An Introduction to belief revision and knowledge representation with 2CNF

Meliza Contreras, Miguel Rodríguez, Pedro Bello, Bárbara M. González

Faculty of Computer Science, Benemérita Universidad Autónoma de Puebla
{mcontreras,mrodriguez,pbello}@cs.buap.mx,barbra.monique.gl@gmail.com

Abstract. The aim of this paper is to show an introduction to the calculation of belief revision that expresses knowledge by a 2CNF. This procedure basically consists of adding a new belief expressed by a clause 2CNF, into the set of beliefs of a knowledge database, preserving its consistency. This generally implies having to remove some clauses from previous knowledge. The strategy employed to remove previous knowledge depends mainly on the consistency and the age of each clause. Finally, an application is presented to interact with the user in the pursute of educational ends.

Keywords: belief revision, 2 CNF, implication digraph, consistency, satisfiability

1. Introduction

The revision and transformation of knowledge is widely recognized as a key problem in knowledge representation and reasoning. Reasons for the importance of this topic are the facts that intelligent systems are gradually developed and refined, and that often the environment of an intelligent system is not static but changes over time [3].

Belief revision studies reasoning with changing information. Traditionally, belief revision techniques have been expressed using classical logic. Recently, the study of knowledge revision has increased since it can be applied to several areas of knowledge. Belief revision is a system that contains a corpus of beliefs which can be changed to accommodate new knowledge that may be inconsistent with any previous beliefs. Assuming the new belief is correct, as many of the previous ones should be removed as necessary so that the new belief can be incorporated into a consistent corpus. This process of adding beliefs corresponds to a non-monotonic logic [2, 7].

The AGM (Alchourrón, Gärdenfors and Makinson) model addresses the problem of belief revision using the tools of mathematical logic [4]. These works are considered the foundation for studying the problem of knowledge exchange. According to the AGM framework, knowledge K is represented by propositional logic theories and new information is represented by the same logic formulas.

One way to represent and check the consistency of a knowledge base is modeled by using the 2SAT problem, which has been shown to be solvable in polynomial time [12].

2. Preliminaries

During the 1970's from artificial intelligence and information technology the concept of "default reasoning" was introduced and defined by Raymond Reiter. This kind of logic sustains that in the absence of any contrary information, it is plausible to conclude X. It is a form of reasoning that takes into account the limitations of the agent and the commonness of things, which is pretty close to the way that everyday reasoning works. Indeed, it is due to this kind of reasoning that we can act in the world.

Well, the notion of plausible or default reasoning led to a vast area now known as non-monotonic logic or common sense, as well as circumscription logic(McCarthy), modal logic (McDermott and Doyle) and autoepistemic logic (Moore and Konolige) [8].

Non-monotonic logic is that form of reasoning under which a conclusion may be recast, retracted or defeated by an increase in information that modifies its premise. For example, the type of inference of everyday life in which people formulate tentative conclusions, reserving the right to withdraw them in light of new information. This logic satisfies the issue considering the defeatable nature of typical inferences of human common sense reasoning. Considering this type of reasoning, a formal and systematic study of cognitive processes that are present in the manipulation of knowledge structures emerges, by which an intelligent agent can draw conclusions in different ways, without having complete information to do so [9].

Before formalizing changes in beliefs, we must consider several issues: every execution of a dynamic model of beliefs must choose a language to represent them. Whatever the chosen language, the question arises of how to represent the corpus (base) of information as well as the operations for the concepts of minimum and maximum length change of the corpus of information. This implies an epistemic theory which considers the changes in knowledge and beliefs of a rational agent. In our case, we use the criteria of rationality to determine the behavior of changes in beliefs; criteria include the minimum change of preexisting beliefs, the primacy of new information and consistency. Thus for belief revision based on the AGM model using these criteria of rationality, three basic operations are used: expansion, contraction and review [4, 10].

Expansion is the operation that models the process of adding new knowledge to the corpus. This can be thought of as the expression of the learning process and is symbolized by the $+$ operator, so it is defined as, $F+p = C(F \cup p)$, where F is the knowledge base, p is the new belief and C is the function that check new knowledge base.

Contraction is the operation that causes a new belief to remove a piece from the corpus of knowledge, because the agent in question must stop having a certain position on this belief. This becomes complicated when there are other beliefs that would need to be abandoned based on the abandonment of the initial belief, so in the end, only the absolutely necessary beliefs would remain. This is symbolized by the operator $-$ and is defined as $F-p = C(F-p)$ where F is the corpus, the new belief p and C is the function that check new knowledge base.

