

Interactive Learning of Scene Context Extractor Using Combination of Bayesian Network and Logic Network

Keum-Sung Hwang and Sung-Bae Cho

Dept. of Computer Science, Yonsei University,
Shinchon-dong, Seodaemun-ku,
Seoul 120-749, Korea
{yellowg, sbcho}@sclab.yonsei.ac.kr

Abstract. The vision-based scene understanding technique that infers scene-interpreting contexts from real-world vision data has to not only deal with various uncertain environments but also reflect user's requests. Especially, learnability is a hot issue for the system. In this paper, we adopt a probabilistic approach to overcome the uncertainty, and propose an interactive learning method using combination of Bayesian network and logic network to reflect user's requirements in real-time. The logic network works for supporting logical inference of Bayesian network. In the result of some learning experiments using interactive data, we have confirmed that the proposed interactive learning method is useful for scene context reasoning.

1 Introduction

Bayesian network (BN) is a useful tool for modeling causal judgment and inference processes [1], and it receives increasing attention in the vision-based scene understanding area, where it recognizes contexts by reasoning detected objects and features from vision data to understand scene. Bayesian network is also robust to real-world situations because the probabilistic approach manages well uncertain data and supports multiple directional inferences.

Scene understanding is the task of understanding a scene beyond single-object recognition. A scene understanding is determined by constructing a description of the scene in terms of concepts provided in a conceptual knowledge base. It still remains a difficult problem because of complexity and uncertainty of real-world.

In the vision-based scene understanding area, automatic learning is important because not only expert knowledge but also domain data collected in real-world are used for modeling inference model. However, it is not easy since the actual environment of a scene understanding agent has various uncertain data and causes continuous user's requests. In this paper, we propose an interactive learning method to adopt user's requests to inference model by learning Bayesian network with logic network (LN). Because Bayesian network training data and interactive data have different features as follows, we exploit the learning method of logic network. The characteristics of Bayesian network training data:

- Complete data: have values of all nodes.
- Large quantity: require a lot of data for training BN parameters. A node requires 100 data for each distinct instantiation of the parent set for 1% error range.
- Case data: each datum is a case of available states (possible values of node), so the gathered data set reflects probabilistic causalities.

The characteristics of interactive data:

- Incomplete data: have values of a few nodes.
- Small quantity: not easy to collect enough data for training BN parameters
- Logical data: beneficial to design structure.
- Changeable: imply user's opinion, so it is profitable to manage them specially without direct modification of BN parameters.

2 Related Works

Some probabilistic approaches are studied recently to solve the vision-based scene understanding problem as follows;

- T. M. Strat *et al.* (1991): Assumes that a target object is defined by several shape models and can be extracted some local features.
- A. Torralba *et al.* (2003): Recognizes scenes using Hidden Markov model from the visual feature vectors.
- B. Neumann *et al.* (2003): Researches for description logics and framework for high-level scene understanding and interpretation, modeling based on detected objects.
- M. Marengoni *et al.* (2003): Selects visual function sets automatically based on hierarchical BN on aerial picture recognition system - Ascender I.
- J. Luo, A. E. *et al.* (2005): Detects natural objects in outdoor scenes based on probabilistic spatial context[2].

However, the proposed probabilistic model demands an enormous amount of training data or expert's assistance. For this reason, they are difficult to adapt several interactive data effectively added by a user-feedback. There were researches for learning the Bayesian model. Adaptive Bayesian network using revised backward propagation method is researched [3], and Bayesian network refinement method that adapts Bayesian networks using minimum description length score metric is proposed [4].

- B.P.L. Lo *et al.* (2003): Adaptive BN: adapts BN using re-training technique - revised backward propagation[3].
- W. Lam (1998): Bayesian network refinement technique: using MDL score metric that minimizes distance between network structure and data-set[4].

As these methods require complete and sufficient amount of Bayesian network data, and it is not suitable to learn interactive data since they have small quantity and many missing data, and they do not concern the features of interactive data, so we propose a learning method with due consideration to them.

3 Interactive Learning of Bayesian Network Module Using Logic Network

In this section, we propose a BN+LN model (combination of Bayesian network and logic network) to expand or update inference model with collected interactive data. Figure 1 shows the proposed model. In the proposed method, a logic network plays a role of supporting inference on the posterior stage of Bayesian network. The computational result of logic network covers that of Bayesian network.

