Partial Inconsistency and Vector Semantics

Michael Bukatin[1], Ralph Kopperman[2], and Steve Matthews[3]

[1] Nokia Corporation
Cambridge, Massachusetts, USA
bukatin@cs.brandeis.edu
[2] Department of Mathematics
City College
City University of New York
New York, New York, USA
rdkcc@ccny.cuny.edu
[3] Department of Computer Science
University of Warwick
Coventry, UK
Steve.Matthews@warwick.ac.uk

Abstract. This short communication lists the key elements of the mathematics of partial inconsistency. The links between partial inconsistency and vector semantics are emphasized. Possible applications to program learning are discussed.

Keywords: Bilattices, bitopology, negative length, negative probability, probabilistic programming, fuzzy sampling, generalized animation

1 Introduction

The traditional mathematical view is that there is only one kind of contradiction and that all contradictions imply each other and everything else. However, there is also rich tradition of studying various kinds of graded or partial contradictions.

There are a number of common motives appearing multiple times in various studies of graded inconsistency. These common motives link a variety of independently done studies together and serve as focal elements of what we call the *partial inconsistency landscape* [3]. We list many of these common motives and their interplay.

An especially important motive is that in the presence of partial inconsistency many otherwise impoverished algebraic structures become groups and vector spaces. In particular, domains for denotational semantics tend to acquire group and vector space structure when partial inconsistency is present.

Known applications include handling of inconsistent information and non-monotonic and anti-monotonic inference. Perhaps even more importantly for the advanced AI, vector semantics is likely to offer new powerful schemes for program learning.

2 Focal Elements of the Partial Inconsistency Landscape

– Various forms of *negative measure* (negative length and distance, negative probability and signed measures, negative membership and signed multisets)

– Bilattices
– Bitopology
– Domains with group and vector space structures
– Non-monotonic and anti-monotonic inference
– Modal and paraconsistent logic and possible world models
– Bilattice pattern and Hahn-Jordan decomposition

3 Partially Inconsistent Interval Numbers

Interval numbers are segments $[a,b]$ on the real line where $a \leq b$. One can extend interval numbers by adding *pseudosegments* $[a,b]$ with the contradictory property that $b < a$. This structure was independently discovered many times and is known under various names including Kaucher interval arithmetic, directed interval arithmetic, generalized interval arithmetic, and modal interval arithmetic (a comprehensive repository of literature on the subject is maintained by Evgenija Popova [10]). Our group tends to call it *partially inconsistent interval numbers*.

There are two partial orders on partially inconsistent interval numbers. The *informational order*, $\sqsubseteq$, is defined by reverse inclusion on interval numbers: $[a,d] \sqsubseteq [b,c]$ iff $a \leq b$ and $c \leq d$. The same formula is used for partially inconsistent interval numbers. The *material order* is component-wise: $[a,b] \leq [c,d]$ iff $a \leq c$ and $b \leq d$.

Addition on interval numbers (and partially inconsistent interval numbers) is defined component-wise: $[a_1,b_1]+[a_2,b_2] = [a_1+a_2, b_1+b_2]$.

The operation of *weak minus* is defined as $-[a,b] = [-b,-a]$. Addition and weak minus are *monotonic* with respect to $\sqsubseteq$.

Consider $-[a,b]+[a,b] = [-b,-a]+[a,b] = [a-b,b-a]$. If $a < b$, then the strict inequality, $[a-b,b-a] \sqsubset [0,0]$, holds. So if $a < b$, $-[a,b]+[a,b]$ approximates $[0,0]$, but is not equal to it, hence interval numbers with weak minus don't form a group.

If one allows pseudosegments, one can define the component-wise *true minus*: $-[a,b] = [-a,-b]$. Partially inconsistent interval numbers with the component-wise addition and the true minus form a group (and a 2D vector space over reals). The true minus maps precisely defined numbers, $[a,a]$, to precisely defined numbers, $[-a,-a]$. Other than that, the true minus maps segments to pseudosegments and maps pseudosegments to segments. The true minus is *anti-monotonic* with respect to $\sqsubseteq$.

