

**Ya.S. Bondarenko, D.O. Rachko, A.O. Rozlyvan**  
*Oles Honchar Dnipro National University*

## **PROBABILISTIC INFERENCE IN BAYESIAN INSURANCE NETWORK**

In this paper, the technique to solve the prediction problem of reparation of the financial losses caused by a road traffic accident is solved. Exact inference is represented using the Sum-Product Variable Elimination algorithm, Sum-Product Variable Elimination algorithm for computing conditional probabilities, Max-Product Variable Elimination algorithm for MAP, Max-Sum-Product Variable Elimination algorithm for marginal MAP. Reasoning patterns are presented graphically and descriptively.

**Keywords:** Bayesian network, marginal distribution, joint distribution, conditional distribution, conditional independence, posterior distribution, MAP query, marginal MAP query.

**Я.С. Бондаренко Д.О. Рачко, А.О. Розливан**  
*Дніпровський національний університет імені Олеся Гончара*

## **ІМОВІРНІСНИЙ ВИСНОВОК В БАЙЄСІВСЬКІЙ МЕРЕЖІ "СТРАХУВАННЯ"**

В роботі запропонована методика розв'язання задачі прогнозування страхового відшкодування майнової шкоди завданої внаслідок дорожньо-транспортної пригоди. Мета клієнта, який укладає договір страхування зі страховою компанією, – уникнути фінансових витрат, пов'язаних із випадковістю настання тих чи інших небажаних подій (страхових випадків): транспортних пригод, смерті застрахованого та ін. До укладання договору страхування клієнт має певний ризик, який може призвести до випадкових втрат, а може й не призвести до них. Уклавши договір страхування зі страховою компанією, клієнт за певну плату купує страховий захист, тобто за плату клієнт уникає можливих втрат, пов'язаних зі страховим випадком, які хоча й малоймовірні, але для клієнта можуть бути катастрофічно великими. При цьому ризик не зникає – його перебирає на себе страхова компанія. Забезпечення фінансової стійкості страхової компанії передбачає зусилля як її керівників, так і економістів, юристів, математиків.

Робота присвячена формуванню ймовірнісного висновку на ймовірнісній графічній моделі – байєсівській мережі – у вигляді спрямованого ациклічного графу, який призначений для моделювання та візуалізації інформації щодо задачі прогнозування страхового відшкодування майнової шкоди завданої внаслідок дорожньо-транспортної пригоди та формування статистичного висновку – прийняття рішення щодо поставленої задачі. Байєсівська мережа дає можливість встановити причинно-наслідкові зв'язки між подіями та визначити ймовірності настання тієї чи іншої ситуації при отриманні нової інформації щодо зміни стану будь-якої змінної мережі.

Досліджено клас алгоритмів, заснованих на виключенні змінних, зокрема, аналітично викладно точний ймовірнісний висновок за допомогою класичного алгоритму виключення змінних, алгоритму виключення змінних при обчисленні умовного розподілу, алгоритму виключення змінних для оцінювання максимуму апостеріорної ймовірності, алгоритму виключення змінних для оцінювання максимуму маргіналь-

ної апостеріорної ймовірності. Моделі міркувань, побудовані за допомогою алгоритмів точного імовірнісного висновку, представлено графічно та описово.

Здобуті результати можуть бути використані в практичній діяльності страхової компанії, зокрема, при прогнозуванні страхового відшкодування та знаходження причин виникнення дорожньо-транспортних пригод.

**Ключові слова:** байєсівська мережа, маргінальний розподіл, спільний розподіл, умовний розподіл, умовна незалежність, оцінка максимуму апостеріорної ймовірності, оцінка максимуму маргінальної апостеріорної ймовірності.

**Я.С. Бондаренко, Д.А. Рачко, А.А. Розливан**

*Дніпровський національний університет імені Олеся Гончара*

## **ВЕРОЯТНОСТНЫЙ ВЫВОД В БАЙЕСОВСКОЙ СЕТИ "СТРАХОВАНИЕ"**

В работе предложена методика решения задачи прогнозирования страхового возмещения материального ущерба, причиненного в случае дорожно-транспортного происшествия. Сформировано точный вероятностный вывод с помощью классического алгоритма исключения переменных, алгоритма исключения переменных для вычисления условного распределения, алгоритма исключения переменных для оценивания максимума апостериорной вероятности, алгоритма исключения переменных для оценивания максимума маргинальной апостериорной вероятности. Модели рассуждений представлены графически и описательно.

**Ключевые слова:** байесовская сеть, маргинальное распределение, совместное распределение, условное распределение, условная независимость, апостериорное распределение вероятностей, оценка максимума апостериорной вероятности, оценка максимума маргинальной апостериорной вероятности.

**Introduction.** The aim of the customer who concludes an insurance contract with the insurance company is to avoid financial losses associated with the occurrence of certain adverse events (insurance cases): traffic accidents, death of the insured, etc. Before concluding the insurance contract, the customer has a certain risk, which may lead to accidental losses or may not lead to them. After concluding the insurance contract with the insurance company, the customer buys insurance coverage for a certain fee, i.e. the customer avoids possible losses related to the insured event, which, although unlikely, can be catastrophically large for the customer. It should be noted that the risk does not disappear: it is taken over by the insurance company for a fee. Therefore, financial risk and the associated risk of bankruptcy are the hallmarks of each insurance company. Ensuring the financial stability of an insurance company involves the efforts of both its managers and economists, lawyers, mathematicians. Mathematical aspects of this problem are studied in risk theory which allows us to quantify the financial risks in the activities of the insurance company. Many excellent books on probabilistic graphical models [1-8] and online courses [9-11] cover many innovative concepts and going through cutting-edge technologies in the implementation of algorithms for doing exact and approximate inference as well as learning to find practical solution of real-world problems in financial and scientific areas.

An adaptive probabilistic network algorithm for learning the insurance network from data was presented in [12]. Prior knowledge in the form of known structure with hidden variables significantly reduced the amount of data required to train the network as well as allowed to improve the learning rate of the Bayesian network.

An insurance recommendation system to predict products for current and prospective customers for which missing data are accustomed was described in [13]. Experimental results showed that Bayesian networks outperformed Low-Rank Matrix Factorization models by generating recommendations based on customer portfolio and policy data.

The Bayesian network structure learning of road accidents was performed by using expert experience and the K2 algorithm, as well as, the Bayesian network parameter learning of road accidents was conducted by using the Expectation-Maximization algorithm that can deal with missing data [14]. Posterior probability reasoning found the causes for road accidents and identified the most likely combination of variable states which led to the occurrence of an accident.

Bayesian network to analyze road accidents and find an effective way to avoid them was designed by using the K2 algorithm for Bayesian network structure learning [15]. This model can not only perform accident diagnosis, but also accident prediction.

Many software tools are available for decision making and reasoning with Bayesian networks under uncertainty. We chose state-of-the-art software SamIam: Sensitivity Analysis, Modeling, Inference and More [6] for modeling and analysis of the Bayesian network, developed the Automated Reasoning Group at the University of California, Los Angeles.

In this paper, the basic algorithms for exact inference, MAP inference, marginal MAP inference in training insurance network are described. The paper consists of six sections. The second section is devoted to the main definitions and theorems. The next sections deal with Insurance network representation, with Exact inferences using Sum-Product and Conditioning algorithms and with MAP and marginal MAP inferences using Max-Product and Max-Sum-Product algorithms.

### Probability queries.

**Definition 1** [5]. A Bayesian network is a pair  $B = (G, P)$  where  $G$  is a directed acyclic graph over the nodes  $X_1, \dots, X_n$ ,  $P$  is a joint distribution that factorizes over  $G$ .

**Definition 2** [5]. Graph  $G$  encodes the set of conditional independence assumptions which connected with a joint distribution  $P$  if each node  $X_i$  is conditionally independent of its nondescendants given its parents:

$$P(X_i | X_1, \dots, X_{i-1}) = P(X_i | Parents_{X_i}), i = 1, \dots, n, \quad (1)$$

where  $X_1, \dots, X_n$  is a topological ordering relative to  $G$ .

**Definition 3** [5]. The distribution of a single random variable  $X_i$ :

$$P(X_i) = \sum_{X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n} P(X_1, \dots, X_n) \quad (2)$$

is called a marginal of the joint probability distribution  $P(X_1, \dots, X_n)$ . The process of computing a marginal from joint distribution is called marginalization.

Let  $G$  be a Bayesian network over the variables  $X_1, \dots, X_n$ . A joint distribution  $P(X_1, \dots, X_n)$  factorizes according to  $G$  if it can be expressed as a product conditional probability distributions (CPDs):

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | \text{Parents}_{X_i}). \quad (3)$$

This equation is called the chain rule for the Bayesian network.

