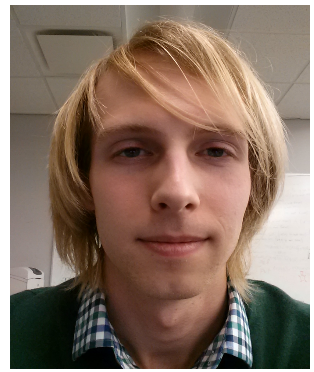
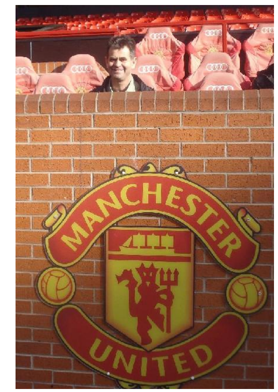
What is the Value of an Action in Ice Hockey? Learning a Q-function for the NHL

Oliver Schulte Zeyu Zhao Kurt Routley Tim Schwartz



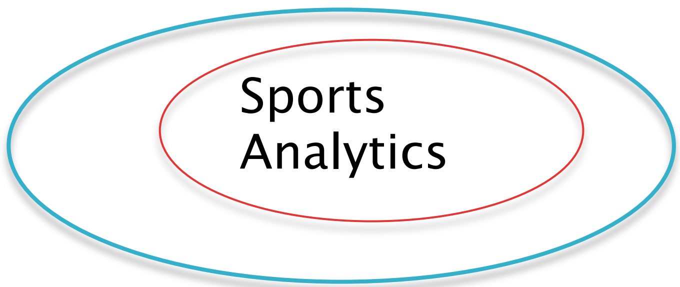
Computing Science/Statistics
Simon Fraser University
Burnaby-Vancouver, Canada

SFU

Big Picture: Sports Analytics meets Reinforcement Learning

- Reinforcement Learning: Major branch of Artificial Intelligence (*not* psychology).
- Studies *sequential decision-making under uncertainty*.
- Studied since the 1950s
  - Many models, theorems, algorithms, software.

Reinforcement Learning



on-line intro text
by Sutton and Barto

Markov Game Models

Markov Game

- Fundamental model type in reinforcement learning: **Markov Decision Process.**
- **Multi-agent version: Markov Game.**
- **Models** dynamics: **e.g. given the current state of a match, what event is likely to occur next?**
- **Application in this paper:**
  1. **value actions.**
  2. **compute player rankings.**

Markov Game Dynamics Example

Home = Colorado
Away = St. Louis

Differential = Home – Away

Initial State
Goal
Differential = 0,
Manpower
Differential = 2,
Period = 1

face-off(
Home,Offensive Zone)

0,2,1
[face-
off(Home,Off.)]

Time in
Sequence
(sec)

0 sec
Alexander Steen
wins Face-off in
Colorado's
Offensive Zine

Markov Game Dynamics Example

GD = 0, MD = 2, P = 1

0,2,1
[face-off(Home,Off.)]

0,2,1
[face-off(Home,Off.)
Shot(Away,Off.)]

Time in
Sequence (sec)

