

We thank all three reviewers for their time and attention to our paper. We will follow all the very helpful suggestions for edits and clarifications. We especially appreciated the high level of engagement with the material, such as the suggestions for research questions to investigate.

Reviewers 1+2: Both reviewers suggest linking player impact scores to predicting other outcomes of interest, such as team standing, performance, and match results. Current sports analytics treat event prediction as a separate problem from evaluating player performance, which they view mainly as a descriptive statistics problem (e.g. +/-) [see http://ai.arizona.edu/mis510/other/SportsDataMining_Book.pdf, references in our paper, and reference [1] provided by Reviewer_5]. We agree that it is a great project to link advanced player analytics to prediction; one might call it the holy grail of sports statistics. This is a great topic for future research with a lot of potential for the UAI community to contribute to sports statistics. The more limited aim of our current paper was to use AI models and techniques to improve player descriptive statistics for ice hockey as an important topic in itself. A common motivation for player evaluation scores are hiring decisions (Moneyball).

Reviewer_1:

p.7 The entropy numbers measure the uncertainty of action-states with respect to the conditional probability of who scores the next goal. Our measurements quantify the decrease in the uncertainty about who scores next as we enrich the Markov model to include more information.

p.9 Table 9 describes how action values change when the model propagates information from one state to another. It shows the average change + s.d. in action values over all contexts. We observe in Table 9 that propagation changes the value of an action because it captures long-term effects. This is especially true for penalties, but for example, even scoring a goal has a long-term effect of increasing the chance of a penalty, as shown in Figure 2b. Results were statistically significant from t-test ($p=2.8 \times 10^{-8}$).

- The data specify which player performs an action. Our model treats all players as exchangeable, so it does not make use of their positions or identities. The model provides a generic value for an action, within context, over all players. When we rank specific players, we can use the generic action value to score their specific action. This is the same approach as in our closest predecessor Schuckers and Curro, and in reference [1] given by Reviewer_5.

Including player positions (and even player identities) in the feature space is a great idea for future work. It does raise computational challenges because it multiplies the state space by a factor of 5.

- Yes, it should be “on policy” setting, as we are evaluating the observed policy by computing Q-values.

The 64.8% goal difference on powerplays shows there is still a significant number of shorthanded goals, as the powerplay team will score goals with conditional probability of 82.4%, and the shorthanded team

will score goals with conditional probability of 17.6%. The rewards for Figure 1 are for expected goals, so the reward for the goal(Home,Offensive) state is $R(s) = 1$, but the STOPPAGE state should have $R(s) = 0$. The boxplots in Figure 2 are created using the standard MATLAB format, with green asterisks added for THoR action values [5], red dots denoting action impact values outside 2.7sd, and the horizontal dashed lines are a cut-off we set to keep the figure small, so not all outliers were shown.

Reviewer_2:

Modeling Penalty Impact. Yes, ultimately it's goals that matter. Penalties are one path to goals that a coach may want understand in more detail. For instance, if a team receives an unusual number of penalties, a coach may want to know which players are responsible and by which actions. Another reason is penalties are major events, so understanding their dynamics contributes to understanding hockey. For example, our finding that giveaways increase penalties appears to be new in the hockey literature.

Reviewer_5:

Thank you for the great basketball reference [1] from last year's Sloan, which we will certainly include. Our work was done independently, but since our submission we've met with one of the [1] authors and we appreciate the overlap in the definition of the impact of an action. We think that this definition is a big idea---even if we are not the first to publish it---that deserves being explored and applied in more than one paper, and in more than one sport. The fact that a prominent sports analytics group independently developed a similar definition seems to confirm that we are on the right track. We believe there is a lot of value in showcasing this approach to the UAI community: The reference [1] is very recent, and uses neither AI terminology nor AI techniques, so our paper will be more accessible to the AI community. Although there is agreement on the definition of action impact, pretty much everything else in our paper is different from [1]. (Three out of our four listed main contributions.) For example, the NHL does not yet have and therefore we do not use spatial tracking data, which is the main focus of [1]. On P.6 of [1] the authors discuss the advantages of using a discrete state space (stochastic consistency) but consider it computationally infeasible for their data. We show that leveraging AI data structures and algorithms makes handling a large discrete state space feasible for ice hockey. Including the local action history in the state space allows us to capture the medium-term effects of actions. This is more important for ice hockey than for basketball because in basketball scoring occurs at much shorter intervals.

-There seems to be a misunderstanding about the functional nature of the state successor relationship. $\text{Occ}(s,s')$ is only a functional relationship near leaf nodes, in less than 1/3 state transitions.

[1] <http://arxiv.org/pdf/1408.0777v1.pdf>