**Theorem 1** [5]. Let  $G$  be a Bayesian network over a set of random variables  $X_1, \dots, X_n$ , and let  $P$  be a joint distribution over the same variables. If the graph  $G$  encodes the set of conditional independence assumptions which connected with a joint distribution  $P$ , then  $P$  factorizes accordingly  $G$ .

**Theorem 2** [5]. Let  $G$  be a Bayesian network over a set of random variables  $X_1, \dots, X_n$ , and let  $P$  be a joint distribution over the same variables. If  $P$  factorizes according to  $G$ , then  $G$  encodes the set of conditional independence assumptions which connected with a joint distribution  $P$ .

A Bayesian network can be used to answer the different types of queries of interest such as a probability query, a MAP query, and a marginal MAP query.

A probability query aims to compute a posterior probability distribution of some query variables  $Y$  given evidence  $E$  on others:

$$P(Y | E = e). \quad (4)$$

A MAP query aims to find the most likely assignment to all of the non-evidence variables  $W$ :

$$\text{MAP}(W | E = e) = \arg \max_w P(W | E = e). \quad (5)$$

A marginal MAP query aims to find the most likely assignment to query variables  $Y$ , marginalizing out over the rest  $Z$ :

$$\text{MAP}(Y | E = e) = \arg \max_Y \sum_Z P(Y, Z | E = e). \quad (6)$$

**Insurance Bayesian Network Representation.** Bayesian network is a directed acyclic graph, whose nodes represent the random variables and whose edges represent the direct influence of one node on another.

The Bayesian network known as Insurance [9] that is used as training Bayesian network throughout this paper is shown in fig. 1. So a driver attention span *Focused* and a driver age *Age* affects deeply on a driving school test result *Good Student*; a driver attention span *Focused*, a driver age *Age* and taking an extreme driving course *Extra Training* influence profoundly on a level of driver proficiency *Driver Quality*; a level of driver skills *Driver Quality* affects strongly on a number of accidents *Driving History*; a driver attention span *Focused* moves deeply on a car-size *Vehicle Size*; a level of driver proficiency *Driver Quality*, a car-size *Vehicle Size*, year of manufacture *Vehicle Year*, an airbag *Airbag* and antilock braking systems *Antilock Brakes* influence profoundly on the severity of accident *Accident*; year of manufacture *Vehicle Year* and severity of accident *Accident* affect strongly on the cost of insurance policy *Cost*.

Let  $X_1, \dots, X_{12}$  be a topological ordering of the nodes relative to the graph  $G$ . The joint distribution  $P(X_1, \dots, X_{12})$  can be decomposed according to the chain rule for probabilities in the following way:

$$\begin{aligned} P(X_1, \dots, X_{12}) = & P(X_{12} | X_1, \dots, X_{11}) P(X_{11} | X_1, \dots, X_{10}) \cdot \\ & \cdot P(X_{10} | X_1, \dots, X_9) P(X_9 | X_1, \dots, X_8) P(X_8 | X_1, \dots, X_7) P(X_7 | X_1, \dots, X_6) \cdot \\ & \cdot P(X_6 | X_1, \dots, X_5) P(X_5 | X_1, \dots, X_4) P(X_4 | X_1, X_2, X_3) P(X_3 | X_1, X_2) P(X_2 | X_1) P(X_1). \end{aligned}$$

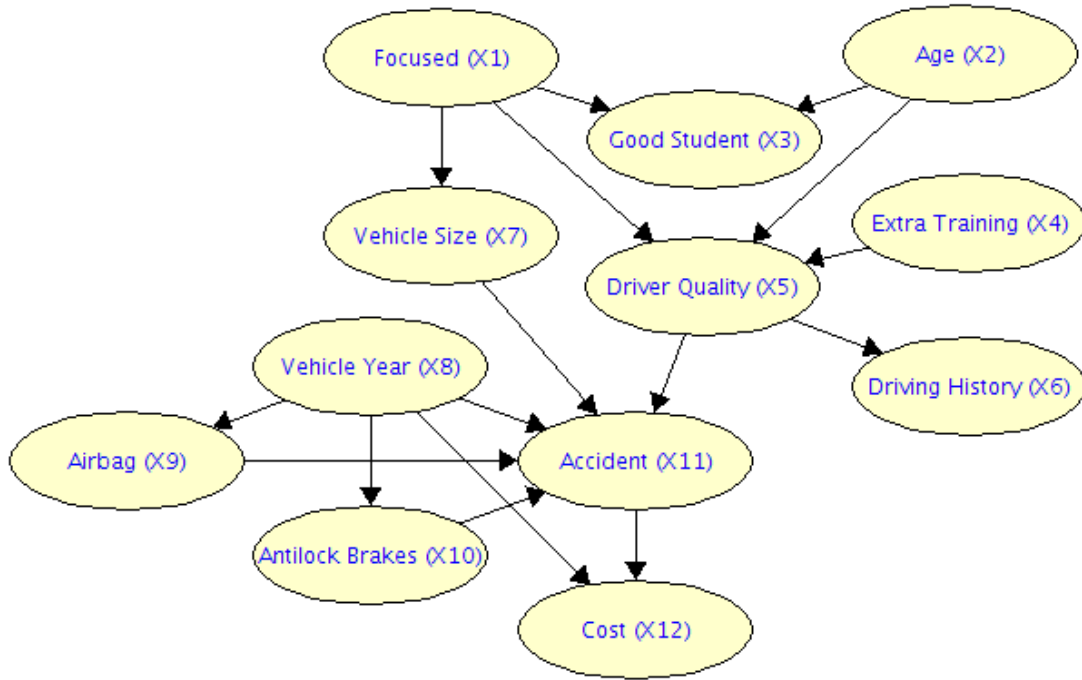
A set of conditional independence assumptions

$$\begin{aligned} & (X_2 \perp X_1), (X_4 \perp X_3, X_2, X_1), (X_5 \perp X_3 | X_4, X_2, X_1), \\ & (X_6 \perp X_4, X_3, X_2, X_1 | X_5), (X_7 \perp X_6, X_5, X_4, X_3, X_2 | X_1), \\ & (X_8 \perp X_7, X_6, X_5, X_4, X_3, X_2, X_1), \\ & (X_9 \perp X_7, X_6, X_5, X_4, X_3, X_2, X_1 | X_8), \\ & (X_{10} \perp X_9, X_7, X_6, X_5, X_4, X_3, X_2, X_1 | X_8), \\ & (X_{11} \perp X_6, X_4, X_3, X_2, X_1 | X_{10}, X_9, X_8, X_7, X_5), \\ & (X_{12} \perp X_{10}, X_9, X_7, X_6, X_5, X_4, X_3, X_2, X_1 | X_{11}, X_8) \end{aligned}$$

allows to factorize the joint distribution  $P(X_1, \dots, X_{12})$  on conditional probability distributions (CPDs):

$$\begin{aligned} P(X_1, \dots, X_{12}) = & P(X_{12} | X_{11}, X_8) P(X_{11} | X_{10}, X_9, X_8, X_7, X_5) P(X_{10} | X_8) P(X_9 | X_8) \cdot \\ & \cdot P(X_8) P(X_7 | X_1) P(X_6 | X_5) P(X_5 | X_1, X_2, X_4) P(X_4) P(X_3 | X_2, X_1) P(X_2) P(X_1). \end{aligned}$$

The major advantage of the Bayesian network representation is that the number of parameters in resulting factorized form is equal to  $2^5 3$ . For the sake of comparison, the number of parameters in the joint distribution is equal to  $2^{10} 3^2 - 1$ .



**Fig. 1. The Insurance network structure**

**Exact Inference in Insurance Bayesian Network.**

We consider the Sum-Product Variable Elimination algorithm and Conditioning algorithm, which use dynamic programming techniques to perform exact inference in a reasonable time.

The basic idea of the Sum-Product Variable Elimination algorithm [5] is that we sum out random variables one at a time. When we sum out any random variable, we multiply all the CPDs that mention that variable, generating a product factor. We sum out the variable from this combined factor, generating a new factor that we enter into set of CPDs to be dealt with.