Revision consists of modifying the set of beliefs when a new belief is incorporated into the previous set so that logical consistency is conserved. I If the set of beliefs is already consistent with the new information, then the review coincides with expansion, but if new knowledge is inconsistent with any previous beliefs, the operation of review must determine the resulting set of beliefs which keeps only the part of the original which would obtain a consistent result, so the original set of beliefs must be modified by eliminating as many beliefs as necessary to ensure that the resulting set, which includes the new belief, is consistent, and is defined as $F * p$ where F is the set of beliefs or knowledge base and p is the new belief.

To address the problems of belief revision, it is useful to consider the model using propositional logic to verify the consistency of the knowledge base in order to analyze results from adding new beliefs which are considered valid, so it is necessary to define the concepts of propositional logic involved as follows: a formula is said to be in conjunctive normal form (CNF) if it is composed of a conjunction of disjunctive clauses and will be true if all its clauses are [1, 6].

A clause is a disjunction of literals, so that each literal stands for any formula composed of a single proposition symbol x (positive literal) or its negation $-x$ (negative literal) or a constant $\perp$ o $\top$.

So any formula F can be translated into an implication digraph (EF), which is a directed graph whose construction is done by taking each of the clauses (x_i, x_j) of the formula, where vertices of the graph are the x_i and $-x_i$. Here, there is a vertex for each variable and another for its negation. For each clause, two edges are generated by applying the following formula: $(-x_i, x_j)$ and $(-x_j, x_i)$. The implication digraph is widely used to ensure if a formula is satisfiable or not [13].

The Satisfiability Problem(SAT) is posed as follows: given a set of variables and a constraint in conjunctive normal form, a truth assignment that satisfies the constraint must be found. In our case, we worked on CNF for 2SAT problem, which means the formula consists of clauses consisting of two literals [5].

To solve the 2SAT, the implication digraph is built and the strongly connected components of the digraph are calculated. It is said that the problem is solvable if and only if no variable and its negation belong to the same strongly connected component. There is a theorem that supports this formalism [11]: F is unsatisfiable if and only if a variable x exists such that there exist trajectories x a $-x$ and $-x$ to x in EF.

3. Knowledge Representation

In artificial intelligence, there are several problems where an initial knowledge base is considered. That is the case in belief revision, which can be considered as a propositional theory.

The review of a set of beliefs F with respect to a new knowledge p (new clause), is denoted by $F * p$. The set of beliefs resulting from the review should be consistent (C), $F * P = C(F * p)$. In the case that by adding new knowledge

p, the set of prior beliefs is inconsistent, it is necessary to find what clause must be removed in order to maintain consistency.

The problem studied in this work is knowledge or belief revision. Simply put, we seek to answer the question: If we have a consistent knowledge F and we want to incorporate new information p, is the resulting knowledge consistent or not?

To answer this question, the problem is viewed under the AGM model, which addresses the problem with the tools of mathematical logic and in particular, knowledge F is represented by propositional logic.

3.1. The approach of the research problem

Let the Knowledge base F be a consistent conjunctive normal formula (CNF) consisting of conjunctions of disjunctions of variables or their negations. In particular, we are considering 2SAT, where in SAT problem CNF consists of clauses consisting of two literals (variables or their negations).

$$F = \{(x_1 \vee x_2) \wedge (x_i \vee x_j) \ldots \wedge (x_m \vee x_n)\} \quad (1)$$

And let $p = (x_a, x_b)$ be the new knowledge clause. By applying this new knowledge p on the base of knowledge, the result is a new knowledge base that may remain consistent or may generate inconsistent knowledge, meaning there is no possible assignment that makes formula (1) satisfiable, where $i, j = 1 \ldots n$, x_i positive or negative. In that case, any of the clauses that cause inconsistency (and therefore non-satisfiability) of the formula F must be removed.

A CNF is true if all its clauses are. Each clause is true if it contains at least one true variable (satisfiability problem).

To represent the problem of belief revision the formula F transforms into EF (extended CNF) considering the associated implication digraph of the form (2) where vertex: one for each literal, edges: two for each clause:

$$(x_i \vee x_j) = \{(-x_i \rightarrow x_j) \wedge (-x_j \rightarrow x_i)\} \quad (2)$$

For example:
Let $F = \{(-x_1, -x_4), (-x_4, -x_3), (-x_1, -x_2), (x_4, x_1), (-x_4, x_1), (x_1, x_2)\}$, if we apply (2) the following implication digraph is generated
$EF = \{(x_1, -x_4), (x_4, -x_1), (x_4, -x_3), (x_3, -x_4), (x_1, -x_2), (x_2, -x_1), (-x_4, x_1), (-x_1, x_4), (x_4, x_1), (-x_1, -x_4)(-x_1, x_2)(-x_, x_1)\}$, see figure 1.

