is in finding strengths and weaknesses of teams. The Q-function can be used to find situations in which a team’s mix of actions provides a substantially different expected result from that of a generic team, but this application is left for future work.

Chapter 3

Domain Description: Hockey Rules and Hockey Data

The rules of hockey are outlined first in Section 3.1 to set the framework for my research. Next, the format of the NHL play-by-play data is shown in Section 3.2, which forms the basis for construction of the Markov Game model. Finally, the process of forming the NHL play-by-play data in a relational database for effective use in my model is outlined in Section 3.3.

3.1 Hockey Rules

In this work, I describe a Markov Game model for ice hockey, specifically in the NHL. To motivate the model, I give a brief overview of rules of play in the NHL. For detailed rules of play in the NHL, refer to [18]. NHL games consist of three periods in regular play, each 20 minutes in duration. A team will try to score more goals than their opponent within three periods in order to win the game. If the game is still tied after three periods, the teams will enter a fourth overtime period, where the first team to score a goal wins the game. If the game is still tied after overtime during the regular season, a shootout will commence. Shootouts consist of 3 rounds where skaters will go one-on-one with the opposing goaltender and try to score a goal in one shot. If the score is still tied after 3 rounds, extra shootout rounds are added. During the playoffs, overtime periods are repeated until a team scores a goal to win the game. In regular play, teams have five skaters and one goalie on the ice, and are said to be at even strength. Penalties result in a player sitting in the penalty box for two, four, or five minutes and the penalized team will be shorthanded, creating a manpower differential between the two teams. When a team receives a penalty, the duration of play during the penalty is referred to as a penalty kill for the penalized team, and as a powerplay for their opponent. Penalty situations are also referred to as special teams situations, as coaches often pick specific players to increase their potential of scoring on the powerplay, or to improve defending the net on a penalty kill. Our Markov Game model is context-dependent, so we can also rank players in special teams situations. As ice hockey is a continuous-flow sport, players are permitted to return to the ice as soon as the duration of their penalty has been reached, ending the special teams situation. A **continuous-flow sport** is a sport where players play over time intervals in a continuous fashion, rather than in discrete series of bounded events, such as in baseball [8]. Teams can also pull their goalie to have an additional player on the ice and improve manpower differential in their favor, with the empty net also increasing the risk of being scored on.

3.2 Data Format

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. An ice hockey **play-by-play event** is an event in an ice hockey game recorded as it occurs in the game. The NHL play-by-play events are recorded in a play-by-play event log, where events are recorded in series as they occur. 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 3.1. A **sequence** in our data is formed from sequential play-by-play events concatenated together, typically starting with a faceoff and ending with a play stoppage indicator. Note that there are regularly only 30 teams in the NHL, but some teams were replaced and moved to new locations, so there are 32 teams recorded in this dataset. We also retrieve player salaries from http://nhlnumbers.com/ and http://www.dropyourgloves.com/ to supplement our analysis.

The events recorded by the NHL from the 2007-2008 regular season and onwards are listed in Table 3.2. There are two types of events: actions performed by players and start and end markers for each play sequence. Throughout my work, events that are

Table 3.1: Size of Dataset

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

player actions from the left column of Table 3.2 will be referred to as **actions** or **action-events**, and start or end of sequence markers from the right column will be referred to as **events**. For each event, the current goal differential GD, manpower differential MD, and period P are scraped from the play-by-play data. Every event is marked with a continuous timestamp, and every action is also marked with a zone Z and which team T, Home or Away, carries out the action. The Markov Game model I present does not make use of the continuous timestamps, although this feature is partially encoded in the play sequence ordering and using P as a context feature. Methods for using the continuous timestamps are discussed in future work.

Table 3.2: NHL Play-By-Play Events Recorded

Action Event	**Start/End Event**
Faceoff	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

3.3 Relational Database Setup

We describe the relational database setup for storing the NHL play-by-play data and creating the Markov Game model. First, the play-by-play data must be scraped from http://www.nhl.com using a custom web crawler. The scraped play-by-play data is written to

CSV files in a relational database format, imported into a relational database, and tables are formed as in Figure 3.1. Information shared by all play-by-play events is stored in a central play-by-play events table. Additional tables are created for each event type, which store event-specific information. To navigate from an event in the play-by-play data to the specific event in its corresponding event table, we reify each event to their event type using a star schema [22]. An ExternalEventId is recorded in the play-by-play table and references the unique identification number for the recorded event type in the event-specific table. For example, an event-specific table for shots on net would appear as in Figure 3.2. The players performing each action are recorded in the event-specific tables, and this information is used to apply the values of each action to each player. We also record useful information, such as shot location and shot type [10], but we do not include these features in our model.

GameId	AwayTeamId	HomeTeamId	ActionSequence	EventNumber	PeriodNumber	EventTime	EventType	ExternalEventId
2013020600	23	21	1	1	1	00:00:00	PERIOD START	2057
2013020600	23	21	1	2	1	00:00:00	FACEOFF	36328
2013020600	23	21	1	3	1	00:00:32	HIT	28310
2013020600	23	21	1	4	1	00:00:50	SHOT	33093
2013020600	23	21	1	5	1	00:01:06	PENALTY	5290
2013020600	23	21	2	6	1	00:01:06	FACEOFF	36329
2013020600	23	21	2	7	1	00:01:27	SHOT	33094
2013020600	23	21	2	8	1	00:01:54	SHOT	33095
2013020600	23	21	2	9	1	00:02:30	HIT	28311
2013020600	23	21	2	10	1	00:02:45	BLOCKED SHOT	17068
2013020600	23	21	2	11	1	00:03:11	STOPPAGE	28271
2013020600	23	21	3	12	1	00:03:11	FACEOFF	36330
2013020600	23	21	3	13	1	00:03:14	BLOCKED SHOT	17069
2013020600	23	21	3	14	1	00:03:25	SHOT	33096
2013020600	23	21	3	15	1	00:04:11	STOPPAGE	28272
2013020600	23	21	4	16	1	00:04:22	FACEOFF	36331
2013020600	23	21	4	17	1	00:05:15	HIT	28312
2013020600	23	21	4	18	1	00:05:29	SHOT	33097
2013020600	23	21	4	19	1	00:05:59	BLOCKED SHOT	17070
2013020600	23	21	4	20	1	00:06:36	STOPPAGE	28273
2013020600	23	21	5	21	1	00:06:36	FACEOFF	36332
2013020600	23	21	5	22	1	00:07:04	STOPPAGE	28274
2013020600	23	21	6	23	1	00:07:04	FACEOFF	36333
2013020600	23	21	6	24	1	00:07:12	HIT	28313
2013020600	23	21	6	25	1	00:07:32	MISSED SHOT	13998

Figure 3.1: Play-by-Play Data in Relational Database

Players exist as an additional entity in the relational database, and are given unique identification numbers to facilitate table joins and quick searching. The player identification numbers match those on http://www.nhl.com and Figure 3.3 shows a sample from the players table. The data we scrape includes additional player features, such as age, size, birthplace, and draft information. While these player features are informative for predicting

ShotId	ShotByTeam	ShootingPlayerId	ShotType	Zone	Distance
1	MTL	8475848	Wrist	offensive	23
2	MTL	8471504	Wrist	offensive	46
3	MTL	8470324	Slap	offensive	50
4	TOR	8474037	Backhand	offensive	35
5	TOR	8470602	Wrist	offensive	59
6	TOR	8470602	Slap	offensive	57
7	MTL	8476851	Wrist	offensive	18
8	TOR	8470867	Wrist	offensive	14
9	MTL	8470324	Slap	offensive	59
10	MTL	8471504	Wrist	offensive	24
11	MTL	8467407	Tip-In	offensive	11
12	TOR	8475098	Wrist	offensive	10
13	TOR	8473548	Wrist	offensive	20
14	MTL	8464975	Wrist	offensive	26
15	TOR	8475098	Wrist	offensive	25
16	TOR	8474037	Wrist	offensive	9
17	TOR	8473548	Snap	offensive	36
18	TOR	8471245	Wrist	offensive	59
19	MTL	8474056	Slap	offensive	50
20	TOR	8471245	Wrist	offensive	14

Figure 3.2: Shot Event Table

player performance [3, 2], I restrict our model to examining play sequences and sequence context features for player analysis in the spirit of descriptive statistics.

The overall entity-relationship diagram for our relational database is shown in Figure 3.4. We see events in the central play-by-play events table are reified to all other event-specific tables through the ExternalEventId in a star schema [22]. All of these tables link to the player, game, and team tables. Event-specific tables for events denoting start and end markers all have similar structures, as their only purpose is for starting or terminating sequences and they are not used to evaluate players. Lines between tables represent foreign key dependencies in our star schema.

Figure 3.5 shows the overall process of our work. We start with a relational database with tables for play-by-play events, players, and event-specific tables. The Markov Game model construction algorithm uses the play-by-play events and event-specific tables to generate a Markov Game model. Internally, we also create a separate table mapping play-by-play events to edges in the Markov Game model. This is to facilitate mapping player

PlayerId	PlayerName	Position	Birthplace	Country	Height	Weight
8462129	Michal Handzus	C	Banská Bystrica, Slovakia	Slovakia	6'5"	215
8464975	Daniel Briere	C	Gatineau, QC, Canada	Canada	5'9"	174
8465058	Michal Rozsival	D	Vlasim, Czech Republic	Czech Republic	6'1"	210
8465914	Francis Bouillon	D	New York City, NY, United States	United States	5'8"	194
8466140	Olli Jokinen	C	Kuopio, Finland	Finland	6'2"	210
8466148	Marian Hossa	R	Stará Lubovna, Slovakia	Slovakia	6'1"	207
8466251	Jason Chimera	L	Edmonton, AB, Canada	Canada	6'3"	213
8467365	John Erskine	D	Kingston, ON, Canada	Canada	6'4"	220
8467407	Brian Gionta	R	Rochester, NY, United States	United States	5'7"	176
8467496	Andrei Markov	D	Voskresensk, Russia	Russia	6'0"	201
8468064	Martin Erat	R	Trebic, Czech Republic	Czech Republic	6'0"	196
8468095	George Parros	R	Washington, PA, United States	United States	6'5"	224
8468208	Joel Ward	R	North York, ON, Canada	Canada	6'1"	226
8468635	Travis Moen	L	Stewart Valley, SK, Canada	Canada	6'2"	210
8468778	Colton Orr	R	Winnipeg, MB, Canada	Canada	6'3"	222
8469501	Chris Thorburn	R	Sault Ste. Marie, ON, Canada	Canada	6'3"	235
8469508	Jay McClement	C	Kingston, ON, Canada	Canada	6'1"	205
8469521	Tomas Plekanec	C	Kladno, Czech Republic	Czech Republic	5'11"	196
8469544	Patrick Sharp	L	Winnipeg, MB, Canada	Canada	6'1"	198
8469639	Brooks Laich	C	Wawota, SK, Canada	Canada	6'2"	210
8469665	Johnny Oduya	D	Stockholm, Sweden	Sweden	6'0"	188
8470085	Paul Ranger	D	Whitby, ON, Canada	Canada	6'3"	210

Figure 3.3: Player Table

actions to action impact values. The value iteration algorithm uses the Markov Game model to generate Q-values for each state in the Markov Game model. We then compute impact values across edges, corresponding to action impact values. The player evaluation pairs action impact values with the players performing the actions. Finally, we group and sort these player action values over games and over seasons to generate player impact scores and player rankings in each NHL season.

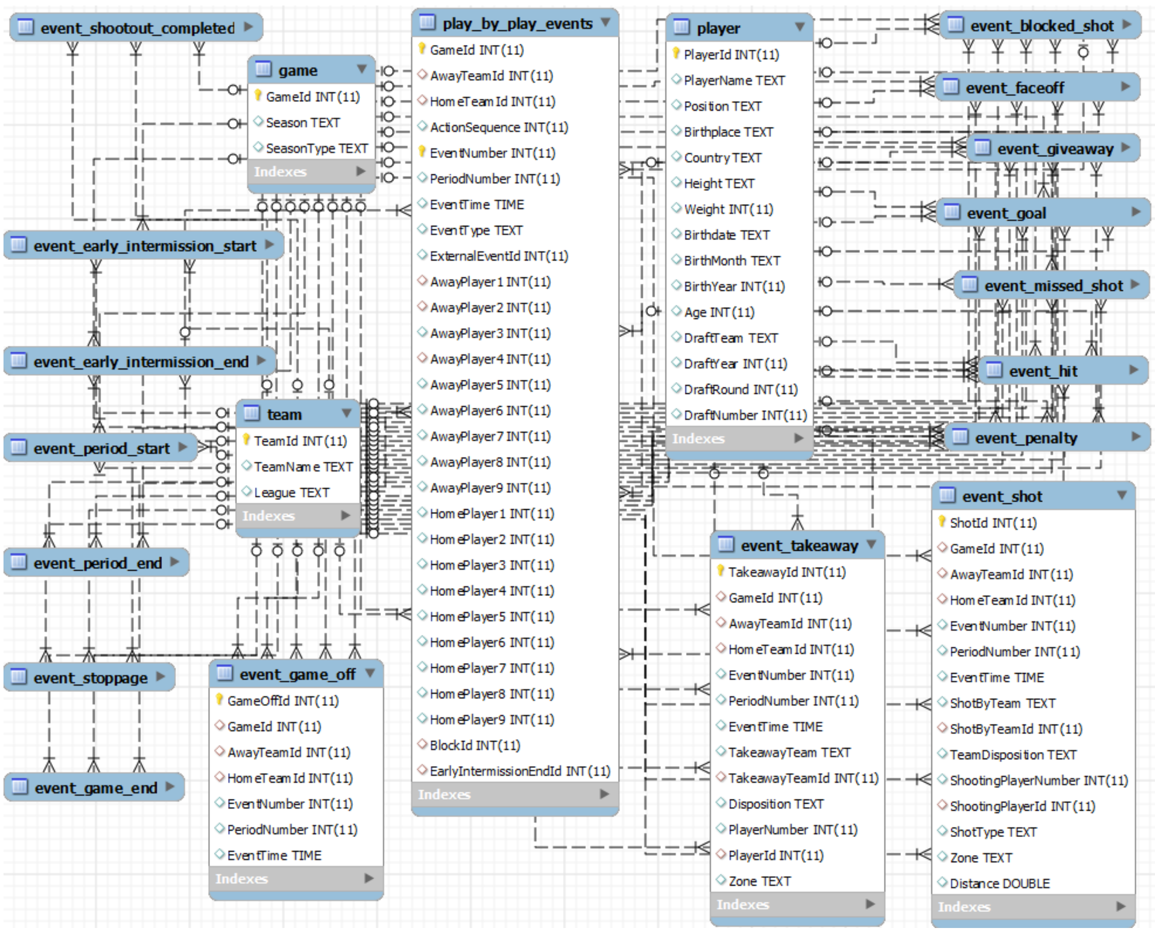

Figure 3.4: Entity-Relationship Diagram

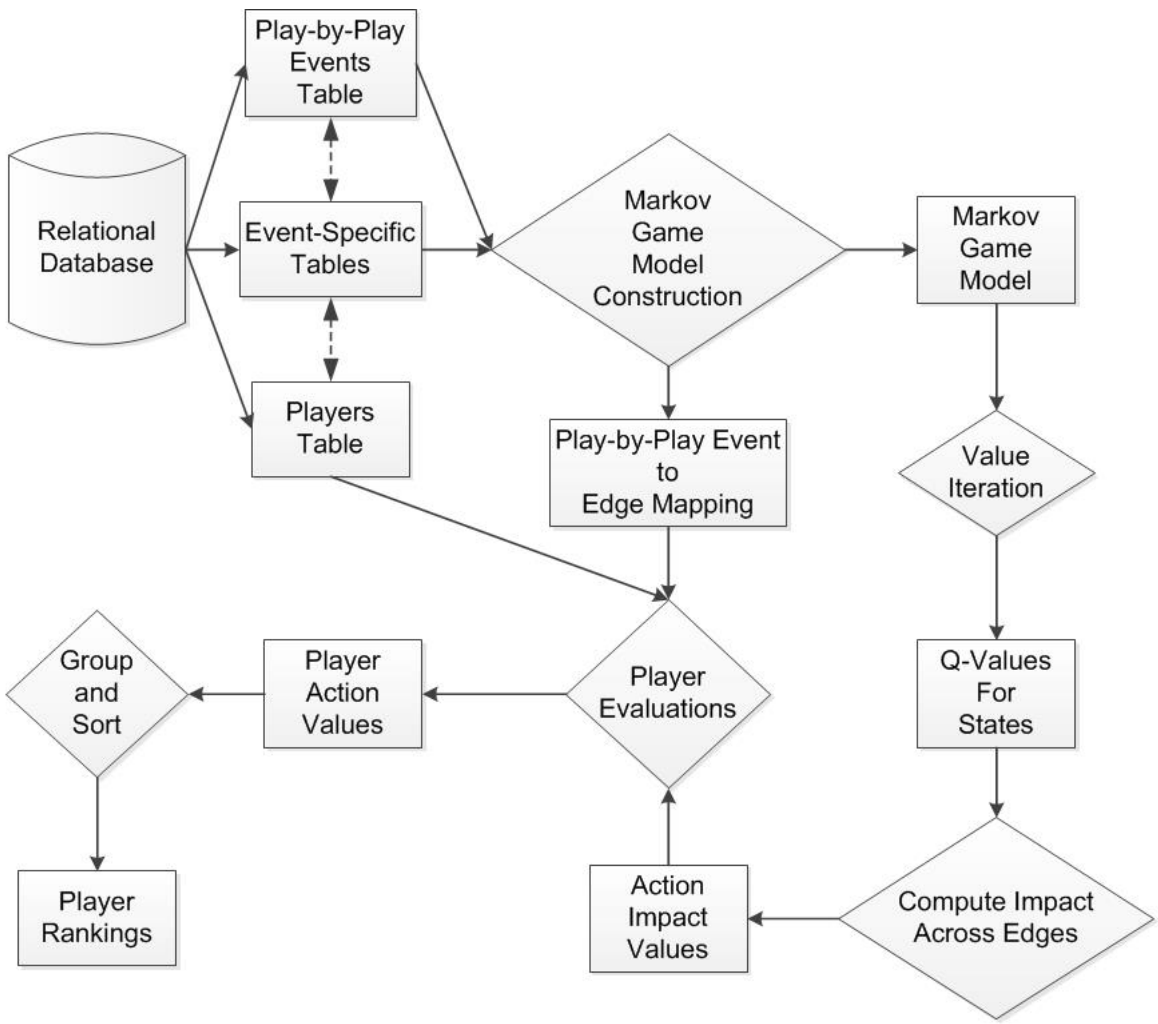

Figure 3.5: System Flow

Chapter 4

Markov Games

In its general form, a Markov Game [12], sometimes called a stochastic game, is defined by a set of states, S, and a collection of action sets, one for each agent in the environment. A **state** captures the information of the current observed gameplay. The **action set** is the set of player action events, a subset of the recorded play-by-play event types, from the left column of Table 3.2. State **transitions** are controlled by the current state and one action from each agent, and model a shift in gameplay from one state to another. Each state transition has an associated **transition probability**, representing the probability of the state transition occurring given the current state. For each agent, there is an associated **reward** function mapping a state transition to a reward. An overview of how our Markov Game model fits the Markov Game schematic is as follows. There are two agents, the Home Team H and the Away Team A. The game is zero-sum, meaning whenever a home team receives a reward, the Away Team receives minus the reward. Therefore we can simply use a single reward value, where positive numbers denote a reward for the home team (the maximizer), and negative number a reward for the Away Team (the minimizer). In each state, only one team performs an action, although not in a turn-based sequence. This reflects the way the NHL records actions, and motivates my choice of value iteration over other score-based computations such as minimax. Thus at each state of the Markov Game, exactly one team, or player in this two-agent game, chooses No-operation, meaning that team does not perform an action. Actions in a game are performed by a player as a member of their team. Since the Markov Game model is designed from the perspective of two teams as agents, actions are described as team actions rather than player actions. In order to evaluate players in Chapter 7, the value of a team action is applied to the player performing the action, measuring their contribution to their team. State transitions follow a semi-episodic model [32] where play moves from episode to episode, and information from past episodes is recorded as a list of *context features*. The past information includes the goal score, manpower, and period as context features, as well as the action history. A sequence in the NHL play-by-play data corresponds to an episode in Markov Decision Process terminology. *Within* each episode/sequence, the Markov Game model is essentially a game tree with perfect information as used in AI game research [22]. The following generic notation is introduced for all states. MDP notation follows [22], and a modification of the Markov Game notation described in [12] is used as follows.

- $Occ(s)$ is the number of occurrences of state s as observed in the play-by-play data. s forms a node in the transition graph of the Markov Game model.
- $Occ(s, s')$ is the number of occurrences of state s being immediately followed by state s' as observed in the play-by-play data. (s, s') forms an edge in the transition graph of the Markov Game model.
- The transition probability function TP is a mapping of $S \times S \to (0, 1]$. It is estimated using the observed transition frequency $\dfrac{Occ(s, s')}{Occ(s)}$.

My Markov Game model extends from previous models in two ways. The first way is by including a larger set of context features. The second way is by including a history of actions, i.e. play sequences, as part of a state, which is a major extension in the level of modelling detail. To build the state space S for the Markov Game model, context features are defined, followed by play sequences.

4.1 State Space: Context Features

Previous work on Markov Process models for ice hockey [33] defined states in terms of hand-selected features that are intuitively relevant for the game dynamics, such as goal differential and penalties. Such features will be referred to as **context features**. Context features remain the same throughout each play sequence, with the exception of manpower differential. For example, a goalie can be pulled by his team during a play sequence and substituted for an additional skater. Penalized players are also allowed to return to the ice once the duration of their penalty has been reached, changing the manpower differential during the play sequence.

A **context state** lists the values of relevant features at a point in the game, and provides the initial information prior to a play sequence occurring. These relevant context features are shown in Table 4.1, together with the range of integer values observed.

Table 4.1: Context Features

Notation	Name	Range
GD	Goal Differential	[-8,8]
MD	Manpower Differential	[-3,3]
P	Period	[1,7]

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 as Number of Home Skaters on Ice - Number of Away Skaters on Ice. A positive manpower differential typically means the home team is on the powerplay (away team is penalized), and a negative manpower differential typically means the home team is shorthanded (away team is on the powerplay). The other occurrences of manpower differentials are when a goalie is pulled and an extra skater comes on the ice, which are not powerplay or shorthanded situations. Period P represents the current period number the play sequence occurs in, typically ranging in value from 1 to 5, but can extend during the playoffs, where extra overtime periods may be necessary. Periods 1 to 3 are the regular play of an ice hockey game, and periods 4 and onwards are for overtime and shootout periods as needed.

Potentially, there are $(17 \times 7 \times 7) = 833$ context states. We derive this count from 17 GD values, 7 MD values, and 7 P values. In this NHL dataset, we observe 450 context states occur at least once. Table 4.2 is a contingency table of the context states that includes statistics for the top-25 most frequent context states over all $590,924$ play sequences, and lists $52,793$ total goals and $89,612$ total penalties. Positive differences are for the home team and negative differences are for the away team. For example, a Goal Difference of 7.1% means the home team is 7.1% more likely to score a goal in that context state than the away team. Similarly, a Penalty Difference of -33.2% means the away team is 33.2% more likely to receive a penalty in that context state than the home team.

Table 4.2: Statistics for Top-25 Most Frequent Context States

Goal Differential	Manpower Differential	Period	Number of Sequences	Number of Goals	Goal Difference	Number of Penalties	Penalty Difference
0	0	1	78,118	5,524	7.1%	11,398	-2.3%
0	0	2	38,315	2,935	7.6%	5,968	-2.9%
0	0	3	30,142	2,050	5.9%	3,149	-2.2%
1	0	2	29,662	2,329	2.0%	4,749	2.2%
1	0	3	25,780	2,076	4.3%	3,025	3.5%
-1	0	2	25,498	1,970	8.6%	4,044	-8.7%
1	0	1	24,721	1,656	5.3%	4,061	3.4%
-1	0	3	22,535	1,751	0.7%	2,565	-18.3%
-1	0	1	20,813	1,444	4.6%	3,352	-8.1%
2	0	3	17,551	1,459	6.9%	2,286	-0.9%
2	0	2	15,419	1,217	2.7%	2,620	2.9%
-2	0	3	13,834	1,077	-2.3%	1,686	-12.6%
0	1	1	12,435	1,442	64.8%	2,006	65.9%
-2	0	2	11,799	882	3.9%	1,927	-15.7%
0	-1	1	11,717	1,260	-54.8%	2,177	-44.7%
3	0	3	10,819	678	0.3%	1,859	1.2%
-3	0	3	7,569	469	7.0%	1,184	-6.3%
0	1	2	7,480	851	57.0%	1,157	25.7%
0	0	4	7,024	721	5.7%	535	-10.7%
0	-1	2	6,853	791	-52.5%	1,160	-37.4%
3	0	2	6,405	472	0.4%	1,184	8.1%
2	0	1	6,057	394	6.1%	1,050	9.1%
1	-1	2	5,716	701	-56.1%	915	-28.1%
1	1	2	5,579	677	58.1%	949	26.7%
-1	1	2	5,252	628	57.6%	831	21.3%

A number of previous papers on hockey dynamics have considered the context features of play sequences [4, 33, 20]. The important trends that are possible to glean from statistics such as those shown in Table 4.2 have been discussed in several papers [24, 33, 1]. Data analysis confirms these observations on our dataset, which is a larger dataset than previously used. Notable findings include the following.

1. While goals and penalties are rare when compared to the total number of actions and events, 24.1% of all play-by-play sequences end in either a goal or a penalty.
   (a) 8.9% of all play-by-play sequences end in a goal.
   (b) 15.2% of all play-by-play sequences end in a penalty.
2. Home team advantage: the same advantages in terms of context features translate into higher scoring rates.
3. Penalties are more frequent than goals, except for the 4th period (cf. [24]).
4. Gaining a powerplay substantially increases the probability of scoring a goal [33].
5. Gaining a powerplay also significantly increases the conditional probability of receiving a penalty [24, 1].
   (a) When the home team goes on the powerplay in Period 1, the conditional probability of the home team receiving a penalty increases from 48.9% to 65.9%.

(b) When the away team goes on the powerplay in Period 1, the conditional probability of the away team receiving a penalty increases from 51.1% to 72.3%.

6. Short-handed goals are surprisingly likely: a manpower advantage translates only into a goal scoring difference of at most 64.8%. (Powerplay for the home team in period 1.)

   (a) If a goal is scored on the powerplay, it is 76.2% likely to be a powerplay goal and 23.8% likely to be a shorthanded goal. We computed this from the full contingency table by summing all goals scored for the team on the powerplay and all goals scored for the shorthanded team and dividing each by total goals scored in special teams situations.

   (b) If the away team is on the powerplay, they can be up to 55% more likely to score the next goal.

   (c) If the home team is on the powerplay, they can be up to 65% more likely to score the next goal.

7. Although it is obvious that goals win games, our contingency table quantifies how scoring a goal significantly increases the probability of winning.

   (a) When the home team scores a goal in Period 2 for a one goal lead, their probability of winning increases from 53.8% to 72.5%. If the home team scores another goal in Period 2 for a two goal lead, the probability of winning increases further to 86.5%.

   (b) When the away team scores a goal in Period 2 for a one goal lead, their probability of winning increases from 46.2% to 66.6%. If the away team scores another goal in Period 2 for a two goal lead, the probability of winning increases further to 84.0%.

8. When observing even-strength, tied game situations, it is interesting to note that there is a slight increase in the goal difference and penalty difference from the 1st period to the 2nd period, but these values fall when moving to the 3rd period. A possible explanation is that players are more cautious in the 1st and 3rd periods during even-strength, tied game situations. Players become more tired in the 2nd period, which

may cause more penalties, and they may also be more willing to take risks to try and score goals.

These context features are useful in modelling hockey dynamics as a Markov process. While such patterns provide interesting and useful insights into hockey dynamics, such as how goal scoring or penalty rates depend on the game context [33], they do not consider action events. This means that analysis at the sequence level does not consider the internal dynamics within each sequence, and that it is not suitable for evaluating the impact of hockey actions. Next, I extend the state space beyond context features to include play sequences of actions.

4.2 State Space: Play Sequences

The state space is extended from only context features to include actions and action histories. The basic set of 8 possible actions is listed in Table 3.2. Each of these actions has two parameters: which team T performs the action a and the zone Z. Zone Z represents the area of the ice rink in which an action takes place. Z can have values Offensive, Neutral, or Defensive, relative to the team performing an action. For example, Z = Offensive relative to the home team is equivalent to Z = Defensive relative to the away team. A specification of an action plus parameters is an **action event**. Using action description language notation [11], action events are written in the form $a(T, Z)$. For example, $\mathit{faceoff}(Home, Neutral)$ denotes the home team wins a faceoff in the neutral zone. Usually the action parameters are omitted from generic notation and a is written for a generic action event.