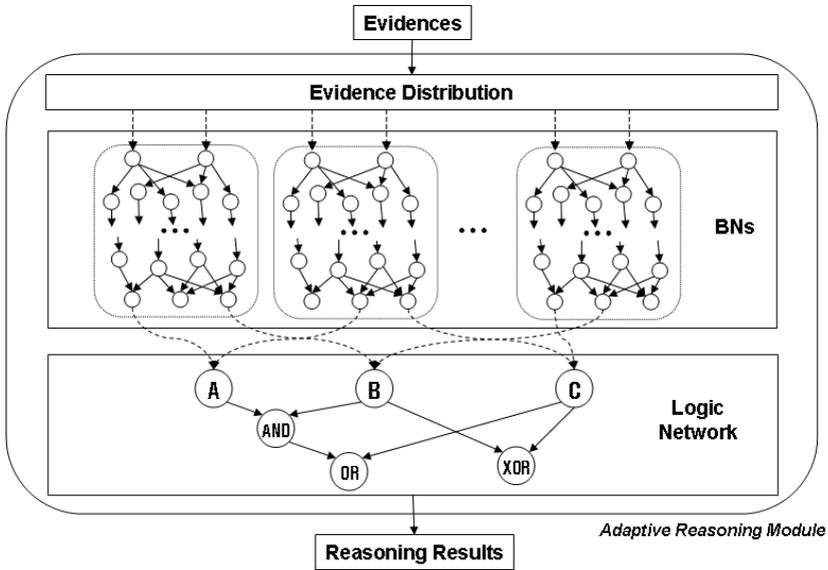


Fig. 1. The BN+LN model

3.1 Bayesian Network

A Bayesian network have a shape of DAG (directed acyclic graph) expressing the relations of nodes and describes a large probabilistic relations with CPTs (conditional probability tables) constrained by the structure. The belief value using the given evidence set E on Bayesian network, $Bel(h)$ is calculated by Bayes' Rule such as a formula (1).

$$Bel(h) = P(h|E) = \frac{P(E|h)P(h)}{P(E)} = \frac{P(E \cap h)}{P(E)}. \tag{1}$$

where h is the hypothesis of a node state. The probability set is computed by a chain Rule such as formula (2).

$$\begin{aligned} P(x_1, x_2, \dots, x_n) &= P(x_1)P(x_2|x_1) \dots \\ &= P(x_1)P(x_2|\pi_2) \dots P(x_n|\pi_n). \end{aligned} \tag{2}$$

where x_i is the i -th node and π_i is the parent of the node i .

3.2 Logic Network

A logic network is a structure for expression of input /outputs or Boolean computations of digital circuit. It has input/output nodes, internal nodes with a logical function and a directed acyclic graph structures that indicates data stream [5]. A node function is composed of multiple inputs and 1 output and a logic function. The internal functions of logic network are listed in Table 1.

Table 1. The internal node function list of logic network. Where v_i means a result value of logic calculation, which is a Boolean type, possible only *true* and *false*. n_i and s_i indicate node and state. *StateEq*(\cdot) and *IsInStateList*(\cdot) work for input value selection.

Function name	Description
<i>NOT</i> (v)	If ($v = false$) then <i>true</i> else <i>false</i>
<i>AND</i> (v_1, v_2, \dots, v_n)	If (every $v_i(1 \leq i \leq n) = true$) then <i>true</i> else <i>false</i>
<i>OR</i> (v_1, v_2, \dots, v_n)	If (any one of $v_i(1 \leq i \leq n) = true$) then <i>true</i> else <i>false</i>
<i>StateEq</i> (n_1)	If (the state of $n_1 = true$ OR the state of $n_1 = yes$) then <i>true</i> else <i>false</i>
<i>StateEq</i> (n_1, s)	If (the state of $n_1 = s$) then <i>true</i> else <i>false</i>
<i>IsInStateList</i> (n_1, n_2)	If (the state of $n_1 = true$) AND (the name of $n_1 \in$ the state list of n_2) then <i>true</i> else <i>false</i>

3.3 Interactive Learning Method

We propose a method that adapts logic network and combine it with Bayesian network to deal with the interactive data collected from the inference module application process. In the proposed method, a logic network works for supporting logical inference of Bayesian network. The detailed operations are given as follows:

- Interaction with user: Interacting with user, collecting new data/feedback from user. In this paper we uses only predefined sentences.
- Reasoning: Inferring contexts that user wants for. If the reasoning is not available the system requires the user to feedback.
- Causality extraction: Defining logical relations of variables. The structure of sentence decides logics by predefined rules.
- Variable extraction: If the variable that is not defined is detected it declares a new variable as evidence-variable or result-variables with its role.
- Adapt logic network: Expressing logical relation of variable as a network and adapt the previous network.
- Update module: Updating the adapted logic network.

4 Experiments

We have applied additional evidences given as interactive data to the proposed interactive learning method of LN+BN model.