0 sec
Alexander
Steen wins
Face-off

16 sec
Matt
Duchen
shoots

Markov Game Dynamics Example

GD = 0, MD = 2, P = 1
P(Away goal) = 32%

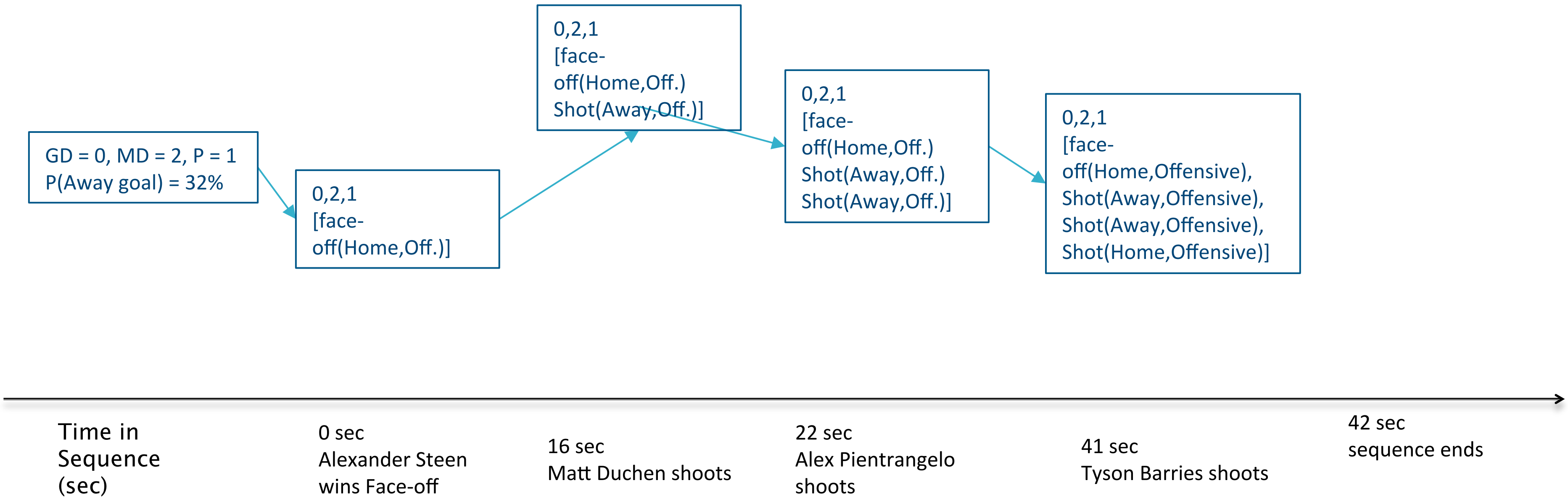
0,2,1
[face-off(Home,Off.)]

0,2,1
[face-off(Home,Off.)
Shot(Away,Off.)]

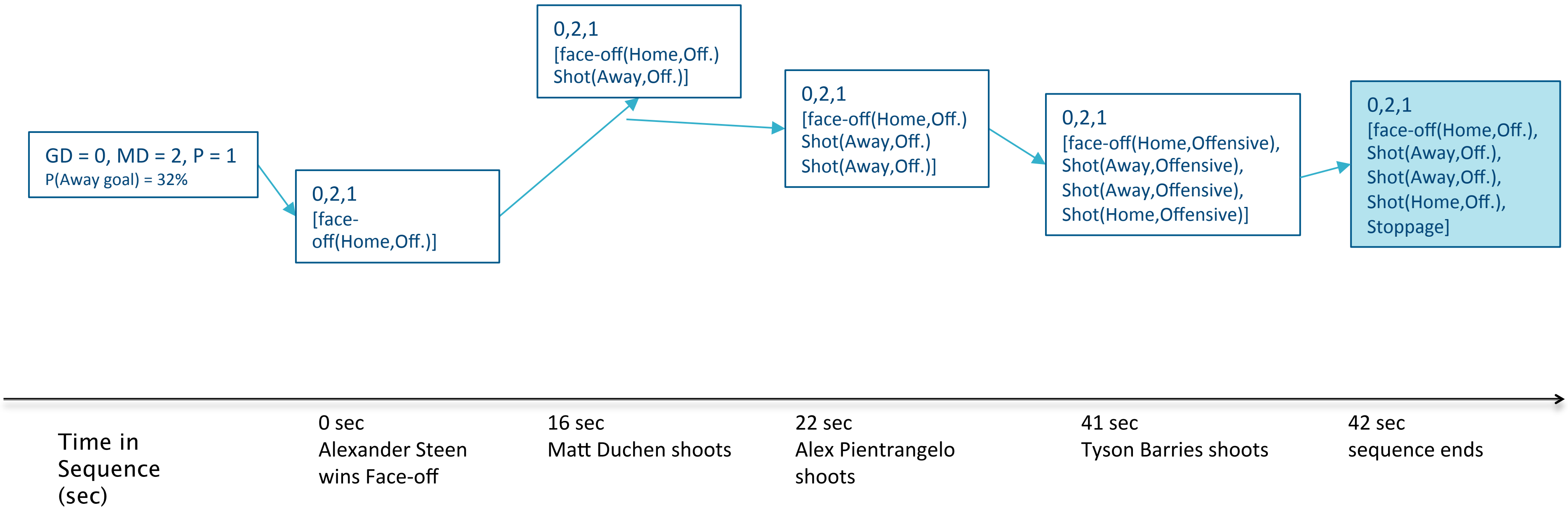
0,2,1
[face-off(Home,Off.)
Shot(Away,Off.)
Shot(Away,Off.)]