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. Table 3.2 displays start and end markers in the right column, noting that shots and faceoffs are also valid start markers, and goals are also valid end markers. The empty history $\emptyset$ is also allowed as a valid play sequence. A **complete** play sequence is a play sequence ending with an end marker. A **state** is a pair $s = \langle \mathbf{x}, h \rangle$ where $\mathbf{x}$ denotes a list of context features and h an action history. State s is formulated as a play sequence consisting of action events $a_1, a_2, \ldots, a_n$ as the **action history**, together with a particular GD, MD, and P as the **context features**. If the sequence is empty, then state s is purely a context state. Table 4.3

shows an example of a NHL play-by-play action sequence in tabular form. Potentially, there are $(7 \times 2 \times 3)^{42} = 42^{42}$ action histories. This is derived from the 7 player actions, 2 teams in a match, and 3 zones where the action can occur. The exponent of 42 is derived from the maximum observed sequence length, shown in Table 4.4. In our NHL dataset, $1,325,809$ states, that is, combinations of context features and action histories, occur at least once. Play-by-play sequence data is stored in SQL tables (see Table 4.3). SQL provides fast retrieval, and native support for the necessary COUNT and SUM operations.

Table 4.3: Sample Play-By-Play Data in Tabular Format

GameId	**Period**	**Sequence Number**	**Event Number**	**Event**
1	1	1	1	PERIOD START
1	1	1	2	faceoff(Away,Neutral)
1	1	1	3	STOPPAGE
1	1	2	4	faceoff(Home,Neutral)
1	1	2	5	shot(Away,Offensive)
1	1	2	6	hit(Away,Neutral)
1	1	2	7	STOPPAGE
1	1	3	8	faceoff(Home,Offensive)
1	1	3	9	goal(Home,Offensive)
1	1	4	10	faceoff(Home,Neutral)
1	1	4	11	shot(Home,Offensive)
1	1	4	12	STOPPAGE
1	1	5	13	faceoff(Away,Defensive)
1	1	5	14	STOPPAGE
1	1	6	15	faceoff(Home,Defensive)
1	1	6	16	hit(Away,Offensive)
1	1	6	17	hit(Home,Offensive)
1	1	6	18	STOPPAGE
1	1	7	19	faceoff(Away,Defensive)
1	1	7	20	hit(Home,Offensive)
...				

It is noteworthy that sequences ending in a goal tend to be longer in length, as also observed by [33], and consist of 5.85 events on average, as shown in Table 4.4. A possible explanation is that longer play sequences have players on the ice for a longer duration, with less time to rest. This can cause players to make mistakes that may lead to goals. Another possible explanation is that goals are often followed by many actions in quick succession. This fact was found by creating a decision tree for the sequence data, as a data mining exercise. One of the few significant findings from the decision tree was that when 3 events happen in quick succession (i.e. under 5 seconds), the last event was most likely to be a goal event. These possible explanations for temporal dependencies on goal scoring could be reinforced by modelling continuous time intervals between events, but this is left as future work.

Table 4.4: Event Sequence Statistics

Sequence Length	Maximum	Average	Variance
Overall	42	4.87	10.95
Sequence ends in a goal	38	5.85	9.66
Sequence ends in a penalty	42	4.10	10.92

4.3 State Transitions

If h is an incomplete play sequence, the play sequence that results from appending a to h is written as $h \star a$, 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. This notation utilizes the fact that context features do not change until an end marker is reached. For example, the goal differential does not change unless a goal event occurs. If h is a complete play sequence, then the state $\langle \mathbf{x}, h \rangle$ has a unique successor $\langle \mathbf{x}', \emptyset \rangle$, where the mapping from $\mathbf{x}$ to $\mathbf{x}'$ is determined by the end marker. For instance, if the end marker is $goal(Home, *)$, then the goal differential increases by 1. A sample of the state transition graph is shown in Figure 4.1. Note that $R(s)$ are the rewards for each state, which will be defined in Section 4.4.

Since the complete action history is encoded in the state, action-state pairs are equivalent to state pairs. Therefore transitions are modeled from state to state only, rather than transitions from state to state given an action, even though our main interest is in the effects of actions. For example, $Q(s \star a)$ is written to denote the expected reward from taking action a in state s, where Q maps states to real numbers, rather than mapping action-state pairs to real numbers, as is more usual. In reinforcement learning terms, this means the Q-function can be computed by value iteration applied to states, rather than on action-state pairs.

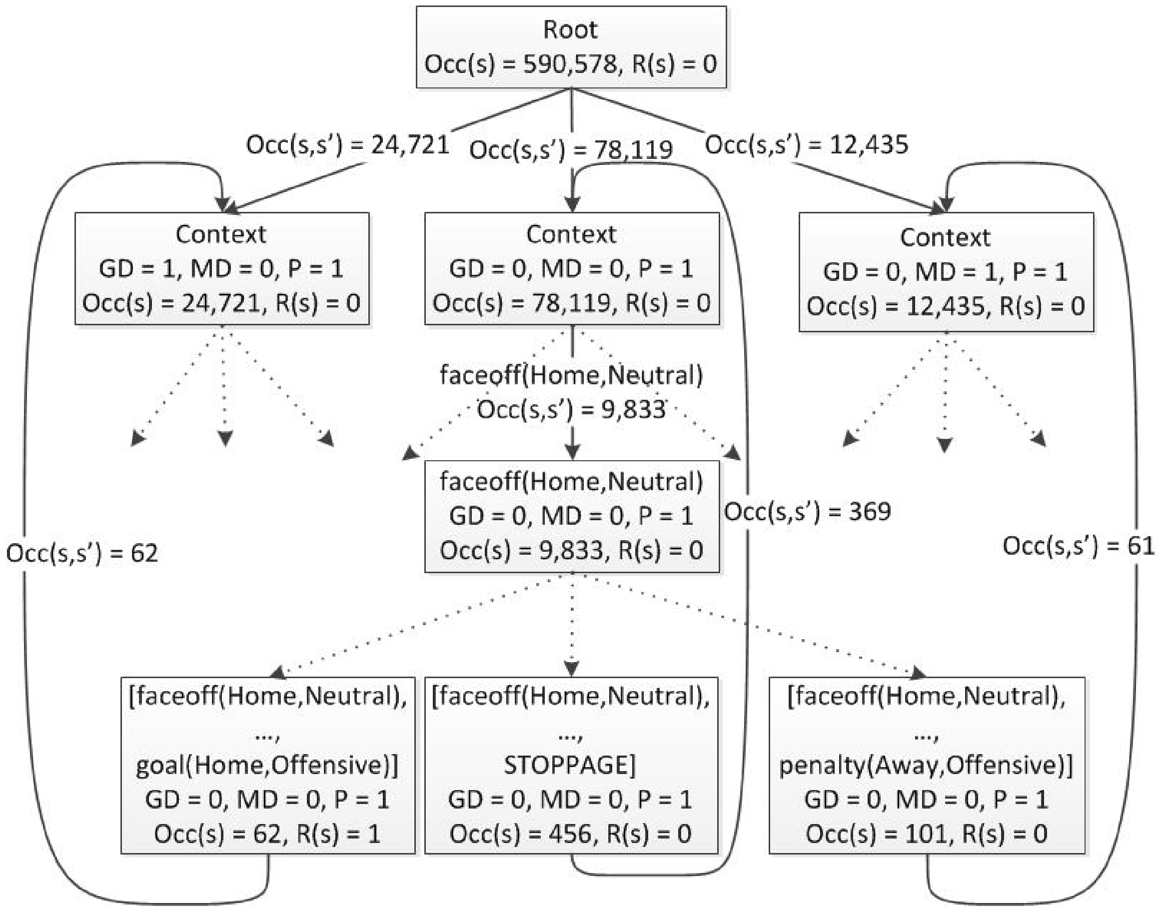

Figure 4.1: State Transition Graph

4.4 Reward Functions

A strength of Markov Game modelling is that value iteration can be applied to many reward functions in the model, depending on the results of interest. The reward functions we use in these experiments are focused on scoring goals, receiving penalties, and winning the match. Receiving penalties can be viewed as a cost rather than a reward, as receiving penalties decreases a team's chances of goal scoring and winning a match. These reward objectives are important events that change the flow of an ice hockey game. The corresponding Q-functions are easily interpreted, and are discussed in Section 6.1 as a precursor to setting up the value iteration computation. Recall that states encode action histories, so rewards are defined as being associated with states only rather than state-action pairs. The objective for the probability of the next goal can be represented in the Markov Game model as follows.

1. For any state s with a complete play sequence that ends in a Home resp. Away goal, set $R_H(s) := 1$ resp. $R_A(s) := 1$. For other states the reward is 0.

2. Any state s with a complete play sequence that ends in a Home or Away goal is an absorbing state (no transitions from this state).

With these definitions, $Q_H(s)$ 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. The cost function for receiving the next penalty and the reward function for winning the match can be represented in exactly the same way.

Our Markov Game model can also be used for computing expected values rather than probabilities. For objectives of expected values, the value of rewards differ from probabilistic objectives, as the reward values of both teams are considered rather than a single team's probability. For example if the objective is the expected number of goals in an Expected Goals Model, the rewards are defined as follows:

1. For any state s with a complete play sequence that ends in a Home goal, set $R_H(s) := 1$. For any state s with a complete play sequence that ends in an Away goal, set $R_A(s) := -1$. For other states the reward is 0.

2. No states are considered as absorbing sequences. As such, this allows the model to be extended for result prediction, but this is left as future work.

The flexibility in reward functions and, therefore, Q-functions (see Section 6.1), allows our Markov Game model to be compared against multiple advanced statistics with respect to different objectives, such as expected goals [15] and winning [25, 20].

Chapter 5

Constructing the Markov Game Model

Once the necessary components of the Markov Game model are well-defined, the main computational challenge is to build a data structure for managing the large state space. The state space is large because each (sub)sequence of actions defines a new state. Since we are modelling the actual hockey dynamics in the "on policy" setting, only action sequences observed in some NHL matches need to be considered, rather than the much larger space of all possible action sequences. This significantly reduces the size of the state transition graph and allows faster execution of the dynamic programming algorithm used on our Markov Game model. As such, the next step is constructing the Markov Game model as a state transition graph. First, an informal description of the construction algorithm is given in Section 5.1. Next, the steps of the algorithm are given in Section 5.2. Finally, we give a short example of the algorithm execution in Section 5.3.

5.1 Informal Description

Plays in the NHL form natural sequences of actions, typically starting with a faceoff and ending with a goal, penalty, or play stoppage. The actions in each play sequence can be viewed as actions performed by each team. In Markov Games, each agent, or team, performs an action to transition to a new state. It is intuitive to then transform these sequences of events into a tree of events, or a game tree, where each subsequent event in a sequence is the child state of the preceding event. We must also account for the context of a play sequence, so the tree must include the starting context of each play sequence as a state. The graph construction is performed as follows: the tree is initialized with a root state, or root node of the graph, where there is no context or sequence information. This is followed by a new node representing the context of the game the play sequence is starting in, but contains no sequence information. Next, the sequence of events follow below the context node, with branches forming as different events occur over multiple sequences. The process is repeated for each new play sequence by starting from the root node and adding new states, or nodes in the graph, as new action sequences are observed. The number of observances at each node is recorded and updated through each iteration. The levels of the sequence tree can be viewed as starting with no information in the first level (root node), adding context information to the second level (context node), and adding observed action histories to the following levels (event nodes).

Actions, such as penalties, often have an effect on the following play sequences. In order to propagate these effects, an edge is added from each leaf node to the context state node of the following play sequence. Each leaf node corresponds to a play sequence ending with a goal or an end marker. This loopback edge causes the state transition graph to become cyclic. As such, adding a loopback edge transforms the graphical model from a tree structure into a multi-agent Markov Decision Process called a Markov Game Model. For an in-depth explanation of Markov Decision Processes, refer to [22]. For more details on Markov Game Models, refer to [12].

5.2 Construction Algorithm

We use a modified version of the classic AD-tree structure [17] to compute and store sufficient statistics over observed action sequences. The AD-tree is a tree of play sequences where a node is expanded only with those successors observed in at least one match. The play sequence tree is augmented with additional edges that model further state transitions; for example, a new action sequence is started after a goal. The augmented AD-tree structure compactly manages sufficient statistics, in this case state transition probabilities and state occurrences. It also supports value iteration updates very efficiently.

The algorithm for Context-Aware Markov Game model construction is shown in Algorithm 1 and is described as follows. The root node initializes the graph, and is an empty node with no context or event information. Values backed up to the root node give a baseline for beliefs about goals, penalties, and winning when there is no match information.

For each node, the context information, that is, goal differential GD, manpower differential MD, and period P, are set when the new node is created, and the new action a is added to the sequence along with the zone Z that a occurs in. Nodes are also assigned unique identification numbers to facilitate table joins when gathering results. The reward $R(s)$ is also applied to each node, and the value of $R(s)$ is dependent on the objective function, as discussed in Section 4.4. The node counts $Occ(s)$ and edge counts $Occ(s, s')$ are applied to each node and edge respectively, and are used to generate transition probabilities TP for the value iteration using observed frequencies. The function $incrementCount(s)$ is used to update node count $Occ(s)$, and $incrementCount(s, s')$ is used to update edge count $Occ(s, s')$. Both functions increment the count by 1. The NHL play-by-play event data records goals, but no separate event for the shot leading to the goal exists. Following [25], this algorithm records the shot leading to the goal in addition to the goal itself by injecting a shot event into the event sequence prior to the goal. In order to facilitate backup computation for winning the match, an additional graph node, signifying a home team win or away team win, is added as a child node from the leaf node corresponding to the last event in the play-by-play data for the match. The state transition graph is stored in two tables in a MySQL database, one table for nodes and another for edges. Nodes are given unique identification numbers, and the edge table references these identification numbers as foreign keys.

5.3 Example

A step-by-step example of the Markov Game model constructions highlights the details of the algorithm. The example will follow sample play-by-play data in Table 5.1, and the construction algorithm will analyze expected goals for rewards. First, the algorithm creates the root node with no context information, and the occurrences are updated. This is shown in Figure 5.1.

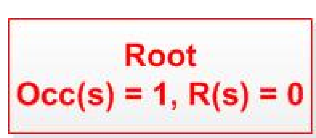

Figure 5.1: Construction: Step 1

Algorithm 1 Context-Aware Markov Game Model Construction

Require: NHL play-by-play data, win data w

```
root = new Node(empty)
for all games g do
    current = root
    previous = null
    lastLeaf = false
    for all events i in game g do
        if current == root then
            incrementCount(root)
            state = i.getStateInformation
            if not root.hasChild(state) then
                root.addChild(state)
            end if
            current = state
            incrementCount(current)
            incrementCount(root, current)
            if lastLeaf == true then
                if not previous.hasChild(current) then
                    previous.addChild(current)
                end if
                incrementCount(previous, current)
                lastLeaf = false
            end if
        end if
        if i.event ==GOAL then
            shotEvent = new Node(i, "SHOT")
            if not current.hasChild(shotEvent) then
                current.addChild(shotEvent)
            end if
            incrementCount(current, shotEvent)
            incrementCount(shotEvent)
            previous = current
            current = shotEvent
        end if
        event = new Node(i)
        if not current.hasChild(event) then
            current.addChild(event)
        end if
        incrementCount(current, event)
        incrementCount(event)
        previous = current
        current = event
```

Algorithm 1 Context-Aware Markov Game Model Construction (continued)

```
    if current.isEndMarker() then
      lastLeaf = true
      previous = current
      current = root
    end if
  end for
  win = new Node(w)
  if not previous.hasChild(win) then
    previous.addChild(win)
  end if
  incrementCount(previous, win)
  incrementCount(win)
end for
```

Table 5.1: Sample Play-By-Play Data

Goal Differential	Manpower Differential	Period	Event Number	Event
0	0	1	1	PERIOD START
0	0	1	2	faceoff(Away,Neutral)
0	0	1	3	hit(Away,Neutral)
0	0	1	4	penalty(Away,Neutral)
0	1	1	5	faceoff(Home,Neutral)
...				

Next, the context node is created. Since the first event of the match is being processed, the context information must be extracted from the first event of the first play sequence. The event is PERIOD START and has context features $GD = 0$, $MD = 0$, and $P = 1$, as this is context in which all ice hockey matches start. The context node is created with this context information, and the action history is empty, as no events are processed yet. Next, we create an edge from the root node to the context node. The occurrences of the context node is updated, as well as the occurrences of the transition edge from the root node to the context node, and the reward is applied to the context node. This will facilitate computing the state transition probabilities for the value iteration computation. Adding the context node and the edge from the root node to the context node is shown in Figure 5.2.

Next, we process the first event of the first play sequence, PERIOD START. A new event node will be created for PERIOD START with the same context information as before, $GD = 0$, $MD = 0$, and $P = 1$. The reward value is set for the event node and the occurrences are updated. An edge from the context node to the event node is created

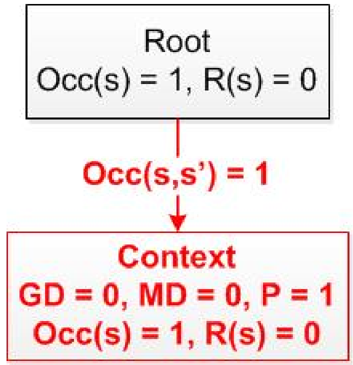

Figure 5.2: Construction: Step 2

signifying the event PERIOD START. The occurrences of this edge are also updated. This step is shown in Figure 5.3.

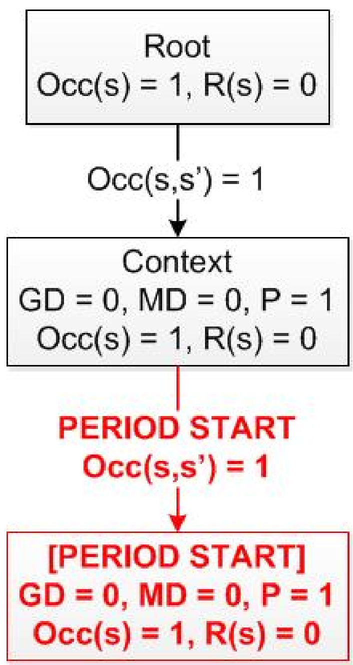

Figure 5.3: Construction: Step 3

Next, the following event, faceoff(Away,Neutral) is processed. Again, a new event node will be created for faceoff(Away,Neutral) with the same context information as before. The event history is appended with the new event, faceoff(Away,Neutral). The reward values and occurrences are also set. An edge from the previous event node to the new event node is created signifying the event faceoff(Away,Neutral) and the occurrences are updated. This step is shown in Figure 5.4. This process is repeated for the following events, hit(Away,Neutral) and penalty(Away,Neutral), as shown in Figure 5.5.

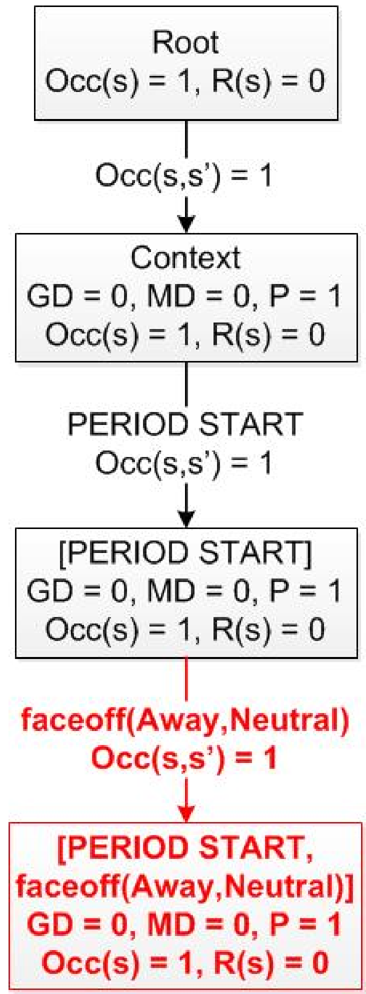

Figure 5.4: Construction: Step 4

Since penalty(Away,Neutral) is an end marker for play-by-play sequences, the current marker returns to the root node. The next play sequence is then processed, starting again from the root node. The occurrences for the root node are updated, and a new context node is created, as the away team penalty creates a manpower differential. The context information is taken from the faceoff(Home,Neutral) starting event of the next play-by-play sequence. An edge from the root node to the new context node is added and the occurrences are updated. A loopback edge is also added from the penalty(Away,Neutral) node of the previous sequence to the context node of the following sequence, to facilitate propagating the effects of the penalty. The occurrences of this loopback edge is also updated, and all these steps are all detailed in Figure 5.6. These steps are repeated for all play sequences, with the addition of adding a node for the win event after all play-by-play events have been processed for a single game. For each game, the process begins again from the root node.

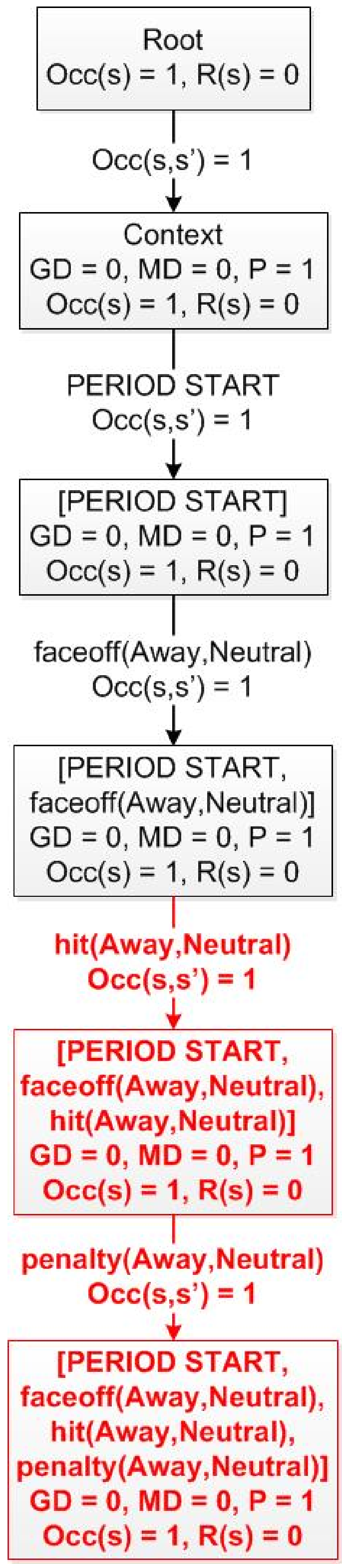

Figure 5.5: Construction: Step 5

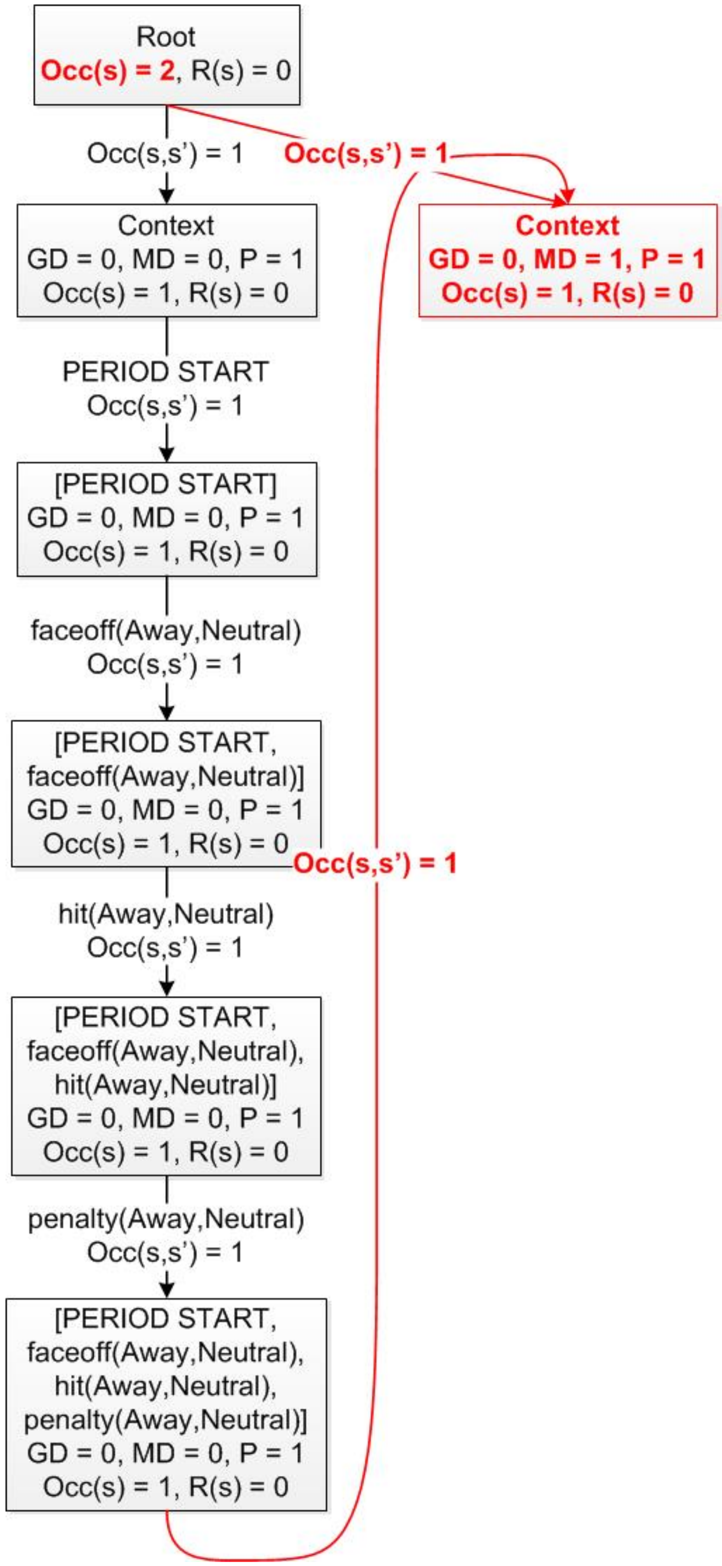

Figure 5.6: Construction: Step 6

Chapter 6

Value Iteration

The next step is to perform reinforcement learning on the Markov Game Model, which will yield valuations of player actions in different context states. We use a dynamic programming value iteration algorithm as the reinforcement learning technique to determine the value of each state in the Markov Game model. State valuation can be performed over many objective functions simultaneously, and is run iteratively until a convergence criterion is met or a maximum number of iterations is reached. We use a relative convergence criterion so the value iteration algorithm will terminate when Q-values for states are only being updated by a small amount, and iteration will continue if the updates are large. Nine objective functions shown in Table 6.1 are used in our work and determine which equations are used for the value iteration Q-function computations. Conditional Probabilities for goal scoring, receiving a penalty, and winning the match can also be derived by combining the probabilistic objectives for the home team and away team. From a hockey perspective, the motivation for learning Q-values for each state is that the Q-values quantify how close each team is to reaching an objective given the current state of gameplay. We can later compute the impact of a player's action by analyzing how performing the action impacts the team's chances of reaching an objective. This can easily be done by looking at the difference in Q-values between states, and will be discussed in Chapter 7.

6.1 Q-functions

The total reward in a state sequence is often computed using a discount factor. In ice hockey, discounting or averaging is not natural. For example, winning the game has the

Table 6.1: Reward Functions

Expected Wins
Probability of the Home Team Winning
Probability of the Away Team Winning
Expected Goals
Probability that the Home Team Scores the Next Goal
Probability that the Away Team Scores the Next Goal
Expected Penalties
Probability that the Home Team Receives the Next Penalty
Probability that the Away Team Receives the Next Penalty

same value for a team regardless of how many actions occurred previously. Goals may be more valuable if they are scored after fewer actions, but this should be an empirical finding from the analysis, not built into the definition of the Q-function. We use an undiscounted Q-function for value iteration in our work, following [29]. Different Q-functions are used depending on the objective being analyzed. For expected values of wins, goals, or penalties, Equation 6.1 is used as the value iteration function. $R(s)$ is initialized based on the event being analyzed as an objective. For example, if the objective is to find the expected goals, $R(s) = 1$ when s corresponds to a goal(Home,*) event, $R(s) = -1$ when s corresponds to a goal(Away,*) event, and $R(s) = 0$ for all other events and states. We use a similar initialization when processing wins and penalties as the objective. Note that $\frac{Occ(s, s')}{Occ(s)}$ forms the transition probability from state s to state s', but $\frac{1}{Occ(s)}$ is factored out to the front of the summation to speed computation time and prevent potential issues with numerical instability.

$$Q_{i+1}(s) = R(s) + \frac{1}{Occ(s)} \sum_{(s,s') \in E} (Occ(s, s') \times Q_i(s')) \tag{6.1}$$

For the probability of the next goal, or next penalty, Equation 6.2 is used as the value iteration function. Here, a can be one of goal or penalty, and T can be one of Home or Away. For example, if the objective find the probability of the next home goal, then a would be goal and T would be Home. All events of type of a are excluded from the first summation in Equation 6.2. This facilitates backing up the value 0 for the opposite T. For example, if

a is goal and T is Home, $a =$ goal(Away,*) is excluded from the summation, equivalent to backing up 0 for goal(Away,*).

$$Q_{i+1}(s) \;=\; \frac{1}{Occ(s)}((\sum_{\substack{(s,s')\in E\\ s'\neq a(*,*)}} \Big(Occ(s,s') \times Q_i(s')\Big)) + (\sum_{\substack{(s,s')\in E\\ s'=a(T,*)}} \Big(Occ(s,s') \times 1\Big))) \tag{6.2}$$

The probability of the home team or away team winning is similar to Equation 6.2 but also includes the reward $R(s) = 1$ for the a being analyzed and $R(s) = 0$ for all other states. This calculation is outlined in Equation 6.3. $R(s)$ can be included in the summation without the sum becoming greater than 1. This is because nodes denoting win events are always leaf nodes with no children by construction.

$$\begin{aligned} Q_{i+1}(s) = R(s) + \frac{1}{Occ(s)}((\sum_{\substack{(s,s')\in E\\ s'\neq a(*,*)}} \Big(Occ(s,s') \times Q_i(s')\Big)) \\ + (\sum_{\substack{(s,s')\in E\\ s'=a(T,*)}} \Big(Occ(s,s') \times 1\Big))) \end{aligned} \tag{6.3}$$

In a single-agent setting with a fixed policy, the value of a state is the expected reward for following the policy from the state. In the game-theoretic setting with two agents, we need to consider the difference in rewards. In a zero-sum game, the value of a state is the final result following optimal play. Intuitively, the value specifies which player has a better position in a state. Since the states in the Markov Game are modelling not optimal play, but actual play in an "on policy" setting, the difference in rewards is the natural counterpart

6.2 Dynamic Programming Algorithm

Recall that since states encode action histories, learning the expected value of states in the Markov Game model is equivalent to learning a Q-function (Section 4.3). In reinforcement learning terms, there is no difference between the value function V and the Q-function in the Markov Game model. Therefore, we apply standard value iteration over states [32] to learn a Q-function for the ice hockey Markov Game model. Algorithm 2 shows pseudocode for a dynamic programming algorithm for value iteration based on the Markov Game model. Separate Q-functions are computed for the Home team and for the Away team when the objective function is probabilistic. Since our model is in the “on policy” setting, there is a fixed policy for the other team. This means the other team can be treated as part of the environment, and reduce the Markov Game to two single-agent Markov Decision Processes for the purpose of value iteration. In our experiments, a relative convergence of 0.0001 is used as the convergence criterion, and 100,000 as the maximum number of steps. Value iteration converges in at most 10,304 iterations in all our experiments. Algorithm 2 uses Equation 6.1 as the Q-function, but can be substituted with other Q-functions mentioned in Section 6.1 to match the objective being analyzed.

