Paraconsistent Sensitivity Analysis for Bayesian Significance Tests

Julio Michael Stern

BIOINFO and Computer Science Dept., University of São Paulo
`jstern@ime.usp.br`

Abstract. In this paper, the notion of degree of inconsistency is introduced as a tool to evaluate the sensitivity of the Full Bayesian Significance Test (FBST) value of evidence with respect to changes in the prior or reference density. For that, both the definition of the FBST, a possibilistic approach to hypothesis testing based on Bayesian probability procedures, and the use of bilattice structures, as introduced by Ginsberg and Fitting, in paraconsistent logics, are reviewed. The computational and theoretical advantages of using the proposed degree of inconsistency based sensitivity evaluation as an alternative to traditional statistical power analysis is also discussed.

Keywords: Hybrid probability / possibility analysis; Hypothesis test; Paraconsistent logic; Uncertainty representation.

1 Introduction and Summary

The Full Bayesian Significance Test (FBST), first presented in [25] is a coherent Bayesian significance test for sharp hypotheses. As explained in [25], [23], [24] and [29], the FBST is based on a possibilistic value of evidence, defined by coherent Bayesian probability procedures. To evaluate the sensitivity of the FBST value of evidence with respect to changes in the prior density, a notion of degree of inconsistency is introduced and used. Despite the possibilistic nature of the uncertainty given by the degree of inconsistency defined herein, its interpretation is similar to standard probabilistic error bars used in statistics. Formally, however, this is given in the framework of the bilattice structure, used to represent inconsistency in paraconsistent logics. Furthermore, it is also proposed that, in some situations, the degree of inconsistency based sensitivity evaluation of the FBST value of evidence, with respect to changes in the prior density, be used as an alternative to traditional statistical power analysis, with significant computational and theoretical advantages. The definition of the FBST and its use are reviewed in Section 2. In Section 3, the notion of degree of inconsistency is defined, interpreted and used to evaluate the sensitivity of the FBST value of evidence, with respect to changes in the prior density. In Section 4, two illustrative numerical examples are given. Final comments and directions for further research are presented in Section 5. The bilattice structure, used to represent inconsistency in paraconsistent logics is reviewed in the appendix.

2   The FBST Value of Evidence

Let $\theta \in \Theta \subseteq \mathcal{R}^p$ be a vector parameter of interest, and $L_x = L(\theta \mid x)$ the likelihood associated to the observed data x, a standard statistical model. Under the Bayesian paradigm the posterior density, $p_x(\theta)$, is proportional to the product of the likelihood and a prior density $p(\theta)$. That is,

$$p_x(\theta) \propto p(\theta)L(\theta \mid x).$$

The (null) hypothesis H states that the parameter lies in the null set Θ_H, defined by

$$\Theta_H = \{\theta \in \Theta \mid g(\theta) \leq 0 \wedge h(\theta) = 0\},$$

where g and h are functions defined in the parameter space. Herein, however, interest will rest particularly upon sharp (precise) hypotheses, i.e., those for which $\dim(\Theta_H) < \dim(\Theta)$.

The posterior surprise, $s(\theta)$, relative to a given reference density $r(\theta)$, is given by

$$s(\theta) = p_x(\theta)/r(\theta).$$

The relative surprise function, $s(\theta)$, was used by several others statisticians, see [19], [20] and [13]. The supremum of the relative surprise function over a given subset Θ_H of the parameter space, will be denoted by $s^*(\Theta_H, p, L_x, r)$, that is,

$$s^*(\Theta_H, p, L_x, r) = \sup_{\theta \in \Theta_H} s(\theta)$$

Despite the importance of making a conceptual distinction between the statement of a statistical hypothesis, H, and the corresponding null set, Θ_H, one often relax the formalism and refers to the hypothesis Θ_H, instead of $H : \theta \in \Theta_H$. In the same manner, when some or all of the argument functions, p, L_x and r, are clear from the context, they may be omitted in a simplified notation and $s^*(\Theta_H)$ or even $s^*(H)$ would be acceptable alternatives for $s^*(\Theta_H, p, L_x, r)$.