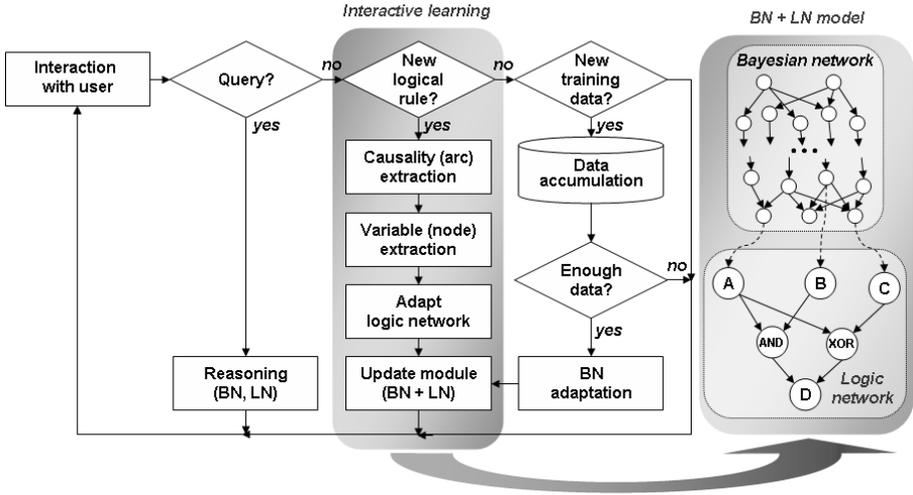


Fig. 2. The proposed interactive learning process

4.1 A Case Study

The first experiment is a case study to observe learning result of the proposed method. We defined a scene domain (11 places and 25 objects) and designed the place-recognition Bayesian network based on detected object (at the left side on figure 3) [6]. Then, we have added 8 logical interactive data for interactive learning of BN+LN model. Table 2 denotes a part of the interactive data used for learning and their extracted logic rules by the proposed interactive learning process. For the smooth experiments we have used predefined sentences and extraction rules. Figure 3 shows an experimental result of adapting given interactive data using the proposed interactive learning method. In the figure, it can be seen that some object nodes and logical inference functions are augmented. In fact, the logic relations of the logic network can be expressed by Bayesian network, but they do not require complex probability values, and a Bayesian network represents inefficiently the node that has many parents, for example "Indoor" node requires $4,096 (= 2^{12})$ CPVs (conditional probability values). The complexity of the number of CPVs is $O(k^N)$ and it is calculated as follows;

$$\prod_{i \in P \cup I} N_i, \quad (3)$$

where P is parents set, I is the self-node, N_k is the number of states of the node k .

4.2 Performance Test

To evaluate the performance of the proposed method, we experimented with interactive data that contain certain agreement/disagreement-evidences for place

Table 2. A part of the 8 interactive data used for learning BN+LN

Object	Interaction Data	Extracted Logic
Elevator	If it is narrow place, has door and no ground objects, and It is linked corridor, it is elevator.	place shape=narrow AND ground object=no AND door=yes AND linked=corridor → elevator. Needs more information about ground object.
Ground object	Air conditioner, garbage can, bookshelf, dresser, chair, lectern, partition, table and castor whiteboard are ground objects.	AirConditioner, garbageCan, bookshelf, dresser, chair, lectern, partition, table, castorWhiteboard → ground-Object
Hall	Hall is linked to corridor and has a door.	NOT (linked=corridor AND door=yes) → NOT hall
Corridor	Corridor is long and linked to an indoor place.	NOT (place shape=long AND linked=indoor) → NOT corridor
Lecture room	Lecture room has a lectern and has a wall whiteboard or a screen.	NOT ((wall whiteboard=yes OR screen=yes) AND lectern=yes) → NOT lectureRoom
Seat place	Seat-place includes chair, bench and sofa.	chair, bench, sofa → seatPlace

