Efficient Value of Information Computation

Ross D. Shachter **Engineering-Economic Systems and Operations Research Dept.**
Stanford University Ross D. Shachter

Systems and Operatio

Stanford University

nnford CA 94305-4023 mic Systems and Operations
Stanford, CA 94305-4023
shachter@stanford.edu shachter@stanford.edu

Abstract

One of the most useful sensitivity analysis One of the most useful sensitivity analysis
techniques of decision analysis is the com-
putation of value of information (or clair-One of the most useful sensitivity analysis
techniques of decision analysis is the com-
putation of value of information (or clair-
vovance) the difference in value obtained by putation of value of information (or clair-
voyance), the difference in value obtained by putation of value of information (or clair-
voyance), the difference in value obtained by
changing the decisions by which some of the
uncertainties are observed. In this paper voyance), the difference in value obtained by
changing the decisions by which some of the
uncertainties are observed. In this paper,
some simple but powerful extensions to prechanging the decisions by which some of the
uncertainties are observed. In this paper,
some simple but powerful extensions to pre-
vious algorithms are introduced which allow uncertainties are observed. In this paper,
some simple but powerful extensions to pre-
vious algorithms are introduced which allow an efficient value of information calculation vious algorithms are introduced which allow
an efficient value of information calculation
on the rooted cluster tree (or strong junction
tree) used to solve the original decision proban efficient value of information calculation
on the rooted cluster tree (or strong junction
tree) used to solve the original decision prob-
lem lem.

Keywords: value of information, clairvoyance, cluster trees, junction trees, decision analysis, influence diagrams.

1 Introduction

The analysis of sequential decision making under un-
The analysis of sequential decision making under un-
certainty is closely related to the analysis of probabilis-The analysis of sequential decision making under uncertainty is closely related to the analysis of probabilistic inference. In fact, much of the research into efficient The analysis of sequential decision making under uncertainty is closely related to the analysis of probabilistic inference. In fact, much of the research into efficient methods for probabilistic inference in expert systems certainty is closely related to the analysis of probabilistic inference. In fact, much of the research into efficient
methods for probabilistic inference in expert systems
has been motivated by the fundamental normative ar tic inference. In fact, much of the research into efficient
methods for probabilistic inference in expert systems
has been motivated by the fundamental normative ar-
guments of decision theory. Previous research has anmethods for probabilistic inference in expert systems
has been motivated by the fundamental normative arguments of decision theory. Previous research has ap-
plied those developments by modifying algorithms for has been motivated by the fundamental normative arguments of decision theory. Previous research has applied those developments by modifying algorithms for efficient probabilistic inference on belief networks to guments of decision theory. Previous research has applied those developments by modifying algorithms for efficient probabilistic inference on belief networks to address decision making problems represented by inplied those developments by modifying algorithms for
efficient probabilistic inference on belief networks to
address decision making problems represented by influence diagrams (Jensen and others 1994; Ndilikilikeaddress decision making problems represented by influence diagrams (Jensen and others 1994; Ndilikilikesha 1991; Shachter and Ndilikilikesha 1993; Shachter and Peot 1992: Shenov 1992) fluence diagrams (Jensen and ot
sha 1991; Shachter and Ndilikil
and Peot 1992; Shenoy 1992).

and Peot 1992; Shenoy 1992).
One of the most useful sensitivity analysis techniques of decision analysis is the computation of value of information (or clairvoyance), the difference in value obof decision analysis is the computation of value of information (or clairvoyance), the difference in value obtained by changing the decisions by which some of the uncertainties are observed (Baiffa 1968). In this paper formation (or clairvoyance), the difference in value obtained by changing the decisions by which some of the uncertainties are observed (Raiffa 1968). In this paper, some simple but powerful extensions to previous algotained by changing the decisions by which some of the uncertainties are observed (Raiffa 1968). In this paper, some simple but powerful extensions to previous algo-
rithms are introduced which allow an efficient value of uncertainties are observed (Raiffa 1968). In this paper,
some simple but powerful extensions to previous algo-
rithms are introduced which allow an efficient value of
information calculation on the rooted cluster tree (or
 some simple but powerful extensions to previous algorithms are introduced which allow an efficient value of information calculation on the rooted cluster tree (or strong junction tree) used to solve the original decision rithms are introduced which allow an efficient value of problem.

Dittmer and Jensen(1997) proposed that multiple
value of information calculations could all be per-Dittmer and Jensen(1997) proposed that multiple value of information calculations could all be per-
formed using the same tree. It is this idea that this Dittmer and Jensen(1997) proposed that multiple value of information calculations could all be performed using the same tree. It is this idea that this paper builds on value of information calculations could all be per-
formed using the same tree. It is this idea that this
paper builds on.

Section 2 presents a brief introduction of influence di-
Section 2 presents a brief introduction of influence di-
agrams and Section 3 reviews the most efficient meth-Section 2 presents a brief introduction of influence di-
agrams and Section 3 reviews the most efficient meth-
ods for solving them Section 4 develops some new Section 2 presents a brief introduction of influence diagrams and Section 3 reviews the most efficient methods for solving them. Section 4 develops some new results which are in applied in Section 5 to efficiently agrams and Section 3 reviews the most efficient methods for solving them. Section 4 develops some new results which are in applied in Section 5 to efficiently perform multiple value of information calculations. Fiods for solving them. Section 4 develops some new
results which are in applied in Section 5 to efficiently
perform multiple value of information calculations. Fi-
nally Section 6 provides some suggestions for future results which are in applied in Section 5 to efficiently
perform multiple value of information calculations. Fi-
nally, Section 6 provides some suggestions for future
research research.

2 Influence Diagrams

2 Influence Diagrams
Influence diagrams are graphical representations for
decision problems under uncertainty. In this section Influence diagrams are graphical representations for
decision problems under uncertainty. In this section
the components and notation of influence diagrams are Influence diagrams are graphical representations for
decision problems under uncertainty. In this section
the components and notation of influence diagrams are
briefly introduced. The graphical structure of the indecision problems under uncertainty. In this section
the components and notation of influence diagrams are
briefly introduced. The graphical structure of the in-
fluence diagram reveals conditional independence and the components and notation of influence diagrams are
briefly introduced. The graphical structure of the in-
fluence diagram reveals conditional independence and
the information available at the time decisions must be briefly introduced. The graphical structure of the influence diagram reveals conditional independence and
the information available at the time decisions must be
taken. This is a cursory introduction and the reader fluence diagram reveals conditional independence and
the information available at the time decisions must be
taken. This is a cursory introduction and the reader
is referred to the relevant literature for more informathe information available at the time decisions must be taken. This is a cursory introduction and the reader is referred to the relevant literature for more information. tion. is referred to the relevant literature for more informa-
tion.
An *influence diagram* is a directed graph network rep-

An *influence diagram* is a directed graph network rep-
resenting a single decision maker's beliefs and prefer-
ences about a sequence of decisions to be made un-An *influence diagram* is a directed graph network rep-
resenting a single decision maker's beliefs and prefer-
ences about a sequence of decisions to be made un-
der uncertainty (Howard and Matheson 1984) The resenting a single decision maker's beliefs and preferences about a sequence of decisions to be made under uncertainty (Howard and Matheson 1984). The nodes in the influence diagram represent variablesences about a sequence of decisions to be made under uncertainty (Howard and Matheson 1984). The nodes in the influence diagram represent variables–
uncertainties (drawn as ovals) decisions (drawn as der uncertainty (Howard and Matheson 1984). The
nodes in the influence diagram represent variables-
uncertainties (drawn as ovals), decisions (drawn as
rectangles) and the criterion values for making decinodes in the influence diagram represent variables-
uncertainties (drawn as ovals), decisions (drawn as
rectangles), and the criterion values for making deci-
sions (drawn as diamonds). The parents of uncertainuncertainties (drawn as ovals), decisions (drawn as
rectangles), and the criterion values for making deci-
sions (drawn as diamonds). The parents of uncertain-
ties and values condition their distributions, while the sions (drawn as diamonds). The parents of uncertainsions (drawn as diamonds). The parents of uncertain-
ties and values condition their distributions, while the
parents of decisions represent those variables that will
be observed before the decision must be made. The ties and values condition their distributions, while the parents of decisions represent those variables that will
be observed before the decision must be made. The
value represents the expected utility of its parents and parents of decisions represent those variables that will
be observed before the decision must be made. The
value represents the expected utility of its parents, and be observed before the decision must be made. The value represents the expected utility of its parents, and decisions are made to maximize this expected utility. When there are multiple value podes the total utility value represents the expected utility of its parents, and
decisions are made to maximize this expected utility.
When there are multiple value nodes, the total utility
is the sum of the utilities for each value (The results decisions are made to maximize this expected utility.
When there are multiple value nodes, the total utility
is the sum of the utilities for each value. (The results in
this paper could also be applied to products (Shachte When there are multiple value nodes, the total utility
is the sum of the utilities for each value. (The results in
this paper could also be applied to products (Shachter
and Peot 1992: Tatman and Shachter 1990)) and Peot 1992; Tatman and Shachter 1990).)

