Evidence Absorption — Experiments on Different Classes of Randomly Generated Belief Networks

L.C. van der Gaag

UU-CS-1994-42
October 1994

Utrecht University

Department of Computer Science

Padualaan 14, P.O. Box 80.089,
3508 TB Utrecht, The Netherlands,
Tel. : ... + 31 - 30 - 531454

Evidence Absorption —
Experiments on Different Classes of Randomly Generated Belief Networks

L.C. van der Gaag

Technical Report UU-CS-1994-42
October 1994

Department of Computer Science
Utrecht University
P.O.Box 80.089
3508 TB Utrecht
The Netherlands

ISSN: 0924-3275

Evidence Absorption — Experiments on Different Classes of Randomly Generated Belief Networks

Linda C. van der Gaag
Utrecht University, Department of Computer Science
P.O. Box 80.089, 3508 TB Utrecht, The Netherlands

Abstract

More and more real-life applications of the belief network framework begin to emerge. As applications grow larger, the networks involved increase in size accordingly. For large belief networks, the computations involved in probabilistic inference tend to become rather time consuming, even so to an unacceptable extent. To address this problem, we have proposed in a previous paper to incorporate the method of *evidence absorption* into Pearl's algorithms for probabilistic inference. In the present paper, the ability of this method to improve on the average-case computational expense of probabilistic inference is illustrated by means of experiments performed on different classes of randomly generated belief networks. Both the set-up of the experiments and the results obtained are detailed. The results from our experiments are shown to reflect to a large extent the use of belief networks incorporating a randomly generated digraph. We comment on this observation by addressing the suitability of using of randomly generated belief networks in this type of experiment.

1 Introduction

More and more knowledge-based systems employing the belief network framework are being developed for various domains of application, most notably for medical diagnosis and prognosis [Andreassen *et al.*, 1987; Heckerman *et al.*, 1992]. The success of the framework can be attributed to its powerful formalism for representing probabilistic knowledge and to its general algorithms for mathematically sound probabilistic inference. As real-life applications of the framework begin to emerge, however, the tendency of the basic algorithms involved to slow down problem solving becomes apparent. Since probabilistic inference is NP-hard [Cooper, 1990], this tendency cannot be denied in general; the algorithms associated with the framework have an exponential worst-case time complexity. Current research therefore aims at improving on the average-case computational performance of these algorithms.

In a previous paper, we have proposed integrating the method of *evidence absorption* into Pearl's algorithms for exact probabilistic inference to save on the computational effort spent on average-case problem solving [van der Gaag, 1993]. The basic idea of this method is to dynamically modify a belief network as evidence becomes available to reflect newly created independencies. Since Pearl's algorithms for probabilistic inference exploit the independencies portrayed by a belief network directly, the incorporation of evidence absorption into these algorithms is expected to speed up problem solving. To gain more insight into this ability of the method of evidence absorption, we have conducted several experiments on different classes of randomly generated belief networks. In this paper we present the results obtained from these experiments.

The paper is organised as follows. Section 2 briefly reviews the belief network framework, and describes the method of evidence absorption and its integration into Pearl's algorithms for probabilistic inference. In Section 3, we outline the set-up of our experiments and present the results obtained. The results from our experiments are discussed in detail in Section 4. In Section 5 we comment on the suitability of using randomly generated belief networks for experiments with the method of evidence absorption. The paper is rounded off with some conclusions in Section 6.

2 Preliminaries

In this section we briefly review the basic notions involved in the belief network framework. We further describe the method of evidence absorption and its integration into Pearl's algorithms for probabilistic inference. These preliminaries serve only as a sketch of the background of our experiments and are not meant to be tutorial; a more elaborate introduction to the belief network framework and to Pearl's algorithms can be found in [Pearl, 1988].

2.1 The Belief Network Framework

The belief network framework comprises a formalism for representing joint probability distributions and a set of algorithms for exact probabilistic inference.

The belief network formalism provides for both a *qualitative representation* and a *quantitative representation* of the joint probability distribution on a set of statistical variables discerned in a domain. The qualitative part of a belief network takes the form of an *acyclic digraph*. Each vertex in this digraph represents a statistical variable that can take one of a set of values. The arcs of the digraph represent interdependencies between these variables. Informally speaking, we take an arc $V \rightarrow W$ in the digraph to represent a direct 'influential' relationship between the variables V and W; the direction of the arc designates W as the 'effect' of V. Absence of an arc between two vertices means that the corresponding variables do not influence each other directly and, hence, are (conditionally) independent. In building a belief network for a domain of application, its graphical part is constructed to reflect as many of the independencies between the variables discerned as possible. Associated with the digraph of a belief network is a numerical assessment of the 'strengths' of the represented relationships: with each vertex is associated a set of conditional probabilities describing the influence of the values of its immediate predecessors on the probabilities of its values. The independencies portrayed by the graphical part of the network and the associated probabilities together model the joint probability distribution for the problem domain at hand.

A belief network is used for making probabilistic statements concerning the variables discerned in the domain of application. To this end, two *algorithms for probabilistic inference* are associated with the formalism: an algorithm for (efficiently) computing probabilities of interest from a belief network, and an algorithm for processing evidence, that is, an algorithm for entering evidence into the network and subsequently (efficiently) computing the revised joint probability distribution given the evidence. Several such algorithms have been developed, the most well-known of which are the algorithms by J. Pearl [Pearl, 1988]. Since we will build on Pearl's algorithms for probabilistic inference, we will briefly outline their basic idea. In doing so, we take an object-oriented point of view. The digraph of a belief network is taken

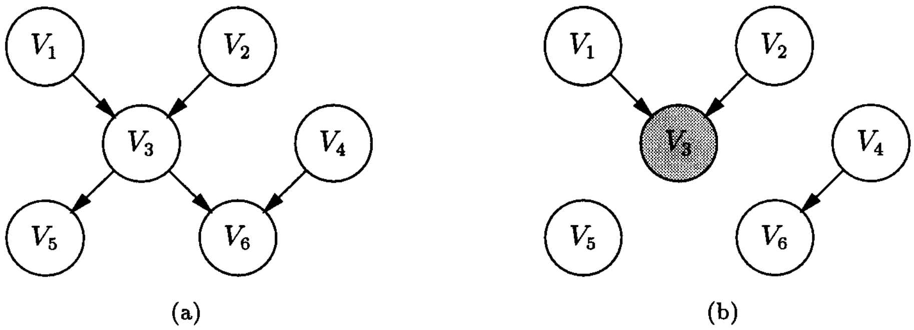

Figure 1: An Example of Evidence Absorption.

as a computational architecture: the vertices of the digraph are autonomous objects having a local processor capable of performing certain probabilistic computations and a local memory in which the associated probabilities are stored; the arcs of the digraph are bi-directional communication channels. Through these communication channels the vertices send each other *parameters* providing information about the represented joint probability distribution and the evidence obtained so far. Each vertex is able to compute the probabilities of its values from the information it receives from its neighbours and its own local probabilities. The details of the computations involved are not relevant for the present paper. It suffices to emphasize that the computational effort spent on probabilistic inference is largely determined by the sparsity of the digraph of the network; in fact, Pearl's algorithms exploit the independencies portrayed by the digraph of a network directly and perform the better from a computational point of view as the digraph is sparser.

2.2 Evidence Absorption and Pearl's Algorithms

The method of *evidence absorption* is applied to a belief network after a piece of evidence has been entered. The method seeks to incorporate the independencies newly created by the evidence explicitly into the network at hand. To this end, the topology of the digraph of the network is modified and the conditional probabilities assessed for its vertices are adjusted. The modification of the graphical part of the network amounts to deleting from the digraph all arcs departing from the vertex for which evidence has been entered. We illustrate this modification by means of an example. Consider a belief network comprising the singly connected digraph shown in Figure 1(a) and suppose that evidence is entered for the variable V_3. The digraph is modified by the method of evidence absorption by deleting all arcs departing from vertex V_3. The modified digraph is shown in Figure 1(b); the evidence is represented by drawing vertex V_3 with shading. As will become clear in the sequel, the modification of the probabilities involved is of no relevance to the present paper. For further details on the method of evidence absorption, we refer to [van der Gaag, 1993].

Since the method of evidence absorption tends to delete arcs from the digraph of a belief network, it is worthwhile to integrate the method into Pearl's algorithms to cut down on the computational expense of further probabilistic inference: the more arcs are deleted from a digraph, the sparser the digraph will become, and, as mentioned before, the computational effort involved in probabilistic inference is largely determined by the sparsity of a network's digraph. The basic idea of integrating the method of evidence absorption into Pearl's algorithms is as follows. When a piece of evidence for a specific variable is entered into the network, the method of evidence absorption is applied. Then, Pearl's algorithms are called upon to actually propagate the evidence. The correctness of the thus extended algorithm derives from the observation that after propagation of the evidence the modified network and the original network model the same updated joint probability distribution and the same independency relation given the processed evidence.