We apply this approach to compute the marginal distribution of query variable *Cost*, eliminating all the nonquery variables:

$$\begin{aligned}
 P(X_{12}) &= \sum_{X_1, \dots, X_{11}} P(X_1, \dots, X_{12}) = \\
 &= \sum_{X_1, \dots, X_{11}} P(X_{12} | X_{11}, X_8) P(X_{11} | X_{10}, X_9, X_8, X_7, X_5) P(X_{10} | X_8) P(X_9 | X_8) \cdot \\
 &\cdot P(X_8) P(X_7 | X_1) P(X_6 | X_5) P(X_5 | X_1, X_2, X_4) P(X_4) P(X_3 | X_2, X_1) P(X_2) P(X_1).
 \end{aligned}$$

Consider eliminating nodes in the order

$$Z = \{X_4, X_6, X_2, X_3, X_1, X_5, X_7, X_9, X_{10}, X_8, X_{11}\},$$

which was found according to the greedy search for constructing an elimination ordering [5]. We can transform summation into the following way:

$$P(X_{12}) = \sum_{X_{11}} \sum_{X_8} P(X_{12} | X_{11}, X_8) P(X_8) \sum_{X_{10}} P(X_{10} | X_8) \sum_{X_9} P(X_9 | X_8) \cdot$$

$$\cdot \sum_{X_7} \sum_{X_5} P(X_{11} | X_{10}, X_9, X_8, X_7, X_5) \sum_{X_1} P(X_1) P(X_7 | X_1) \cdot \\ \cdot \sum_{X_3} \sum_{X_2} P(X_3 | X_2, X_1) P(X_2) \sum_{X_6} P(X_6 | X_5) \sum_{X_4} P(X_5 | X_4, X_2, X_1) P(X_4).$$

Firstly, we compute the factors by eliminating variable  $X_4$ :

$$\psi_1(X_1, X_2, X_4, X_5) = P(X_5 | X_1, X_2, X_4) P(X_4), \\ \tau_1(X_1, X_2, X_5) = \sum_{X_4} \psi_1(X_1, X_2, X_4, X_5).$$

Secondly, we compute the factors by eliminating variable  $X_6$ :

$$\psi_2(X_6, X_5) = P(X_6 | X_5), \\ \tau_2(X_5) = \sum_{X_6} \psi_2(X_6, X_5).$$

Thirdly, we compute the factors by eliminating variable  $X_2$ :

$$\psi_3(X_5, X_3, X_2, X_1) = P(X_3 | X_1, X_2) P(X_2) \tau_1(X_1, X_2, X_5), \\ \tau_3(X_5, X_3, X_1) = \sum_{X_2} \psi_3(X_5, X_3, X_2, X_1).$$

Fourthly, we compute the factors by eliminating variable  $X_3$ :

$$\psi_4(X_5, X_3, X_1) = \tau_3(X_5, X_3, X_1), \\ \tau_4(X_5, X_1) = \sum_{X_3} \psi_4(X_5, X_3, X_1).$$

Fifthly, we compute the factors by eliminating variable  $X_1$ :

$$\psi_5(X_7, X_5, X_1) = P(X_1) P(X_7 | X_1) \tau_4(X_5, X_1), \\ \tau_5(X_7, X_5) = \sum_{X_1} \psi_5(X_7, X_5, X_1).$$

Sixthly, we compute the factors by eliminating variable  $X_5$ :

$$\psi_6(X_{11}, X_{10}, X_9, X_8, X_7, X_5) = P(X_{11} | X_{10}, X_9, X_8, X_7, X_5) \tau_2(X_5) \tau_5(X_7, X_5), \\ \tau_6(X_{11}, X_{10}, X_9, X_8, X_7) = \sum_{X_5} \psi_6(X_{11}, X_{10}, X_9, X_8, X_7, X_5).$$

Seventhly, we compute the factors by eliminating variable  $X_7$ :

$$\psi_7(X_{11}, X_{10}, X_9, X_8, X_7) = \tau_6(X_{11}, X_{10}, X_9, X_8, X_7), \\ \tau_7(X_{11}, X_{10}, X_9, X_8) = \sum_{X_7} \psi_7(X_{11}, X_{10}, X_9, X_8).$$

Eighthly, we compute the factors by eliminating variable  $X_9$ :

$$\psi_8(X_{11}, X_{10}, X_9, X_8) = \tau_7(X_{11}, X_{10}, X_9, X_8) P(X_9 | X_8), \\ \tau_8(X_{11}, X_{10}, X_8) = \sum_{X_9} \psi_8(X_{11}, X_{10}, X_9, X_8).$$

Ninthly, we compute the factors by eliminating variable  $X_{10}$ :

$$\psi_9(X_{11}, X_{10}, X_8) = \tau_8(X_{11}, X_{10}, X_8) P(X_{10} | X_8),$$

$$\tau_9(X_{11}, X_8) = \sum_{X_{10}} \psi_9(X_{11}, X_{10}, X_8).$$

Tenthly, we compute the factors by eliminating variable  $X_8$ :

$$\begin{aligned} \psi_{10}(X_{12}, X_{11}, X_8) &= \tau_9(X_{11}, X_8)P(X_{12}|X_{11}, X_8)P(X_8), \\ \tau_{10}(X_{12}, X_{11}) &= \sum_{X_8} \psi_{10}(X_{12}, X_{11}, X_8). \end{aligned}$$

Finally, we compute the factors by eliminating variable  $X_{11}$ :

$$\begin{aligned} \psi_{11}(X_{12}, X_{11}) &= \tau_{10}(X_{12}, X_{11}), \\ \tau_{11}(X_{12}) &= \sum_{X_{11}} \psi_{11}(X_{12}, X_{11}). \end{aligned}$$

and distribution of node  $X_{12}$  has the following form:

$$P(X_{12}) = \tau_{11}(X_{12}).$$

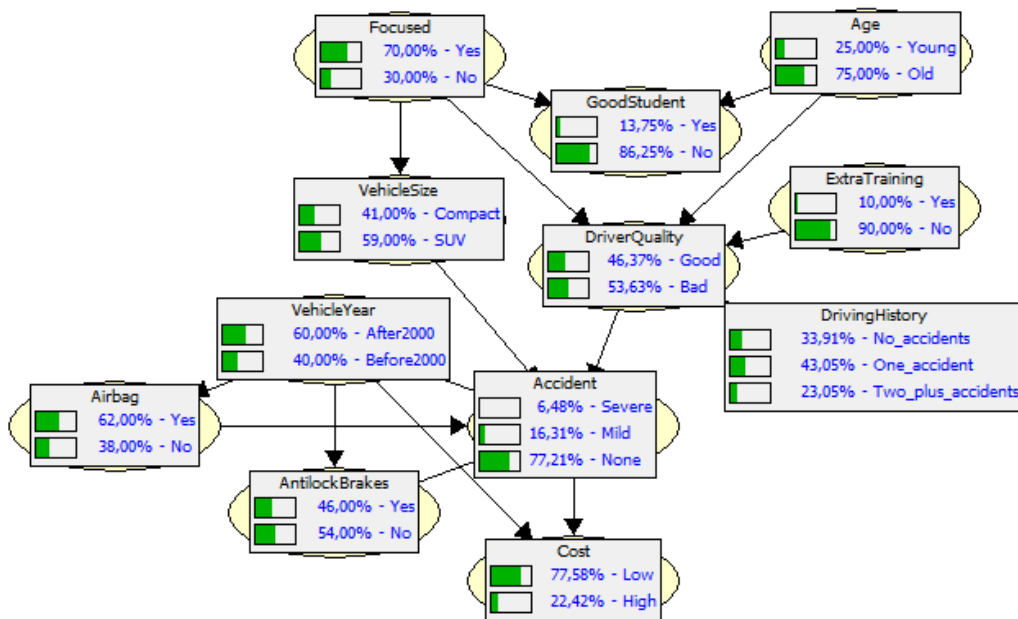
It is convenient to summarize the enumerated steps in table 1. The potential benefit of Variable Elimination for the query  $P(X_{12})$  is that the arithmetic complexity is equal to  $2^4 3^3 + 12$  basic operations (sums and products). In contrast, doing a summation on all the variables in factorized joint distribution is required  $12 \cdot 2^{10} 3^2 - 2$  operations.

Table 1

**Variable Elimination for the query  $P(X_{12})$**

Step	Variable eliminated	CPDs and factors used	New factor
1	$X_4$	$P(X_5 X_1, X_2, X_4), P(X_4)$	$\tau_1(X_1, X_2, X_5)$
2	$X_6$	$P(X_6 X_5)$	$\tau_2(X_5)$
3	$X_2$	$P(X_3 X_1, X_2), P(X_2),$ $\tau_1(X_1, X_2, X_5)$	$\tau_3(X_5, X_3, X_1)$
4	$X_3$	$\tau_3(X_5, X_3, X_1)$	$\tau_4(X_5, X_1)$
5	$X_1$	$P(X_1), P(X_7 X_1), \tau_4(X_5, X_1)$	$\tau_5(X_7, X_5)$
6	$X_5$	$P(X_{11} X_{10}, X_9, X_8, X_7, X_5),$ $\tau_2(X_5), \tau_5(X_7, X_5)$	$\tau_6(X_{11}, X_{10}, X_9, X_8, X_7)$
7	$X_7$	$\tau_6(X_{11}, X_{10}, X_9, X_8, X_7)$	$\tau_7(X_{11}, X_{10}, X_9, X_8)$
8	$X_9$	$\tau_7(X_{11}, X_{10}, X_9, X_8),$ $P(X_9 X_8)$	$\tau_8(X_{11}, X_{10}, X_8)$
9	$X_{10}$	$\tau_8(X_{11}, X_{10}, X_8), P(X_{10} X_8)$	$\tau_9(X_{11}, X_8)$
10	$X_8$	$\tau_9(X_{11}, X_8), P(X_{12} X_{11}, X_8),$ $P(X_8)$	$\tau_{10}(X_{12}, X_{11})$
11	$X_{11}$	$\tau_{10}(X_{12}, X_{11})$	$\tau_{11}(X_{12})$





**Fig. 2. Marginal distributions of all variables in the Insurance network**

An alternative approach to exact inference is based on the idea of conditioning.