Algorithm 2 Dynamic Programming for Value Iteration

Require: Markov Game model, convergence criterion c, maximum number of iterations M

$lastValue = 0$
$currentValue = 0$
$converged = false$
for $i = 1; i \leq M; i \leftarrow i + 1$ **do**
    for all states s in the Markov Game model **do**
        if $converged == false$ **then**
            $Q_{i+1}(s) = R(s) + \frac{1}{Occ(s)} \sum_{(s,s') \in E} (Occ(s, s') \times Q_i(s'))$
            $currentValue = currentValue + |Q_{i+1}(s)|$
        end if
    end for
    if $converged == false$ **then**
        if $\frac{currentValue - lastValue}{currentValue} < c$ **then**
            $converged = true$
        end if
    end if
    $lastValue = currentValue$
    $currentValue = 0$
end for

6.3 Example

To illustrate the dynamic programming algorithm for value iteration, a step-by-step example is given starting with the sample graph in Figure 6.1. Each node is shown with the node identification number, the action leading to the node, the occurrences of the node, and the current Q-value for the node. This example uses expected goals as the objective function being learned.

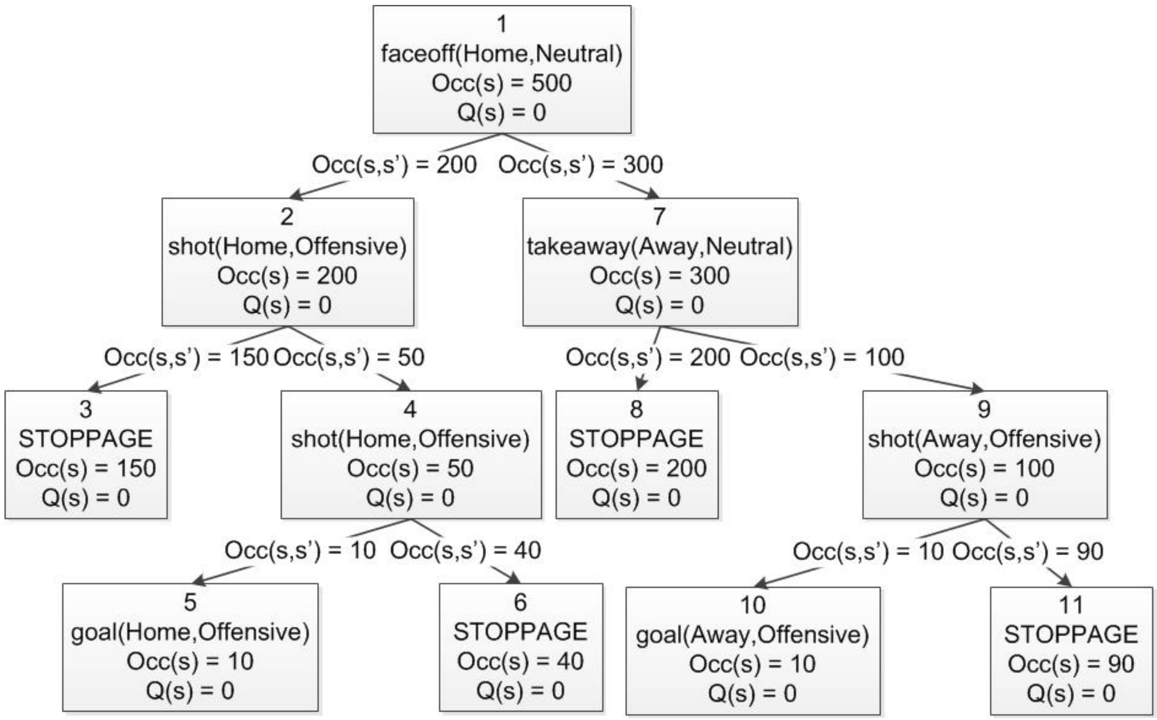

Figure 6.1: Value Iteration Example: Initial Graph

The first two steps of the dynamic programming algorithm for value iteration will set the values of nodes 5 and 10. Node 5 will be given Q-value $Q(5) = 1$, as $R(5) = 1$ for goal(Home,*) events and node 5 has no children. Node 10 on the other hand will be given Q-value $Q(10) = -1$, as $R(10) = -1$ for goal(Away,*) and node 10 has no children. The second step will update the values of nodes 4 and 9. Both nodes 4 and 9 correspond to shot events, and $R(s) = 0$ for all non-goal events in an expected goals model. The calculations for nodes 4 and 9 will include the non-zero values of nodes 5 and 10 respectively. Recall that the transition probability is calculated as $\frac{Occ(s,s')}{Occ(s)}$, where s is the parent node and s' is the child node. As such, the Q-value for node 4 becomes $Q(4) = \frac{10}{50} \times Q(5) + \frac{40}{50} \times Q(6)$ which is $Q(4) = 0.2$. The calculation for node 9 follows the same pattern. The updated Q-values are highlighted in Figure 6.2. The nodes whose children all had Q-values of 0 in the initial graph do not have their Q-value updated in this first step.

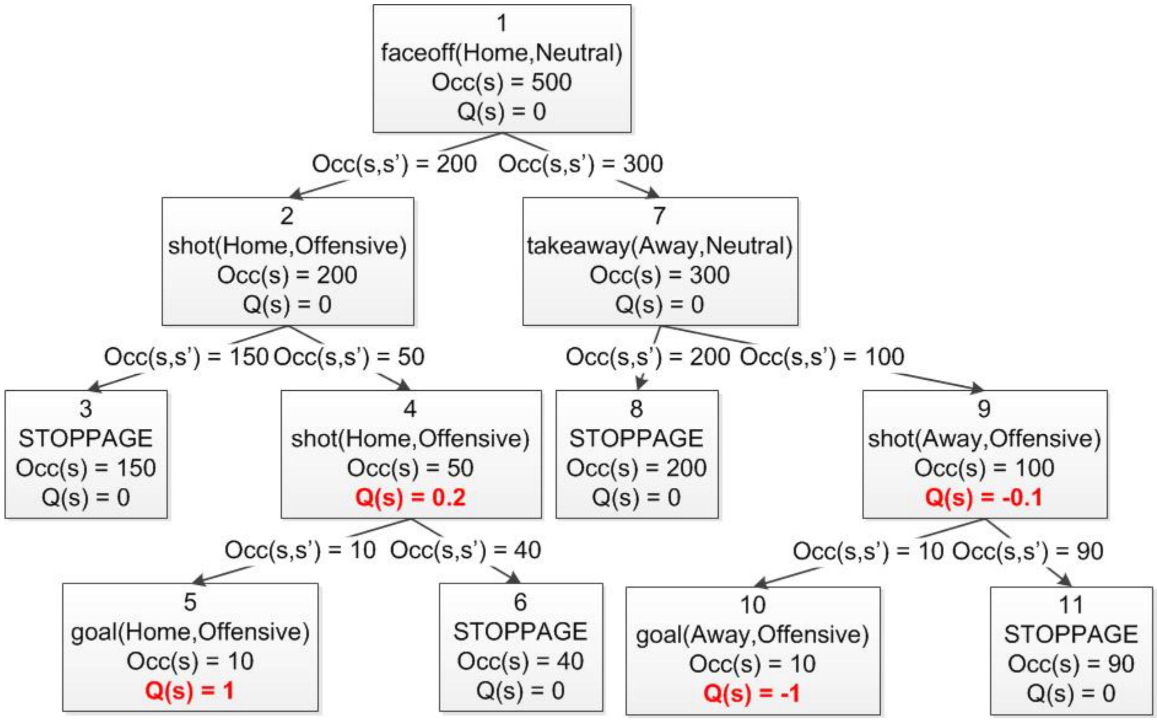

Figure 6.2: Value Iteration Example: First and Second Step

The third value iteration step will update the values of nodes 2 and 7, as shown in Figure 6.3. The Q-value for node 2 will become $Q(2) = \frac{50}{200} \times Q(4) + \frac{150}{200} \times Q(3)$ which is $Q(2) = 0.05$. The Q-value for node 7 is learned in a similar fashion.

The final step of the value iteration will back up the goal values all the way to the faceoff node (the root node in this example), as shown in Figure 6.4. Due to the choice of transition probabilities along the paths to the home and away goals in this example, the faceoff node has a net Q-value of $Q(1) = 0$.

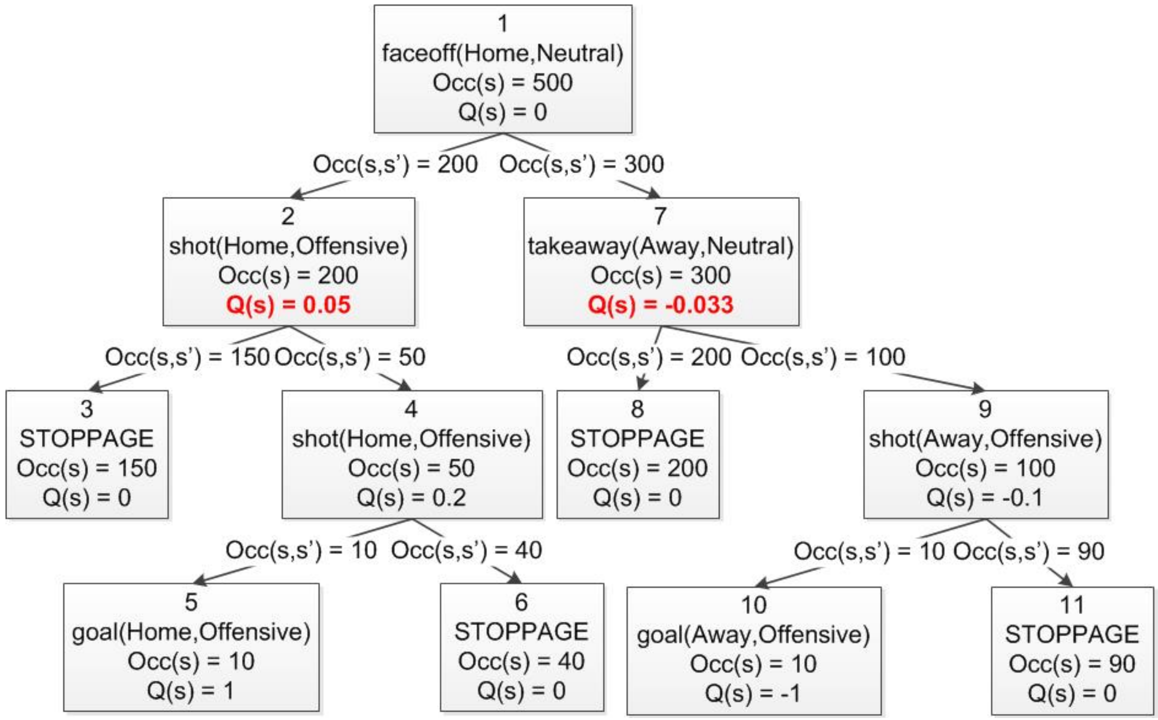

Figure 6.3: Value Iteration Example: Third Step

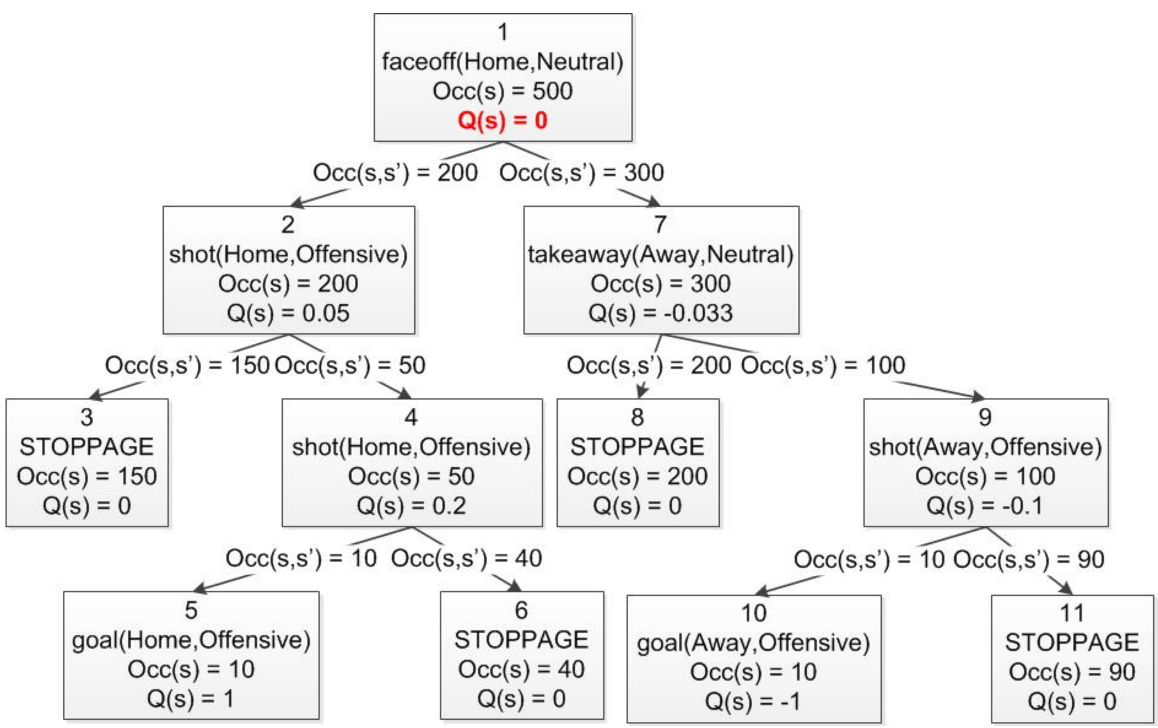

Figure 6.4: Value Iteration Example: Final Step

Chapter 7

Valuing Actions and Players

All player actions, with the exception of giveaways and some penalties, are volitional, meaning it is a clear choice made by the player. Therefore, evaluating a player's actions measures the effectiveness of the choices made by the player. There are multiple methods for calculating action and player values that can be derived from the Markov Game model. The Markov Game model is necessary for computing the impact of a player's actions, as it preserves the opposing objectives of both agents, the home and away team. We can then measure how a player's action impacts each team's probability or expectation of reaching the objective. We begin by discussing how action values are computed in Section 7.1. The choice of action valuation will fuel the calculation of player valuations in Section 7.2.

7.1 Valuing Actions

Due to the formulation of the Markov Game model and multiple Q-functions, there are four approaches for assigning values to actions. The first approach is shown in Equation 7.1 and applies to both probabilistic and expected value computations. This equation denotes the impact of an action a as the value of the single, unique state s reached by performing action a. The problem with this approach is that it only looks at the value of the state prior to performing action a, so it does not capture the information of how the game flow has changed as a result of performing action a. Another issue is for probabilistic cases, where only the information of one team will be included in this calculation. Performing an action may have an impact not only on the team performing the action, but on their opponent as well, and while an action may be slightly good for the team performing the action, it may have an even better impact on their opponent. This means the net impact of the action, computed as the difference in impact for both teams, was not beneficial for the team performing the action. This issue does not exist for expected value methods, as the net impact on both teams is included in the state value. For probabilistic models, we need to examine the net impact of actions between both teams.

$$impact(s, a) = Q(s) \tag{7.1}$$

The second approach is shown in Equation 7.2. To perform the computation for the away team, the negative of Equation 7.2 is used, as positive values will be relative to the team performing the action. Again, Equation 7.2 only applies to probabilistic objectives, and not expected value models. This equation captures the information of both teams, but there is still missing information of how the action changed the game flow. To solve this, the change in information from one state to the next needs to be captured.

$$impact(s, a_H) = Q_H(s) - Q_A(s) \tag{7.2}$$

The third approach is shown in Equation 7.3. Performing a particular action a can be mapped to a unique edge (s, s'). Thus, the change in Q-values between s' and s captures how performing the action changes the flow of the game. While this equation solves the missing game flow information prevalent in Equation 7.1 and Equation 7.2, there is still the issue of missing team information in probabilistic models. Again, expected value models do not have this issue, as they capture the information of both teams in the computation of Q-values. For probabilistic models, a hybrid approach between Equation 7.2 and Equation 7.3 to solve both issues.

$$impact(s, a) = Q(s') - Q(s) \tag{7.3}$$

The fourth and final approach is shown in Equation 7.4, which is a hybrid approach between Equation 7.2 and Equation 7.3. Once again, the equation is shown relative to the home team. To compute the impact value for the away team, the negative of Equation 7.4 is used. So, if a player is playing for the home team when he performs an action, we apply the impact of the action to the player as it is computed in Equation 7.4. If the player is playing for the away team when he performs an action, we apply the negative of Equation 7.4 to the player. This method is used only for probabilistic models, and we use Equation 7.4 for the impact values for actions reported in Chapter 8 and Chapter 9. The net change in Q-value for both teams is captured in this equation, as well as the difference in impact between both teams, which captures the true change in game flow. We will use Equation 7.4 for all reported action and player values in our results.

$$impact(s, a_H) = \big(Q_H(s') - Q_H(s)\big) - \big(Q_A(s') - Q_A(s)\big) \tag{7.4}$$

In order to evaluate players over games and over seasons, the impact values from state to state are needed, as they give a fine-grained analysis of how the player impacted the game flow. For each player, we sum the player's action impact values over a game to get the player's net game impact. This quantifies if the player had a net positive contribution to the objective during the game, or a net negative contribution to the objective. Summing these net game impacts over a season will give a player's net season impact, and is useful for evaluating the player's performance.

7.2 Valuing Players

It is intuitive that action values, computed as impact(s,a_H) from Equation 7.4, must be applied to players to measure player contributions. The problem is then in determining how the action impact values are to be assigned to players. For valuing players, there are three approaches that can be chosen from:

1 Player Apply the action impact value only to the player performing the action.

2 Players Apply the action impact value to the player performing the action, and the negative of the action impact value to the opponent who may be involved in the action. It may be the case that only one player is involved in an action and no opponent is involved, in which case only the player performing the action has the action impact value applied.

All Players Apply the action impact value to the player performing the action and all his teammates present on the ice, and the negative of the action impact value to all opponents on the ice.

Other options for computing player valuations are to standardize player impact scores with respect to the player's team, as well as to the number of minutes or games played by a player, but this is left as future work. The problem with the third approach is that it is difficult to differentiate the contribution of the player from his teammates. Consider that coaches pick four forward lines each containing three forwards, and three defensive lines each containing two defensemen. These players will play together throughout the duration of the game, with some adjustments during special teams situations. These lines often do not differ much throughout the entire hockey season. As such, it becomes difficult to differentiate the contributions between players who play together quite often [14].

Only the first approach is used my player valuations, although comparing the first approach to the second approach would be an interesting study. Since the first approach has been chosen, player impact scores for each season need to be generated. I generate player impact scores for each season in the following steps:

1. For each action in each game, apply the action value to the player performing the action.
2. Sum a player's action values over each game.
3. Sum a player's game action values over each season.

7.2.1 Example

A step-by-step example of player valuation is given, using Sidney Crosby as an example. The action values we use in this example are for the probability of the next goal as the objective function. First, the impact values corresponding to the actions taken by Sidney Crosby during a game are joined together into a single table, as in Figure 7.1. The fields FromNodeId and ToNodeId denote the states s and s' respectively used in the edge (s, s') corresponding to action a.

Next, these impact values are summed over each game to give a net game impact score for Sidney Crosby. This is shown in Figure 7.2. It is clear that even top-tier players such as Sidney Crosby can have games with a net positive performance, as well as games with a net negative performance. On average, he has a positive contribution to his team, generating 0.35 goals per game.

GameId	EventNumber	Player	FromNodeId	ToNodeId	Action	ActionValue
2014020509	4	Sidney Crosby	7931	24421	SHOT	0.0169
2014020509	24	Sidney Crosby	62	72	FACEOFF	0.0088
2014020509	25	Sidney Crosby	72	3625	HIT	0.0021
2014020509	47	Sidney Crosby	62	590	FACEOFF	0.0168
2014020509	53	Sidney Crosby	1325010	1325011	SHOT	0.0000
2014020509	68	Sidney Crosby	1325022	1325023	BLOCKED SHOT	0.0000
2014020509	74	Sidney Crosby	62	939	FACEOFF	-0.0052
2014020509	80	Sidney Crosby	13085	313965	FACEOFF	-0.1141
2014020509	102	Sidney Crosby	391	392	FACEOFF	0.0129
2014020509	114	Sidney Crosby	391	3749	FACEOFF	0.0116
2014020509	198	Sidney Crosby	84771	84772	FACEOFF	-0.0127
2014020509	204	Sidney Crosby	468	12522	SHOT	0.1422
2014020509	222	Sidney Crosby	1325113	1325114	SHOT	0.0000
2014020509	251	Sidney Crosby	187	219	FACEOFF	0.0024
2014020509	263	Sidney Crosby	178	238	FACEOFF	0.0017
2014020509	265	Sidney Crosby	8319	167216	SHOT	0.0071
2014020509	294	Sidney Crosby	258	259	FACEOFF	0.0299
2014020499	11	Sidney Crosby	2	15	FACEOFF	-0.0009
2014020499	59	Sidney Crosby	2273	2876	FACEOFF	0.0238
2014020499	65	Sidney Crosby	2273	2876	FACEOFF	0.0238
2014020499	81	Sidney Crosby	1198	2251	FACEOFF	0.0153
2014020499	113	Sidney Crosby	115419	1323584	MISSED SHOT	-0.0653
2014020499	168	Sidney Crosby	1028	1815	FACEOFF	0.0300
2014020499	190	Sidney Crosby	2302	25553	FACEOFF	-0.0455

Figure 7.1: Sidney Crosby: Individual Action Values

Finally, the net game impact values are summed over an entire season to generate a season impact score for Sidney Crosby. Games are grouped by the season and season type, that is, regular season and playoff games. As the data is stored in a relational database, this is easily performed with simple SQL queries. The results of this are shown in Figure 7.3. We observe that Sidney Crosby has consistently had a high impact on goal scoring across seasons, which explains why he is one of the highest paid players in the NHL.

The impact results reported in Chapter 9 will be the net season impact values for players during the regular season. We do not report results for the playoffs, as these games form a smaller portion of our dataset.

GameId	Player	Number of Actions	Net Game Impact
2014020509	Sidney Crosby	17	0.1186
2014020499	Sidney Crosby	20	-0.5956
2014020485	Sidney Crosby	26	2.0481
2014020468	Sidney Crosby	16	0.9150
2014020406	Sidney Crosby	16	-0.5101
2014020390	Sidney Crosby	13	0.5967
2014020376	Sidney Crosby	11	-0.2541
2014020364	Sidney Crosby	19	-0.3554
2014020349	Sidney Crosby	18	0.0473
2014020339	Sidney Crosby	19	0.6793
2014020328	Sidney Crosby	25	0.3361
2014020312	Sidney Crosby	20	-0.1775
2014020299	Sidney Crosby	11	0.1986
2014020292	Sidney Crosby	17	-0.1106
2014020272	Sidney Crosby	15	0.7753
2014020255	Sidney Crosby	22	-0.0656
2014020244	Sidney Crosby	26	0.0715
2014020222	Sidney Crosby	13	0.3506
2014020203	Sidney Crosby	6	1.0423
2014020192	Sidney Crosby	18	0.7754
2014020179	Sidney Crosby	15	0.1567
2014020159	Sidney Crosby	14	0.5798
2014020140	Sidney Crosby	16	0.7301

Figure 7.2: Sidney Crosby: Net Game Impact

Season	Season Type	Player	Net Season Impact
2014-2015	Regular Season	Sidney Crosby	10.43
2013-2014	Regular Season	Sidney Crosby	24.23
2013-2014	Playoffs	Sidney Crosby	2.24
2012-2013	Regular Season	Sidney Crosby	19.47
2012-2013	Playoffs	Sidney Crosby	3.21
2011-2012	Regular Season	Sidney Crosby	6.75
2011-2012	Playoffs	Sidney Crosby	0.03
2010-2011	Regular Season	Sidney Crosby	14.92
2009-2010	Regular Season	Sidney Crosby	28.12
2009-2010	Playoffs	Sidney Crosby	2.59
2008-2009	Regular Season	Sidney Crosby	33.44
2008-2009	Playoffs	Sidney Crosby	8.97
2007-2008	Regular Season	Sidney Crosby	20.88
2007-2008	Playoffs	Sidney Crosby	5.33

Figure 7.3: Sidney Crosby: Net Season Impact

Chapter 8

Hardware and Evaluation

The hardware used in the data collection, model construction, and value iteration computation are summarized in Section 8.1. Next, we evaluate our model with two lesion studies. In the lesion studies, we remove different features from our model to examine the benefit of retaining or removing the features. The first lesion study in Section 8.2 removes context features and examines the entropy of the state transition graph. We perform this study to justify using the full set of context features. The second lesion study in Section 8.3 examines propagation effects by adding specific loopback edges from the basic AD-Tree structure. We perform this study to justify how we use loopback edges to capture medium-term effects of actions. Finally, in Section 8.4 we evaluate our computation of action impact values.

8.1 Hardware

NHL play-by-play data was obtained from http://www.nhl.com using the Selenium WebDriver with Python 2.7.6 on a 64-bit Ubuntu 14.0.4 LTS Virtual Machine with 4.8GB RAM and an Intel Core i7-2670QM CPU @ 2.20GHz $\times$ 8. Markov Game Model construction and value iteration computation was performed using Java Version 8 Update 25 on 64-bit Windows 7 with 12GB RAM and an Intel Core i7-2670QM CPU @ 2.20GHz $\times$ 8. The state transition graph is stored in a MySQL 5.6.13 database using two tables for nodes and edges.

8.2 Lesion Study: Feature Space

To motivate the use of all context features, that is, GD, MD, and P, a lesion study is performed by adding or removing different parts of the context in the full state transition graph. The sizes of each graph are shown in Table 8.1. As expected, adding more context features increases both the number of nodes and the number of edges in the graph. Transforming the state transition graph from including no context features to including GD, MD, and P as context features increases the number of nodes by 45.9%.

Table 8.1: Size of State Transition Graphs with Different Features

Graph Type	**Number of Nodes**	**Number of Edges**
No Context	909,010	1,134,364
Only *MD*	1,009,536	1,267,020
Only *P*	1,019,702	1,272,599
Only *GD*	1,089,324	1,359,503
MD and *P*	1,125,678	1,412,071
GD and *MD*	1,200,924	1,506,962
GD and *P*	1,208,618	1,508,240
Full Context	1,325,809	1,662,504

In order to justify this large increase in the state space, we must examine how adding context features increase the information in the model. This is done by computing the entropy of each model. Entropy is computed for each state as in Equation 8.1.

$$H[s] = -\sum_{T} p_T(s) \ln p_T(s) \tag{8.1}$$

We compute the entropy with respect to the conditional probability of the next goal for the home and away teams. The model entropy for state transition graphs using a different set of features are shown in Table 8.2. Again, we used state transition graphs with no context feature, 1 context feature, 2 context features, and the full set of context features. As expected, the full state transition graph containing all context features has the lowest entropy, as it contains the most information. There is clearly an information gain when moving from no context to any level of context, underlining the importance of context when analyzing player actions in ice hockey. Table 8.2 also clearly shows the benefit of using all context features to evaluate actions, as this model has the lowest uncertainty. Manpower differential is the context feature having the greatest impact on model uncertainty.

Table 8.2: Entropy of State Transition Graphs with Different Features

Graph Type	**Entropy**
No Context	0.9781
Only *P*	0.9756
Only *GD*	0.9739
Only *MD*	0.9727
GD and *P*	0.9706
MD and *P*	0.9699
GD and *MD*	0.9681
Full	0.9644

8.3 Lesion Study: Effects of Propagation

The transition graph construction algorithm facilitates changing the possible state transitions. We modify the state transitions in our experiments to study how different propagation models affect the impact of actions. To analyze this effect, we examine the Probability of the Next Goal Scored. Specifically, we consider three different transitions graphs of increasing density, and their sizes are shown in Table 8.3.