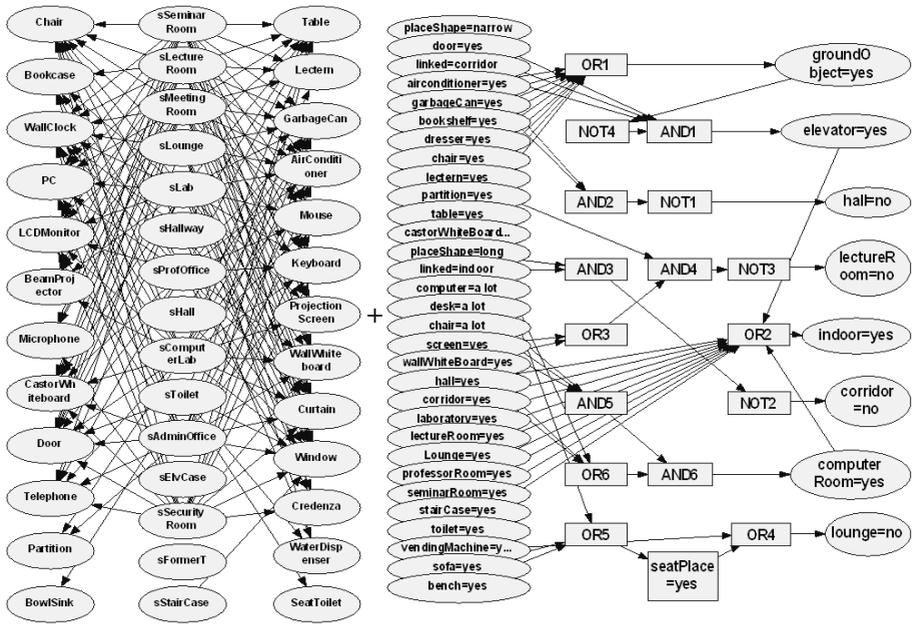


Fig. 3. The designed Bayesian network (left) and the learned logic network (right)

recognition. Where, an agreement-evidence indicates the evidence that is sufficient for an affirmation of fact, and a disagreement-evidence denotes the evidence that is sufficient for negation of a fact. Especially, a negation-logic is usually not used in Bayesian network, because it requires a complex Bayesian network model, while logic network requires a simple model. We experimented each case when k (=evidence size) evidences of 25 objects are discovered. We used totally 3 agreement-evidences and 79 disagreement-evidences for 11 places. When the evidence size is 4 we adopted 2,300 random cases since the number of evidence combination is too large.

Table 3 shows the results, in which the performance of the proposed method is better than the original Bayesian network at all cases except the case of 1-evidence because the interactively-learned logic rules support more accurate inference with relatively smaller evidences.

Table 3. The experimental results with agreement/disagreement-evidences. Abbreviations: Precision = $TP/(TP + FP)$, Recall = $TP/(TP + FN)$, TP =True positive, FP =false positive, FN =false negative, TN =true negative error.

Runs	25 (${}_{25}C_1$)		300 (${}_{25}C_2$)	
Evidence size	1		2	
Method	BN	LN+BN	BN	LN+BN
FP	72%	72%	48%	45%
TN	28%	28%	52%	55%
TP	100%	100%	97%	100%
FN	0%	0%	3%	0%
Precision	58%	58%	67%	69%
Recall	100%	100%	97%	100%
Runs	2,300 (${}_{25}C_3$)		random 2,300	
Evidence size	3		4	
Method	BN	LN+BN	BN	LN+BN
FP	33%	26%	17%	7%
TN	67%	74%	83%	93%
TP	93%	100%	90%	100%
FN	7%	0%	10%	0%
Precision	74%	80%	84%	94%
Recall	93%	100%	90%	100%

5 Concluding Remarks

We propose an interactive learning method using logic network to apply user's requests to Bayesian inference model. The experimental result shows that the proposed interactive learning method is useful for incremental scene context extraction by supporting addition of new context nodes and logical inference rules, so it causes performance improvement. The proposed method might be a good complement for interactive learnable Bayesian network.

In the future work, we would like to apply the proposed method to the more complex and practical applications. We can also compare the method to various Bayesian network adaptation techniques.

Acknowledgement. This research was supported by the Ministry of Information and Communication, Korea under the Information Technology Research Center support program supervised by the Institute of Information Technology Assessment, IITA-2005-(C1090-0501-0019).

References

1. K. B. Korb and A. E. Nicholson, *Bayesian Artificial Intelligence*, Chapman & Hall/CRC, 2003.
2. J. Luo, et al., "A Bayesian network-based framework for semantic image understanding," *Pattern Recognition*, vol. 38, pp. 919-934, 2005.
3. B.P.L. Lo et al., "Adaptive Bayesian networks for video processing," *Int Conf. on Image Processing*, pp.889-892, 2003.
4. Wai Lam, "Bayesian network refinement via machine learning approach," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 240-251, March 1998.
5. P.G. Bishop, "Estimating PLC logic program reliability," *Safety-critical Systems Symposium*, Birmingham, 17-19 Feb., 2004.
6. Y.-S. Song, S.-B. Cho, and I.-H. Suh, "Activity-object Bayesian networks for detecting occluded objects in uncertain indoor environment," *Lecture Notes in Computer Science*, vol. 3683, pp. 937-944, 2005.
7. B. Neumann, *A conceptual framework for high-level vision*, Bericht, FB Informatik, FBI-HH-B245/03, 2003.
8. A. Torralba, et al., "Context-based vision system for place and object recognition," *Int. Conf. Computer Vision*, pp. 273-280, 2003.