4 Bilattices

A bilattice is a set equipped with two lattice structures defining two partial orders, the *material order*, $\leq$, and the *informational order*, $\sqsubseteq$, and an involution monotonic with respect to $\sqsubseteq$, antimonotonic with respect to $\leq$, and preserving appropriate lattice structures. Additional axioms are often imposed.

Bilattices were introduced by Matthew Ginsberg [4] to provide a unified framework a variety of inferences schemes used in AI, such as *non-monotonic inference*, inference with uncertainty, etc. They are now ubiquitous in the studies of partial and graded inconsistency.

The simplest example of a bilattice is the four-valued logic: $f < \bot < t, f < \top < t$, $\bot \sqsubset f \sqsubset \top, \bot \sqsubset t \sqsubset \top$.

Partially inconsistent interval numbers form a bilattice. Sometimes one wants both orders to form complete lattices. This can be achieved by allowing a and b to also take $-\infty$ and $+\infty$ as values, or by confining a and b within a segment $[A,B]$, in both cases sacrificing the property of partially inconsistent interval numbers being a group.

5 Bitopology

A bitopology is a set equipped with two topologies. There are at least three ways bitopologies occur in studies of partial inconsistency. The connections between partial inconsistency and bitopological Stone duality via the notion of d-frame are explored in [6]. A fuzzy bitopology valued in lattice L is a fuzzy topology valued in the bilattice L^2 (in particular, an ordinary bitopology is a topology valued in the four-valued logic) [11]. Finally, in the context of bitopological groups and anti-monotonic inverse the following situation is typical: two topologies, T and T^{-1}, are group dual of each other, the multiplication is continuous with respect to both topologies, and the inverse is a bicontinuous map from (X,T,T^{-1}) to its bitopological dual, (X,T^{-1},T) [1].

In particular, consider the upper topology on reals (the open rays, $(x,+\infty)$, and the empty set) and the dual lower topology (the open rays $(-\infty,x)$ and the empty set). The minus (which is anti-monotonic with respect to $\leq$) is a bi-continuous function from (lower, upper) bitopology to (upper, lower) bitopology, and vice versa.

Consider the (lower, upper) bitopology on the real line. Define the bilattice isomorphism between the d-frame elements, i.e. pairs $\langle L,U \rangle$ of the respective open sets, and partially inconsistent interval numbers. A pair $\langle L,U \rangle$ is a pair of open rays, $\langle (-\infty,a),(b,+\infty) \rangle$ (a and b are allowed to take $-\infty$ and $+\infty$ as values). This pair corresponds to a partially inconsistent interval number $[a,b]$. Consistent, i.e. non-overlapping, pairs of open rays ($a \leq b$) correspond to segments. Total, i.e. covering the whole space, pairs of open rays ($b < a$) correspond to pseudosegments.

It is natural to ascribe negative length $b-a$ to pseudosegments and to associate with them generalized characteristic functions taking the value -1 in the (b,a) interval and 0 outside of that interval (signed multisets allowing negative degree of membership).

6 Negative probability and vector semantics

One can think about probabilistic programs as transformers from the probability distributions on the space of inputs to the probability distributions on the space of outputs. Dexter Kozen showed that it is fruitful to replace the space of probability distributions by the space of signed measures [7]. One defines $\nu < \mu$ iff $\mu - \nu$ is a positive measure. The space of signed measures is a vector lattice (a Riesz space) and a Banach space, so people call this structure a Banach lattice. Denotations of programs are continuous linear operators with finite norms. The probabilistic powerdomain is embedded into the positive cone of this Banach lattice. The structure of Hilbert space on signed measures can be obtained via reproducing kernel methods (see Chapter 4 of [2]).