The contour or level sets, $C(\varphi, p, L_x, r)$, of the relative surprise function, and the Highest Relative Surprise Set (HRSS), $D(\varphi, p, L_x, r)$, at a given level φ, are given by

$$C(\varphi, p, L_x, r) = \{\theta \in \Theta \mid s(\theta) = \varphi\} \quad , \quad D(\varphi, p, L_x, r) = \{\theta \in \Theta \mid s(\theta) > \varphi\}$$

The FBST value of evidence against a hypothesis H, $\mathrm{Ev}(H)$ or $\mathrm{Ev}(\Theta_H)$, is defined by

$$\mathrm{Ev}(\Theta_H, p, L_x, r) = \int_{T(\Theta_H, p, L_x, r)} p_x(\theta)\, d\theta \ , \ \text{where}$$
$$T(\Theta_H, p, L_x, r) = D(s^*(\Theta_H), p, L_x, r)$$

The tangential HRSS $T(\Theta_H)$, or $T(H)$, contains the points in the parameter space whose surprise, relative to the reference density, is higher than that of

any other point in the null set Θ_H. When the uniform reference density, $r(\theta) \propto 1$, is used, $T(\Theta_H)$ is the Posterior's Highest Density Probability Set (HDPS) tangential to the null set Θ_H.

The role of the reference density in the FBST is to make $\mathrm{Ev}(H)$ implicitly invariant under suitable transformations of the coordinate system. Invariance, as used in statistics, is a metric concept. The reference density is just a compact and interpretable representation for the reference metric in the original parameter space. This metric is given by the geodesic distance on the density surface, see [7] and [24]. The natural choice of reference density is an uninformative prior, interpreted as a representation of no information in the parameter space, or the limit prior for no observations, or the neutral ground state for the Bayesian operation. Standard (possibly improper) uninformative priors include the uniform and maximum entropy densities, see [11], [18] and [21] for a detailed discussion.

The value of evidence against a hypothesis H has the following interpretation: "Small" values of $\mathrm{Ev}(H)$ indicate that the posterior density puts low probability mass on values of θ with high relative surprise as compared to values of $\theta \in \Theta_H$ thus providing weak evidence against hypothesis H. On the other hand, if the posterior probability of $T(\Theta_H)$ is "large", that is for "large" values of $\mathrm{Ev}(H)$, values of θ with high relative surprise as compared to values of $\theta \in \Theta_H$, have high posterior density. The data provides thus strong evidence against the hypothesis H. Furthermore, the FBST is "Fully" coherent with the Bayesian likelihood principle, that is, that the information gathered from observations is represented by (and only by) the likelihood function.

3   Prior Sensitivity and Inconsistency

For a given likelihood and reference density, let, $\eta = \mathrm{Ev}(\Theta_H, p, L_x, r)$ denote the value of evidence against a hypothesis H, with respect to prior p. Let $\eta', \eta'' \ldots$ denote the evidence against H with respect to priors $p', p'' \ldots$. The degree of inconsistency of the value of evidence against a hypothesis H, induced by a set of priors, $\{p, p', p'' \ldots\}$ can be defined by the index

$$I\{\eta, \eta', \eta'' \ldots\} \;=\; \max\{\eta, \eta', \eta'' \ldots\} - \min\{\eta, \eta', \eta'' \ldots\}$$

This intuitive measure of inconsistency can be made rigorous in the context of paraconsistent logic and bilattice structures, see the appendix. If $\eta = \mathrm{Ev}(H)$ is the value of evidence against H, the value of evidence in favor of H is defined by $\overline{\eta} = \overline{\mathrm{Ev}}(H) = 1 - \mathrm{Ev}(H)$. The point $x = \langle \overline{\eta}, \eta \rangle$ in the unit square bilattice, represents herein a single evidence, see the appendix. Since $\mathrm{BI}(x) = 0$, such a point is consistent. It is also easy to verify that for the multiple evidence values, the definition of degree of inconsistency given above, is the degree of inconsistency of the knowledge join of all the single evidence points in the unit square bilattice,

$$I(\eta, \eta', \eta'' \ldots) \;=\; \mathrm{BI}\left(\langle \overline{\eta}, \eta \rangle \sqcup_k \langle \overline{\eta}', \eta' \rangle \sqcup_k \langle \overline{\eta}'', \eta'' \rangle \ldots\right) \;.$$

