A graphical method for simplifying Bayesian Games

Peter Thwaites (University of Leeds)

Joint work with Jim Q. Smith (University of Warwick)

GDRR July 2013
Motivation:

If the influence diagram (ID) depicting a Bayesian game is common knowledge (CK) to its players, then they can often use the irrelevance properties it encodes to discover a simpler game which still embodies all their optimal decision strategies.

However many Bayesian games are too asymmetric to be fully represented by an ID.


This talk presents the first use of the CEG for the depiction & analysis of a Bayesian game.
The Game:

The next 2 slides describe a 2-player Bayesian game with a small amount of asymmetry in its structure.
The game is then presented as an Extensive Form (EF) ID.

Extensive Form means that any chance or decision variables whose values are known to a player when they come to make a decision are connected to this decision (node) by directed informational edges. Any variables whose values are not known at this point are not.

We briefly show how the ID can be simplified to form a Parsimonious ID, before introducing a CEG-representation of the game.
X₁: A message \( \alpha \) is sent to A via one of 2 routes. This information is secure (not CK), but B observes which of the 2 routes \( \alpha \) is sent by.

D₁(A): A replies (message \( \beta \)) using one of 2 codes. B observes which code A uses, but without a key has no guarantee of cracking it.

D₁(B): B uses one of 2 techniques to try and decode message \( \beta \). Which technique B uses is observed by A.

X₂: B attempting to decode message \( \beta \) may send a warning to the originator of message \( \alpha \). The probability of it doing so depends on which of the 2 routes \( \alpha \) was sent by, which code A used for \( \beta \), and which technique B uses to try and decode \( \beta \). Whether or not this warning is sent is observed by both A & B.

X₃: Either B fails to decode \( \beta \) with utility \((U^A_3, U^B_3) = (+10, 0)\), or succeeds with utility \((U^A_3, U^B_3) = (-10, +10)\). B’s success depends on which code A used & on which technique B used. B’s success or otherwise is observed by A.

D₂(A): A sends a second message \( \gamma \) to the originator of \( \alpha \). The information contained in this message is false, but this fact is only known to A and the message’s intended recipient (not CK). A uses one of 2 codes. B observes which code A uses, but without a key has no guarantee of cracking it.

D₂(B): B uses one of 2 techniques to try and decode message \( \gamma \).

X₄: Either B fails to decode \( \gamma \) with utility \((U^A_4, U^B_4) = (0, 0)\), or succeeds.
We look at just one asymmetry:

If B failed to decode $\beta$ but succeeds in decoding $\gamma$ then he believes the information contained in the latter message with utility $(U_A^4, U_B^4) = (+20, +10)$. If B succeeded in decoding $\beta$ he does not believe the information in $\gamma$ with utility $(U_A^4, U_B^4) = (0, 0)$. B's success depends on which code A used and on which technique B used.

Note that $U_B^4(X_3 = 1, X_4 = 2)$ (B fails to decode $\beta$ but succeeds in decoding $\gamma$) is positive because B **believes** the information contained in $\gamma$. $U_A^4(X_3 = 1, X_4 = 2) > U_B^4(X_3 = 1, X_4 = 2)$ as A has successfully planted false information on B.

Note that if B succeeds in decoding $\beta$ then the decisions made at $D_2(A)$, $D_2(B)$ have no influence on $U$, and neither does the outcome of $X_4$. 
Initial ID

X

D

D

A

(B)

A

(B)

X

X

X

X

X

D

D

2

2

1

1

(Smith 1996)
In the ID we work back through the decision nodes from the utility node to the root node \((X_1)\). For each decision node, we look at whether we can write down a conditional independence statement of the form:

\[
U \perp Q^{\text{irr}}(D) \mid (D,Q^{\text{suff}}(D))
\]

where \(Q^{\text{suff}}(D)\) are called the sufficient parents of \(D\), and \(Q^{\text{irr}}(D)\) the irrelevant parents of \(D\).

