Propagation of Belief Functions through Frames of Discernment : Application to Context Computing

Bastien Pietropaoli, Michele Dominici, Frédéric Weis

Propagation of Belief Functions through Frames of Discernment : Application to Context Computing

Bastien Pietropaoli and **Michele Dominici**
INRIA, Rennes-Bretagne Atlantique,
Campus Universitaire de Beaulieu
35042 Rennes Cedex, France

Frédéric Weis
IRISA, Université de Rennes 1,
Campus Universitaire de Beaulieu
35042 Rennes Cedex, France

Abstract

In smart homes, many context attributes, i.e. small pieces of context, have to be computed from sensors in order to provide adapted services. We observe that many of those context attributes can be deduced from others. This paper presents how methods to propagate belief functions from one frame of discernment to another can be used to compute contexts. It enables the creation of evidential networks, preventing from recomputing many times the same evidence and enabling the construction of higher levels of abstraction often difficult to obtain directly from sensors. The proposed method is based on an intuitive way of thinking evidential networks considering mappings between frames of discernment. An example where this method is essential is presented.

Introduction

Context-aware applications have to sense the environment in order to adapt themselves and provide with contextual services. This is the case of Smart Homes equipped with sensors and augmented appliances. However, sensors can be numerous, heterogeneous and unreliable. Thus the data fusion is complex and requires a solid theory to handle those problems. For this purpose, we adopted the belief functions theory (BFT) (Shafer 1976).

The aim of the data fusion, in our case, is to compute small pieces of context we call *context attributes*. Those context attributes are diverse and could be for example the presence in a room, the number of people in a room or even that someone may be sleeping in a room. As shown in Figure 1, certain context attributes are based on the same sensors. Thus, the motion sensor in the presented example is used twice as evidence.

Since the BFT requires a substantial amount of computations, it could be a great idea to reduce as much as possible the number of evidence required to compute a context attribute. Moreover, the number of possible worlds, i.e. the number of possible states for a context attribute, is also an important source of computation. Thus, reducing the number of possible worlds we are working on is also important.

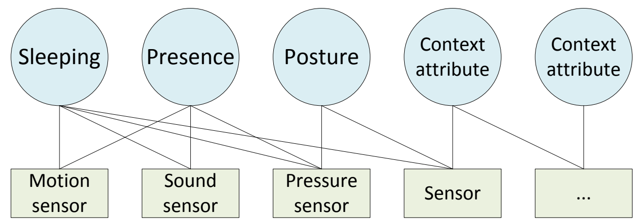

Figure 1: Example of context attributes associated to sensors. Some sensors are used many times by different context attributes.

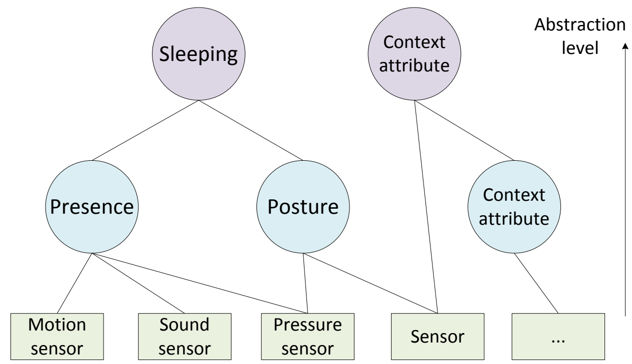

Figure 2: Example of network applied to a set of context attributes.

It is especially problematic when working on embedded systems, which may be the case when trying to observe context in smart homes. Thus, with this objective in mind, it is possible to observe that some context attributes could be used to compute others. By doing this, the number of gathered and combined evidence for each context attribute could be drastically reduced. In the example presented in Figure 1, the fact that someone may be sleeping can be easily inferred from the presence and the posture of this person. Then, the set of context attributes in Figure 1 can be transformed to obtain the network of Figure 2 if the sets of possible worlds for "Presence" and "Posture" are seen as subsets of "Sleeping".

The paper is organized as follows: the next section presents the basics of the BFT required to understand the presented methods and their advantages; it is followed by the presentation of methods used to propagate belief functions. An example where the application of these methods is useful is then presented. Finally, the paper is concluded.

Basics of BFT

In this section, we present only the basics required to understand the propagation of belief functions from one frame of discernment to another and the interest of doing so.

Frame of Discernment

In the BFT (Shafer 1976), the first thing that should be defined is a set of possible "worlds" called the *frame of discernment*. It is often noted:

$$\Omega = \{\omega_1, \omega_2, ..., \omega_n\} \tag{1}$$

These worlds have to be exclusive (i.e. $\omega_i \cap \omega_j = \emptyset$ if $i \neq j$) and if possible exhaustive meaning that the true state of the world has always been defined. To give an example, we can define a set of possible postures for someone as $\Omega = \{Seated, Standing, LyingDown\}$. In our work, each context attribute is represented by a frame of discernment.