Local Transitions Only State transitions occur only within a play sequence, not across play sequences.

Penalty Transitions State transitions occur from penalty leaf nodes to successor context nodes, in addition to state transitions in the local state transition graph.

Full Transition Graph Includes loopback edges from all leaf nodes to context nodes of the following play sequences, as defined in Section 4.2, in addition to the state transitions in the penalty graph.

As we are only modifying the state transitions, the states are preserved and the number of graph nodes is equal across all three state transition graphs. The large change in the number of edges from the state transition graph with penalty transitions to the full state transition graph is expected, as sequences ending in penalties only make up 15.2% of our

Table 8.3: Size of State Transition Graphs

	Local	**Penalty**	**Full**
Number of Nodes	1,325,809	1,325,809	1,325,809
Number of Edges	1,325,808	1,382,780	1,662,504

dataset. Other loopback transitions are derived from goals, stoppages, and other sequence end markers. Action impact changes value depending on the state transition graph. With the local transition graph, value iteration computes the impact of an action on the current play sequence only. Thus the Q-value differential for context states, with the initial empty play sequence, can be obtained from Table 4.2. The average differences in action values, as well as the standard deviation of the differences, are shown in Table 8.4. While the aggregate effects provide insight into medium-term hockey dynamics, they do not reflect the considerable context dependence shown by the standard deviations of the impact differentials. We observe the standard deviation is greater than the average change, showing propagating effects of actions creates a wide range of action values. The penalty transition graph propagates to the next sequence the effect of penalties only. Propagating the effect of penalties changes most the estimation of the impact of penalties. This change reflects that receiving a penalty lowers the chances of scoring the next goal. Less obviously, winning a faceoff in the offensive zone has a relatively high positive indirect impact on scoring the next goal, via increasing the probability of a penalty against the opposing team. The effect of winning an offensive zone faceoff can also be seen in Figure 8.2. Comparing the full transition graph with penalty propagation only, it is still observed that the strongest average impact change is for penalties. This shows that penalties have ripple effects on goals via events other than penalties. Clearly, including the full state transitions provides further insight into the value of actions compared to analyzing actions in a local sequence or as standalone actions.

8.4 Action Impact Values

The main quantity considered is the **impact** of an action as a function of context (= Markov state). We use Equation 7.4 to calculate the impact of home team actions, and the negative of Equation 7.4 to calculate the impact of actions by the away team. In a zero-sum

Table 8.4: Difference In Action Impact Values for Next Goal Scored, Across Transition Graphs

	Full vs. Penalty		**Penalty vs. Local**	
	Average Change	**Standard Deviation**	**Average Change**	**Standard Deviation**
Blocked Shot	0.0001	0.0210	-0.0003	0.0126
Faceoff (Defensive)	-0.0030	0.0455	-0.0018	0.0225
Faceoff (Neutral)	0.0013	0.0464	0.0006	0.0203
Faceoff (Offensive)	0.0038	0.0432	0.0024	0.0260
Giveaway	-0.0003	0.0245	-0.0001	0.0142
Hit	0.0000	0.0194	-0.0001	0.0126
Missed Shot	-0.0001	0.0218	0.0003	0.0130
Penalty	**-0.0190**	0.0278	**-0.0235**	0.0337
Shot	0.0002	0.0191	0.0002	0.0103
Takeaway	0.0006	0.0245	0.0003	0.0146

game, the state value is usually defined as the final result following optimal play [22]. Intuitively, the value specifies which player has a better position in a state. Since the Markov Game model presented is not modelling optimal play, but actual play in an "on policy" setting, the expected difference in rewards is the natural counterpart. The impact quantity measures how performing an action in a state affects the expected reward difference. Figure 8.1 shows a boxplot for the action impact values as they range over different contexts, i.e., states in the Markov Game model. (Boxplots produced with MATLAB R2014a.) While the Q-values are based on the frequency of states, all states are weighted equally in discussing the properties of the Q-function. The boxplot does not include Q-values for states whose frequency is below 5%. It is clear from Figure 8.1 that *depending on the context and event history, the value of an action can vary greatly.* The context-dependence is observed for both scoring goals and receiving penalties.

8.4.1 Impact on Scoring the Next Goal

All actions, with the exception of faceoffs won in the offensive zone, have at least one state where the action has a positive impact, and another state with a negative impact. We use our impact rating for evaluating player performance rather than generating strategies, but by analyzing action values, players and coaches can learn when to perform certain actions in particular contexts and when to avoid them. A positive impact value means a player's action leads to the player's team being more likely to score the next goal. A negative impact value means a player's actions causes the opposing team to be more likely to score the next goal. Players performing actions with a positive impact improve their team's chances of scoring the next goal. Conversely, players performing actions with a negative impact improve their opponent's chances of scoring the next goal. Examples of context-dependence include the following.

(1) Blocking the first shot on net when killing a penalty is bad $(impact = -0.0864)$, but blocking the second shot on net is very good $(impact = 0.1399)$.

(2) Receiving a penalty when on the powerplay is very bad $(impact = -0.1789)$, but if a player on the penalty kill can goad their opponent into an offsetting penalty, it is slightly good $(impact = 0.0474)$. These two impact values for penalties are not symmetric because receiving a penalty is bad in general, as it decreases the number of players on the bench by 1 man. This means a penalized player's teammates have to play longer on the penalized player's behalf, and may become tired.

(3) During overtime, if the opposing team wins the faceoff in the neutral zone and a player takes the puck away from his opponent in the neutral zone, it has a very high impact on goal scoring $(impact = 0.2919)$. If a player's team is already up by 3 goals, takeaways they perform in their own zone can be very bad $(impact = -0.2544)$.

(4) If a player's team wins the faceoff in the neutral zone, but then gives the puck away in their own zone, it is very bad $(impact = -0.1874)$. If a player gives the puck away in their zone in the first period after their opponent has just taken a shot on net, it can often lead to a positive impact $(impact = 0.1184)$.

The THoR player ratings compute the impact of actions based on goals that immediately follow the action ([13, 26, 25]; see Section 2). The values given for each action in [13] are displayed as an asterisk in Figure 8.1. The THoR values agree with our median impact values in terms of whether an action generally has positive or negative impact. For example, penalties are known to generally be good for the opposing team, and shots are good for the shooter's team. THoR values are close to our median Markov model values in 6 out of 10 cases. The exceptions are blocked shots, faceoffs won in the offensive zone, penalties, and shots. [13] makes an adjustment to blocked shots and shots based on averages, which may cause these two actions to be greatly overvalued. This comparison suggests that THoR aggregates action values over many contexts that the Markov Game explicitly models.

Another comparison of context-aware action values versus fixed action values is to quantify the information lost by ignoring context in terms of the entropy of the Next Goal probabilities. The context-unaware Next Goal probability for an action event, is the marginal probability obtained from action-state probabilities by summing out all states where the action is taken. For all action events, this marginal probability of the next goal for the away team is 47% and 48%. This leads to an average context-unaware entropy of 0.9741 with standard deviation of only 0.0012. The average of the context-aware entropies is 0.9582; but these entropies show considerable variance, ranging smoothly from 0 to 1, with a large standard deviation of 0.1482.

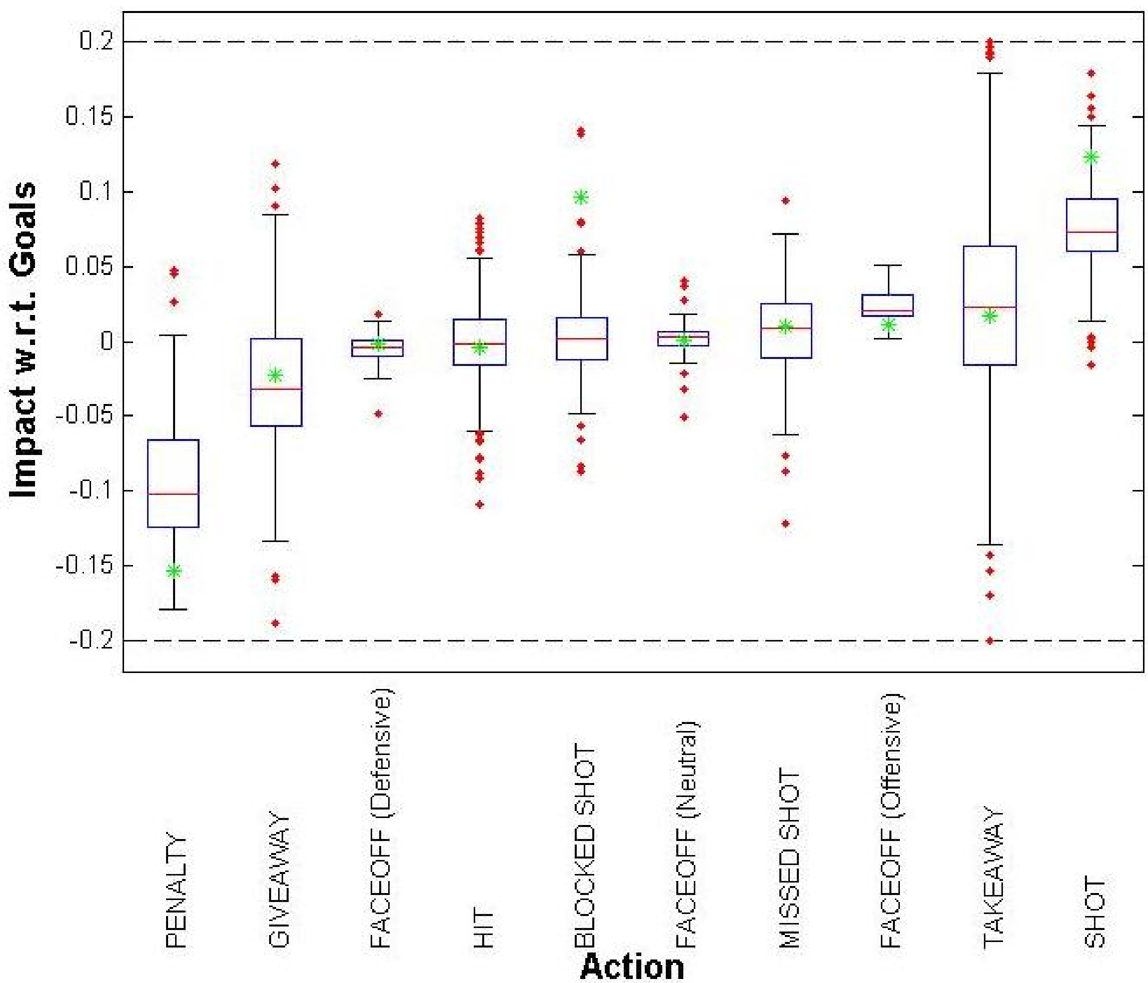

Figure 8.1: Impact on the probability of scoring the next goal. Higher numbers are *better* for the team that performs the action. Green asterisks represent the action values used in THoR [25].

8.4.2 Impact on Receiving Penalties

The range of action values with the probability of the next penalty as the objective function is shown in Figure 8.2. A positive impact value means a player's action leads to the player's team being more likely to receive the next penalty. A negative impact value means a player's actions causes the opposing team to be more likely to receive the next penalty. Players performing actions with a positive impact on penalties hurt their team, as they become more likely to receive the next penalty. Conversely, players performing actions with a negative impact improve their team's chance of gaining a powerplay. Again, it is observed that the impact of actions on penalties varies greatly with context. Winning faceoffs in the Offensive Zone and takeaways tend to cause the opponent to receive a penalty. Giveaways and goals tend to be followed by a penalty for the player's team. This finding is consistent with the observation that there are more penalties for teams who are leading their opponent with respect to goals [24]. Similarly, teams who are trailing behind their opponent with respect to goals tend to receive less penalties. A possible explanation is referees are reluctant to penalize a trailing team, but more likely to penalize a leading team, suggesting a levelling bias in penalty calling.

8.4.3 Impact on Winning

Action values with respect to impact on winning are observed in Figure 8.3. It is clear that penalties have a negative impact on winning, and it is observed in Table 4.2 that penalties affect goal scoring rates, which in turn affects a team's probability of winning. Shots on net and goals have a positive impact on winning, showing quantitatively that goals win games. Takeaways tend to leads to goals for a team, and giveaways lead to goals against a team, so their respective positive and negative values are valid. All events, with the exception of goals, have both positive and negative occurrences. It is interesting to note that there are contexts when even taking a shot on net can increase a team's probability of winning by up to 10%. In these cases, the shot likely leads to a goal.

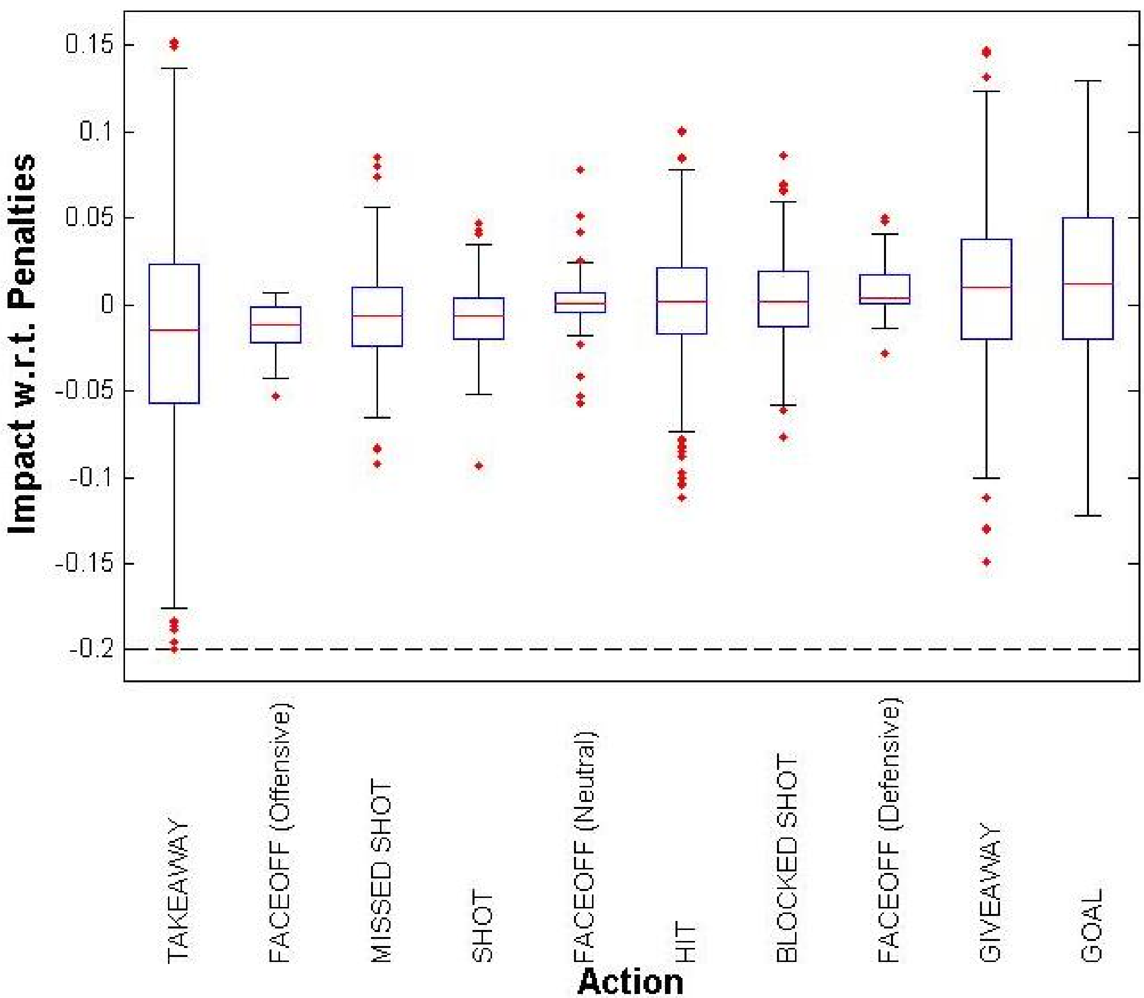

Figure 8.2: Impact on the probability of receiving the next penalty. Higher numbers are *worse* for the team that performs the action.

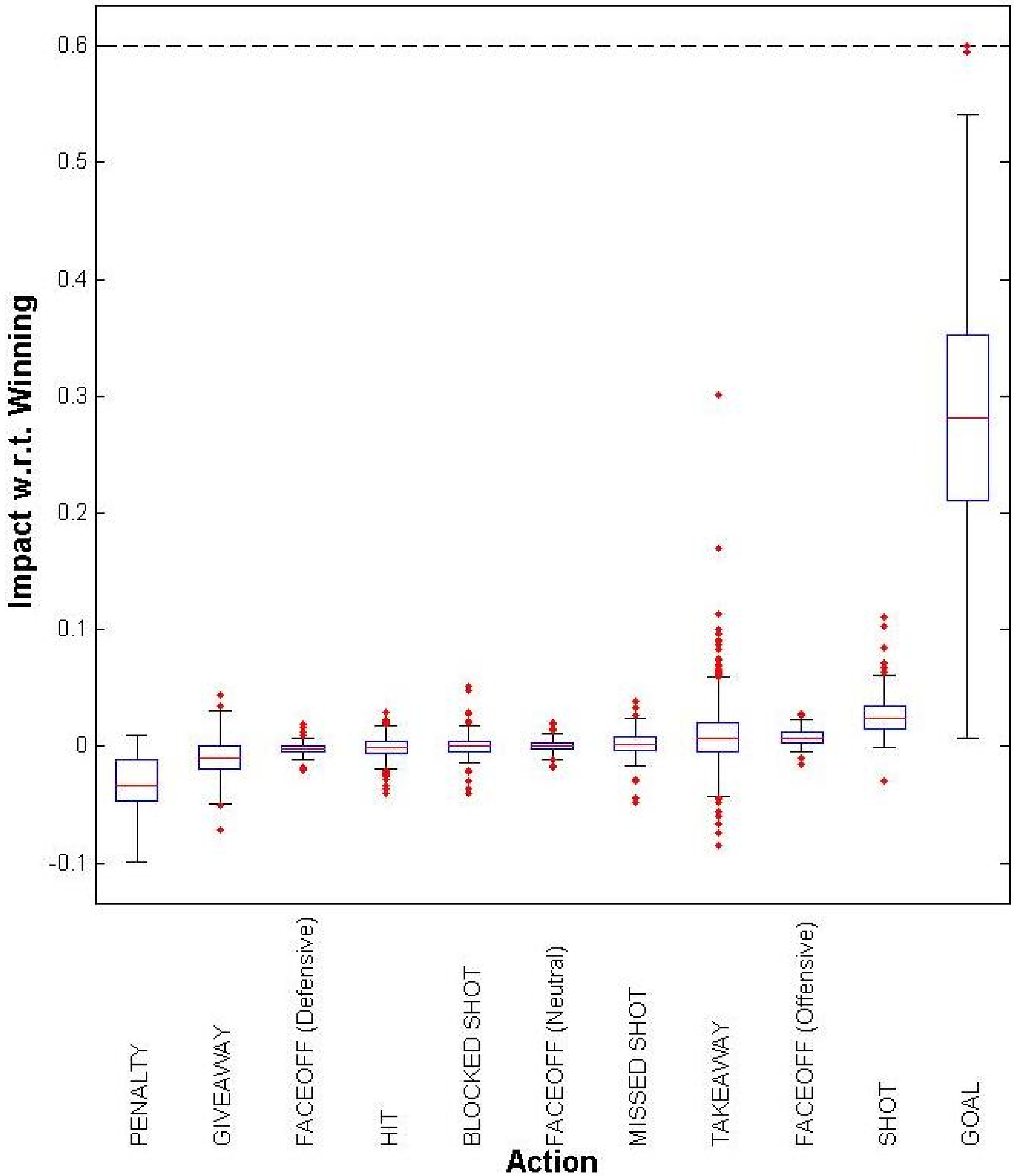

Figure 8.3: Impact on the probability of winning. Higher numbers are *better* for the team that performs the action.

Chapter 9

Results

Computing the net game impact will be useful for rapid post-game analysis, and will be beneficial for coaches picking which players to dress for future matches. Our value iteration algorithm is necessary for this, as action values cannot be computed directly from entire game statistics. General managers can also examine net season impact for players to determine monetary valuations of players. For our results, we rank players according to different objective functions. Player rankings with respect to goals are covered in Section 9.1. For penalties, players are ranked in Section 9.2. Finally, we rank players according to their win impact in Section 9.3. We also examine special teams rankings as a subset of win impact ratings in Section 9.4. We give a comparison of our win impact scores with current advanced statistics player valuations in Section 9.5.

9.1 Player Rankings: Goals

We compare impact on the Probability of the Next Goal Scored with three other player ranking metrics and statistics: points earned, salary, and +/-. We apply the impact value of each action to players as they perform that action. A player's impact scores are first aggregated over a match in a season, then over all matches in a single season to produce a season impact score. Player impact scores with respect to goals for the 2013-2014 season are shown in Table 9.1. Player impact scores with respect to goals for other seasons are shown in Appendix A. When examining player impact scores across seasons, we notice that player impacts change across seasons, suggesting player performance can improve or diminish across seasons. This is a counterargument to [20], whose AGV metric for players was strongly correlated across seasons. The average player impact score with respect to goals for the 2013-2014 season was 5.33, meaning the average player contributes to 5.33 goals. Players with high impact on goal tend to have high salaries as they help their team to produce more goals. However, the magnitude of player salaries vary greatly regardless of the magnitude of their impact score, and the median salary in the NHL is $2.4 million USD. Since these players have a high impact on goals, they also tend to have a positive +/- rating. Jason Spezza is an anomaly, as he has the highest impact score but a very negative +/- score. This is due to his team performing poorly overall in the 2013-2014 season, and the team overall had a goal differential of -29, one of the lowest goal differentials that season. This example shows that impact scores distinguish a player who generally performs useful actions but happens to be on a poor team. The negative +/- score also hides Jason Spezza's contribution to goal scoring, whereas our impact metric clearly shows his contribution to goal scoring. Ryan Johansen is also an anomaly in regards to his salary, which is only $810,000 USD and is a much lower salary than the salaries of the other players in the same ranking. This shows the impact score is useful for general managers who are evaluating players and looking for bargain players with high impact. We also observe that Sidney Crosby has double the salary of players with similar ranking, suggesting that he is overpaid for how many goals he generates. It is interesting to note the lack of defensemen in the top-20 players. This could be due to low offensive output by defensemen, and more offensive events are recorded by forwards than by defensemen.

Figure 9.1 shows that next goal impact correlates well with points earned. A point is earned for each goal or assist by a player. Since assists are not recorded as events in the NHL play-by-play event logs used in our Markov Game model, the correlation suggests including events other than goals in our Markov Game model helps to capture some of the assist information.

9.2 Player Rankings: Penalties

Table 9.2 displays player impact with respect to Next Penalty Received. High impact numbers indicate a tendency to cause penalties for a player's own team, or prevent penalties for the opponent. The Q-function impact numbers with respect to penalties are compared to Penalties in Minutes (PIM), +/-, and salary. Players with high Q-function numbers

Table 9.1: 2013-2014 Top-20 Player Impacts For Goals

Name	**Position**	**Goal Impact**	**Points**	**+/-**	**Salary**
Jason Spezza	C	29.64	66	-26	$5,000,000
Jonathan Toews	C	28.75	67	25	$6,500,000
Joe Pavelski	C	27.20	79	23	$4,000,000
Marian Hossa	RW	26.12	57	26	$7,900,000
Patrick Sharp	LW	24.43	77	12	$6,500,000
Sidney Crosby	C	24.23	104	18	$12,000,000
Claude Giroux	C	23.89	86	7	$5,000,000
Tyler Seguin	C	23.89	84	16	$4,500,000
Max Pacioretty	LW	22.54	60	8	$4,000,000
Patrice Bergeron	C	22.26	62	38	$4,550,000
Jamie Benn	LW	22.08	79	21	$5,000,000
Ryan O'Reilly	C	22.06	64	-1	$6,500,000
Ryan Johansen	C	21.96	63	4	$810,000
Valterri Filppula	C	21.70	58	5	$4,000,000
Patrick Marleau	C	20.92	70	0	$6,900,000
Matt Duchene	C	20.67	70	8	$3,750,000
John Tavares	C	20.45	66	-6	$5,000,000
Zach Parise	LW	19.93	56	10	$12,000,000
David Backes	C	19.25	57	14	$4,750,000
Derek Stepan	C	19.11	57	12	$2,300,000

have high penalty minutes as we would expect. They also have low +/-, which shows the importance of penalties for scoring chances. Their salaries tend to be lower. There are however notable exceptions, such as Dion Phaneuf and Dustin Byfuglien, who draw high salaries although their actions have a strong tendency to incur penalties. Dion Phaneuf has however been a regular at NHL All-Star matches, and is a highly valued player, suggesting he may offset his tendencies for penalties with effective play-making. This is verified by observing his impact on goal scoring and winning, where he generates 2.95 goals and contributes to 2.95 wins.

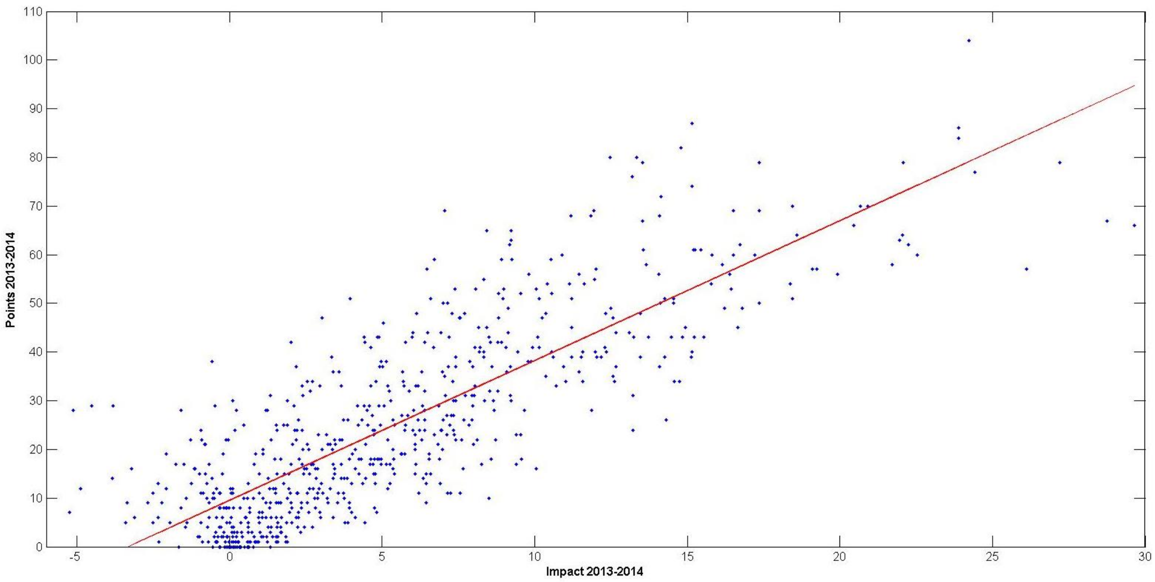

Figure 9.1: 2013-2014 Player Goal Impact Vs. Season Points

9.3 Player Rankings: Wins

Finally, the third objective of impact on winning is observed. The top-25 player impact scores with respect to winning in the 2013-2014 regular season are shown in Table 9.3. As expected, these players have above average salaries, with Sidney Crosby and Zach Parise having the highest salaries of $12,000,000 USD. There are some notable exceptions such as Ryan Johansen and Sean Monahan with below average salaries of $810,000 USD and $925,000 USD respectively. These two players are playing well above their value and are a bargain for the teams that own them. All players in this table have a high number of goals, points, shots, and takeaways. We do observe that +/- also varies between positive and negative values, even though these players have high contributions to winning. This suggests that applying +/- values to all players on the ice obscures their actual contribution, and may be incorrectly applying a negative value to those actually making a positive contribution. Those with a negative +/- rating are playing on teams who perform poorly overall, and while the actions of these players have a large positive impact on winning, they may have decreasing +/- due to their teammates. Results for the top-25 and bottom-25 players in other seasons are recorded in Appendix C.