For each \(D\) in turn we remove the edges from members of \(Q^{\text{irr}}(D)\) into \(D\). The resulting graph is called the parsimonious ID. This ID embodies both players’ optimal decision strategies.
Doing this for $D_2(B)$
we get ...
Doing this for $D_2(A)$
we get ...
We can see that $X_2$ is barren so it can be removed.
Looking at $D_1(B)$ & $D_1(A)$
we get ...

\[ X_1 \]

\[ D_1(A) \]
\[ D_1(B) \]
\[ X_3 \]
\[ D_2(A) \]
\[ D_2(B) \]
\[ X_4 \]

\[ U \]

\[ X_1 \] is now barren, so can be removed.
Parsimonious ID

Does not explicitly depict the asymmetry in the problem
Given the Parsimonious ID, both players can calculate their optimal decision strategies using any of the established techniques for solving IDs.

First B maximises his expected utility at $D_2(B)$. Then A maximises his expected utility at $D_2(A)$, assuming that B has acted rationally. Then B maximises his expected utility at $D_1(B)$, assuming ... Lastly A maximises his expected utility at $D_1(A)$, assuming ...

The resultant expressions obtained by B and A are the same as those we would get if we treated this as a single-decision-maker decision problem.

The calculations and resultant expressions are very messy because of the asymmetry described on slide 5. This asymmetry is associated with the variables $X_3$, $X_4$ and $U$, so it affects the arithmetic right from the start.
CEGs

1. CEGs are designed for asymmetric problems.

2. Unlike trees they depict the full conditional independence structure of the problem, through their colouring (local structure) and coalescence (global structure).

3. Unlike IDs they explicitly depict both the complete outcome space, and the asymmetries in the problem.
CEGs and our game

On the next slide we show a naive CEG-representation of the problem.

It is naive in the sense that it tells us only a little more than the ID – it does not yet encode the asymmetry from slide 5. This is encoded on the following slide.

These slides have also been drawn out in more detail than is strictly necessary.

It is sufficient for any CEG-representation to show certain key aspects of the model - the colouring of the vertices associated with $X_3$; the grouping and colouring of the $X_4$ vertices; and the asymmetric grouping of the utility vertices that we see once the asymmetry has been encoded.

These four aspects correspond to the four non-trivial conditional independence statements associated with the model.

The CEG is EF in the same sense that the ID depiction is EF.

The CEG is assumed to be CK.

We show how the CEG simplifies to a Parsimonious CEG, noting as we do so, the CEG-analogue of Shachter’s barren-node deletion for IDs.
A naive CEG representation does not encode the asymmetry of the problem.
Iteration 1: Asymmetrising the CEG

If $X_3 = 2$ (B succeeds in decoding $\beta$) then the decisions made at $D_2(A)$ & $D_2(B)$ have no influence on $U$, and neither does the outcome of $X_4$. 
Consider node 1: All routes for which \( X_3 = 1, D_2(A) = 1 \) and \( D_2(B) = 1 \) pass through this node, and no other routes. Similar results apply to the other \( X_4 \) nodes. So using the Markov property on these nodes we get that

\[
(X_4, U) \perp \!
\!
\!
\!
\perp (X_1, D_1(A), D_1(B), X_2) \mid (X_3 = 1, D_2(A), D_2(B))
\]
In particular:
\[ U \perp (X_1, D_1(A), D_1(B), X_2) \mid (X_3=1, D_2(A), D_2(B)) \]

But this is of the form:
\[ U \perp Q_{\text{irr}}(D_2(B)) \mid (D_2(B), Q_{\text{suff}}(D_2(B))) \]

So \( Q_{\text{irr}}(D_2(B)) = \{X_1, D_1(A), D_1(B), X_2\} \) are irrelevant to B for the purposes of optimising his utility at \( D_2(B) \), and the 32 \( D_2(B) \) nodes can be replaced by just 2 nodes.