The basic idea of the Conditioning algorithm [5] is that we construct a network for each assignment  $x_j$  of the random variable  $X_j$  on which to condition. These networks have identical structures, but different parameters. We run the Sum-Product Variable Elimination algorithm in each of them to obtain a probability  $P(X_i, X_j = x_j)$  for the query variable  $X_i$  reduced on the conditioning variable  $X_j$ . We add up these probabilities to obtain  $P(X_i)$ :

$$P(X_i) = \sum_{x_j} P(X_i, X_j = x_j) = \sum_{x_j} P(X_i | X_j = x_j) P(X_j = x_j).$$

We apply this approach to compute the marginal distribution of query variable  $Cost$ , where joint distribution  $P(X_1, \dots, X_{12})$  is factorized on CPDs in the following way:

$$\begin{aligned} P(X_{12}) &= \sum_{X_{11}} \sum_{X_1, \dots, X_{10}} P(X_1, \dots, X_{12}) = \\ &= \sum_{X_{11}} \sum_{X_1, \dots, X_{10}} P(X_{12} | X_{11}, X_8) P(X_{11} | X_{10}, X_9, X_8, X_7, X_5) P(X_{10} | X_8) P(X_9 | X_8) P(X_8) \cdot \\ &\quad \cdot P(X_7 | X_1) P(X_6 | X_5) P(X_5 | X_1, X_2, X_4) P(X_4) P(X_3 | X_2, X_1) P(X_2) P(X_1). \end{aligned}$$

Then summations can be transformed into the following way:

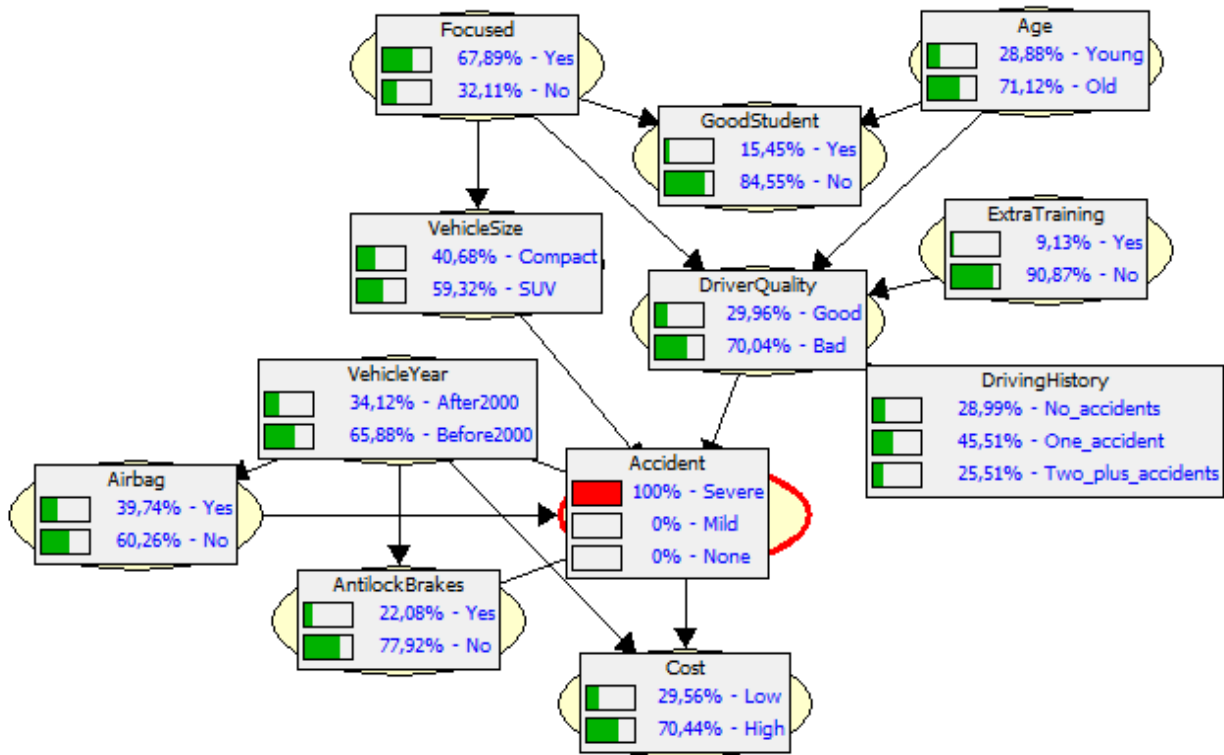
$$\begin{aligned} P(X_{12}) &= \sum_{X_{11}} \sum_{X_8} P(X_{12} | X_{11}, X_8) P(X_8) \sum_{X_{10}} P(X_{10} | X_8) \sum_{X_9} P(X_9 | X_8) \cdot \\ &\quad \cdot \sum_{X_7} \sum_{X_5} P(X_{11} | X_{10}, X_9, X_8, X_7, X_5) \sum_{X_1} P(X_1) P(X_7 | X_1) \cdot \\ &\quad \cdot \sum_{X_3} \sum_{X_2} P(X_3 | X_2, X_1) P(X_2) \sum_{X_6} P(X_6 | X_5) \sum_{X_4} P(X_5 | X_4, X_2, X_1) P(X_4). \end{aligned}$$

Then summations are transformed for each assignment of  $X_{11}$  :

$$\begin{aligned}
 P(X_{12}, X_{11} = x_{11}) = & \sum_{X_8} P(X_{12} | X_{11} = x_{11}, X_8) P(X_8) \sum_{X_{10}} P(X_{10} | X_8) \sum_{X_9} P(X_9 | X_8) \cdot \\
 & \cdot \sum_{X_7} \sum_{X_5} P(X_{11} = x_{11} | X_{10}, X_9, X_8, X_7, X_5) \sum_{X_1} P(X_1) P(X_7 | X_1) \cdot \\
 & \cdot \sum_{X_3} \sum_{X_2} P(X_3 | X_2, X_1) P(X_2) \sum_{X_6} P(X_6 | X_5) \sum_{X_4} P(X_5 | X_4, X_2, X_1) P(X_4).
 \end{aligned}$$

The number of operations performed by the conditioning is no less than the number of operations performed by Variable Elimination. But conditioning is undemanding of essential amounts of memory. We need to perform exact inference on the conditioned network several times, once for each assignment, and amass the results. Conditioning doesn't store intermediate results in memory and it costs additional time for repetition of computations. So, the trade-off of conditioning versus variable elimination is a time-space trade-off.

To improve the performance of conditioning, we can alternate elimination and condition. We conduct the internal summation on all random variables preceding the conditioning variable  $X_{11}$  in the above-mentioned expression. We reduce the network on conditioning variable  $X_{11}$  afterwards and receive a set of simplified networks. In each network, we use variable elimination on the rest of the random variables and aggregate the results from different networks.



**Fig. 3. The posterior probability distributions for all of the non-evidence variables when the evidence variable *Accident* is *Severe***

### MAP Inference in Insurance Network.

The MAP problem aims to find an assignment to the set of variables whose joint probability is the highest possible. The basic idea of the Max-Product Variable Elimination algorithm [5] is that we maximize random variables one at a time. When we maximize any random variable, we multiply all the CPDs that mention that variable, generating a product factor. We maximize the variable from this factor, generating a new factor that we enter into a set of CPDs to be dealt with.

We apply this approach to compute the most likely assignment to the entire network variables:

$$\begin{aligned} & \max_{X_1, \dots, X_{12}} P(X_1, \dots, X_{12}) = \\ & = \max_{X_1, \dots, X_{12}} P(X_{12} | X_{11}, X_8) P(X_{11} | X_{10}, X_9, X_8, X_7, X_5) P(X_{10} | X_8) P(X_9 | X_8) \cdot \\ & \cdot P(X_8) P(X_7 | X_1) P(X_6 | X_5) P(X_5 | X_1, X_2, X_4) P(X_4) P(X_3 | X_2, X_1) P(X_2) P(X_1). \end{aligned}$$

Consider eliminating nodes in the order

$$Z = \{X_4, X_6, X_2, X_3, X_1, X_5, X_7, X_9, X_{10}, X_8, X_{11}, X_{12}\},$$

which was found according to the greedy search for constructing an elimination ordering [5]. We have no query variable so that all variables are eliminated.