Figure 1 and Peot 1992; Tatman and Shachter 1990).)
Consider the influence diagram shown in Figure 1 from
Dittmer and Jensen(1997). There are four uncertain-Consider the influence diagram shown in Figure 1 from
Dittmer and Jensen(1997). There are four uncertain-
ties A B C and E three decisions D_1, D_2 and D_3 Dittmer and Jensen(1997). There are four uncertainties, A, B, C , and E , three decisions, D_1, D_2 , and D_3 ,

Figure 1: The example influence diagram from Dittmer and Jensen(1997).

Figure 2: The influence diagram from Figure 1 with clairvoyance on B before D_1 is chosen.

and a single value, U . The decisions are ordered in the graph and information available at the time of one and a single value, U . The decisions are ordered in the graph and information available at the time of one decision is remembered for subsequent decisions the and a single value, U . The decisions are ordered in
the graph and information available at the time of one
decision is remembered for subsequent decisions, the
no forgetting principle. For example, D_1 and C are obthe graph and information available at the time of one
decision is remembered for subsequent decisions, the
no forgetting principle. For example, D_1 and C are ob-
served before both D_2 and D_2 , while D_2 , E decision is remembered for subsequent decisions, the

<u>no forgetting</u> principle. For example, D_1 and C are ob-

served before both D_2 and D_3 , while D_2 , E, and A are

observed before only D_2 . None of the var no forgetting principle. For example, D_1 and C are ob-
served before both D_2 and D_3 , while D_2 , E, and A are
observed before only D_3 . None of the variables are ob-
served before D_1 is chosen. Not all of t served before both D_2 and D_3 , while D_2 , E , and A are observed before only D_3 . None of the variables are observed before D_1 is chosen. Not all of the observations are really needed or *requisite* for a d observed before only D_3 . None of the variables are observed before D_1 is chosen. Not all of the observations are really needed or *requisite* for a decision. For example, although five of the variables are observed served before D_1 is chosen. Not all of the observations are really needed or *requisite* for a decision. For example, although five of the variables are observed before D_3 is chosen, A is the only requisite observation–once A has been observed the other variables provide no ple, although five of the variables are observed before D_3 is chosen, A is the only requisite observation–once A has been observed, the other variables provide no additional information. Similarly C is the only requi- D_3 is chosen, A is the only requisite observation-once
A has been observed, the other variables provide no
additional information. Similarly, C is the only requi-
site observation for D_2 . The diagram can be analyzed A has been observed, the other variables provide no
additional information. Similarly, C is the only requi-
site observation for D_2 . The diagram can be analyzed
to determine the maximal expected utility. If the utiladditional information. Similarly, C is the only requisite observation for D_2 . The diagram can be analyzed to determine the maximal expected utility. If the util-
ity does not represent dollars we could convert it to site observation for D_2 . The diagram can be analyzed
to determine the maximal expected utility. If the util-
ity does not represent dollars, we could convert it to
dollars by applying the inverse of the utility functio to determine the maximal expected utility. If the utility does not represent dollars, we could convert it to dollars by applying the inverse of the utility function that maps from dollars to utility.

Modulars by applying the inverse of the distribution
that maps from dollars to utility.
We can solve a different decision problem without
changing any of the distributions in the uncertainties chat maps from donars to during.
We can solve a different decision problem without
changing any of the distributions in the uncertainties
and values by changing the informational assumptions We can solve a different decision problem without
changing any of the distributions in the uncertainties
and values by changing the informational assumptions.
For example in Figure 2 B is now observed before changing any of the distributions in the uncertainties
and values by changing the informational assumptions.
For example, in Figure 2 B is now observed before and values by changing the informational assumptions. a
For example, in Figure 2 B is now observed before w
 D_1 is chosen. The expected utility from this diagram timust be at least as much as from the earlier diagram For example, in Figure 2 B is now observed before D_1 is chosen. The expected utility from this diagram timust be at least as much as from the earlier diagram hecause of this extra information the opportunity to D_1 is chosen. The expected utility from this diagram
must be at least as much as from the earlier diagram
because of this extra information, the opportunity to
observe B . The influence diagram makes it explicit must be at least as much as from the earlier diagram
because of this extra information, the opportunity to
observe B . The influence diagram makes it explicit
what information is available and when it is available because of this extra information, the opportunity to observe B . The influence diagram makes it explicit what information is available and when it is available in the two diagrams. This extra value leads to a differobserve B . The influence diagram makes it explicit ence in dollar values called the value of information or in the two diagrams. This extra value leads to a difference in dollar values called the *value of information* or *value of clairvoyance*. Technically, the value of information is only approximated by this difference (Bai ence in dollar values called the *value of information* or
value of clairvoyance. Technically, the value of infor-
mation is only approximated by this difference (Raiffa
1968) but we will work with this approximated val *value of clairvoyance*. Technically, the value of infor-
mation is only approximated by this difference (Raiffa
1968), but we will work with this approximated value.
Without any new assessments the decision problem mation is only approximated by this difference (Raiffa
1968), but we will work with this approximated value.
Without any new assessments, the decision problem
can thus be solved many times varying the informate 1968), but we will work with this approximated value.
Without any new assessments, the decision problem
can thus be solved many times, varying the informa-
tional assumptions for one variable at a time. This is Without any new assessments, the decision problem
can thus be solved many times, varying the informa-
tional assumptions for one variable at a time. This is
the process this paper seeks to perform efficiently can thus be solved many times, varying the inform
tional assumptions for one variable at a time. This
the process this paper seeks to perform efficiently. tional assumptions for one variable at a time. This is
the process this paper seeks to perform efficiently.
Another influence diagram example that will appear

In this paper is shown in Figure 3 (Jensen and oth-
in this paper is shown in Figure 3 (Jensen and oth-
ers 1994) This diagram has four value nodes whose Another influence diagram example that will appear
in this paper is shown in Figure 3 (Jensen and oth-
ers 1994). This diagram has four value nodes, whose

Figure 3: The example influence diagram from Jensen
et al(1994) Figure 3: The
et al(1994).

functions are summed to obtain the expected utility.

functions are summed to obtain the expected utility.
The influence diagram has been developed as a practi-
cal representation for a decision problem, and to that The influence diagram has been developed as a practical representation for a decision problem, and to that The influence diagram has been developed as a practical representation for a decision problem, and to that end there are several semantic restrictions, which are described in detail elsewhere (Howard and Matheson cal representation for a decision problem, and to that
end there are several semantic restrictions, which are
described in detail elsewhere (Howard and Matheson
1984: Shachter 1986) In particular, we cannot obend there are several semantic restrictions, which are
described in detail elsewhere (Howard and Matheson
1984; Shachter 1986). In particular, we cannot ob-
serve the descendant of a decision before making the described in detail elsewhere (Howard and Matheson
1984; Shachter 1986). In particular, we cannot ob-
serve the descendant of a decision before making the
decision since the decision can affect its descendants 1984; Shachter 1986). In particular, we cannot observe the descendant of a decision before making the decision, since the decision can affect its descendants. The one excention is when the descendant represents a serve the descendant of a decision before making the decision, since the decision can affect its descendants.
The one exception is when the descendant represents a constraint and is a deterministic function of the decidecision, since the decision can affect its descendants.
The one exception is when the descendant represents a
constraint and is a deterministic function of the deci-
sion and its requisite observations. But this case coul The one exception is when the descendant represents a constraint and is a deterministic function of the decision and its requisite observations. But this case could constraint and is a deterministic function of the decision and its requisite observations. But this case could
be modeled as a value node (with certain cases hav-
ing probibitive value) instead of as an observation and sion and its requisite observations. But this case could
be modeled as a value node (with certain cases hav-
ing prohibitive value) instead of as an observation and
thus we can exclude it without loss of generality be modeled as a value node (with certain cases h
ing prohibitive value) instead of as an observation ε
thus we can exclude it without loss of generality. thus we can exclude it without loss of generality.
 3 Rooted Cluster Trees

Efficient algorithms have been developed to solve decision problems represented as influence diagrams. Efficient algorithms have been developed to solve decision problems represented as influence diagrams.
These algorithms build an auxiliary structure called
a rooted cluster tree or strong junction tree. Previous cision problems represented as influence diagrams.
These algorithms build an auxiliary structure called
a rooted cluster tree or strong junction tree. Previous
work has suggested how value of information calcula-These algorithms build an auxiliary structure called
a rooted cluster tree or strong junction tree. Previous
work has suggested how value of information calcula-
tions could be performed efficiently on such a tree a rooted cluster tree or strong junction tree. Previous work has suggested how value of information calculations could be performed efficiently on such a tree.

work has suggested how value of miorihation calcula-
tions could be performed efficiently on such a tree.
Although the influence diagram can be solved directly
(Shachter 1986) the most efficient procedures work on Although the influence diagram can be solved directly
(Shachter 1986), the most efficient procedures work on
related graphical structures (Jensen and others 1994. Although the influence diagram can be solved directly
(Shachter 1986), the most efficient procedures work on
related graphical structures (Jensen and others 1994;
Ndilikilikesha 1991: Shachter and Ndilikilikesha 1993; (Shachter 1986), the most efficient procedures work on related graphical structures (Jensen and others 1994;
Ndilikilikesha 1991; Shachter and Ndilikilikesha 1993;
Shachter and Peot 1992; Shenov 1992). This paper related graphical structures (Jensen and others 1994;
Ndilikilikesha 1991; Shachter and Ndilikilikesha 1993;
Shachter and Peot 1992; Shenoy 1992). This paper Ndilikilikesha 1991; Shachter and Ndilikilikesha 1993;
Shachter and Peot 1992; Shenoy 1992). This paper
considers one of those graphical structures, the rooted
cluster tree, a slight generalization of the strong iunc-Shachter and Peot 1992; Shenoy 1992). This paper
considers one of those graphical structures, the rooted
cluster tree, a slight generalization of the strong junc-
tion tree cluster tree, a slight generalization of the strong junction tree.