3 The Experiments and Their Results

In the previous section, we have seen that the method of evidence absorption tends to delete arcs from the digraph of a belief network and never inserts any new arcs. The most interesting question to address now is what impact applying the method of evidence absorption has on the topology of the digraph of a belief network as successive evidence is entered, since this impact can be related directly to the computational expense involved in further probabilistic inference.

From a theoretical point of view, the best case and the worst case are easily identified. The *worst* case would be a digraph for which evidence is obtained only for vertices without any outgoing arcs. In this case, applying the method of evidence absorption is pointless: there are no arcs deleted from the digraph and further computations are just as expensive as when evidence absorption had not been applied. It is worth noting, however, that the method of evidence absorption would not weigh heavily on the computational effort spent on probabilistic inference: the only additional work required would be a simple check on a vertex' successor set. In the *best* case, the method of evidence absorption causes the digraph of the network to fall apart into components of size one only for a single piece of evidence: this would be a digraph having the shape of a tree of depth one for which evidence is entered for the root vertex.

The above observations are very general. To gain more insight into the impact of the method of evidence absorption in situations which were not predesigned, we have conducted several experiments on different classes of randomly generated belief networks. In this section, we discuss these experiments and their results. In Section 3.1, the software developed for use in the experiments is described. In Section 3.2, we outline the various experiments performed. Section 3.3 details the results obtained from these experiments. An interpretation and detailed discussion of the results are deferred to Section 4.

3.1 The Generators Used

As outlined before, the aim of our experiments is to gain insight into the impact of the method of evidence absorption on the average-case computational expense involved in probabilistic inference. Since this impact derives from the way the method modifies the graphical part of a belief network and not from the modification of the associated conditional probabilities, we have designed our experiments to apply to the graphical part of a network only. In each of the experiments, we have generated a set of acyclic digraphs. Each digraph is generated to comprise n vertices, $n \geq 1$, and m arcs, $n-1 \leq m \leq \frac{1}{2}n \cdot (n-1)$. In each digraph, k pieces of evidence are entered for which the method of evidence absorption is applied, $1 \leq k \leq n$. The integers n, m and k differ in the various experiments and will be further detailed in Section 3.2.

3.1.1 The Graph Generator

For generating acyclic digraphs comprising a pre-specified number of vertices and arcs, a *graph generator* has been written in Common Lisp. At the basis of the generator lies an ordered list of vertices $V_1, V_2, \ldots$. For a digraph of n vertices, the first n elements are selected from this list. Subsequently, m arcs are created from the selected list of vertices, using the random number generator facilities of our Common Lisp system. For each arc, two vertices are selected at random from the list of vertices $V_1, \ldots, V_n$ of the digraph; the pair is taken to represent an arc directed from the lower-ordered vertex to the higher-ordered one. Note that this approach prevents the introduction of cycles into the digraph.

In our experiments, only *connected* digraphs have been considered. We have decided to introduce the bias of connectivity into our experiments because digraphs occurring in real-life belief networks tend to be connected. We would like to note that this bias is commonly used in this type of experiment [Suermondt & Cooper, 1990]. To incorporate the connectivity bias, the graph generator is furnished with a *connectivity test.* This test is employed in the following sense. For a digraph with m arcs, the specified number of arcs is generated in the manner described above. After having inserted m arcs into the digraph, the resulting graph is tested for connectivity. If the graph is not yet connected, then additional arcs are created and inserted, one at a time until the digraph has become connected. Subsequently, arcs are selected at random from the thus generated digraph and deleted until the desired number of arcs is arrived at; however, before eliminating a selected arc it is verified whether the digraph will remain connected.

A digraph generated by the graph generator described above is not truly random in the sense of the theory of random graphs [Bollobas, 1985]; randomness has been destroyed as a result of the incorporation of the bias of connectivity. In addition, we observe that of the class of connected acyclic digraphs of specific size not all graphs have equal probability of being generated; the way the connectivity bias has been implemented accounts for the loss of uniformity of generation. As these properties may seem unwished-for, we would like to note that the generation of random digraphs is a complex research issue in theoretical computer science. Rather than contribute to this line of research, our aim has been to create a practical generator that would be 'random enough' for our purposes. As will become clear in the sequel, the results obtained from our experiments indicate that our generator indeed achieves a high degree of randomness. In addition, it is noted that generators reported to have been used for similar experiments exhibit properties closely resembling the properties mentioned above; we refer for example to the graph generator used in the experiment with loop cutsets reported by H.J. Suermondt and G.F. Cooper [Suermondt & Cooper, 1990].

3.1.2 The Evidence Generator

To study the impact of *repeated* application of the method of evidence absorption, in each experiment we have entered k pieces of evidence into each of the digraphs generated. Recall that we are interested primarily in the impact of the method of evidence absorption on a digraph's topology and not so much in the precise probabilistic computations involved in propagating the evidence. We therefore have modelled entering a piece of evidence by selecting a vertex from the set of vertices of the digraph at hand and applying the modifying operation of the method of evidence absorption to the digraph's topology only.

Vertices modelling pieces of evidence are selected by an *evidence generator*. This generator selects vertices from the digraph at hand either randomly or with one of two different biases. These biases concern the location in the digraph of the vertices for which evidence is entered. One bias aims at selecting vertices located in the upper part of the digraph; in the sequel, this bias will be referred to as the *upper bias*. The other bias aims at selecting vertices located in the lower part of the digraph; this bias will be termed the *lower bias*. These biases have been introduced into the evidence generator because it is expected that the location in the digraph of the vertices for which evidence is entered plays a major role in the impact of the method of evidence absorption on a digraph's topology. Note that for diagnostic applications, the vertices for which evidence is entered tend to be located in the lower part of the digraph whereas for prognostic applications these vertices are more likely to be located in the upper part of the network.

The two biases mentioned above are implemented in the evidence generator as a two-stage selection process. The process of selecting vertices located in the upper part of the digraph starts with selecting a vertex V_i from the ordered list of vertices $V_1, \ldots, V_n$ of the digraph in a random fashion, $1 \leq i \leq n$; the vertex modelling the piece of evidence then is selected among the vertices $V_1, \ldots, V_i$. The property that this two-stage selection process implements the wished-for bias derives from the observation that the ordering of the vertices employed by the graph generator coincides with a topological ordering of the vertices of the generated digraph. The bias for selecting vertices located in the lower part of the digraph is modelled in a similar fashion.

3.2 The Experiments Performed

Before presenting the results, we outline the experiments performed and the objectives we had in designing each of these experiments.

Experiments with the method of evidence absorption are performed in a search space spanned by four parameters:

- the *number of vertices* of the digraphs investigated;
- the *number of arcs* of these digraphs;
- the *number of pieces of evidence* that are entered;
- the *location in the digraph* of these pieces of evidence.

In the first experiment, the aim is to develop a feel for the influence of the various parameters on the behaviour of the topology of a digraph under evidence absorption. To this end, a coarse-meshed net is laid over the search space and the combinations of parameter values corresponding with the meshes of this net are investigated. The number of vertices of the digraphs generated is fixed to fifty — due to the use of randomly generated digraphs, we feel that it is the ratio of the numbers of vertices and arcs that is of significance rather than the number of vertices in itself. The pieces of evidence entered into the digraphs are selected in a *random* fashion: the impact of the location in the digraph of the evidence is studied in a separate experiment. The remaining two parameters defining our search space, that is, the number of arcs of the digraphs and the number of pieces of evidence entered, are varied.

We investigate digraphs of fifty vertices comprising fourty-nine arcs (that is, *singly connected* digraphs), and digraphs of fifty vertices with seventy-five, one hundred, one hundred and fifty and two hundred and fifty arcs (that is, *multiply connected* digraphs), respectively. We have refrained from investigating more densely connected digraphs, as in real-life applications the digraphs of belief networks tend to be sparse [Pearl, 1988]. The numbers of pieces of evidence investigated are one, ten and twenty-five.

Before describing which information is collected from this experiment, we recall that the aim of incorporating the method of evidence absorption into Pearl's algorithms is to improve on the average-case computational expense involved in probabilistic inference. Since the computational effort spent on inference depends on the sparsity of the digraph of the belief network at hand, we determine from the experiment the minimum and maximum number of deleted arcs as well as the average number of arcs deleted over all digraphs investigated. In addition, we observe that if the digraph of a belief network has fallen apart into separate components, then further probabilistic inference can be restricted to single components. We therefore also determine from the experiment the average number of components after evidence absorption and the average sizes of the minimum and the maximum component. In addition, the cumulated count of the sizes of all components arising in the experiment is determined.