Table 9.2: 2013-2014 Top-20 Player Impacts For Penalties

Name	Position	Penalty Impact	PIM	+/-	Salary
Chris Neil	RW	62.58	211	-10	$2,100,000
Antoine Roussel	LW	54.26	209	-1	$625,000
Radko Gudas	D	53.34	152	2	$575,000
Dion Phaneuf	D	52.52	144	2	$5,500,000
Zac Rinaldo	C	48.65	153	-13	$750,000
Rich Clune	LW	47.08	166	-7	$525,000
Tom Sestito	LW	46.34	213	-14	$650,000
Tom Wilson	RW	46.12	151	1	$925,000
Zack Smith	C	44.55	111	-9	$1,500,000
David Perron	LW	42.49	90	-16	$3,500,000
Steve Downie	RW	41.28	106	1	$2,750,000
Dustin Byfuglien	RW	40.88	86	-20	$5,750,000
P.K. Subban	D	40.36	81	-4	$3,750,000
Mark Stuart	D	38.98	101	11	$1,800,000
Ryan Garbutt	LW	38.89	106	10	$600,000
Kevin Bieksa	D	38.76	104	-8	$5,000,000
David Backes	C	38.54	119	14	$4,750,000
Matt Carkner	D	38.05	149	-10	$1,500,000
Wayne Simmonds	RW	37.49	106	-4	$2,800,000
Kyle Quincey	D	35.99	88	-4	$4,000,000

9.4 Player Rankings: Special Teams

Advanced statistics in special teams situations were first covered in [15], and due the context-inclusive nature of my Markov Game model, it is easy to also examine player performance in special teams contexts. Special teams situations are gameplay instances where there is a manpower differential between teams. Powerplay situations are where one team has a manpower advantage over the other, and penalty kill or shorthanded situations are where one team has a manpower disadvantage to the other. For coaches, it is crucial to put players on the ice that will maximize their team's chance of winning while on the powerplay, and prevent the other team from winning while shorthanded. Coaches will typically pick a short list of players from their game roster to perform in special teams situations. As such, modelling player impact with respect to wins during special teams

Table 9.3: 2013-2014 Top-25 Player Impacts For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Joe Pavelski	C	10.77	41	79	225	56	23	$4,000,000
Jonathan Toews	C	10.60	27	67	192	51	25	$6,500,000
Jason Spezza	C	10.24	23	66	223	47	-26	$5,000,000
Marian Hossa	RW	9.56	29	57	238	73	26	$7,900,000
Sidney Crosby	C	9.49	36	104	259	41	18	$12,000,000
John Tavares	C	9.48	24	66	188	55	-6	$5,000,000
Claude Giroux	C	9.16	28	86	223	43	7	$5,000,000
Valtteri Filppula	C	8.99	25	58	131	53	5	$4,000,000
Nicklas Backstrom	C	8.73	18	79	196	54	-20	$6,000,000
Patrick Sharp	LW	8.70	34	77	306	41	12	$6,500,000
Patrick Marleau	LW	8.64	33	70	285	50	0	$6,900,000
Anze Kopitar	C	8.63	29	70	200	42	34	$7,500,000
Zach Parise	LW	8.54	29	56	245	46	10	$12,000,000
Jamie Benn	LW	8.38	34	79	279	70	21	$5,000,000
Ryan Johansen	C	8.36	33	63	237	39	3	$810,000
Max Pacioretty	LW	8.24	39	60	270	28	8	$4,000,000
Derek Stepan	C	7.79	17	57	199	50	12	$2,300,000
T.J. Oshie	RW	7.75	21	60	152	62	19	$4,000,000
Tyler Seguin	C	7.63	37	84	294	67	16	$4,500,000
Matt Duchene	C	7.49	23	70	217	40	8	$3,750,000
Bryan Little	C	7.41	23	64	170	38	8	$4,000,000
Brad Richards	C	7.25	20	51	259	33	-8	$9,000,000
Sean Monahan	C	7.23	22	34	140	26	-20	$925,000
Ryan O'Reilly	C	7.21	28	64	201	83	-1	$6,500,000
Patrice Bergeron	C	7.18	30	62	243	49	38	$4,550,000

situations is important. Results for powerplay situations are covered in Section 9.4.1 and penalty kill situations are covered in Section 9.4.2. These win impact scores are a subset of the general win impact scores reported in Section 9.3, as they focus on particular game contexts.

9.4.1 Powerplay

The top-25 player impact scores during powerplay situations in the 2013-2014 regular season are shown in Table 9.4. PPTOI is an acronym used by the NHL for Powerplay Time on Ice. Sidney Crosby has the highest win impact score on the powerplay, generating 4.73 wins from powerplay situations alone. Torey Krug is an interesting find in this list, as he isn't found in the top-25 players for win impact in all contexts, and isn't typically a name heard in discussions of top-tier players. It is clear by analyzing players in different contexts that key players can be found for different gameplay situations.

Table 9.4: 2013-2014 Top-25 Player Impacts For Winning in Powerplay Situations

Name	Position	Winning Impact in Powerplays	PPTOI	Goals	Points
Sidney Crosby	C	4.73	343.0	36	104
Jason Spezza	C	4.03	229.3	23	66
Claude Giroux	C	3.23	307.3	28	86
Shea Weber	D	3.22	246.6	23	56
Valtteri Filppula	C	2.99	252.4	25	58
Pavel Datsyuk	C	2.92	141.7	17	37
Jonathan Toews	C	2.90	238.0	27	67
Niklas Kronwall	D	2.69	262.4	8	48
Zach Parise	LW	2.63	223.5	29	56
Alex Ovechkin	RW	2.59	392.8	51	79
Ryan O'Reilly	C	2.59	213.3	28	64
Torey Krug	D	2.51	198.7	14	40
Rick Nash	LW	2.45	161.8	26	39
Joe Pavelski	C	2.44	288.3	41	79
Marian Hossa	RW	2.43	157.4	29	57
Eric Staal	C	2.41	263.9	21	61
Tyler Seguin	C	2.38	301.1	37	84
Patrick Marleau	LW	2.29	296.3	33	70
Ryan Kesler	C	2.28	268.8	25	43
Nicklas Backstrom	C	2.27	301.1	18	79
Mikko Koivu	C	2.24	224.2	11	54
Oliver Ekman-Larsson	D	2.15	327.4	15	44
Henrik Zetterberg	LW	2.14	158.7	16	48
Bryan Little	C	2.07	218.5	23	64
Radim Vrbata	RW	2.04	224.0	20	51

9.4.2 Penalty Kill

The top-25 players on the penalty kill with respect to winning are shown in Table 9.5. SHTOI is an acronym used in the NHL for Shorthanded Time On Ice. It is interesting to note that Matt Duchene has one of the highest winning impacts while shorthanded, given that he has only played 18.9 minutes on the penalty kill and most other players in this class have played at least 100 minutes on the penalty kill. This cannot be attributed to a statistical fluke, as Matt Duchene played in 71 out of 82 games throughout the 2013-2014 regular season. This finding suggests Matt Duchene's coach should be more willing to put him on the ice during penalty kill situations. Other players, such as Brad Marchand and Dan Hamhuis, are expected in this list, as they are known to be some of the best performers of the hip-check (a technique for hitting players) in the NHL.

Table 9.5: 2013-2014 Top-25 Player Impacts For Winning in Shorthanded Situations

Name	Position	Winning Impact in Penalty Kill	SHTOI	Goals	Points
Brandon Sutter	C	1.78	187.3	13	26
Dominic Moore	C	1.74	122.3	6	18
Cal Clutterbuck	RW	1.65	129.7	12	19
Ondrej Palat	LW	1.57	167.9	23	59
Marian Hossa	RW	1.48	92.8	29	57
Brad Marchand	LW	1.43	127.1	25	53
Matt Read	RW	1.41	230.3	22	40
Brandon Dubinsky	C	1.36	163.6	16	50
Matt Duchene	C	1.29	18.9	23	70
Jaden Schwartz	LW	1.26	123.1	25	56
Mikael Backlund	C	1.25	147.1	18	39
Joe Pavelski	C	1.20	130.7	41	79
Marc Staal	D	1.17	145.1	3	14
Fedor Tyutin	D	1.16	190.8	4	26
Francois Beauchemin	D	1.15	223.0	4	17
Tyler Johnson	C	1.14	166.3	24	50
Matt Cooke	LW	1.10	206.8	10	28
Shea Weber	D	1.10	205.6	23	56
Artem Anisimov	C	1.06	169.6	22	39
Dan Hamhuis	D	1.06	246.0	5	22
Jordan Eberle	RW	1.05	55.2	28	65
Anze Kopitar	C	1.05	164.9	29	70
Adam Henrique	C	1.05	169.4	25	43
Patrick Dwyer	RW	1.04	136.4	8	22
Brian Gionta	RW	1.03	130.0	18	40

9.5 Advanced Statistics Comparison

Our win impact score easily lends itself for comparison to other advanced statistics. We start by comparing our win impact score with Added Goal Value in Section 9.5.1. Next, we compare with a popular advanced statistic, Total Hockey Rating, in Section 9.5.2.

9.5.1 Win Impact vs. Added Goal Value (AGV)

The Added Goal Value (AGV) [20] metric is a measurement of how a goal contributes to the value of winning. As such, it can naturally be compared against the impact score with respect to winning. The comparison of win impact versus AGV is shown in Table 9.6. The win impact values vary greatly from the AGV values, and the reasons for this are twofold. Firstly, for the cases where AGV is much greater than the win impact values, consider goals that occur in long play-by-play sequences. The events near the end of longer play-by-play sequences will be rarer and have occurrences close to 1. As such, when the value iteration is performed, from an MDP perspective, the edges between nodes, with edge and node occurrences both equal to 1, are deterministic and changing state does not change the game flow. Therefore, the winning impact of goals in these situations will appear to be 0. When calculating sum of player scores, these goals will make no contribution to the player win impact score. The second case is where AGV is much smaller than win impact values. As AGV only includes goal events, other player actions are not considered. The Markov Game model includes all player actions, and the win impact of these player actions are also included in the player win impact scores, not only goals. This will cause the win impact value to be much greater than AGV. We observe that Alex Ovechkin may have the highest impact when only observing goals scored, but when all actions are applied as in my Markov Game model, Joe Pavelski has the highest win impact per game in this selection of players.

Table 9.6: Impact vs. AGV

Name	AGV per game	Games Played	AGV	Win Impact (2013-2014)	Win Impact per game
Alex Ovechkin	13.68%	78	10.67	4.56	5.85%
Steven Stamkos	13.03%	37	4.82	2.92	7.89%
Jeff Skinner	10.28%	71	7.30	4.80	6.76%
Corey Perry	10.09%	81	8.17	5.60	6.91%
James Neal	9.54%	59	5.63	6.72	11.39%
Gustav Nyquist	9.15%	57	5.22	2.75	4.82%
Sidney Crosby	8.93%	80	7.14	9.49	11.86%
Phil Kessel	8.67%	82	7.11	4.52	5.51%
Max Pacioretty	8.61%	73	6.29	8.24	11.29%
Kyle Okposo	8.45%	71	6.00	6.95	9.79%
Joe Pavelski	8.34%	82	6.84	10.77	13.13%
Jeff Carter	8.20%	72	5.90	6.75	9.38%
Mike Cammalleri	8.06%	63	5.08	6.89	10.94%
Evgeni Malkin	7.85%	60	4.71	4.74	7.90%
Pavel Datsyuk	7.68%	45	3.46	5.77	12.82%

9.5.2 Win Impact vs. Total Hockey Rating (THoR)

The impact score with respect to winning is compared with the Total Hockey Rating (THoR) [25] wins created metric in Table 9.7. Our impact score with respect to winning is computed using Equation 7.4. The wins created metric reported in [25] spanned both the

2010-2011 and 2011-2012 regular seasons. The THoR values often agree with the average of the win impact scores across the 2010-2011 and 2011-2012 seasons for forwards. For defensemen, the THoR values and win impact scores vary greatly. This is due to the logistic regression approach used in [25] giving more weight to players who spend more time on the ice, and defensemen typically spend more time on the ice than forwards. The reason for this is that there are generally 6 defensemen and 12 forwards dressed for each team in a match.

Table 9.7: Impact vs. THoR

Name	Position	THoR Wins Created (2010-2012)	Win Impact (2010-2011)	Win Impact (2011-2012)
Alexander Steen	C	6.72	5.39	2.63
Pavel Datsyuk	C	6.32	3.15	8.05
Tyler Kennedy	C	6.05	4.43	2.26
Patrice Bergeron	C	5.95	4.61	8.63
Patric Hornqvist	RW	5.88	3.12	3.39
Kimmo Timonen	D	5.73	3.20	0.24
Ray Whitney	LW	5.62	4.86	7.98
Evgeni Malkin	C	5.57	4.03	11.91
Ryan Kesler	C	5.53	9.48	9.54
Jonathan Toews	C	5.50	11.24	8.36
Daniel Sedin	LW	5.47	6.28	3.66
Joe Pavelski	C	5.42	7.28	9.57
Jeff Skinner	C	5.07	4.89	2.78
Anze Kopitar	C	4.93	8.74	6.95
Sidney Crosby	C	4.92	5.13	2.47
Drew Doughty	D	4.07	3.35	0.26
Tom Gilbert	D	3.32	-0.79	0.35
Fedor Tyutin	D	3.13	0.13	2.23
Mark Giordano	D	3.08	-0.69	1.15
Andrej Meszaros	D	2.82	2.61	1.57
Brent Seabrook	D	2.63	-0.24	4.04
Ryan McDonagh	D	2.50	0.78	0.92
Niklas Kronwall	D	2.48	1.08	2.34
Lubomir Visnovsky	D	2.48	2.91	-0.99
Paul Martin	D	2.27	2.68	0.79
Tobias Enstrom	D	2.23	1.31	1.03
Erik Karlsson	D	2.22	3.60	3.56
Zdeno Chara	D	2.18	3.20	2.35
Michael Sauer	D	1.95	0.74	-0.02

Chapter 10

Conclusion

In our research, we construct a Markov Game Model for a massive set of NHL play-by-play events with a rich state space that utilizes much of the information in the data. Tree-based data structures support efficient parameter estimation and storage. Value iteration computes the values of each action given its context and sequence history—the Q-function of the model. Compared to previous work that assigns a single value to actions, the 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. The Q-function provides knowledge about hockey dynamics by quantifying how much which action matters when. Propagating action effects across sequences utilizes the ordering of play sequences in a game, rather than treating sequences as an unordered independent set. Analysis of the computed Q-function shows the impact of an action varies greatly with context, and medium-term ripple effects make a difference. The Markov Game model is applied to evaluate the performance of players in terms of their actions' total impact. Action impact scores are calculated for players with respect to different objective functions. Impact scores for the next goal correlate with points and +/- statistics. The impact of players on the next penalty has not been previously considered, and shows some surprises, as some highly-paid players hurt their team by causing penalties. Another potential application for context-aware performance evaluation is in finding strengths and weaknesses of teams: The Q-function can be used to find situations in which a team's mix of actions provides a substantially different expected result from a generic team. 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.

10.1 Future Work

The NHL data provides a rich dataset for real-world event modelling. A number of further AI techniques can be applied to utilize even more of the available information than our Markov Game model does. A promising direction is to extend this Markov Game model, which is discrete with data about continuous quantities. These include (i) the time between events, —which requires a continuous time Markov Game model, (ii) the absolute game time of the events, (e.g. “minute 15”), (iii) location of shots [10] (however, reported shot locations are noisy [23]). Continuous time models could also incorporate player shift changes that occur within play sequences, and determine optimal shift lengths for different players. Player information, such as age [3, 2] and salary [5], has also been shown to affect player performance, and methods for augmenting our model with this information could be an interesting study.

The use of reinforcement learning techniques has been mainly for finding patterns in a rich data set, in the spirit of descriptive statistics and data mining. Another goal is to *predict* a player or team's future performance based on past performance using machine learning techniques. For example, is it possible to predict a player's performance in the next season based on the previous seasons? Machine learning methods aim to provide reliable generalization, and can be combined with dynamic programming for predicting future performance [32]. For example, sequence modelling would be able to generalize from play sequence information. A promising model class are Piecewise Constant Conditional Intensity Models for continuous time event sequences [7, 19]. These models are especially well suited for sequences with a large set of possible events, such as our action events. Another promising machine learning approach is to combine regression techniques with action value summations to determine player valuations.

Bibliography

[1] Jason Abrevaya and Robert McCulloch. Reversal of fortune: a statistical analysis of penalty calls in the national hockey league. *Journal of Quantitative Analysis in Sports*, 10(2):207–224, 2014. 23

[2] James A. Brander, Yeung Louisa, and Egan Edward J. Estimating the effects of age on nhl player performance. *Journal of Quantitative Analysis in Sports*, 10(2):241259, 2014. 16, 75

[3] Alex Bryson, Rafael Gomez, and Tingting Zhang. All-star or benchwarmer? relative age, cohort size and career success in the nhl. 2014. 16, 75

[4] Samuel Buttrey, Alan Washburn, and Wilson Price. Estimating nhl scoring rates. *Journal of Quantitative Analysis in Sports*, 7(3), 2011. 10, 23

[5] Philippe Cyrenne. Salary inequality, team success and the superstar effect. Technical report, The University of Winnipeg, 2014. 75

[6] R.B. Gramacy, S.T. Jensen, and M. Taddy. Estimating player contribution in hockey with regularized logistic regression. *Journal of Quantitative Analysis in Sports*, 9:97–111, 2013. 8

[7] Asela Gunawardana, Christopher Meek, and Puyang Xu. A model for temporal dependencies in event streams. In *Advances in Neural Information Processing Systems*, pages 1962–1970, 2011. 75

[8] Adam Hipp and Lawrence Mazlack. Mining ice hockey: Continuous data flow analysis. In Estonia Pascal Lorenz University of Haute Alsace France Ulrich Norbisrath, University of Tartu, editor, *IMMM 2011: The First International Conference on Advances in Information Mining and Management*, pages 31–36, 2011. 13

[9] Nobuyoshi Hirotsu and Michael Wright. Using a markov process model of an association football match to determine the optimal timing of substitution and tactical decisions. *Journal of the Operational Research Society*, 53(1):88–96, 2002. 10

[10] Ken Krzywicki. Shot quality model: A logistic regression approach to assessing nhl shots on goal, January 2005. 15, 75

[11] H.J. Levesque, F. Pirri, and R. Reiter. Foundations for the situation calculus. *Linköping Electronic Articles in Computer and Information Science*, 3(18), 1998. 25

[12] Michael 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. 1, 7, 20, 21, 31

[13] D. Lock and M. Schuckers. Beyond +/-: A rating system to compare nhl players., 2009. Presentation at joint statistical meetings. 3, 8, 59

[14] B. Macdonald. An improved adjusted plus-minus statistic for nhl players. In *Proceedings of the MIT Sloan Sports Analytics Conference*, March 2011. 2, 8, 50

[15] B. Macdonald. An expected goals model for evaluating nhl teams and players. In *MIT Sloan Sports Analytics Conference*, March 2012. 29, 68

[16] Brian Macdonald. A regression-based adjusted plus-minus statistic for nhl players. *Journal of Quantitative Analysis in Sports*, 7(3):29, 2011. 8

[17] Andrew W. Moore and Mary S. Lee. Cached sufficient statistics for efficient machine learning with large datasets. *Journal of Artificial Intelligence Research*, 8:67–91, 1998. 4, 5, 31

[18] National Hockey League. National hockey league official rules 2014-2015, 2014. 12

[19] Ankur P Parikh, Asela Gunawardana, and Christopher Meek. Conjoint modeling of temporal dependencies in event streams. In *UAI Bayesian Modelling Applications Workshop*. Citeseer, 2012. 75

[20] Stephen Pettigrew. Assessing the offensive productivity of nhl players using in-game win probabilities. In *9th Annual MIT Sloan Sports Analytics Conference*, 2015. 2, 3, 8, 9, 23, 29, 64, 71

[21] Dan T. Rosenbaum. Measuring how nba players help their teams win, April 2004. 8

[22] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, 2010. 7, 15, 16, 21, 31, 58

[23] M. Schuckers and B. Macdonald. Accounting for rink effects in the national hockey league's real time scoring system. *ArXiv e-prints*, Dec 2014. 75

[24] Michael Schuckers and Lauren Brozowski. Referee analytics: An analysis of penalty rates by national hockey league officials. In *MIT Sloan Sports Analytics Conference*, 2012. 23, 61

[25] Michael Schuckers and James 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. xi, 2, 4, 8, 9, 29, 32, 59, 60, 72, 73

[26] Michael E. Schuckers, Dennis F. Lock, Chris Wells, C. J. Knickerbocker, and Robin H. Lock. National hockey league skater ratings based upon all on-ice events: An adjusted minus/plus probability (ampp) approach. Unpublished manuscript., 2011. 59

[27] Oliver Schulte and Kurt Routley. Aggregating predictions vs. aggregating features for relational classification. In *Computational Intelligence and Data Mining (CIDM), 2014 IEEE Symposium on*, pages 121–128, Dec 2014. 2

[28] Robert P. Schumaker, Osama K. Solieman, and Hsinchun Chen. Research in sports statistics. In *Sports Data Mining*, volume 26 of *Integrated Series in Information Systems*, pages 29–44. Springer US, 2010. 2

[29] Anton Schwartz. A reinforcement learning method for maximizing undiscounted rewards. In *Proceedings of the tenth international conference on machine learning*, volume 298, pages 298–305, 1993. 41

[30] Gagan Sidhu and Brian Caffo. Moneybarl: Exploiting pitcher decision-making using reinforcement learning. *The Annals of Applied Statistics*, 8(2):926–955, June 2014. 10

[31] Nathan Spagnola. The complete plus-minus: A case study of the columbus blue jackets. Master's thesis, University of South Carolina, 2013. 2, 8

[32] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning : an introduction*. MIT Press, Cambridge, Mass., 1998. 1, 8, 21, 42, 75

[33] A.C. Thomas, S.L. Ventura, S. Jensen, and S. Ma. Competing process hazard function models for player ratings in ice hockey. *The Annals of Applied Statistics*, 7(3):1497–1524, 2013. 10, 21, 23, 25, 26

[34] Joshua Weissbock. *Forecasting Success in the National Hockey League using In-Game Statistics and Textual Data*. PhD thesis, University of Ottawa, 2014. 2

Appendix A

Player Rankings: Goals

We report player rankings with respect to impact scores for the probability of the next goal. We compute action values for the probability of the next goal using Equation 7.4. For each player, we sum up the values of his actions over a game, and then over a season, to compute a net season impact for the player. The reported values are net impact values over the regular season being observed. A positive impact score with respect to goals means the player generates goals for their team. A negative impact score means the player causes goals to be scored against their team. We observe that some top-tier players, such as Jason Spezza, have consistently high performance across seasons. When observing other players in the top-25 and bottom-25 player rankings, it is clear that player performance can vary across seasons. Average player values are found by taking the average of all players' net impact value.

A.1 2014-2015

Player impact scores with respect to the probability of the next goal are recorded for the first 512 games of the 2014-2015 regular season. The average player generated 1.95 goals for his team. The top-25 players are shown in Table A.1 and the bottom-25 players are shown in Table A.2.

A.2 2013-2014

The average player generated 4.28 goals for his team during the 2013-2014 regular season. Table A.3 shows the top-25 players with the highest goal impact scores and generate goals for their team. Table A.4 shows the bottom-25 players with the lowest goal impact scores, meaning their actions cause their opponent to score goals.

Table A.1: 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
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

A.3 2012-2013

The average player generated 2.89 goals for his team during the 2012-2013 regular season. Table A.5 shows the top-25 players with the highest goal impact scores and generate goals for their team. Table A.6 shows the bottom-25 players with the lowest goal impact scores, meaning their actions cause their opponent to score goals.

A.4 2011-2012

The average player generated 4.13 goals for his team during the 2011-2012 regular season. Table A.7 shows the top-25 player impact scores for goal scoring. Table A.8 shows the bottom-25 player impact scores.

Table A.2: 2014-2015 Bottom-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Jan Hejda	D	-6.97	0	5	-5	14	$3,250,000
Chris Neil	RW	-4.73	4	7	3	4	$2,100,000
Hampus Lindholm	D	-4.39	4	15	11	14	$925,000
Alex Goligoski	D	-4.18	1	15	1	12	$4,800,000
Willie Mitchell	D	-3.41	1	2	-6	9	$4,250,000
Joe Vitale	C	-3.08	3	6	-3	17	$950,000
Ryan Suter	D	-3.03	1	22	3	12	$11,000,000
Dylan Olsen	D	-2.83	1	6	-1	7	$700,000
Matt Stajan	C	-2.78	1	4	1	5	$3,625,000
Matt Hunwick	D	-2.78	0	7	8	10	$600,000
Cody McLeod	LW	-2.76	2	5	2	7	$1,150,000
Manny Malhotra	C	-2.69	0	1	-4	9	$850,000
Thomas Hickey	D	-2.64	2	11	3	12	$750,000
Brad Malone	C	-2.53	0	0	-10	12	$600,000
Adam Larsson	D	-2.50	1	4	-3	4	$900,000
Brenden Dillon	D	-2.38	0	5	-2	5	$1,250,000
Steve Downie	RW	-2.14	5	17	8	4	$1,000,000
Mikhail Grabovski	C	-2.03	5	12	3	14	$4,000,000
Jesse Joensuu	LW	-2.02	2	4	-8	9	$1,000,000
Lauri Korpikoski	LW	-1.99	3	10	-13	12	$2,300,000
Rob Scuderi	D	-1.96	0	5	5	5	$4,000,000
Chris Phillips	D	-1.95	0	2	-1	6	$2,500,000
Travis Hamonic	D	-1.91	3	7	0	9	$2,500,000
Jim Slater	C	-1.87	1	3	2	3	$1,600,000
Chris Kreider	LW	-1.78	5	15	6	6	$2,350,000

A.5 2010-2011

The average player generated 3.99 goals for his team during the 2010-2011 regular season. Table A.9 shows the top-25 player impact scores. Table A.10 shows the bottom-25 player impact scores.

A.6 2009-2010

The average player generated 3.99 goals for his team during the 2009-2010 regular season. Table A.11 shows the top-25 player impact scores. Table A.12 shows the bottom-25 player impact scores.

Table A.3: 2013-2014 Top-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Jason Spezza	C	29.64	23	66	-26	47	$5,000,000
Jonathan Toews	C	28.75	27	67	25	51	$6,500,000
Joe Pavelski	C	27.20	41	79	23	56	$4,000,000
Marian Hossa	RW	26.12	29	57	26	73	$7,900,000
Patrick Sharp	LW	24.43	34	77	12	41	$6,500,000
Sidney Crosby	C	24.23	36	104	18	41	$12,000,000
Claude Giroux	C	23.89	28	86	7	43	$5,000,000
Tyler Seguin	C	23.89	37	84	16	67	$4,500,000
Max Pacioretty	LW	22.54	39	60	8	28	$4,000,000
Patrice Bergeron	C	22.26	30	62	38	49	$4,550,000
Jamie Benn	LW	22.08	34	79	21	70	$5,000,000
Ryan O'Reilly	C	22.06	28	64	-1	83	$6,500,000
Ryan Johansen	C	21.96	33	63	3	39	$810,000
Valtteri Filppula	C	21.70	25	58	5	53	$4,000,000
Patrick Marleau	LW	20.92	33	70	0	50	$6,900,000
Matt Duchene	C	20.67	23	70	8	40	$3,750,000
John Tavares	C	20.45	24	66	-6	55	$5,000,000
Zach Parise	LW	19.93	29	56	10	46	$12,000,000
David Backes	C	19.25	27	57	14	32	$4,750,000
Derek Stepan	C	19.11	17	57	12	50	$2,300,000
Bryan Little	C	18.58	23	64	8	38	$4,000,000
Brad Richards	C	18.45	20	51	-8	33	$9,000,000
Anze Kopitar	C	18.45	29	70	34	52	$7,500,000
Logan Couture	C	18.37	23	54	21	39	$3,000,000
Nicklas Backstrom	C	17.36	18	79	-20	54	$6,000,000

A.7 2008-2009

The average player generated 3.99 goals for his team during the 2008-2009 regular season. Table A.13 shows the top-25 player impact scores. Table A.14 shows the bottom-25 player impact scores.

A.8 2007-2008

The average player generated 3.94 goals for his team during the 2007-2008 regular season. Table A.15 shows the top-25 player impact scores. Table A.16 shows the bottom-25 player impact scores.