A set of variables is called a *cluster*. A tree of clus-
A set of variables is called a *cluster*. A tree of clus-
ters is called a *cluster tree* (or join tree) if every de-A set of variables is called a *cluster*. A tree of clusters is called a *cluster tree* (or join tree) if every decision or uncertainty appears somewhere in the tree¹. A set of variables is called a *cluster*. A tree of clusters is called a *cluster tree* (or join tree) if every decision or uncertainty appears somewhere in the tree¹,

¹If the cluster tree were not being constructed to compute value of information, it might be worthwhile to exclude vari-¹If the cluster tree were not being constructed to compute value of information, it might be worthwhile to exclude variables determined to be extraneous, but here it is desirable to keep all of the variables in the mode value of information, it might be worthwishles determined to be extraneous, but
to keep all of the variables in the model.

Figure 4: Rooted cluster tree for the influence diagram Figure 5: Rooted cluster tree for the inf
from Figure 1 from Dittmer and Jensen(1997) from Figure 3 from Jensen et al(1994). from Figure 1 from Dittmer and Jensen(1997)

each uncertainty and its parents appear together in at
least one cluster, and any variable that appears in two each uncertainty and its parents appear together in at
least one cluster, and any variable that appears in two
different clusters appears in all of the clusters on the each uncertainty and its parents appear together in at
least one cluster, and any variable that appears in two
different clusters appears in all of the clusters on the
path between them. Corresponding to the notation in least one cluster, and any variable that appears in two
different clusters appears in all of the clusters on the
path between them. Corresponding to the notation in different clusters appears in all of the clusters on the path between them. Corresponding to the notation in Jensen et al(1994), there are two *potential functions* associated with each cluster C a probability potenpath between them. Corresponding to the notation in
Jensen et al(1994), there are two *potential functions*
associated with each cluster C, a probability poten-
tial ϕ_C and a utility potential ψ_C . This paper will Jensen et al(1994), there are two <u>potential functions</u>
associated with each cluster C, a probability poten-
tial, ϕ_C , and a utility potential, ψ_C . This paper will
introduce and present a minimal amount of this nota associated with each cluster C, a probability potential, ϕ_C , and a utility potential, ψ_C . This paper will introduce and present a minimal amount of this notation instead focusing on other extensions to Jensen et tial, ϕ_C , and a utility potential, ψ_C . This paper will
introduce and present a minimal amount of this nota-
tion, instead focusing on other extensions to Jensen et
al(1994) All of the tables in the influence diagram introduce and present a minimal amount of this notation, instead focusing on other extensions to Jensen et al(1994). All of the tables in the influence diagram are incorporated into these potential functions. al(1994). All of the tables in the influence diagram are

 T_{max} and T_{max} . The cluster tree is *rooted* if the arcs between clusters
The cluster tree is *<u>rooted</u>* if the arcs between clusters
are directed so that one cluster the *root cluster* has The cluster tree is <u>rooted</u> if the arcs between clusters
are directed so that one cluster, the <u>root cluster</u>, has
no children, and all of the other clusters have exactly The cluster tree is <u>rooted</u> if the arcs between clusters
are directed so that one cluster, the <u>root cluster</u>, has
no children, and all of the other clusters have exactly
one child. It is useful to distinguish between clu are directed so that one cluster, the <u>root cluster</u>, has no children, and all of the other clusters have exactly one child. It is useful to distinguish between clusters and variables by their location relative to the root one child. It is useful to distinguish between clusters one child. It is useful to distinguish between clusters
and variables by their location relative to the root.
Cluster C is <u>inward</u> of another cluster C' in a rooted
cluster tree if C is either the root cluster or between and variables by their location relative to the root.
Cluster C is <u>inward</u> of another cluster C' in a rooted
cluster tree if C is either the root cluster or between
the root cluster and C' In that case C' is said to Cluster C is <u>inward</u> of another cluster C' in a rooted
cluster tree if C is either the root cluster or between
the root cluster and C'. In that case C' is said to
be *outward* of C. If all clusters containing a variable cluster tree if C is either the root cluster or between
the root cluster and C'. In that case C' is said to
be <u>outward</u> of C. If all clusters containing a variable
A are outward of some cluster containing a variable B A are outward of some cluster order of the container of the contact of the method itself the root cluster and C'. In that case C' is said to from C's outward neighbors, and marginalize all variable A are outward of some be <u>outward</u> of C. If all clusters containing a variable A are outward of some cluster containing a variable B then A is *strictly outward* of B and B is *strictly inward* of A . If all clusters containing A A are outward of some cluster containing a variable B
then A is *strictly outward* of B and B is *strictly inward*
of A . If all clusters containing A either contain B
or are outward of a cluster containing then A is *strictly outward* of B and B is *strictly inward*
of A. If all clusters containing A either contain B
or are outward of a cluster containing B, then A is
weakly outward of B and B is *weakly inward* of A For of A. If all clusters containing A either contain B or are outward of a cluster containing B , then A is *weakly outward* of B and B is *weakly inward* of A . For or are outward of a cluster containing B , then A is *weakly outward* of B and B is *weakly inward* of A . For example in Figure 5, k is strictly outward of h , strictly inward of i and neither weakly inward weakly outward of B and B is weakly inward of A. For
example in Figure 5, k is strictly outward of h, strictly
inward of j, and neither weakly inward nor weakly
outward of a inward of j , and neither weakly inward nor weakly outward of q .

There are other restrictions that have been developed
for rooted cluster trees, but for simplicity only the fol-
lowing, new definition will be presented here. A rooted for rooted cluster trees, but for simplicity only the folfor rooted cluster trees, but for simplicity only the following, new definition will be presented here. A rooted cluster tree is *properly constructed* for an influence di-
agram if lowing, new o
cluster tree is
agram if

- D must be chosen before D' ;
- 2. decision *D* is strictly inward of decision *D* only if
 D must be chosen before *D'*;

2. decision *D* is weakly inward of uncertainty *A* if *A* is a descendant of *D* in the influence diagram. D must be chosen before D;
decision D is weakly inward of uncertainty A if λ
is a descendant of D in the influence diagram;
- 2. decision *D* is weakly inward of uncertainty *A* if *A* is a descendant of *D* in the influence diagram;
3. decision *D* is not strictly inward of uncertainty *A* if *A* will be observed before *D* is chosen: if a descendant of D in the influence diagretical
decision D is not strictly inward of uncertaint A will be observed before D is chosen; 3. decision D is not strictly inward of uncertainty A
if A will be observed before D is chosen;
4. decision D and its requisite observations are all
contained in some cluster: and
- If A will be observed before D is contained in some cluster; and

Figure 5: Rooted cluster tree for the influence diagram
from Figure 3 from Jensen et al(1994). Figure 5: Rooted cluster tree for the influ
from Figure 3 from Jensen et al(1994).

5. any variable A strictly inward of decision D and
also in a cluster with D is observed when D is any variable A strictly inward of decision D and also in a cluster with D is observed when D is chosen chosen.

chosen.
Rooted cluster trees properly constructed for the in-
fluence diagrams from Section 2 are shown in Figure 4 Rooted cluster trees properly constructed for the influence diagrams from Section 2 are shown in Figure 4
and Figure 5 The influence diagram's value can then Rooted cluster trees properly constructed for the in-
fluence diagrams from Section 2 are shown in Figure 4
and Figure 5. The influence diagram's value can then
be determined by making a single sweep through the fluence diagrams from Section 2 are shown in Figure 4
and Figure 5. The influence diagram's value can then
be determined by making a single sweep through the
rooted cluster tree toward the root, as summarized be determined by making a single sweep through the rooted cluster tree toward the root, as summarized in Algorithm 1. The marginalization operator is debe determined by making a single sweep through the rooted cluster tree toward the root, as summarized in Algorithm 1. The marginalization operator is described in Jensen et al(1994). rooted cluster tree toward the
in Algorithm 1. The marginal
scribed in Jensen et al(1994).

scribed in Jensen et al(1994).
 Algorithm 1 (Value Calculation) This algorithm

commutes the ontimal expected value on a properly con-**Algorithm 1 (Value Calculation)** This algorithm computes the optimal expected value on a properly constructed moted cluster tree $\bf Algorithm~1~(Value~Calc~corrected~rooted~cluster~tree.$ computes the optimal expected value on a properly constructed rooted cluster tree.
Visit each cluster C in the tree working inward from

 t is the leaves toward the root.
 t is the leaves toward the root. That is, choose any cluster
 t o wisit whose outward neighbors have glready been wis-Visit each cluster C in the tree working inward from
the leaves toward the root. That is, choose any cluster
to visit whose outward neighbors have already been vis-
ited When wisiting a cluster incorparate the undates the leaves toward the root. That is, choose any cluster
to visit whose outward neighbors have already been vis-
ited. When visiting a cluster, incorporate the updates
from C's outward neighbors, and marginalize all varifor visit whose outward neighbors have already been visited. When visiting a cluster, incorporate the updates from C's outward neighbors, and marginalize all vari-
ables that do not amear in C's inward neighbor in an from C 's outward neighbors, and marginalize all variables that do not appear in C 's inward neighbor in an