We can transform this expression into the following way:

$$\begin{aligned} & \max_{X_1, \dots, X_{12}} P(X_1, \dots, X_{12}) = \\ & = \max_{X_{12}, X_{11}, X_8} P(X_{12} | X_{11}, X_8) P(X_8) \max_{X_{10}} P(X_{10} | X_8) \max_{X_9} P(X_9 | X_8) \\ & \max_{X_7} \max_{X_5} P(X_{11} | X_{10}, X_9, X_8, X_7, X_5) \max_{X_1} P(X_1) P(X_7 | X_1) \\ & \max_{X_3} \max_{X_2} P(X_3 | X_2, X_1) P(X_2) \max_{X_6} P(X_6 | X_5) \max_{X_4} P(X_5 | X_4, X_2, X_1) P(X_4). \end{aligned}$$

Firstly, we compute the factors by eliminating variable  $X_4$ :

$$\begin{aligned} \psi_1(X_1, X_2, X_4, X_5) &= P(X_5 | X_1, X_2, X_4) P(X_4), \\ \tau_1(X_1, X_2, X_5) &= \max_{X_4} \psi_1(X_1, X_2, X_4, X_5). \end{aligned}$$

Secondly, we compute the factors by eliminating variable  $X_6$ :

$$\begin{aligned} \psi_2(X_6, X_5) &= P(X_6 | X_5), \\ \tau_2(X_5) &= \max_{X_6} \psi_2(X_6, X_5). \end{aligned}$$

Thirdly, we compute the factors by eliminating variable  $X_2$ :

$$\begin{aligned} \psi_3(X_5, X_3, X_2, X_1) &= P(X_3 | X_1, X_2) P(X_2) \tau_1(X_1, X_2, X_5), \\ \tau_3(X_5, X_3, X_1) &= \max_{X_2} \psi_3(X_5, X_3, X_2, X_1). \end{aligned}$$

Fourthly, we compute the factors by eliminating variable  $X_3$ :

$$\begin{aligned} \psi_4(X_5, X_3, X_1) &= \tau_3(X_5, X_3, X_1), \\ \tau_4(X_5, X_1) &= \max_{X_3} \psi_4(X_5, X_3, X_1). \end{aligned}$$

Fifthly, we compute the factors by eliminating variable  $X_1$ :

$$\begin{aligned}\psi_5(X_7, X_5, X_1) &= P(X_1)P(X_7|X_1)\tau_4(X_5, X_1), \\ \tau_5(X_7, X_5) &= \max_{X_1} \psi_5(X_7, X_5, X_1).\end{aligned}$$

Sixthly, we compute the factors by eliminating variable  $X_5$ :

$$\begin{aligned}\psi_6(X_{11}, X_{10}, X_9, X_8, X_7, X_5) &= P(X_{11}|X_{10}, X_9, X_8, X_7, X_5)\tau_2(X_5)\tau_5(X_7, X_5), \\ \tau_6(X_{11}, X_{10}, X_9, X_8, X_7) &= \max_{X_5} \psi_6(X_{11}, X_{10}, X_9, X_8, X_7, X_5).\end{aligned}$$

Seventhly, we compute the factors by eliminating variable  $X_7$ :

$$\begin{aligned}\psi_7(X_{11}, X_{10}, X_9, X_8, X_7) &= \tau_6(X_{11}, X_{10}, X_9, X_8, X_7), \\ \tau_7(X_{11}, X_{10}, X_9, X_8) &= \max_{X_7} \psi_7(X_{11}, X_{10}, X_9, X_8).\end{aligned}$$

Eighthly, we compute the factors by eliminating variable  $X_9$ :

$$\begin{aligned}\psi_8(X_{11}, X_{10}, X_9, X_8) &= \tau_7(X_{11}, X_{10}, X_9, X_8)P(X_9|X_8), \\ \tau_8(X_{11}, X_{10}, X_8) &= \max_{X_9} \psi_8(X_{11}, X_{10}, X_9, X_8).\end{aligned}$$

Ninthly, we compute the factors by eliminating variable  $X_{10}$ :

$$\begin{aligned}\psi_9(X_{11}, X_{10}, X_8) &= \tau_8(X_{11}, X_{10}, X_8)P(X_{10}|X_8), \\ \tau_9(X_{11}, X_8) &= \max_{X_{10}} \psi_9(X_{11}, X_{10}, X_8).\end{aligned}$$

Tenthly, we compute the factors by eliminating variable  $X_8$ :

$$\begin{aligned}\psi_{10}(X_{12}, X_{11}, X_8) &= \tau_9(X_{11}, X_8)P(X_{12}|X_{11}, X_8)P(X_8), \\ \tau_{10}(X_{12}, X_{11}) &= \max_{X_8} \psi_{10}(X_{12}, X_{11}, X_8).\end{aligned}$$

Eleventh, we compute the factors by eliminating variable  $X_{11}$ :

$$\begin{aligned}\psi_{11}(X_{12}, X_{11}) &= \tau_{10}(X_{12}, X_{11}), \\ \tau_{11}(X_{12}) &= \max_{X_{11}} \psi_{11}(X_{12}, X_{11}).\end{aligned}$$

Finally, we compute the factors by eliminating variable  $X_{12}$ :

$$\begin{aligned}\psi_{12}(X_{12}) &= \tau_{11}(X_{12}), \\ \max_{X_1, \dots, X_{12}} P(X_1, \dots, X_{12}) &= \max_{X_{12}} \psi_{12}(X_{12}).\end{aligned}$$

The computations generate the factors shown in table 2.

Now we consider the Traceback-MAP algorithm, which allows us to identify the MAP assignment once the variable elimination process is complete:

$$X_1^*, \dots, X_{12}^* = \arg \max_{X_1, \dots, X_{12}} P(X_1, \dots, X_{12}).$$

We begin by computing  $X_{12}^* = \arg \max_{X_{12}} \psi_{12}(X_{12})$ . The value  $X_{12}^*$  is the value of  $X_{12}$

in the most likely assignment to all the variables  $X_1, \dots, X_{12}$ . Given  $X_{12}^*$ , we can compute  $X_{11}^* = \arg \max_{X_{11}} \psi_{11}(X_{12}^*, X_{11})$ . The value  $X_{11}^*$  is the value of  $X_{11}$  in the most

likely complete assignment  $X_1, \dots, X_{12}$ . We use the same procedure for the remaining variables. So

$$\begin{aligned}
 X_8^* &= \arg \max_{X_8} \psi_{10}(X_{12}^*, X_{11}^*, X_8), & X_{10}^* &= \arg \max_{X_{10}} \psi_9(X_{11}^*, X_{10}, X_8^*), \\
 X_9^* &= \arg \max_{X_9} \psi_8(X_{11}^*, X_{10}^*, X_9, X_8^*), & X_7^* &= \arg \max_{X_7} \psi_7(X_{11}^*, X_{10}^*, X_9^*, X_8^*, X_7), \\
 X_5^* &= \arg \max_{X_5} \psi_6(X_{11}^*, X_{10}^*, X_9^*, X_8^*, X_7^*, X_5), & X_1^* &= \arg \max_{X_1} \psi_5(X_7^*, X_5^*, X_1), \\
 X_3^* &= \arg \max_{X_3} \psi_4(X_5^*, X_3, X_1^*), & X_2^* &= \arg \max_{X_2} \psi_3(X_5^*, X_3^*, X_2, X_1^*), \\
 X_6^* &= \arg \max_{X_6} \psi_2(X_6, X_5^*), & X_4^* &= \arg \max_{X_4} \psi_1(X_1^*, X_2^*, X_4, X_5^*).
 \end{aligned}$$

Table 2

**Max-Product Variable Elimination**

Step	Variable eliminated	CPDs and factors used	New factor
1	$X_4$	$P(X_5 X_1, X_2, X_4), P(X_4)$	$\tau_1(X_1, X_2, X_5)$
2	$X_6$	$P(X_6 X_5)$	$\tau_2(X_5)$
3	$X_2$	$P(X_3 X_1, X_2), P(X_2),$ $\tau_1(X_1, X_2, X_5)$	$\tau_3(X_5, X_3, X_1)$
4	$X_3$	$\tau_3(X_5, X_3, X_1)$	$\tau_4(X_5, X_1)$
5	$X_1$	$P(X_1), P(X_7 X_1), \tau_4(X_5, X_1)$	$\tau_5(X_7, X_5)$
6	$X_5$	$P(X_{11} X_{10}, X_9, X_8, X_7, X_5),$ $\tau_2(X_5), \tau_5(X_7, X_5)$	$\tau_6(X_{11}, X_{10}, X_9, X_8, X_7)$
7	$X_7$	$\tau_6(X_{11}, X_{10}, X_9, X_8, X_7)$	$\tau_7(X_{11}, X_{10}, X_9, X_8)$
8	$X_9$	$\tau_7(X_{11}, X_{10}, X_9, X_8),$ $P(X_9 X_8)$	$\tau_8(X_{11}, X_{10}, X_8)$
9	$X_{10}$	$\tau_8(X_{11}, X_{10}, X_8), P(X_{10} X_8)$	$\tau_9(X_{11}, X_8)$
10	$X_8$	$\tau_9(X_{11}, X_8), P(X_{12} X_{11}, X_8),$ $P(X_8)$	$\tau_{10}(X_{12}, X_{11})$
11	$X_{11}$	$\tau_{10}(X_{12}, X_{11})$	$\tau_{11}(X_{12})$
12	$X_{12}$	$\tau_{11}(X_{12})$	$\max_{X_1, \dots, X_{12}} P(X_1, \dots, X_{12})$

The most likely assignment  $X_1^*, \dots, X_{12}^*$  is represented in fig. 3.

