TOWARDS A LOGIC OF PERISHABLE PROPOSITIONS

Flávio Soares Corrêa da Silva
Universidade de São Paulo, SP BRAZIL
Institute of Mathematics and Statistics
Department of Computer Science
Rua do Matão, 1010 – São Paulo (SP) BRAZIL 05504-900
Email: fcs@ime.usp.br
URL: http://www.ime.usp.br/~fcs

Abstract

In this article we consider propositions whose information about truth evaluation is gradually lost with time, e.g. information about financial markets, meteorological conditions, or - as presented in our prototypical example - car traffic status in big cities. We propose functions of loss of information, and show how they can be used to characterise the obsolescence of propositional information in a logical system.

Keywords: probabilistic propositional logic, artificial intelligence.

Introduction

The work presented here has been motivated by an engineering problem, which was proposed within the context of the project SIDAM (SIDAM - Sistemas de Informação Distribuídos para Agentes Móveis - Distributed Information Systems for Mobile Agents - is a research project funded by FAPESP (98/06138-2) and developed within the Department of Computer Science, University of Sao Paulo) [ESSRM00]: car traffic jams create ubiquitous and costly problems in big cities like Sao Paulo (Brazil), and one way to reduce these problems is by disseminating real time information about traffic status, e.g. by means of wireless communication systems, so that drivers have effective means to re-plan their routes to avoid problematic spots.

The local Traffic Engineering Company (CET) has developed an experimental system to generate information about traffic status at strategic points: data is collected and processed at every 30 minutes, and then broadcasted by radio and television stations and sent to pager systems and to the World Wide Web (to get the most recent information for the main avenues in Sao Paulo, check http://200.19.93.5/internew/index1.html).

The problem is that traffic status can change quite drastically within 30 minutes. Thus, the reliability of the information being consulted by a final user varies significantly across time (if a driver is unlucky and gets information that is 29 minutes old, s/he may be badly misled rather than aided in his/her re-routing procedures by that information).

We propose a framework to link information to measures of reliability, so that users of that information will at least know what they are buying. We restrict ourselves to finite propositional information - to avoid some computational tractability issues - and start with a logical system in which propositions are associated to probabilistic estimates of their truth evaluations. Then we consider what happens when

1. The probability measures change across time (which is what happens when e.g. we change from off-peak hours to rush hours), and when
2. The events of interest themselves change across time (thus taking into account the inherent dynamics of the system).

Preliminaries - Probabilistic Propositional Logic

The basic language in use here is the propositional version of a simple logic extensively explored in [CorreadaSilva92, CdSRH93, CorreadaSilva96, CorreadaSilva96b, CDSC01].

Given a finite set of atomic propositions $P = \{p_1, \ldots, p_n\}$ and the standard connectives $\{\neg, \vee, \wedge, \rightarrow\}$ we build propositional expressions the usual way. Each proposition is evaluated to $\{T, F\}$ to produce a classical prepositional theory, and we have 2^n alternative evaluations for the set of expressions (the evaluations for some expressions may of course coincide in different evaluations for the whole set of expressions - e.g. in all 2^n evaluations we find that the evaluation of $(p_1 \vee \neg p_1)$ is equal to T). We call the set $W = \{T, F\}^P$ the set of alternative evaluations, or possible worlds for the expressions based on P.

Now we build a partition $\sigma = \{S_1, \ldots S_m\}$ of W - i.e. a collection of subsets of W such that $S_i \cap S_j = \{\}, i \neq j$, and $\bigcup_1^m S_i = W$.

Then, we attach to each subset S_i a probability estimate $\mu(S_i) \in [0,1]$, such that $\sum_1^m \mu(S_i) = 1$.