3.2. Expansion Process

To solve the problem of belief revision, we assume that the formula F which represents the initial knowledge base is satisfiable, we transform the formula F into EF and after adding a new clause p, the satisfiability of the new knowledge base is checked If it is unsatisfiable, the elimination process begins, removing

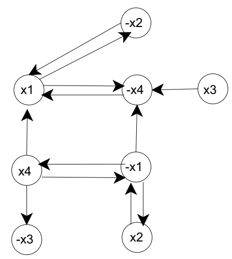

Fig. 1. Implication Digraph of formula F

one clause in the knowledge base at a time exhaustively and getting the number of inconsistent variables generated, until eventually the clause with fewest inconsistencies is removed.

Given the formula (1) the expansion process as follows:

1. Transform the formula F using (2) to create the representation of EF

$$\begin{aligned} EF = \{&(-x_1 \vee x_2) \wedge (-x_2 \vee x_1) \wedge (-x_i \vee x_j) \wedge (-x_j \vee x_i) \\ &\ldots (-x_m \vee x_n) \wedge (-x_n \vee x_m)\} \end{aligned} \tag{3}$$

2. Store data in a linked list L of the expression EF (3) that represents an implication digraph like the one shown below, see figure 2
3. Let TX the sets that represent the consistency of each literal, where the status of each one is known. These sets are calculated as follows $TX[x_i] = x_i, L[x_i] \cup L[L[x_i]]$ for each $L[x_i]$ that does not belong to the set. x_i is said to be inconsistent if in the whole $TX[x_i]$ there is a variable x_j and its negation $-x_j$.
4. If in the calculation of the set TX, the case exists that for some x_i, $TX[x_i]$ is inconsistent and $T[-x_i]$ is inconsistent then the formula F is unsatisfiable.
5. When adding new knowledge p, the new knowledge base will be $F + p$ (expansion process).
6. We evaluate this new knowledge base $C(F * p)$. If the result is unsatisfiable, then we apply the contraction process $F - p$ on the knowledge base.

3.3. Contraction Process

After adding new knowledge p, if the formula is unsatisfiable, it is necessary to discard a clause from the knowledge base. In this case, we are witnessing a contraction.

The aspect of complexity appears when we see that within the set of beliefs there are other accepted beliefs which cause the entire set of beliefs F to be unsatisfiable and therefore the knowledge is not consistent. In these cases it will be necessary to abandon other beliefs. We apply here the concept of "possible worlds", so we calculate the number of inconsistencies generated by eliminating each clause and finally select the one that causes the fewest possible.

The contraction of a set of beliefs F with respect to a sentence p, is denoted $F - p$.

For $i = 1, \ldots n$ where n is the number of clauses of the knowledge base F, we calculate the number of inconsistencies r_i generated by removing the clause (x_i, x_j), and then select as candidate clause the one that generates the fewest number of inconsistencies, if two or more clauses exist which share the fewest inconsistencies, we eliminate the clause that represents the earliest knowledge, thus generating the new knowledge base $F - p$.

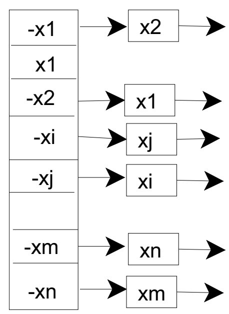

Fig. 2. Linked list of EF

Example 1: New knowledge base is added and the knowledge base continues to be consistent (process of expansion $F + p$).

Let $F = \{(-x_1, -x_4), (-x_4, -x_3), (-x_1, -x_2), (x_4, x_1), (-x_4, x_1)\}$ a satisfiable knowledge base that a new knowledge p is added and it is represented by the clause (x_1, x_2) thus the formula F is transformed to EF using the formula(2). $EF = \{(x_1, -x_4), (x_4, -x_1), (x_4, -x_3), (x_3, -x_4), (x_1, -x_2), (x_2, -x_1), (-x_4, x_1), (-x_1, x_4), (x_4, x_1), (-x_1, -x_4), (-x_1, x_2), (-x_2, x_1)\}$, see figure 3.