The aim of our second experiment is to study, in isolation, the influence of the number of arcs on the behaviour of a digraph's topology under evidence absorption. In this experiment, we once more consider digraphs comprising fifty vertices each. Again, the pieces of evidence entered into these digraphs are generated in a random fashion. We have now fixed the number of pieces of evidence entered to ten. The number of arcs of the digraphs is varied from fifty up to one hundred and fifty, increasing by two. From this experiment, we determine the average number of arcs deleted over all generated digraphs, the average number of components resulting after evidence absorption and the average sizes of the minimum and the maximum component.

The third experiment is similar to the second one in the sense that its aim also is to study the impact of one of the parameters defining the search space for experimentation in isolation: it is the number of pieces of evidence that is varied in this experiment. We once more generate digraphs comprising fifty vertices; the number of arcs of the digraphs generated has been fixed to one hundred. The pieces of evidence entered into these digraphs are generated randomly. The number of pieces of evidence entered is varied from one up to twenty-five. From this experiment, the same information is collected as from the second experiment.

The fourth experiment has been designed to investigate the influence of the location in the digraph of the vertices for which evidence is entered. To this end, we repeat the first experiment twice: once using the bias of selecting vertices located in the lower part of the digraph and once using the bias of selecting vertices located in the upper part of the digraph. We do not repeat the second and third experiment using these biases: we feel that repeating these experiments would only reveal the same relationships between the parameters of the search space involved. From the fourth experiment, we collect the same information as from the second and third experiment.

In each experiment, we have created several sets of one hundred digraphs, one set for each combination of parameter values investigated. We are aware that this number is not large enough to allow any statistically sound conclusions to be drawn. However, for each combination investigated, there is such a huge number of different digraphs that generating and testing a representative number of digraphs would simply not be feasible.

3.3 The Results of the Experiments

In this section, we present the results obtained from our experiments with the method of evidence absorption as outlined above.

3.3.1 The First Experiment

Before presenting the results of our first experiment, we recall that in this experiment we have generated digraphs comprising fifty vertices; the pieces of evidence entered into these digraphs have been generated *randomly.* The experiment is composed of several tests in which we have varied the number of arcs of the digraphs and the number of pieces of evidence entered.

Test 1
We have generated three sets of *singly connected* digraphs. For the first set, we have selected *one* piece of evidence for each digraph; for the second set, the number of pieces of evidence equals *ten*; for the third set, this number equals *twenty-five.* To each digraph, the method of evidence absorption has been applied for the selected pieces of evidence. For the modified digraphs, we have found the following statistics:

Number of pieces of evidence entered	1	10	25
Minimum number of arcs deleted	0	3	16
Maximum number of arcs deleted	4	18	33
Average number of deleted arcs	0.79	9.58	24.81
Average number of components	1.79	10.58	25.81
Average size of the minimum component	23.65	1	1
Average size of the maximum component	47.67	27.4	10.75

In addition to these statistics, Figure 2 presents the cumulated count of numbers of components per component size.

Test 2
We have generated three sets of *multiply connected* digraphs comprising *seventy-five* arcs. For the first set, we have selected *one* piece of evidence for each digraph; for the second set, the number of pieces of evidence equals *ten*; for the third set, this number equals *twenty-five.* To each digraph, the method of evidence absorption has been applied for the pieces of evidence selected. For the modified digraphs, we have found the following statistics:

Number of pieces of evidence entered	1	10	25
Minimum number of arcs deleted	0	6	20
Maximum number of arcs deleted	5	29	46
Average number of deleted arcs	1.35	14.89	36.51
Average number of components	1.44	5.95	16.36
Average size of the minimum component	31.43	1.52	1
Average size of the maximum component	49.5	44.5	30.18

In addition to these statistics, Figure 3 presents the cumulated count of numbers of components per size.

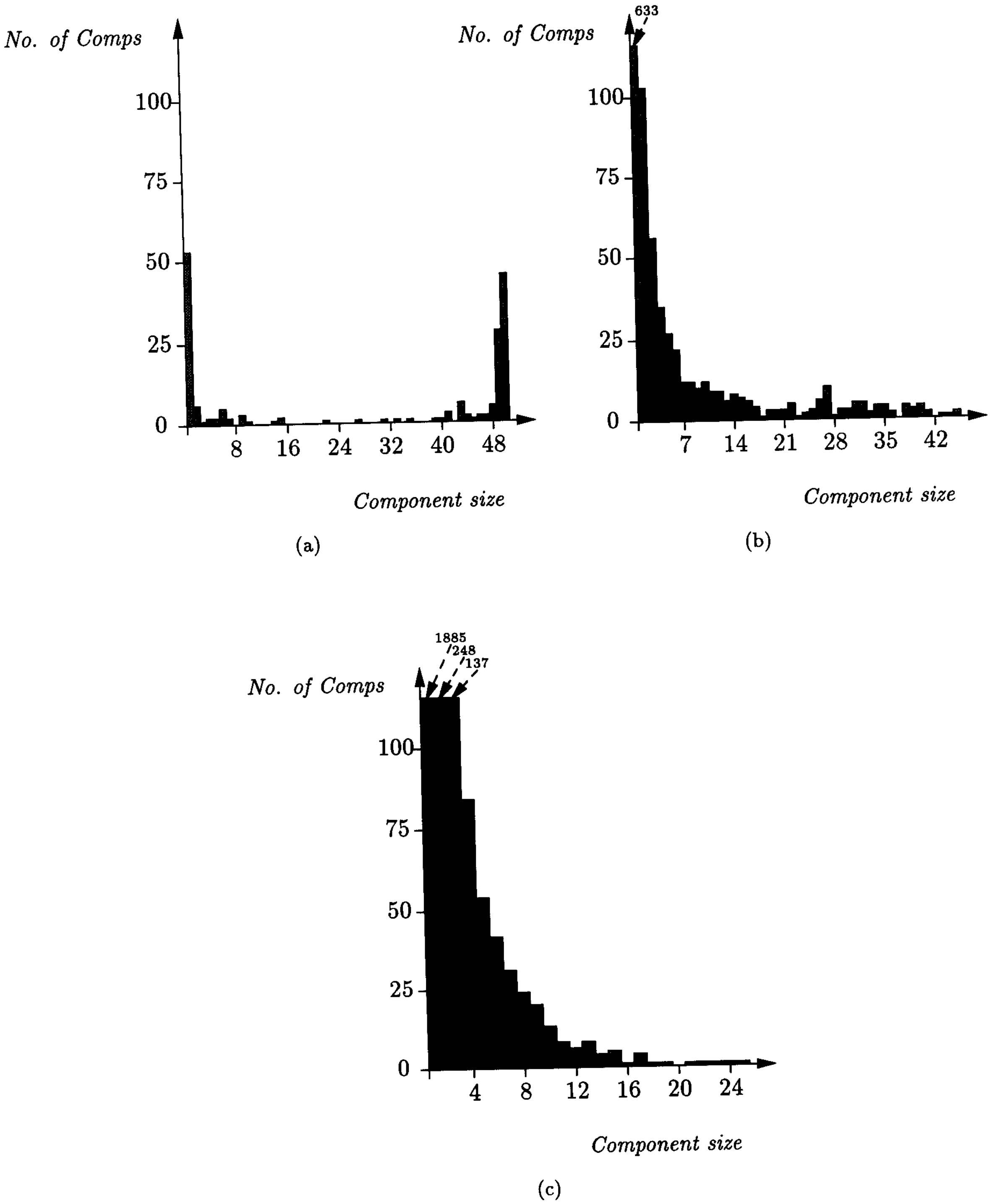

Figure 2: Cumulated Count for Digraphs with 50 Vertices and 49 Arcs, and 1 (a), 10 (b) and 25 (c) Pieces of Evidence Entered.

Test 3
We have generated three sets of *multiply connected* digraphs comprising *one hundred* arcs. For the first set, we have selected *one* piece of evidence for each digraph; for the second set, the number of pieces of evidence equals *ten*; for the third set, this number equals *twenty-five*. To each digraph, the method of evidence absorption has been applied for the selected pieces of evidence. For the modified digraphs, we have found the following statistics:

Number of pieces of evidence entered	1	10	25
Minimum number of arcs deleted	0	9	36
Maximum number of arcs deleted	7	31	67
Average number of deleted arcs	2.08	20.59	49.45
Average number of components	1.32	4.55	12.61
Average size of the minimum component	34.82	2.48	1
Average size of the maximum component	49.67	46.38	37.74

In addition to these statistics, Figure 4 presents the cumulated count of numbers of components per size.