As shown in [29], the value of evidence in favor of a composite hypothesis $H = A \vee B$, is the most favorable value of evidence in favor of each of its terms, i.e., $\overline{\mathrm{Ev}}(H) = \max\{\overline{\mathrm{Ev}}(A), \overline{\mathrm{Ev}}(B)\}$. This makes $\overline{\mathrm{Ev}}$ a possibilistic (partial) support structure coexisting with the probabilistic support structure given by the posterior probability measure in the parameter space, see [10] and [29]. The degree of inconsistency for the evidence against H induced by multiple changes of the prior can be used as an index of imprecision or fuzziness of the value of evidence $\mathrm{Ev}(H)$. Moreover, it can also be interpreted within the possibilistic context of the partial support structure given by the evidence. Some of the alternative ways of measuring the uncertainty of the value of evidence $\mathrm{Ev}(H)$, such as the empirical power analysis have a dual possibilistic / probabilistic interpretation, see [28] and [22]. The degree of inconsistency has also the practical advantage of being "inexpensive", i.e., given a few changes of prior, the calculation of the resulting inconsistency requires about the same work as computing $\mathrm{Ev}(H)$. In contrast, an empirical power analysis requires much more computational work than it is required to compute a single evidence.

4   Numerical Examples

In this paper we will concentrate on two simple model examples: the Hardy-Weinberg (HW) Equilibrium Law model and Coefficient of Variation model. The HW Equilibrium is a genetic model with a sample of n individuals, where x_1 and x_3 are the two homozygote sample counts and $x_2 = n - x_1 - x_3$ is the hetherozygote sample count. The parameter vector for this trinomial model is $\theta = [\theta_1, \theta_2, \theta_3]$ and the parameter space, the null hypothesis set, the prior density, likelihood function and the reference density are given by:

$$\Theta \;=\; \{\theta \geq 0 \mid \theta_1 + \theta_2 + \theta_3 = 1\} \;\;,\;\;\;\; \Theta_H \;=\; \{\theta \in \Theta \mid \theta_3 = (1 - \sqrt{\theta_1})^2\}$$

$$p(\theta) \;\propto\; 1 \;\;,\;\;\;\; L(\theta \mid x) \propto \theta_1^{x_1}\theta_2^{x_2}\theta_3^{x_3} \;\;,\;\;\;\; r(\theta) \;\propto\; 1$$

For the Coefficient of Variation model, a test for the coefficient of variation $C = \mu\sqrt{\rho}$, of a normal variable with mean μ and precision $\rho = 1/\sigma^2$, the parameter space, the null hypothesis set, the maximum entropy prior, the reference density, and the likelihood density are given by:

$$\Theta \;=\; \{\theta = [\mu, \rho] \in \mathcal{R} \times \mathcal{R}_+\} \;\;,\;\;\;\; \Theta_H \;=\; \{\theta \in \Theta \mid \mu\sqrt{\rho} = c\}$$

$$p(\theta) \;\propto\; 1/\rho \;\;,\;\;\;\; r(\theta) \;\propto\; 1$$

$$L_x(\theta \mid x) \propto \sqrt{\rho}\, \exp\left(-\frac{n}{2}\rho\,(\mu - m)^2\right)\exp\left(-\frac{n}{2}v\,\rho\right)\rho^a$$

$$x = [x_1 \ldots x_n] \;\;,\;\;\;\; m = \frac{1}{n}\sum_{i=1}^{n} x_i \;\;,\;\;\;\; v = \frac{1}{n}\sum_{i=1}^{n}(x_i - m)^2 \;\;,\;\;\;\; a = \frac{n-1}{2}$$