At the end, the root cluster computes two scalar updates, ϕ_{\emptyset} representing the probability of the evidence At the end, the root cluster computes two scalar up-
dates, ϕ_{\emptyset} representing the probability of the evidence
and ψ_{\emptyset} , where $\psi_{\emptyset}/\phi_{\emptyset}$ is the expected utility of the op-
timal strategy. For value of info dates, ϕ_{\emptyset} representing the probability of the evidence
and ψ_{\emptyset} , where $\psi_{\emptyset}/\phi_{\emptyset}$ is the expected utility of the op-
timal strategy. For value of information calculations,
this latter quantitu can be use and ψ_{\emptyset} , where $\psi_{\emptyset}/\phi_{\emptyset}$ is the expected utility of the optimal strategy. For value of information calculations,
this latter quantity can be used directly or it can be
converted to units of dollars (by applyi timal strategy. For value of information calculations,
this latter quantity can be used directly or it can be
converted to units of dollars (by applying the inverse
utility function). converted to units of dollars (by applying the inverse

agram if

1. decision D is strictly inward of decision D' only if

¹ cxample, the tree in Figure 6 has uncertainty B added

² cxample, the tree in Figure 6 has uncertainty B added utility function).
This algorithm is generalized in Dittmer and Jensen
(1997) to perform multiple value of information calcu-This algorithm is generalized in Dittmer and Jensen
(1997) to perform multiple value of information calcu-
lations with only one cluster tree. The variable(s) to This algorithm is generalized in Dittmer and Jensen
(1997) to perform multiple value of information calculations with only one cluster tree. The variable(s) to
be observed earlier are added to inward clusters. For (1997) to perform multiple value of information calculations with only one cluster tree. The variable(s) to to the three clusters where it did not appear before.

Example, the tree in Figure 5 has uncertainty *D* added
to the three clusters where it did not appear before.
It is not exactly clear how this expanded cluster tree
should be processed According to Dittmer and Jensen should be processed. According to Dittmer and Jensen
should be processed. According to Dittmer and Jensen
(1997) "As mentioned earlier" a control structure is It is not exactly clear how this expanded cluster tree
should be processed. According to Dittmer and Jensen
(1997), "As mentioned earlier, a control structure is
associated with the (strong) iunction tree. This strucshould be processed. According to Dittmer and Jensen
(1997), "As mentioned earlier, a control structure is
associated with the (strong) junction tree. This struc-
ture handles the order of marginalization, and there- (1997) , "As mentioned earlier, a control structure is associated with the (strong) junction tree. This structure handles the order of marginalization, and there-
fore we can use the expanded iunction tree (and the associated with the (strong) junction tree. This structure handles the order of marginalization, and there-
fore we can use the expanded junction tree (and the
associated control structure) in Figure 7c to marginalture handles the order of marginalization, and therefore we can use the expanded junction tree (and the associated control structure) in Figure 7c to marginal-

Figure 6: Expanded rooted cluster tree for the influ-Figure 6: Expanded rooted cluster tree for the influ-
ence diagram from Figure 1. Those clusters changed
from the rooted cluster tree in Figure 4 are shaded Figure 6: Expanded rooted cluster tree for the influ-
ence diagram from Figure 1. Those clusters changed
from the rooted cluster tree in Figure 4 are shaded.

ence diagram from Figure 1.

ize B from any clique of our choice. After B has been ize B from any clique of our choice. After B has been
marginalized from a clique, the table space reserved for
 B in cliques closer to the strong root is obsolete. Clever ize B from any clique of our choice. After B has been
marginalized from a clique, the table space reserved for
 B in cliques closer to the strong root is obsolete. Clever
use of the control structures will prevent ca marginalized from a clique, the table space reserved for B in cliques closer to the strong root is obsolete. Clever use of the control structures will prevent calculations to take place in the remaining table expansions B in cliques closer to the strong root is obsolete. Clever
use of the control structures will prevent calculations
to take place in the remaining table expansions, and
the number of table operations in the remaining sub use of the control structures will prevent calculations
to take place in the remaining table expansions, and
the number of table operations in the remaining subtree equals that of an ordinary strong junction tree."

4 New Results

The definition for a properly constructed rooted clus-The definition for a properly constructed rooted cluster tree introduced in Section 3 allows the derivation of some simple but powerful results that will be approved. The definition for a properly constructed rooted cluster tree introduced in Section 3 allows the derivation of some simple but powerful results that will be applied to perform value of information calculations in ter tree introduced in Section 3 allows the derivation
of some simple but powerful results that will be ap-
plied to perform value of information calculations in
Section 5 But first it will be helpful to build the best of some simple but powerful results that will be applied to perform value of information calculations in Section 5. But first it will be helpful to build the best possible rooted cluster trees for the original influence possible to perform value of information calculations in Section 5. But first it will be helpful to build the best possible rooted cluster trees for the original influence diagram.

Figure 8: Finding requisite observations for the influ-
ence diagram from Figure 3 Figure 8: Finding requisite obse
ence diagram from Figure 3.

The first step in building a cluster tree is recognizing
which observations are requisite for the different de-The first step in building a cluster tree is recognizing
which observations are requisite for the different de-
cisions – Although the BayesBall algorithm (Shachter The first step in building a cluster tree is recognizing
which observations are requisite for the different de-
cisions. Although the BayesBall algorithm (Shachter
1998) is fast (linear time in the size of the graph) it which observations are requisite for the different decisions. Although the BayesBall algorithm (Shachter 1998) is fast (linear time in the size of the graph), it is conservative in computing requisite observations ascisions. Although the BayesBall algorithm (Shachter 1998) is fast (linear time in the size of the graph), it is conservative in computing requisite observations, as-
suming that the value sets are nested. A less conservais conservative in computing requisite observations, assuming that the value sets are nested. A less conservais conservative in computing requisite observations, assuming that the value sets are nested. A less conservative algorithm can be fashioned by teaming BayesBall with the reductions in Tatman and Shachter(1990) suming that the value sets are nested. A less conservative algorithm can be fashioned by teaming BayesBall with the reductions in Tatman and Shachter(1990).

with the reductions in Tatman and Shachter(1990).
 Algorithm 2 (Requisite Observations) This al-
 aorithm determines the requisite observations for each **Algorithm 2 (Requisite Observations)** This algorithm determines the requisite observations for each decision in an influence diagram as a prelude to proper **Algorithm 2 (Requisite Observations)** This algorithm determines the requisite observations for each decision in an influence diagram as a prelude to proper construction of a rooted cluster tree. It runs in time gorithm determines the requisite observations for each decision in an influence diagram as a prelude to proper construction of a rooted cluster tree. It runs in time $O((\text{number of decisions})(\text{graph size}))$ decision in an influence diagram as a prelude to proper
construction of a rooted cluster tree. It runs in time
 $O((\text{number of decisions})(\text{graph size}))$.
Visit each decision D_i in reverse chronological order,
 $i = m - 1$ Let V; be the set of valu

 $\begin{array}{l} \n\text{Using } \mathcal{L}(\mathcal{L}) \leq \mathcal{L}(\mathcal{L$ Visit each decision D_i in reverse chronological order,
 $i = m, \ldots, 1$. Let V_i be the set of value descendants

of D in the current diagram. Run the BayesBall al-

oorithm on V_i given D_i and L_i the variables obs $i = m, \ldots, 1$. Let V_i be the set of value descendants
of D in the current diagram. Run the BayesBall al-
gorithm on V_i given D_i and I_i , the variables observed
before D_i is chosen and let R_i be the requisite o of D in the current diagram. Run the BayesBall algorithm on V_i given D_i and I_i , the variables observed
before D_i is chosen, and let R_i be the requisite obser-
vations (not including D_i). Benlace D_i by a chan gorithm on V_i given D_i and I_i , the variables observed
before D_i is chosen, and let R_i be the requisite obser-
vations (not including D_i). Replace D_i by a chance

Figure 9: Moral graph based on the modified version
of the influence diagram from Figure 1. The value Figure 9: Moral graph based on the modified version
of the influence diagram from Figure 1. The value
node has been removed, requisite informational arcs Figure 9: Moral graph based on the modified version
of the influence diagram from Figure 1. The value
node has been removed, requisite informational arcs
are drawn as heavy lines, and the moralizing arc is of the influence diagram from Figure 1. The value
node has been removed, requisite informational arcs
are drawn as heavy lines, and the moralizing arc is
drawn as dashed line node has been removed
are drawn as heavy lindrawn as dashed line.