Mass Function

Once the frame of discernment is created, a *mass function* representing the degree of belief associated to each subset of Ω is defined such that:

$$\begin{gathered} m : 2^{\Omega} \longmapsto [0,1] \\ \sum_{A \subseteq \Omega} m(A) = 1 \end{gathered} \tag{2}$$

Each subset $A \subseteq \Omega$ with $m(A) > 0$ is called a *focal element*. When a focal element contains several worlds, the degree of belief given by the mass function cannot be distributed in any way between those worlds. The degree of belief is thus completely associated to the fact that the true state of the world is contained in the focal element but cannot be more specific. Thus, the mass function can represent uncertainty using degrees of belief but also imprecision using non-atomic subsets of Ω. Those mass functions are used to express uncertainty on which "world" is the true state of a context attribute.

Building mass functions

To build mass functions, it is possible to use methods exploiting statistics (Aregui and Denoeux 2008). However, we wanted a method requiring only few experiments.

As described in (Pietropaoli, Dominici, and Weis 2012), for each couple (sensor/context attribute), we build a set of mass functions associating a mass function to any raw data given by a sensor. The set of mass functions is built after few experiments by observing the sensor behavior. It could be also totally based on intuition as human beliefs. (For more details on how this building method and the way it can be improved, please refer to (Pietropaoli, Dominici, and Weis 2012)).

Even if this method does not create perfect beliefs, it enables the use of many sensors to build evidence on context attributes. The imperfections are then compensated by the accumulation of evidence.

Accumulating evidence

The theory of belief functions is about accumulating evidence on what is going on. To do this, the most common rule of combination is the Dempster's rule given by:

$$m_1 \oplus m_2(A) = \frac{1}{1-K} \sum_{B \cap C = A \neq \emptyset} m_1(B) m_2(C) \tag{3}$$

$$\text{with } K = \sum_{B \cap C = \emptyset} m_1(B) m_2(C) \tag{4}$$

This combination rule offers a quick convergence when evidence is not in conflict. However, its complexity is not acceptable for embedded systems if the number of possible worlds increases too much and if many mass functions have to be combined altogether. There exist many other rules of combination in the literature. A generalization is also proposed by E. Lefevre et al. (Lefevre, Colot, and Vannoorenberghe 2002). Unfortunately, the complexity of all these rules is equivalent or even worse than the complexity of Dempster's rule. The reduction of the size of frames of discernment we are working on and the reduction of the number of combinations of evidence is thus an advantage for all these rules.

Propagation of belief functions

The idea of optimizing the computations required by belief functions by dividing the frame of discernment in subsets is not new (Shafer, Shenoy, and Mellouli 1987). In this section, we present a common method consisting in the propagation of belief functions based on a simple "one-to-one" correspondence between subsets of possible worlds but also how evidential mappings enable the use of heuristic knowledge links between frames of discernment.

Vacuous extension of frames

Let's consider two frames of discernment Ω and Θ supposed to describe the same context attribute. Let m be a mass function defined on the frame Ω that we want to propagate to the frame Θ. We can now consider the two possible cases: either $card(\Omega) \leq card(\Theta)$, either $card(\Theta) \leq card(\Omega)$. In simple words, one frame of discernment is more precise than the other.

In the first case, as both frames describe the same set of possible worlds, the elements of Ω are unions of elements of Θ (i.e. for each $A \subseteq \Omega$, there exist $B_i \in \Theta$ such that $A = \bigcup B_i$). The propagation of m to Θ is then called the *vacuous extension* of m and is denoted by $m^{\uparrow\Theta}$. It is given by:

$$m^{\uparrow\Theta}(A) = \begin{cases} m(B) & \text{if } A = B \subseteq \Omega \\ 0 & \text{otherwise.} \end{cases} , \forall A \subseteq \Theta \tag{5}$$

This propagation brings no new evidence and the two mass functions are strictly equivalent in terms of information.