Figure 1 displays the elements of a value of evidence against the hypothesis, computed for the HW (Left) and Coefficient of Variation (Right) models. The

null set, Θ_H, is represented by a dashed line. The contour line of the posterior, delimiting the tangencial set, $T(\Theta_H)$, is represented by a solid line. The posterior unconstrained maximum is represented by "o" and the posterior maximum constrained to the null set is represented by "*".
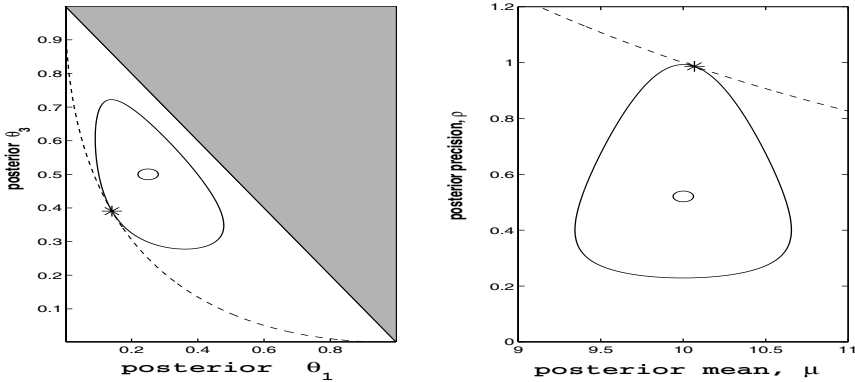
Fig. 1. FBST for Hardy-Weinberg (L) and Coefficient of Variation (R)

In order to perform the sensitivity analysis several priors have to be used. Uninformative priors are used to represent no previous observations, see [16], [21] and [31] for a detailed discussion.

For the HW model we use as uniformative priors the uniform density, that can be represented as $[0,0,0]$ observation counts, and also the standard maximum entropy density, that can be represented as $[-1,-1,-1]$ observation counts. Between these two uninformative priors, we also consider perturbation priors corresponding to $[-1,0,0]$, $[0,-1,0]$ and $[0,0,-1]$ observation counts. Each of these priors can be interpreted as the exclusion of a single observation of the corresponding type from the data set, $x_1, \ldots x_n$.

Finally, we consider the dual perturbation priors corresponding to $[1,0,0]$, $[0,1,0]$ and $[0,0,1]$ observation counts. The term dual is used meaning that instead of exclusion, these priors can be interpreted as the inclusion of a single artificial observation of the corresponding type, x_{n+1}, in the data set.

The examples in the top part of Table 1 are given by size and proportions, $[x_1, x_2, x_3] = n * [0.25, 0.5, 0.25]$, where the HW hypothesis is true.

For the Coefficient of Variation model we use as uninformative priors the uniform density, for the mean, and either the standard maximum entropy density, $p(\theta) \propto 1/\rho$, or the uniform, $p(\theta) \propto 1$, for the precision. We also consider (with uniform prior) perturbations by the inclusion in the data set of an artificial observation, x_{n+1}, at fixed quantiles of the predictive posterior, in this case, at three standard deviations below or above the mean, $x_{n+1} = m \pm 3\,\sigma$.

The examples in the bottom part of Table 2 are given by $cv = 0.1$, size n, and the sufficient statistics $m = 10$ and $std = 1.2$, where the hypothesis is false.

Table 1. HW and CV models, $\mathrm{Ev}(H)$ for several priors and sample sizes

HW model; sample $= n[1/4, 1/2, 1/4]$; H: equilibrium, true										
$p(\theta) \setminus n =$	8	16	32	64	128	$p(\theta) \setminus n =$	8	16	32	64	128
$[0, 0, 0]$	0.00	0.00	0.00	0.00	0.00	$[-1, -1, -1]$	0.13	0.04	0.02	0.01	0.00
$[1, 0, 0]$	0.05	0.03	0.02	0.01	0.00	$[-1, 0, 0]$	0.12	0.04	0.02	0.01	0.00
$[0, 1, 0]$	0.07	0.03	0.02	0.01	0.00	$[0, -1, 0]$	0.09	0.04	0.02	0.01	0.00
$[0, 0, 1]$	0.05	0.03	0.02	0.01	0.00	$[0, 0, -1]$	0.12	0.04	0.02	0.01	0.00