Max-Product Variable Elimination algorithm is generally applied to compute the most likely assignment to *all of the non-evidence variables* to fill in unknown information.

For instance, if we let  $W = \{X_1, X_2, X_4, X_5, X_7, X_8, X_9, X_{10}, X_{11}\}$  be the subset of query variables in the network and  $E = \{X_3, X_6, X_{12}\}$  be the subset of the random variable in our model, then  $MAP(W|E = e)$  is represented in fig. 4.

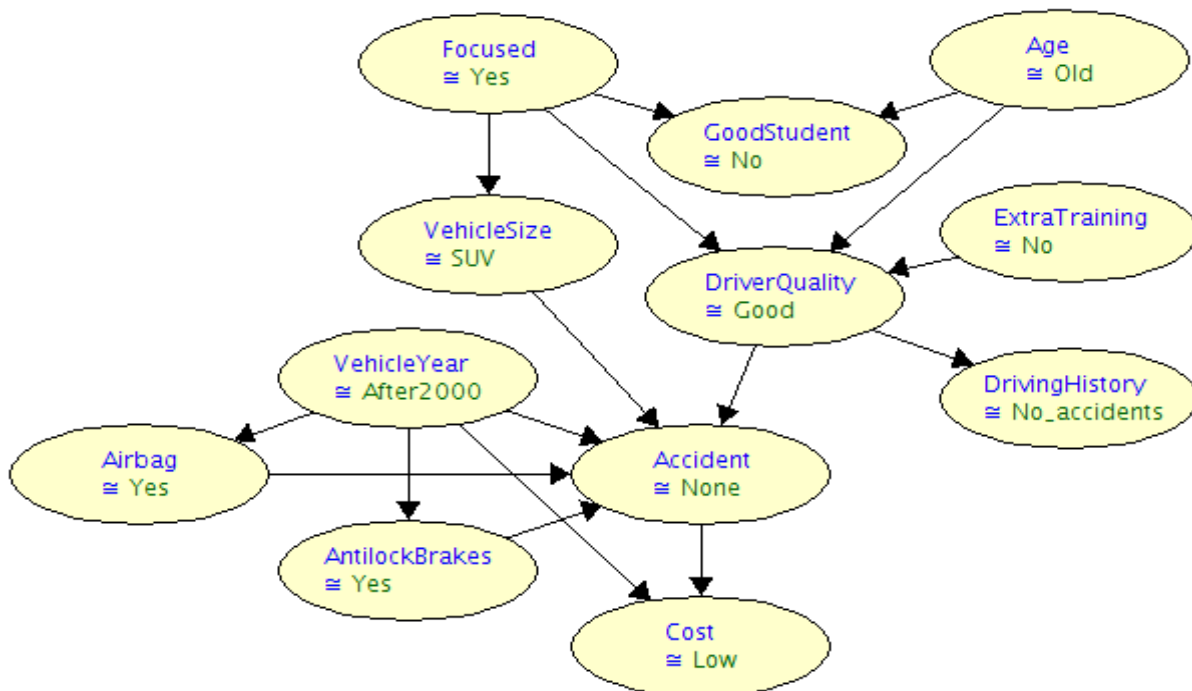


Fig. 4. The most likely assignment to all variables in the Insurance network

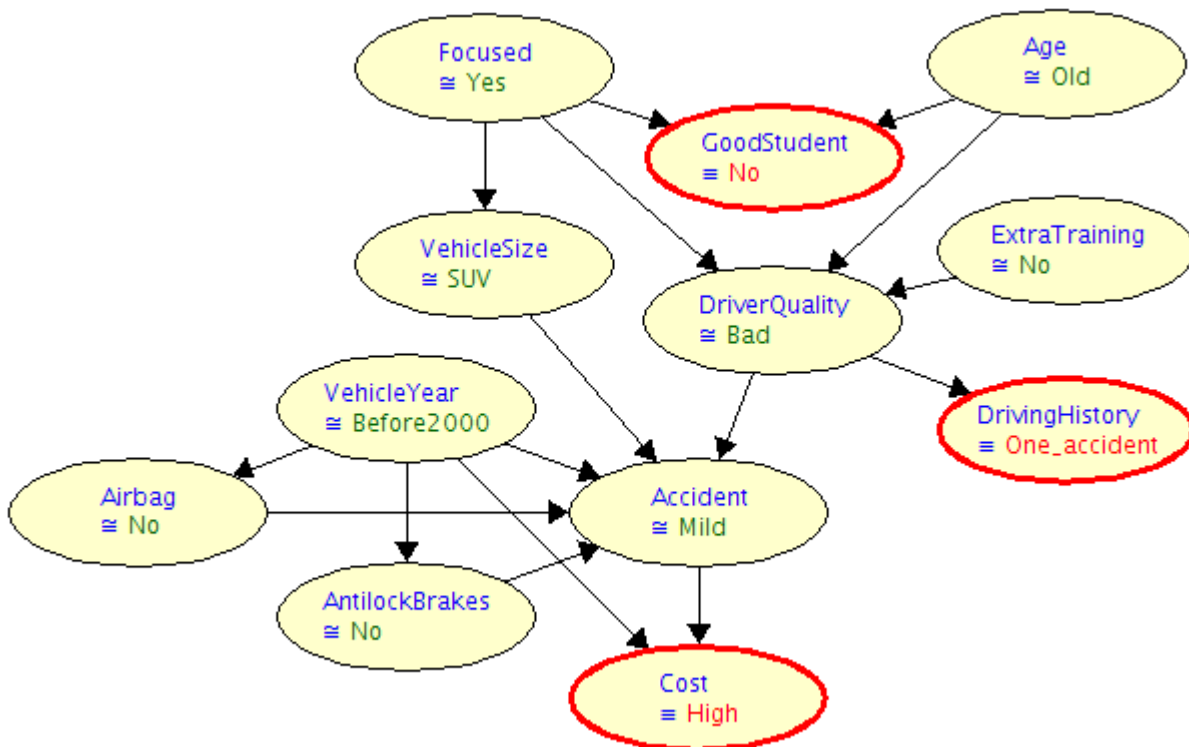


Fig. 5. The most likely assignment to the variables *Focused*, *Age*, *Vehicle Size*, *Driving Quality*, *Extra Training*, *Vehicle Year*, *Airbag*, *Antilock Brakes*, *Accident* given evidence *Good Student*, *Driving History*, *Cost*

### Marginal MAP Inference in Insurance Network.

The marginal MAP query contains elements of both a probability query and a MAP query. The goal of the marginal MAP problem is to find the assignment to the set of variables whose conditional marginal probability is maximal:

$$MAP(Y|E=e) = \arg \max_Y \sum_Z P(Y, Z|E=e).$$

in other words, to find an assignment that maximizes the posterior probability of query variables, marginalizing over the rest of ones.