Test 4
We have generated three sets of *multiply connected* digraphs comprising *one hundred and fifty* arcs. For the first set, we have selected *one* piece of evidence for each digraph; for the second set, the number of pieces of evidence equals *ten*; for the third set, this number equals *twenty-five*. To each digraph, the method of evidence absorption has been applied for the selected pieces of evidence. For the modified digraphs, we have found the following statistics:

Number of pieces of evidence entered	1	10	25
Minimum number of arcs deleted	0	7	56
Maximum number of arcs deleted	10	47	96
Average number of deleted arcs	2.95	29.44	74.82
Average number of components	1.09	2.92	9.18
Average size of the minimum component	45.59	6.39	1
Average size of the maximum component	49.91	48.06	41.74

In addition to these statistics, Figure 5 presents the cumulated count of numbers of components per size.

Test 5
We have generated three sets of *multiply connected* digraphs comprising *two hundred and fifty* arcs. For the first set, we have selected *one* piece of evidence for each digraph; for the second set, the number of pieces of evidence equals *ten*; for the third set, this number equals *twenty-five*. To each digraph, the method of evidence absorption has been applied for the pieces of evidence selected. For the modified digraphs, we have found the following statistics:

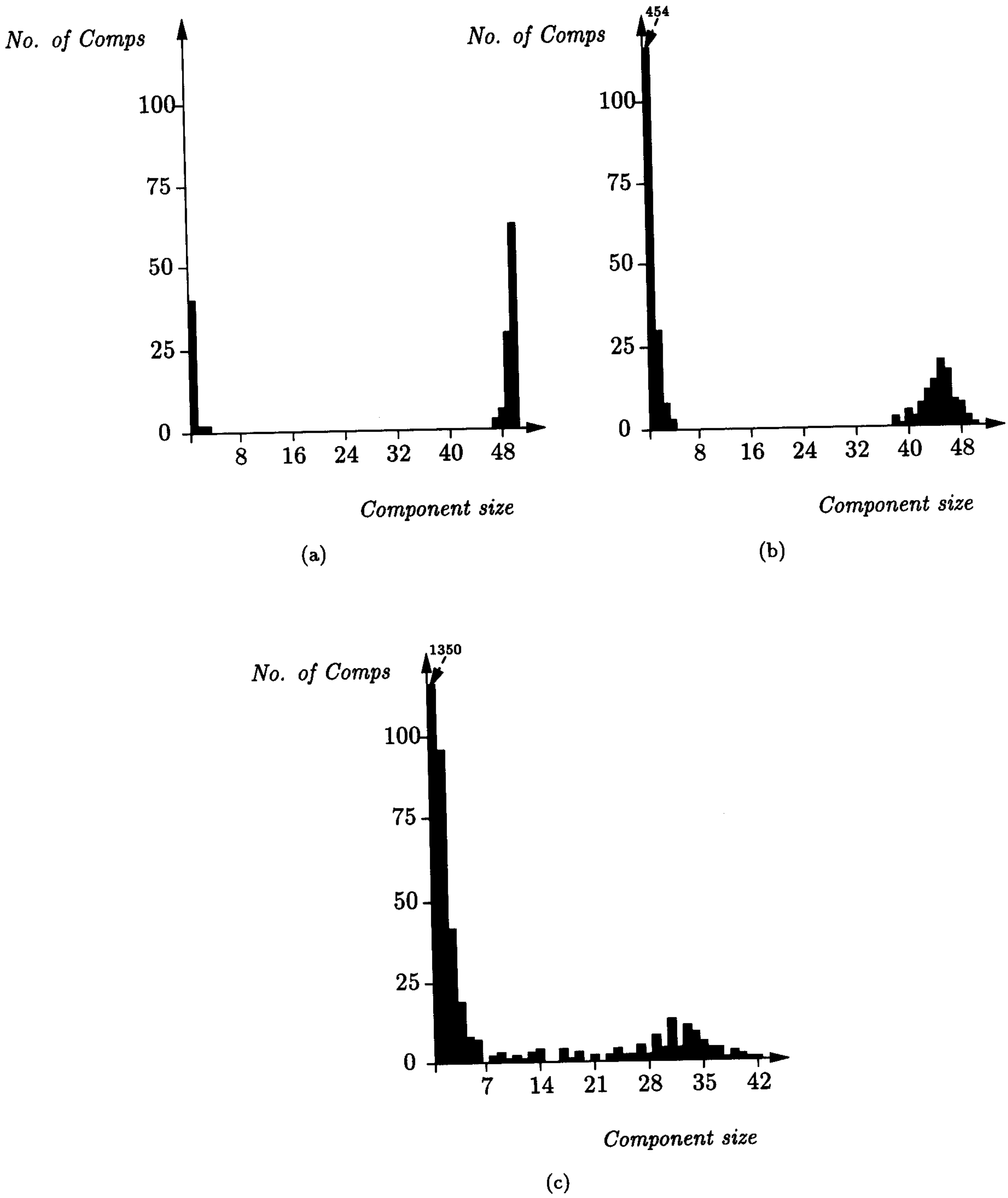

Figure 3: Cumulated Count for Digraphs with 50 Vertices and 75 Arcs, and 1 (a), 10 (b) and 25 (c) Pieces of Evidence Entered.

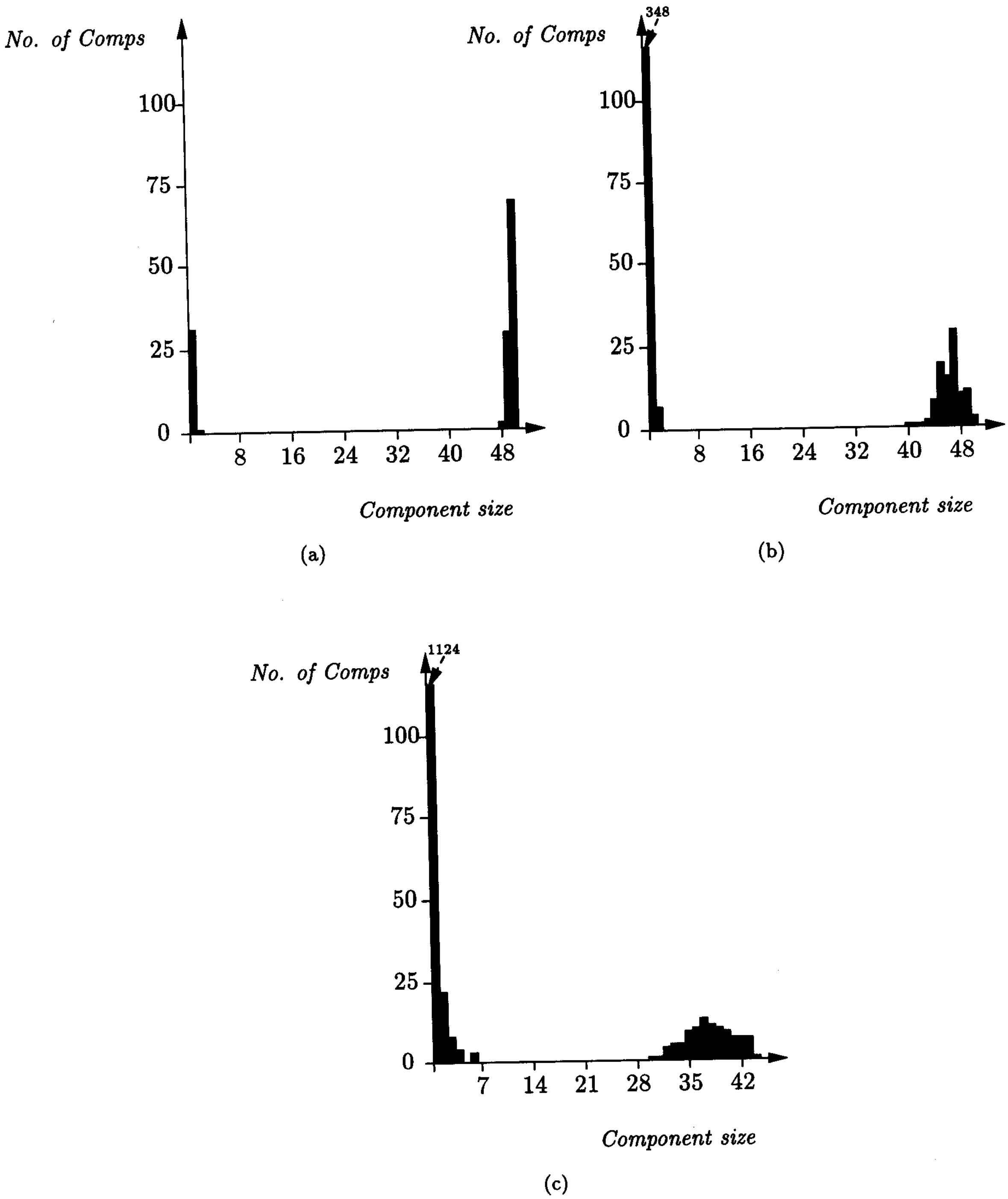

Figure 4: Cumulated Count for Digraphs with 50 Vertices and 100 Arcs, and 1 (a), 10 (b) and 25 (c) Pieces of Evidence Entered.