CV model; suff.stat: $m = 10, std = 1.2$; H: $cv = 0.1$, false										
$p(\rho) \setminus n =$	16	32	64	128	256	$x_{n+1} \setminus n =$	16	32	64	128	256
$\propto 1$	0.45	0.69	0.91	0.99	1.00	$m + 3\sigma$	0.38	0.66	0.89	0.99	1.00
$\propto 1/\rho$	0.64	0.79	0.94	0.99	1.00	$m - 3\sigma$	0.55	0.75	0.92	0.99	1.00

In order to get a feeling of the asymptotic behavior of the evidence and the inconsistency, the calculations are repeated for the same sufficient statistics but for sample sizes, n, taking values in a convenient range. In Figure 2, the maximum and minimum values of evidence against the hypothesis H, among all choices of priors used in the sensitivity analysis, are given by the interpolated dashed lines. For the HW model, Table 1 and Figure 2 top, the sample size ranged from $n = 8$ to $n = 128$. For the Coefficient of Variation model, Table 1 and Figure 2 bottom, the sample size ranged from $n = 16$ to $n = 256$. In Figure 2, the induced degree of inconsistency is given by the vertical distance between the dashed lines. The interpretation of the vertical interval between the lines in Figure 2 (solid bars) is similar to that of the usual statistical error bars. However, in contrast with the empirical power analysis developed in [28] and [22], the uncertainty represented by these bars does not have a probabilistic nature, being rather a possibilistic measure of inconsistency, defined in the partial support structure given by the FBST evidence, see [29].
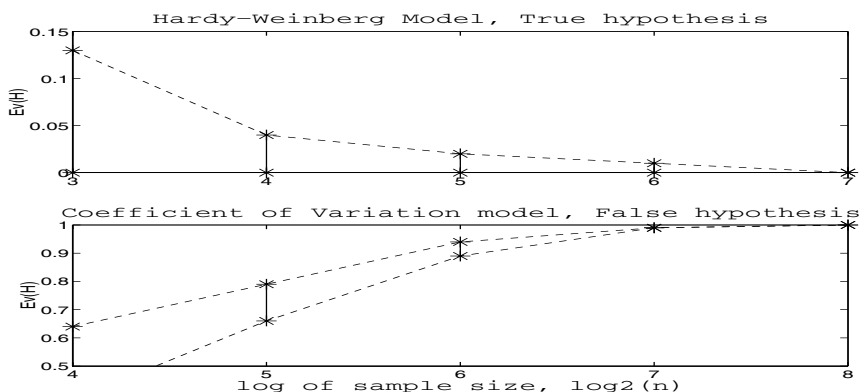
Fig. 2. Sensitivity Analysis for $\mathrm{Ev}(H)$

5   Directions for Further Research and Acknowledgements

For complex models, the sensitivity analysis in the last section can be generalized using perturbations generated by the inclusion of single artificial observations created at (or the exclusion of single observations near) fixed quantiles of some convenient statistics, $t(x)$, of the predictive posterior. Perturbations generated by the exclusion of the most extreme observations, according to some convenient criteria, could also be considered. For the sensitivity analysis consistency when the model allows the data set to be summarized by some sufficient statistics in the form of L-estimators, see [4], section 8.6. The asymptotic behavior of the sensitivity analysis for several classes of models and perturbations is the subject of forthcoming articles.

Finally, perturbations to the reference density, instead of to the prior, could be considered. One advantage of this approach is that, when computing the evidence, only the integration limit, i.e. the threshold s^*, is changed, while the integrand, i.e. the posterior density, remains the same. Hence, when computing $Ev(H)$, only little additional work is required for the inconsistency analysis.