 $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
Figure 10: Moral graph based on the modified version
of the influence diagram from Figure 3 . Value nodes Figure 10: Moral graph based on the modified version
of the influence diagram from Figure 3. Value nodes
have been removed requisite informational arcs are Figure 10: Moral graph based on the modified version
of the influence diagram from Figure 3. Value nodes
have been removed, requisite informational arcs are
drawn as heavy lines, and moralizing arcs are drawn of the influence diagram from Figure 3. Value nodes
have been removed, requisite informational arcs are
drawn as heavy lines, and moralizing arcs are drawn
as dashed lines drawn as heavy lines, and moralizing arcs are drawn as dashed lines.

node "policy" with R_i as parents and proceed to the next earlier decision $node$ "policy" with R_i
next earlier decision.

next earlier decision.
This algorithm is applied to the two influence diagrams
from Section 2 as shown in Figure 7 and Figure 8 This algorithm is applied to the two influence diagrams
from Section 2, as shown in Figure 7 and Figure 8.
In the figures, the value descendants for a particular This algorithm is applied to the two influence diagrams
from Section 2, as shown in Figure 7 and Figure 8.
In the figures, the value descendants for a particular
decision are highlighted and the decision and its obfrom Section 2, as shown in Figure 7 and Figure 8.
In the figures, the value descendants for a particular
decision are highlighted, and the decision and its ob-
servations are shaded. In Figure 7, it can be seen that In the figures, the value descendants for a particular decision are highlighted, and the decision and its observations are shaded. In Figure 7, it can be seen that the requisite observations are $R_2 = \{A\}$, $R_3 = \{C\}$ decision are highlighted, and the decision and its observations are shaded. In Figure 7, it can be seen that
the requisite observations are $R_3 = \{A\}$, $R_2 = \{C\}$,
and $R_1 = \emptyset$. In Figure 8, the requisite observations
 the requisite observations are $R_3 = \{A\}$, $R_2 = \{C\}$, and $R_1 = \emptyset$. In Figure 8, the requisite observations are $R_4 = \{g, D_2\}$, $R_3 = \{f\}$, $R_2 = \{e\}$, and $R_1 = \{b\}$. Note the different sets of value descendants.

The next step is to generate the moral graph of the modified diagram. These moral graphs are shown in The next step is to generate the moral graph of the
modified diagram. These moral graphs are shown in
Figure 9 and Figure 10. The heavy shaded arcs cor-The next step is to generate the moral graph of the modified diagram. These moral graphs are shown in Figure 9 and Figure 10. The heavy shaded arcs correspond to requisite observations the dashed lines are modified diagram. These moral graphs are shown in Figure 9 and Figure 10. The heavy shaded arcs correspond to requisite observations, the dashed lines are Figure 9 and Figure 10. The heavy shaded arcs correspond to requisite observations, the dashed lines are moralizing arcs (added between parents with a child in common) and the value nodes have been removed respond to requisite observations, the dashed lines are
moralizing arcs (added between parents with a child
in common), and the value nodes have been removed
(after any corresponding moralizing arcs were added) moralizing arcs (added between parents with a child
in common), and the value nodes have been removed
(after any corresponding moralizing arcs were added).

(after any corresponding moralizing arcs were added).
Rooted cluster tree can now be properly constructed
hased on the moral graphs of the modified diagrams aborted cluster tree can now be properly constructed
based on the moral graphs of the modified diagrams.
There is no efficient algorithm to generate such trees Rooted cluster tree can now be properly constructed
based on the moral graphs of the modified diagrams.
There is no efficient algorithm to generate such trees,
but the structure of the moral graph guides the process based on the moral graphs of the modified diagrams.
There is no efficient algorithm to generate such trees,
but the structure of the moral graph guides the process
(Jensen and others 1994). It can be shown, however, but the structure of the moral graph guides the process

that the method just presented can always yield at that the method just presented can always yield at least some properly constructed rooted cluster trees. Frank is enterpreferent constructed rooted cluster trees.
Theorem 1 (Requisite Observations) Al

gorithm 2 can be applied to an influence diagram to yield a rooted cluster tree properly constructed for the diagram.

Proof: It is sufficient to show how one such rooted **Proof:** It is sufficient to show how one such rooted cluster tree could be properly constructed for any in-
fluence diagram. At each step of the algorithm let **Proof:** It is sufficient to show how one such rooted cluster tree could be properly constructed for any influence diagram. At each step of the algorithm, let Ω be the non-value variables relevant to V as detercluster tree could be properly constructed for any in-
fluence diagram. At each step of the algorithm, let
 Q be the non-value variables relevant to V_i as deter-
mined by the BavesBall algorithm (If we are building fluence diagram. At each step of the algorithm, let Q be the non-value variables relevant to V_i as determined by the BayesBall algorithm. (If we are building a rooted cluster tree for potential value of informa- Q be the non-value variables relevant to V_i as determined by the BayesBall algorithm. (If we are building a rooted cluster tree for potential value of informamined by the BayesBall algorithm. (If we are building
a rooted cluster tree for potential value of informa-
tion queries also add to Q any descendants of D_i that
could be observed. Otherwise, apparently extranea rooted cluster tree for potential value of informa-
tion queries also add to Q any descendants of D_i that
could be observed. Otherwise, apparently extrane-
ous variables will not appear in the constructed tree tion queries also add to Q any descendants of D_i that
could be observed. Otherwise, apparently extrane-
ous variables will not appear in the constructed tree.)
Now let Q_i be those nodes in Q for the first time could be observed. Otherwise, apparently extrane-
ous variables will not appear in the constructed tree.)
Now let Q_i be those nodes in Q for the first time,
 $Q_i = Q \setminus (Q_{i+1} + \cdots + Q_m)$ Finally let Q_0 be any ous variables will not appear in the constructed tree.)
Now let Q_i be those nodes in Q for the first time,
 $Q_i = Q \setminus (Q_{i+1} \cup ... \cup Q_m)$. Finally, let Q_0 be any
nodes relevant to R_1 that have not been included in Now let Q_i be those nodes in Q for the first time,
 $Q_i = Q \setminus (Q_{i+1} \cup \ldots \cup Q_m)$. Finally, let Q_0 be any

nodes relevant to R_1 that have not been included in
 Q_1 $Q_i = Q \setminus (Q_i)$
nodes relevan
 $Q_1, \ldots, Q_m.$ nodes relevant to R_1 that have not been included in Q_1, \ldots, Q_m .
If the value sets are nested, that is, $V_1 \supseteq \ldots \supseteq V_m$, then the rooted cluster tree shown in Figure 11a is

properly constructed. Otherwise, $V_{i+1} \not\supseteq V_{i+2}$ if and then the rooted cluster tree shown in Figure 11a is
properly constructed. Otherwise, $V_{i+1} \not\supseteq V_{i+2}$ if and
only if $V_{i+1} \cap V_{i+2} = \emptyset$. Suppose that $V_{i+1} \cap V_{i+2} = \emptyset$
but $V_{i} \supseteq (V_{i+1} \cup V_{i+2})$. In that case, properly constructed. Otherwise, $V_{i+1} \not\supseteq V_{i+2}$ if and
only if $V_{i+1} \cap V_{i+2} = \emptyset$. Suppose that $V_{i+1} \cap V_{i+2} = \emptyset$
but $V_i \supseteq (V_{i+1} \cup V_{i+2})$. In that case, then the partial
tree shown in Figure 11b is properly only if $V_{i+1} \cap V_{i+2} = \emptyset$. Suppose that $V_{i+1} \cap V_{i+2} = \emptyset$
but $V_i \supseteq (V_{i+1} \cup V_{i+2})$. In that case, then the partial
tree shown in Figure 11b is properly constructed. \Box tree shown in Figure 11b is properly constructed. \Box

of course, the purpose of this exercise is to generate
more efficient rooted cluster trees. Examples of such
for the influence diagrams from Section 2 are shown Of course, the purpose of this exercise is to generate
more efficient rooted cluster trees. Examples of such
for the influence diagrams from Section 2 are shown
in Figure 12 and Figure 13. They are indeed more more efficient rooted cluster trees. Examples of such
for the influence diagrams from Section 2 are shown
in Figure 12 and Figure 13. They are indeed more
efficient than the rooted cluster trees in Figure 4 and in Figure 12 and Figure 13. They are indeed more in Figure 12 and Figure 13. They are indeed more efficient than the rooted cluster trees in Figure 4 and Figure 5, respectively, reducing the size of the cluster state spaces efficient than t
Figure 5, respe
state spaces. state spaces.
The rest of this section contains the derivation of three

state spaces.
The rest of this section contains the derivation of three
simple but powerful results, based on the definition of
properly constructed rooted cluster tree. First, the in-The rest of this section contains the derivation of three
simple but powerful results, based on the definition of
properly constructed rooted cluster tree. First, the in-
wardmost cluster with a particular decision must co simple but powerful results, based on the definition of properly constructed rooted cluster tree. First, the in-
wardmost cluster with a particular decision must con-
tain its requisite observations properly constructed rooted clus
wardmost cluster with a particu
tain its requisite observations.

tain its requisite observations.
 Lemma 1 (Current Requisite Observations)
 Given a rooted cluster tree properly constructed for Lemma 1 (Current Requisite Observations)
Given a rooted cluster tree properly constructed for

Figure 12: A more efficient rooted cluster tree for the
influence diagram from Figure 1. Figure 12: A more efficient rooted cl
influence diagram from Figure 1.