Time in Sequence (sec)	0 sec Alexander Steen wins Face-off	16 sec Matt Duchen shoots	22 sec Alex Pientrangelo shoots	41 sec Tyson Barries shoots	42 sec sequence ends

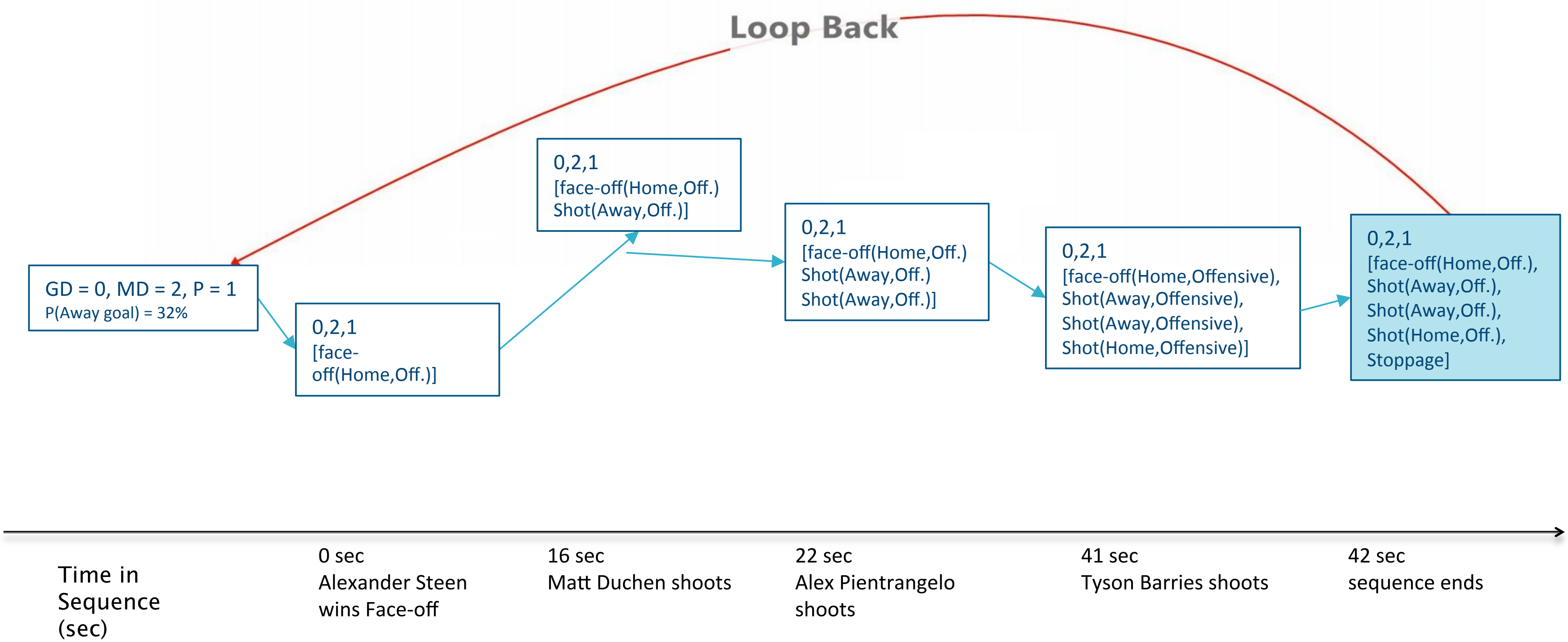
Markov Game Dynamics Example



Markov Game Dynamics Example



Markov Game Dynamics Example

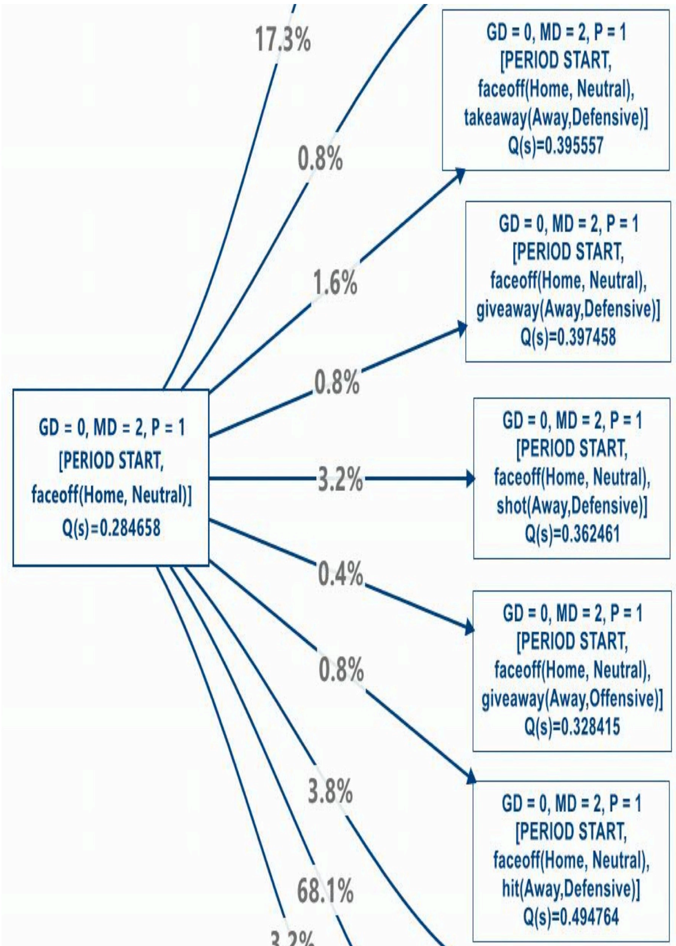


Markov Game Description

- Two agents, Home and Away.
- Zero-sum: if Home earns a reward of r, then Away receives -r.
- Rewards can be
  - win match
  - **score goal**
  - receive penalty (cost).

Learning Markov Game Parameters

Markov Game Transition Probabilities = Parameters



Big Data: Play-by-play 2007-2015

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

Big Model: 1.3 M states

Action Values

Player Performance Evaluation

Expected rewards

- Key quantity in Markov game models: the **total expected reward** for a player given the current game state.
  - Written V(s).
- **Looks ahead** over all possible game continuations.

transition probabilities $\xrightarrow{\text{dynamic programming}}$ total expected reward of state s

