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

<table>
<thead>
<tr>
<th></th>
<th>Local</th>
<th>Penalty</th>
<th>Full</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Edges</td>
<td>1,325,809</td>
<td>1,382,780</td>
<td>1,662,504</td>
</tr>
</tbody>
</table>

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
sequence occurs in the same powerplay, the second se-
sequence is ignored in the lookahead (unless it also ends in a
penalty). The full transition graph propagates the informa-
tion about the manpower advantage to the next sequence.
Comparing the full transition graph with penalty propaga-
tion 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 De-
pending on Propagation Model.