Figure 13: A more efficient rooted cluster tree for the
influence diagram from Figure 3 Figure 13: A more efficient rooted cl
influence diagram from Figure 3.

an influence diagram, all requisite observed variables
for decision D are contained in the inwardmost clusan influence diagram, all requisite observed variables
for decision D are contained in the inwardmost clus-
ter containing D. Furthermore, any variables in both an influence diagram, all requisite observed variables
for decision D are contained in the inwardmost clus-
ter containing D . Furthermore, any variables in both
that cluster and the nert inward cluster are observed for decision D are contained in the inwardmost cluster containing D . Furthermore, any variables in both that cluster and the next inward cluster are observed when D is chosen ter containing $D.$ F
that cluster and the
when D is chosen.

when D is chosen.
Proof: By proper construction, any variable observed
before D is chosen must be weakly inward of D and any **Proof:** By proper construction, any variable observed
before D is chosen must be weakly inward of D and any
requisite observation must be contained in a cluster **Proof:** By proper construction, any variable observed
before D is chosen must be weakly inward of D and any
requisite observation must be contained in a cluster
with D On the other hand if A is not observed befor before D is chosen must be weakly inward of D and any
requisite observation must be contained in a cluster
with D . On the other hand, if A is not observed before
 D is chosen and strictly inward of D then it mu requisite observation must be contained in a cluster
with D. On the other hand, if A is not observed before
D is chosen and strictly inward of D then it must not
be contained in that cluster \Box D is chosen and strictly inward of D then it must not
be contained in that cluster. \Box

Next, when an uncertainty becomes observable before decision D is chosen, it will not become requisite unless it is weakly outward to D . Next, when an uncertainty becomes observable before
decision D is chosen, it will not become requisite unless
it is weakly outward to D .
Theorem 2 (Newly Requisite Observations)
Given a rooted cluster tree properly

Theorem 2 (Newly Requisite Observations)
Given a rooted cluster tree properly constructed for an
influence diagram if uncertainty A is not weakly out-**Theorem 2 (Newly Requisite Observations)**
Given a rooted cluster tree properly constructed for an
influence diagram, if uncertainty A is not weakly out-
ward of decision D nor in any clusters with D then if Given a rooted cluster tree properly constructed for an influence diagram, if uncertainty A is not weakly outward of decision D nor in any clusters with D then if A were to be observed before D were chosen it wo ward of decision D nor in any clusters with D then if A were to be observed before D were chosen it would

not be requisite for ^D.

Proof: By proper construction and Lemma 1, the util-**Proof:** By proper construction and Lemma 1, the utility from D is weakly outward from D and all variables in common between the inwardmost cluster containing **Proof:** By proper construction and Lemma 1, the utility from D is weakly outward from D and all variables in common between the inwardmost cluster containing D and the next inward cluster are observed when D is in common between the inward
most cluster containing D and the next inward cluster are observed when
 D is chosen. Therefore, the utility is separated in the clus-D and the next inward cluster are observed when D is
chosen. Therefore, the utility is separated in the clus-
ter tree from A by observations for D, and the utility
from D is conditionally independent of A given the obchosen. Therefore, the utility is separated in the cluster tree from A by observations for D , and the utility from D is conditionally independent of A given the observations for D (Jensen and others 1990a: Lauri ter tree from A by observations for D, and the utility
from D is conditionally independent of A given the ob-
servations for D (Jensen and others 1990a; Lauritzen
and others 1990) from *D* is conditionally independent of *A* given the observations for *D* (Jensen and others 1990a; Lauritzen and others 1990). \Box

Finally, when an uncertainty stops being observable Finally, when an uncertainty stops being observable
before decision D is chosen, all of the observations now
requisite for D are weakly inward Finally, when an uncertainty stops
before decision D is chosen, all of the
requisite for D are weakly inward.

requisite for *D* are weakly inward.
**Proposition 1 (Previously Requisite Observa-
tions)** Given a rooted cluster tree properly constructed **Proposition 1 (Previously Requisite Observations)** Given a rooted cluster tree properly constructed for an influence diagram where uncertainty A is ob-**Proposition 1 (Previously Requisite Observations)** Given a rooted cluster tree properly constructed for an influence diagram where uncertainty A is observed before decision D is chosen, then if A were **tions)** Given a rooted cluster tree properly constructed for an influence diagram where uncertainty A is observed before decision D is chosen, then if A were not to be observed before D were chosen, all variables served before decision D is chosen, then if A were
not to be observed before D were chosen, all variables
which would be requisite observations for D are weakly
inward of D not to be obser
which would be
inward of D .

inward of D.
Proof: When properly constructed, all variables observed before decision D is chosen (not just the requi-**Proof:** When properly constructed, all variables observed before decision D is chosen (not just the requisite ones) are weakly inward of decision $D \Box$ **Proof:** When properly constructed, all variables of served before decision D is chosen (not just the requisite ones) are weakly inward of decision D . \Box site ones) are weakly inward of decision $D. \Box$
5 Computing the Value of

Information

Information
The new results from Section 4 can now be applied
to perform value of information calculations on the The new results from Section 4 can now be applied
to perform value of information calculations on the
rooted cluster tree for the original influence diagram. The new results from Section 4 can now be applied
to perform value of information calculations on the
rooted cluster tree for the original influence diagram.
First a method is presented for computing the value to perform value of information calculations on the rooted cluster tree for the original influence diagram.
First a method is presented for computing the value of a decision problem when an uncertainty is already First a method is presented for computing the value observed. This is then generalized to computing the of a decision problem when an uncertainty is already
observed. This is then generalized to computing the
value when there is an earlier observation, and then
when there is a later observation observed. This is then generalized
value when there is an earlier observation.
when there is a later observation.

when there is an earlier observation, and then
when there is a later observation.
Suppose that an uncertainty has already been ob-
served such as a in the influence diagram shown in Suppose that an uncertainty has already been observed, such as a in the influence diagram shown in Figure 3. By Theorem 2 it can be requisite only for Suppose that an uncertainty has already been observed, such as a in the influence diagram shown in Figure 3. By Theorem 2 it can be requisite only for decisions weakly inward in the tree shown in Figure 13. served, such as a in the influence diagram shown in Figure 3. By Theorem 2 it can be requisite only for decisions weakly inward in the tree shown in Figure 13. By exploiting the probabilistic heritage of the decision Figure 3. By Theorem 2 it can be requisite only for
decisions weakly inward in the tree shown in Figure 13.
By exploiting the probabilistic heritage of the decision
algorithm (Jensen and others 1990b: Lauritzen and decisions weakly inward in the tree shown in Figure 13.
By exploiting the probabilistic heritage of the decision
algorithm (Jensen and others 1990b; Lauritzen and
Spiegelhalter 1988), the rooted cluster tree is ideally By exploiting the probabilistic heritage of the decision
algorithm (Jensen and others 1990b; Lauritzen and
Spiegelhalter 1988), the rooted cluster tree is ideally
suited to solve this problem. First, the evidence is algorithm (Jensen and others 1990b; Lauritzen and
Spiegelhalter 1988), the rooted cluster tree is ideally
suited to solve this problem. First, the evidence is
stored in the probability potential in a cluster contain-Spiegelhalter 1988), the rooted cluster tree is ideally suited to solve this problem. First, the evidence is stored in the probability potential in a cluster containing a , say the inwardmost cluster containing a . Now suited to solve this problem. First, the evidence is stored in the probability potential in a cluster containing a . Now Algorithm 1 could be run incorporating this evidence stored in the probability potential in a cluster containing a . Now Algorithm 1 could be run incorporating this evidence.

Algorithm 1 could be run incorporating this evidence.
Suppose, however, that Algorithm 1 had already been
run before a was observed. No problem-only the clusrugorium 1 could be full incorporating this evidence.
Suppose, however, that Algorithm 1 had already been
run before a was observed. No problem–only the clus-
ters between the inwardmost cluster containing a and Suppose, however, that Algorithm 1 had already been
run before a was observed. No problem-only the clus-
ters between the inwardmost cluster containing a and
the root need to be visited. All of the other calcuters between the inwardmost cluster containing a and the root need to be visited. All of the other calcuters between the inwardmost cluster containing a and
the root need to be visited. All of the other calcu-
lations are unchanged! We could perform this same
operation even if a were not observed precisely prothe root need to be visited. All of the other calculations are unchanged! We could perform this same operation even if a were not observed precisely, pro-
vided we had some imperfect observation about a replations are unchanged! We could perform this same
operation even if a were not observed precisely, pro-
vided we had some imperfect observation about a rep-
resented by a likelihood function operation even if a were not observed precisely, provided we had some imperfect observation about a represented by a likelihood function.

Now consider the case in which an uncertainty will be

Figure 14: Effective rooted cluster trees for value of
information calculations on the influence diagram from
Figure 1 when B is observed before decisions are made Figure 14: Effective rooted cluster trees for value of information calculations on the influence diagram from Figure 1 when B is observed before decisions are made.
Those clusters changed from the rooted cluster tree in information calculations on the influence diagram from
Figure 1 when B is observed before decisions are made.
Those clusters changed from the rooted cluster tree in
Figure 12 are shaded Figure 1 when B is observed before decisions are made.
Those clusters changed from the rooted cluster tree in Figure 12 are shaded.