The basic idea of the Max-Sum-Product Variable Elimination algorithm [5] is that we eliminate the random variables  $Z$  by summing them out exploiting Sum-Product Variable Elimination and eliminate the random variables  $Y$  by maximizing them out utilizing Max-Product Variable Elimination. We apply this approach to compute the most likely assignment of query variables *Focused, Age, Extra Training, Vehicle Year*:

$$\max_{X_1, X_2, X_4, X_8} \sum_{X_3, X_5, X_6, X_7, X_9, X_{10}, X_{11}, X_{12}} P(X_1, \dots, X_{12}).$$

Consider eliminating nodes in the order

$$Z = \{X_3, X_6, X_{12}, X_{11}, X_{10}, X_9, X_7, X_5, X_8, X_4, X_2, X_1\}.$$

We can transform this summation into the following way:

$$\begin{aligned} & \max_{X_1, X_2, X_4, X_8} \sum_{X_5} P(X_5|X_4, X_2, X_1) \sum_{X_7} P(X_7|X_1) \sum_{X_9} P(X_9|X_8) \sum_{X_{10}} P(X_{10}|X_8) \cdot \\ & \cdot \sum_{X_{11}} P(X_{11}|X_{10}, X_9, X_8, X_7, X_5) \sum_{X_{12}} P(X_{12}|X_{11}, X_8) \sum_{X_6} P(X_6|X_5) \sum_{X_3} P(X_3|X_1, X_2). \end{aligned}$$

Firstly, we compute the factors by eliminating variable  $X_3$ :

$$\begin{aligned} \psi_1(X_1, X_2, X_3) &= P(X_3|X_1, X_2), \\ \tau_1(X_1, X_2) &= \sum_{X_3} \psi_1(X_1, X_2, X_3). \end{aligned}$$

Secondly, we compute the factors by eliminating variable  $X_6$ :

$$\begin{aligned} \psi_2(X_5, X_6) &= P(X_6|X_5), \\ \tau_2(X_5) &= \sum_{X_6} \psi_2(X_5, X_6). \end{aligned}$$

Thirdly, we compute the factors by eliminating variable  $X_{12}$ :

$$\begin{aligned} \psi_3(X_8, X_{11}, X_{12}) &= P(X_{12}|X_{11}, X_8), \\ \tau_3(X_8, X_{11}) &= \sum_{X_{12}} \psi_3(X_8, X_{11}, X_{12}). \end{aligned}$$

Fourthly, we compute the factors by eliminating variable  $X_{11}$ :

$$\begin{aligned} \psi_4(X_5, X_7, X_8, X_9, X_{10}, X_{11}) &= P(X_{11}|X_{10}, X_9, X_8, X_7, X_5) \tau_3(X_8, X_{11}), \\ \tau_4(X_5, X_7, X_8, X_9, X_{10}) &= \sum_{X_{11}} \psi_4(X_5, X_7, X_8, X_9, X_{10}, X_{11}). \end{aligned}$$

Fifthly, we compute the factors by eliminating variable  $X_{10}$ :

$$\psi_5(X_5, X_7, X_8, X_9, X_{10}) = \tau_4(X_5, X_7, X_8, X_9, X_{10}) P(X_{10}|X_8),$$

$$\tau_5(X_5, X_7, X_8, X_9) = \sum_{X_{10}} \psi_5(X_5, X_7, X_8, X_9, X_{10}).$$

Sixthly, we compute the factors by eliminating variable  $X_9$ :

$$\begin{aligned} \psi_6(X_5, X_7, X_8, X_9) &= \tau_5(X_5, X_7, X_8, X_9)P(X_9|X_8), \\ \tau_6(X_5, X_7, X_8) &= \sum_{X_9} \psi_6(X_5, X_7, X_8, X_9). \end{aligned}$$

Seventhly, we compute the factors by eliminating variable  $X_7$ :

$$\begin{aligned} \psi_7(X_1, X_5, X_7, X_8) &= \tau_6(X_5, X_7, X_8)P(X_7|X_1), \\ \tau_7(X_1, X_5, X_8) &= \sum_{X_7} \psi_7(X_1, X_5, X_7, X_8). \end{aligned}$$

Eighthly, we compute the factors by eliminating variable  $X_5$ :

$$\begin{aligned} \psi_8(X_1, X_2, X_4, X_5, X_8) &= \tau_2(X_5)\tau_7(X_1, X_5, X_8)P(X_5|X_4, X_2, X_1), \\ \tau_8(X_1, X_2, X_4, X_8) &= \sum_{X_5} \psi_8(X_1, X_2, X_4, X_5, X_8). \end{aligned}$$

Ninthly, we compute the factors by eliminating variable  $X_8$ :

$$\begin{aligned} \psi_9(X_1, X_2, X_4, X_8) &= \tau_8(X_1, X_2, X_4, X_8)P(X_8), \\ \tau_9(X_1, X_2, X_4) &= \max_{X_8} \psi_9(X_1, X_2, X_4, X_8). \end{aligned}$$

Tenthly, we compute the factors by eliminating variable  $X_4$ :

$$\begin{aligned} \psi_{10}(X_1, X_2, X_4) &= \tau_9(X_1, X_2, X_4)P(X_4), \\ \tau_{10}(X_1, X_2) &= \max_{X_4} \psi_{10}(X_1, X_2, X_4). \end{aligned}$$

Eleventh, we compute the factors by eliminating variable  $X_2$ :

$$\begin{aligned} \psi_{11}(X_1, X_2) &= \tau_{10}(X_1, X_2)\tau_9(X_1, X_2)P(X_2), \\ \tau_{11}(X_1) &= \max_{X_2} \psi_{11}(X_1, X_2). \end{aligned}$$

Finally, we compute the factors by eliminating variable  $X_1$ :

$$\psi_{12}(X_1) = \tau_{11}(X_1)P(X_1), \quad \tau_{12} = \max_{X_1} \psi_{12}(X_1).$$

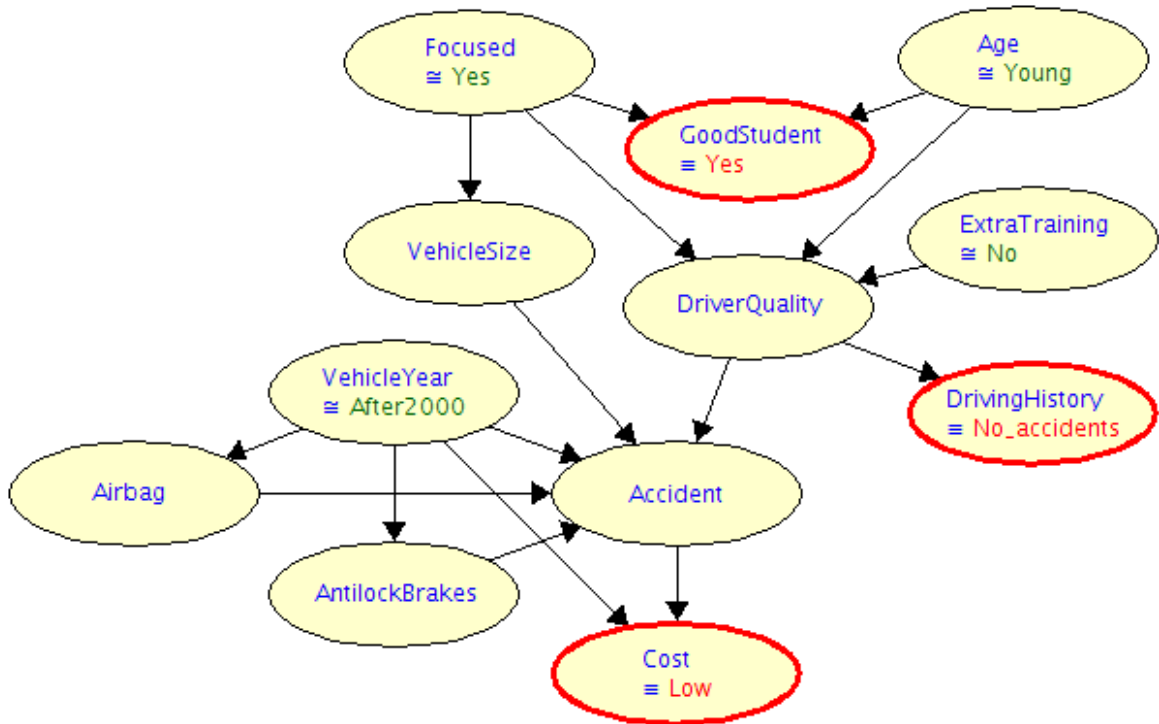
The computations generate the factors shown in table 3. We identify the marginal MAP assignment once the variable elimination process is completed:

$$\begin{aligned} X_1^* &= \arg \max_{X_1} \psi_{12}(X_1), \quad X_2^* = \arg \max_{X_2} \psi_{11}(X_1^*, X_2), \\ X_4^* &= \arg \max_{X_4} \psi_{10}(X_1^*, X_2^*, X_4), \quad X_8^* = \arg \max_{X_8} \psi_9(X_1^*, X_2^*, X_4^*, X_8). \end{aligned}$$

Max-Sum-Product Variable Elimination algorithm is typically applied to compute the most likely assignment to *query variables* given the evidence to obtain unknown knowledge.