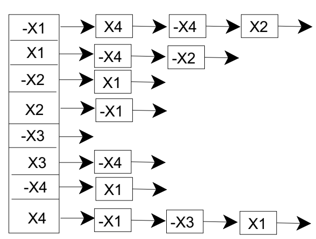

Fig. 3. Linked list of knowledge base of example 1

Below the sets $TX[x_i]$ are calculated for each variable
$TX[-x_1] = \{-x_1, x_4, -x_4, x_2, -x_3, x_1, -x_2\}$ inconsistent
$TX[x_1] = \{x_1, -x_4, -x_2\}$
$TX[-x_2] = \{-x_2, x_1, -x_4\}$
$TX[x_2] = \{x_2, -x_1, x_4, -x_4, -x_2, -x_3, x_1\}$ inconsistent
$TX[-x_3] = \{-x_3\}$
$TX[x_3] = \{x_3, -x_4, x_1, -x_2\}$
$TX[-x_4] = \{-x_4, x_1, -x_2\}$
$TX[x_4] = \{x_4, -x_1, -x_3, x_1, -x_4, x_2, -x_2\}$ inconsistent

In the computation of the TX sets does not exist contradiction so that the new knowledge is consistent and the formula F is satisfiable. The new knowledge base is $F + p$.

Example 2: New knowledge is added and the knowledge base is not consistent, so prior knowledge must be removed (contraction process $F - p$).

Let it be $F = \{(-x_1, -x_4), (-x_4, -x_3), (-x_1, -x_2), (x_4, x_1), (-x_4, x_1), (x_1, x_2)\}$, a satisfiable knowledge base, when a new knowledge p that is represented by the clause (x_2, x_4) is added, the formula F is converted in EF by the formula (2).

Lets it be $EF = \{(x_1, -x_4), (x_4, -x_1), (x_4, -x_3), (x_3, -x_4), (x_1, -x_2), (x_2, -x_1), -x_4, x_1), (-x_1, x_4), (x_4, x_1), (-x_1, -x_4), (-x_1, x_2), (-x_2, x_1), (-x_2, x_4), (-x_4, x_2)\}$, we make the representation of the implication digraph, see Figure 4.

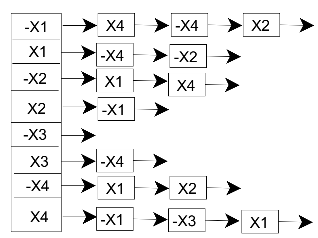

Fig. 4. Linked list of knowledge base for a new knowledge $(-x_2, x_4)(-x_4, x_2)$

Next, we calculate the sets $TX[x_i]$ for each variable
$TX[-x_1] = -x_1, x_4, -x_4, x_2, -x_3, x_1, -x_2$ inconsistent
$TX[x_1] = x_1, -x_4, -x_2, x_2, x_4, -x_1, -x_3$ inconsistent
$TX[-x_2] = -x_2, x_1, x_4, -x_4, -x_1, -x_3, x_2$ inconsistent
$TX[x_2] = x_2, -x_1, x_4, -x_4, -x_2, -x_3, x_1$ inconsistent
$TX[-x_3] = -x_3$
$TX[x_3] = x_3, -x_4, x_1, x_2, -x_2, x_4, -x_1, -x_3$ inconsistent
$TX[-x_4] = -x_4, x_1, x_2, -x_2, -x_1, -x_3, x_4$ inconsistent
$TX[x_4] = x_4, -x_1, -x_3, x_1, -x_4, x_2, -x_2$ inconsistent

When the sets TX are calculated, we observe that exist a contradiction for the variables x_1, x_2, x_4 ,so the knowledge base that is represented by the formula

F is not satisfiable, then is necessary modify the knowledge base by removing some previous clauses, considering the new knowledge is valid.

We consider remove the first clause and calculate the sets TX to know the number of unconsistent variables, similarly we remove the remaining clauses except the clause the represents the new knowledge p. In the table 1 are showed the results for each clauses.

Table 1. Inconsistent variables when the corresponding clause is removed

#clause	*Clause to remove*	*Inconsistent variables*
1	$(-x_1, -x_4)$	$-x_1, x_2, x_3, -x_4$
2	$(-x_4, -x_3)$	*Contradiction*
3	$(-x_1, -x_2)$	$-x_1, -x_2, x_4$
4	(x_4, x_1)	$x_1, -x_2, x_4$
5	$(-x_4, x_1)$	$x_1, -x_2, x_3, -x_4$
6	(x_1, x_1)	*Contradiction*

In the table 1, we can observe that the clause 2 and 6 are not candidate to be removed, because the formula F will be continue unsatisfiable, so the strategy used for removing knowledge involve removing the clause that generate the less number of inconsistencies that be the most old because the order by these arrived, in this case the candidate clause to be removed would be the clause 3 $(-x_1, -x_2)$ that make to the formula F with new knowledge $k * p$, satisfiable.