% observed earlier, but has not yet been observed, such as B in Figure 2. Again it is possible to exploit the % observed earlier, but has not yet been observed, such as B in Figure 2. Again it is possible to exploit the well-known properties of cluster trees. To compute the observed earlier, but has not yet been observed, such as B in Figure 2. Again it is possible to exploit the well-known properties of cluster trees. To compute the value of the decision problem in which B will be obas B in Figure 2. Again it is possible to exploit the well-known properties of cluster trees. To compute the value of the decision problem in which B will be observed earlier, cycle through all of the possible values well-known properties of cluster trees. To compute the value of the decision problem in which B will be observed earlier, cycle through all of the possible values of B performing the calculations each time as though value of the decision problem in which B will be observed earlier, cycle through all of the possible values of B , performing the calculations each time as though B were observed. The potentials computed can then served earlier, cycle through all of the possible values
of B , performing the calculations each time as though
 B were observed. The potentials computed can then
be summed, thereby incorporating the probability disof B , performing the calculations each time as though B were observed. The potentials computed can then be summed, thereby incorporating the probability distribution over the different possible values of B . If this B were observed. The potentials computed can then summing occurs immediately after the optimal policy tribution over the different possible values of B . If this summing occurs immediately after the optimal policy for D_i is computed, then this is the value of observing B before D_i is chosen summing occurs immedifor D_i is computed, then B before D_i is chosen.

 B before D_i is chosen.

One can think of this as "effectively" adding B to the

clusters inward to the inwardmost cluster containing O before D_i is enosen.
One can think of this as "effectively" adding B to the
clusters inward to the inwardmost cluster containing
 D_i as shown in Figure 14. In Figure 14a, this is used One can think of this as "effectively" adding B to the clusters inward to the inwardmost cluster containing D_i , as shown in Figure 14. In Figure 14a, this is used to compute the value of observing B before D_i and clusters inward to the inwardmost cluster containing D_i , as shown in Figure 14. In Figure 14a, this is used
to compute the value of observing B before D_1 , and
in Figure 14b before D_2 . Figure 14c is not different D_i , as shown in Figure 14. In Figure 14a, this is used
to compute the value of observing B before D_1 , and
in Figure 14b before D_2 . Figure 14c is not different
from Figure 12 because B would not be requisite if to compute the value of observing B before D_1 , and
in Figure 14b before D_2 . Figure 14c is not different
from Figure 12 because B would not be requisite if
it were observed before D_2 . This can be recognized in Figure 14b before D_2 . Figure 14c is not different
from Figure 12 because B would not be requisite if
it were observed before D_3 . This can be recognized
immediately from the rooted cluster tree in Figure 12 from Figure 12 because B would not be requisite if
it were observed before D_3 . This can be recognized
immediately from the rooted cluster tree in Figure 12 in which B is inward of D_3 . Note that unlike Figure 4, immediately from the rooted cluster tree in Figure 12
in which B is inward of D_3 . Note that unlike Figure 4,
in which the tree has been "expanded," this approach
sums over cases doing the same work but there is no in which B is inward of D_3 . Note that unlike Figure 4,
in which the tree has been "expanded," this approach
sums over cases, doing the same work, but there is no
need to store the larger tables and it uses the origin in which the tree has been "expanded," this approach
sums over cases, doing the same work, but there is no
need to store the larger tables, and it uses the original
rooted cluster tree! sums over cases, doing the same work, but there is no
need to store the larger tables, and it uses the original
rooted cluster tree!

Now consider the influence diagram shown in Figure 3.
Observing a earlier yields the effective rooted cluster
tree shown in Figure 15a Uncertainty a would be Now consider the influence diagram shown in Figure 3.
Observing a earlier yields the effective rooted cluster
tree shown in Figure 15a. Uncertainty a would be
requisite for D_1 but not for any of the later decisions Observing *a* earlier yields the effective rooted cluster
tree shown in Figure 15a. Uncertainty *a* would be
requisite for D_1 but not for any of the later decisions.
Suppose instead that *i* were observed earlier. It c tree shown in Figure 15a. Uncertainty *a* would be requisite for D_1 but not for any of the later decisions.
Suppose instead that *j* were observed earlier. It cannot be observed before D_2 since it is a descendant of requisite for D_1 but not for any of the later decisions.
Suppose instead that j were observed earlier. It cannot
be observed before D_1 since it is a descendant of D_1 . Suppose instead that j were observed earlier. It cannot
be observed before D_1 since it is a descendant of D_1 .
It is not requisite for D_2 or D_4 since it is not inward
of either but it would be requisite for $D_$ be observed before D_1 since it is a descendant of D_1 .
It is not requisite for D_2 or D_4 since it is not inward
of either, but it would be requisite for D_3 as shown in
Figure 15b of either, but it would be requisite for D_3 as shown in Figure 15b.

Figure 15b.
Now suppose that a variable is observed later rather
than earlier Consider C in Figure 1 and suppose that Now suppose that a variable is observed later rather
than earlier. Consider C in Figure 1 and suppose that
it is no longer observed before D_2 is chosen. From Now suppose that a variable is observed later rather
than earlier. Consider C in Figure 1 and suppose that
it is no longer observed before D_2 is chosen. From T
Proposition 1 the observations now requisite for D_2 ti than earlier. Consider C in Figure 1 and suppose that
it is no longer observed before D_2 is chosen. From
Proposition 1, the observations now requisite for D_2
are inward so the solution is to run Algorithm 2 to it is no longer observed before D_2 is chosen. From
Proposition 1, the observations now requisite for D_2
are inward, so the solution is to run Algorithm 2 to

Figure 15: Some effective rooted cluster trees for value
of information calculations on the influence diagram
from Figure 3 Figure 15: Some et
of information cal
from Figure 3.

Figure 16: Effective rooted cluster trees for value of information calculations on the influence diagram from Figure 1 when the observation of either C or E is delayed.

For D_2 instead of summing. Similarly, if A were not consider tree!
Now consider the influence diagram shown in Figure 3. as in Figure 16b. Finally, if E were not observed for figure how inward D_2 must effectively move up as in figure how inward D_2 must effectively move up as in
Figure 16a. Only now maximize over different cases
for D_2 instead of summing. Similarly, if A were not figure how inward D_2 must effectively move up as in
Figure 16a. Only now maximize over different cases
for D_2 instead of summing. Similarly, if A were not
observed for D_2 . D_2 can be effectively moved inward Figure 16a. Only now maximize over different cases
for D_2 instead of summing. Similarly, if A were not
observed for D_3 , D_3 can be effectively moved inward
as in Figure 16b. Finally, if E were not observed for for D_2 instead of summing. Similarly, if A were not observed for D_3 , D_3 can be effectively moved inward as in Figure 16b. Finally, if E were not observed for D_2 there is no change since E is not requisite for observed for D_3 , D_3 can be effectively moved inward
as in Figure 16b. Finally, if E were not observed for
 D_3 there is no change, since E is not requisite for D_3 .
Finally, a similar process can be done for the

 D_3 there is no change, since E is not requisite for D_3 .
Finally, a similar process can be done for the diagram
in Figure 3. Figure 17a shows the effective rooted clus-Finally, a similar process can be done for the diagram
in Figure 3. Figure 17a shows the effective rooted clus-
ter tree when f is no longer observed before D_3 and
Figure 17b shows the effective tree when e is no longe in Figure 3. Figure 17a shows the effective rooted cluster tree when f is no longer observed before D_3 and Figure 17b shows the effective tree when e is no longer observed before D_3 Figure 17b shows the effective tree when e is no longer observed before D_2 .

6 Conclusions and Future Research

6 Conclusions and Future Research
This paper has developed improved value of informa-
tion calculations over previous work in two respects This paper has developed improved value of information calculations over previous work in two respects.
First, it improves the rooted cluster trees used to solve tion calculations over previous work in two respects.
First, it improves the rooted cluster trees used to solve

Figure 17: Some effective rooted cluster trees for value
of information calculations on the influence diagram 50
from Figure 3 Figure 17: Some et
of information cal
from Figure 3.

for the value of a decision problem. Second, it develfor the value of a decision problem. Second, it develops methods for reusing the original tree in order to perform multiple value of information calculations for the value of a decision problem. Second, it devel
ops methods for reusing the original tree in order to
perform multiple value of information calculations. perform multiple value of information calculations.
There are several opportunities for further research.

When a particular variable is observed at multiple ear-There are several opportunities for further research.
When a particular variable is observed at multiple earlier decisions it should be possible to reuse some of the
calculations Also this approach exploits the special When a particular variable is observed at multiple ear-
lier decisions it should be possible to reuse some of the
calculations. Also, this approach exploits the special
properties of changing the time when a single uncerlier decisions it should be possible to reuse some of the calculations. Also, this approach exploits the special properties of changing the time when a single uncer-
tainty becomes observed. It would be useful if the calculations. Also, this approach exploits the special
properties of changing the time when a single uncer-
tainty becomes observed. It would be useful if the
method could be generalized to solve the decision probproperties of changing the time when a single uncertainty becomes observed. It would be useful if the method could be generalized to solve the decision prob-
lem with any set of informational assumptions from tainty becomes observed. It would be useful if the
method could be generalized to solve the decision prob-
lem with any set of informational assumptions from
the original rooted cluster tree. If that could be done method could be generalized to solve the decision prob-
lem with any set of informational assumptions from
the original rooted cluster tree. If that could be done
efficiently then the original decision problem could be lem with any set of informational assumptions from
the original rooted cluster tree. If that could be done
efficiently, then the original decision problem could be
solved from the most convenient cluster tree efficiently, then the original decision problem could be solved from the most convenient cluster tree.