Given any subset $E \subseteq W$, the internal and external probability estimates for E (respectively, $\mu_*(E)$ and $\mu^*(E)$) are given by:

$$\mu_*(E) = \sum_{Si \subseteq E} \mu(S_i)$$

$$\mu^*(E) = \sum_{Si \cap E \neq \{\}} \mu(S_i)$$

It is not difficult to show that if $E = \bigcup_{\{i1,\ldots ir\} \subseteq \{1,\ldots,m\}} S_{ij}$, then $\mu_*(E) = \mu^*(E) = \sum_{\{i1,\ldots ir\} \subseteq \{1,\ldots,m\}} S_{ij}$.

Given any propositional expression φ, we have $E(\varphi) \subseteq W$ defined as the set of evaluations in which φ is T. Thus, we have the probability estimates for φ given by:

$$\mu_*(\varphi) = \sum_{Si \subseteq E(\varphi)} \mu(S_i)$$

$$\mu^*(\varphi) = \sum_{Si \cap E(\varphi) \neq \{\}} \mu(S_i)$$

It is also not difficult to show that, for any expression φ, we have:

$$\mu_*(\varphi) + \mu^*(\neg\varphi) = 1$$

If probability estimates were "non-perishable items", given the probability estimates for subsets of W, we could evaluate at any time the probability estimates for any propositional expression φ using the formulae above.

Let us consider however that probability estimates change smoothly with time, i.e. after an infinitesimal interval of time the estimates $\mu(S_i)$ may have suffered infinitesimal variations.

In this case, given an expression φ, and given an elapsed time t from the moment the corresponding probabilities $\mu_*(S_i)$ and $\mu^*(S_i)$ were evaluated, we should account for the possible changes occurred in those estimated probabilities.

This is what we take into account in the next sessions.

Propositions With Smooth Variations of Probability Estimates

In a similar vein to what was proposed in [CdSV00], we define two probability estimates $\mu, \mu' : \sigma \rightarrow [0,1]$ to be δ-neighbors if $\mu(S_i) - \mu(S_i) \leq \dfrac{\delta}{m-1}$ for all $S_i \in \sigma = \{S_1, \ldots S_m\}$.

Given any two δ-neighbors μ and μ', we have the following results for any propositional expression φ:

$$\mu_*(\varphi) - \mu'_*(\varphi) \leq \delta$$

$$\mu^*(\varphi) - \mu'^*(\varphi) \leq \delta$$

We now define $\delta(t)$ as a non-decreasing function of $t, t \geq 0$ being a representation of linear time. We assume that $\delta(0) = 0$.

We assume that, when we reach time t, a probability estimate μ may have changed to any of its δ(t)-neighbors. If we have no means to identify what is the prevailing probability estimate at time t, the best we can do is characterise the smallest interval which we can assure that contains the updated probability estimate. The intervals $\mu_*^t(\varphi)$ and $\mu^{*t}(\varphi)$ can be defined as below:

$$\mu_*^t(\varphi) = [\max\{0, \mu_*(\varphi) - \delta(t)\}, \min\{1, \mu_*(\varphi) + \delta(t)\}]$$

$$\mu^{*t}(\varphi) = [\max\{0, \mu^*(\varphi) - \delta(t)\}, \min\{1, \mu^*(\varphi) + \delta(t)\}]$$

In [Vit00] some experimental results are presented based on this model, employing linear-by-parts δ-neighborhood functions. It is shown that, if the linear slope of $\delta(t)$ is adjusted from off-peak to rush hour traffic situations, this framework can be used to provide car drivers with appropriate estimates of the reliability of traffic information.

While the interval $[\mu_*^t(\varphi), \mu^{*t}(\varphi)]$ is not equal to [0,1], we have some information about the possible probabilities of $E(\varphi)$. We define the perishing time of φ

as being the smallest t such that $[\mu_*^{t}(\varphi), \mu*'(\varphi)] = [0,1]$. If information feed occurs at time intervals t' larger than the perishing time t of some proposition φ, then there are going to be moments (between t and t') when we will not be able to obtain any useful information about truth evaluations of φ.

Throughout [CdSV00,Vit00] and till this section of the present article, it is considered that only probability estimates can become obsolete with time. However, it can be the case that events of interest also change: some event may need to occur at time 0 - represented by a propositional expression $\varphi(0)$ - so that another event $\varphi(t)$ can occur at time t.