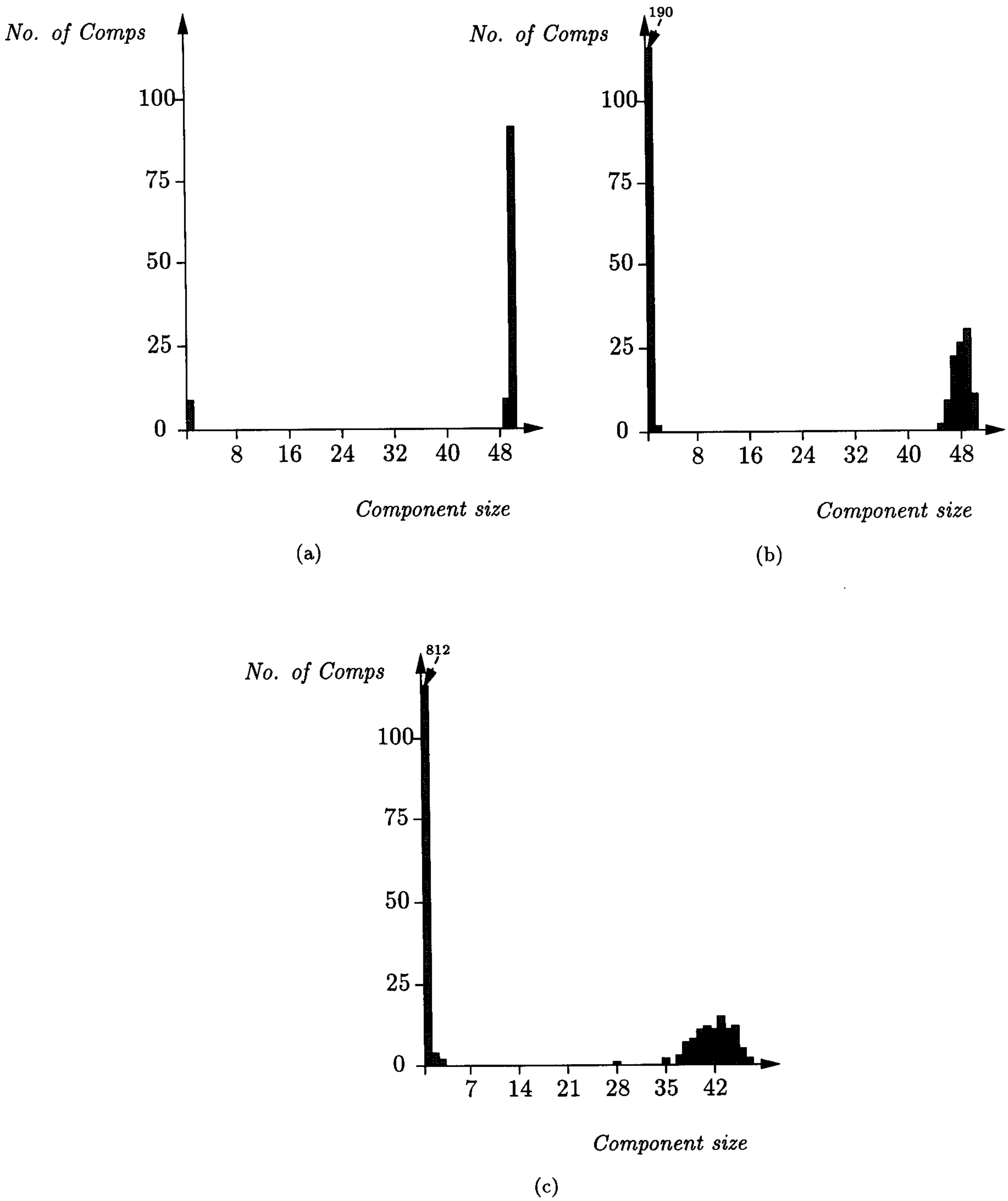

Figure 5: Cumulated Count for Digraphs with 50 Vertices and 150 Arcs, and 1 (a), 10 (b) and 25 (c) Pieces of Evidence Entered.

Number of pieces of evidence entered	1	10	25
Minimum number of arcs deleted	0	29	84
Maximum number of arcs deleted	14	84	151
Average number of deleted arcs	5.64	49.32	125.73
Average number of components	1.07	2.17	5.76
Average size of the minimum component	46.57	1.57	2.96
Average size of the maximum component	49.93	48.83	45.33

In addition to these statistics, Figure 6 presents the cumulated count of numbers of components per size.

3.3.2 The Second Experiment

Before presenting the results of our second experiment, we recall that the aim of this experiment has been to study in isolation the influence of the number of arcs on the behaviour of a digraph's topology under evidence absorption. In this experiment, we have generated several sets of digraphs comprising fifty vertices each; the pieces of evidence entered into these digraphs have been generated *randomly*. We have varied the number of arcs of the generated digraphs from fifty up to one hundred and fifty, increasing by two for each set. To each digraph generated, we have applied the method of evidence absorption for ten pieces of evidence selected. For the modified digraphs, we have found the statistics summarized in Figure 7; Figure 7 (a) shows the average number of deleted arcs, in Figure 7(b) the average number of components of the modified digraphs is shown, and Figures 7(c) and 7(d) plot the average sizes of the minimum and maximum component of the modified digraphs, respectively.

3.3.3 The Third Experiment

Before presenting the results of our third experiment, we recall that the aim of this experiment has been to study in isolation the influence of the number of pieces of evidence entered on the behaviour of a digraph's topology under evidence absorption. In this experiment, we have generated several sets of digraphs comprising fifty vertices each; we have fixed the number of arcs of these digraphs to one hundred. The pieces of evidence entered into these digraphs have been generated *randomly*; the number of pieces of evidence entered is varied from one up to twenty-five, increasing by one for each set of digraphs. To each digraph generated, we have applied the method of evidence absorption for the pieces of evidence selected. For the modified digraphs, we have found the statistics summarized in Figure 8; Figure 8 (a) shows the average number of deleted arcs, in Figure 8(b) the average number of components of the modified digraphs is shown, and Figures 8(c) and 8(d) plot the average sizes of the minimum and maximum component of the modified digraphs, respectively.

3.3.4 The Fourth Experiment

Before presenting the results of our fourth experiment, we recall that the aim of this experiment has been to investigate the influence of the location in the digraph of the vertices for which evidence is entered. In this experiment, we have generated several sets of digraphs comprising fifty vertices each; the number of arcs of the digraphs, the number of pieces of evidence entered, and the bias for selecting the evidence are varied.