These correspond to the 2 combinations of \( Q_{\text{suff}}(D_2(B)) = \{X_3=1, D_2(A)\} \).
Iteration 2

2 nodes only

D1(A)

D1(B)

D2(A)

D2(B)

X1

X2

X3

X4

U1

U2

U3
Consider now node 2: All routes for which $X_3 = 1$ and $D_2(A) = 1$ now pass through this node, and no other routes.

A similar result applies to the other $D_2(B)$ node.

So using the Markov property on these nodes we get that

$$(D_2(B), X_4, U) \perp\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
\!
In particular:
\[ U \perp \perp (X_1, D_1(A), D_1(B), X_2) \mid (X_3=1, D_2(A)) \]

But this is of the form:
\[ U \perp \perp Q_{\text{irr}}(D_2(A)) \mid (D_2(A), Q_{\text{suff}}(D_2(A))) \]

So \( Q_{\text{irr}}(D_2(A)) = \{X_1, D_1(A), D_1(B), X_2\} \) are irrelevant to A for the purposes of optimising his utility at \( D_2(A) \), and the 16 \( D_2(A) \) nodes can be replaced by just 1 node.

This corresponds to \( Q_{\text{suff}}(D_2(A)) = \{X_3=1\} \).

Also the coloured pairs of \( X_3 \) nodes now have identical outgoing paths, rather than just identical outgoing edges. The pairs can therefore be coalesced.
We can now see that the $X_2$ nodes are barren. We can simply remove them and run the edges from the $D_1(B)$ nodes straight into the $X_3$ nodes.
Iteration 4

\[ D_1(A) \rightarrow D_1(B) \rightarrow X_3 \rightarrow D_2(A) \rightarrow D_2(B) \rightarrow X_4 \]

\[ U_1 \rightarrow U_2 \rightarrow U_3 \]
We can see that B should make the same decision at nodes 3 & 5, and the same decision at nodes 4 & 6.
Hence $X_1$ is irrelevant to B for the purposes of optimising his utility at $D_1(B)$ and the 4 $D_1(B)$ nodes can be replaced by 2 nodes.
Iteration 5

X_1 \rightarrow D_1(A) \rightarrow X_2 \rightarrow D_1(B) \rightarrow X_3 \rightarrow D_2(A) \rightarrow X_4

X_3 \rightarrow D_2(B) \rightarrow X_4

U_1 \rightarrow X_4

U_2 \rightarrow X_4

U_3 \rightarrow X_4
We can see that A should make the same decision at both $D_1(A)$ nodes. Hence $X_1$ is irrelevant to A for the purposes of optimising his utility at $D_1(A)$ and the 2 $D_1(A)$ nodes can be replaced by 1 node.
Iteration 6
The $X_1$ node is clearly barren, so can be removed.
Explicitly depicts all aspects of the problem
Given the Parsimonious CEG, both players can easily calculate their optimal decision strategies using a rollback technique adapted from decision trees.

As the CEG explicitly depicts all aspects of the problem, the inherent asymmetry is automatically represented in the resultant expressions.

The maximum expected utility for A for example, is given by:

\[
U(A) = \text{Max}_{d_1(A)} \left[ \text{Max}_{d_1(B)} \left[ p(X_3=1|d_1(A),d_1(B)) \right] \right.
\]
\[
\times \text{Max}_{d_2(A)} \left[ \text{Max}_{d_2(B)} \left[ \Sigma_{x_4} p(x_4|d_2(A),d_2(B)) U(X_3=1,x_4) \right] \right]
\]
\[
+ p(X_3=2|d_1(A),d_1(B)) U(X_3=2) \right]
\]
Conclusion:
Many Bayesian games are too asymmetric to be fully represented by an ID.

The CEG has been designed for the representation & analysis of asymmetric problems.

It transpires that CEGs can be used to model & analyse Bayesian games very easily.
Some references:

JQ Smith (1996) “Plausible Bayesian Games” in *Bayesian Statistics 5*, ed Bernardo et al