In the next section we build the concept of neighbor events. With this in hand, given an event $\varphi(t)$ occurring at time t, we can identify the collection of events $\varphi_i(0)$ that are neighbors of $\varphi(t)$, i.e. that have the possibility of evolving to $\varphi(t)$. Then we can estimate the probabilities of those events at time 0, thus estimating the probability of having $\varphi(t)$ at time t.

To clarify our point, we observe that this corresponds to a change of perspective on how information becomes obsolete across time: by fixing the event of interest and taking into account the δ-neighbors of the probability estimate measured at time 0, we look forward in time, accounting for the possible changes that may occur in that estimate; alternatively, when we admit that the present event of interest $\varphi(t)$ can be the result of one of its neighboring events $\varphi_i(0)$ and take as approximation for the probability of $\varphi(t)$ the measures for $\bigcup \varphi_i(0)$, we look back in time, accounting for past events that may have evolved to the present event of interest.

Probabilities of Smooth Variations of Propositions

In our setting, an event of interest is characterised by a propositional expression: given an expression φ, the set of possible worlds $E(\varphi)$ in which φ is T is the event of interest, whose probability we are interested in estimating.

A neighboring event E' must be characterised in terms of some measure of similarity between events. In [EGGR97,DPEGG97,EGG00] we find an interesting proposal to account for similarity between sets, acknowledged to be originally authored by Enrique Ruspini.

Similarity is constructed based on the concept of triangular norm. A triangular norm in the real interval [0,1] is any binary operation Δ with the following properties:

- $\Delta : [0,1]^2 \rightarrow [0,1]$
- $x\Delta(y\Delta z) = (x\Delta y)\Delta z$
- $x\Delta y = y\Delta x$
- $x \geq y, z \in [0,1] \Rightarrow x\Delta z \geq y\Delta z$
- $1\Delta x = x$
- $0\Delta x = 0$

Triangular norms are the standard class of operations used to capture the notion of fuzzy conjunction. Given two sets of possible worlds $S_i, S_j \in W$, it seems natural to say that S_i and S_j are perfectly similar if $S_i \cap S_j = S_i = S_j$. This notion of similarity is extended by means of triangular norms to approximately similar sets in terms of a similarity measure $s(S_i, S_j)$ as follows:

- $s : (2^W)^2 \rightarrow [0,1]$
- $s(S_i, S_j) = 1 \mathit{iff} S_i = S_j$
- $s(S_i, S_j) = s(S_j, S_i)$
- $s(S_i, S_j)\Delta s(S_j, S_k) \overset{\leq}{=} s(S_i, S_k)$

Similarity measures are also defined for singleton sets. Recalling that $W = \{T, F\}^P$, given two worlds $w, w' \in W$ we have:

$$s(w, w') = s(\{w\}, \{w'\})$$

An implicative similarity $I(S_i, S_j)$ extends the notion of set subsumption: if $S_j \subseteq S_i$ we say that S_j is perfectly subsumed by S_i. Borrowing from [EGGR97,DPEGG97,EGG00], we define implicative similarity as:

- $I : (2^W)^2 \rightarrow [0,1]$
- $I(S_i, S_j) = \min_{w \in S_j} \{\max_{w' \in S_i} \{s(w, w')\}\}$

We now propose an additional relation, based on the concept of implicative similarity. We define symmetric implicative similarity Is as:

- $Is : (2^W)^2 \rightarrow [0,1]$
- $Is(S_i, S_j) = I(S_i, S_j)\Delta I(S_j, S_i)$

Symmetric implicative similarity measures the extent to which two subsets of W can be regarded as logically equivalent.