References

1. Abe,J.M. Avila,B.C. Prado,J.P.A. (1998). *Multi-Agents and Inconsistence.* IC-CIMA'98. 2nd International Conference on Computational Intelligence and Multimidia Applications. Traralgon, Australia.
2. Alcantara,J. Damasio,C.V. Pereira,L.M. (2002). Paraconsistent Logic Programs. JELIA-02. 8th European Conference on Logics in Artificial Intelligence. *Lecture Notes in Computer Science,* 2424, 345–356.
3. Arieli,O. Avron,A. (1996). Reasoning with Logical Bilattices. *Journal of Logic, Language and Information*, 5, 25–63.
4. Arnold,B.C. Balakrishnan,N. Nagaraja,H.N. (1992). *A First Course in Order Statistics.* NY: Wiley.
5. C.M.Barros, N.C.A.Costa, J.M.Abe (1995). Tópicos de Teoria dos Sistemas Ordenados. *Lógica e Teoria da Ciência*, 17,18,19. IEA, Univ. São Paulo.
6. N.D.Belnap (1977). A useful four-valued logic, pp 8–37 in G.Epstein, J.Dumm. *Modern uses of Multiple Valued Logics.* Dordrecht: Reidel.
7. Boothby,W. (2002). *An Introduction to Differential Manifolds and Riemannian Geometry.* Academic Press, NY.
8. N.C.A.Costa, V.S.Subrahmanian (1989). Paraconsistent Logics as a Formalism for Reasoning about Inconsistent Knowledge Bases. *Artificial Inteligence in Medicine*, 1, 167–174.

9. Costa,N.C.A. Abe,J.M. Subrahmanian,V.S. (1991). Remarks on Annotated Logic. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 37, 561–570.

10. Darwiche,A.Y. Ginsberg,M.L. (1992). *A Symbolic Generalization of Probability Theory.* AAAI-92. 10th Natnl. Conf. on Artificial Intelligence. San Jose, USA.

11. Dugdale,J.S. (1996). *Entropy and its Physical Meaning.* Taylor-Francis,London.

12. Epstein,G. (1993). *Multiple-Valued Logic Design.* Inst.of Physics, Bristol.

13. M.Evans (1997). Bayesian Inference Procedures Derived via the Concept of Relative Surprise. *Communications in Statistics*, 26, 1125–1143.

14. M. Fitting (1988). Logic Programming on a Topological Bilattice. *Fundamentae Informaticae*, 11, 209–218.

15. Fitting,M. (1989). Bilattices and Theory of Truth. *J. Phil. Logic*, 18, 225–256.

16. M.H.DeGroot (1970). *Optimal Statistical Decisions.* NY: McGraw-Hill.

17. Ginsberg,M.L. (1988). Multivalued Logics. *Computat. Intelligence*, 4, 265–316.

18. Gokhale,D.V. (1999). On Joint Conditional Enptropies. *Entropy Journal*,1,21–24.

19. Good,I.J. (1983). *Good Thinking.* Univ. of Minnesota.

20. Good,I.J. (1989). Surprise indices and p-values. *J. Statistical Computation and Simulation*, 32, 90–92.

21. Kapur,J.N.(1989). *Maximum Entropy Models in Science Engineering.* Wiley, NY.

22. Lauretto,M. Pereira,C.A.B. Stern,J.M. Zacks,S. (2004). Comparing Parameters of Two Bivariate Normal Distributions Using the Invariant FBST. To appear, Brazilian Journal of Probability and Statistics.

23. Madruga,M.R. Esteves,L.G. Wechsler,S. (2001). On the Bayesianity of Pereira-Stern Tests. *Test*, 10, 291–299.

24. Madruga,M.R. Pereira,C.A.B. Stern,J.M. (2003). Bayesian Evidence Test for Precise Hypotheses. *Journal of Statistical Planning and Inference,* 117,185–198.

25. Pereira,C.A.B. Stern,J.M. (1999). Evidence and Credibility: Full Bayesian Significance Test for Precise Hypotheses. *Entropy Journal*, 1, 69–80.