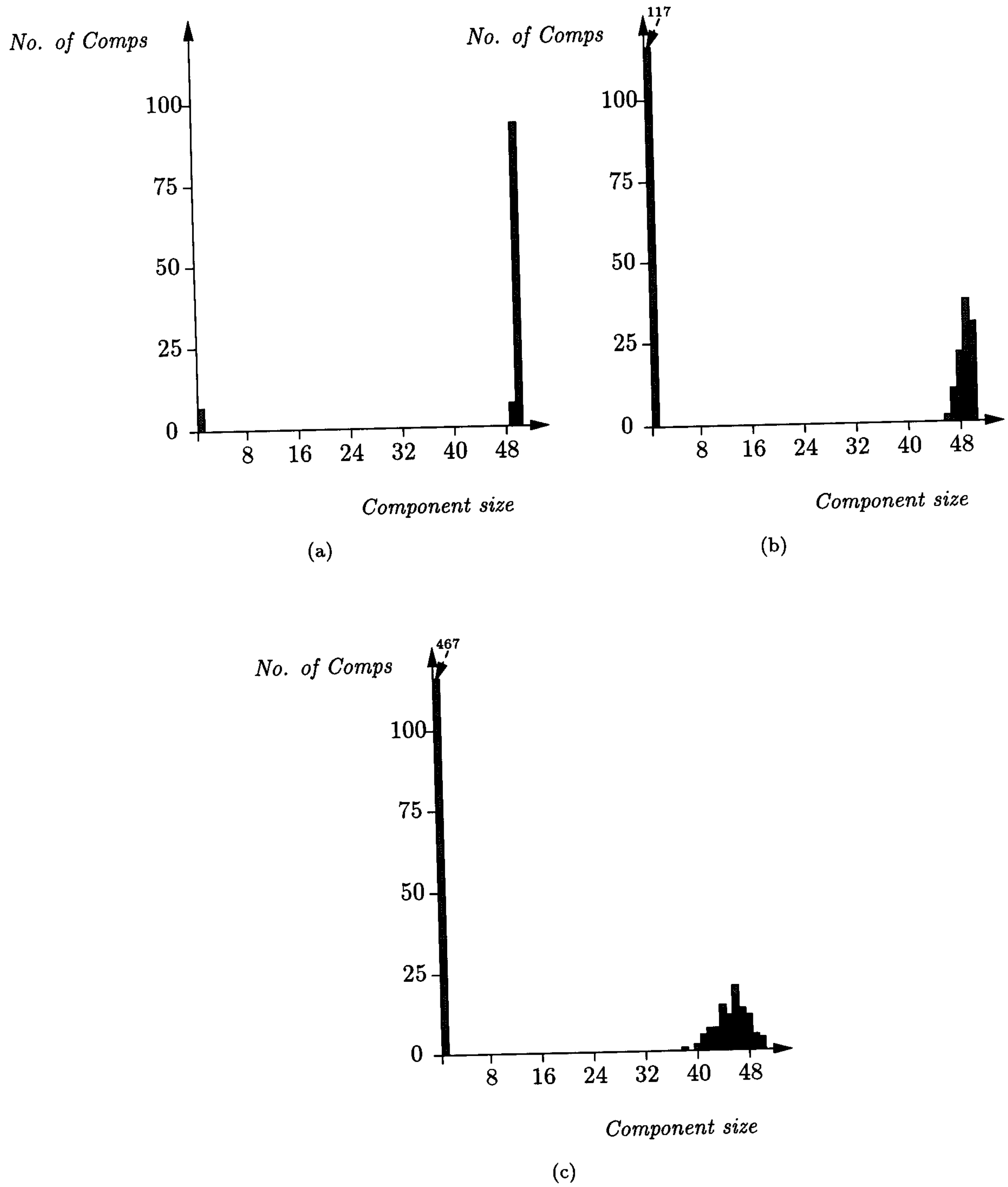

Figure 6: Cumulated Count for Digraphs with 50 Vertices and 250 Arcs, and 1 (a), 10 (b) and 25 (c) Pieces of Evidence Entered.

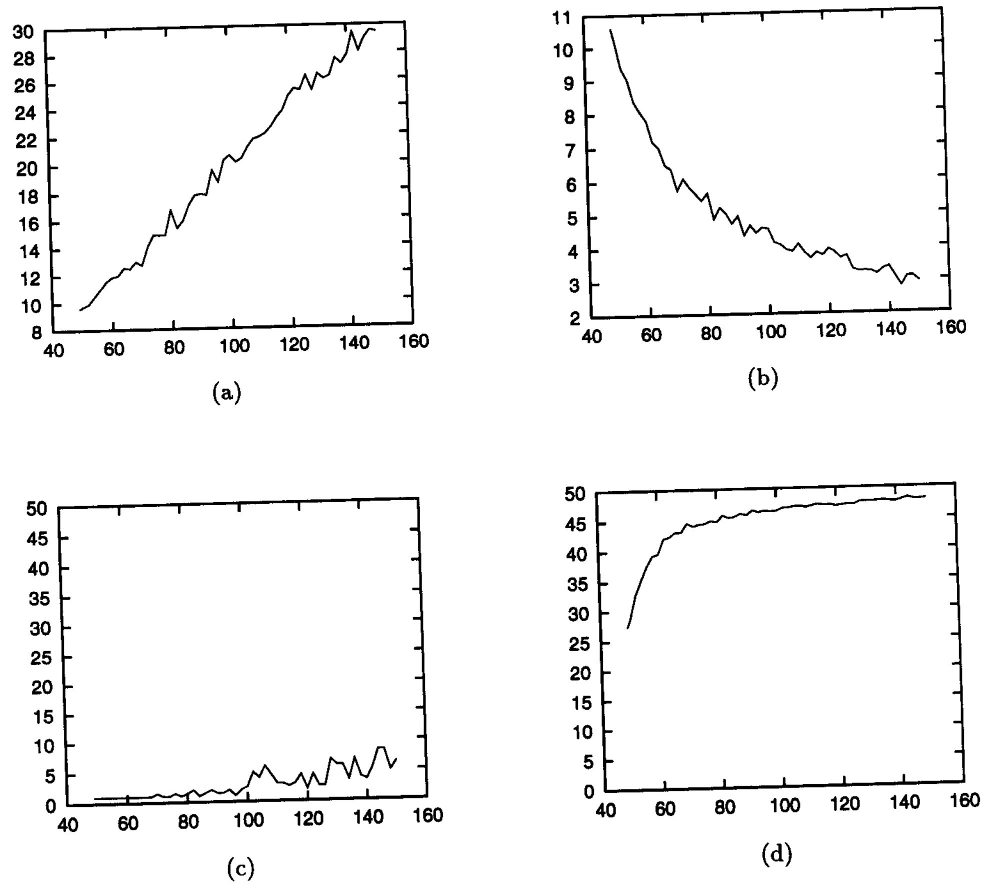

Figure 7: The Results of the Second Experiment — (a) The Average Number of Deleted Arcs, (b) The Average Number of Components, (c) The Average Size of the Minimum Component, (d) The Average Size of the Maximum Component

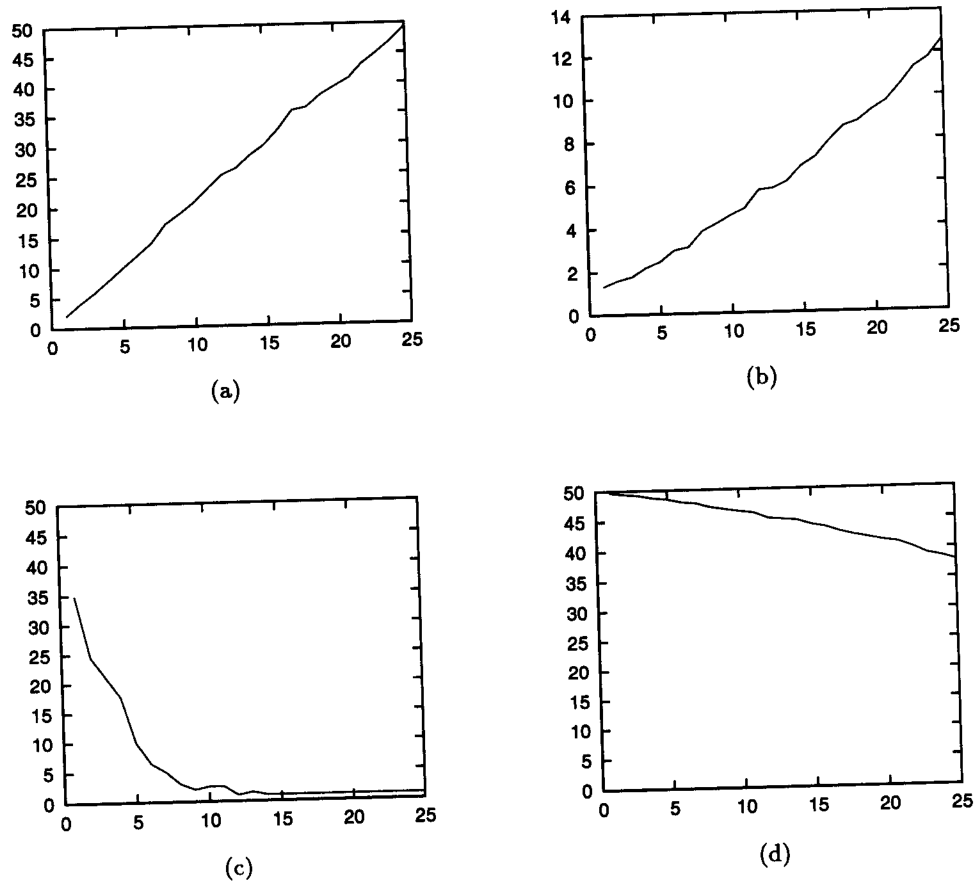

Figure 8: The Results of the Third Experiment — (a) The Average Number of Deleted Arcs, (b) The Average Number of Components, (c) The Average Size of the Minimum Component, (d) The Average Size of the Maximum Component

Test 1
We have generated six sets of *singly connected* digraphs. For the first set, we have selected *one* piece of evidence for each digraph using the bias for the lower part of the digraph; for the second set, the number of pieces of evidence equals *ten*; for the third set, this number equals *twenty-five*. To each digraph, the method of evidence absorption has been applied for the selected pieces of evidence. For the fourth, fifth, and sixth set, we have performed the same tests now using the upper bias. For the modified digraphs, we have found the following statistics:

Number of pieces of evidence entered	1		10		25	
Bias used	*lower*	*upper*	*lower*	*upper*	*lower*	*upper*
Average number of deleted arcs	0.45 (57%)	1.54 (195%)	5.15 (54%)	14.24 (149%)	16.11 (65%)	32.85 (132%)
Average number of components	1.45 (81%)	2.54 (142 %)	6.15 (58%)	15.24 (144%)	17.11 (66%)	33.85 (131%)
Average size of the minimum component	34.66	8.1	1.98	1	1	1
Average size of the maximum component	48.71	46.04	34.21	21.43	17.85	7.55

Note that the table also indicates the average number of deleted arcs as a percentage of the average number of deleted arcs found when selecting the vertices to obtain evidence for in a random fashion.