We are now in position to define the neighborhood of an event. Given a set $E \subseteq W$, we define the symmetric neighborhood of E - denoted as $N*$ - as:

- $N^*: 2^W x[0,1] \to 2^W$
- $N^*(E,\alpha) = \bigcup\{S_i \subseteq W : Is(S_i, E) \geq \alpha\}$

Thus, $N^*(E,\alpha)$ denotes the largest set E^* such that $Is(E^*, E) \geq \alpha$.

Notice that there are no means to characterise a unique "lower bound" for sets similar to E based on symmetric similarity measures.

If we define $\alpha(t)$ as a non-increasing function of $t, t \geq 0$, such that $\alpha(0) = 1$, then given an event of interest E occurring at time t, we can determine the event $E^* \supseteq E$, that is the least determined event that can have occurred at time 0 and from which E may have evolved.

Since $E^* \supseteq E$, then $\mu^*(E^*) \geq \mu^*(E)$ and $\mu_*(E^*) \geq \mu_*(E)$ for any probability measure μ. With all this in hand, we have a safe way to determine an upper bound for the probability of a propositional expression φ occurring at time t being evaluated to T, given probability estimates generated at time 0. The "algorithm" is as follows:

- determine the event of interest at time $t, E(\varphi)$
- determine the largest event that may have occurred at time 0 to evolve to $E(\varphi)$ at time $t, E^* = N^*(E(\varphi), \alpha(t))$
- determine the upper bound for the probability of E^* at time $t, \hat{\mu}^{*\prime}(\varphi) = \mu^*(E^*)$.

Since we have no means to characterise a unique "lower bound" for E at time 0, we do not have means to construct a sound definition of a lower bound for the probability estimate of φ being evaluated to T at time t considering that this event may have evolved from a different event at time 0, unless we work by contraposition:

- negate the original expression, thus producing $\neg\varphi$
- determine the event $E(\neg\varphi)$
- determine the largest event that may have occurred at time 0 to evolve to $E(\neg\varphi)$ at time $t, E^{\neg *} = N^*(E(\neg\varphi), \alpha(t))$
- determine the upper bound for the probability of $E^{\neg *}$ at time t, given by $\hat{\mu}^{*\prime}(\neg\varphi) = \mu(E^{\neg *})$.
- determine the lower bound for the probability of $E(\varphi)$ at time t, given by $\hat{\mu}_*^{t}(\varphi) = 1 - \hat{\mu}^{*\prime}(\neg\varphi)$.

Hence, we have two alternative approaches to account for the "perishability" of information: we can either look forward in time and consider that probability estimates can change as time goes by, or we can look back in time and consider that events of interest can be the result of previous different occurring events. There is no clear way of choosing which approach will give the most accurate results, and one possible way to circumvent the problem can be the adoption of a "most conservative" choice, i.e. choose the solutions that minimise the possibility of error.

In the next section we give a numerical example, to make our discussion more concrete.

An Illustrative Example

As an illustration, let us consider the set $P = \{p1, p2, p3, p4\}$. We employ the following notation for possible worlds: a possible world is represented by the smallest natural number whose digits are precisely the indices of the elements of P whose truth evaluation is set to T in that world. Hence,

- the world in which $p1 = p2 = p3 = p4 = F$ is denoted as 0,
- the world in which $p1 = p2 = p3 = p4 = T$ is denoted as 1234,
- the world in which $p1 = p2 = F$ and $p3 = p4 = T$ is denoted as 34,

and so forth. Then, we have the following set of possible worlds:

$$W = \{0,1,2,3,4,12,13,14,23,24,34,123,124,134,234,1234\}$$

Now we define

$$\sigma = \{S1 = \{0,1,2,3,4\}, S2 = \{12,13,14\},$$
$$S3 = \{23,24\}, S4 = \{34\},$$
$$S5 = \{123,124,134,234,1234\}\}$$

Let us say that at time 0 we have $\mu(S1) = 0.4, \mu(S2) = \mu(S4) = 0.1, \mu(S3) = \mu(S5) = 0.2$. We now consider the following propositional expression:

$(p1 \quad p2 \quad p3 \quad p4)$
$(\ p1 \quad ((p2 \quad ((p3 \quad p4)$
$(\ p3 \quad p4))) \quad (\ p2 \quad p3 \quad p4)))$

The expression was carefully constructed so that $E(\ ) \quad \{12,23,24,34\}$.