Table A.4: 2013-2014 Bottom-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Rich Clune	LW	-5.24	3	7	-7	11	$525,000
Andrew MacDonald	D	-5.11	4	28	-22	23	$575,000
Willie Mitchell	D	-4.87	1	12	14	10	$3,500,000
Matt Carkner	D	-4.57	0	3	-10	9	$1,500,000
Jacob Trouba	D	-4.52	10	29	4	31	$925,000
Chris Neil	RW	-3.83	8	14	-10	16	$2,100,000
Patrick Maroon	LW	-3.80	11	29	11	15	$575,000
Mike Brown	RW	-3.40	2	5	-9	8	$725,000
Tom Sestito	LW	-3.36	5	9	-14	8	$650,000
Chuck Kobasew	RW	-3.34	2	2	1	10	$434,000
Mark Fraser	D	-3.34	1	2	-15	7	$1,275,000
Johnny Oduya	D	-3.21	3	16	11	15	$3,300,000
Tim Gleason	D	-3.10	1	6	-21	9	$4,500,000
Colton Orr	RW	-3.03	0	0	-3	3	$925,000
Mike Weber	D	-2.68	1	9	-29	11	$1,500,000
Patrick Bordeleau	LW	-2.52	6	11	-1	8	$1,000,000
Nick Schultz	D	-2.50	0	5	-13	2	$3,600,000
Mark Stuart	D	-2.33	2	13	11	6	$1,800,000
Matt Greene	D	-2.33	2	6	6	1	$3,250,000
John Scott	LW	-2.30	1	1	-12	3	$750,000
Nate Guenin	D	-2.21	1	9	3	12	$600,000
Ville Leino	LW	-2.18	0	15	-16	23	$4,000,000
Travis Moen	LW	-2.06	2	12	2	15	$1,850,000
Dmitry Kulikov	D	-2.06	8	19	-26	29	$3,000,000
Mark Fistric	D	-1.94	1	5	9	5	$900,000

Table A.5: 2012-2013 Top-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Jonathan Toews	C	20.84	23	48	28	56	$6,000,000
Sidney Crosby	C	19.47	15	56	26	15	$7,500,000
Pavel Datsyuk	C	16.86	15	49	26	15	$6,700,000
John Tavares	C	16.13	28	47	-2	27	$4,000,000
Patrick Kane	RW	15.94	23	55	11	36	$6,000,000
Dustin Brown	RW	14.86	18	29	6	17	$3,500,000
Mikko Koivu	C	14.75	11	37	2	26	$5,400,000
Zach Parise	LW	14.21	18	38	2	24	$12,000,000
Claude Giroux	C	13.92	13	48	-7	17	$3,500,000
Matt Duchene	C	13.86	17	43	-12	44	$3,250,000
Logan Couture	C	13.82	21	37	7	31	$2,750,000
Derek Stepan	C	13.78	18	44	25	34	$875,000
Rick Nash	LW	13.39	21	42	16	19	$7,600,000
Jamie Benn	LW	12.95	12	33	-12	41	$4,500,000
Mark Letestu	C	12.64	13	27	7	27	$600,000
Sam Gagner	C	12.15	14	38	-6	23	$3,200,000
Matt Moulson	LW	12.09	15	44	-3	20	$3,000,000
Corey Perry	RW	12.08	15	36	10	18	$4,875,000
Jason Pominville	RW	11.70	14	34	1	38	$5,500,000
Ryan Callahan	RW	11.47	16	31	9	23	$4,000,000
Kyle Turris	C	11.43	12	29	6	28	$1,600,000
David Krejci	C	11.30	10	33	1	20	$5,250,000
Ryan Getzlaf	C	10.85	15	49	14	27	$6,125,000
Alexander Steen	LW	10.81	8	27	5	13	$3,567,000
Steven Stamkos	C	10.77	29	57	-4	24	$8,000,000

Table A.6: 2012-2013 Bottom-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Jay Harrison	D	-4.04	3	10	-10	12	$750,000
Nicklas Grossmann	D	-3.62	1	4	-1	6	$3,500,000
John Erskine	D	-3.58	3	6	10	4	$1,500,000
Michael Del Zotto	D	-3.44	3	21	6	10	$2,200,000
Roman Polak	D	-2.78	1	6	-2	8	$2,450,000
Deryk Engelland	D	-2.69	0	6	5	7	$525,000
Ian Cole	D	-2.46	0	1	-4	1	$875,000
Tuomo Ruutu	LW	-2.43	4	9	-6	12	$4,000,000
Erik Gudbranson	D	-2.33	0	4	-22	2	$900,000
Kevin Westgarth	RW	-2.30	2	4	1	4	$700,000
Brad Malone	C	-2.24	1	2	-7	6	$788,000
Chris Thorburn	RW	-2.20	2	4	-5	3	$850,000
George Parros	RW	-2.11	1	2	-15	5	$925,000
Andrej Meszaros	D	-2.04	0	2	-9	2	$4,750,000
Jared Boll	RW	-1.98	2	6	1	1	$950,000
Tyson Strachan	D	-1.92	0	4	-13	11	$600,000
Zenon Konopka	C	-1.91	0	0	-4	5	$850,000
Paul Bissonnette	LW	-1.86	0	6	2	1	$725,000
Robyn Regehr	D	-1.79	0	4	-4	1	$4,000,000
Mike Weber	D	-1.77	1	7	3	5	$1,000,000
Jordin Tootoo	RW	-1.57	3	8	0	11	$1,700,000
John Scott	LW	-1.50	0	0	-1	8	$600,000
Jordie Benn	D	-1.49	1	6	-4	8	$500,000
Ryan Whitney	D	-1.48	4	13	-7	9	$5,500,000
Adam Larsson	D	-1.44	0	6	4	17	$925,000

Table A.7: 2011-2012 Top-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Evgeni Malkin	C	33.78	50	109	18	52	$9,000,000
Zach Parise	LW	28.98	31	69	-5	65	$6,000,000
Jason Spezza	C	26.32	34	84	11	64	$8,000,000
John Tavares	C	26.11	31	81	-6	99	$900,000
Claude Giroux	C	24.62	28	93	6	50	$2,750,000
Ryan Kesler	C	24.12	22	49	11	43	$5,000,000
Loui Eriksson	RW	24.11	26	71	18	50	$4,100,000
Joe Pavelski	C	23.68	31	61	18	73	$4,000,000
Jonathan Toews	C	23.53	29	57	17	82	$6,000,000
David Krejci	C	22.87	23	62	-5	43	$4,000,000
Patrice Bergeron	C	22.50	22	64	36	55	$5,900,000
Teemu Selanne	RW	22.14	26	66	-1	28	$4,000,000
Steven Stamkos	C	22.09	60	97	7	42	$8,000,000
Jason Pominville	RW	21.82	30	73	-7	45	$5,500,000
David Backes	C	21.70	23	53	14	50	$4,500,000
Rick Nash	LW	21.06	30	59	-10	62	$7,500,000
Logan Couture	C	21.02	31	65	2	61	$788,000
Radim Vrbata	RW	20.96	35	61	22	36	$3,000,000
Alex Ovechkin	RW	20.74	38	65	-8	34	$9,000,000
Jamie Benn	LW	20.64	26	63	15	56	$670,000
Marian Gaborik	RW	20.08	41	76	15	30	$7,500,000
Pavel Datsyuk	C	19.98	19	67	21	97	$6,700,000
Phil Kessel	RW	19.62	37	82	-10	39	$6,000,000
Ilya Kovalchuk	RW	19.29	37	83	-9	42	$11,000,000
Patrick Marleau	LW	19.19	30	64	10	38	$6,900,000

Table A.8: 2011-2012 Bottom-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Scott Hannan	D	-7.15	2	12	-10	13	$1,000,000
Travis Hamonic	D	-5.26	2	24	6	37	$875,000
Colin White	D	-5.10	1	4	-5	10	$1,000,000
Rostislav Klesla	D	-4.68	3	13	13	13	$2,975,000
Jake Gardiner	D	-4.17	7	30	-2	34	$875,000
Zac Rinaldo	C	-3.96	2	9	-1	4	$560,000
John Carlson	D	-3.27	9	32	-15	31	$788,000
Jared Boll	RW	-3.11	2	3	-8	7	$750,000
Jody Shelley	LW	-3.04	0	1	-6	4	$1,100,000
Robyn Regehr	D	-2.72	1	5	-12	11	$4,000,000
Radek Dvorak	RW	-2.69	4	21	-16	42	$1,500,000
Jay Pandolfo	LW	-2.65	1	3	-14	14	$600,000
Tim Gleason	D	-2.45	1	18	12	18	$3,500,000
Stu Bickel	D	-2.40	0	9	2	5	$600,000
Adam Pardy	D	-2.32	0	3	-5	9	$2,000,000
Chris Pronger	D	-2.29	1	12	1	7	$7,600,000
Jason Demers	D	-2.26	4	13	-8	11	$1,100,000
Justin Faulk	D	-2.08	8	22	-16	32	$790,000
Jared Cowen	D	-2.04	5	17	-4	28	$900,000
Brett Clark	D	-1.99	2	15	-26	26	$1,300,000
Warren Peters	C	-1.96	1	5	-15	24	$497,000
Adam Hall	RW	-1.91	2	7	-11	12	$600,000
Arron Asham	RW	-1.84	5	16	-5	13	$775,000
Pavel Kubina	D	-1.77	3	15	-2	18	$3,500,000
Sheldon Brookbank	D	-1.77	3	13	12	12	$800,000

Table A.9: 2010-2011 Top-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Ryan Kesler	C	29.28	41	73	24	65	$5,000,000
Jonathan Toews	C	27.96	32	76	25	93	$6,500,000
Eric Staal	C	26.73	33	76	-10	64	$7,500,000
Jeff Carter	C	25.92	36	66	27	40	$5,500,000
Alex Ovechkin	RW	24.78	32	85	24	48	$9,000,000
Joe Pavelski	C	23.86	18	63	9	50	$4,000,000
Claude Giroux	C	23.73	25	76	20	48	$765,000
Patrick Sharp	LW	23.01	34	71	-1	64	$4,100,000
Tomas Plekanec	C	21.97	22	57	8	43	$5,000,000
Brad Richards	C	21.45	28	77	1	47	$7,800,000
Anze Kopitar	C	21.29	25	73	25	62	$6,000,000
Henrik Zetterberg	LW	21.14	24	80	-1	54	$7,750,000
Jason Spezza	C	21.05	21	57	-7	52	$8,000,000
John Tavares	C	20.82	29	67	-15	75	$900,000
Michael Grabner	RW	20.35	34	52	14	69	$765,000
Steven Stamkos	C	20.29	45	91	3	40	$875,000
Brandon Dubinsky	C	19.43	24	54	-3	48	$2,000,000
Paul Stastny	C	19.26	22	57	-7	52	$6,600,000
Stephen Weiss	C	18.84	21	49	-9	44	$3,200,000
Mike Santorelli	C	18.80	20	41	-17	32	$600,000
Rick Nash	LW	18.69	32	66	2	47	$7,500,000
Jarome Iginla	RW	18.14	43	86	0	40	$7,000,000
Thomas Vanek	LW	18.10	32	73	2	43	$6,400,000
Bryan Little	C	17.81	18	48	11	80	$1,650,000
Patrick Marleau	LW	17.72	36	71	-5	34	$6,900,000

Table A.10: 2010-2011 Bottom-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Andrew Alberts	D	-5.88	1	7	0	7	$1,300,000
Theo Peckham	D	-5.62	3	13	-5	30	$550,000
Matt Martin	LW	-5.54	5	14	-13	25	$615,000
Chris Phillips	D	-5.50	1	9	-35	26	$3,500,000
Cody McLeod	LW	-4.56	5	8	-7	7	$1,000,000
Brian Lee	D	-3.84	0	3	-10	5	$875,000
Jim Vandermeer	D	-3.84	2	14	-15	27	$2,300,000
Sean O'Donnell	D	-3.03	1	18	8	15	$1,300,000
Matt Carkner	D	-2.83	1	7	0	9	$700,000
Derek Joslin	D	-2.82	2	9	4	7	$500,000
Bryan Allen	D	-2.81	4	17	-1	24	$3,100,000
Brad Staubitz	RW	-2.81	4	9	-5	6	$550,000
Jamal Mayers	RW	-2.80	3	14	3	24	$600,000
Jared Boll	RW	-2.75	7	12	-2	14	$700,000
Jassen Cullimore	D	-2.54	0	8	4	7	$328,000
Cory Sarich	D	-2.43	4	17	11	15	$3,700,000
J.P. Dumont	RW	-2.33	10	19	2	35	$4,000,000
John Erskine	D	-2.31	4	11	1	11	$1,250,000
Kevin Westgarth	RW	-2.31	0	3	-6	5	$500,000
Chris Neil	RW	-2.20	6	16	-14	35	$2,000,000
Greg Zanon	D	-2.20	0	7	-5	21	$2,000,000
Jonas Holos	D	-2.19	0	6	-3	9	$624,000
Mike Brown	RW	-2.13	3	8	1	11	$550,000
Andreas Lilja	D	-2.13	1	7	-15	6	$600,000
Matthew Corrente	D	-2.10	0	6	-5	0	$817,500

Table A.11: 2009-2010 Top-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Sidney Crosby	C	28.12	51	109	15	43	$9,000,000
Jonathan Toews	C	25.66	25	68	22	69	$850,000
Anze Kopitar	C	24.00	34	81	6	36	$6,000,000
Paul Stastny	C	22.72	20	79	2	59	$6,600,000
Jeff Carter	C	22.21	33	61	2	43	$5,000,000
Nicklas Backstrom	C	22.08	33	101	37	54	$850,000
Alex Ovechkin	RW	22.01	50	109	45	66	$9,000,000
Henrik Zetterberg	LW	21.91	23	70	12	53	$7,400,000
Phil Kessel	RW	21.87	30	55	-8	30	$4,500,000
Matt Cullen	C	21.66	16	48	-7	54	$2,800,000
Travis Zajac	C	21.38	25	67	22	51	$2,750,000
Vincent Lecavalier	C	21.19	24	70	-16	33	$10,000,000
Jason Spezza	C	21.12	23	57	0	39	$8,000,000
Stephen Weiss	C	21.07	28	60	-7	63	$3,000,000
Mikko Koivu	C	20.81	22	71	-2	55	$3,300,000
Patrice Bergeron	C	20.79	19	52	6	55	$5,000,000
Jason Arnott	C	20.74	19	46	0	30	$4,500,000
Eric Staal	C	20.28	29	70	4	45	$6,000,000
Steven Stamkos	C	20.21	51	95	-2	47	$875,000
Ilya Kovalchuk	RW	20.14	41	85	10	34	$7,500,000
Tomas Plekanec	C	19.99	25	70	5	46	$2,750,000
Pavel Datsyuk	C	19.84	27	70	17	132	$6,700,000
Derek Roy	C	19.82	26	69	9	51	$3,500,000
Patrick Marleau	LW	19.56	44	83	21	53	$6,300,000
Brad Richards	C	19.21	24	91	-12	57	$7,800,000

Table A.12: 2009-2010 Bottom-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Matt Greene	D	-6.24	2	9	4	4	$2,750,000
Anton Volchenkov	D	-5.02	4	14	2	17	$3,200,000
Darcy Hordichuk	LW	-4.95	1	2	-7	7	$771,000
Douglas Murray	D	-4.74	4	17	3	24	$2,500,000
Matt Carkner	D	-4.58	2	11	0	19	$500,000
Ryan O'Byrne	D	-4.49	1	4	-3	7	$725,000
Craig Rivet	D	-3.88	1	15	-6	17	$3,500,000
Andreas Lilja	D	-3.76	1	2	-2	1	$1,250,000
Andrei Markov	D	-3.72	6	34	11	34	$5,750,000
Mike Lundin	D	-3.64	3	13	-4	17	$433,000
Kevin Klein	D	-3.47	1	11	-13	31	$800,000
Jonathan Ericsson	D	-3.44	4	13	-15	13	$900,000
Jared Boll	RW	-3.33	4	7	-8	10	$550,000
Nick Boynton	D	-3.26	1	8	5	14	$1,500,000
Zenon Konopka	C	-2.94	2	5	-11	7	$500,000
Josh Gorges	D	-2.91	3	10	2	20	$1,000,000
Derek Boogaard	LW	-2.61	0	4	-12	6	$930,000
Brandon Prust	LW	-2.47	5	14	9	11	$525,000
Luca Caputi	LW	-2.44	2	8	-1	8	$284,000
Adam Pardy	D	-2.29	2	9	-3	16	$700,000
Milan Lucic	LW	-2.27	9	20	-7	12	$685,000
Adam Foote	D	-2.24	0	9	8	11	$3,250,000
Andrew Peters	LW	-2.23	0	0	-5	0	$500,000
Shane Hnidy	D	-2.21	2	14	-6	11	$750,000
Dean Arsene	D	-2.15	0	0	-3	0	$292,000

Table A.13: 2008-2009 Top-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Sidney Crosby	C	33.44	33	103	3	56	$9,000,000
Jeff Carter	C	32.44	46	84	23	72	$4,500,000
Eric Staal	C	27.65	40	75	15	55	$5,000,000
Jarome Iginla	RW	26.02	35	89	-2	35	$7,000,000
Derek Roy	C	25.47	28	70	-5	52	$3,500,000
Alex Ovechkin	RW	24.33	56	110	8	60	$9,000,000
Vincent Lecavalier	C	24.13	29	67	-9	51	$7,167,000
Rick Nash	LW	23.97	40	79	11	70	$6,500,000
Zach Parise	LW	23.81	45	94	30	34	$2,500,000
Todd White	C	23.68	22	73	-9	57	$2,350,000
Chris Drury	C	22.78	22	56	-8	48	$7,100,000
Mike Richards	C	22.71	30	80	22	83	$5,400,000
Pavel Datsyuk	C	22.08	32	97	34	89	$6,700,00
Henrik Zetterberg	LW	21.79	31	73	13	42	$2,900,000
Mike Ribeiro	C	20.58	22	78	-4	67	$5,000,000
Saku Koivu	C	20.36	16	50	4	38	$4,750,000
Jonathan Toews	C	20.25	34	69	12	54	$850,000
Ryan Getzlaf	C	20.14	25	91	5	55	$4,500,000
Tomas Plekanec	C	19.96	20	39	-9	42	$1,800,000
Mikko Koivu	C	19.80	20	67	2	63	$3,300,000
Scott Gomez	C	19.54	16	58	-2	57	$8,000,000
Jason Blake	LW	19.28	25	63	-2	53	$4,500,000
Jason Pominville	RW	19.23	20	66	-4	43	$1,375,000
R.J. Umberger	C	19.17	26	46	-10	43	$3,000,000
Anze Kopitar	C	19.09	27	66	-17	49	$765,000

Table A.14: 2008-2009 Bottom-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Boris Valabik	D	-5.68	0	5	-14	19	$729,000
Jim Vandermeer	D	-4.75	1	7	1	14	$2,300,000
Chris Neil	RW	-4.32	3	10	-13	22	$1,200,000
Tom Poti	D	-4.21	3	13	3	24	$3,500,000
Zack Stortini	RW	-4.16	6	11	-3	3	$600,000
Bret Hedican	D	-3.93	1	6	-7	6	$805,000
Freddy Meyer	D	-3.89	4	9	-19	11	$575,000
Cory Sarich	D	-3.86	2	20	12	13	$3,400,000
John-Michael Liles	D	-3.47	12	39	-19	22	$3,700,000
Denis Gauthier	D	-3.39	2	4	-11	10	$1,931,000
Ben Eager	LW	-3.39	11	15	1	12	$601,000
Radek Martinek	D	-3.39	6	10	-16	47	$1,200,000
Luke Schenn	D	-3.34	2	14	-12	33	$2,975,000
Eric Godard	RW	-3.27	2	4	-3	2	$725,000
Steve Downie	RW	-3.03	3	6	-2	9	$274,000
Eric Boulton	LW	-3.02	3	13	-3	12	$600,000
Kyle Quincey	D	-2.90	4	38	-5	17	$500,000
Cam Janssen	RW	-2.84	1	4	-5	1	$550,000
Colton Orr	RW	-2.78	1	5	-15	9	$550,000
Krys Barch	RW	-2.71	4	9	1	9	$575,000
Theo Peckham	D	-2.63	0	0	-1	2	$174,000
Karl Alzner	D	-2.61	1	5	-1	10	$594,000
Ladislav Smid	D	-2.55	0	11	-6	13	$952,381
Darcy Hordichuk	LW	-2.53	4	5	1	11	$750,000
Jonas Frogren	D	-2.53	1	7	0	12	$1,230,000

Table A.15: 2007-2008 Top-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Rick Nash	LW	29.88	38	69	2	56	$6,500,000
Vincent Lecavalier	C	27.30	40	92	-17	52	$7,167,000
Jarome Iginla	RW	26.72	50	98	27	47	$7,000,000
Henrik Zetterberg	LW	26.53	43	92	30	53	$2,700,000
Marian Hossa	RW	26.11	29	66	-14	66	$7,000,000
Jason Spezza	C	24.60	34	92	26	44	$5,000,000
Alex Ovechkin	RW	24.41	65	112	28	68	$984,000
Pavel Datsyuk	C	24.01	31	97	41	144	$6,700,000
Evgeni Malkin	C	23.08	47	106	16	69	$984,000
Daniel Alfredsson	RW	22.83	40	89	15	72	$4,690,670
Mike Richards	C	22.68	28	75	14	46	$942,000
Jeff Carter	C	20.93	29	53	6	56	$942,400
Sidney Crosby	C	20.88	24	72	18	35	$850,000
Eric Staal	C	20.50	38	82	-2	56	$4,500,000
Daymond Langkow	C	20.04	30	65	16	52	$2,442,000
Alex Kovalev	RW	19.67	35	84	18	47	$4,500,000
Scott Gomez	C	19.49	16	70	3	77	$10,000,000
Ilya Kovalchuk	RW	19.45	52	87	-12	49	$5,432,000
Patrick Sharp	LW	18.74	36	62	23	44	$825,000
Mike Modano	C	18.44	21	57	-11	86	$4,250,000
Chris Drury	C	18.35	25	58	-3	64	$7,100,000
Marian Gaborik	RW	18.19	42	83	17	41	$6,500,000
Jarret Stoll	C	18.10	14	36	-23	39	$2,200,000
Marc Savard	C	17.92	15	78	3	47	$5,000,000
Paul Stastny	C	17.73	24	71	22	54	$685,000

Table A.16: 2007-2008 Bottom-25 Player Impact Scores For Goals

Name	Position	Goal Impact	Goals	Points	+/-	Takeaways	Salary
Anders Eriksson	D	-4.61	1	18	-5	28	$1,500,000
Kyle McLaren	D	-4.55	3	11	3	11	$2,500,000
Chris Neil	RW	-4.52	6	20	-3	21	$1,100,000
Zack Stortini	RW	-4.37	3	12	3	5	$475,000
Hal Gill	D	-4.36	3	24	6	24	$2,075,000
Krys Barch	RW	-4.30	1	3	-3	10	$475,000
Riley Cote	LW	-3.66	1	4	2	5	$476,000
John Erskine	D	-3.62	2	9	1	20	$525,000
George Parros	RW	-3.60	1	5	3	9	$525,000
Colton Orr	RW	-3.54	1	2	-13	14	$525,000
Aaron Downey	RW	-3.50	0	3	0	3	$525,000
Milan Jurcina	D	-3.39	1	9	4	15	$850,000
Staffan Kronwall	D	-3.38	0	0	-2	4	$112,000
Braydon Coburn	D	-3.26	9	36	17	36	$942,400
Craig Weller	RW	-3.25	3	11	-7	10	$475,000
Matt Bradley	RW	-3.23	7	18	1	21	$700,000
Ruslan Salei	D	-3.22	6	30	-4	13	$3,025,000
Ryan Hollweg	LW	-3.07	2	4	-12	10	$495,000
Greg Zanon	D	-3.02	0	5	-5	36	$700,000
Nicklas Grossmann	D	-3.00	0	7	10	8	$675,000
Jaroslav Modry	D	-2.94	1	9	-9	12	$1,200,000
Cory Sarich	D	-2.89	2	7	2	27	$3,900,000
Jeff Cowan	LW	-2.88	0	1	-5	13	$725,000
Jack Johnson	D	-2.63	3	11	-19	23	$2,150,000
Eric Godard	RW	-2.61	1	2	-8	1	$472,000

Appendix B

Player Rankings: Penalties

We report player impact scores with respect to the probability of receiving the next penalty. We compute action values for the probability of receiving the next penalty using Equation 7.4. Average player values are found by taking the average of all players' net impact value. Recall that while penalties are the "reward" from the perspective of the Q-function, they are actually a cost rather than a reward. This is because when players receive penalties, it has a negative effect on their team. As such, having a high or positive impact with respect to penalties is bad, and having a low or negative impact is good. A low or negative impact score means the player's actions are more likely to cause their opponent to receive a penalty. An interesting trend is that the average net impact score across a single season has been decreasing, from 15.26 penalties in the 2007-2008 regular season to 11.38 penalties during the 2013-2014 regular season. The 2012-2013 season would appear to break this trend, with 7.35 penalties generated on average, but there was a lockout during this season. The lockout caused 510 games to be removed from the season schedule, so there were less opportunities for players to generate penalties. This declining trend in penalties generated may suggest that either referees are more reluctant to call penalties, or players are behaving less recklessly and are less likely to incur penalties.

B.1 2014-2015

The average player caused their team to receive 5.51 penalties during the first 512 games of the 2014-2015 regular season. Table B.1 shows the top-25 player impact scores. Table B.2 shows the bottom-25 player impact scores.

Table B.1: 2014-2015 Top-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Steve Downie	RW	40.90	135	50	8	20	$1,000,000
Derek Dorsett	RW	27.07	77	77	-2	9	$2,000,000
Dion Phaneuf	D	26.14	65	97	11	77	$8,000,000
Kevin Bieksa	D	24.40	48	52	-3	59	$4,000,000
Scott Hartnell	LW	23.68	39	70	-11	20	$5,000,000
Antoine Roussel	LW	22.77	76	43	-3	29	$1,600,000
Cody McLeod	LW	22.35	76	95	2	18	$1,150,000
Evgeni Malkin	C	22.26	56	11	2	10	$9,500,000
Tom Wilson	RW	21.43	69	69	-1	13	$925,000
Mark Borowiecki	D	20.68	55	106	0	49	$600,000
Simon Despres	D	19.85	50	113	9	49	$900,000
Milan Lucic	LW	19.68	62	105	3	12	$6,000,000
Mark Stuart	D	19.13	37	74	6	74	$2,750,000
Brenden Dillon	D	19.03	34	60	-2	54	$1,250,000
Eric Gryba	D	17.77	62	62	5	18	$1,200,000
Brooks Orpik	D	17.68	32	127	5	91	$6,500,000
Mike Weber	D	17.21	39	72	-5	53	$1,500,000
Dustin Byfuglien	RW	17.17	53	102	2	21	$5,750,000
Brad Marchand	LW	17.13	38	21	9	7	$4,500,000
P.K. Subban	D	16.76	36	36	9	58	$7,000,000
Shea Weber	D	16.48	26	74	14	68	$14,000,000
Jori Lehtera	C	16.46	20	16	13	15	$3,250,000
Brandon Prust	LW	15.99	73	51	4	15	$2,500,000
David Clarkson	RW	15.73	39	80	-3	11	$4,750,000
Dalton Prout	D	15.69	40	76	-10	43	$1,050,000

B.2 2013-2014

The average player caused their team to receive 11.38 penalties during the 2013-2014 regular season. Table B.3 shows the top-25 player impact scores. Table B.4 shows the bottom-25 player impact scores.

B.3 2012-2013

The average player caused his team to receive 7.35 penalties during the 2012-2013 regular season. Table B.5 shows the top-25 player impact scores. Table B.6 shows the bottom-25 player impact scores.

Table B.2: 2014-2015 Bottom-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Frans Nielsen	C	-3.61	4	23	-1	42	$3,000,000
Patrick Marleau	LW	-3.37	4	39	-2	9	$7,000,000
Derek Stepan	C	-3.02	2	13	4	3	$3,850,000
Sean Monahan	C	-2.68	2	27	6	18	$850,000
Tyler Bozak	C	-2.06	4	34	-2	26	$4,000,000
Marcus Johansson	C	-1.85	0	28	0	6	$2,175,000
Adam Cracknell	RW	-1.85	2	41	-8	5	$600,000
Mikhail Grabovski	C	-1.79	2	7	3	14	$4,000,000
Mikael Granlund	C	-1.50	4	25	0	19	$900,000
Colton Sceviour	C	-1.40	0	30	2	16	$600,000
Cam Fowler	D	-1.38	0	24	2	44	$4,000,000
Patrick Kane	RW	-1.34	2	12	9	8	$6,500,000
Shayne Gostisbehere	D	-1.19	0	0	-2	1	$925,000
Jason Pominville	RW	-1.15	4	11	3	11	$6,000,000
Jonas Brodin	D	-1.15	2	10	13	37	$833,000
Bogdan Yakimov	C	-1.05	0	2	-1	0	$793,000
Derek Roy	C	-1.05	2	3	0	10	$1,000,000
Cam Atkinson	RW	-0.89	8	29	-7	16	$1,175,000
Mark Letestu	C	-0.88	0	8	-1	6	$1,300,000
Jamie McBain	D	-0.85	0	10	-1	10	$550,000
Mikhail Grigorenko	C	-0.83	0	6	-1	2	$925,000
Andrew Campbell	D	-0.80	0	2	-1	0	$550,000
Joe Thornton	C	-0.78	4	13	2	11	$6,750,000
Eriah Hayes	RW	-0.78	2	7	-2	4	$668,000
Jordan Martinook	LW	-0.76	0	4	1	0	$733,000

B.4 2011-2012

The average player caused his team to receive 11.51 penalties during the 2011-2012 regular season. Table B.7 shows the top-25 player impact scores. Table B.8 shows the bottom-25 player impact scores.

B.5 2010-2011

The average player caused his team to receive 12.24 penalties during the 2010-2011 regular season. Table B.9 shows the top-25 player impact scores. Table B.10 shows the bottom-25 player impact scores.

Table B.3: 2013-2014 Top-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Chris Neil	RW	62.58	211	253	-10	18	$2,100,000
Antoine Roussel	LW	54.26	209	146	-1	49	$625,000
Radko Gudas	D	53.34	152	273	2	138	$575,000
Dion Phaneuf	D	52.52	144	227	2	156	$5,500,000
Zac Rinaldo	C	48.65	153	231	-13	12	$750,000
Rich Clune	LW	47.08	166	132	-7	9	$525,000
Tom Sestito	LW	46.34	213	121	-14	12	$650,000
Tom Wilson	RW	46.12	151	197	1	13	$925,000
Zack Smith	C	44.55	111	175	-9	28	$1,500,000
David Perron	LW	42.49	90	116	-16	20	$3,500,000
Steve Downie	RW	41.28	106	70	1	18	$2,750,000
Dustin Byfuglien	RW	40.88	86	213	-20	80	$5,750,000
P.K. Subban	D	40.36	81	135	-4	125	$3,750,000
Mark Stuart	D	38.98	101	229	11	160	$1,800,000
Ryan Garbutt	LW	38.89	106	141	10	47	$600,000
Kevin Bieksa	D	38.76	104	144	-8	130	$5,000,000
David Backes	C	38.54	119	273	14	56	$4,750,000
Matt Carkner	D	38.05	149	58	-10	59	$1,500,000
Wayne Simmonds	RW	37.49	106	132	-4	35	$2,800,000
Kyle Quincey	D	35.99	88	87	-4	106	$4,000,000
Mark Giordano	D	35.68	63	73	12	103	$4,000,000
Matt Hendricks	LW	35.41	112	180	-11	59	$1,850,000
Scott Hartnell	LW	34.47	103	155	11	39	$6,000,000
Evander Kane	LW	34.47	66	173	-7	36	$4,500,000
Scottie Upshall	LW	34.39	73	134	1	25	$3,500,000

B.6 2009-2010

The average player caused his team to receive 13.20 penalties during the 2009-2010 regular season. Table B.11 shows the top-25 player impact scores. Table B.12 shows the bottom-25 player impact scores.