Test 2
We have generated six sets of *multiply connected* digraphs comprising *seventy-five* arcs. For the first set, we have selected *one* piece of evidence for each digraph using the bias for the lower part of the digraph; for the second set, the number of pieces of evidence equals *ten*; for the third set, this number equals *twenty-five*. To each digraph, the method of evidence absorption has been applied for the pieces of evidence selected. For the fourth, fifth and sixth set, we have performed the same tests now using the bias for the upper part of the digraph. For the modified digraphs, we have found the following statistics:

Number of pieces of evidence entered	1		10		25	
Bias used	*lower*	*upper*	*lower*	*upper*	*lower*	*upper*
Average number of deleted arcs	0.70 (52%)	2.34 (173%)	7.89 (53%)	21.37 (144%)	24.67 (68%)	50.28 (138%)
Average number of components	1.24 (86%)	1.65 (115%)	2.89 (49%)	9.22 (155%)	9.1 (56%)	26.97 (165%)
Average size of the minimum component	39.22	21.58	7.97	1	1	1
Average size of the maximum component	49.75	49.33	47.72	40.97	40.39	18.7

Test 3
We have generated six sets of *multiply connected* digraphs comprising *one hundred* arcs. For the first set, we have selected *one* piece of evidence for each digraph using the bias for the lower part of the digraph; for the second set, the number of pieces of evidence equals *ten*; for the third set, this number equals *twenty-five*. To each digraph, the method of evidence absorption has been applied for the pieces of evidence selected. For the fourth, fifth and sixth

set, we have performed the same tests now using the upper bias. For the modified digraphs, we have found the following statistics:

Number of pieces of evidence entered	1		10		25	
Bias used	*lower*	*upper*	*lower*	*upper*	*lower*	*upper*
Average number of deleted arcs	0.97 (47%)	3.00 (144%)	11.06 (54%)	29.31 (142%)	33.81 (68%)	68.26 (138%)
Average number of components	1.06 (80%)	1.57 (119%)	2.1 (46%)	7.46 (164%)	6.48 (51%)	23.63 (187%)
Average size of the minimum component	47.56	25.01	15.71	1	1	1
Average size of the maximum component	49.93	49.43	48.81	43.38	44.15	25.40

Test 4
We have generated six sets of *multiply connected* digraphs comprising *one hundred and fifty* arcs. For the first set, we have selected *one* piece of evidence for each digraph using the bias for the lower part of the digraph; for the second set, the number of pieces of evidence equals *ten*; for the third set, this number equals *twenty-five*. To each digraph, the method of evidence absorption has been applied for the pieces of evidence selected. For the fourth, fifth and sixth set, we have performed the same tests now using the upper bias. For the modified digraphs, we have found the following statistics:

Number of pieces of evidence entered	1		10		25	
Bias used	*lower*	*upper*	*lower*	*upper*	*lower*	*upper*
Average number of deleted arcs	1.27 (43%)	4.67 (158%)	16.50 (56%)	43.40 (147%)	50.32 (67%)	100.70 (135%)
Average number of components	1.02 (94%)	1.41 (129%)	1.46 (50%)	5.88 (201%)	3.45 (38%)	1.89 (206%)
Average size of the minimum component	49.02	30.4	30.41	1.49	3.94	1
Average size of the maximum component	49.98	49.59	49.53	45.09	47.53	31.86

Test 5
We have generated six sets of *multiply connected* digraphs comprising *two hundred and fifty* arcs. For the first set, we have selected *one* piece of evidence for each digraph using the bias for the lower part of the digraph; for the second set, the number of pieces of evidence equals *ten*; for the third set, this number equals *twenty-five*. To each digraph, the method of evidence absorption has been applied for the pieces of evidence selected. For the fourth, fifth and sixth set, we have performed the same tests now using the bias for the upper part of the digraph. For the modified digraphs, we have found the following statistics:

Number of pieces of evidence entered	1		10		25	
Bias used	*lower*	*upper*	*lower*	*upper*	*lower*	*upper*
Average number of deleted arcs	2.55 (45%)	7.61 (135%)	27.80 (56%)	72.54 (147%)	81.29 (65%)	168.21 (134%)
Average number of components	1.01 (94%)	1.3 (121%)	1.09 (50%)	4.53 (209%)	1.56 (27%)	16.13 (280%)
Average size of the minimum component	49.51	35.3	46.08	2.47	27.95	1
Average size of the maximum component	49.99	49.7	49.91	46.47	49.44	34.87

4 Discussion of the Results

In the previous section, we have outlined the experiments we have performed with the method of evidence absorption and presented the results obtained from these experiments. Here, we will closely examine these results.

4.1 The Number of Deleted Arcs

Applying the method of evidence absorption amounts to deleting from a digraph all arcs departing from a vertex for which evidence has been entered. We begin our discussion by considering the average numbers of deleted arcs found in the various experiments.

From a theoretical point of view, we observe that in a digraph comprising n vertices and m arcs, the average number of arcs departing from a vertex equals $\frac{m}{n}$. When applying the method of evidence absorption for one piece of evidence, the number of deleted arcs therefore is expected to approximate this ratio. Since deleting the arcs departing from one vertex does not influence the number of arcs departing from any of the other vertices in the digraph, we find that for k pieces of evidence the number of deleted arcs is expected to approximate $k \cdot \frac{m}{n}$. For a given digraph, this formula indicates a linear relation between the number of pieces of evidence entered and the number of arcs deleted by evidence absorption. The results of our experiments confirm this observation. For example, Figure 8(a) shows a linear increase in the number of deleted arcs for an increasing number of pieces of evidence entered. From the formula $k \cdot \frac{m}{n}$, we further observe that the number of arcs deleted by evidence absorption for a fixed number of pieces of evidence is related linearly to the total number of arcs comprised in the digraph at hand. This observation is also confirmed by our experiments. For example, Figure 7(a) indicates a linear increase in the number of deleted arcs for an increasing total number of arcs.

When using either the lower bias or the upper bias in selecting vertices to enter evidence for, the number of arcs deleted by the method of evidence absorption deviates from the number of arcs deleted when selecting vertices in a random fashion: the experiments show that when the lower bias is used the number of deleted arcs decreases, and that this number increases when the upper bias is used. These results are easily explained by considering the probability of selecting a vertex without any departing arcs in a digraph. Note that the more such vertices are selected, the fewer arcs will be deleted by evidence absorption. We recall that the lower and upper biases have been implemented building on a topological ordering of the vertices of a digraph. It will be evident that a vertex that is assigned a higher number in this ordering has a higher probability of having no arcs departing from it than a vertex that is assigned a low number in the ordering. As a consequence, when applying the lower bias there is a higher probability of selecting a vertex without any departing arcs than when applying the upper bias.

For the lower bias, we observe that the number of deleted arcs varies between 43% and 68% of the number of arcs that is deleted when vertices to enter evidence for are selected randomly. Note that this percentage increases as more pieces of evidence are entered into the digraph. This increase can be explained from the observation that the more pieces of evidence need to be generated, the higher in the digraph these vertices will have to be selected. A similar observation applies to the upper bias where the number of deleted arcs varies between 131% and 195% of the number of arcs that is deleted when vertices to enter evidence for are selected randomly.

4.2 The Components of the Modified Digraphs

The main aim of incorporating the method of evidence absorption into Pearl's algorithms for probabilistic inference is to improve on the computational effort spent on average-case problem solving. The impact of the method is most marked if the digraph of the network has fallen apart into separate components since then further probabilistic inference can be restricted to single components. We therefore closely examine the average numbers of components and their respective sizes found in the experiments.

In singly connected digraphs, any arc deletion causes the digraph to fall apart into separate components. For multiply connected digraphs, however, we observe that the more arcs are comprised in the digraph, the higher the probability that there is more than one path between two vertices and that arc deletion does not cause the digraph to fall apart. A theoretical analysis of the behaviour of the topology of a randomly generated multiply connected digraph under evidence absorption is far from trivial and would require further research. Since we feel that the method's impact in view of randomly generated digraphs may not be representative for its impact on real-life belief networks, we refrain from such an analysis. To nevertheless explain the results concerning the numbers of components of the modified digraphs obtained from our experiments and the sizes of these components, we compare the behaviour of the topology of a randomly generated digraph under evidence absorption with the behaviour of the topology of a random digraph under arc deletion. We would like to recall that the digraphs involved in our experiments are not truly random; moreover, the arcs to be deleted are not selected entirely at random either. However, we feel that our experiments incorporate enough randomness to justify such a comparison.