We adopt as a triangular norm the relation min, and define the following similarity relations between pairs of possible worlds for $t \quad 0$:

s(w,w')	0	1	2	3	4	12	13	14	23	24	34	123	124	134	234	1234
0	1	.9	.8	.7	.6	.5	.4	.3	.2	.3	.4	.5	.6	.7	.8	.9
1	.9	1	.9	.8	.7	.6	.5	.4	.3	.2	.3	.4	.5	.6	.7	.8
2	.8	.9	1	.9	.8	.7	.6	.5	.4	.3	.2	.3	.4	.5	.6	.7
3	.7	.8	.9	1	.9	.8	.7	.6	.5	.4	.3	.2	.3	.4	.5	.6
4	.6	.7	.8	.9	1	.9	.8	.7	.6	.5	.4	.3	.2	.3	.4	.5
12	.5	.6	.7	.8	.9	1	.9	.8	.7	.6	.5	.4	.3	.2	.3	.4
13	.4	.5	.6	.7	.8	.9	1	.9	.8	.7	.6	.5	.4	.3	.2	.3
14	.3	.4	.5	.6	.7	.8	.9	1	.9	.8	.7	.6	.5	.4	.3	.2
23	.2	.3	.4	.5	.6	.7	.8	.9	1	.9	.8	.7	.6	.5	.4	.3
24	.3	.2	.3	.4	.5	.6	.7	.8	.9	1	.9	.8	.7	.6	.5	.4
34	.4	.3	.2	.3	.4	.5	.6	.7	.8	.9	1	.9	.8	.7	.6	.5
123	.5	.4	.3	.2	.3	.4	.5	.6	.7	.8	.9	1	.9	.8	.7	.6
124	.6	.5	.4	.3	.2	.3	.4	.5	.6	.7	.8	.9	1	.9	.8	.7
134	.7	.6	.5	.4	.3	.2	.3	.4	.5	.6	.7	.8	.9	1	.9	.8
234	.8	.7	.6	.5	.4	.3	.2	.3	.4	.5	.6	.7	.8	.9	1	.9
1234	.9	.8	.7	.6	.5	.4	.3	.2	.3	.4	.5	.6	.7	.8	.9	1

Just as an additional illustration, this gives us the following implicative similarities between pairs of elements of at $t \quad 0$:

I(A,B)	B=S1	S2	S3	S4	S5
A=S1	1	.7	.5	.4	.5
S2	.5	1	.8	.7	.3
S3	.3	.7	1	.9	.4
S4	.2	.5	.8	1	.5
S5	.5	.4	.7	.9	1

Implicative similarities determine symmetric implicative similarities as follows for $t \quad 0$:

Is(A,B)	B=S1	S2	S3	S4	S5
A=S1	1	.5	.3	.2	.5
S2		1	.7	.5	.3
S3			1	.8	.4
S4				1	.5
S5					1

Let us adopt now the following linear functions for (t) and (t):

$$(t) \quad 0.1 \quad t$$
$$(t) \quad 1 \quad 0.2 \quad t$$

This will give us the following results for t=0,1,2,3,4,5:

	t=0	1	2	3	4	5
$E*$	12,23, 24,34	3,4, 12,13, 14,23, 24,34, 123,124	1,2, 3,4, 12,13, 14,23, 24,34, 123,124, 134,234	0,1, 2,3,4, 12,13, 14,23, 24,34, 123,124, 134,234, 1234	0,1, 2,3,4, 12,13, 14,23, 24,34, 123,124, 134,234, 1234	0,1, 2,3,4, 12,13, 14,23, 24,34, 123,124, 134,234, 1234
$*_t$	.4	.5	.6	.7	.8	.9
$\overset{t}{*}$	.3	.2	.1	0	0	0
$\hat{}*_t$	.4	1	1	1	1	1
$\hat{}\overset{t}{*}$	.3	0	0	0	0	0

This means that, if we take into account only the variation of probability estimates, will not perish before time t=5, whereas if we take into account the variations of propositions, will perish as early as at time t=1. A conservative analysis could then advise us to "look back in time" with these data, to ensure quality of information being provided to users of an information system based on these estimates.