Q-values and Action Impact

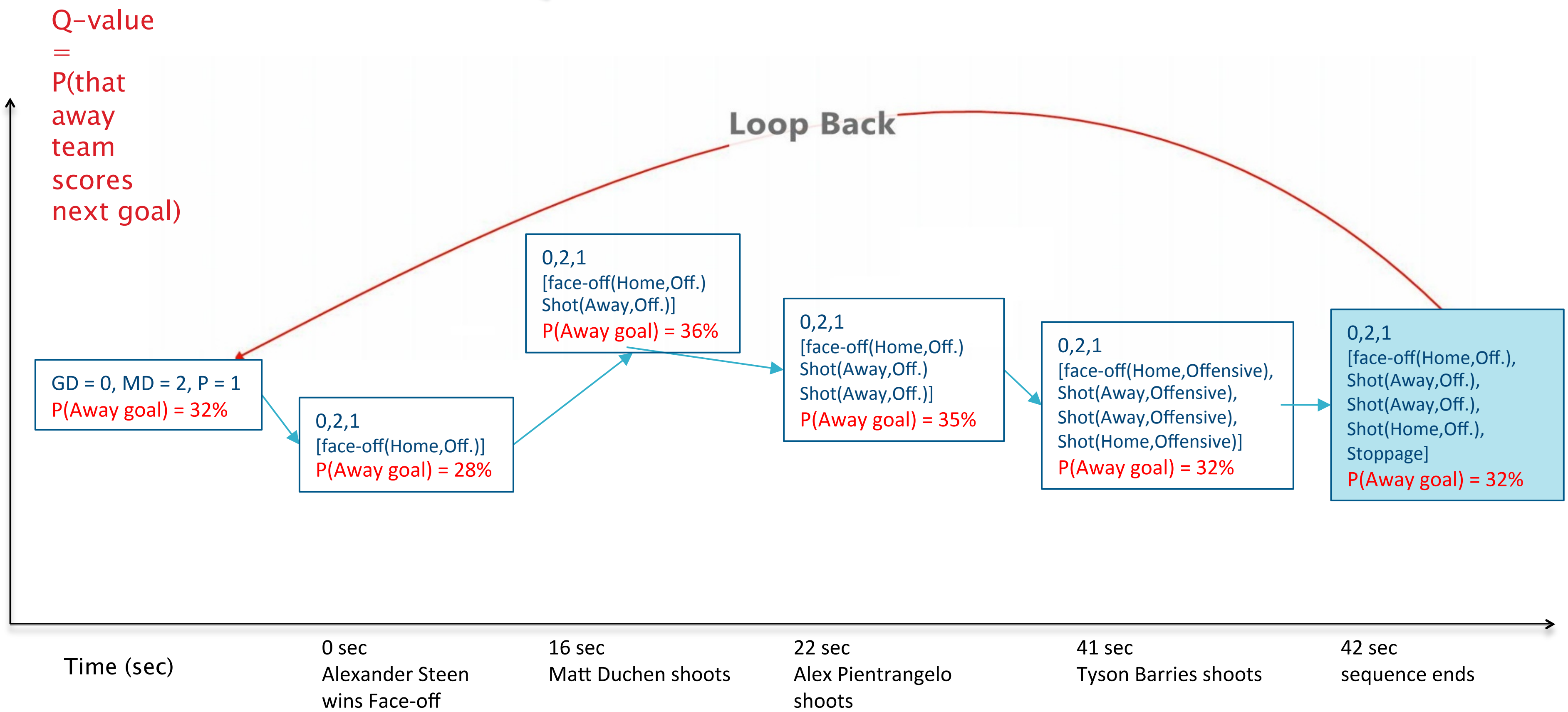
- Q(s,a) = the expected total reward if action a is executed in state s.
- The **action-value function.**

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

Expected reward after action

Expected reward before action

Q-value Ticker

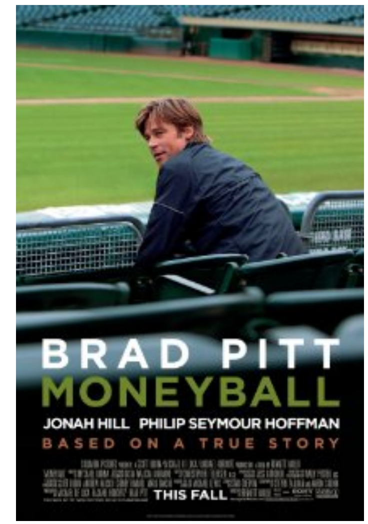


Advantages of Impact Value

- Context-Aware.
  - e.g. goals more valuable in ties than when ahead.
- Look Ahead:
  - e.g. penalties➔ powerplay ➔ goals but not immediately.

Computing Player Impact

1. From the Q-function, compute impact values of state-action pairs.
2. For each action that a player takes in a game state, find its impact value.
3. Sum player action impacts over all games in a season. (Like +/-).



Results 2014-2015 1st half

- The Blues' STL line comes out very well.
- Tarasenko is under-valued, St. Louis increased his salary 7-fold.