B.7 2008-2009

The average player caused his team to receive 14.62 penalties during the 2008-2009 regular season. Table B.13 shows the top-25 player impact scores. Table B.14 shows the bottom-25 player impact scores.

Table B.4: 2013-2014 Bottom-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Patrick Eaves	RW	-2.42	2	43	-7	5	$1,200,000
Markus Granlund	C	-1.56	0	3	2	1	$743,000
Jason Zucker	LW	-1.37	2	16	2	11	$900,000
Ryan Hamilton	LW	-1.22	0	3	-2	1	$600,000
Daniel Paille	LW	-1.20	6	70	9	30	$1,300,000
Elias Lindholm	C	-1.10	4	51	-14	17	$1,475,000
Mark Cundari	D	-0.92	0	10	-4	4	$600,000
Sean Monahan	C	-0.92	8	42	-20	23	$925,000
Aaron Palushaj	RW	-0.83	0	3	-1	0	$15,000
Chris Porter	LW	-0.66	0	68	-3	3	$650,000
Philip Samuelsson	D	-0.56	0	5	-1	7	$640,000
Freddie Hamilton	C	-0.56	2	24	-5	3	$640,000
Joe Piskula	D	-0.52	0	2	1	0	$550,000
Zach Trotman	D	-0.46	0	2	0	0	$690,000
Jerry D'Amigo	RW	-0.45	0	23	-1	4	$810,000
Derek Grant	C	-0.45	4	32	-3	11	$660,000
Andrew Alberts	D	-0.44	0	6	1	7	$600,000
Zach Boychuk	LW	-0.42	0	17	2	2	$550,000
Taylor Fedun	D	-0.37	0	1	-1	3	$675,000
Martin St Pierre	C	-0.34	0	1	0	1	$8,000
Denis Grebeshkov	D	-0.34	2	3	0	9	$285,000
Justin Florek	LW	-0.34	0	4	1	0	$690,000
Mike Santorelli	C	-0.32	6	23	9	44	$550,000
Zach Redmond	D	-0.29	0	7	1	11	$715,000
Ben Smith	RW	-0.26	2	39	2	58	$575,000

B.8 2007-2008

The average player caused his team to receive 15.26 penalties during the 2007-2008 regular season. Table B.15 shows the top-25 player impact scores. Table B.16 shows the bottom-25 player impact scores.

Table B.5: 2012-2013 Top-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Steve Ott	C	37.25	93	187	3	21	$3,200,000
Chris Neil	RW	34.54	144	206	0	10	$2,000,000
Zdeno Chara	D	33.59	70	101	14	64	$6,000,000
Brandon Prust	LW	33.03	110	87	11	26	$3,000,000
Colton Orr	RW	33.03	155	78	4	17	$1,000,000
Rich Clune	LW	32.37	113	159	3	6	$525,000
David Backes	C	31.11	62	158	5	41	$3,750,000
B.J. Crombeen	RW	30.27	112	44	4	15	$1,050,000
Wayne Simmonds	RW	28.71	82	72	-7	16	$2,000,000
Alexandre Burrows	RW	27.46	52	52	15	15	$2,000,000
Scott Hartnell	LW	27.41	70	68	-5	21	$3,200,000
Brenden Dillon	D	26.74	65	133	1	74	$690,000
Cody McLeod	LW	26.43	83	106	4	17	$1,150,000
Zenon Konopka	C	26.43	117	23	-4	9	$850,000
Kimmo Timonen	D	26.40	36	31	3	79	$3,000,000
Adam McQuaid	D	24.76	60	62	0	43	$1,400,000
P.K. Subban	D	24.63	57	51	12	49	$2,000,000
Evander Kane	LW	24.55	80	147	-3	22	$3,000,000
Mike Brown	RW	24.54	123	87	-7	13	$725,000
Jay Harrison	D	24.23	51	83	-10	110	$750,000
Ryane Clowe	LW	23.60	93	99	1	19	$4,000,000
Keith Yandle	D	23.59	54	18	4	40	$5,000,000
Dion Phaneuf	D	23.47	65	131	-4	91	$6,500,000
Milan Lucic	LW	23.08	75	139	8	16	$4,250,000
Zack Smith	C	23.01	56	97	-9	22	$775,000

Table B.6: 2012-2013 Bottom-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Jared Spurgeon	D	-2.53	4	46	1	55	$535,000
Justin Braun	D	-2.16	6	53	-5	51	$1,000,000
Benn Ferriero	RW	-1.90	0	4	0	2	$700,000
Mikhail Grigorenko	C	-1.51	0	2	-1	2	$925,000
Cam Atkinson	RW	-1.49	4	24	9	16	$838,000
Ryan Nugent-Hopkins	C	-1.32	8	27	3	27	$925,000
Jake Gardiner	D	-1.22	0	13	0	15	$875,000
Nathan Beaulieu	D	-1.20	0	3	5	4	$925,000
Francis Wathier	LW	-1.05	0	5	0	0	$154,000
Erik Gustafsson	D	-0.91	2	15	-1	39	$576,500
Andreas Lilja	D	-0.75	0	12	-1	7	$201,000
Chad Ruhwedel	D	-0.75	0	7	0	8	$925,000
Stephane Da Costa	C	-0.67	0	8	-3	1	$234,000
David Rundblad	D	-0.65	0	7	-5	9	$900,000
Joe Colborne	C	-0.64	2	11	-1	0	$875,000
Filip Forsberg	C	-0.59	0	3	-5	0	$925,000
Joakim Andersson	C	-0.58	8	17	2	12	$638,000
Matt Moulson	LW	-0.53	4	24	-3	28	$3,000,000
Cody Goloubef	D	-0.53	0	10	-3	7	$875,000
John-Michael Liles	D	-0.50	4	45	-1	52	$4,250,000
Kevin Klein	D	-0.50	9	66	-1	89	$1,350,000
Jason Akeson	RW	-0.43	0	0	2	0	$715,000
Ed Jovanovski	D	-0.43	0	3	-4	8	$4,250,000
Brett Carson	D	-0.39	0	4	-1	12	$139,000
Rickard Rakell	C	-0.38	0	4	-2	0	$925,000

Table B.7: 2011-2012 Top-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Derek Dorsett	RW	57.10	235	199	-11	40	$575,000
Zac Rinaldo	C	55.30	232	175	-1	19	$560,000
Nick Foligno	LW	53.78	124	196	2	30	$1,550,000
Scott Hartnell	LW	52.72	136	188	19	37	$3,700,000
Chris Neil	RW	52.72	178	271	-10	20	$2,000,000
P.K. Subban	D	51.96	119	105	9	113	$875,000
Steve Ott	C	49.63	156	278	5	27	$3,300,000
Zenon Konopka	C	45.07	193	54	-4	18	$700,000
Shawn Thornton	LW	44.61	154	91	-7	12	$800,000
Corey Perry	RW	44.03	127	69	-7	44	$5,375,000
Milan Lucic	LW	43.96	135	201	7	26	$4,000,000
Steve Downie	RW	42.69	137	105	-6	25	$1,900,000
Brandon Dubinsky	C	41.81	110	207	16	36	$3,750,000
Kyle Quincey	D	41.18	89	101	-1	94	$3,250,000
Cody McLeod	LW	40.91	164	123	0	13	$1,200,000
Raffi Torres	LW	40.91	83	128	3	21	$1,750,000
Sheldon Souray	D	40.29	73	55	11	90	$2,400,000
James Neal	LW	40.00	87	108	6	15	$3,500,000
Mark Stuart	D	39.76	98	198	-4	182	$1,600,000
Cal Clutterbuck	RW	39.73	103	288	-4	31	$1,500,000
Brenden Morrow	LW	38.97	97	130	1	33	$4,100,000
Pierre-Alexandre Parenteau	RW	38.66	89	99	-8	23	$1,250,000
Shane O'Brien	D	38.34	105	138	2	86	$1,100,000
Brandon Prust	LW	38.28	156	144	-1	51	$800,000
Patrick Kaleta	RW	38.12	116	139	-5	46	$955,000

Table B.8: 2011-2012 Bottom-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Adam Henrique	C	-2.80	7	83	8	57	$588,000
Antti Miettinen	RW	-2.06	0	38	-5	11	$949,000
Brandon Manning	D	-1.60	0	6	1	2	$715,000
Mark Scheifele	C	-0.91	0	3	0	1	$925,000
Radek Martinek	D	-0.89	0	5	-3	7	$2,200,000
Mark Letestu	C	-0.87	8	60	-9	19	$650,000
Brett Sterling	LW	-0.85	0	1	-1	0	$242,000
Chad Rau	C	-0.83	0	2	-1	5	$650,000
Taylor Chorney	D	-0.69	0	0	-1	4	$735,000
Greg Nemisz	RW	-0.64	0	3	1	1	$875,000
Kyle Wilson	C	-0.58	0	7	-1	1	$156,000
Chris Porter	LW	-0.53	9	121	-1	4	$600,000
Patrice Cormier	C	-0.53	0	12	1	0	$613,000
Ben Smith	RW	-0.47	0	6	-5	5	$605,000
Stephen Gionta	C	-0.46	0	1	1	1	$525,000
Steven Zalewski	C	-0.42	0	11	-2	1	$136,000
Zach Boychuk	LW	-0.40	0	16	-3	2	$788,000
Jeff Taffe	LW	-0.40	0	7	2	2	$226,000
Brandon Saad	LW	-0.36	0	0	0	1	$618,000
J.T. Brown	RW	-0.35	0	2	2	3	$925,000
Brenden Dillon	D	-0.29	0	4	0	3	$640,000
Philippe Cornet	LW	-0.28	0	2	0	0	$28,000
Mark Mancari	RW	-0.24	0	5	0	0	$37,000
Peter Regin	C	-0.23	2	6	3	3	$1,050,000
Brett Carson	D	-0.21	0	0	-2	1	$329,000

Table B.9: 2010-2011 Top-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Zenon Konopka	C	64.05	284	109	-14	51	$700,000
Theo Peckham	D	60.18	198	196	-5	123	$1,075,000
Chris Neil	RW	57.46	210	258	-14	21	$2,000,000
Derek Dorsett	RW	54.24	184	195	-15	37	$550,000
Cody McLeod	LW	53.56	189	147	-7	20	$1,000,000
Steve Ott	C	52.30	183	252	-9	44	$2,950,000
Brad Staubitz	RW	47.78	173	133	-5	18	$600,000
Steve Downie	RW	46.73	171	94	8	21	$1,850,000
Corey Perry	RW	45.02	104	64	9	41	$5,375,000
Paul Gaustad	C	44.81	101	128	7	44	$2,500,000
Brent Burns	D	44.69	98	133	-10	106	$3,800,000
P.K. Subban	D	44.23	124	110	-8	106	$875,000
Sean Avery	LW	42.43	174	115	-4	24	$4,000,000
Scott Hartnell	LW	42.39	142	168	14	38	$4,200,000
Cody McCormick	C	42.32	142	108	2	60	$500,000
Jared Boll	RW	40.70	182	144	-2	18	$700,000
Alexander Semin	RW	39.29	71	27	22	6	$6,000,000
Matt Cooke	LW	38.97	117	189	14	38	$1,800,000
B.J. Crombeen	RW	38.48	154	93	-18	15	$885,000
Milan Lucic	LW	38.25	121	167	28	22	$4,000,000
Steve Montador	D	37.22	83	86	16	138	$1,550,000
Ryan O'Byrne	D	36.98	75	179	-7	131	$1,400,000
Travis Hamonic	D	36.89	103	118	4	117	$875,000
Jarkko Ruutu	LW	36.18	97	131	-2	29	$1,300,000
Zack Smith	C	36.02	120	129	-11	27	$583,000

Table B.10: 2010-2011 Bottom-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Jerome Samson	RW	-2.48	0	13	0	3	$500,000
Jake Muzzin	D	-2.10	0	22	-2	2	$615,000
Mats Zuccarello	LW	-2.08	4	56	3	19	$850,000
Kyle Wellwood	C	-1.80	0	8	10	7	$650,000
Yannick Weber	D	-1.80	14	44	0	33	$638,000
Nicklas Lidstrom	D	-1.43	20	49	-2	92	$6,200,000
Andrei Loktionov	C	-1.31	2	6	2	8	$324,000
Christopher Tanev	D	-1.27	0	10	0	32	$900,000
Spencer Machacek	RW	-1.20	0	22	-2	2	$165,000
Joel Perrault	C	-1.04	0	11	-1	2	$559,000
Derek Smith	D	-0.92	0	5	3	11	$152,000
Noah Welch	D	-0.86	0	10	-1	3	$136,000
Jared Spurgeon	D	-0.85	2	37	-1	45	$510,000
Oskars Bartulis	D	-0.63	4	8	-4	13	$571,000
Stephane Da Costa	C	-0.57	0	3	-1	0	$78,000
Roman Wick	RW	-0.56	0	2	-4	3	$131,000
Taylor Chorney	D	-0.54	4	13	-5	17	$785,000
Ryan Potulny	C	-0.48	0	5	-1	1	$151,000
Brayden Schenn	C	-0.44	0	12	-1	1	$900,000
Michael Grabner	RW	-0.39	10	21	14	26	$765,000
Patrick Rissmiller	LW	-0.38	0	17	-1	2	$1,312,000
Joe Colborne	C	-0.36	0	0	1	1	$875,000
Nick Leddy	D	-0.33	4	26	-3	46	$900,000
Travis Morin	C	-0.31	0	1	0	2	$510,000
Cody Hodgson	C	-0.29	0	4	1	2	$875,000

Table B.11: 2009-2010 Top-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Steve Downie	RW	75.79	208	140	14	20	$600,000
Colton Orr	RW	60.04	239	119	-4	26	$1,000,000
Zenon Konopka	C	57.60	265	109	-11	23	$500,000
Matt Carkner	D	56.81	190	127	0	125	$500,000
Scott Hartnell	LW	55.15	155	138	-6	30	$4,200,000
Chris Neil	RW	52.84	175	245	-1	8	$2,000,000
Sean Avery	LW	52.65	160	145	0	17	$4,000,000
Daniel Carcillo	LW	51.28	207	194	5	17	$938,000
Steve Ott	C	49.47	153	251	-14	19	$1,500,000
Alexandre Burrows	RW	48.73	121	97	34	54	$2,000,000
B.J. Crombeen	RW	48.02	168	82	-5	31	$860,000
Corey Perry	RW	47.46	111	93	0	34	$6,500,000
Evgeni Malkin	C	47.40	100	58	-6	25	$9,000,000
Ryane Clowe	LW	46.88	131	157	0	33	$3,500,000
Cody McLeod	LW	45.22	138	197	-13	28	$900,000
Michal Rozsival	D	44.58	78	136	3	130	$6,000,000
Andy Sutton	D	44.47	107	197	-10	204	$3,438,000
Cam Janssen	RW	44.28	190	73	-3	2	$550,000
Rene Bourque	RW	43.61	88	91	7	44	$1,400,000
Mike Rupp	LW	41.71	120	198	5	16	$850,000
Zack Stortini	RW	41.68	155	144	3	30	$700,000
Shawn Thornton	LW	41.19	141	110	-9	17	$550,000
Jamal Mayers	RW	40.73	131	69	-3	30	$1,400,000
Jarkko Ruutu	LW	39.86	121	134	-2	45	$1,300,000
Wayne Simmonds	RW	39.74	116	126	22	32	$585,000

Table B.12: 2009-2010 Bottom-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Milan Hejduk	RW	-3.32	10	30	6	21	$4,000,000
Brad Richards	C	-3.02	14	16	-12	22	$7,800,000
Benn Ferriero	RW	-1.72	8	17	4	3	$635,000
Brock Trotter	LW	-1.49	0	0	-1	2	$79,000
Warren Peters	C	-1.18	2	27	1	4	$206,000
Bryan Bickell	LW	-1.17	5	22	4	2	$500,000
MacGregor Sharp	C	-0.98	0	2	0	0	$91,000
Mathieu Darche	LW	-0.95	4	39	2	12	$458,000
Nolan Baumgartner	D	-0.91	2	10	7	26	$416,000
Ivan Vishnevskiy	D	-0.90	0	0	-2	1	$84,500
Nick Spaling	C	-0.88	0	13	3	9	$738,000
Mikkel Boedker	LW	-0.88	0	17	2	1	$875,000
Maksim Mayorov	LW	-0.80	0	10	-1	1	$96,000
Pavol Demitra	RW	-0.62	0	18	3	12	$4,000,000
Chris Conner	RW	-0.50	0	6	-1	4	$500,000
Kaspars Daugavins	LW	-0.47	0	2	0	0	$8,000
David Desharnais	C	-0.47	0	2	-1	2	$525,000
Casey Borer	D	-0.45	0	4	-1	5	$68,000
Eric Tangradi	LW	-0.44	0	3	0	0	$728,000
Brandon Sutter	C	-0.44	2	55	-1	45	$875,000
Ryan Vesce	RW	-0.41	0	13	-1	2	$205,000
Ben Guite	RW	-0.36	4	14	-3	4	$162,000
Steven Zalewski	C	-0.33	0	1	-2	1	$81,000
T.J. Hensick	C	-0.32	0	2	0	1	$309,500
Andrei Loktionov	C	-0.31	0	0	0	0	$604,000

Table B.13: 2008-2009 Top-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Daniel Carcillo	LW	62.22	254	151	-15	20	$850,000
Mike Komisarek	D	59.84	121	191	0	207	$1,900,000
Scott Hartnell	LW	59.68	143	104	14	36	$4,700,000
Shane O'Brien	D	59.67	196	81	5	51	$1,025,000
Evgeny Artyukhin	RW	59.00	151	249	1	16	$890,000
David Backes	C	57.22	165	204	-3	52	$2,500,000
Cody McLeod	LW	53.26	162	194	-11	32	$523,000
Jarkko Ruutu	LW	51.96	144	148	0	37	$1,251,000
Ryan Getzlaf	C	51.51	121	134	5	43	$4,500,000
Alexandre Burrows	RW	51.04	150	68	23	38	$525,000
Steve Ott	C	50.70	135	220	3	23	$1,350,000
Eric Boulton	LW	50.67	176	77	-3	13	$600,000
Pavel Kubina	D	50.02	94	91	-15	133	$5,000,000
Steve Montador	D	49.63	143	95	17	60	$800,000
Boris Valabik	D	49.44	132	78	-14	68	$729,000
Colton Orr	RW	48.87	193	133	-15	14	$550,000
Raitis Ivanans	LW	48.29	145	125	-8	9	$600,000
Mike Commodore	D	48.17	100	201	11	162	$4,300,000
Ben Eager	LW	45.44	161	95	1	11	$601,000
Dominic Moore	C	44.66	92	75	-2	37	$900,000
Jordin Tootoo	RW	44.23	124	129	-15	14	$975,000
Corey Perry	RW	44.23	109	109	10	13	$4,500,000
B.J. Crombeen	RW	43.89	148	87	-9	14	$550,000
Arron Asham	RW	43.74	155	143	0	10	$640,000
Rob Blake	D	43.65	110	66	15	14	$5,000,000

Table B.14: 2008-2009 Bottom-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Mike Sillinger	C	-4.43	0	7	-5	4	$2,300,000
Niklas Hjalmarsson	D	-1.69	0	15	4	22	$644,000
Chris Durno	LW	-1.62	0	6	0	0	$15,000
Anze Kopitar	C	-1.41	32	85	-17	56	$765,000
Clay Wilson	D	-1.27	0	6	-3	1	$63,000
Dustin Jeffrey	C	-1.20	0	8	4	3	$500,000
Kyle Cumiskey	D	-0.95	0	2	-2	3	$475,000
Kevin Quick	D	-0.86	0	1	0	3	$38,000
Josh Hennessy	C	-0.82	0	1	0	0	$8,000
Niklas Hagman	LW	-0.76	4	22	-5	24	$3,000,000
Joseph Motzko	RW	-0.66	0	1	1	1	$36,000
Marian Gaborik	RW	-0.62	2	15	3	5	$7,500,000
Kevin Porter	C	-0.59	4	37	-2	13	$895,000
David Van Der Gulik	LW	-0.54	0	6	-1	2	$475,000
Kyle Greentree	LW	-0.51	0	1	-1	1	$18,000
Jiri Tlusty	LW	-0.45	0	5	0	5	$850,000
Ben Lovejoy	D	-0.42	0	1	0	1	$638,000
Chris Porter	LW	-0.32	0	3	-1	3	$875,000
Janis Sprukts	C	-0.31	0	0	0	0	$3,000
Patrick Rissmiller	LW	-0.30	0	5	-2	0	$290,000
Raymond Sawada	RW	-0.27	0	12	-1	1	$44,000
Ryan Potulny	C	-0.25	0	5	2	1	$59,000
Tim Stapleton	C	-0.23	0	2	-3	3	$27,000
Trevor Lewis	C	-0.23	0	4	0	4	$850,000
Karl Alzner	D	-0.20	2	23	-1	54	$875,000

Table B.15: 2007-2008 Top-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Daniel Carcillo	LW	83.08	324	109	1	10	$525,000
Dion Phaneuf	D	76.26	182	194	12	88	$942,000
Chris Neil	RW	72.73	199	204	-3	10	$1,100,000
Jared Boll	RW	61.21	226	135	-4	13	$545,000
David Clarkson	RW	59.74	183	150	1	12	$555,000
Alexandre Burrows	RW	58.35	179	80	11	45	$475,000
Adam Burish	RW	57.07	214	89	-13	66	$575,000
Chris Pronger	D	55.85	128	74	-1	99	$6,250,000
Pavel Kubina	D	55.05	116	121	5	166	$5,000,000
Shane O'Brien	D	54.53	154	128	-2	100	$875,000
Scott Hartnell	LW	52.61	159	110	2	32	$5,200,000
Cory Sarich	D	52.36	135	157	2	64	$3,900,000
Zack Stortini	RW	51.69	201	99	3	17	$475,000
Corey Perry	RW	51.25	108	95	12	11	$494,000
Adam Foote	D	49.39	107	93	2	148	$4,600,000
Steve Staios	D	49.35	121	81	-14	187	$2,900,000
Tuomo Ruutu	LW	48.50	91	171	4	19	$2,250,000
Zdeno Chara	D	48.15	114	223	14	78	$7,500,000
Steve Ott	C	47.27	147	182	2	26	$800,000
Nick Boynton	D	45.25	125	93	-9	93	$2,903,000
Riley Cote	LW	45.03	202	60	2	4	$476,000
George Parros	RW	44.52	183	91	3	9	$525,000
Kris Draper	C	43.87	68	90	-2	25	$2,128,000
Jarkko Ruutu	LW	42.65	138	134	3	19	$1,150,000
Jay Bouwmeester	D	42.36	72	105	-5	118	$2,250,000

Table B.16: 2007-2008 Bottom-25 Player Impact Scores For Penalties

Name	Position	Penalty Impact	PIM	Hits	+/-	Blocked Shots	Salary
Frans Nielsen	C	-1.78	0	4	1	8	$560,000
Dustin Boyd	C	-1.51	6	39	-11	14	$539,000
Tomas Plihal	C	-1.39	4	20	4	5	$353,000
Jack Skille	RW	-1.28	0	15	1	4	$850,000
T.J. Hensick	C	-0.93	2	4	-4	6	$318,000
Adam Pineault	RW	-0.90	0	1	-2	0	$13,000
Dan Girardi	D	-0.84	14	179	0	123	$550,000
Claude Giroux	C	-0.76	0	0	-2	1	$850,000
Rob Schremp	C	-0.75	0	0	-1	0	$25,000
Pavol Demitra	RW	-0.49	24	17	9	48	$4,500,000
Thomas Pock	D	-0.48	0	1	-2	3	$79,000
Jon Sim	LW	-0.47	2	4	-1	0	$1,000,000
Jay Leach	D	-0.47	0	1	-1	2	$11,000
Brendan Bell	D	-0.37	0	0	-2	2	$64,000
Jonathan Filewich	RW	-0.28	0	2	-2	1	$52,000
Martin St Pierre	C	-0.23	0	3	-3	1	$45,000
Darryl Boyce	C	-0.22	0	1	0	0	$213,000
Lawrence Nycholat	D	-0.18	0	4	1	2	$18,000
Ilya Zubov	C	-0.18	0	3	0	0	$5,000
Kyle Greentree	LW	-0.16	0	2	-1	1	$23,000
Chris Higgins	LW	-0.15	22	76	0	65	$1,500,000
Brandon Nolan	C	-0.15	0	8	-2	1	$36,000
Lukas Kaspar	LW	-0.13	0	3	-2	1	$56,000
Pascal Pelletier	LW	-0.13	0	12	-2	2	$40,000
Connor James	LW	-0.11	2	3	-2	0	$76,000

Appendix C

Player Rankings: Wins

We show net player impact scores over a single season with respect to the probability of winning the game. We compute action values for the probability of winning using Equation 7.4. Again, average player values are found by taking the average of all players' net impact value. Players with high impact scores perform actions that make their team more likely to win. Players with low impact scores perform actions that make their team more likely to lose.

C.1 2014-2015

The average player generates 0.70 wins during the first 512 games of the 2014-2015 regular season. Table C.1 shows the top-25 player impact scores. Table C.2 shows the bottom-25 player impact scores.

C.2 2013-2014

The average player generated 1.58 wins for his team during the 2013-2014 regular season. Table C.3 shows the top-25 player impact scores. Table C.4 shows the bottom-25 player impact scores.

C.3 2012-2013

The average player generated 0.97 wins for his team during the 2012-2013 regular season. Table C.5 shows the top-25 player impact scores. Table C.6 shows the bottom-25 player impact scores.

Table C.1: 2014-2015 Top-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Jori Lehtera	C	6.48	8	25	57	21	13	$3,250,000
Jonathan Toews	C	5.89	13	29	86	19	9	$6,500,000
Vladimir Tarasenko	RW	5.56	20	37	121	20	18	$900,000
Jason Spezza	C	5.17	6	25	72	25	-11	$4,000,000
Henrik Zetterberg	LW	5.14	7	30	113	21	-1	$7,500,000
Kyle Okposo	RW	5.03	8	29	116	18	-4	$3,500,000
Joe Thornton	C	4.44	8	30	69	28	2	$6,750,000
Joe Pavelski	C	4.37	16	29	126	22	5	$6,000,000
Patrick Marleau	LW	4.08	7	27	107	19	-2	$7,000,000
Logan Couture	C	4.04	13	29	110	16	2	$6,000,000
Gustav Nyquist	RW	3.96	14	22	84	15	-7	$1,050,000
Jaromir Jagr	RW	3.73	5	20	78	25	-12	$3,500,000
Steven Stamkos	C	3.73	16	33	103	14	-2	$8,000,000
Ryan Kesler	C	3.69	12	27	109	20	-1	$5,000,000
Brent Burns	D	3.64	10	27	103	16	-3	$5,760,000
Nathan MacKinnon	C	3.56	5	21	103	19	-9	$925,000
Sean Monahan	C	3.50	11	22	97	23	6	$925,000
Frans Nielsen	C	3.47	6	17	73	23	-1	$3,000,000
Sidney Crosby	C	3.31	10	37	96	18	12	$12,000,000
Claude Giroux	C	3.29	11	41	132	21	12	$10,000,000
Martin Hanzal	C	3.22	6	17	57	16	1	$3,250,000
Tomas Plekanec	C	3.21	10	23	89	15	6	$5,000,000
Alex Ovechkin	RW	3.20	16	28	163	18	5	$10,000,000
Kris Letang	D	3.20	8	23	94	17	7	$7,250,000
Pavel Datsyuk	C	3.10	13	25	67	16	4	$10,000,000

C.4 2011-2012

The average player generated 1.56 wins for his team during the 2011-2012 regular season. Table C.7 shows the top-25 player impact scores. Table C.8 shows the bottom-25 player impact scores.

C.5 2010-2011

The average player generates 1.43 wins during the 2010-2011 regular season. Table C.9 shows the top-25 player impact scores. Table C.10 shows the bottom-25 player impact scores.

C.6 2009-2010

The average player contributed to 1.42 wins for his team during the 2009-2010 regular season. Table C.11 shows the top-25 player impact scores. Table C.12 shows the bottom-25 player impact scores.