Acknowledgments

Acknowledgments
This paper has benefited from the comments, sugges-
tions, and ideas of friends and students, most notably This paper has benefited from the comments, suggestions, and ideas of friends and students, most notably Mark Peot and Prakash Shenov This paper has benefited from the
tions, and ideas of friends and stu
Mark Peot and Prakash Shenoy.

References

References
Dittmer, S. L. and F. Jensen. "Myopic Value of Infor-
mation in Influence Diagrams" In Uncertainty in Dittmer, S. L. and F. Jensen. "Myopic Value of Information in Influence Diagrams." In Uncertainty in Artificial Intelligence: Proceedings of the Thir-Dittmer, S. L. and F. Jensen. "Myopic Value of Information in Influence Diagrams." In **Uncertainty in**
Artificial Intelligence: Proceedings of the Thir-
teenth Conference eds D Geiger and P P Shenov mation in Influence Diagrams." In Uncertainty in Artificial Intelligence: Proceedings of the Thirteenth Conference, eds. D Geiger and P P Shenoy. 142-149. San Francisco. CA: Morgan Kaufmann. 1997. Artificial Intelligence: Proceedings of the Thir-

Francisco, CA: Morgan Kaufmann, 1997.
142-149. San Francisco, CA: Morgan Kaufmann, 1997.
Howard, R. A. and J. E. Matheson. "Influence Dia-F42-145. San Francisco, OA. Morgan Raumiann, 1557.
Howard, R. A. and J. E. Matheson. "Influence Diagrams." In The Principles and Applications of
Decision Analysis eds. R. A. Howard and J. E. Howard, R. A. and J. E. Matheson. "Influence Diagrams." In **The Principles and Applications of Decision Analysis**, eds. R. A. Howard and J. E. Matheson. II. Menlo Park. CA: Strategic Decisions grams." In **The Principles and Applications of Decision Analysis**, eds. R. A. Howard and J. E. Matheson. II. Menlo Park, CA: Strategic Decisions Group 1984 Decision Ana
Matheson. II.
Group, 1984. Matheson. II. Menlo Park, CA: Strategic Decisions
Group, 1984.
Jensen, F., F. V. Jensen, and S. L. Dittmer. "From

Influence Diagrams to Junction Trees." In Uncer-
tainty in Artificial Intelligence: Proceedings of Influence Diagrams to Junction Trees." In Uncertainty in Artificial Intelligence: Proceedings of
the Tenth Conference eds. B Lopez de Mantaras Influence Diagrams to Junction Trees." In Uncertainty in Artificial Intelligence: Proceedings of
the Tenth Conference, eds. R Lopez de Mantaras
and D Poole 367-373 San Mateo CA: Morgan Kauftainty in Artificial Intelligence: Proceedings of
the Tenth Conference, eds. R Lopez de Mantaras
and D Poole. 367-373. San Mateo, CA: Morgan Kauf-
mann 1994 the Tenth C
and D Poole.
mann, 1994. mann, 1994.
Jensen, F. V., S. L. Lauritzen, and K. G. Olesen.

"Bayesian Updating in Causal Probabilistic Networks Jensen, F. V., S. L. Lauritzen, and K. G. Olesen.
"Bayesian Updating in Causal Probabilistic Networks
by Local Computations." **Comp. Stats. Q.** 4
(1990a): 269-282 "Bayesian Updating
by Local Computa
(1990a): 269-282. (1990a): 269-282.
Jensen, F. V., K. G. Olesen, and S. K. Andersen.

(1990a): 209-202.
Jensen, F. V., K. G. Olesen, and S. K. Andersen.
"An algebra of Bayesian belief universes for knowledge
based systems" Networks 20 (1990b): 637-659 Jensen, F. V., K. G. Olesen, and S. K. Anders
"An algebra of Bayesian belief universes for knowled
based systems." **Networks** 20 (1990b): 637-659. based systems." **Networks** 20 (1990b): 637-659.
Lauritzen, S. L., A. P. Dawid, B. N. Larsen, and H.-G.

Based systems. Treeworks 20 (19909). 691-699.
Lauritzen, S. L., A. P. Dawid, B. N. Larsen, and H.-G.
Leimer. "Independence properties of directed Markov
fields." Networks 20 (1990): 491-505 Lauritzen, S. L., A. P. Dawid, B. N. Larse
Leimer. "Independence properties of dire
fields." **Networks** 20 (1990): 491-505. fields." **Networks** 20 (1990): 491-505.
Lauritzen, S. L. and D. J. Spiegelhalter. "Local com-

putations with probabilities on graphical structures
putations with probabilities on graphical structures
and their application to expert systems." **JBSS B** Lauritzen, S. L. and D. J. Spiegelhalter. "Local computations with probabilities on graphical structures
and their application to expert systems." **JRSS B**
50 (2 1988): 157-224 putations with probab
and their application t
50 (2 1988): 157-224.

So (2 1988): 157-224.
50 (2 1988): 157-224.
Ndilikilikesha, P. **Potential Influence Diagrams**.
University of Kansas, School of Business, 1991. Work-
ing Paper 235. Ndilikilikesha, F
University of Ka
ing Paper 235.

Enversity of Raifsas, School of Dusiness, 1991. Working Paper 235.
Raiffa, H. **Decision Analysis**. Reading, MA:
Addison-Wesley 1968 mg 1 aper 255.
Raiffa, H. **Decision**
Addison-Wesley, 1968. Raiffa, H. Decision Analysis. Reading, MA:
Addison-Wesley, 1968.
Shachter, R. D. "Evaluating Influence Diagrams."

Ops. Rsrch. 34 (November-December 1986): 871- 882.

Space Ratch, St (November-December 1960).
882.
Shachter, R. D. "Bayes-Ball: The Rational Pastime
(for Determining Irrelevance and Requisite Informa-Soz.
Shachter, R. D. "Bayes-Ball: The Rational Pastime
(for Determining Irrelevance and Requisite Informa-
tion in Belief Networks and Influence Diagrams)." In Shachter, R. D. "Bayes-Ball: The Rational Pastime
(for Determining Irrelevance and Requisite Informa-
tion in Belief Networks and Influence Diagrams)." In
Hatter Lincertainty in Artificial Intelligence: Proceed-(for Determining Irrelevance and Requisite Information in Belief Networks and Influence Diagrams)." In Uncertainty in Artificial Intelligence: Proceed-
ings of the Fourteenth Conference, 480-487. San tion in Belief Networks and Influence Diagrams)." In
Uncertainty in Artificial Intelligence: Proceed-
ings of the Fourteenth Conference, 480-487. San
Francisco, CA: Morgan Kaufmann, 1998 Uncertainty in Artificial Intelligence:
ings of the Fourteenth Conference, 48
Francisco, CA: Morgan Kaufmann, 1998. Francisco, CA: Morgan Kaufmann, 1998.
Shachter, R. D. and P. M. Ndilikilikesha. "Using Po-

tential Influence Diagrams for Probabilistic Inference Shachter, R. D. and P. M. Ndilikilikesha. "Using Potential Influence Diagrams for Probabilistic Inference
and Decision Making." In Uncertainty in Artificial
Intelligence: Proceedings of the Ninth Confertential Influence Diagrams for Probabilistic Inference
and Decision Making." In **Uncertainty in Artificial**
Intelligence: Proceedings of the Ninth Confer-
ence 383-390 San Mateo CA: Morgan Kaufmann and Decision Making." In Uncertainty in Artificial
Intelligence: Proceedings of the Ninth Confer-
ence, 383-390. San Mateo, CA: Morgan Kaufmann,
1993 1993.

Shachter, R. D. and M. A. Peot. "Decision Making
Shachter, R. D. and M. A. Peot. "Decision Making
Using Probabilistic Inference Methods" In Uncer-1999.
Shachter, R. D. and M. A. Peot. "Decision Making
Using Probabilistic Inference Methods." In Uncer-
tainty in Artificial Intelligence: Proceedings of Using Probabilistic Inference Methods." In Uncertainty in Artificial Intelligence: Proceedings of the Eighth Conference, 276-283. San Mateo, CA: Using Probabilistic Inference Methods." In Uncertainty in Artificial Intelligence: Proceedings of the Eighth Conference, 276-283. San Mateo, CA: Morgan Kaufmann 1992 tainty in Artificial Intell
the Eighth Conference,
Morgan Kaufmann, 1992.

Shenoy, P. P. "Valuation-Based Systems for Bayesian
Shenoy, P. P. "Valuation-Based Systems for Bayesian
Decision Analysis" Ops. Bsrch. 40 (3 1992): 463-Shenoy, P. P. "Valuation-Based Systems for Bayesian
Decision Analysis." **Ops. Rsrch.** 40 (3 1992): 463-
484 484.

Eccision Analysis. Ups. 1816.. 40 (5 1332). 463-484.
484.
Tatman, J. A. and R. D. Shachter. "Dynamic Pro-Tatman, J. A. and R. D. Shachter. "Dynamic Programming and Influence Diagrams." **IEEE SMC** 20 (2.1990): 365-379 gramming and Influence Diagrams." IEEE SMC 20 $(2\ 1990)$: 365-379.