26. Pereira,C.A.B. Stern,J.M. (2001). Model Selection: Full Bayesian Approach. *Environmetrics*, 12, 559–568.

27. Perny,P. Tsoukias,A. (1998). *On the Continuous Extension of a Four Valued Logic for Preference Modelling.* IPMU-98. 7th Conf. on Information Processing and Management of Uncertainty in Knowledge Based Systems. Paris, France.

28. Stern,J.M. Zacks,S. (2002). Testing the Independence of Poisson Variates under the Holgate Bivariate Distribution, The Power of a new Evidence Test. *Statistical and Probability Letters*, 60, 313–320.

29. Stern,J.M. (2003). Significance Tests, Belief Calculi, and Burden of Proof in Legal and Scientific Discourse. Laptec-2003, 4th Cong. Logic Applied to Technology. *Frontiers in Artificial Intelligence and its Applications*, 101, 139–147.

30. Zadeh,L.A. (1987). *Fuzzy Sets and Applications.* Wiley, NY.

31. Zellner,A. (1971). *Introduction to Bayesian Inference in Econometrics.* NY:Wiley.

Appendix: Bilattices

Several formalisms for reasoning under uncertainty rely on ordered and lattice structures, see [5], [6], [8], [9], [14], [15], [17], [30] and others. In this section we recall the basic bilattice structure, and give an important example. Herein, the presentations in [2] and [3], is followed.

Given two complete lattices, $\langle C, \leq_c \rangle$, and $\langle D, \leq_d \rangle$, the bilattice $B(C, D)$ has two orders, the knowledge order, $\leq_k$, and the truth order, $\leq_t$, given by:

$$B(C, D) = \langle C \times D, \leq_k, \leq_t \rangle$$
$$\langle c_1, d_1 \rangle \leq_k \langle c_2, d_2 \rangle \Leftrightarrow c_1 \leq_c c_2 \text{ and } d_1 \leq_d d_2$$
$$\langle c_1, d_1 \rangle \leq_t \langle c_2, d_2 \rangle \Leftrightarrow c_1 \leq_c c_2 \text{ and } d_2 \leq_d d_1$$

The standard interpretation is that C provides the "credibility" or "evidence in favor" of a hypothesis (or statement) H, and D provides the "doubt" or "evidence against" H. If $\langle c_1, d_1 \rangle \leq_k \langle c_2, d_2 \rangle$, then we have more information (even if inconsistent) about situation 2 than 1. Analogously, if $\langle c_1, d_1 \rangle \leq_t \langle c_2, d_2 \rangle$, then we have more reason to trust (or believe) situation 2 than 1 (even if with less information).

For each of the bilattice orders we define a join and a meet operator, based on the join and the meet operators of the single lattices orders. More precisely, $\sqcup_t$ and $\sqcap_t$, for the truth order, and $\sqcup_k$ and $\sqcap_k$, for the knowledge order, are defined by the folowing equations:

$$\langle c_1, d_1 \rangle \sqcup_t \langle c_2, d_2 \rangle = \langle c_1 \sqcup_c c_2, d_1 \sqcap_d d_2 \rangle \quad ,$$
$$\langle c_1, d_1 \rangle \sqcap_t \langle c_2, d_2 \rangle = \langle c_1 \sqcap_c c_2, d_1 \sqcup_d d_2 \rangle$$
$$\langle c_1, d_1 \rangle \sqcup_k \langle c_2, d_2 \rangle = \langle c_1 \sqcup_c c_2, d_1 \sqcup_d d_2 \rangle \quad ,$$
$$\langle c_1, d_1 \rangle \sqcap_k \langle c_2, d_2 \rangle = \langle c_1 \sqcap_c c_2, d_1 \sqcap_d d_2 \rangle$$

Negation type operators are not an integral part of the basic bilattice structure. Ginsberg (1988) and Fitting (1989) require of possible "negation", $\neg$ and "conflation", $-$, operators to be compatible with the bilattice orders, and to satisfy the double negation property:

Ng1: $x \leq_k y \Rightarrow \neg x \leq_k \neg y$, Ng2: $x \leq_t y \Rightarrow \neg y \leq_t \neg x$, Ng3: $\neg\neg x = x$.
Cf1: $x \leq_k y \Rightarrow -y \leq_k -x$, Cf2: $x \leq_t y \Rightarrow -x \leq_t -y$, Cf3: $--x = x$.