We consider the generation of a random digraph by successive addition of arcs between randomly selected vertices [Bollobas, 1985]. It will be evident that the more arcs are added to a digraph in the making, the more likely it is to become connected. A well-known result from random graph theory is that a random digraph with n vertices is almost always connected if it comprises $O(n \cdot \log n)$ arcs or more. Moreover, a random digraph with between $O(n)$ and $O(n \cdot \log n)$ arcs typically comprises one large component of $O(n)$ vertices, called the *giant component*, and many small components of size at most $O(\log n)$ each. Now consider adding to a digraph having the topology just desribed an arc between two randomly selected vertices. We distinguish between three situations:

- the new arc connects two vertices comprised in the giant component — the probability that this situation will occur is rather high and even increases as the giant component increases in size;
- the new arc connects one vertex from within the giant component and one vertex from within one of the tiny components — the probability that this situation will occur is fairly small and even diminishes as the giant component grows; note that since adding such an arc results in the giant component encapsulating a tiny one, we have that the probability that the giant component will increase in size is inversely proportional to its current size;
- the new arc connects two vertices not yet comprised in the giant component — the probability that this situation will occur is small and even diminishes as the giant component grows.

We now observe that the behaviour of the topology of a random digraph under arc deletion is dual to its behaviour under arc addition. From this observation we have that by successive arc deletion a connected random digraph will at first stay connected until it has shrunk to comprise approximately $O(n \cdot \log n)$ arcs. Further arc deletion will tend to yield a topology in which one giant component can be discerned and many tiny ones.

The digraphs generated in our experiments with the method of evidence absorption are rather sparse and therefore are likely to exhibit the behaviour outlined above. The cumulated counts of component sizes shown in the Figures 3, 4, 5, and 6 clearly reflect the giant-component topology. The behaviour of the giant component itself is seen most markedly in Figures 7(d) and 8(d). Figure 7(d) shows that as the number of arcs of the generated digraphs increases, the size of the giant component rapidly rises to approximate the number of vertices of the digraphs; note that the amount of increase in size of the giant component for an increase in the number of arcs is inversely proportional to the size the component already has. Figure 8(d) shows that as the number of pieces of evidence entered, and hence the number of deleted arcs, increases, the giant component slowly decreases in size. Figures 7(b) and 8(b) depict the average number of components found in our experiments. Figure 7(b) shows that the number of components rapidly decreases as the number of arcs of the digraphs, and hence the size of the giant component, increases; Figure 8(b) shows that the number of components increases as the number of pieces of evidence entered increases. Both Figure 7(c) and 8(c) demonstrate that the size of the minimum component, and hence the size of the tiny components, is very small compared to the size of the giant component.

5 The Use of Randomly Generated Belief Networks

The aim of the experiments reported in this paper has been to gain insight into the ability of the method of evidence absorption to improve on the computational expense involved in probabilistic inference. We have chosen to perform our experiments on randomly generated digraphs to avoid the risk of *fine-tuning* the set-up of the experiments to the method to be investigated. A close examination of the results obtained from the experiments has revealed several interesting properties of the method of evidence absorption. Yet, from the discussion in the previous section it will be evident that these properties to a large extent derive from applying the method to randomly generated digraphs — in fact, the results obtained from our experiments cannot be exploited for drawing detailed conclusions as to the method's behaviour on belief networks that do not incorporate a random digraph. Unfortunately, the digraphs found in present-day real-life belief networks do not exhibit a random topology.

There are several alternative set-ups for experimentation with the method of evidence absorption possible that do not involve using randomly generated belief networks. One alternative is to apply the method to a range of existing real-life belief networks. Experiments with the method of evidence absorption on real-life belief networks would give accurate insight into the method's true ability. At present, however, only few full-scaled, real-life belief networks are available, rendering extensive experiments on such networks practically infeasible. Another set-up for experimentation is to closely examine existing belief networks and derive general properties the digraph of a real-life belief network is expected to have — experiments then are performed on randomly generated networks exhibiting these properties. However, although most present-day belief networks have been designed for the same task, namely the task of diagnosis, and therefore are expected to share some characteristics, they show a large variety in their digraph's topology [Wessels, 1994]. This variety can be explained to some extent by the differences in the respective domains. Yet, the variety in topologies can at least partly be attributed to there being no consensus as to the wished-for properties of a belief network other than most general ones. In fact, most existing networks are taylored to non-standardized state-of-the-art methods for reasoning with a belief network which tend to impose rather strong restrictions on the topology of the graphical part of the network. Since research on reasoning methods rapidly progresses, future belief networks may very well differ considerably from present-day networks. We conclude that at present it is not possible to draw any decisive conclusions as to the properties a realistic belief network is expected to have. Despite this observation, we do not expect that future belief networks will comprise digraphs of random topology. We feel that as applications grow larger, the digraphs involved will tend to have a topology in which subgraphs with a high degree of connectivity can be discerned modeling different focal areas of attention of the domain at hand; these dense subgraphs will tend to be loosely interconnected. As long as this tendency is not confirmed by full-scale real-life networks, however, setting up experiments along these lines runs the risk of fine-tuning to the method to be investigated.

From the above observations, we conclude that although the use of randomly generated belief networks in our experiments may leave much to be desired, at present it seems to be the only feasible set-up for experimentation.

6 Conclusions

As more and more real-life applications of the belief network framework begin to emerge, it is becoming apparent that the basic algorithms involved in probabilistic inference tend to slow down problem solving. Recent research therefore aims at improving these basic algorithms. In a previous paper, we proposed the method of evidence absorption to this end. To gain some insight in the ability of this method to improve on the computational expense involved in inference, we performed several experiments on different classes of randomly generated belief networks. Unfortunately, the results obtained from these experiments to a large extent reflect the use of randomly generated belief networks and do not provide for drawing detailed conclusions as to the method's behaviour on real-life networks.

In present-day real-life applications of the belief network framework, the networks involved exhibit a large variety in their digraph's topology. Since the impact of applying the method of evidence absorption on probabilistic inference to a large extent is determined by the topological properties of the digraph of the network at hand, it will have to be decided for each belief network separately whether or not applying evidence absorption is expected to be advantageous. To this end, a simple investigation of the location in the network's digraph of the vertices for which evidence is likely to be entered suffices; note that it is not so much the number of outgoing arcs of these vertices that determines the impact of applying evidence absorption as these vertices' ability to let the digraph fall apart into components of approximately equal size upon evidence absorption. We would like to emphasize that applying the method of evidence absorption does not weigh on the computational complexity of probabilistic inference.

We have mentioned before that due to the use of randomly generated belief networks the results obtained from our experiments with the method of evidence absorption do not provide for drawing detailed conclusions as to the method's behaviour on real-life networks. Although our experiments were designed to investigate the behaviour of one specific method only, we feel that a similar observation applies to using randomly generated belief networks in other experiments in which the topology of the digraph of a network plays a central role. As long as there is no well-established insight into the properties of real-life belief networks other than most general ones, however, using randomly generated networks may be the only feasible method for experimentation. Yet, care has to be taken in drawing conclusions from the results of such experiments.

Acknowledgment

I am indebted to Han La Poutré for sharing his knowledge of random graphs and for thus contributing to the interpretation of the results obtained from the experiments reported in this paper. Also, I am very grateful to Peter Lucas for his help in improving the presentation of the results and for his many helpful comments on earlier versions of the paper.

References

[Andreassen et al., 1987] S. Andreassen, M. Woldbye, B. Falck, S.K. Andersen (1987). MUNIN - A causal probabilistic network for interpretation of electromyographic findings, *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, pp. 366 – 372.

[Bollobas, 1985] B. Bollobas (1985). *Random Graphs*, Academic Press, London.

[Cooper, 1990] G.F. Cooper (1990). The computational complexity of probabilistic inference using Bayesian belief networks, *Artificial Intelligence*, vol. 42, pp. 393 – 405.

[Heckerman *et al.*, 1992] D.E. Heckerman, E.J. Horvitz, B.N. Nathwani (1992). Toward normative expert systems. Part I: the Pathfinder project, *Methods of Information in Medicine*, vol. 31, pp. 90 – 105.

[Pearl, 1988] J. Pearl (1988). *Probabilistic Reasoning in Intelligent Systems. Networks of Plausible Inference*, Morgan Kaufmann, Palo Alto.

[Suermondt & Cooper, 1990] H.J. Suermondt & G.F. Cooper (1990). Probabilistic inference in multiply connected belief networks using loop cutsets, *International Journal of Approximate Reasoning*, vol. 4, pp. 283 – 306.

[van der Gaag, 1993] L.C. van der Gaag (1993). *Evidence Absorption for Belief Networks*, Technical Report CS-RUU-93-35, Utrecht University.

[Wessels, 1994] M.L. Wessels (1994). *Belief Networks in Diagnostic Applications: an Overview*, in preparation.