Table C.2: 2014-2015 Bottom-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Chris Neil	RW	-2.05	4	7	18	4	3	$2,100,000
Jan Hejda	D	-2.02	0	5	30	14	-5	$3,250,000
Brenden Dillon	D	-1.59	0	5	34	5	-2	$1,250,000
Willie Mitchell	D	-1.54	1	2	30	9	-6	$4,250,000
Hampus Lindholm	D	-1.52	4	15	50	14	11	$925,000
Ryan Suter	D	-1.39	1	22	62	12	3	$11,000,000
Matt Hunwick	D	-1.30	0	7	28	10	8	$600,000
Joe Vitale	C	-1.28	3	6	27	17	-3	$950,000
Brad Malone	C	-1.22	0	0	22	12	-10	$600,000
Adam Larsson	D	-1.21	1	4	21	4	-3	$900,000
Manny Malhotra	C	-1.10	0	1	28	9	-4	$850,000
Jay Bouwmeester	D	-1.05	1	4	38	8	-5	$5,000,000
Jesse Joensuu	LW	-0.97	2	4	18	9	-8	$1,000,000
Erik Gudbranson	D	-0.96	1	5	48	0	2	$2,250,000
David Schlemko	D	-0.92	1	4	24	1	-5	$1,275,000
Rob Scuderi	D	-0.92	0	5	18	5	5	$4,000,000
Jeff Petry	D	-0.89	3	8	67	16	-20	$3,075,000
Alex Goligoski	D	-0.88	1	15	42	12	1	$4,800,000
Chris Phillips	D	-0.86	0	2	20	6	-1	$2,500,000
Tim Jackman	RW	-0.85	2	4	40	5	-2	$638,000
Derek MacKenzie	C	-0.81	3	5	28	7	-6	$1,300,000
Matt Stajan	C	-0.78	1	4	12	5	1	$3,625,000
Jim Slater	C	-0.77	1	3	22	3	2	$1,600,000
Jason Garrison	D	-0.76	3	17	56	11	10	$5,000,000
Matthew Carle	D	-0.73	3	8	38	10	7	$5,750,000

C.7 2008-2009

The average player contributed to 1.39 wins for his team during the 2008-2009 regular season. Table C.13 shows the top-25 player impact scores. Table C.14 shows the bottom-25 player impact scores.

C.8 2007-2008

The average player generated 1.37 wins for his team during the 2007-2008 regular season. Table C.15 shows the top-25 player impact scores. Table C.16 shows the bottom-25 player impact scores.

Table C.3: 2013-2014 Top-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Joe Pavelski	C	10.77	41	79	225	56	23	$4,000,000
Jonathan Toews	C	10.60	27	67	192	51	25	$6,500,000
Jason Spezza	C	10.24	23	66	223	47	-26	$5,000,000
Marian Hossa	RW	9.56	29	57	238	73	26	$7,900,000
Sidney Crosby	C	9.49	36	104	259	41	18	$12,000,000
John Tavares	C	9.48	24	66	188	55	-6	$5,000,000
Claude Giroux	C	9.16	28	86	223	43	7	$5,000,000
Valtteri Filppula	C	8.99	25	58	131	53	5	$4,000,000
Nicklas Backstrom	C	8.73	18	79	196	54	-20	$6,000,000
Patrick Sharp	LW	8.70	34	77	306	41	12	$6,500,000
Patrick Marleau	LW	8.64	33	70	285	50	0	$6,900,000
Anze Kopitar	C	8.63	29	70	200	42	34	$7,500,000
Zach Parise	LW	8.54	29	56	245	46	10	$12,000,000
Jamie Benn	LW	8.38	34	79	279	70	21	$5,000,000
Ryan Johansen	C	8.36	33	63	237	39	3	$810,000
Max Pacioretty	LW	8.24	39	60	270	28	8	$4,000,000
Derek Stepan	C	7.79	17	57	199	50	12	$2,300,000
T.J. Oshie	RW	7.75	21	60	152	62	19	$4,000,000
Tyler Seguin	C	7.63	37	84	294	67	16	$4,500,000
Matt Duchene	C	7.49	23	70	217	40	8	$3,750,000
Bryan Little	C	7.41	23	64	170	38	8	$4,000,000
Brad Richards	C	7.25	20	51	259	33	-8	$9,000,000
Sean Monahan	C	7.23	22	34	140	26	-20	$925,000
Ryan O'Reilly	C	7.21	28	64	201	83	-1	$6,500,000
Patrice Bergeron	C	7.18	30	62	243	49	38	$4,550,000

Table C.4: 2013-2014 Bottom-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Mike Weber	D	-1.98	1	9	47	11	-29	$1,500,000
Rich Clune	LW	-1.95	3	7	29	11	-7	$525,000
Willie Mitchell	D	-1.84	1	12	73	10	14	$3,500,000
Andrew MacDonald	D	-1.66	4	28	92	23	-22	$575,000
Jacob Trouba	D	-1.61	10	29	121	31	4	$925,000
Matt Carkner	D	-1.49	0	3	50	9	-10	$1,500,000
Mark Fraser	D	-1.32	1	2	11	7	-15	$1,275,000
Colton Orr	RW	-1.31	0	0	12	3	-3	$925,000
Tim Gleason	D	-1.20	1	6	43	9	-21	$4,500,000
Matt Greene	D	-1.17	2	6	38	1	6	$3,250,000
Mike Brown	RW	-1.10	2	5	45	8	-9	$725,000
Travis Moen	LW	-1.06	2	12	56	15	2	$1,850,000
Daniel Cleary	RW	-1.03	4	8	64	12	-11	$1,750,000
Mark Pysyk	D	-0.97	1	7	51	12	-11	$900,000
Tom Sestito	LW	-0.96	5	9	31	8	-14	$650,000
Chuck Kobasew	RW	-0.95	2	2	37	10	1	$434,000
Patrick Maroon	LW	-0.84	11	29	93	15	11	$575,000
Tom Gilbert	D	-0.81	3	28	93	16	-5	$900,000
Ville Leino	LW	-0.81	0	15	38	23	-16	$4,000,000
Barret Jackman	D	-0.71	3	15	83	25	11	$3,250,000
Johnny Oduya	D	-0.70	3	16	81	15	11	$3,300,000
Julien Brouillette	D	-0.68	1	2	4	1	3	$550,000
T.J. Galiardi	LW	-0.66	4	17	100	23	-13	$1,250,000
Tyson Strachan	D	-0.65	0	2	7	4	-2	$550,000
Sergei Gonchar	D	-0.63	2	22	89	10	-12	$5,000,000

Table C.5: 2012-2013 Top-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Jonathan Toews	C	8.53	23	48	143	56	28	$6,000,000
Patrick Kane	RW	7.76	23	55	138	36	11	$6,000,000
Pavel Datsyuk	C	7.00	15	49	107	56	21	$6,700,000
Sidney Crosby	C	6.91	15	56	124	15	26	$7,500,000
Mikko Koivu	C	6.58	11	37	127	26	2	$5,400,000
Logan Couture	C	6.53	21	37	151	31	7	$2,750,000
John Tavares	C	6.42	28	47	162	27	-2	$4,000,000
Dustin Brown	RW	6.32	18	29	142	17	6	$3,500,000
Claude Giroux	C	6.13	13	48	137	17	-7	$3,500,000
Zach Parise	LW	6.13	18	38	182	24	2	$12,000,000
Matt Duchene	C	5.95	17	43	132	44	-12	$3,250,000
Corey Perry	RW	5.65	15	36	128	18	10	$4,875,000
Derek Stepan	C	5.54	18	44	108	34	25	$875,000
Jason Pominville	Rw	5.51	14	34	118	38	1	$5,500,000
Mark Letestu	C	5.46	13	27	92	27	7	$600,000
Rick Nash	LW	5.41	21	42	176	19	16	$7,600,000
Andy McDonald	LW	5.41	7	21	86	13	-2	$4,200,000
Artem Anisimov	C	5.28	11	18	68	17	-6	$1,875,000
Alexandre Burrows	RW	5.19	13	24	140	19	15	$2,000,000
Sam Gagner	C	5.17	14	38	113	23	-6	$3,200,000
Jamie Benn	LW	4.99	12	33	110	41	-12	$4,500,000
Jakob Silfverberg	RW	4.97	10	19	134	20	9	$900,000
Alexander Steen	LW	4.86	8	27	129	13	5	$3,567,000
Steven Stamkos	C	4.81	29	57	157	24	-4	$8,000,000
Tyler Bozak	C	4.74	12	28	61	37	-1	$1,400,000

Table C.6: 2012-2013 Bottom-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
John Erskine	D	-1.61	3	6	32	4	10	$1,500,000
Tuomo Ruutu	LW	-1.16	4	9	30	12	-6	$4,000,000
Roman Polak	D	-1.04	1	6	39	8	-2	$2,450,000
Paul Bissonnette	LW	-1.03	0	6	13	1	2	$725,000
Brad Stuart	D	-0.96	0	6	39	15	4	$3,600,000
Ryan O'Byrne	D	-0.92	2	6	26	10	-4	$2,000,000
Jay Harrison	D	-0.88	3	10	54	12	-10	$750,000
Kevin Westgarth	RW	-0.87	2	4	16	4	1	$700,000
Nicklas Grossmann	D	-0.84	1	4	21	6	-1	$3,500,000
Michael Del Zotto	D	-0.74	3	21	81	10	6	$2,200,000
Chris Thorburn	RW	-0.71	2	4	13	3	-5	$850,000
Ian Cole	D	-0.70	0	1	10	1	-4	$875,000
Adam Larsson	D	-0.68	0	6	30	17	4	$925,000
Erik Gudbranson	D	-0.67	0	4	49	2	-22	$900,000
Braydon Coburn	D	-0.66	1	5	38	12	-10	$4,000,000
Cam Barker	D	-0.66	0	2	19	0	-3	$800,000
Colton Gillies	C	-0.66	1	2	17	3	1	$650,000
Fedor Tyutin	D	-0.66	4	22	56	9	9	$4,000,000
Zenon Konopka	C	-0.63	0	0	18	5	-4	$850,000
George Parros	RW	-0.63	1	2	16	5	-15	$925,000
Ladislav Smid	D	-0.61	1	4	30	8	-1	$2,250,000
Jordie Benn	D	-0.60	1	6	31	8	-4	$525,000
Sami Salo	D	-0.60	2	17	48	13	5	$4,000,000
Tanner Glass	LW	-0.59	1	2	38	1	-11	$1,100,000
Robyn Regehr	D	-0.58	0	4	27	1	-4	$4,000,000

Table C.7: 2011-2012 Top-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Zach Parise	LW	12.15	31	69	293	65	-5	$6,000,000
Evgeni Malkin	C	11.91	50	109	339	52	18	$9,000,000
Jason Spezza	C	11.09	34	84	232	64	11	$8,000,000
Loui Eriksson	RW	10.88	26	71	187	50	18	$4,100,000
Teemu Selanne	RW	9.98	26	66	209	28	-1	$4,000,000
John Tavares	C	9.82	31	81	286	99	-6	$900,000
Joe Pavelski	C	9.57	31	61	269	73	18	$4,000,000
Ryan Kesler	C	9.54	22	49	220	43	11	$5,000,000
Claude Giroux	C	9.51	28	93	242	50	6	$2,750,000
Marian Gaborik	RW	9.40	41	76	276	30	15	$7,500,000
Tyler Seguin	C	9.19	29	67	242	30	34	$900,000
Ilya Kovalchuk	RW	9.08	37	83	310	42	-9	$6,000,000
Radim Vrbata	RW	8.83	35	61	230	36	22	$3,000,000
Patrice Bergeron	C	8.63	22	64	191	55	36	$5,900,000
Stephen Weiss	C	8.57	20	57	149	56	5	$4,000,000
Jason Pominville	RW	8.42	30	73	235	45	-7	$5,500,000
Jonathan Toews	C	8.36	29	57	185	82	17	$6,000,000
Steven Stamkos	C	8.30	60	97	303	42	7	$8,000,000
Patrick Kane	RW	8.21	22	65	249	77	7	$6,000,000
Pavel Datsyuk	C	8.05	19	67	164	97	21	$6,700,000
Ray Whitney	LW	7.98	23	76	182	34	24	$3,000,000
Rick Nash	LW	7.97	30	59	306	62	-19	$7,500,000
David Krejci	C	7.85	23	62	145	43	-5	$4,000,000
Alex Ovechkin	RW	7.74	38	65	303	34	-8	$9,000,000
Jarome Iginla	RW	7.71	32	67	251	52	-10	$7,000,000

Table C.8: 2011-2012 Bottom-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Scott Hannan	D	-2.26	2	12	49	13	-10	$1,000,000
Colin White	D	-2.04	1	4	33	10	-5	$1,000,000
Jake Gardiner	D	-1.97	7	30	79	34	-2	$875,000
Travis Hamonic	D	-1.91	2	24	124	37	6	$875,000
Robyn Regehr	D	-1.68	1	5	49	11	-12	$4,000,000
Brett Clark	D	-1.32	2	15	61	26	-26	$1,300,000
Bryan Allen	D	-1.19	1	14	87	22	-1	$3,150,000
Marc-Edouard Vlasic	D	-1.19	4	23	119	18	12	$3,500,000
Carl Gunnarsson	D	-1.15	4	19	89	34	-9	$1,400,000
Zac Rinaldo	C	-1.10	2	9	54	4	-1	$560,000
Zenon Konopka	C	-1.09	3	5	34	7	-4	$700,000
Jared Boll	RW	-1.08	2	3	35	7	-8	$750,000
Lubomir Visnovsky	D	-0.99	6	26	110	22	5	$5,000,000
Justin Faulk	D	-0.98	8	22	101	32	-16	$790,000
Rostislav Klesla	D	-0.92	3	13	87	13	13	$2,975,000
Kris Russell	D	-0.91	6	12	56	16	12	$1,300,000
Stu Bickel	D	-0.91	0	9	22	5	2	$600,000
Radek Dvorak	RW	-0.88	4	21	83	42	-16	$1,500,000
John McCarthy	LW	-0.87	0	0	14	2	-2	$525,000
Mike Weber	D	-0.83	1	5	51	14	-19	$900,000
Brad Stuart	D	-0.83	6	21	96	22	16	$3,750,000
Artem Anisimov	C	-0.83	16	36	132	37	12	$1,875,000
Pavel Kubina	D	-0.82	3	15	75	18	-2	$3,500,000
Jay Pandolfo	LW	-0.80	1	3	44	14	-14	$600,000
Scottie Upshall	LW	-0.80	2	5	53	8	-3	$3,500,000

Table C.9: 2010-2011 Top-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Jonathan Toews	C	11.24	32	76	233	93	25	$6,500,000
Ryan Kesler	C	9.48	41	73	260	65	24	$5,000,000
Alex Ovechkin	RW	9.32	32	85	367	48	24	$9,000,000
Patrick Sharp	LW	9.17	34	71	268	64	-1	$4,100,000
Brad Richards	C	9.13	28	77	272	47	1	$7,800,000
Anze Kopitar	C	8.74	25	73	233	62	25	$6,000,000
Jeff Carter	C	8.60	36	66	335	40	27	$5,500,000
Eric Staal	C	8.59	33	76	296	64	-10	$7,500,000
Rick Nash	LW	8.31	32	66	305	47	2	$7,500,000
Claude Giroux	C	8.07	25	76	169	48	20	$765,000
Steven Stamkos	C	7.79	45	91	272	40	3	$875,000
Tomas Plekanec	C	7.62	22	57	227	43	8	$5,000,000
Stephen Weiss	C	7.32	21	49	172	44	-9	$3,200,000
Joe Pavelski	C	7.28	18	63	275	50	9	$4,000,000
Mike Santorelli	C	7.20	20	41	193	32	-17	$600,000
Paul Stastny	C	7.07	22	57	181	52	-7	$6,600,000
Olli Jokinen	C	6.98	17	54	208	54	-17	$3,000,000
Jason Spezza	C	6.84	21	57	188	52	-7	$8,000,000
Henrik Zetterberg	LW	6.73	24	80	306	54	-1	$7,750,000
Mikko Koivu	C	6.72	17	62	191	64	4	$3,700,000
John Tavares	C	6.70	29	67	241	75	-15	$900,000
Jarome Iginla	RW	6.62	43	86	289	40	0	$7,000,000
Bryan Little	C	6.59	18	48	158	80	11	$1,650,000
Michael Grabner	RW	6.51	34	52	227	69	14	$765,000
Brad Boyes	RW	6.36	17	55	178	30	13	$4,500,000

Table C.10: 2010-2011 Bottom-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Matt Martin	LW	-2.38	5	14	59	25	-13	$615,000
Theo Peckham	D	-2.04	3	13	41	30	-5	$550,000
Chris Phillips	D	-1.77	1	9	81	26	-35	$3,500,000
Brian Lee	D	-1.36	0	3	35	5	-10	$875,000
Andrew Alberts	D	-1.33	1	7	21	7	0	$1,300,000
Cody McLeod	LW	-1.32	5	8	73	7	-7	$1,000,000
Jim Vandermeer	D	-1.24	2	14	57	27	-15	$2,300,000
Niklas Hjalmarsson	D	-1.21	3	10	64	32	13	$3,500,000
Jonas Holos	D	-1.18	0	6	36	9	-3	$624,000
Jamal Mayers	RW	-1.17	3	14	61	24	3	$600,000
Andrew MacDonald	D	-1.03	4	27	71	49	10	$500,000
Andreas Lilja	D	-1.02	1	7	31	6	-15	$600,000
John Erskine	D	-0.92	4	11	58	11	1	$1,250,000
Micheal Haley	C	-0.90	2	3	13	7	-4	$500,000
Mike Komisarek	D	-0.90	1	10	48	14	-8	$6,000,000
Chris Neil	RW	-0.84	6	16	105	35	-14	$2,000,000
Tom Gilbert	D	-0.79	6	26	106	43	-14	$5,500,000
Colin White	D	-0.78	0	6	50	21	-2	$3,000,000
Matt Carkner	D	-0.76	1	7	40	9	0	$700,000
Paul Bissonnette	LW	-0.72	0	0	17	2	5	$600,000
Mark Giordano	D	-0.69	8	43	165	25	-8	$1,075,000
Derek Joslin	D	-0.68	2	9	34	7	4	$500,000
Douglas Murray	D	-0.68	1	14	102	21	6	$2,500,000
Petr Prucha	RW	-0.65	0	1	10	4	0	$1,100,000
Francis Lessard	RW	-0.64	0	0	6	1	0	$254,000

Table C.11: 2009-2010 Top-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Sidney Crosby	C	10.47	51	109	298	43	15	$9,000,000
Jonathan Toews	C	10.24	25	68	202	69	22	$850,000
Alex Ovechkin	RW	9.33	50	109	368	66	45	$9,000,000
Anze Kopitar	C	8.91	34	81	259	36	6	$6,000,000
Mikko Koivu	C	8.84	22	71	246	55	-2	$3,300,000
Henrik Zetterberg	LW	8.68	23	70	309	53	12	$7,500,000
Stephen Weiss	C	8.44	28	60	180	63	-7	$3,000,000
Jason Spezza	C	8.44	23	57	165	39	0	$8,000,000
Pavel Datsyuk	C	8.42	27	70	203	132	17	$6,700,000
Ilya Kovalchuk	RW	8.23	41	85	290	34	10	$6,000,000
Steven Stamkos	C	8.06	51	95	297	47	-2	$875,000
Tomas Plekanec	C	7.99	25	70	216	46	5	$2,750,000
Jeff Carter	C	7.96	33	61	319	43	2	$5,000,000
Matt Cullen	C	7.91	16	48	195	54	-7	$2,800,000
Brad Richards	C	7.69	24	91	284	57	-12	$7,800,000
Patrice Bergeron	C	7.54	19	52	184	55	6	$5,000,000
Vincent Lecavalier	C	7.48	24	70	295	33	-16	$10,000,000
Marian Gaborik	RW	7.36	42	86	272	25	15	$7,500,000
Nicklas Backstrom	C	7.32	33	101	222	54	37	$850,000
Phil Kessel	RW	7.25	30	55	297	30	-8	$4,500,000
Jarome Iginla	RW	7.17	32	69	257	45	-2	$7,000,000
Patrick Marleau	LW	7.08	44	83	274	53	21	$6,300,000
Paul Stastny	C	7.01	20	79	199	59	2	$6,600,000
Rick Nash	LW	7.01	33	67	254	41	-2	$7,000,000
Scott Gomez	C	7.00	12	59	180	52	1	$8,000,000

Table C.12: 2009-2010 Bottom-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Andrei Markov	D	-2.37	6	34	85	34	11	$5,750,000
Matt Greene	D	-2.24	2	9	57	4	4	$2,750,000
Nick Boynton	D	-1.98	1	8	50	14	5	$1,500,000
Darcy Hordichuk	LW	-1.51	1	2	21	7	-7	$771,000
Adam Foote	D	-1.42	0	9	25	11	8	$3,250,000
Matt Carkner	D	-1.42	2	11	87	19	0	$500,000
Ryan O'Byrne	D	-1.37	1	4	27	7	-3	$725,000
Andreas Lilja	D	-1.10	1	2	19	1	-2	$1,250,000
Paul Martin	D	-1.09	2	11	21	10	10	$4,500,000
Jonathan Ericsson	D	-1.08	4	13	55	13	-15	$900,000
Jared Boll	RW	-1.05	4	7	56	10	-8	$550,000
Adam Pardy	D	-1.04	2	9	40	16	-3	$700,000
Dean Arsene	D	-1.03	0	0	4	0	-3	$292,000
Zenon Konopka	C	-1.02	2	5	41	7	-11	$500,000
Brad Staubitz	RW	-0.95	3	6	24	3	0	$500,000
Brandon Prust	LW	-0.94	5	14	44	11	9	$525,000
Mike Lundin	D	-0.93	3	13	42	17	-4	$433,000
Craig Rivet	D	-0.92	1	15	63	17	-6	$3,500,000
Matt Martin	LW	-0.88	0	2	10	1	-1	$665,000
Ruslan Salei	D	-0.86	1	6	22	1	-1	$3,275,000
Josh Gorges	D	-0.83	3	10	52	20	2	$1,000,000
Brendan Witt	D	-0.80	2	5	25	10	-18	$1,959,000
Brett Clark	D	-0.79	3	20	75	20	6	$3,500,000
Christoph Schubert	D	-0.77	2	7	73	25	-6	$900,000
Milan Lucic	LW	-0.72	9	20	72	12	-7	$685,000

Table C.13: 2008-2009 Top-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Sidney Crosby	C	11.16	33	103	238	56	3	$9,000,000
Jeff Carter	C	10.52	46	84	342	72	23	$4,500,000
Eric Staal	C	9.54	40	75	372	55	15	$5,000,000
Zach Parise	LW	9.01	45	94	364	34	30	$2,500,000
Vincent Lecavalier	C	8.54	29	67	291	51	-9	$7,167,000
Derek Roy	C	8.38	28	70	221	52	-5	$3,500,000
Pavel Datsyuk	C	8.15	32	97	248	89	34	$6,700,000
Jonathan Toews	C	8.13	34	69	195	54	12	$850,000
Chris Drury	C	8.12	22	56	219	48	-8	$7,100,000
Ryan Getzlaf	C	8.04	25	91	227	55	5	$4,500,000
Jarome Iginla	RW	8.02	35	89	289	35	-2	$7,000,000
Alex Ovechkin	RW	7.99	56	110	528	60	8	$9,000,000
Mike Ribeiro	C	7.79	22	78	163	67	-4	$5,000,000
Rick Nash	LW	7.65	40	79	263	70	11	$6,500,000
Mike Richards	C	7.58	30	80	238	83	22	$5,400,000
Saku Koivu	C	7.42	16	50	123	38	4	$4,750,000
Patrick Marleau	LW	7.29	38	71	251	45	16	$6,300,000
Henrik Zetterberg	LW	7.18	31	73	309	42	13	$2,900,000
Teemu Selanne	RW	7.18	27	54	186	23	-3	$3,250,000
Scott Gomez	C	7.15	16	58	271	57	-2	$8,000,000
Jason Blake	LW	7.04	25	63	302	53	-2	$4,500,000
Olli Jokinen	C	6.99	29	57	236	38	-12	$5,250,000
Todd White	C	6.98	22	73	150	57	-9	$2,350,000
Alexander Semin	RW	6.87	34	79	223	73	25	$4,200,000
Shane Doan	RW	6.80	31	73	230	48	5	$4,550,000

Table C.14: 2008-2009 Bottom-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Scott Niedermayer	D	-2.06	14	59	178	75	-8	$6,750,000
Duncan Keith	D	-1.78	8	44	173	50	33	$1,600,000
Boris Valabik	D	-1.74	0	5	16	19	-14	$729,000
Zack Stortini	RW	-1.47	6	11	28	3	-3	$600,000
Bret Hedican	D	-1.22	1	6	40	6	-7	$805,000
Cory Sarich	D	-1.15	2	20	57	13	12	$3,400,000
Denis Gauthier	D	-1.15	2	4	36	10	-11	$1,931,000
Garnet Exelby	D	-1.08	0	7	42	27	-2	$1,400,000
Dan Hinote	RW	-1.07	1	5	24	11	-7	$1,000,000
Chris Neil	RW	-1.06	3	10	59	22	-13	$1,200,000
Steve Downie	RW	-1.05	3	6	26	9	-2	$585,000
Luke Schenn	D	-0.99	2	14	102	33	-12	$875,000
Mike Brown	RW	-0.98	2	4	44	9	-7	$523,000
Darcy Hordichuk	LW	-0.97	4	5	26	11	1	$750,000
Ben Eager	LW	-0.93	11	15	80	12	1	$601,000
Aaron Voros	LW	-0.90	8	16	66	6	-9	$1,200,000
Cam Janssen	RW	-0.88	1	4	22	1	-5	$550,000
Colton Orr	RW	-0.87	1	5	40	9	-15	$550,000
Ruslan Salei	D	-0.86	4	21	93	18	-15	$3,025,000
Jim Vandermeer	D	-0.83	1	7	31	14	1	$2,300,000
Shane O'Brien	D	-0.82	0	10	39	19	5	$1,025,000
Eric Godard	RW	-0.82	2	4	20	2	-3	$725,000
Brad May	LW	-0.81	1	7	32	13	0	$600,000
Ladislav Smid	D	-0.81	0	11	33	13	-6	$952,381
Krys Barch	RW	-0.78	4	9	27	9	1	$575,000

Table C.15: 2007-2008 Top-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Rick Nash	LW	10.93	38	69	329	56	2	$5,500,000
Henrik Zetterberg	LW	10.12	43	92	358	53	30	$2,700,000
Jarome Iginla	RW	9.24	50	98	338	47	27	$7,000,000
Alex Ovechkin	RW	8.95	65	112	446	68	28	$984,000
Evgeni Malkin	C	8.81	47	106	272	69	16	$984,000
Vincent Lecavalier	C	8.45	40	92	318	52	-17	$7,167,000
Pavel Datsyuk	C	8.37	31	97	264	144	41	$6,700,000
Marian Hossa	RW	8.25	29	66	264	66	-14	$7,000,000
Sidney Crosby	C	8.14	24	72	173	35	18	$850,000
Ilya Kovalchuk	RW	7.94	52	87	283	49	-12	$5,500,000
Brad Boyes	RW	7.93	43	65	207	33	1	$1,600,000
Anze Kopitar	C	7.91	32	77	201	52	-15	$850,000
Jason Spezza	C	7.82	34	92	210	44	26	$5,000,000
Chris Drury	C	7.63	25	58	220	64	-3	$7,100,000
Mike Richards	C	7.62	28	75	212	46	14	$942,000
Alex Kovalev	RW	7.61	35	84	230	47	18	$4,500,000
Daniel Alfredsson	RW	7.51	40	89	217	72	15	$4,690,670
Scott Gomez	C	7.42	16	70	242	77	3	$10,000,000
Patrick Sharp	LW	7.31	36	62	209	44	23	$825,000
Daymond Langkow	C	7.28	30	65	201	52	16	$2,442,000
Jeff Carter	C	7.12	29	53	260	56	6	$942,400
Mike Modano	C	6.91	21	57	200	86	-11	$4,250,000
Patrick Kane	RW	6.87	21	72	191	49	-5	$3,725,000
Eric Staal	C	6.82	38	82	310	56	-2	$4,500,000
Joe Thornton	C	6.64	29	96	178	55	18	$6,670,000

Table C.16: 2007-2008 Bottom-25 Player Impact Scores For Winning

Name	Position	Winning Impact	Goals	Points	Shots	Takeaways	+/-	Salary
Hal Gill	D	-2.72	3	24	86	24	6	$2,075,000
Kyle McLaren	D	-2.62	3	11	39	11	3	$2,500,000
Jack Johnson	D	-1.96	3	11	81	23	-19	$2,150,000
Ruslan Salei	D	-1.67	6	30	111	13	-4	$3,025,000
Ryan Hollweg	LW	-1.65	2	4	59	10	-12	$495,000
Anders Eriksson	D	-1.50	1	18	50	28	-5	$1,500,000
Milan Jurcina	D	-1.50	1	9	58	15	4	$850,000
Braydon Coburn	D	-1.49	9	36	113	36	17	$942,400
Zack Stortini	RW	-1.48	3	12	38	5	3	$506,000
Krys Barch	RW	-1.47	1	3	23	10	-3	$475,000
Nick Schultz	D	-1.32	2	15	52	20	9	$1,850,000
Branislav Mezei	D	-1.26	2	4	38	9	-13	$850,000
Cam Barker	D	-1.16	6	18	42	5	-3	$1,595,000
Craig Weller	RW	-1.15	3	11	72	10	-7	$475,000
Ladislav Smid	D	-1.14	0	4	45	18	-15	$617,000
Aaron Downey	RW	-1.12	0	3	15	3	0	$525,000
Nicklas Grossmann	D	-1.11	0	7	34	8	10	$675,000
Cory Sarich	D	-1.08	2	7	57	27	2	$3,900,000
Danny Richmond	D	-1.08	0	0	2	0	-5	$151,000
Riley Cote	LW	-1.07	1	4	17	5	2	$476,000
Tom Gilbert	D	-1.06	13	33	98	33	-6	$907,000
Staffan Kronwall	D	-1.05	0	0	9	4	-2	$112,000
Filip Kuba	D	-1.03	6	31	113	23	-8	$3,000,000
Colton Orr	RW	-0.98	1	2	24	14	-13	$525,000
George Parros	RW	-0.98	1	5	30	9	3	$525,000