The Hahn-Jordan decomposition, $\mu^+ = \mu \vee 0, \mu^- = \mu \wedge 0, \mu = \mu^+ + \mu^-$, holds, since it's a theorem for all lattice-ordered groups. Defining $\nu \sqsubseteq \mu$ iff $\nu^+ \leq \mu^+$ and $\nu^- \leq \mu^-$, one also obtains $\mu = \mu^+ \sqcup \mu^-$, making this an instance of the "bilattice pattern" [3].

Possible world models indexed by probability distributions or by signed measures are quite fruitful in this context [3].

7 Vector semantics for program learning?

There is a strong hope that the ability to take linear combinations of programs can give rise to new schemes of program learning. We summarize our current understanding of the situation.

For $0 < \alpha < 1$ and **random** being a generator of uniformly distributed reals between 0 and 1, **if random** $< \alpha$ **then P else Q** yields a linear combination of programs **P** and **Q**. To allow negative coefficients one needs to consider computing a negative and a positive channel in parallel (computations are marked as negative or positive).

The situations when one can consider linear combinations of single execution runs should be especially attractive. Probabilistic programming via MCMC sampling is one example of this situation. To implement linear combinations of probabilistic programs with positive coefficients one can simply execute those programs in parallel, and the values of coefficients can be controlled by changing the relative execution speed of those programs. To allow negative coefficients one again needs to use a negative and a positive channel, not unlike the schemes used in retina (see pages 65 and 173 of [9]).

Fuzzy samplings where points are taken with real coefficients might be even more attractive (one can think about them as generalized animations, where points might be indexed by a more sophisticated index set than a discretized rectangle). Here one might allow negative coefficients and avoid the need for a separate channel (speaking in terms of animation this means that 0 is at some grey level, between black and white).

There are some indications that these formalisms are quite expressive, and that they allow for powerful schemes of program learning.

Among recent examples of how expressive these formalisms might be is a system which can solve a CAPTCHA as an inverse problem to computer graphics via an intelligent animation scheme controlled by a set of MCMC-sampled parameters [8].

For an example of a powerful program learning scheme for probabilistic programs using matrix decomposition and a context-free grammar of models see [5].

References

1. Andima, S., Kopperman, R., Nickolas, P.: An Asymmetric Ellis Theorem. Topology and Its Applications 155, 146–160 (2007)
2. Berlinet, A., Thomas-Agnan, C.: Reproducing Kernel Hilbert Spaces in Probability and Statistics. Kluwer Academic Publishers, Boston (2001)
3. Bukatin, M., Kopperman, R., Matthews, S.: Partial Inconsistency Landscape: an Overview. 28th Summer Conference on Topology and Its Applications, Nipissing University, July 2013. `http://at.yorku.ca/cgi-bin/abstract/cbgy-77`
4. Ginsberg, M.: Multivalued Logics: a Uniform Approach to Inference in Artifcial Intelligence. Computational Intelligence 4(3), 256–316 (1992)

5. Grosse, R., Salakhutdinov, R., Freeman, W., Tenenbaum, J.: Exploiting Compositionality to Explore a Large Space of Model Structures. Conference on Uncertainty in Artificial Intelligence (2012)
6. Jung, A., Moshier, M.A.: On the Bitopological Nature of Stone Duality. Technical Report CSR-06-13. School of Computer Science, University of Birmingham (December 2006)
7. Kozen, D.: Semantics of Probabilistic Programs. Journal of Computer and System Sciences 22 (3), 328–350 (1981)
8. Mansinghka, V., Kulkarni, T., Perov, Yu., Tenenbaum, J.: Approximate Bayesian Image Interpretation using Generative Probabilistic Graphics Programs (2013). `http://probcomp.csail.mit.edu/gpgp`
9. Marr, D.: Vision. W.H.Freeman and Company, New York (1982)
10. Popova, E.: The Arithmetic on Proper & Improper Intervals (a Repository of Literature on Interval Algebraic Extensions). `http://www.math.bas.bg/~epopova/directed.html`
11. Rodabaugh, S.: Functorial Comparisons of Bitopology with Topology and the Case for Redundancy of Bitopology in Lattice-valued Mathematics. Applied General Topology 9(1), 77–108 (2008)