Discussion

In this article we proposed a simple system to deal with information that can become obsolete. We introduced a propositional logic in which propositions become unreliable as time passes, and proposed the concept of perishability of propositions: a proposition perishes when we loose all probabilistic information about its truth evaluation.

This system can be useful to safeguard users from unreliable information. Many interesting problems arise from the utilisation of this system and shall be studied for future presentation:

- alternative queries: many interesting queries can be posed to this system, and their corresponding algorithms must be constructed and implemented, e.g. what is the possibly most probable event at any given time t, given $\delta(t)$ and $\alpha(t)$?
- computational tractability issues.
- extensibility to first order languages.
- proof systems development for logics of perishable information.

We are particularly interested in this last issue, and future articles shall be devoted to the following topic: consider that we have time as linearly dependent of the length of proofs. Then, the lengthier the proof the less useful its results (from either of the viewpoints considered here), and if the information used within a proof perishes before the end of the proof then that proof will also perish. We are at the moment studying how to compute non-perishable proofs.

Bibliography

[CorreadaSilva92] Correa da Silva, F. S.: Automated Reasoning With Uncertainties. PhD thesis (1992), University of Edinburgh, Scotland.

[CdSRH93] Correa da Silva, F. S., Robertson, D. S., Hesketh, J.: Automated Reasoning With Uncertainties. Chapter 5 of Knowledge Representation and Uncertainty, edited by M. Masuch and L. Pawlos, Springer Verlag LNAI 808 (1993).

[CorreadaSilva96] Correa da Silva, F. S.: On Proof- and Model-based Techniques for Reasoning with Uncertainty. International Journal of Intelligent Systems, v. 10, pgs. 561-570 (1995).

[CorreadaSilva96b] Correa da Silva, F. S.: On Reasoning With and Reasoning About Uncertainty in Artificial Intelligence. Proceedings of European Summer Meeting of the Association of Symbolic Logic (1996).

[CDSC01] Correa da Silva, F. S., Carbogim, D. V.: Facts, Annotations, Arguments, and Reasoning. New Generation Computing, vol. 19 (1), pgs. 1-22 (2001).

[CdSV00] Correa da Silva, F. S., Vitorio, S. A.: Remarks on Reasoning with Progressively Obsolete Information. Proceedings of the Brazilian Symposium of Artificial Intelligence / Ibero-American Conference of Artificial Intelligence (2000).

[EGGR97] Esteva, F., Garcia, P., Godo, L., Rodriguez, R.: A Modal Account of Similarity-Based Reasoning. International Journal of Approximate Reasoning, vol. 16, pgs. 235-260 (1997).

[DPEGG97] Dubois, D., Prade, H., Esteva, F., Garcia, P., Godo, L.: A Logical Approach to Interpolation Based on Similarity Relations. International Journal of Approximate Reasoning, vol. 17, pgs. 1-36 (1997).

[EGG00] Esteva, F., Garcia, P., Godo, L.: Similarity-Based Reasoning. Discovering Worlds With Fuzzy Logic: Perspectives and Approaches to Formalisation of Human-Consistent Logical Systems, edited by V. Novak and I. Perfilieva, Springer Verlag (2000).

[ESSRM00] Endler, M., Silva, D. M., Silva e Silva, F., Rocha, R. A., Moura, M. A.: Project SIDAM: Overview and Preliminary Results. Annals of 2nd Workshop on Wireless Communications (2000), Brazil.

[Vit00] Vitorio, S. A.: Obsolescencia de Informacao. MSc dissertation, Department of Computer Science, University of Sao Paulo (2000), Brazil.