Hence, negation should reverse trust, but preserve knowledge, and conflation should reverse knowledge, but preserve trust. If the double negation property is not satisfied (Ng3 or Cf3) the operators are called weak (negation or conflation).

The "unit square" bilattice, $\langle [0,1] \times [0,1], \leq, \leq \rangle$ has been routinely used to represent fuzzy or rough pertinence relations, logical probabilistic annotations, etc. Examples can be found in [1], [9], [12], [27], [30] and others. The lattice $\langle [0,1], \leq \rangle$ is the standard unit interval, where the join and meet, $\sqcup$ and $\sqcap$ coincide with the max and min operators. The standard negation and conflation operators are defined by $\neg \langle c, d \rangle = \langle d, c \rangle$, $- \langle c, d \rangle = \langle 1-c, 1-d \rangle$.

In the unit square bilattice the "truth", "false", "inconsistency" and "indetermination" extremes are t, f, $\top$, $\bot$, whose coordinates are given in Table 3. As a simple example, let region R be the convex hull of the four vertices n, s, e and w, given in Table 3. Points kj, km, tj and tm are the knowledge and truth join and meet, over $r \in R$.

In the unit square bilattice, the degree of trust and degree of inconsistency for a point $x = \langle c, d \rangle$ are given by a convenient linear reparameterization of $[0,1]^2$, to $[-1, +1]^2$ defined by

Table 2. Coordinates $\langle c, d\rangle$ and $\langle \mathrm{BT}, \mathrm{BI}\rangle$ for example A1

	f	$\perp$	$\top$	t	w	n	s	e	tm	km	kj	tj
c	0	0	1	1	1/4	1/2	1/2	3/4	1/4	1/4	3/4	3/4
d	1	0	1	0	1/2	3/4	1/4	1/2	3/4	1/4	3/4	1/4
BT	−1	0	0	1	−1/4	−1/4	1/4	1/4	−1/2	0	0	1/2
BI	0	−1	1	0	−1/4	1/4	−1/4	1/4	0	−1/2	1/2	0
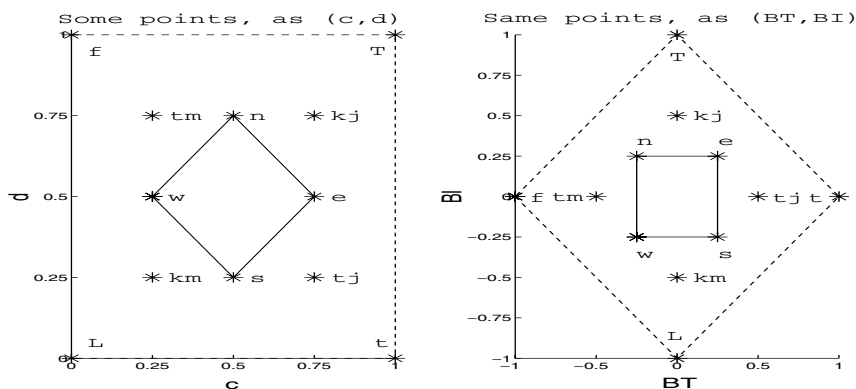
Fig. 3. Points in Table 3, using (c, d) and $(\mathrm{BT}, \mathrm{BI})$ coordinates

$$\mathrm{BT}\left(\langle c, d\rangle\right) = c - d \ , \quad \mathrm{BI}\left(\langle c, d\rangle\right) = c + d - 1 \ .$$

Figure 3 shows the points in Table 3 in the unit square bilattice, also using the trust-inconsistency reparameterization.