To illustrate the propagation presented here, we can consider the two frames $\Omega = \{yes, no\}$ associated to the presence in a room and $\Theta = \{0, 1, 2+\}$ associated to the number of people in the same room (where "2+" means "two or more people"). It is easy to create a direct transformation of subsets of Ω into subsets of Θ. For example, we could have the following correspondence:

$$\begin{aligned} \{yes\} &= \{1 \cup 2+\} \\ \{no\} &= \{0\} \end{aligned}$$

As a consequence, it is possible to transform mass functions using the presented operations. For example, consider the following mass function

$$\begin{aligned} m(\{yes\}) &= 0.3 \\ m(\{no\}) &= 0.1 \\ m(\{yes \cup no\}) &= 0.6 \end{aligned}$$

By using the vacuous extension, it would become

$$\begin{aligned} m^{\uparrow\Theta}(\{1 \cup 2+\}) &= 0.3 \\ m^{\uparrow\Theta}(\{0\}) &= 0.1 \\ m^{\uparrow\Theta}(\{0 \cup 1 \cup 2+\}) &= 0.6 \end{aligned}$$

Restriction of frames

In the second case, as Θ is less precise than Ω, the propagation of m to Θ is equivalent to the conditioning of m given by:

$$m(A|B) = \begin{cases} \sum_{C \subseteq \bar{B}} m(A \cup C) & \text{for } A \subseteq B \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Let $m_{\downarrow\Theta}$ denote the *restriction* of m to Θ, then it is given by:

$$m_{\downarrow\Theta}(A) = m(A|\Theta) \tag{7}$$

As elements of Θ, in this case, are unions of elements of Ω, the resulting mass function can be less informative than the original one.

Those two ways of propagating belief functions from one frame to another are based on the direct correspondence that exists between the elements of the two frames. Those tools are very useful if it is possible to divide a huge frame of discernment in subsets of possible worlds. However, in some cases, the link between frames of discernment is not trivial and requires a more flexible approach.

Evidential mappings

As previously seen, some context attributes can be linked and it is possible to infer belief on a frame of discernment from belief on another frame of discernment. Unfortunately, the vacuous extension and the coarsening are valid only if one frame is a subset of the other one.

For instance, if we consider the posture of someone, defined by the frame $\Omega = \{Seated, Standing, LyingDown\}$ and if that person is sleeping, defined by $\Theta = \{yes, no\}$, it is not easy to create a direct correspondence between the worlds. As a matter of fact, a person could be sleeping seated in a sofa or reading a book, thus not sleeping, lying down on a bed. Anyway, the likelihood that someone may be sleeping while lying down is way higher than if he or she is standing or even seated...

Evidential mappings enable the use of heuristic knowledge to create a link between frames of discernment (Liu, Hughes, and McTear 1992). Thus, it is possible to propagate belief from any frame of discernment to another if there exists any known link between those frames.

The idea here is thus to create some kind of mass function on Θ for each subset of possible worlds of Ω. For example, we could define a mass function on the fact that the person may be sleeping, knowing that he or she is seated:

$$\begin{aligned} m(\{yes\}|Seated) &= 0.1 \\ m(\{no\}|Seated) &= 0.7 \\ m(\{yes \cup no\}|Seated) &= 0.2 \end{aligned}$$

This mass function corresponds to the heuristic knowledge linking the value "Seated" of the context attribute "Posture" to the context attribute "Sleeping". Since it is defined as a mass function, it is possible to use any already existing method to build mass functions such as methods based on intuition (Pietropaoli, Dominici, and Weis 2012) or methods based on statistics (Aregui and Denoeux 2008) by observing the correlation between the worlds in different frames of discernment.

By defining such a mass function for each nonempty subset of Ω, we obtain the *Basic Matrix* (BM) that characterizes completely the propagation of mass functions defined on Ω to the frame of discernment Θ.

Let's now consider a general case with $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ and $\Theta = \{\theta_1, \theta_2, ..., \theta_k\}$, two frames of discernment. The BM, in our case, is defined by:

$$M_{\Omega\to\Theta} = \begin{pmatrix} m(\{\theta_1\}|\omega_1) & \cdots & m(\{\theta_1\}|\Omega) \\ \vdots & \ddots & \vdots \\ m(\{\Theta\}|\omega_1) & \cdots & m(\{\Theta\}|\Omega) \end{pmatrix} \tag{8}$$

The element of the BM denoted by $M_{\Omega\to\Theta}(A, B)$ with $A \subseteq \Theta$ and $B \subseteq \Omega$ corresponds to $m(A|B)$

The propagation of any mass function m defined on Ω to Θ is obtained using the following formula for each A, $\emptyset \neq A \subseteq \Theta$:

$$m_{\Omega\to\Theta}(A) = \sum_{B \subseteq \Omega} m(B) M_{\Omega\to\Theta}(A, B) \tag{9}$$

The empty set is considered as a special case where the propagation of its mass always results in a mass given to itself. Thus we always have:

$$m_{\Omega\to\Theta}(\emptyset) = m(\emptyset) \tag{10}$$

In terms of computability, the worst case is in $O(2^{n+k})$ where n and k are the number of possible worlds defined in the two frames Ω and Θ respectively. However, in terms of implementation, as it is possible to consider only the focal elements of mass functions (Osswald 2012), it is possible to get rid of all the zeros of the basic matrix. The complexity increases with the "complexity" of heuristic knowledge linking the frames of discernment.