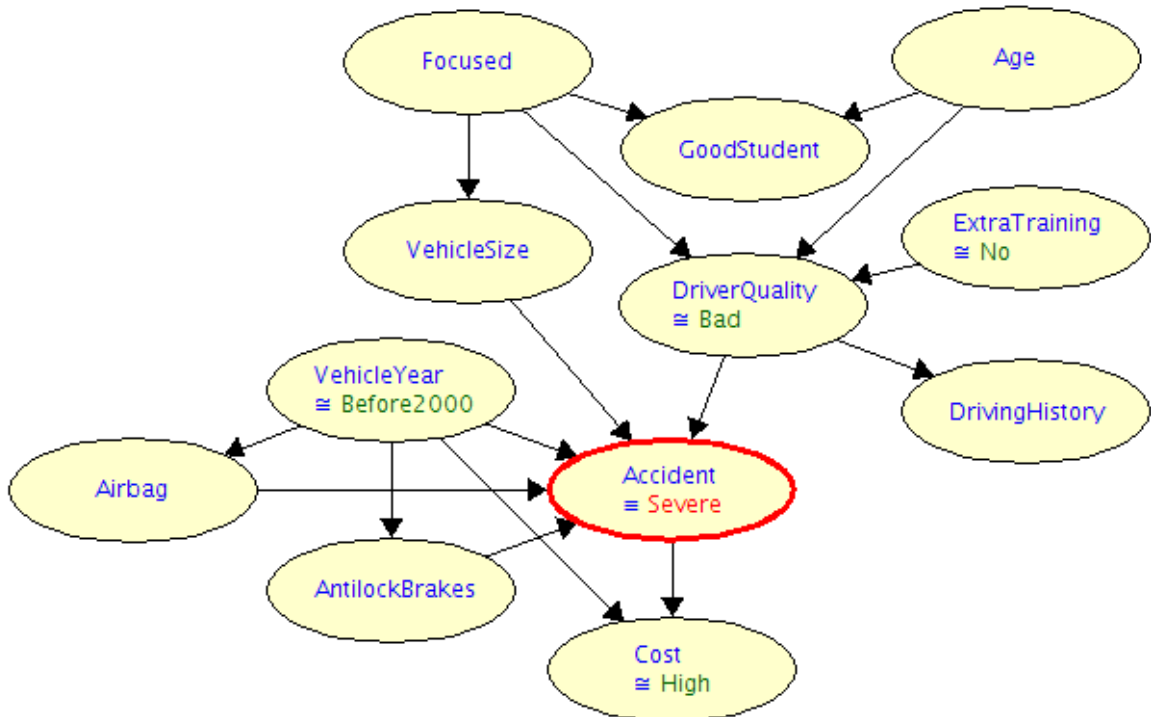
For instance, if we let  $Y = \{X_1, X_2, X_4, X_8\}$  be subsets of query variables in the network,  $Z = \{X_5, X_7, X_9, X_{10}, X_{11}\}$  be a subset of random variables that are neither to be maximized, nor to be observed, but marginalized given the evidence  $E = \{X_3, X_6, X_{12}\}$  in our model, then  $MAP(Y|E=e)$  is represented in fig. 5.





**Fig. 6.** The most likely assignment to the variables *Focused*, *Age*, *Extra Training*, *Vehicle Size* given evidence *Good Student*, *Driving History*, *Cost*

In addition, if we let  $Y = \{X_4, X_5, X_8, X_{12}\}$ ,  $Z = \{X_1, X_2, X_3, X_6, X_7, X_9, X_{10}\}$  and  $E = \{X_{11}\}$  be subsets in our model, then  $MAP(Y|E = e)$  is represented in figure 6.



**Fig. 7.** The most likely assignment to the variables *Extra Training*, *Driving Quality*, *Vehicle Year*, *Cost* given evidence *Accident*

**Max-Sum-Product Variable Elimination**

Step	Variable eliminated	CPDs and factors used	New factor
1	$X_3$	$P(X_3 X_1, X_2)$	$\tau_1(X_1, X_2)$
2	$X_6$	$P(X_6 X_5)$	$\tau_2(X_5)$
3	$X_{12}$	$P(X_{12} X_{11}, X_8)$	$\tau_3(X_8, X_{11})$
4	$X_{11}$	$P(X_{11} X_{10}, X_9, X_8, X_7, X_5),$ $\tau_3(X_8, X_{11})$	$\tau_4(X_5, X_7, X_8, X_9, X_{10})$
5	$X_{10}$	$\tau_4(X_5, X_7, X_8, X_9, X_{10}), P(X_{10} X_8)$	$\tau_5(X_5, X_7, X_8, X_9)$
6	$X_9$	$\tau_5(X_5, X_7, X_8, X_9), P(X_9 X_8)$	$\tau_6(X_5, X_7, X_8)$
7	$X_7$	$\tau_6(X_5, X_7, X_8), P(X_7 X_1)$	$\tau_7(X_1, X_5, X_8)$
8	$X_5$	$\tau_2(X_5), \tau_7(X_1, X_5, X_8),$ $P(X_5 X_4, X_2, X_1)$	$\tau_8(X_1, X_2, X_4, X_8)$
9	$X_8$	$\tau_8(X_1, X_2, X_4, X_8), P(X_8)$	$\tau_9(X_1, X_2, X_4)$
10	$X_4$	$\tau_9(X_1, X_2, X_4), P(X_4)$	$\tau_{10}(X_1, X_2)$
11	$X_2$	$\tau_1(X_1, X_2), \tau_{10}(X_1, X_2), P(X_2)$	$\tau_{11}(X_1)$
12	$X_1$	$\tau_{11}(X_1), P(X_1)$	$\tau_{12}$

**Conclusion.** We considered the technique to solve the prediction problem of reparation of the material damage caused by road traffic accidents and showed how to find variable assignments that have maximal probability under given evidence. We examined one class of exact algorithms for this problem based on variable elimination. Sum-Product Variable Elimination can be used to find *the marginal distribution of query variables* if the model includes a set of variables that are to be marginalized. Conditioning can be used to find *the posterior distribution of the query variable* if the model includes a set of variables that are to be marginalized given the evidence. Max-Product Variable Elimination can be applied to find *an assignment that maximizes the posterior probability of query variables* if the model contains query variables that are to be maximized given the evidence. Max-Sum-Product Variable Elimination can be used to find *an assignment that maximizes the posterior probability of query variables, marginalizing over the rest of the ones* if the model includes variables that are neither to be maximized, nor to be observed, but marginalized given the evidence.

## References

1. **Pearl, J.** Probabilistic Inference in Intelligent Systems: Networks of Plausible Inference. [Text] – Morgan Kaufmann, 1988.
2. **Pearl, J.** Causality: Models, Reasoning and Inference [Text]. Cambridge University Press, 2000.
3. **Korb, K.B. Nicholson, A.E.** Bayesian artificial intelligence [Text]. Chapman & Hall/CRC Press LLC, 2004.
4. **Bishop, C.M.** Pattern Recognition and Machine Learning [Text]. Springer, 2007.
5. **Koller, D., Friedman, N.** Probabilistic Graphical Models: Principles and Techniques. [Text] The MIT Press, 2009.
6. **Darwiche, A.** Modeling and Reasoning with Bayesian networks [Text]. Cambridge University Press, 2009.
7. **Barber, D.** Bayesian Reasoning and Machine Learning [Text]. Cambridge University Press, 2012.
8. **Murphy, K.P.** Machine Learning: A Probabilistic Perspective [Text]. The MIT Press, 2012.
9. **Koller, D.** Probabilistic Graphical Models 1: Representation // Stanford University. URL: <https://www.coursera.org/learn/probabilistic-graphical-models>. Online. Accessed 1-Sep-2020.
10. **Koller, D.** Probabilistic Graphical Models 2: Inference // Stanford University. URL: <https://www.coursera.org/learn/probabilistic-graphical-models>. Online. Accessed 1-Sep-2020.
11. **Koller, D.** Probabilistic Graphical Models 3: Learning // Stanford University. URL: <https://www.coursera.org/learn/probabilistic-graphical-models>. Online. Accessed 1-Sep-2020.
12. **Binder, J., Koller, D., Russell, S., Kanazawa, K.** Adaptive Probabilistic Networks with Hidden Variables [Text], Machine Learning, vol. 29, no. 2-3, pp. 213-244, 1997.
13. **Qazi, M., Fung, G.M., Meissner, K.J., Fontes, E.R.** An Insurance Recommendation System Using Bayesian Networks, Proceedings of the Eleventh ACM Conference on Recommender Systems, pp. 274–278, 2017.
14. **Xin, Zou, Wen, Long Yue.** A Bayesian Network Approach to Causation Analysis of Road Accidents Using Netica [Text], Journal of Advanced Transportation, Vol. 2017, pp. 1-18, 2017.
15. **Zhu Xu, Yi Jiang, Fan Lin, Long Dai,** The Analysis and Prevent in Traffic Accidents Based on Bayesian Network [Text]`, Advanced Engineering Forum, Vol. 1, pp. 21-25, 2011.

*Received 17.09. 2020.*