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

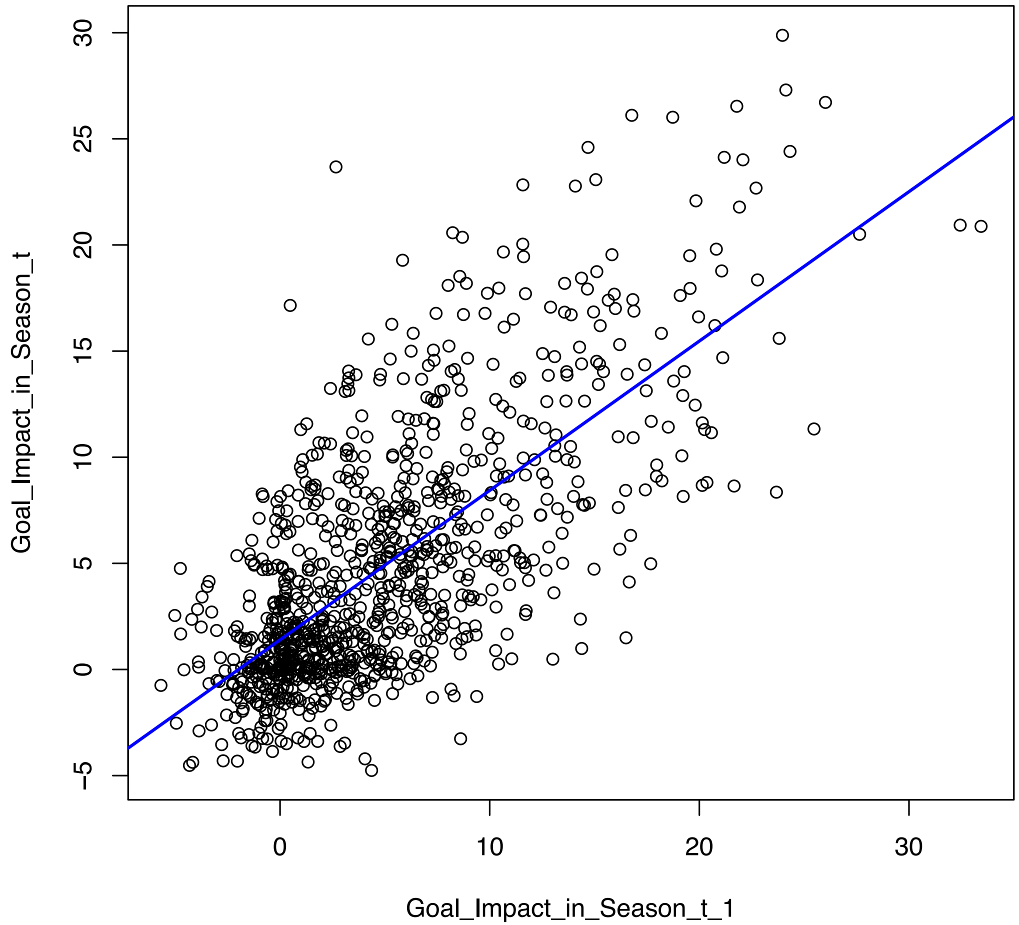
Results 2013-2014 Season

Name	Goal Impact	Points	+/-	Salary
Jason Spezza	29.64	66	-26	$5,000,000
Jonathan Toews	28.75	67	25	$6,500,000
Joe Pavelski	27.20	79	23	$4,000,000
Marian Hossa	26.12	57	26	$7,900,000
Patrick Sharp	24.43	77	12	$6,500,000
Sidney Crosby	24.23	104	18	$12,000,000
Claude Giroux	23.89	86	7	$5,000,000
Tyler Seguin	23.89	84	16	$4,500,000

Jason Spezza: high goal impact, low +/-.

- plays very well on poor team (Ottawa Senators).
- Requested transfer for 2014-2015 season.

Consistency Across Seasons



Correlation coefficient = 0.703
Follows Pettigrew(2015)

Related Work

- Routley and Schulte, UAI 2015
  - Values of Ice Hockey Actions, compares with THoR (Schuckers and Curro 2015).
  - Ranks players by impact on goals and *penalties*.
- Pettigrew, Sloan 2015.
  - reward = win.
  - estimates impact of goal on win probability given score differential, manpower differential, game time.
- Cervone et al., Sloan 2014.
  - Conceptually similar but for **basketball**.
  - our impact function = their EPVA.
  - uses spatial tracking data.

Conclusion

- Reinforcement Learning ➔ Model of Game Dynamics.
- Connects advanced machine learning with sports analytics.
- Application in this paper:
  - use Markov game model to **quantify impact** of a player's action (on expected reward).
  - use total impact values to rank players.
- Impact value
  - is aware of context.
  - looks ahead to game future trajectory.
- Total impact value is consistent across seasons.