Supplementary Material for "A Markov Game Model for Valuing Player Actions in Ice Hockey"

Kurt Routley
School of Computing Science
Simon Fraser University
Vancouver, BC, Canada
kdr4@sfu.ca

Oliver Schulte
School of Computing Science
Simon Fraser University
Vancouver, BC, Canada
oschulte@cs.sfu.ca

1 LESION STUDY

We evaluate the components of our full model by considering simpler models.

1.1 USING AVERAGE ACTION VALUES

We compare context-aware action values vs. fixed action values (as in THoR) in terms of the entropy of the Next Goal conditional probabilities. This quantifies the information lost by ignoring context. Table 1 shows the average entropy for context-aware and context-unaware probabilities. The *context-unaware* Next Goal probability for an action event, is the marginal probability obtained from action-state probabilities by averaging over all states where the action is taken. The marginal probability of the next goal leads to an average context-unaware entropy of 0.9971. The average of the context-aware entropies is 0.9540. The entropy improvement is statistically significant according to the paired t-test (p $= 2.8 \times 10^{-8}$). Moreover, the variance of the context-aware entropy is considerable, which means that Next Goal predictions are in many states even more informative than the average entropy shown in Table 1.

1.2 EXAMING PROPAGATION EFFECTS

The transition graph construction algorithm facilitates changing the possible state transitions. We utilize this in our experiments to study how different propagation models affect the impact of actions on Next Goal Scored. Specifically, we consider three different transitions graphs of increasing density, their sizes shown in Table 2. The number of states/nodes 1,325,809 is the same for all graphs.

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.

Full Transition Graph Includes loopback edges from all leaf nodes to context nodes, as defined in Section 4.2. of the main paper.

Table 2: Size of State Transition Graphs

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

Action impact changes value depending on the state transition graph. The average differences in action values of the same states across different transition graphs, as well as the standard deviation of the differences, are shown in Table 3. The table shows that the estimated impact on who scores the next goal changes as more information is propagated between states.

Penalty vs. Local. With the local transition graph, value iteration computes the impact of an action on the current play sequence only. This means that the next play sequence is not considered during look-ahead. In hockey terms, with the local transition graph, the model is not aware that a penalty is followed by a powerplay. The local Q-value differential for context states, with the initial empty play sequence, can be obtained from Table 4 of the main paper (last two columns). The penalty transition graph propagates to the next sequence the effect of penalties only. This means that the next play sequence is considered during look-ahead only if the current sequence ends with a penalty. 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 2.

Full vs. Penalty. In hockey terms, with the penalty transition graph, the model is aware that a penalty is followed by a *single* powerplay sequence. But if more than one

Table 1: Context-Aware vs. Context-Unaware Entropies.

Action	Context-Unaware Probability Of Next Goal	Context-Unaware Entropy	Average Context-Aware Entropy	Context-Aware Standard Deviation
Blocked Shot	0.4840	0.9993	**0.9455**	0.1981
Faceoff (Defensive)	0.4828	0.9991	**0.9913**	0.0539
Faceoff (Neutral)	0.5025	1.0000	**0.9944**	0.0541
Faceoff (Offensive)	0.5335	0.9968	**0.9876**	0.0578
Giveaway	0.4907	0.9997	**0.9271**	0.0283
Hit	0.4985	1.0000	**0.9462**	0.0233
Missed Shot	0.5178	0.9991	**0.9413**	0.0280
Penalty	0.4442	0.9910	**0.9833**	0.0219
Shot	0.5673	0.9869	**0.8951**	0.0386
Takeaway	0.5125	0.9995	**0.9279**	0.0276
Average Entropy Over Actions		0.9971	**0.9540**	

sequence occurs in the same powerplay, the second sequence is ignored in the lookahead (unless it also ends in a penalty). The full transition graph propagates the information about the manpower advantage to the next sequence. Comparing the full transition graph with penalty propagation only, we still find the strongest average impact change for penalties. The simplest explanation of this result is that in hockey, the effect of penalties often goes beyond a single play sequence, and the full transition graph captures more of this medium-term effect.

While the aggregate differential effects show that more propagation leads to more informative results on average, the variance of the impact differentials show that in many states, propagation provides even more information than the averages in Table 3 suggest.

Table 3: Action Impact *Differences* For The Next Goal Depending on Propagation Model.

	Full vs. Penalty		Penalty vs. Local	
Action	Average	Std. Dev.	Average	Std. Dev.
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