4. Conclusions

We present an introduction to the problem of belief revision using this 2CNF representation of the propositional calculus and generating consistent sets for each variable allows us to determine if knowledge is satisfiable or not.

We have a simple method based on the elimination of the clause that generates the fewest inconsistencies by adding new knowledge p, this thanks to the calculation of set TX and the implication generated by the implication graph.

We used the principles of expansion and contraction of AGM model and the criteria of the assumptions of "possible worlds".

Although the deleting process is exhaustive, the execution time can be better if we will use a distributed environment where each candidate clause will be verified in a processor.

It also created a simple application that allows people interested in the subject to evaluate a knowledge base for teaching purposes. It permits us to interact with the user to make tests about beliefs revision. The figure 5 shows the output of the application where a user load the consistent previous knowledge from a file or graphically build the implications digraph, to see figure 6, next the user can test the knowledge base to add new knowledge by clauses and the system will show the result, if it is not satisfiable then some old clauses will be removed and the new knowledge base will be show.

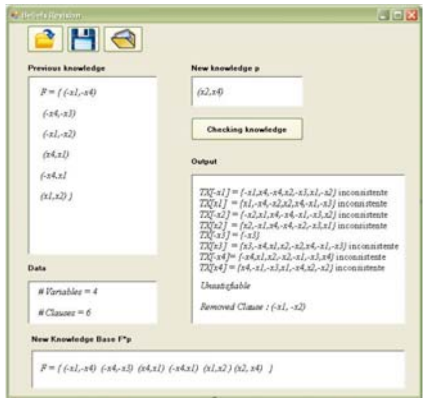

Fig. 5. **Main Screen of the application**

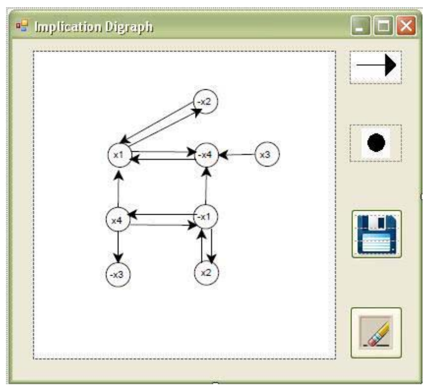

Fig. 6. **Screen for generate implication digraph**

References

1. De Ita, G., Bello, P., Contreras, M.: New Polynomial Classes for #2SAT Established Via Graph-Topological Structure. Jour. Engineering Letters, 15(2),250-258 (2007).
2. Delgrande, J.P., Schaub, T.: Two Approaches to Merging Knowledge Bases. In 9th European Conference on Logics in Artificial Intelligence, September, Lisbon Portugal (2004).
3. Eiter,T., Gottlob, G: On the Complexity of Propositional Knowledge Base Revision, Updates, and Counterfactuals. Artificial Intelligence. 57, 227-270(1992).
4. Ferme, E.: Revisión de Creencias, Revista Iberoamericana de Inteligencia Artificial.11, 17-39 (2007)
5. Ginsberg, M.L., Conunterfactuals. Artificial Intelligence, 30,35-79(1986).
6. Guilln,C., López, A., De Ita,G: Diseño de Algoritmos Combinatorios para # SAT y su Aplicacin al Razonamiento Proposicional. Technical report (2005)

7. Katsuno, H., Mendelzon, A. O.: On the Difference between Updating a Knowledge Base and Revising it. In Proceeding KB-91,387-395(1991).
8. Nebel. B.: A Knowledge Level Analysis of Belief Revision. In Proceeding KR-89, 301-311(1989).
9. Nebel, B.: Belief Revision and Default Reasoning:Syntax-Based Approaches. In Procceding KR-91,417-428 (1991).
10. Paniagua, E: Contribucin a la Representacin y Generacinde Planes con Incertidumbre. Doctoral Thesus. chapter 5. pp 113-210 Universidad Politécnica de Cataluña (1998)
11. Papadimitriou,C.: Computacional Complexity Addison-Wesley Publishing Company San Francisco(1994)
12. Roth D., On the hardness of approximate reasoning.Artificial Intelligence. 82, 273-302 (1996)
13. Wahlström, M.: A Tighter Bound for Counting Max-Weight Solutions to 2SAT Instances. LNCS, vol 5018, 202-213 Springer-Verlag (2008)