Subsets	$\{Se\}$	$\{St\}$	$\{Ly\}$	$\{Se \cup St\}$	$\{Se \cup Ly\}$	$\{St \cup Ly\}$	$\{Se \cup St \cup Ly\}$
$\{yes\}$	0.1	0	0.8	0	0.1	0	0
$\{no\}$	0.7	1	0	0	0.5	0	0
$\{yes \cup no\}$	0.2	0	0.2	1	0.4	1	1

Table 1: Basic matrix linking the posture of someone to the fact that he or she may be sleeping. "Se", "St" and "Ly" stands for "Seated", "Standing" and "LyingDown" respectively.

Example: Posture→ Sleeping

Let's now come back to our example of propagating the belief on the posture of someone to the fact that he or she may be sleeping. We can use the basic matrix shown in Table 1. Let's take an example mass function defined on the posture of someone:

$$\begin{aligned} m(\{Seated\}) &= 0.4 \\ m(\{LyingDown\}) &= 0.1 \\ m(\{Seated \cup LyingDown\}) &= 0.2 \\ m(\{\Omega\}) &= 0.3 \end{aligned}$$

If we propagate this mass function to the fact that the person may be sleeping, we obtain:

$$\begin{aligned} m_{\Omega\to\Theta}(\{yes\}) &= 0.4 * 0.1 + 0.1 * 0.8 + 0.2 * 0.1 \\ &= 0.14 \\ m_{\Omega\to\Theta}(\{no\}) &= 0.4 * 0.7 + 0.2 * 0.5 \\ &= 0.38 \\ m_{\Omega\to\Theta}(\{\Theta\}) &= 0.48 \end{aligned}$$

In this example, m is quite indecisive on the posture of the person, thus $m_{\Omega\to\Theta}$ is indecisive as well.

In the example given in Figure 2, the mass functions of the "Posture" and the "Presence" are themselves the result of a combination of mass functions based on sensor measures. They are propagated to the frame of discernment "Sleeping" and can be combined using any combination rule.

If one is concerned with the independence of evidence, for example when a sensor has been used twice or more to compute context attributes, the cautious rule of combination may be helpful (Denoeux 2006).

Conclusion

In this paper, we presented an approach of the evidential network for the computation of context attributes. The presented methods have many advantages:

- they enable the construction of higher level abstractions sometimes difficult to obtain directly from sensors or using joint mass functions;
- they reuse evidence by propagating entire mass functions from a level of abstraction to another;
- they enable the computation of beliefs on huge sets of possible worlds by dividing them in smaller subsets;
- they are flexible in the way it is built, based on statistics or on few experiments and intuition;
- they are completely independent of any combination rule and any decision making process.

Keeping in mind that our project ask for computing dozens of context attributes for a smart home, if possible within a sensor network or a sensor node network with very limited computation capabilities, the effective gain has to be tested with a high scale evidential network. Since an implementation of the theory is already working, testing it on high scales will eventually be a subject of interest in future work.

Acknowledgments

The authors would like to thank EDF that has funded this research and Guérain Le Bars from Université de Rennes 1 (France) for his contribution to this work during his internship. Thanks also to Nathanaël Prémillieu and Benjamin Lesage for their remarks and the discussion about this work.

References

Aregui, A., and Denoeux, T. 2008. Constructing consonant belief functions from sample data using confidence sets of pignistic probabilities. *International Journal of Approximate Reasoning* 49:575–594.

Denoeux, T. 2006. The cautious rule of combination for belief functions and some extensions. In *Information Fusion, 2006 9th International Conference on*, 1 –8.

Lefevre, E.; Colot, O.; and Vannoorenberghe, P. 2002. Belief function combination and conflict management. *Information Fusion* 3(2):149–162.

Liu, W.; Hughes, J. G.; and McTear, M. F. 1992. Representing heuristic knowledge in d-s theory. In *Proceedings of the Eighth international conference on Uncertainty in artificial intelligence*, UAI'92, 182–190. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.

Osswald, C. 2012. Controling the number of focal elements - some combinatorial considerations. In Denoeux, T., and Masson, M.-H., eds., *Belief Functions*, volume 164 of *Advances in Soft Computing*, 135–143. Springer.

Pietropaoli, B.; Dominici, M.; and Weis, F. 2012. Belief inference with timed evidence. In *Belief Functions: Theory and Applications*, volume 164 of *Advances in Intelligent and Soft Computing*. Springer Berlin / Heidelberg. 409–416.

Shafer, G.; Shenoy, P. P.; and Mellouli, K. 1987. Propagating belief functions in qualitative markov trees. *Int. J. Approx. Reasoning* 1(4):349–400.

Shafer, G. 1976. *A Mathematical Theory of Evidence*. Princeton: Princeton University Press.